NEAR EAST UNIVERSITY

Faculty of Engineering

Department of Computer Engineering

APPLICATIONS OF NEURAL NETWORKS

Graduation Project
COM- 400

Student: AHMAD SHAHIN

Supervisor: Assoc.Prof.Dr.Adnan Khashman

Nicosia - 2004

AKNOWLEDGEMENT

First, I am so thankful to most gracious Allah, the Almighty, who enabled me to complete my project.

I would like to thank my parents, brothers, sisters and all my relatives for their encouragements and endless supports, without them I would never be what I am now so a lot of thanks for them.

I should not forget to thank my intellectual and conscientious **Assoc. Prof. Dr Adnan Khashman** for inspiring and endurance me to complete this project.

With homologizing, I would like to thank my dear university (N.E.U.) that never forgot its big unmistakable and contribution in my development and composing my self.

And special thanks to Miss. Nazan Ulu she was always beside me and she helped me too much to get over a lot of problems during my studying time.

I want to say thanks to all of my friends who helped me in finalizing my project and especially to all I have living with, for their standing beside me. Really they were my shadow I twigged in complete my project.

My brothers Majdi, Mohammad, Anwar and Abdullah Shahin, Hani Fares, Ali Abu Kharma, Adnan Khader, Ahmad Eed, Ibrahim Kuzu, Rami Waleed, Abdulmajeed Mohammad, Sohaib Haseeb, Ayman Abu Dayeh, Yousef Abu Khuruj, Mazhar Ghuniem, Mohammad Barghothi, Majdi Sultan, Ahmad Bustami, Fares Namrooti and all my friends in Cyprus, Turkey, Jordan and Palestine.

ABSTRACT

This project explores the theoretical and particular underpinning of Neural Networks and its applications, the reader of this project will come away with an appreciation for the basic concepts of Neural Network., and with an idea about Neural Networks fields and the use of its applications.

This project reports on the area of Neural Networks and how it becomes more important in both undergraduate and graduates in computer science and engineering. This will provide the fundamental conceptual necessary to confront the rapidly developing of the world.

This project includes a general background and history about the Neural Networks and some early examples, some applications areas is included in this project.

In this project several model and applications of Neural Networks available; like Neural Networks in medicine, in sport, and other applications, with real examples.

TABLE OF CONTENTS

INTRODUCTION

1. Overview

Neural Networks are collections of mathematical models that emulate some of the observed properties of biological nervous systems and draw on the analogies of adaptive biological learning. The key element of the Neural Network paradigm is the novel structure of the information processing system. It is composed of a large number of highly interconnected processing elements that are analogous to neurons and are tied together with weighted connections that are analogous to synapses.

Learning in biological systems involves adjustments to the synaptic connections that exist between the neurons. This is true of Neural Networks as well. Learning typically occurs by example through training, or exposure to a truthed set of input/output data where the training algorithm iteratively adjusts the connection weights (synapses). These connection weights store the knowledge necessary to solve specific problems.

Although Neural Networks have been around since the late 1950's, it wasn't until the mid-1980's that algorithms became sophisticated enough for general applications. Today Neural Networks are being applied to an increasing number of real- world problems of considerable complexity. They are good pattern recognition engines and robust classifiers, with the ability to generalize in making decisions about imprecise input data. They offer ideal solutions to a variety of classification problems such as speech, character and signal recognition, as well as functional prediction and system modeling where the physical processes are not understood or are highly complex. Neural Networks may also be applied to control problems, where the input variables are measurements used to drive an output actuator, and the network learns the control function. The advantage of Neural Networks lies in their resilience against distortions in the input data and their capability of learning. They are often good at solving problems that are too complex for conventional technologies (e.g., problems that do not have an algorithmic solution or for which an algorithmic solution is too complex to be found) and are often well suited to problems that people are good at solving, but for which traditional methods are not.

There are multitudes of different types of Neural Networks. Some of the more popular include the multilayer perceptron which is generally trained with the backpropagation of error algorithm, learning vector quantization, radial basis function, Hopfield, and Kohonen, to name a few. Some Neural Networks are classified as feedforward while others are recurrent (i.e., implement feedback) depending on how data is processed through the network. Another way of classifying Neural Networks types is by their method of learning (or training), as some Neural Networks employ supervised training while others are referred to as unsupervised or self-organizing. Supervised training is analogous to a student guided by an instructor. Unsupervised algorithms essentially perform clustering of the data into similar groups based on the measured attributes or features serving as inputs to the algorithms. This is analogous to a student who derives the lesson totally on his or her own. Neural Networks can be implemented in software or in specialized hardware.

2. Description of Thesis Structure

The first chapter introduction to neural network you will see how neural network work related to the brain, why neural network now? And a brief history abouit neural networks.

The second chapter will describe the architecture of neural networks and the way of learning which are supervised learning and unsupervised learning,

The third chapter will discusses the most important applications of a neural networks such as:

Business applications Neural network red-flags police officers with potential for misconduct, Credit scoring with BrainMaker neural network software, BrainMaker strikes gold! And managing jury summoning with BrainMaker.

Manufacturing applications BrainMaker performs non-destructive concrete strength testing, neural network optimize enzyme synthesis, the use of neural networks in testing plastic quality, BrainMaker tracks beer quality and using neural networks to determine steam quality.

Sport applications Predicting thoroughbreds finish time with BrainMaker neural networks, BrainMaker predicts the order of finish in horseracing and selecting winning dogs with BrainMaker neural networks.

Science applications Using a neural network to measure air quality, using a neural network to predict **EL** Nino, BrainMaker analysis of Transmembrane-spanning protein helices and BrainMaker predicts Rainfall.

The forth chapter will be about the medical application of neural network classifies breast cancer cells with BrainMaker neural network, Brain maker improves hospital treatment and reduces expenses, diagnose heart attacks with BrainMaker neural network software and BrainMaker order medical laboratory tests for ER.

3. The aim of this project

The following aims and objectives are to be met throughout the work presented in this project these aims can be summarized as :

1. To know neural networks definitions and the use of them.
2. To know neural networks learning ways, some example of them, and how to build a neural network.
3. To investigate the structure of neural networks and applications in the real life.
4. To be able to describe the relation between neural networks and the human brain.
5. To show how neural networks can be used in medicals field.

CHAPTER ONE

INTRODUCTION TO NEURAL NETWORKS

1.1 Overview

This chapter present Neural Networks In general, and it contains the history of Neural Networks, the future of it., what is Neural Networks and shows that Neural network model consists of a set of computational units at certain time unit examines its input and computes signed number called an activation as its out put.

Each connection has signed number called a weight that determines whether an activation that travels along it influences the receiving cell to produce a similar or a different activation. According to the sign (+ or -) of the weight.

The connection and their weights are the important parameter in any model. They determine the behavior of a model, and can be compared to instruction in a conventional computer program.

1.2 Definition of Neural Networks

Neural Networks have a large appeal to many researchers due to their great closeness to the structure of the brain, a characteristic not shared by more traditional system.\

In an analogy to the brain, an entity made up of interconnected neurons, neural networks are made up of interconnected processing elements called units. Which respond in parallel to a set of input signals given to each. The unit is the equivalent of its brain counterpart, the neurons.

1.3 How Neural Networks Work?

They point out that neural nets can use extrapolation for classification problems or interpolation for computational problems. This paper specifically sets out to explain how a neural net can solve computational problems. The authors choose examples from chaos theory to demonstrate how neural nets can be used for predicting time series. Several examples of chaotic systems in physics and the natural sciences are provided. The authors cite the works of others that have demonstrated that two hidden layers are sufficient to solve most problems.

The authors use backpropagation to train the neural net. The process of backpropagation is explained thoroughly. The neural net takes the inputs and the outputs of the training data set and assigns weights to the connections. Next, the neural net calculates the sum of squared errors.

Iterations occur which adjust the weights of the connections in order to minimize the sum of squared errors. After this, the non-training data set inputs are entered into the neural net so that the neural net can calculate the predicted outputs. The authors used the Glass-Mackey equation to generate time series. They generated data sets with 500 observations. They then applied neural nets to make predictions for events. They used a neural net to make predictions at various future observations (t + P, where P is the number of periods ahead of observation t). Although they found that the neural net because less accurate as the number of periods in the future increased, their results for the neural net were still an order of magnitude more accurate at increased intervals than the other methods they employed.

The authors demonstrate graphically why the neural net is successful for computational problems. The neural net uses a squashing function that is flat then descends in a curve to a lower flat surface. The neural net then adds a squashing function that has the same absolute value but is of opposite sign to the original squashing function; this sum looks like a ridge. The next step is to add the sum of two squashing functions that are perpendicular to the original summed squashing functions. This results in a bump. The edges of the bump are not flat, however. So the neural net adds a bias term (or set of bias terms) in order to flatten the edges of the bump. The greatest height of the bump or lowest point in the bump will be the optimized solution to the problem.

1.4 Why Neural Network Now?

Current technology has run into a lot of bottlenecks-sequential processing, for one. When a computer can handle information only one small piece at a time, there are limits to how fast you can push a lot of information through. Even with many processors working in a parallel, much time is wasted waiting for sequential operation to complete. It's also difficult to write programs that can use parallel operation effectively.

1.5 What is a Neural Network?

Neural Networks have a large appeal to many researchers due to their great closeness to the structure of the brain, a characteristic not shared by more traditional system.\

In an analogy to the brain, an entity made up of interconnected neurons, neural networks are made up of interconnected processing elements called units. Which respond in parallel to a set of input signals given to each. The unit is the equivalent of its brain counterpart, the neurons.

There are many definitions for a Neural Network

1-*Neural network*: is a massively parallel-distributed processor that has a natural propensity for experiential knowledge and making it available for use.

2-*Neural network*: is machine that is designed to model the way in which the brain performs a particular task or function of interest, the network is usually importanted using electronic components or simulated in software on digital computers.

3-***Neural networks*:** are also referred to in the literature as neurocomputers, connectionist network, parallel-distributed processors, etc.

4-***Neural networks*:** are different paradigms for computing:

A-Von Neumann machines are based on the processing/memory abstraction of human information processing.

B-Neural networks are based on the parallel architecture of animal brains.

5-***Neural networks*:** are forms of multiprocessor computer system, with

A-Simple processing elements.

B-A high degree of interconnection.

C-Simple scalar messages.

D-Adaptive interaction between elements.

6-***Neural network*:** a mathematical model composed of a large number of processing elements organization in to layers.

1.6 Benefits of Neural Networks

The use of neural networks offers the following useful properties and capabilities:

1-Nonlinearity. A neuron is basically a nonlinear device. Consequently, a neural network, made up of an interconnection of neurons is itself nonlinear.

2-Input – output mapping. Popular paradigm of learning called supervised learning involves the modification of the synaptic weights of a neural network by applying a set of labeled training samples or task examples.

3-Adaptively. Neural networks have a built-in capability to adapt their synaptic weights to changes in the surrounding environment.

4-Evidential response. In the context of pattern classification, a neural network can be designed to provide information may be used to reject ambiguous patterns, should they arise, and thereby improve the classification performance of the network.

5-Contextual information. Knowledge is represented by the very structure and activation state of a neural network.

1.7 History of Neural Networks

The earliest work in neural computing goes back to the 1940's when McCulloch and Pitts introduced the first neural network computing model. In the 1950's, Rosenblatt's work resulted in a two-layer networks, the perceptron, which was capable of learning certain classifications by adjusting connection weights. Although the perceptron was successful in classifying certain patterns, it had a number of limitations. The perceptron was not able to solve the classic XOR (exclusive or) problem. Such limitations led to the decline of the field of neural networks. How ever, the perceptron had laid foundations for later work in neural computing. {1}.

In the early 1980's, researchers showed renewed interest in neural networks. Recent work includes Boltzmann machines, Hopfield nets, competitive learning models, multilayer networks, and adaptive resonance theory models. {1}.

The first artificial neurons were produced in 1943 by the neurophysiologist Warren McCulloch and the logician Walter Pits. But the technology available at that time did not allow them to do too much. {1}.

1.8 Future of Neural Networks

A great deal of research is going on in neural networks worldwide. This ranges from basic research into new and more efficient learning algorithms, to networks, which can respond to temporally varying patterns, to techniques for implementing neural networks directly in silicon. Already one chip commercially available exists, but it does not include adaptation. And is working on the learning problem. Production of a learning chip would allow the application of this technology to a whole range of problems where the price of a

PC and software cannot be justified. There is particular interest in sensory and sensing applications: nets, which learn to interpret real-world sensors and learn about their environment.{2}

1.9 Where are Neural Networks Applicable?

Or are they just a solution in search of a problem? Neural networks cannot do anything that cannot be done using traditional computing techniques, **BUT** they can do some things, which would otherwise be very difficult. In particular, they can form a model from their training data (or possibly input data) alone. This is particularly useful with sensory data, or with data from a complex (e.g. chemical, manufacturing, or commercial) process. There may be an algorithm, but it is not known, or has too many variables. It is easier to let the network learn from examples.

1.10 Neural Networks Fields of Application:

In investment analysis:

To attempt to predict the movement of stocks currencies etc., from previous data. There, they are replacing earlier simpler linear models.

In signature analysis:

As a mechanism for comparing signatures made (e.g. in a bank) with those stored. This is one of the first large-scale applications of neural networks in the USA, and is also one of the first to use a neural network chip.

In process control:

There are clearly applications to be made here: most processes cannot be deterring computable algorithms. Newcastle University Chemical Engineering Department is working with industrial partners (such as Zeneca and BP) in this area.

In monitoring:

Networks have been used to monitor the state of aircraft engines. By monitoring vibration levels and sound, early warning of engine problems can be given.

In marketing:

Networks have been used to improve marketing mail shots. One technique is to run a test mail shot, and look at the pattern of returns from this. The idea is to find a predictive mapping from the data known about the clients to how they have responded. This mapping is then used to direct further mail shots.

1.11 Summary

This chapter showed some generals information about Neural Networks and the way the process act between the units and the neurons, and it provided the historical background for the Neural Networks just like the beginning of this in 1940's and its history through the years, and it supported some fields of Neural Networks and where it can be used in the real world.

CHAPTER TWO

ARCHETECTURE OF NEURAL NETWORKS

2.1 Overview

This chapter will explain the learning of Neural Networks with some of the most important example for each one, and it has some graphics to show the way of the connections between the neurons, and the Neural Networks trainings again with some figures, and how to build a Neural Network.

A neural network consists of four main parts:

1. Processing units, where each unit has certain activation level at any point in time.
2. Weighted interconnections between the various processing units, which determine how the activation of one unit leads to input for another unit.
3. An activation rule which acts on the set of input signals at a unit to produce a new output signal, or activation.
4. Optionally, a learning rule that specifies how to adjust the weights for a given input/output pair.

2.2 Neural Network Learning

Among the many interesting properties of a neural network, the property that is of primary signification is the ability of the network to learn from its environment, and to improve its performance through learning; the improvement in performance takes place over time in accordance with some prescribed measure. A neural network learns about its environment through an iterative process of adjustments applied to its synaptic weights and thresholds. Ideally, the network becomes more knowledgeable about its environment after each iteration of the learning process. Machine learning refers to computer models that improve their performance in signification ways based upon data.

2.2.1 Supervised learning

This is usually performed with feedforward nets where training patterns are composed of two parts, an input vector and an output vector, associated with the input and output nodes respectively. A training cycle consists of the following steps. An input vector is presented at the inputs together with a set of desired responses, one for each node, at the output layer. A forward pass is done and the errors or discrepancies, between the desired and actual response for each node in the output layer, are found. These are then used to determine weight changes in the net according to the prevailing rule. The term `supervised' originates from the fact that the desired signals on individual output nodes are provided by an external `teacher'. The best-known examples of this technique occur in the backpropagation algorithm, the delta rule and Perceptron rule. See Fig. 2.1

2.2.1.1 Supervised learning divided into two parts:

1-Feedback nets:

A-Back propagation through time

B-Real time recurrent learning

C-Recurrent extended kalman filter

2-Feed forward –only net: -

A-Perceptron

B-Adeline, Madeline

C-Time delay neural network

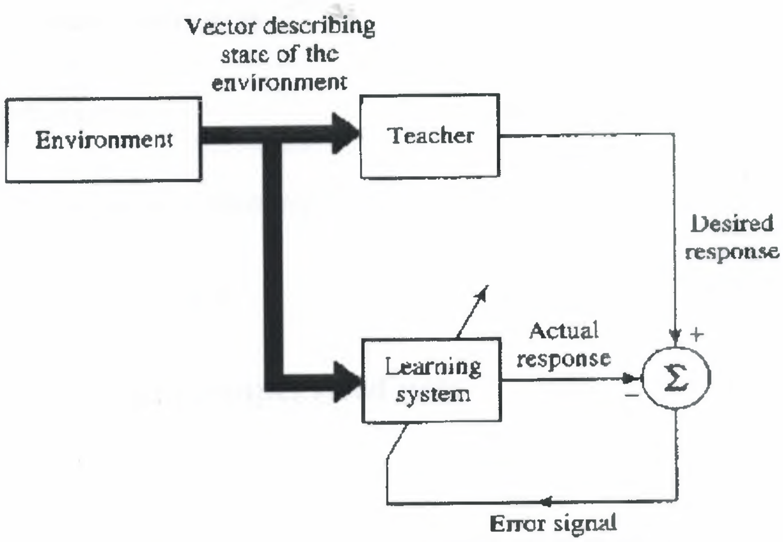


Fig. 2.1 supervised learning

2.2.2 Unsupervised learning

This is usually found in the context of recurrent and competitive nets. There is no separation of the training set into input and output pairs. Typically a training cycle will consist of the following steps: a vector is applied to the visible nodes (or in the case of competitive learning, the input nodes); the net is allowed to reach equilibrium; weight changes are made according to some prescription. It is the amalgamation of input-output pairs, and hence the disappearance of the external supervisor providing target outputs, that gives this scheme its name. This kind of learning is sometimes referred to as self-organization. See Fig. 2.2

2.2.2.1 Unsupervised divided into two parts:

1-Feedback nets:

A-Discrete hop filed

B-Analog adaptive resonance theory

C-Additive gross berg

2-Feed forward –only nets

A-Learning matrix

B-Linear associative memory

C-Counter propagation

2.2.3 Applications for unsupervised nets

Clustering data:

Exactly one of a small number of output units comes on in response to an input.

Reducing the dimensionality of data:

Data with high dimension (a large number of input units) is compressed into a lower dimension (small number of output units).

Although learning in these nets can be slow, running the trained net is very fast - even on a computer simulation of a neural net.

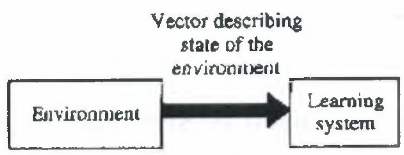


Fig. 2.2 unsupervised learning

2.2.4 Hebbian learning

Hebb's postulate of learning is the oldest and most famous of all learning rules; it is named in honor of the neurophysiologist Hebb (1949). Quoting from Hebb's book, the organization of behavior (1949, p.62): {3}

When an axon of cell A is near enough to a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic changes take place in one or both cells such that A's efficiency as one of the cells firing B is increased.

Hebb proposed this change as a basis of associative learning (at the cellular level), which would result in an enduring modification in the activity pattern of a spatially distributed "assemble of never cells."

2.3 Network Architectures

The manner in which the neurons of a neural network are structured is intimately linked with the algorithm used to train network.

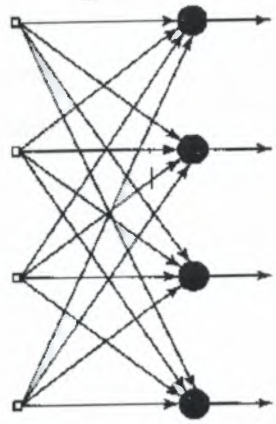
In general we may identify four different classes of network architectures:

2.3.1 Single-Layer Feedforward Networks

A layered neural network is a neurons organized in the form of layers. In the simplest form of a layered network, we just have an input layer of source nodes that project onto an output layer of neurons (computation nods), but not vice versa.

In other words, this network is strictly of a feedforward type. It is illustrated in fig. 2.3 for the case of four nodes in both the input layers. Such a network is called a single-layer network, with the designation "single layer" referring to the output layer of source nodes, because no computation is performed there. A linear associative memory is an example of a single-layer neural network.

In such an application, the network associates an output pattern (vector) with an input pattern (vector), and information is stored in the network by virtue of modification made to the synaptic weights of the network.{3}



In put layer of source node   out put layer of neurons

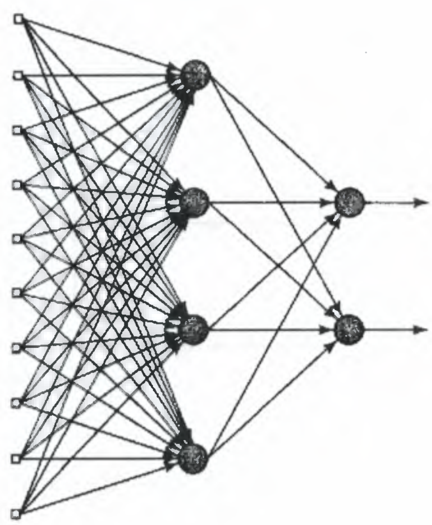
Fig 2.3 feedforward network with a single layer of neurons.

2.3.2 Multilayer Feedforward Networks

The second class of a feedward neural network distinguishes itself by the presence of one or more hidden layers, whose computation nodes are correspondingly called hidden neurons or hidden units. The function of the hidden is to intervene between the external input and the network output. By adding one more hidden layer, the network is enabled to extract higher-order statistics, for (in a rather loose sense) the network acquires a global perspective despite is local connectivity by virtue of the extra set of synaptic connections and the extra dimension of neural interaction. The ability of hidden neurons to extract higher-order statistics is particularly valuable when the size of the input layer is large.

The source node in the input layer of the network supply respective elements of the activation pattern (input vector), which constitute the input signals applied to the neurons (computation nodes) in the second layer (i.e., the first hidden layer). The output signals of the second layer are used as input to the third layer, and so on for the rest of the network. Typically, the neurons in each layer of the network have as their inputs the output signals of the preceding layer only.

The neural network of Fig. 2.4 is said to be fully connected in the sense that every node in each layer of the network is connected to every other node in the adjacent forward layer. If, however, some of the communication links (synaptic connections) are missing from the network, we say that the network is partially connected. A form of partially connected multilayer feedforward network of particular interest is a locally connected network. {3}



In put layer	layer of	layer
of sources	hidden	out put
nodes	neurons	neurons

Fig 2.4 one hidden layer and output

2.3.3 Recurrent Networks

A recurrent neural network distinguishes itself from a feedforward neural network in that it has at least one feedback loop. For example, a recurrent network may consist of a single layer of neurons with each neuron feeding its output signal back to the inputs of all the other neurons, as illustrated in the architectural graph of Fig 2.5.

In the structure depicted in this figure there are no self-feedback loops in the network; self-feedback refers to a situation where the output of a neuron is fed back to its own input. The recurrent network illustrated in fig. 2.5 also has no hidden neurons.{2}

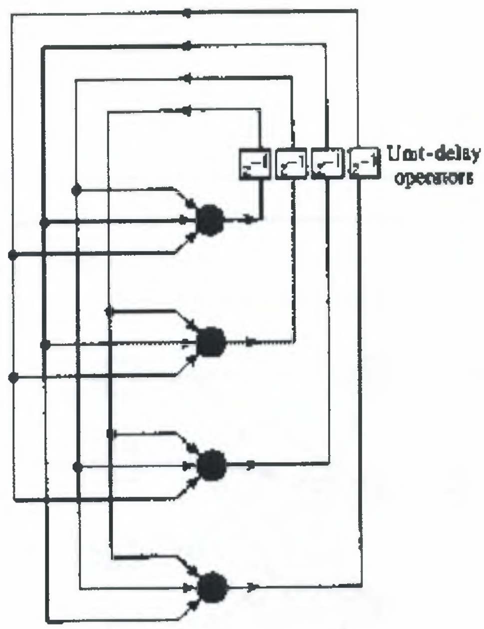


Fig 2.5 Recurrent network with no self-feedback loops and no hidden neurons

2.3.4 Lattice Structures

A lattice consists of a one-dimensional, two-dimensional, or higher-dimensional array of neurons with a corresponding set of source nodes that supply the input signals to the array; the dimension of the lattice refers to the number of dimensions of the space in which the graph lies. The architectural graph of Fig. 2.6 depicts a one-dimensional lattice of 3 neurons fed from a layer of 3 source nodes. The hidden layer learns to recode (or to provide a representation for) the inputs. More than one hidden layer can be used.

The architecture is more powerful than single-layer networks: it can be show that any mapping can be learned, given two hidden layers (of units). The units are a little more complex than those in the original perception: their input/output graph is

As a function:

$Y = 1 / (1+ \exp(-k. (\&sum\ W_{in} * X_{in}))$

The graph shows the output for *k=0.5, 1, and 10,* as the activation varies from -10 to 10.

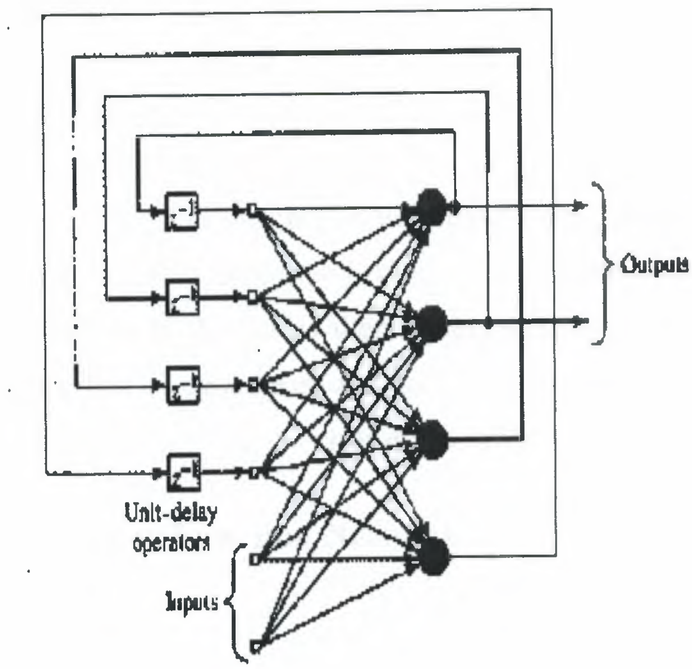


Fig 2.6 input layer of source nodes{3}

2.3.5 Typical Radial Basis Function Architecture:

Like BP, RBF nets can learn arbitrary mappings: the primary difference is in the hidden layer.

RBF hidden layer units have a receptive field, which has a *center:* that is, a particular input value at which they have a maximal output. Their output tails off as the input moves away from this point.

2.4 Neural Networks Training

2.4.1 Training Back Propagation Networks

The weight change rule is a development of the perception-learning rule. Weights are changed by an amount proportional to the error at that unit times *the output of the unit* feeding into the weight.

Running the network consists of: -

Forward pass:

The outputs are calculated and the error at the output units calculated.

Backward pass:

The output unit error is used to alter weights on the output units. Then the error at the hidden nodes is calculated (by back-propagating the error at the output units through the weights), and the weights on the hidden nodes altered using these values.

For each data pair to be learned a forward pass and backwards pass is performed.

This is repeated over and over again until the error is at a low enough level

2.4.2 Training Radial Basis Function Networks.

RBF networks are trained by

*Deciding on how many hidden units there should be

*Deciding on their centers and the sharp nesses (standard deviation)

*Training up the output layer.

Generally, the centers and SDs are decided on first by examining the vectors in the training data. The output layer weights are then trained using the Delta rule. BP is the most widely applied neural network technique. RBFs are gaining in popularity.

Nets can be

*Trained on classification data (each output represents one class), and then used directly as classifiers of new data.

RBFs have the advantage that one can add extra units with centers near parts of the input, which are difficult to classify. Both BP and RBFs can also be used for processing time-varying data: one can consider a window on the data: {4}

2.4.3 Back-Propagated Delta Rule Networks (BP)

Is a development from the simple Delta rule in which extra *hidden layers* (layers additional to the input and output layers, not connected externally) are added? The network topology is constrained to be feed forward: i.e. loop-free - generally connections are allowed from the input layer to the first (and possibly only) hidden layer; from the first hidden layer to the second... and from the last hidden layer to the output layer.

2.4.4 Kohonen clustering Algorithm:

In the winner-take-all competitive learning network, only connections to the winner are updated and updating of the weights dose not rely in any way on the spatial relations among the units in the competition layer. Therefore, as expected, the winner-take-all network cannot develop any spatial organization.

Kohonen demonstrated the formation of a topographic map of the first type by unsupervised self-organization. Unit in the output layer start off by responding randomly to the input signal. Kohonen identified two key mechanisms for a network to self-organization spatially.

1- Locates the unit that best responds to the given input. This unit is called the winning unit.

2- Modify the connections to the winning unit and connections to units in its neigh borhood.

2.5 Visualizing Processes in Neural Networks

An insightful method to overcome the weakness in our present understanding of knowledge representation inside a neural network is to resort to the experimental use of visualization of the learning process. By so doing we are merely recognizing the fact that representing information-bearing data by visual means is the essence of scientific visualization.

Indeed, such an approach permits the human eye-brain system to perceive and infer visual information between the neural network simulator and the user, this is all the more so, given the enhanced facilities for the interactive manipulation of imaging and display processes that are presently available. Yet it has to be said that the use of graphical display of the learning process experienced by a neural network has not received the attention it deserves. The Hinton diagram involves drawing columns of squares, with each column representing the synaptic weights and threshold of a particular neuron in the network.

The size of each square represents the magnitude of a certain synaptic weight; the color of the square, black or white, indicates the polarity of the weight, positive or negative, respectively. The various columns of square are positioned in the diagram so as to maintain correspondence with network architecture. Consider, for example, the simple, two-layer feedforward network of Fig.2.7.a The Hinton diagram for this network is shown in Fig. 2.7.b Starting from the bottom of the diagram, the first column of square represents the two weights and threshold of the top hidden neuron.

The second column of square represents the two weights and threshold of the bottom-hidden neuron. Finally, the third column of square represents the two weighs and weights and threshold for the output neuron. The top row of square in the diagram represents the thresholds of the individual neurons. A limitation of the Hinton diagram, however, is that it dose not explicitly reveal the network topology in relation to the synaptic weight data. We usually find it highly informative to have a graphical representation that explicitly integrates synaptic weight values with network topology, for than we are able to see how the internal representation of synaptic weights relates to the particular problem the neural network is learning.

Such an objective is met by the bond diagram, according to which the synaptic weights are display as "bond" between the nodes of the network. In particular, the stronger a synaptic connection between two nodes is, the longer would the bond be.{4}

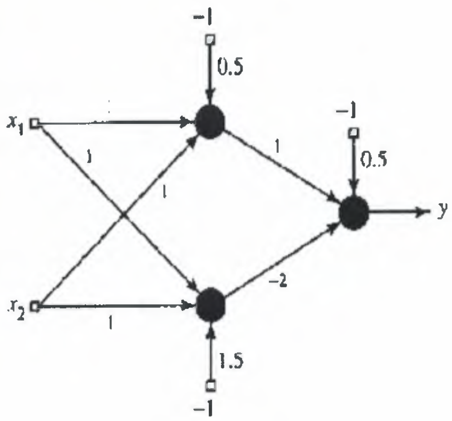


Fig 2.7.a Neural Network

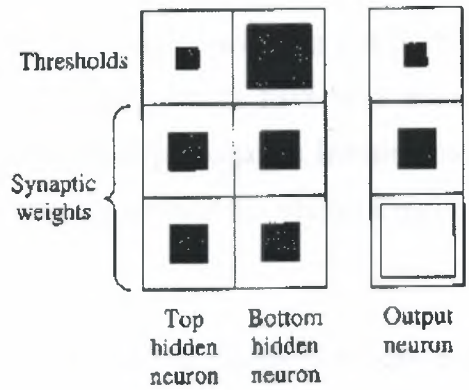


Fig.2.7.b Hinton Diagram

2.6 Building a Neural Network

Since 1958, when psychologist Frank Rosenblatt proposed the "Perceptron," a pattern recognition device with learning capabilities, the hierarchical neural network has been the most widely studied form of network structure. A hierarchical neural network is one that links multiple neurons together hierarchically. The special characteristic of this type of network is its simple dynamics. That is, when a signal is input into the input layer, it is propagated to the next layer by the interconnections between the neurons. Simple processing is performed on this signal by the neurons of the receiving layer prior to its being propagated on to the next layer. This process is repeated until the signal reaches the output layer completing the processing process for that signal. The manner in which the various neurons in the intermediary (hidden) layers process the input signal will determine the kind of output signal it becomes (how it is transformed). As you can see, then, hierarchical network dynamics are determined by the weight and threshold parameters of each of their units. If input signals can be transformed to the proper output signals by adjusting these values (parameters), then hierarchical networks can be used effectively to perform information processing. Since it is difficult to accurately determine multiple parameter values, a learning method is employed. This involves creating a network that randomly determines parameter values. This network is then used to carry out input-to-output transformations for actual problems. The correct final parameters are obtained by properly modifying the parameters in accordance with the errors that the network makes in the process. Quite a few such learning methods have been proposed. Probably the most representative of these is the error back-propagation learning method proposed by D. E. Rumelhart et al. in 1986. This learning method has played a major role in the recent Neurocomputing boom.

The back-propagation paradigm has been tested in numerous applications including bond rating, mortgage application evaluation, protein structure determination, backgammon playing, and handwritten digit recognition. Choosing the right methodology, or backpropagation algorithm, is another important consideration. In working with the financial applications, many have found that the back-propagation algorithm can be very slow. Without using advanced learning techniques to speed the process up, it is hard to effectively apply backpropagation to real-world problems. Overfitting of a neural network model is another area, which can cause beginners difficulty. Overfitting happens when an ANN model is trained on one set of data, and it learns that data too well. This may cause the model to have poor generalization abilities - the model may instead give quite poor results for other sets of data. For an in-depth coverage of other neural network models and their learning algorithms, please refer to the Technical Reading at the end of this User's Guide, the Technical Reference (sold separately), those papers listed in the Reference, or any other reference books on neural networks and relevant technology.{4}

2.7 Summary

This chapter has good information about neural networks because it showed the way of learning in neural networks with the most important application, and it described the networks model using neural network system, and the building of neural networks by showing the connection between the layers which have the input and the output layers and in some example there is a hidden layers connected. Everything was described with a figure to make it easy to be understandable.

CHAPTER THREE

APPLICATIONS OF NEURAL NETWORKS

3.1 Overview

This chapter is the main chapter in this project, it shows some of neural networks application with some explained example for them just like business field like the red-flag police officer, science field as neural networks is being used to measure air quality and many more things in science, and also this chapter shows how neural network can be used in sport and industrial sections.

3.2 Applications of Neural Networks

3.2.1 Business

3.2.1.1 Neural Network Red-Flags Police Officers With Potential

Neural network to forecast which officers on the force are potential candidates for misbehavior. The Department's Internal Affairs Division used neural networks to study 200 officers who had been terminated for disciplinary reasons and developed a database of pattern-like characteristics, behaviors, and demographist found among the 200 police officers.

Neural network then compared current Department officers against the pattern gleaned from the 200-member control group and produced a list of officers who, by virtue of matching the pattern or sharing questionable characteristics to some degree, were deemed to be "at risk. To counterbalance the inherent "dispassion" of the neural network, the Department closely examined the net's findings to ensure that officers who are clear anomalies, and thus don't warrant being on the list, are removed from consideration.

3.2.1.2 Credit Scoring with Neural Network Software

The credit scoring neural network was trained on no more than 100 loan applications yet achieved a 75-80% success rate. One day's work by an operator familiar with the BrainMaker software package was required to build, train and test the credit scoring neural network. Except for showing a greater bias towards approving weak loan applications, the neural network's loan classification rate was identical to that achieved using a commercial credit-scoring scheme. The input data for the credit scoring with neural network software study consisted of information typically found on loan applications. The outcomes of those loans were classified as delinquent, charged-off, or paid-off. The actual outputs from the network were 0 to 1 ratings for the three alternatives.

Once the network was built, it was subjected to two training trials. In the first trial, the data was arranged in random order and the first 75 applications were used to train the network. The remaining 50 applications were then evaluated using the trained network. The network misclassified 10 of the 50 applications in the sample for an 80% success rate. In short, the network favored approving loan applications. More traditional and much more costly, credit scoring method used by 82% of all banks, resulted in a 74% success rate. The credit scoring method proved to be more conservative than the neural network in granting credit in the second trial, the data was rearranged in different random order and the first 100 applications were used to train the network. The remaining 25 applications were then evaluated using the trained network. The network misclassified 6 of the 25 applications in the sample for a 76% success rate. Classifications of good loans as bad and of bad loans as good were equal at 12% each. The credit scoring method for this sample of 25 applications also misclassified 6 of the 25 applications.{5}

3.2.1.3 Neural Network Strikes Gold!

Neural networks are a form of artificial intelligence and are currently being applied in the medical field to screen for cancer and other diseases. Neural pattern recognition programs are also common in the financial world and in the optical character recognition and speech recognition programs. In the case of mining, we can say the technology has proved effective in step out drilling an existing mine.{4}

3.2.1.4 Managing Jury Summoning with Neural Network

The Intelligent Summoner from MEA allows a courthouse to dramatically reduce the number of jurors called for potential service, saving an average of 25-40% of the cost of jurors. Courts often waste money by calling (and paying for) more potential jurors to show up and wait around than are needed for the cases being heard. The Montgomery Court House in Norristown, PA saves $70 million annually using this system. The Intelligent Summoner determines the number of jurors needed for the next day at a specific courthouse.

The program is custom-tailored for a specific courthouse by MEA with information about that courthouse. At least one year's worth of past information about the courthouse is needed for custom tailoring. The information can be entered into the Intelligent Summoner system and includes dates, judges, types of cases, and number of jurors used at a particular courthouse. A file is written to disk, which is then sent to MEA. MEA then creates and trains a neural network using neural network, which learns the needs of that courthouse. The trained neural network is sent back to the courthouse and the file is read onto the PC. From then on, the system is ready for use at that courthouse by anyone with typing skills.

To use the system, a daily survey of tomorrow's judges, trial types, and size of jury panel is input with the data-entry portion of the program. The program will immediately provide the total number of jurors that should be called for tomorrow.{5}

3.2.2 Manufacturing

3.2.2.1 Neural Networks Optimize Enzyme Synthesis

A neural network has been trained to predict the outcome of a chemical reaction controlled by molar ratios, temperature, pressure, amount of enzyme and stirring speed.

Used the BrainMaker program to train their neural network to predict the amount of desired product and by-product, which would be formed after 22 hours of reaction time. An excellent correlation between predicted yields and experimental results was found. The neural network saves time and money by predicting the results of chemical reactions so that the most promising conditions can then be verified in the lab, rather than performing a large number of experiments to gain the same information.

Initially 16 experiments were performed to identify the most important parameters controlling the process. The molar ratio between fatty acid and glucoside, reaction temperature, pressure, amount of enzyme, and stirring speed were varied. The synthesis yielded ethyl 6-O-dodecanoyl D-glucopyranoside. This experimental data was used to train the neural network to output the amount of the 6-O monoester and a diester by-product, represented as a percentage of yields.

The neural network had three layers:

1. Five input layer neurons.

2. Four hidden layer neurons.

3. Tow output layer neurons.

It was trained using the back propagation algorithm with the sigmoid threshold neuron function. Twelve facts were used to train the network to an accuracy of 96% for the outputs. In only a few minutes, all facts were learned. The trained network was then asked to make four predictions on data it hadn't seen before. The network predictions were compared to experimental observations. Very good correlations were found.

The average deviation between the network and the experiments was 4% (percentage of yield), ranging between 2% and 7% difference. These deviations are within the normal experimental error of synthesis after being tested; the network was put to work evaluating thousands of possible conditions in order to find the most optimum. Using a simple algorithm, a test file was generated containing all of the possible values, totaling 9900 cases.

The computer- generated test file contained values for each parameter which were both within and without of the training value's range. The entire file ran through the network in 7 minutes and the predictions were saved in a file. Using a search function, predictions for specified yields were selected. Only three cases were found to predict more than 88% monoester with a less than 4% formation of the diester. One of these cases was tested in the lab and the results were close to experimental observation. The network had predicted 88.1% monoester and the experiment yielded 86.2%. The network predicted 4.0% diester the experiment yielded 4.8%. Finally, the 9900 predictions were again searched, but this time with additional restrictions more suitable for large-scale chemical processing. Again, the experimental results were very close to the yields predicted by the network{6}

3.2.2.2 The Use of Neural Networks in Testing Plastic Quality

Using neural network to predict the quality of plastics to be used in windshields. The lab tests for brittleness and elasticity take up to 4 hours. Through use of a neural network they were able to save research time, processing time, and money.{4}

3.2.2.3 Neural Network Tracks Beer Quality

Neural network has used to identify the organic content of their own and their competitors' beer vapors with 96% accuracy. This allows them to assure consistent quality for their customers, and keep track of any changes made by their competitors. "Contrary to my earlier beliefs, neural network applications to natural product classification have proven to be of considerable propriety interest to the company. As a result, we have decided not to publish any information concerning our activities at this time. I am sure you can understand our position we will state, however, that recent efforts have been quite rewarding and that work will be continuing in this area of endeavor."{7}

3.2.2.4 Using Neural Networks to Determine Steam Quality

Canada has developed the **INSIGHT** steam quality monitor, an instrument used to measure steam quality and mass Flowrate. Steam Quality and Mass Flowrate is the energy injected into the ground in an oil recovery project, for example.

The improvement obtained by using the trained network was immediately apparent using a conventional linear program, the standard error of estimate (RMS of deviations about the ideal line) for steam quality and mass Flowrate are 28% and 0.59 kg/s. using the trained neural network, the standard error was 8.2% and 0.34kg/s. A common test set of 26 sets of input data was used and the network was trained on an additional 100 facts. Later, a similar network was trained and tested all of the **INSIGHT** monitor calibration data obtained to date (i.e. data from tests at four different facilities collected over a period of seven years using a minimum of six to a maximum of nine different monitors). Here, the standard error of estimate for steams quality and mass Flowrate were 7.7% and 0.4kg/s, respectively.

Recently neural network has successes trained to return methanol, gasoline and water contents from the RF reflectance spectra of mixtures of these three components. Currently they are investigating the application of a neural network to a-spectroscopy and to the interpretation of on-line chemical sensor signals.{8}

3.2.3 Sport

3.2.3.1 Predicting Thoroughbreds Finish Time with Neural Networks

When Don Emmons' neural network picked the winning horse in 17 out of 22 thoroughbred races at Detroit Race Course, he was astounded that a $195 program (BrainMaker) running on a PC could do so well. "I am amazed at the ability of software to consistently include the winning horse in three of its picks," said Don.

Designing a neural network is largely a matter of defining the problem well in your own mind. The most difficult aspects are deciding what information you're going to use and gathering it. There are several known methods of successfully predicting horserace winners with neural networks BrainMaker Professional provides a program, which automates the design of “competition”, networks such a horserace predictors. The program "Compete" designs, tests and runs a network based upon comparisons of all the competing items, comparing them two at a time. The one that wins the largest number of

2-item comparisons is the overall winner. Each item is rated for its overall likelihood of winning. Another design approach, which can be used with standard neural network, uses the full number of competing items as input, such as ten horses. There are two difficulties with this. First, not all races have the same number of horses running. Second, There will be a lot of inputs and outputs, making it more difficult to train. Has designed approaches the problem by looking at each horse individually and predicting how well it might do in the race. A separate neural network is trained for each horse with past performance information. Then the neural networks are given current information and the ones with the best results are selected as likely winners. When Don selected three horses as the possible winner of the race, 77% of the time one of them was the winner.

Information may be gathered from the Racing Form or from a computerized service such as the Equine Line. Neural network will read several different databases or a plain text file. Each horse has its own file and every past race is on a different line in the file. Don started out with four pieces of information: post position, the number of horses in the race, the horse's finish time and the track record. A little work was done with the numbers before the network was trained. The finish time was divided by the track record to provide a common denominator between all horserace. Linear regressions were done on sets of three consecutive races in order to rate the recent performance as improving, staying the same or getting worse. The final design included the postposition, the number of horses, the finish time as a percent of track record, and the change in recent performance as inputs. The output was the predicted finish time of the horse's next race expressed as a percentage of track records.{9}

NOTES: Times are expressed as a percentage of track records. The change in performance is recalculated with a linear regression equation is used to produce a slope which describes the horse's recent performances as improving, not changing or getting worse. While the network is being trained, the inputs (the left portion of the diagram) represent the information from a past race. The output is what the horse did at its next race. At least eight races worth of information are needed for training. The program presents the races to the network one a time, over and over, until the network learns what the horse has done in the past. Don typically let the training continue until the network output numbers, which were 95% accurate. The training can be done just before the race since it takes about five minutes to complete. After the network is trained, it can be used to predict an upcoming race Information about the horse's last race is input, and the output is a prediction of what the horse will do at this (the next) race. In order to yield the best results from this network, a race selection system was created:

1. No races with maiden horses or allowance races are used.
2. No races with more than nine horses are used.
3. Every horse must have at least eight past performances
4. Best odds for profit are with races that have at least a 37% long shot to favorite ratio.

3.2.3.2 Neural Network Predicts the Order of Finish in Horseracing

Twenty years ago, before he had access to computer technology, only imagined being able to predict the order of finish in a six furlong claiming race. As a teenager, his father dragged him to the races to perform the legwork of running money to betting windows. Over the years, Rich watched his father lose a small fortune. Armed with the belief that predicting the most probable horse in a race should be easier than putting a man on the moon, and having seen too much to be a gambler, he set to work on a solution. An engineer by profession, Mr. Janava worked on the problem for seven years using calculators and some early IBM 360 machines, but a major obstacle was the need for better technology enabling the required non-linear optimized solutions.

Four years ago after being exposed to neural net analysis, Rich discovered the tool required to accomplish his goal. Using neural network Competitor and Lotus macros, he has developed and automated a method for predicting the order of finish of six furlong claiming races at a Philadelphia racetrack. So far in the first 300 races, 39% of the winners have been predicted at odds, which average better than 4.5 to 1. The real power is in three horsebox exactas and four-horse box triple wagering which are both hitting better than 35% of the time. Rich says the neural network Competitor formulation is the only software he is aware of which sets up the fact files in the manner necessary to formulate and solve this type of problem. It also provides the convenient ranking of results for the many hundreds of races necessary to train and solve this problem.

According to Mr. Janava, the key to his success lies in the fact that he has limited his focus to one type of race and one racetrack. As any horseplayer knows, every type of race is affected by different racing variables. Every racetrack is also different. So far Janava has thoroughly explored and compared nine major racetracks. He explains, "If you try to generalize too many types of races and race tracks, you will lose the fine edge necessary to be successful at any race, distance, or track." In other words, the key to success is to become a specialist.

To make a generalized prediction, Mr. Janava analyzed hundreds of quality six-furlong races using the standard Racing Form paper. He started with an initial set of over 50 variables and determined that 24 variables were statistically significant. These include variables related to speed, horse position during previous races, class, earnings, recent activity, in-the-money percentages, and post position in today's and previous races. Several variables in each category are also combined to yield composite variables, which extend the total number of variables considered for each horse to 31.

After much thought, Rich derived an approach to calculating probability functions for each significant variable. For fact file generation and training, non-linear-regression-fitted probability values are used as inputs to Competitor. While using raw data values instead of probabilities for training has not been explored, Rich believes that raw data should work. In order to produce a sufficiently generalized network, Rich trains using at least 200 races.

Rich says that, had it not been for the neural network software, his efforts would have been nearly impossible. He intends to constantly refine his analyses with additional races during his yearly winter recess, and looks forward to a neural network Competitor version that runs race fact files in batch mode.

3.2.3.3 Selecting Winning Dogs with Neural Networks

Mr. Derek Anderson (Lakewood, CO) has trained neural networks that assist him in picking winning dogs at the racetrack. He trained the neural networks with two months of race results found in the daily racing booklets. Once trained, he runs the current day's race information through seven neural networks. He adds up the dogs' "scores" from his neural networks and places them in predicted finish order. Whenever the first place dog is ahead by at least ten neural network points over the second place dog, he bets on the winner. He claims 94% accuracy with this method, but he can bet on only a third of the races.

Mr. Anderson input information for approximately 300 races for the training file. The neural network looks at the statistics for three dogs at a time and outputs which of the three dogs did best. If there are eight dogs in a race, he must group the dogs in all possible combinations of three: dogs A, B and C; dogs A, B and D; dogs A, B and E; etc. For each race, there are 56 combinations, or sets of input data.

The data Mr. Anderson uses include the winning time of the race, the time that each dog took to finish the race, the time that dog reached each of four positions in the race (out of box, first corner, backstretch, outside corner) as well as comments about the dog's behavior. The behavior was classified as one of fifteen types such as ran wide, bumped, hit, and ran inside. He presented these pieces of information for each dog for each of the last eight races the dog ran. His networks have 504 inputs (21 statistics * 3 dogs * 8 races = 504).

Mr. Anderson has designed six basic neural networks with these 504 inputs. The difference in the six networks is the output. One neural network has three outputs which represent which dog is best: dog A, dog B or dog C. The dog, which was best, gets 1 the others get a 0. Another has three outputs, which represent which dog did worst of the three. The dog, which did the worst, gets a 0, the others get a 1. Another network has three outputs, which represent whether the dog was in the top three finishing positions. These dogs get a 1 the others get a 0. An opposite network outputs if the dog was not in the top three. Another pair of networks output whether the dog was in the last three to finish the race.

Because the dog racing information is not available in computer format, Derek spent a lot of time doing data entry. When it's time to predict a race, Derek runs the data through all of his networks and adds up the score for each dog. The scores range from 0 to 25 most of the time. The dog with the highest score is the winner.{9}

3.2.4 Science

3.2.4.1 Using a Neural Network to Measure Air Quality

Researchers Eugene Yee and Jim Ho at the Defense Research Establishment Suffield, Chemical & Biological Defense Section, in Alberta, Canada have trained a neural network to recognize classify and characterize aerosols of unknown origin with a high degree of accuracy. Their results hold considerable promise for applications to rapid real-time air monitoring in the areas of occupational health and air pollution standards.

Their research applied a neural network to the recognition and classification of environmental, bacterial, and artificial aerosols on the basis of the aerodynamic particle size distribution. Because of their variability, aerosols are difficult to recognize using conventional pattern recognition techniques. However, the health effects posed by airborne industrial, bacterial, and viral particles depend critically on the ability to recognize, characterize and classify these particles on the basis of their particle size distribution functions. The input data was constructed from aerodynamic particle size distribution functions (PSDF) obtained from 11 different aerosol populations. The PSDF's were measured with an aerodynamic particle size, which determines the aerodynamic diameter of individual aerosol particles, by measuring the transit time of the particles between two spots generated by a laser velocimeter that employs a polarized laser light source. Size distributions were classified into 11 categories depending on the source of the aerosol particles generating the distribution.

It was found that a recognition rate of 100% was obtained for the training set using neural networks with three or more hidden units and that there was a smaller number of passes through the training data with an increase in the number of hidden units in the network. There was virtually no increase in the learning times of the networks with more than 10 hidden neurons. In addition, the performance of the networks did not deteriorate when the number of hidden units was increased beyond 10.

Experiments were also conducted to study the performance characteristics of the neural network as a function of the quality of data used for the training set and the test set and of the inclusion of random noise in the connection strengths of the trained network. Results showed that the neural network was more suited than conventional methods for classification of signals from systems where one is confronted with ignorance of the statistical characteristics of the noise corrupting the signals.{10}

3.2.4.2 Using a Neural Network to Predict El Nino

Since January of 1991, a research team at the National Oceanic and Atmospheric Administration in Boulder, Colorado has been training a neural network to predict El Nino. According to head researcher, the purpose of the study was to determine if a neural network could accurately predict warm and cold events in the Pacific Ocean, and to compare the prediction capabilities of the neural network to other methods, particularly the Persistence method.

Researchers defined an El Nino or warm event as a departure of more than 1 standard deviation larger than the long term mean in various regions of the Eastern Pacific ocean. If the standard deviation was 1 standard deviation below the long term mean, it was a cold event. While the Persistence method is often used to make weather and climate predictions, it is unable to forecast change or predict the onset of a new situation. The neural network on the other hand was able to show a correlation between the prediction of El Nino and the actual occurrences of warm and cold events in the Pacific. The neural net proved to be a useful device for predicting out to about six months, and depending on the input data, could possibly be useful to the fishing industry.

Researchers used input data found in the Comprehensive Ocean Atmospheric Data Set **(COADS). COADS** are worldwide oceans data giving the sea surface on a monthly basis since 1884. Because warm events occur every five to seven years, and because each event is unique in terms of duration, onset and decay, the statistical character of each even is quite varied. As a result, an event is difficult to predict by any means.

One part of the research study was to determine the best set of data. "The set of data we used to predict things over the last year is probably not the ideal set, and we will be using a different set in the future." Because the team used most of the available data for training, only 10% of the data was left for validation and this remaining 10% may not even encompass a period in which El Nino occurred. The network was trained using the standard sigmoid transfer function. Using the genetic algorithm method, the team studied learning rates and tolerances to determine the best set for the data set they were using. They also varied the number of hidden neurons to determine the optimum number, but have not yet gone to more than 1 layer.

Testing went as follows: In January of 1991, the team started predicting skill scores actually the RMS differences between the actual ocean temperature and the predicted temperature for up to six months ahead Then in February of 1991, they again predicted (on the basis of current data) for 1 to 6 months ahead. They continued in this manner up until the present time. "Those were quite good in the sense that the RMS skill differences were in the order of less than a degree averaged over a long period." However one of the problems was that neither the Persistence method nor the neural net did a thorough job of predicting the onset of the warm or cold event. This fact leads to speculate that the data was not sufficient for the purpose and that it should include not only at the sea surface temperature in Region 4 of the Pacific, but also at least a nine or ten year the history of it. In the means time he plan to employ a rather unusual validation process. The team will train using all the data they have and then they're going to find skill scores for the same period of time again using all the data they have.{7}

3.2.4.3 Neural Network Predicts Rainfall

The need for accurate local rainfall prediction is readily apparent when considering the many benefits such information would provide for river control, reservoir operations, forestry interests, flash flood watches, etc. While the data required to make such predictions has been available for quite some time, the complex, ever-changing relationships among the data and its effect on the probability, much less the quantity, of rain has often proved difficult using conventional computer analysis. The use of a neural network, however, which learns rather than analyzes these complex relationships, has shown a great deal of promise in accomplishing the goal of predicting both the probability and quantity of rain in a local area to an accuracy of 85%.

Using neural network software (a hydro meteorologist from the National Weather Service in Fort Worth, Texas) has developed such a model. Nineteen meteorological variables (e.g. moisture, lift, instability, potential energy, etc.) were used to develop two networks for quantitative predictions one for the warm season and one for the cool season. Two additional networks for probability predictions were also generated. Another completely different program, written in C, was developed to allow both the quantitative and the probability networks to run simultaneously with the results appearing on the same computer monitor. Results to date have been outstanding. In the quantitative model, five categories were used to group the rainfall data (0.01 to 0.49 inches, 0.5 to 0.99 inches, 1.0 to 1.99 inches, etc.) Different tolerances were allowed for each range. For example, the tolerance for the first category was 0.2 inches while the tolerance for the higher categories ranged from 0.25 to 0.5 inches. Predictions for the quantitative models have been accurate in a range of 74% to 100% for the five categories with an overall accuracy of 83%. The probability model used the criteria that a prediction of 30% probability or higher had to correspond to a rainfall of 0.10 inches or more. Otherwise the network output would be considered in error. The accuracy achieved to date for this model is 94% which, when combined with the quantitative results, gives an overall accuracy of 85%. Sensitivity analysis was performed on the input variables to determine which had the most effect on the output. This will allow the developers to refine the models and improve the accuracy. Since there is six additional sites in Texas that will be included in future studies, means of further automating both the data gathering and neural network operations are being investigated to improve the cost and allow the technology to be used more economically. Although only two years of training and testing data were available, the results achieved to date are believed to be reliable and consistent enough to be used for forecasting guidance. Since this was the original goal of the project, the use of neural networks to predict local rainfall is now expanding to locations in other parts of the country.{8}

3.2.5 Industrial

3.2.5.1 Section Thickness

We shall adopt 1 micron as a reasonable section thickness. Sections this thick have a number of useful properties. First, an electron beam can penetrate such sections easily. Sections from .25 to .5 microns are now recommended for routine use with beam energies as low as 100 kev, and tissue sections ranging up to 10 microns thick have been examined (although the higher beam energies required are more expensive. Most current work is done at lower energies tens to hundreds of kev rather than a few mev). Second, large sections (on the order of one square centimeter) can be prepared fairly easily and significantly larger sections seem possible. Third, use of 1 micron sections avoids the use of thin sections (.1 micron or thinner) which are more fragile, more numerous, more difficult to section, and more prone to produce artifacts from buckling, warping, and tearing. {5}

3.2.5.2 Section Support

Sections are generally supported in an electron microscope on a grid of metal, which has holes in it the holes allow the beam of electrons to pass freely through the specimen. If the holes are too large, the section will collapse through them under its own weight. To prevent this, a continuous film of support material is sometimes used to add strength, although even very thin support films tend to blur and obscure the specimen. Even a section supported by a thin film will eventually collapse if the hole is large enough. A wide variety of grids with holes of different sizes and shapes are available. Slot-shaped holes are often used, and are commonly available with slot-widths ranging from over 1000 microns down to about 20 microns. Supporting films are usually used with slot-widths over 100 microns.

Clearly, if a section is laid out on a slot-grid only the parts over one of the slots can be viewed and so a large portion of the section is effectively lost. Three methods for avoiding this problem seem possible.

First, we could arrange matters so that an unsupported specimen would not collapse. Given the size of section we are considering (several centimeters) this approach is probably only feasible in a micro-gravity environment (for example, in an orbital facility). While this would clearly prevent the section from collapsing under its own weight, the additional cost of providing a laboratory in low earth orbit would be substantial. We will not consider this possibility further.

Second, we could arrange matters so that the hidden portion of the section is not of interest first pre- sectioning the specimen into 1 millimeter slices and then interleaving these with 1 millimeter "fill" slices could do this. The combined layered material (somewhat like Neapolitan ice cream) can then be embedded and sectioned. The resulting sections would have alternate 1-millimeter stripes of "fill" and tissue. If the sections were laid out on the slot grid so that the metal of the grid directly supported the "fill" while the tissue was over the slot, then all of the tissue could be examined. This seems to require fairly large slot widths (1 millimeter in this example) and doubles the volume required during later sectioning steps a minor disadvantage.

The third approach would be to move the section on the slot grid after the visible portions of the section had been examined. This is not normally done because present slots are wide enough to allow the full area of interest to be examined. There seems no reason in principle, however, why a 1-micron section could not be lifted from the surface of the slot grid and moved over.

There are many techniques that involve lifting a section off a glass slide (after viewing with a light microscope) and then re-embedding and re- sectioning it for viewing with the electron microscope. Simply moving a section (without re-embedding or re-sectioning) would seem to be an easier operation. This approach does not require large slots; more conventional slot widths of perhaps slightly more than 100 microns could be used and would provide good support.

While no one has yet demonstrated feasibility, there has been no pressing need to do so many possible mechanisms for lifting the section from the grid is possible. {5}

3.2.5.3 Feasibility of Large Sections

The logistics of converting a brain into a series of 1-micron sections that can be viewed under an electron microscope requires some thought. Current techniques can reliably produce 1 micron thick sections which are 12 by 16 millimeters in size which is "...about 200 times larger than those used in electron microscopy. There is today no great need to make larger sections even if they could be made; imaging them in an electron microscope and then analyzing the resulting images (by eye) would be tedious using current techniques. There are no theoretical barriers to producing larger sections, and there is no reason to presume the practical barriers cannot be dealt with no one has done so because no one has really wanted to. Many commercially available microtomes are mechanically accurate enough to produce 1-micron sections from blocks a few decimeters on a side. Suitable glass knives 40 mm long have been made and larger are quite possible. Recent work on diamond coatings should lead to very high quality low cost diamond coated knives (Sumitomo Electric of Japan has already made a tweeter with a 1 micron diamond coating). Because of the large size and the requirement that reproducible serial sections be produced, it seems likely that diamond knives will be required.

If the blade and microtone are suitable, then the only remaining obstacle must be the specimen block and large block faces can deform under the pressure of sectioning. The usual solution is to use a harder embedding media, and this has produced quite satisfactory results.

With care, it might be possible to extend this method to larger section sizes. An alternative method would be to provide direct bracing to the face of the specimen block being sectioned.

The principle should be familiar to anyone who has seen a meat slicer in a deli in which the meat slides along a flat plate, and is sectioned by a blade, which is parallel to and slightly above the plane of the flat plate. The meat itself is quite soft and could not be sectioned easily with a "free hand" knife blade because it would deform too easily, but with the aid of the meat-slicer it can be converted into very thin uniform sections with little difficulty. A similar design in microtones would replace the "free blade" and unsupported block face of the conventional microtone with an optically flat support block against which to lay the specimen, and an optically flat knife perhaps a fraction of a millimeter beyond the edge of the support block and 1 micron above the plain of the support block. The specimen block would then be slid along the support block and into the knife even a soft embedding should produce good results. {5}

3.8 Summary

This chapter showed how neural networks deals with some problem in different fields to explain it and solve it, for example in manufacturing field neural networks can be used to test the quality of some products, and in the other hand in science section Neural Networks can be used to Predicts rainfalls, and also we can use neural networks to solve many more problems.

CHAPTER FOUR

NEURAL MEDICAL SYSTEM

4.1 Overview

In this chapter the main aim is to demonstrate some basic and some important neural network applications in medicine, and it will show how much is neural networks useful in medical field as it is in some other fields, and it will discuss the clinical application also.

4.2 Biological Foundation of Neurocomputing

Neurocomputing involves processing information by means of changing the states of networks formed by interconnecting extremely large numbers of simple processing elements, which interact with one another by exchanging signals. Networks such as the one just described are called neural networks, in the sense that they represent simplified models of natural nerve or neural networks.

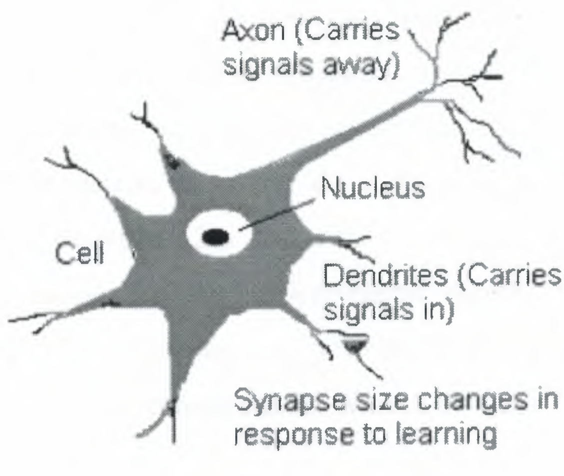


Fig. 4. 1. A simple Neuron Cell

The basic processing element in the nervous system is the neuron. The human brain is composed of about 1011 of over 100 types. Tree-like networks of nerve fiber called dendrites are connected to the cell body or soma, where the cell nucleus is located. Extending from the cell body is a single long fiber called the axon, which eventually branches into strands and sub strands, and is connected to other neurons through synaptic junctions, or synapses. The transmission of signals from one neuron to another at synapses is a complex chemical process in which specific transmitter substances are released from the sending end of the junction. The effect is to rise to lower the electrical potential inside the body of the receiving cell. If the potential reaches a threshold, a pulse is sent down the axon we then say the cell has "fired".{9}

4.3 Medical

4.3.1 Classify Breast Cancer Cells with Neural Network Software

A human, who decides the degree of cancer present, traditionally examines breast cancer cells under a microscope. People are inconsistent in these judgments from day to day and from person to person. A neural network that classifies breast cancer cells has been developed. Initial comparisons showed that neural network is in good agreement with human observer cancer classifications.

Cancer cells are measured with the CAS-100 (Cell Analysis System). There are 17 inputs to the neural network, which represent morph metric features such as density and texture. There are four network outputs representing nuclear grading. The cancerous nucleus is graded as being well, moderate, or poorly differentiated, or as benign. Correct grade assignments were made between 52% and 89% of the time on cases not seen during training the lower success rate (for well differentiated) may have been due to the smaller percentage of this type in the training set. In addition, heterogeneity is much lower in well-differentiated tumors. Cancerous nuclei were classified within one grade of the correct grade.{11}

4.3.2 Neural Network Improves Hospital Treatment

A new hospital information and patient prediction system has improved the quality of care, reduced the death rate and saved millions of dollars in resources at Anderson Memorial Hospital in South Carolina. The **CRTS/QURI** system uses neural networks trained with BrainMaker to predict the severity of illness and use of hospital resources. The **CRTS/QURI** system's goal is to provide educational information and feedback to physicians and others to improve resource efficiency and patient care quality.{11}

The first study showed that the program was directly responsible for saving half a million dollars in the first fifteen months even though the program only included half of the physicians and three diagnoses. Since then, the number of diagnoses and physicians included in the program has increased. The quality of care has improved such that there are fewer deaths, fewer complications, and a lower readmission rate. Expenses have been reduced by fewer unnecessary tests and procedures, lowered length of stays, and procedural changes. The reported success has motivated several other hospitals to join in the program and has provided the impetus to begin a quality program with the state of South Carolina.{10}

Individually trained neural networks learn how to classify and predict the severity of illness for particular diagnoses so that quality and cost issues can be addressed fairly. After attempts to use regression analysis to predict severity levels for several diagnoses failed, Epstein turned to the neural network program for a new approach and taught his neural networks to classify and predict severity with 95% accuracy. The neural networks are also used to predict the mode of discharge - routine through death - for particular diagnoses. Training information is based upon the length of stay in the hospital, which has a direct relationship to the severity of the illness (acuity). The neural network uses variables of seven major types:

1. Diagnosis.
2. Complications/co morbidity.
3. Body systems involved (e.g., cardiac and respiratory).

4. Procedure codes and their relationships (surgical or no surgical).

5. General health indicators (smoking, obesity, anemia, etc.).

6. Patient demographics (race, age, sex, etc.).

7. Admission category.

4.3.3 Diagnose Heart Attacks with Neural Network Software

When a patient complaining of chest pains is received by the emergency room, it is no simple matter to diagnose a heart attack. Merely examining the patient and performing an electrocardiogram is not often enough. If a patient is suspected of having experienced a heart attack, several blood samples are drawn over the next 4 to 48 hours. Patients with heart tissue damage will have various cardiac enzymes appear in their blood. There is a characteristic pattern of the change in enzyme levels during the short period after a heart attack. Using all three techniques (EKG, exam, and blood analysis), a doctor can diagnose and treat heart attack patients. Neural network methods were found to correlate closely with expert human analysis, providing another opinion doctors can use to make a correct and timely diagnosis.{10}

Michigan has designed a neural network that recognizes cases of acute myocardial infarction (AMI, commonly called heart attack) using the cardiac enzyme data from series of tests on patients. The input consisted of two sequential cardiac enzyme tests and the number of hours between the tests. The output was "1" if the patient had a heart attack and "0" if the patient did not. The network was trained with 185 examples from 47 patients using blood tests that were not more than 48 hours apart. There were a total of 21 inputs and 1 output as shown below. The network was trained to a 10% error tolerance on all training data.

The neural network was then tested on 53 new sets of data. The data represented sets of serial cardiac enzyme data for ten patients with AMI and eight patients without AMI. Neural network outputs of less than 10% probability of AMI were classified as no-AMI cases outputs of at least 90% probability of AMI were classified as AMI cases. Outputs between 10% and 89% were interpreted as ambiguous or uncertain. The neural network's diagnosis was then compared to three experts. One evaluated patients on the basis of **ECHO/EKG** changes. Another used the cardiac enzyme data. A third used autopsy reports. The network agreed with 100% of the AMI cases diagnosed by the cardiac enzyme expert, and 93% of the non-AMI cases. The 7% difference occurred where the network was uncertain. The network agreed with 86% of the AMI cases diagnosed by the EKG expert, and 33% of the non-AMI cases. In one case the EKG data was misleading due to multiple past heart attacks.{10}

In another case the network was uncertain. The network agreed with the autopsy results in 92% of the AMI cases, and 67% of the non-AMI cases. In one case the network was uncertain, and in another the heart had been damaged but by another cause.{10}

4.4 Neural Network Orders Medical Laboratory Tests for ER

When a patient appears at the emergency room door it is sometimes an hour until a doctor can see him or her. It may be another hour until the lab can do the ordered tests. In order to save patient waiting time, developed a neural network program that can order the lab tests as soon as the patient is admitted. Up to 38 labs the neural network can order tests.

Not only does the neural network save up to two hours of patient waiting time, it can reduce expenses. When the pilot system was tested, it reduced the number of tests that were ordered by 10-15%. It could save half a million dollars a year. The neural network is able to reduce the number of tests for two reasons. First, medical records are used as examples for training the neural network. It can be determined which tests were actually necessary in retrospect, so the neural network can be trained to order only the pertinent tests. Second, nurses had been given the blanket permission to order tests and they tended to order even more than doctors. The neural network has 67 inputs that include patient demographics and symptoms. This information is gathered when the patient is admitted and placed in the medical record. There are 38 outputs, each representing a different test that might be ordered. The pilot system neural network was trained on 250 patients from past hospital medical records. When the pilot system was tested on new patients it was found to be about 95% accurate, most of the time the neural network would order most of the necessary tests sometimes it did not order enough, but usually the doctor would only need to call the lab and order another test on the already collected specimen.{11}

4.5 Clinical Application

Patients with persistent, intractable pain, exemplified by low-back and lower-extremity pain after multiple surgeries are managed in several ways. Anatomical procedures may be used to correct an abnormality thought to cause the pain. Ablative procedures may be performed to destroy neural pathways that transmit pain signals. Augmentative procedures that superimpose electrical stimulation or medications that act on the intact nervous system may also be administered. Augmentative procedures, which are reversible and amenable to straightforward trial before devices are permanently implanted, are used increasingly because of their obvious advantages.

Spinal cord stimulation is a neuroaugmentative procedure: electrodes are implanted over the dorsal surface of the spinal cord and electrical pulses are delivered from an implantable generator. At stimulus amplitudes adequate to relieve pain, this produces a parenthesis. Overlap of the topography of a patient's pain by parenthesis is a necessary condition for pain relief, and there is a highly significant association between overlap and patient ratings of pain relief. The position of the implanted electrode(s) largely determines the location of stimulation parenthesis. To maximize the likelihood of achieving the desired overlap and to minimize the need for repeated surgical revisions of electrode position, manufacturers have developed arrays of multiple contacts.

As the number of contacts has increased, however, the task of adjustment (thoroughly testing the available anode and cathode assignments) has grown enormously. For a simple contact pair, there are only two assignments to test; for four contacts, there are 50; for eight contacts, there are 6050. The time required of the patient, the physician, and the staff to test these combinations is substantial. If adjustments are made manually, representing the data quantitatively for comparisons and retesting is cumbersome and becomes prohibitively time consuming. This has motivated the development of computerized methods {11}.

Even using computerized methods that require minimal patient supervision after initial training, the task of adjustment can become prohibitively time consuming if all possible combinations are to be tested exhaustively. For an eight-electrode array, even at a rate of 50 combinations per hour, it would take 121 hours to test all the possibilities. The performance of different electrode combinations may be assumed to be interdependent, and thus representative subsets may be defined and search strategies developed so that not all combinations need to be tested. Validating this assumption may be difficult for large numbers of contacts and for complex (two-dimensional) geometries. In the present study, the available 50 combinations for an array of four electrodes is a manageable number; they have been tested exhaustively, allowing validation of our models.

To date linear discriminate methods have yielded the finding that a particular electrode configuration ("guarded cathode" or "split anode," that is, cathode[s] flanked by anode[s] above and below) is preferred by our patients to a statistically significant degree.{8} In this prior work, as in the present study, the importance of overlap of pain by stimulation parenthesis in achieving the clinical goal of pain relief is assumed; this has been established before and is beyond the scope of the present study {11}.

4.6 Ethical Issues in the Use of Neural Network-based Methodologies for Image Interpretation in Medicine.

The concern about the ethical implications of the use of artificial intelligence techniques in medicine is ongoing. On the one hand, the use of artificial intelligence increasingly provides opportunities to facilitate and enhance the work of medical experts and ultimately to improve the efficiency and quality of medical care. On the other hand, the debate about the appropriate level and the role of intelligent decision support has become more complex, as technical, organizational and social issues become intertwined. In this paper we use a research project that applies neural network-based methodologies as an opportunity to study the ethical issues that may rise from the application of artificial intelligence techniques in a medical context.

Advances in neuron computing have opened the way for the establishment of decision support systems which are able to learn complex associations by example. It is acknowledged that the appropriate use of the neural network-based methodologies in medical problem solving could be very effective to improve the efficiency and the quality of medical care. The growing number of projects that employ neural network based methodologies in medical care makes necessary to examine the level and the role that neural networks will play in the development of automatic diagnostic systems.

This section considers the use of neural networks in image processing and knowledge representation in image interpretation systems from an ethical perspective. The diagnostic attributes of such systems are based on the result of the feature extraction procedure from digital images, in conjunction with previously available medical knowledge, in order to expand this knowledge. A key issue in the use of neural network methods in a medical application of this kind is that it is unclear how decisions are reached. Most neural networks suffer from the opaqueness of their learned associations. In medical applications, this black box nature may make clinicians reluctant to utilize a neural network application, no matter how great the claims made for its performance. Thus, there is need to enhance neural networks with rule extraction capabilities. In addition, it is necessary to examine how productivity can be increased and how quality can be assured. This examination addresses the specification of the problem, the development of appropriate representations for the network input and output information and the preparation of the training, testing and validation data.

Using the acquisition, processing, storage, dissemination and use stages of an information life cycle as a basis for the different stages for the development and use of neural network medical imaging applications, this paper elaborates on the broader ethical implications of the use of neural applications in medicine. These include issues of interpretation, coordination between the technology and the human expert, validation of results and professional responsibilities. Although we concentrate on the use of neural network-based methodologies, the ethical issues discussed in this paper are relevant to a broader spectrum of artificial intelligence and information technology applications in health care.

To illustrate this broader set of topics, the paper concludes with suggestions for future research in the area of medical informatics that can support ethical practice.

4.7 Simulation of mind

American scientists since 1940's. Attempting to understand the human brain, have developed mathematical and computer models called neural networks, which try to duplicate the computational power of the human nervous system. For every human behavior or perceptual activity vision, memory and language--the brain enlists dynamic interacting population of neurons (nerve cells) into coordinated activity to perform the specific task at hand. Dr. James. A. Anderson is a neural network pioneer who has made contributions to understanding mental computation using models based loosely on the architecture of the nervous system. His work over several decades, much of it funded by the National Science Foundation, has aided progress in the fields of cognitive science and neuroscience and has been useful for building "smart machine and in other order forms of artificial intelligence.{6}

With funding from the NSF's human cognitive and perception program, part of the BCS (behavioral, & cognitive science) division, Dr. Anderson, chair of Brown University's Department of Cognitive and Linguistic Sciences and colleagues have recently studied neural network modeling of human reaction times .For more than a century, psychologists have studied the patterns seen in the time it takes a person to produce an answer to a problem in an effort to better understand the details of mental operation .this venerable technique benefits from new combine with computer simulations .A major part of this project involves looking at the way humans solve simple arithmetic problems, a surprisingly difficult task both humans and artificial neural networks. Instead of actually computing the answers to problems the way a digital computer would, humans (and networks) seem to learn elementary arithmetic very differently, by memorizing facts and estimating answers. Although these human strategies can be error-prone and slow, natural extensions of them can give rise to the powerful but poorly understood faculty of mathematical intuition as well as the ability to reason "intuitively" about complex systems.{7}

Another NSF funded project in which Dr. Anderson is involved is funded through the LIS (Learning and Intelligent Systems) initiative which promotes studied on intelligence in humans, animals, and artificial systems. Anderson and colleagues are involved in study on "Adaptive Cortical Computation in the Visual Domain~~ which is investigating the long range spatial interactions among used in this study is called informally the "network of networks and is an attempt to bridge the huge gap in scale between single processing ailments (neurons) and entire brain regions that may contain hundreds of millions of cooperating neurons. This model suggests the neurons can work together to build larger and larger functional grouping and that larger, and even larger grouping may form the smaller ones using similar rules of formation.{11}

In this year of NSF-funded research James Anderson has proven repeatedly that, in his own words, "cognitive science can, in fact, be immensely practical in the right situation. "Network models similar to those he and his colleagues first developed with NSF support to simulate the human nervous system have become the foundation for the artificial neural networks now used routinely in many pattern based applications such as credit verification, medical diagnosis, and speech recognition. Anderson has also collaborated with companies such as Texas Instruments to improve military electronics.

One project required a means of analyzing a confusing flood of radar signal data .the radar data was processed by neural network designed to simplify the complex, as humans do, by breaking information into manageable blocks of data. Humans use ''concepts'' as away of simplifying and understanding a complex environment .the techniques used for radar analysis were based directly on neural network models for human concepts formation .he and colleagues at Distributed Data Systems, Inc have used this idea to develop a "smart" radar analysis system for the U.S Navy. The NSF has also supported Anderson and colleagues by providing facilities for these neural network simulations as well as for related work helping to understand the mechanics of the human mind.{9}

4.8 Classify Breast Cancer Cells

Breast cancer cells are traditionally examined under a microscope by a human, who decides the degree of cancer present. People are inconsistent in these judgements from day to day and from person to person.

A BrainMaker neural network that classifies breast cancer cells has been developed. The system was developed by Andrea Dawson, MD of the University of Rochester Medical Center, Richard Austin, MD of the University of California at San Francisco, and David Weinberg, MD, PhD of the Brigham and Women's Hospital and Harvard Medical School of Boston. Initial comparisons showed that BrainMaker is in good agreement with human observer cancer classifications.{12}

Cancer cells are measured with the CAS- 100 (Cell Analysis System, Elmhurst, IL). There are 17 inputs to the neural network which represent morphometric features such as density and texture. There are four network outputs representing nuclear grading. The cancerous nucleus is graded as being well, moderate, or poorly differentiated, or as benign. Correct grade assignments were made between 52% and 89% of the time on cases not seen during training. The lower success rate (for well differentiated) may have been due to the smaller percentage of this type in the training set. In addition, heterogeneity is much lower in well-differentiated tumors. Cancerous nuclei were classified within one grade of the correct grade.{12}

4.9 Diagnose Heart Attacks

When a patient complaining of chest pains is received by the emergency room, it is no simple matter to diagnose a heart attack. Merely examining the patient and performing an electrocardiogram (EKG) is not often enough. If a patient is suspected of having experienced a heart attack, several blood samples are drawn over the next 4 to 48 hours. Patients with heart tissue damage will have various cardiac enzymes appear in their blood.

There is a characteristic pattern of the change in enzyme levels during the short period after a heart attack. Using all three techniques (EKG, exam, and blood analysis), a doctor can diagnose and treat heart attack patients. Neural network methods were found to correlate closely with expert human analysis, providing another opinion doctors can use to make a correct and timely diagnosis.

A physician at St. Joseph Mercy Hospital in Michigan designed a neural network hat recognizes cases of acute myocardial infarction (AMI, commonly called heart attack) using the cardiac enzyme data from series of tests on patients. [9] The input consisted of two sequential cardiac enzyme tests and the number of hours between he tests. The output was "1"if the patient had a heart attack and "0" if the patient id not. The network was trained with 185 examples from 47 patients using blood tests that were not more than 48 hours apart. There were a total of 21 inputs and 1 output as shown below. The network was trained to a 10% error tolerance on all training data.{12}
The neural network was then tested on 53 new sets of data. The data represented sets of serial cardiac enzyme data for ten patients with AMI and eight patients without AJVII. Neural network outputs of less than 10% probability of AMI were classified as no-AMI cases. Outputs of at least 90% probability of AMI were classified as AMI cases. Outputs between 10% and 89% were interpreted as ambiguous or uncertain.{12}

The neural network's diagnosis was then compared to three experts. One evaluated patients on the basis of ECHO/EKG changes. Another used the cardiac enzyme data. A third used autopsy reports. The network agreed with 100% of the AMI cases diagnosed by the cardiac enzyme expert, and 9300 of the non-AMI cases. The 7% difference occurred where the network was uncertain. The network agreed with 86% of the AMI cases diagnosed by the EKG expert, and 33% of the non-A]\41 cases. In one case the EKG data was misleading due to multiple past heart attacks. In another case the network was uncertain. The network agreed with the autopsy results in 92% of the AMI cases, and 67% of the non-AMI cases. In one case the network was uncertain, and in another the heart had been damaged but by another cause.{12}

4.10 Rainmaker Orders Medical Laboratory Tests for ER

When a patient appears at the emergency room door it is sometimes an hour until a doctor can see him or her. It may be another hour until the lab can do the ordered tests. In order to save patient waiting time, Dr. Steven Berkov of Kaiser Hospital in Walnut Creek, California, developed a neural network program that can order the lab tests as soon as the patient is admitted. Up to 38 lab tests can be ordered by the neural network.{13}

Not only does the neural network save up to two hours of patient waiting time, it can reduce expenses. When the pilot system was tested, it reduced the number of tests that were ordered by 1 0-15%. Dr. Berkov says it could save half a million dollars a year. The neural network is able to reduce the number of tests for two reasons. First, medical records are used as examples for training the neural network. It can be determined which tests were actually necessary in retrospect, so the neural network can be trained to order only the pertinent tests. Second, nurses had been given the blanket permission to order tests and they tended to order even more than doctors. The neural network has 67 inputs that include patient demographics and symptoms. This information is gathered when the patient is admitted and placed in the medical record. There are 38 outputs, each representing a different test that might be ordered. The pilot system neural network was trained on *250* patients from past hospital medical records.

When the pilot system was tested on new patients it was found to be about *95%* accurate, according to Dr. Berkov. Most of the time the neural network would order most of the necessary tests. Sometimes it did not order enough, but Dr. Berkov explained that usually the doctor would only need to call the lab and order another test on the already collected specimen.

4.11 Prostate Cancer Detection

4.11.1 A New Way to Detect Prostate Cancer

Using a new information age technology to analyze very small changes in the chemistry of the human body, scientists have developed a new, reliable index that detects the presence prostate cancer cells before they start to spread to the pelvis and other organs that surround a man's prostate The new Risk Index included in ProstAsure® is a neural network-derived analysis of the changes in the body's biochemistry detected in human cells by tumor markers known to be associated with prostate cancer.{13}

The Risk Index has been shown to be clinically effective based on the analysis and study of the biochemistries of more than 2200 men. Its unique advantage is that it appears to be able to tell a physician, better than other procedures available for detecting prostate cancer, when a biopsy should be performed to diagnose this cancer at an early stage when the best chance of effective medical treatment can be provided. Physicians today are taking advantage of PSA and other new types of PSA-related tests to find cancers earlier than was possible even ten years ago. For this reason, and due to the availability of better therapies for treating cancer, the number of prostate surgeries dropped by half, to 116,000, in the past ten years (Monika Guttman, USA Weekend, June 11-13, 1999, plO).{10}

Being tested at the first sign of suspicious symptoms is very important and can be crucial in helping the doctor determine whether or not to perform a biopsy. It can help to catch the cancer while it is still confined to the prostate gland, before it spreads to other parts of the body.

4.11.2 A Better Way to Detect Prostate Cancer Early

Bioinformation technologies, specifically neural networks, facilitate the analysis of biological changes taking place deep inside of human cells long before cancer cells invade the domain of healthy cells to become dominant in the make-up of the prostate. This type of neural network-derived Index of risk examines groups of markers produced by a tumor to find specific patterns that may help confirm the activities and presence of diseased cells as they begin to grow in the prostate.

4.12 Electronic/Artificial Noses

Electronic/artificial noses are being developed as systems for the automated detection and classification of odors, vapors, and gases. An electronic nose is generally composed of a chemical sensing system (e.g., sensor array or spectrometer) and a pattern recognition system (e.g., artificial neural network). At Pacific Northwest National Laboratory, we are developing electronic noses based on artificial neural network technology for the automated identification of volatile chemicals for environmental and medical applications.

4.12.1 When Are Neural Networks Used?

Neural Networks have been applied to an increasing number of real-world problems of considerable complexity. Their most important advantage is in solving problems that are too complex for conventional technologies . problems that do not have an algorithmic solution or for which an algorithmic solution is too complex to be found. In general, because of their abstraction from the biological brain, Neural Networks are well suited to problems that people are good at solving but computers are not. These problems include pattern recognition and forecasting. However, unlike the human capability in pattern recognition, the Neural Networks capability is not affected by factors such as fatigue, working conditions, emotional state, and compensation.

4.12.2 Electronic/Artificial Noses Description

The figure below illustrates the basic schematic of an electronic nose. During operation, a chemical vapor or odor is blown over the sensor array, the sensor signals are digitized and fed into the computer, and the ANN (implemented in software) then identifies the chemical. The benefits of electronic noses include compactness, portability, real-time analysis, and automation.

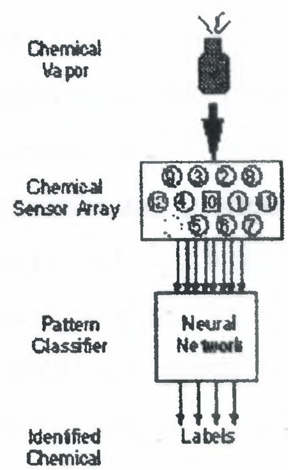


Figure 3.2: Schematic of an electronic nose

4.13 Picking up the Scent

4.13.1 Detecting illnesses in the body through an electronic, artificial nose

Imagine a physician detecting infection in a patient's body by smelling a wound, or a surgeon accessing injuries in a soldier located hundreds of miles away by smelling" the person's laceration through use of telemedicine and an artificial nose.

It's not as far fetched as you may think. In fact, extensive research on development of electronic, artificial noses is being conducted at the Pacific Northwest National Laboratory for use in the automated detection and classification of odors, vapors and gases for medical, environmental and industrial applications.

4.13.2 Giving the Physician a Sixth Sense

The electronic, artificial nose could prove to be a valuable diagnostic tool. Using a hand-held apparatus equipped with a sensor, the physician would scan for odors from the body, such as breath, wounds and body fluids, to identify possible problems. For example, odors in the breath can be indicative of gastrointestinal problems, sinus problems, infections, diabetes, and liver problems, Like wise, infected wounds and tissues emit distinctive odors that can be detected by an electronic nose. Odors coming from body fluids such as blood and urine can indicate liver and bladder problems.

Currently, an electronic nose for examining wound infections is being tested at South Manchester University Hospital. In similar applications, the technology has been used to track glucose levels in diabetics, determine ion levels in body fluids, and detect pathological conditions such as tuberculosis.

But, the medical applications don't end there. Researchers at Pacific Northwest are investigating a more futuristic use of the technology for telesurgery. In this application, the electronic nose would identify odors in the remote surgical environment. These identified odors would then be electronically transmitted to another site where an odor generation system would recreate them-giving the physician a better understanding of intemal injuries in the soldier.

4.13.3 The Nose Knows, Putting Technology to Work

An electronic nose is generally composed of a chemical sensing system, such as a sensor array or spectrometer, and a pattern recognition system like an artificial neural network. At Pacific Northwest, researchers are developing electronic noses based on neural network technology for the automated identification of volatile chemicals for environmental and medical applications.

Neural Network is an information processing paradigm that was inspired by the way biological nervous systems, such as the brain, process information. The key element of this paradigm is the novel structure of the information processing system. It's composed of a large number of highly interconnected processing elements, called neurons, working in unison to solve specific problems.

Neural Network like people, learn by example. A neural network is configured for an application such identifying chemical vapors through a learning process. Learning in biological systems involves adjustments to the synaptic connections that exist between the neurons. This is true of Neural Networks as well. For the electronic nose, the Neural Network leams to identify the various chemicals or odors by example. {13}

4.13.4 Benefiting Real-World Applications

Neural networks have been applied to an increasing number of real-world problems of considerable complexity. Their most important advantage is in solving problems that are too complex for conventional technologies-problems that do not have an algorithmic solution or for which an algorithmic solution is too complex to be found. In general, because of their abstraction from the biological brain, neural networks are well suited to problems that people are good at solving but computers are not.

These problems include pattern recognition and forecasting. However, unlike humans, neural networks are not affected by factors such as fatigue, working conditions, emotional state and compensation.

The neural network approach to information processing has several benefits that are applicable to electronic noses. They are trained by examples instead of rules, eliminate issues associated with human fatigue and habituation, are automated, and enable rapid identification. Neural networks also enable analysis of sensor values in real time.

4.13.5 Cross-Cutting Applications

Pacific Northwest also is applying electronic, artificial noses to perform environmental restoration and waste management in a cost-effective manner. These units are portable, inexpensive systems capable of real-time identification of contaminants in the field for: identification of toxic wastes, analysis of fuel mixtures, detection of oil leaks, identification of household odors, monitoring air quality and factory emission, and testing ground water for odors.

That same technology can be applied to the food industry as well. Electronic, artificial noses could be used to augment or replace panels of human experts or to reduce the amount of analytical chemistry that is performed in food production, especially when qualitative results are sufficient

4.13.6 Summary

This chapter described the medical neural network with a good examples to make easy to understanding by the readers, it explained too many things about the artificial nose and about the clinical applications and too many more things all ion medical field of neural networks applications.

CONCLUSION

My project is about "Applications of Neural Network". So that firstly we discussed a lot of points about neural network and also the subjects, which have related with the application of neural network and during our studies we found the definition of neural network is a massively parallel-distributed processor that has a natural propensity for experiential knowledge and making it available for use. And the kinds of neural network learning such as: supervised learning, unsupervised learning, Hebbian learning. And how neural network is related to statistical method and how neural network is better than statistical method that was in chapter two.

Successful work in elucidating the behavior of individual synapses has led to increased interest in networks of synapses. Tests of complex theories of network function require significant advances in our knowledge of the actual connectivity of real networks. Today, we can analyze only small numbers of neurons by hand and must infer the topology of large networks by indirect evidence. By automating the analysis process we can extend our knowledge to networks of significant size using currently available techniques and hardware. If we use the technology that will be available in 10 to 20 years, if we increase the budget to about one billion dollars, and if we use specially designed special purpose hardware then we can determine the structure of an organ that has long been of the greatest interest to all humanity, the human brain.

We can and should begin work on automated analysis of a "small" neural system today. Not only will it improve our too sketchy knowledge of real neural networks, but the understanding (and the software) that such preliminary efforts provide will be directly applicable to the more ambitious projects that will inevitably follow. As shown here, success on a small project can be scaled up to success on much larger projects up to and including the human brain..

Now we are about the main subject “Applications of Neural Network. Firstly we have to know what are the benefits of these applications? And if there differences between the old methods and new methods (by using neural network) In this project we made a lot of studies about applications of neural network and we found by using neural network every thing will be easily and you can do what ever you want because Neural networks cannot do anything that cannot be done using traditional computing techniques, but they can do some things, which would otherwise be very difficult. And neural network is based on the human brain design. So those neural networks have a lot of applications we wrote about four applications like business, sport, medical, and manufacturing. Now we would like to remind two examples for the application of neural network.

First example: about Medical

Neural Network Improves Hospital Treatment

A new hospital information and patient prediction system has improved the quality of care, reduced the death rate and saved millions of dollars in resources at Anderson Memorial Hospital in South Carolina. The **CRTS/QURI** system uses neural networks trained with BrainMaker to predict the severity of illness and use of hospital resources. The **CRTS/QURI** system's goal is to provide educational information and feedback to physicians and others to improve resource efficiency and patient care quality.

The first study showed that the program was directly responsible for saving half a million dollars in the first fifteen months even though the program only included half of the physicians and three diagnoses. Since then, the number of diagnoses and physicians included in the program has increased. The quality of care has improved such that there are fewer deaths, fewer complications, and a lower readmission rate. Expenses have been reduced by fewer unnecessary tests and procedures, lowered length of stays, and procedural changes. The reported success has motivated several other hospitals to join in the program and has provided the impetus to begin a quality program with the state of South Carolina.

Individually trained neural networks learn how to classify and predict the severity of illness for particular diagnoses so that quality and cost issues can be addressed fairly. After attempts to use regression analysis to predict severity levels for several diagnoses failed, Epstein turned to the neural network program for a new approach and taught his neural networks to classify and predict severity with 95% accuracy. The neural networks are also used to predict the mode of discharge - routine through death - for particular diagnoses. Training information is based upon the length of stay in the hospital, which has a direct relationship to the severity of the illness (acuity). The neural network uses variables of seven major types:

1. Diagnosis.

2. Complications/comorbidity.

3. Body systems involved (e.g., cardiac and respiratory).

4. Procedure codes and their relationships (surgical or nonsurgical).

5. General health indicators (smoking, obesity, anemia, etc.).

6. Patient demographics (race, age, sex, etc.).

7. Admission category.

Second example: about manufacturing

Using Neural Networks to Determine Steam Quality

Canada has developed the **INSIGHT** steam quality monitor, an instrument used to measure steam quality and mass Flowrate. Steam Quality and Mass Flowrate is the energy injected into the ground in an oil recovery project, for example.

The improvement obtained by using the trained network was immediately apparent using a conventional linear program, the standard error of estimate (RMS of deviations about the ideal line) for steam quality and mass Flowrate are 28% and 0.59 kg/s. using the trained neural network, the standard error was 8.2% and 0.34kg/s. A common test set of 26 sets of input data was used and the network was trained on an additional 100 facts.

Later, a similar network was trained and tested all of the **INSIGHT** monitor calibration data obtained to date (i.e. data from tests at four different facilities collected over a period of seven years using a minimum of six to a maximum of nine different monitors). Here, the standard error of estimate for steams quality and mass Flowrate were 7.7% and 0.4kg/s, respectively. Recently neural network has successes trained to return methanol, gasoline and water contents from the RF reflectance spectra of mixtures of these three components. Currently they are investigating the application of a neural network to a-spectroscopy and to the interpretation of on-line chemical sensor signals.

REFERENCES

[1] Academic writing from the World Wide Web:

http://www.pnl.gov

[2] Academic writing. From the World Wide Web:

http://www.pioneermed.gov

[3] Academic writing from the World Wide Web:

http://www.cs.stir.ac.uk

[4] Jones, Stephen, Neural Networks and the Computational Brain or Matters Relating to Artificial Intelligence.

[5] Some specific models of artificial neural nets, Lecture Notes.

[6] Cariani, P. "Emergence and Artificial Life." In C. G. Langton, C. Taylor, J. D. Farmer and S. Rasmussen, eds., Artificial Life II. Addison-Wesley. (1991).

[7] O'Reilly, R. C., & McClelland, J. L., "The Self-Organisation of Spatially Invariant Representations." Technical Report PDP.CNS.92.5. Department of Psychology, CMU.

[8] Rochester, N., J. H. Holland, et al. "Tests on a Cell Assembly Theory of the Action of the Brain, Using a Large Digital Computer." In J.A. Anderson and E. Rosenfeld, eds., 1988, Neurocomputing. Cambridge, Mass.: MIT Press. (1956).

[9] Sima, J., and Orponen, P. "Computing with continuous-time Liapunov systems," 33rd ACM STOC, 2001

[10]Frasconi P., Gori M., and Sperduti A., ``A Framework for Adaptive Data Structures Processing", IEEE Transactions on Neural Networks. *Vol. 9, No. 5, pp. 768-786, 1998.*

[11]Frasconi P., Goller C., Gori M., Kuchler A., and Sperduti A., ``From Sequences to Data Structures: Theory and Applications", in ``Dynamical Recurrent Networks", Eds. J. Kolen and S.C. Kremer, IEEE Press.

[12] de Dombal F.T., Leaper D.J., Staniland J.R., McCann A.P., Horrocks J.C. *Acute Abdominal Pain*, Brit. Med. J. 2: 9-13, (1972).

[13] Schmidt, G. Thews. 1983. Physiologie des Menschen. Springer-Verlag. Berlin –