NEAR EAST UNIVERSITY

Faculty of Engineering

Department of Computer Engineering

NEURAL NETWORK IN MEDICINE

Graduation Project
COM- 400

Student: Tarek Ahmad (990971)

Supervisor: Assoc.Prof.Dr.Adnan Khashman

Nicosia - 2003

ACKNOWLEDGMENT

It takes more than just a student to write a graduation project, especially a project like "Neural Network in Medicine". So I would like to take a moment to acknowledge who contributed their talent and expertise to make this project.

I am indebted to my supervisor Dr.Adnan Khashman, thank you for your suggestions and advice, your patience and persistence, your hard work and your perseverance.

My greatest thanks to my parents, who gave me the courage to get my education supported me in all achievement. My father was always and is now an example to follow as an educator. My mother who sacrificed some success in her life to give me the opportunity to work in this project, also my siblings Nedal, Nesren, Sheren, and Mohammad, who contributed to teaching a lot in their lives; were always encouraging me to succeed in achieving high goals.

I am very grateful to my university, and engineering department, the family which accepted me in 1999. I express my gratitude to all university, faculty and staff with whom I have a good fortune.

Last but not last; I must convey my utmost gratitude to all friends of mine and energy to help with this project.

ABSTRACT

Artificial neural networks have been used in a number of different ways in medicine and medically related fields. This project using neural networks in medicine. This involved the use to solve optimization and dynamical control problems. A general framework for artificial neural networks models is introduced first. Then the main feed-forward and feedback models are presented a number of theoretical and practical aspects of the application of neural networks are presented in this project. Firstly the biological neuron is presented and utilized to try and determine the relationship between the artificial neural network and human secondly, conditions relating to network learning by supervised and unsupervised are examined, with the finding that both the number of hidden nodes, network architecture and the initial conditions of the network are important in determining if a neural network will learn a particular problem.

The next involves attempting to concept a diagnosis prediction via an artificial neural network knowledge base. The aim of this approach the major application of medical information has been the metabolic disease diagnosis. Unfortunately this aim is not realized within the realms of this project, due to problems in training the neural networks, the final aim involves a much more complex system, medical diagnostic aides. This acts as further confirmation of the limitations in the neural network in medicine.

TABLE OF CONTENTS

ACKNOWLEDGEMENT I
ABSTRACT ii
TABLE OF CONTENTS iii
INTRODUCTION 1
CHAPTER ONE: INTRODUCTION TO NEURAL NETWORKS 3
- 1.1. Overview 3
- 1.2. Artificial Neural Network 3
- 1.3. A Brief History of Neural Network 4
- 1.4. What is a Neural Network? 6
- 1.5. Why Use a Neural Network? 7
- 1.6. Are There Any Limits to Neural Networks? 8
- 1.7. Artificial Neurons and How They Work 8
- 1.8. The Biological Neuron 10
- 1.9. Analogy between Human and Artificial Neural Network 12
- 1.10. Summary 13

CHAPTER TWO: THE STRUCTURE OF NEURAL NETWORK 14
- 2.1.Overview 14
- 2.2. Architecture of Neural Network 14
  - 2.2.1. Feed-forward Network 14
  - 2.2.2. Feedback Networks 14
- 2.3. Perceptron 15
  - 2.3.1. History of Perceptron 16
- 2.4. Teaching an Artificial Neural Network 17
  - 2.4.1. Supervised Learning 17
  - 2.4.2. Unsupervised Learning 18
- 2.5. Design a Neural Network 19
  - 2.5.1. Layers 20
  - 2.5.2. Communication and Types of Connection 21
    - 2.5.2.1. Inter-Layer Connection 21
    - 2.5.2.2. Intra-Layer Connection 22
  - 2.5.3. Learning 22
    - 2.5.3.1. Off-Line or On Line 23
    - 2.5.3.2. Learning Laws 24
- 2.6. The Back Propagation Algorithm 25
- 2.7. What can you do With an N.N. and What not? 26
- 2.8. Where Are Neural Network Being Used? 28
- 2.9. Summary 28

CHAPTER THREE: NEURAL NETWORK IN MEDICINE 30
3.1. Overview 30
3.2. Integration of Neural Networks and Knowledge-Based System in Medicine 30
3.3. Artificial Neural Network Based Cardiovascular Modeling 32
3.4. Clinical Application 34
3.5. Ethical Issues in the Use of Neural Network-Based Methodologies 35
3.6. A Neural Network Model for Metabolic Disease Diagnosis 36
3.7. Neural Network for Plasma Torch Operation and Process Control 37
3.8. Application of Neural Network in Immunology 38
3.9. Summary 40
CHAPTER FOURE: MEDICAL DIAGNOSTIC AIDES 41
4.1.Overview 41
4.2. Medical Analysis and Diagnosis by Neural Network 41
4.3. Diagnosis by Growing Neural Network 43
4.4. Diagnosis by Rule Based Networks 44
4.5.Modeling and Diagnosing the Cardiovascular System 45
4.6. Neural Networks in Diagnostic Decision Support Systems 46
4.7. Neural Network Model for Metabolic Disease Diagnosis 49
4.8.Summary 50
CONCLUSION 51
REFERENCES 52

INTRODUCTION

Let us first summarize the most important features of the neural networks found in the brain. Firstly the brain contains many billions of very special kinds of cell - these are the nerve cells or neurons. These cells are organized into a very complicated intercommunicating network. Typically each neuron is physically connected to tens of thousands of others. Using these connections neurons can pass electrical signals between each other. These connections are not merely on or off - the connections have varying strength which allows the influence of a given neuron on one of its neighbors to be either very strong, very weak (perhaps even no influence) or anything in between. Furthermore, many aspects of brain function, particularly the learning process, are closely associated with the adjustment of these connection strengths. Brain activity is then represented by particular patterns of firing activity amongst this network of neurons. It is this simultaneous cooperative behavior of very many simple processing units which is at the root of the enormous sophistication and computational power of the brain.

Artificial neural networks are modeled after the brain. They typically consist of many hundreds of simple processing units which are wired together in a complex communication network. Each unit or node is a simplified model of a real neuron which fires (sends off a new signal) if it receives a sufficiently strong input signal from the other nodes to which it is connected. The strength of these connections may be varied in order for the network to perform different tasks corresponding to different patterns of node firing activity. This structure is very different from traditional computers.

Chapter one is devoted to artifical neural network introduction of the general methods from neural network theory, and provide a breif history of neural network. It also introduces artifical neurons and how it works.

Chapter two studies the structure of neural network. It introduces architecture of neural networks ,this view feed-forward and leads to feedback neural networks. The last subsection of this chapter introduces the supervised neural network and unsupervised neural networks.

Chapter three is specialized in medical application and describe some fields where we can find the neural network in medicine and introduces artificial neural network based cardiovascular modeling.

Chapter four studies the medical diagnostic aids , i.e. Medical analysis and diagnosis by neural networks. and introduced with neural network model for metabolic disease diagnosi. Furthermore it is shown modiling and diagnosing the cardivascular systems.

The aim of this project are:

- To retrieve some information about the artifical neural network.
- To investigate and learn about neural network,the emergence,structure and application in real ;life.
- To form on medical application of neural network.
- To demonstrate a medical neural network application in real life,namely medical diagnostic aides.

CHAPTER ONE

INTRODUCTION TO NEURAL NETWORK

1.1 Overview

Neural networks are an information processing technique based on the way, in this chapter will present an introduction to artificial neural network and provide a brief history of neural network and describe why use a neural network, in this section which is the introduction of neural network, I am going to explain artificial neurons and how they work and introduce a hint of the of the biological neurons.

1.2 Artificial Neural Network

An artificial neural network (ANN) is an information-processing paradigm inspired by the way the densely interconnected, parallel structure of the mammalian brain processes information. Artificial 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 ANN 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 ANNs 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 ANNs have been around since the late 1950s [1], it wasn't until the mid-1980s, that algorithms became sophisticated enough for general applications. Today ANNs 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 and character and signal recognition, as well as functional prediction and system modeling where the physical processes are not understood or are highly complex. ANNs 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 ANNs lies in their resilience against distortions in the input data and their ability to learn. 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.

1.3 A Brief History of Neural Network

Neural network simulations appear to be a recent development. However, this field was established before the advent of computers, and has survived at least one major setback several areas. Many important advances have been boosted by the use of inexpensive computer emulations. Following an initial period of enthusiasm, the field survived a period of frustration and disrepute. During this period when funding and professional support was minimal, important advances were made by relatively few researchers. These pioneers were able to develop convincing technology which surpassed the limitations identified by Minsky and Papert. Minsky and Papert, published a book (in 1969) in which they summed up a general feeling of frustration (against neural networks) among researchers, and was thus accepted by most without further analysis. Currently, the neural network field enjoys a resurgence of interest and a corresponding increase in funding.

First Attempts: There were some initial simulations using formal logic. McCulloch and Pitts (1943) [2], developed models of neural networks based on their understanding of neurology. These models made several assumptions about how neurons worked. Their networks were based on simple neurons which were considered to be binary devices with fixed thresholds. Another attempt was by using computer simulations. Two groups (Farley and Clark, 1954; Rochester, Holland, Haibit and Duda, 1956) [3]. The first group IBM researchers maintained closed contact with neuroscientists at McGill University. So whenever their models did not work, they consulted the neuroscientists.

This interaction established a multidisciplinary trend which continues to the present day, but psychologists and engineers also contributed to the progress of neural network simulations. Rosenblatt (1958) stirred considerable interest and activity in the field when he designed and developed the Perceptron. The Perceptron had three layers with the middle layer known as the association layer. This system could learn to connect or associate a given input to a random output unit. Another system was the ADALINE (ADAptive Linear Element) which was developed in 1960[4], by Widrow and Hoff of Stanford University. The ADALINE was an analogue electronic device made from simple components. The method used for learning was different to that of the Perceptron; it employed the Least-Mean-Squares (LMS) learning rule.

In 1969 Minsky and Papert wrote a book in which they generalized the limitations of single layer Perceptrons to multilayered systems. In the book they said: "...our intuitive judgment that the extension (to multilayer systems) is sterile". The significant result of their book was to eliminate funding for research with neural network simulations. The conclusions supported the disenchantment of researchers in the field. As a result, considerable prejudice against this field was activated.

Klopf (A. Henry Klopf) in 1972 developed a basis for learning in artificial neurons based on a biological principle for neuronal learning called heterostasis. Werbos (Paul Werbos 1974) developed and used the back-propagation learning method, however several years passed before this approach was popularized. Back-propagation nets are probably the most well known and widely applied of the neural networks today. In essence, the back-propagation net. Is a Perceptron with multiple layers, a different threshold function in the artificial neuron, and a more robust and capable learning rule. A Mari (A. Shun-Ichi 1967) was involved with theoretical developments: he published a paper which established a mathematical theory for a learning basis (error-correction method) dealing with adaptive pattern classification. While Fukushima (F. Kunihiko) developed a step wise trained multilayered neural network for interpretation of handwritten characters. The original network was published in 1975 and was called the Cognitron.

Progress during the late 1970s and early 1980s was important to the re-emergence on interest in the neural network field. Several factors influenced this movement. For example, comprehensive books and conferences provided a forum for people in diverse fields with specialized technical languages, and the response to conferences and publications was quite positive. The news media picked up on the increased activity and tutorials helped disseminate the technology. Academic programs appeared and courses were introduced at most major Universities (in US and Europe). Attention is now focused on funding levels throughout Europe, Japan and the US and as this funding becomes available, several new commercial with applications in industry and financial institutions are emerging.

Today, significant progress has been made in the field of neural networks-enough to attract a great deal of attention and fund further research. Advancement beyond current commercial applications appears to be possible, and research is advancing the field on many fronts. Neurally based chips are emerging and applications to complex problems developing. Clearly, today is a period of transition for neural network technology.

1.4 What is a Neural Network?

First of all, when we are talking about a neural network, we should more properly say "artificial neural network" (ANN), because that is what we mean most of the time. Biological neural networks are much more complicated than the mathematical models we use for ANNs. But it is customary to be lazy and drop the "A" or the "artificial". An Artificial Neural Network (ANN) is an information processing paradigm that is 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 is composed of a large number of highly interconnected processing elements (neurons) working in unison to solve specific problems. ANNs, like people, learn by example. ANN is configured for a specific application, such as pattern recognition or data classification, through a learning process. Learning in biological systems involves adjustments to the synaptic connections that exist between the neurons. This is true of ANNs as well.

A neural network is a system composed of many simple processing elements operating in parallel whose function is determined by network structure, connection strengths, and the processing performed at computing elements or nodes, a neural network is a massively parallel distributed processor that has a natural propensity for storing experiential knowledge and making it available for use. It resembles the brain in two respects:

1. Knowledge is acquired by the network through learning, exactly what is a neural network?
2. Interneuron connection strengths known as synaptic weights are used to store the knowledge.

ANNs 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, ANNs are well suited to problems that people are good at solving, but for which computers are not. These problems include pattern recognition and forecasting (which requires the recognition of trends in data).

1.5 Why Use a Neural Network?

Neural networks, with their remarkable ability to derive meaning from complicated or imprecise data, can be used to extract patterns and detect trends that are too complex to be noticed by either humans or other computer techniques. A trained neural network can be thought of as an "expert" in the category of information it has been given to analyze. This expert can then be used to provide projections given new situations of interest and answer "what if" questions. Other advantages include:

Adaptive learning: An ability to learn how to do tasks based on the data given for training or initial experience.

Self-Organization: An ANN can create its own organization or representation of the information it receives during learning time.

Real Time Operation: A.N.N. computations maybe carried out in parallel,and special hardware devices are being designed and manufactured which take advantage of this.

1.6 Are There Any Limits to Neural Networks?

The major issues of concern today are the scalability problem, testing, verification, and integration of neural network systems into the modern environment. Neural network programs sometimes become unstable when applied to larger problems. The defense, nuclear and space industries are concerned about the issue of testing and verification. *The mathematical theories used to guarantee the performance of an applied* neural network are *still under development. The solution for the time being may be to* train and test these intelligent systems much as we do for humans. Also there are some more practical problems like:

- The operational problem encountered when attempting to simulate the parallelism of neural networks. Since the majority of neural networks are simulated on sequential machines, giving rise to a very rapid increase in processing time requirements as size of the problem expands. Solution: implement neural networks directly in hardware, but these need a lot of development still.
- Instability to explain any results that they obtain. Networks function as "black boxes" whose rules of operation are completely unknown.

1.7 Artificial Neurons And How They Work?

A neural net is a physical (as in electronics) or virtual (a computer program) collection of nodes or neurons each in some way connected to the other. Each neuron has several inputs and several outputs. Input starts out as the message from an array of sensors. This message is often passed through associate nets which, in a vision system, do a lot of the pre processing of the signal before it is passed on to neurons based on the McCulloch and Pitt's model.

Warren McCulloch and Walter Pitts were a team of neurophysiologist and logician who in 1943 built a model involving resistors and amplifiers which mimicked what was known about natural neurons. Neurons take weighted inputs and then, depending on the result, either fire or do not. This firing is then passed onto several other neurons which take this input and, according to the weighting, act or do not act. The whole model is a network of interconnected cells, each affecting the next.

Eventually the signal in a neural net reaches an output stage. This can be a value (male, female) or an array of output (sound or a picture). At first this result will be near random until the net has been trained, and trained correctly. The net has to receive enough information through the input to be able to make the correct assumptions. For example, one neural net was being used by the military to aid the recognition of tanks. A net was given different pictures of tanks and had to decide whether they were Russian or American. Each time the net got it wrong the net would learn and reorganize its connections and weights. Eventually the net was achieving perfect results. Other pictures from the same set as the training photographs where also correctly sorting into American and Russian. Problems arose when a new set of photos of the same tanks was given to the machine. This time it went back to making mistakes. This was puzzling until someone pointed out the times of the days the photos were taken. The shadows on the trees and tanks fell at a different angle on the American photos and the Russian photos so the computer was sorting the photos by time of day and not by shape of tank. After different sets of photos were used at varying times of day the net learnt the error of its ways and went back to being correct most of the time.

This could be said to be an example of the old adage, garbage in, garbage out. In this case it shows that for a correct assumption to be reached by man or machine all the correct information must be available. If there are inconsistencies or we cannot work out why we get an answer we put it down to common sense. This is the same as neural nets. The weighting of neurons leads to the guessing of answers some of the time, using information known to fill in the gaps.

In the future computers may be a hybrid of Neural net and conventional Turing based computing. Conventional computing has the advantage of being logical and fast in known mathematical problems. Neural nets are not good at number crunching, much as the human brain finds sums harder to handle that music. Instead they excel in pattern recognition, in tasks that require filtering and analyzing data.

What should be pointed out is that current neural networks are about as intelligent as a stupid insect. Neural computing, despite its history, is still a young subject and yet has to be fully understood with true precision. Having said that it has already produced complex results and is being used in many different fields. In the future HAL type computers could be totally possible. Our emotions are our motivation for the things we do; with an artificial intelligence these motivations might be totally different. As with the rest of a life, there is no reason why we our creations should have to take natures.

Future man made intelligences may live their entire lives in environments alien to the human mind. They could exist in different bodies and spend their time thinking about things we would deem unimportant. They could be specialist intelligences possibly not directly comparable to our intelligence. For an example of how different a possible intelligence can be you only need look at the second closest intelligence on the planet, dolphins. The dolphins' and whales' worlds are radically different; it is a world of the oceans. We now know that dolphins and whales have a symbolic language, if a lot simpler than human language. Dolphins give each other name but, like whales, spend time on navigation as we spend on trying to manipulate tools. Different environments put different priorities on a creature living in it; it is currently hard for us to imagine what form of intelligence a creature living purely in the data sphere would take.

1.8 The Biological Neuron

The brain is a collection of about 10 billion interconnected neurons. Each neuron is a cell that uses biochemical reactions to receive process and transmit information, within humans there are many variations on this basic type of neuron, further complicating man's attempt at electrically replicating the process of thinking. Yet, all natural neurons have the same four basic components; the power at the human mind comes from the sheer numbers of these basic components and the multiple connections between them. These components are known by their biological names- dendrites, nucleus, axon, and synapses. Figure 1.1 shows schematic of biological neuron.

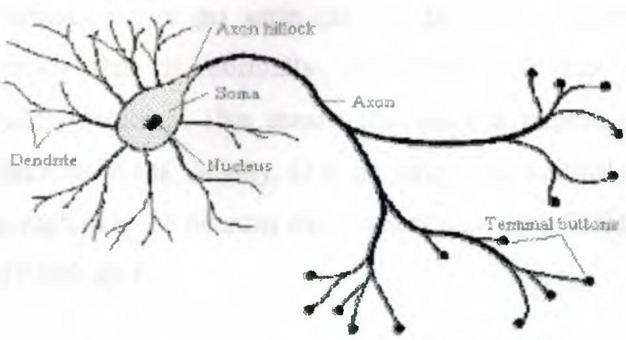

Figure 1.1 Schematic of biological neuron

A neuron's dendrites tree is connected to a thousand neighboring neurons. When one of those neurons fire, a positive or negative charge is received by one of the dendrites. The strengths of all the received charges are added together through the processes of spatial and temporal summation. Spatial summation occurs when several weak signals are converted into a single large one, while temporal summation converts a rapid series of weak pulses from one source into one large signal. The aggregate input is then passed to the soma (cell body). The soma and the enclosed nucleus don't play a significant role in the processing of incoming and outgoing data. Their primary function is to perform the continuous maintenance required to keep the neuron functional. The part of the soma that does concern itself with the signal is the axon hillock. If the aggregate input is greater than the axon hillock's threshold value, then the neuron fires, and an output signal is transmitted down the axon. The strength of the output is constant, regardless of whether the input was just above the threshold, or a hundred times as great. The output strength is unaffected by the many divisions in the axon; it reaches each terminal button with the same intensity it had at the axon hillock. This uniformity is critical in an analogue device such as a brain where small errors can snowball, and where error correction is more difficult than in a digital system.

Each terminal button is connected to other neurons across a small gap called a synapse. The physical and neurochemical characteristic of each synapse determines the strength and polarity of the new input signal. This is where the brain is the most flexible, and the most vulnerable. Changing the constitution of various neuron-transmitter chemicals can increase or decrease the amount of stimulation that the firing axon imparts on the neighboring dendrite. Altering the neurotransmitters can also change whether the stimulation is excitatory or inhibitory. Many drugs such as alcohol and LSD have dramatic effects on the production or destruction of these critical chemicals. The infamous nerve gas sarin can kill because it neutralizes a chemical (acetyl cholinesterase) that is normally responsible for the destruction of a neurotransmitter (acetylcholine). This means that once a neuron fires, it keeps on triggering all the neurons in the vicinity. One no longer has control over muscles, and suffocation ensues, each of these neurons can connect with up to 200,000 other neurons, although 1,000 to 10,000 are typical.

1.9 Analogy Between Human and Artificial Neural Network

Artificial neural nets were originally designed to model in some small way the functionality of the biological neural networks which are a part of the human brain. Our brains contain about 10^{11}neurons. Each biological neuron consists of a cell body, a collection of dendrites which bring electrochemical information into the cell and an axon which transmits electrochemical information out of the cell.

A neuron produces an output along its axon ie it fires when the collective effect of its inputs reaches a certain threshold. The axon from one neuron can influence the dendrites of another neuron across junctions called synapses. Some synapses will generate a positive effect in the dendrite, ie one which encourages its neuron to fire, and others will produce a negative effect, ie one which discourages the neuron from firing. A single neuron receives inputs from perhaps 10^5synapses and the total number of synapses in our brains may be of the order of 10^{16}. It is still not clear exactly how our brains learn and remember but it appears to be associated with the interconnections between the neurons (ie at the synapses).

Artificial neural network try to model this low level functionality of the brain. This contrasts with high level symbolic reasoning in artificial intelligence which tries to model the high level reasoning processes of the brain. When we think we are conscious of manipulating concepts to which we attach names (or symbols) e.g for people or objects. We are not conscious of the low level electrochemical processes which are going on underneath. The argument for the neural net approach to AI is that, if we can model the low level activities correctly, the high level functionality may be produced as an emergent property.

A single software artificial neuron consists of a processing element which has a number of input connections, each with an associated weight, a transfer function which determines the output, given the weighted sum of the inputs, and the output connection itself.The network may be trained by adjusting the weights associated with the connections in the net to try and obtain the required outputs for given inputs from a training set. Note that the threshold values and the weights can be adjusted together by adding an extra connection to each neuron with an input value of -1 and a weight representing the threshold. The neuron then fires if the sum is greater than zero. It can be seen from the above that there is an analogy between biological (human) and artificial neural nets. The analogy is summarised below.

Table 1.1 Analogy between human and artifical neuron

Human	Artifical
Neuron	Processing Element
Dendrites	Combining Function
Cell Body	Transfer Function
Axons	Element Output
Synapses	Weights

However, it should be stressed that the analogy is not a strong one. Biological neurons and neuronal activity are far more complex than might be suggested by studying artificial neurons. Real neurons do not simply sum the weighted inputs and the dendritic mechanisms in biological systems are much more elaborate.

1.10 Summary

This chapter presented definitions of artificial neural networks, and provided brief history of neural networks since 1943 until today. And after that I have moved to give a hint about how artificial neurons they work? And I have described a different analogy between biological (human) and artificial neural networks.

CHAPTER TWO

THE STRUCTURE OF NEURAL NETWORK

2.1 Overview

This chapter has a number of objectives. First i want to introduce the architecture of neural networks and, I define what is the meaning of learning rule, explain the perceptron network and its learning rule, and tell you how to initialize and simulate perceptron networks, also covers supervised learning and unsupervised learning networks, and provide designing a neural network, and given a application of neural networks.

2.2 Architecture of Neural Networks

In this section which is the Architecture Neural Network, am going to give a hint about how Neural Network has been designed, and how the signals travel and what kind of networks does it have. Here Neural Network is been classified in two categories, which are Feed-forward Network and Feed back Network.

2.2.1 Feed-forward Networks

Feed-forward A.N.Ns (figure 2.1) allows signals to travel one way only, from input to output. There is no feedback (loop) i.e. the output of any layer does not affect.

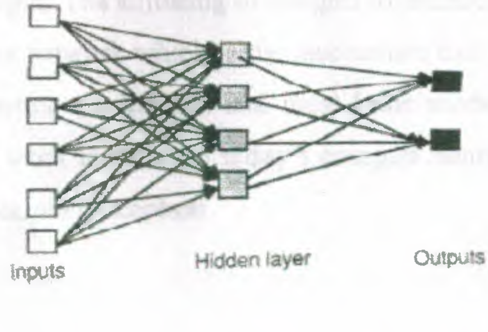

Figure 2.1 A simple feed-forward network.

2.2.2 Feedback Networks

Feedback networks can have signals traveling in both directions by introducing loops in the network. Feedback networks are very powerful and can get extremely complicated. Feedback networks are dynamic; their 'stat' is changing continuously until they reach an equilibrium point. They remain at the equilibrium point until the input change and a new equilibrium needs to be found. Feedback architectures are also referred to as interactive or recurrent, although the latter term is often used to denote feedback connections in single-layer organizations.

2.3 Perceptron

In the early days of artificial intelligence research, Frank Rosenblatt devised a machine called the perceptron that operated much in the same way as the human mind. Albeit, it did not have what could be called a "mental capacity", it could "learn" - and that was the breakthrough needed to pioneer today's current neural network technologies.

For a long time since perceptrons were discovered, neural networks were looked up to as powerful machines that can be trained to perform various tasks. The proof of convergence, that the perceptron will not only find suitable weights but that it will find it in finite time, was very appealing.

A perceptron is a connected network that simulates an associative memory. The most basic perceptron is composed of an input layer and output layer of nodes, each of which are fully connected to the other. Assigned to each connection is a weight which can be adjusted so that, given a set of inputs to the network, the associated connections will produce a desired output. The adjusting of weights to produce a particular output is called the "training" of the network which is the mechanism that allows the network to learn. Perceptrons are among the earliest and most basic models of artificial neural networks, yet they are at work in many of today's complex neural net applications see figure (2.2) describe the simple perceptron.

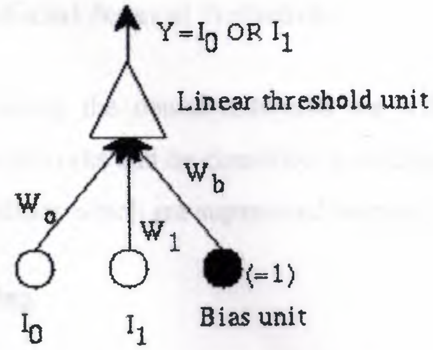

Figure 2.2 The simple perceptron

2.3.1 History of Perceptron

Frank Rosenblatt invented the perceptron in 1957 at the Cornell Aeronautical Laboratory in an attempt to understand human memory, learning, and cognitive processes. On 23 June 1960, he demonstrated the Mark I Perceptron, the first machine that could "learn" to recognize and identify optical patterns.

Rosenblatt's work was a progression from the biological neural studies of noted neural researchers such as D.O. Hebb and the works of Warren McCulloch and Walter Pitts. McCulloch and Pitts had been the first to describe the concept of neural networks. They developed the MP neuron, which was based on the point that a nerve will fire an impulse only if its threshold value is exceeded. This model was somewhat of a scanning device which read pre-defined input- output associations to determine its final output [4]. MP neurons had fixed thresholds and did not allow for learning. They were "hard-wired logic devices, [which] proved that networks of simple neuron-like elements could compute "[5].

Since the MP neuron did not have the mechanisms for learning, it was extremely limited in modeling the functions of the more flexible and adaptive human nervous system. D.O. Hebb suggested that "when an axon of cell A is near enough to excite cell B and repeatedly, or persistently, takes part in firing it, some growth process or metabolic change takes place in one or both cells, such that A is efficiency as one of the cells firing B is increased" [6]. This implied a "learning" network model where not only could the network make associations, but it could also tailor its responses by adjusting the weight on its connections between neurons.

2.4 Teaching an Artificial Neural Network.

In this part of teaching the neural networks we will go through the learning methods of N.N., neural networks can be classified according to the way they learn into two kinds of learning methods which are supervised learning and unsupervised learning.

2.4.1 Supervised Learning

This section contains a full mathematical description of how supervised neural networks learn (train). The most frequently used and effective supervised learning algorithm known in the world of neural networks is the "Back-Error Propagation Algorithm" or Back-Prop for short. The type of neural networks this learning algorithm requires is "the feed forward neural networks". It is for this reason they are also known as "back-propagation neural networks. Being a supervised learning algorithm, the back-error propagation relies on a teacher which is a set of example pairs of patterns, the basic idea of the way this algorithm works is the following:

First a pair from the training data set is chosen randomly. The input pattern of the pair is given to the network at the input layer by assigning each signal of the pattern to one neuron on this layer. Then, the network passes these signals forward to the neurons on the next layer (hidden layer). But, how is this done? For each neuron on the hidden layer, a Net Input value is computed, by doing the sum over the products of the output of each neuron on the input layer (which is the original signal itself) by the weight of the connection that connects it to the neuron on the hidden layer.

When all the neurons on this layer have received a Net Input, the next step for each of these neurons is to compute, from its Net Input, an activation value which is also considered as its output. This process is done using a transfer function, usually the sigmoid function in the following way:

$$OL_{pi} = 1 / (1 + e^{-Net}L_{pi})$$

Then, these outputs are passed forward to the next layer and the same processes of computing net inputs and activations are done, until the output layer of the neural network is reached. The output values of the neurons on the output layer are taken as one pattern of signals which is considered as the actual output pattern of the network.

The actual output pattern that the network produces for each input pattern is compared to the target output pattern it should have produced which is simply the second element of the example pair chosen randomly at the beginning of the whole process. An error value is computed using the actual and target patterns as follows:

$$E_p = å \ (O_{pi} - T_{pi})^2$$

Where

E_p: is the error value that corresponds to example pair p

O_{pi}: is the output value of neuron on the output layer of the network

T_{pi}: is the i'th signal value on the target output pattern of example pair p

If the value of this error is zero, there will be no need to make any changes in the connectivity state. However, if the error value is not zero, some changes are to be made in the weights of the connections in the network reduce this error.

2.4.2 Unsupervised Learning

Unsupervised learning is a process when the network is able to discover statistical regularities in its input space and automatically develops different modes of behaviour to represent different classes of inputs (in practical applications some 'labelling' is required after training, since it is not known at the outset which mode of behaviour will be associated with a given input class). Kohonen's self-organizing (topographic) map neural networks use this type of learning.

In unsupervised or self-organized learning, the network is not given any external Indication as to what the correct responses should be nor whether the generated responses are right or wrong. It is simply exposed to the various input-output pairs and it learns by the environment, that is, by detecting regularities in the structure of input patterns.

So, unsupervised learning aims at finding a certain kind of regularity in the data represented by the exemplars. Roughly speaking, regularity means that much less data are actually required to approximately describe or classify the exemplars than the amount of data in exemplars. Examples exploiting data regularity include vector quantization for data compression and Karhunen-Loeve expansion (often referred to as principal component analysis) for dimension reduction.

In unsupervised learning, a simple Hebbian rule (correlation rule) may be applied to calculate weight changes. Energy-minimizing networks provide a recent example of unsupervised learning that makes interesting use of a two-phrase learning method. Competitive learning rules is another class of learning rules used in unsupervised neural networks. Adaptive resonance theory (ART) combines competitive and Hebbian rules together and uses feedback from the output layer to the input layer to ensure a consistent Categorization. In an ART system, connections run in both directions, from input to output nodes and vice versa. Competitive learning is used to change weights on connections from the input to the output layer in creating groupings of the input patterns.

Hebbian pattern-association learning is used to change weights on connections from the output to the input layer. As a result, an input pattern evokes a pattern on the output layer, which in turn projects the prototype of the winning group back into the input layer.

Every N.Ns goes through three operative phases:

1- Learning (training) phase – network learns on the training sample, the weights are being adjusted in order to minimize the objective function (for example RMS - root mean Square error).

2- Testing phase – network is tested on the testing sample while the weights are fixed.

3- Operative (recall) phase – NN is applied to the new cases with unknown results.

2.5 Design a Neural Network

The developer must go through a period of trial and error in the design decisions before coming up with a satisfactory design. The design issues in neural networks are complex and are the major concerns of system developers.

Designing a neural network consists of:

- Arranging neurons in various layers.

- Deciding the type of connections among neurons for different layers, as well as among the neurons within a layer.
- Deciding the way a neuron receives input and produces output.
- Determining the strength of connection within the network by allowing the networks learn the appropriate values of connection weights by using a training data set.

2.5.1 Layers

Biologically, neural networks are constructed in a three dimensional way from microscopic components. These neurons seem capable of nearly unrestricted interconnections. This is not true in any man-made network. Artificial neural networks are the simple clustering of the primitive artificial neurons. This clustering occurs by creating layers, which are then connected to one another. How these layers connect may also vary. Basically, all artificial neural networks have a similar structure of topology. Some of the neurons interface the real world to receive its inputs and other neurons provide the real world with the network's outputs. All the rest of the neurons are hidden form view.

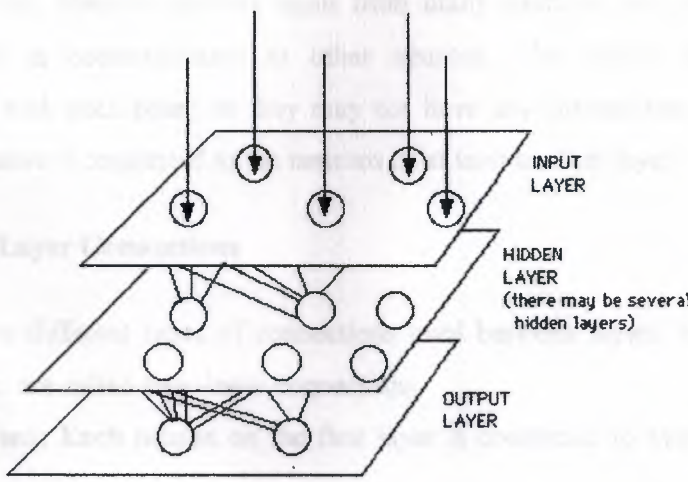

Figure2.3 The process of designing neural networks.

As the figure above shows, the neurons are grouped into layers the input layer consist of neurons that receive input form the external environment. The output layer consists of neurons that communicate the output of the system to the user or external environment. There are usually a number of hidden layers between these two layers; the figure above shows a simple structure with only one hidden layer.

When the input layer receives the input its neurons produce output, which becomes input to the other layers of the system. The process continues until a certain condition is satisfied or until the output layer is invoked and fires their output to the external environment.

To determine the number of hidden neurons the network should have to perform its best, one are often left out to the method trial and error. If you increase the hidden number of neurons too much you will get an over fit, that is the net will have problem to generalize. The training set of data will be memorized, making the network useless on new data sets.

2.5.2 Communication and Types of Connections

Neurons are connected via a network of paths carrying the output of one neuron as input to another neuron. These paths is normally unidirectional, there might however be a two-way connection between two neurons, because there may be another path in reverse direction. Neuron receives input from many neurons, but produces a single output, which is communicated to other neurons. The neuron in a layer may communicate with each other, or they may not have any connections. The neurons of one layer are always connected to the neurons of at least another layer.

2.5.2.1 Inter-Layer Connections

There are different types of connections used between layers; these connections between layers are called inter-layer connections.

Fully connected: Each neuron on the first layer is connected to every neuron on the second layer.

Partially connected: A neuron of the first layer does not have to be connected to all neurons on the second layer.

Feed forward: The neurons on the first layer send their output to the neurons on the second layer, but they do not receive any input back form the neurons on the second layer.

Bi-directional: There is another set of connections carrying the output of the neurons of the second layer into the neurons of the first layer.

Hierarchical: if a neural network has a hierarchical structure, the neurons of a lower layer may only communicate with neurons on the next level of layer.

Resonance: the layers have bi-directional connections, and they can continue sending messages across the connections a number of times until a certain condition is achieved.

2.5.2.2 Intra-Layer Connections

In more complex structures the neurons communicate among themselves within a layer, this is known as intra-layer connections. There are two types of intra-layer connections.

Recurrent: the neurons within a layer are fully- or partially connected to one another. After these neurons receive input form another layer, they communicate their outputs with one another a number of times before they are allowed to send their outputs to another layer. Generally some conditions among the neurons of the layer should be achieved before they communicate their outputs to another layer.

On-center/off surround: A neuron within a layer has excitatory connections to itself and its immediate neighbors, and has inhibitory connections to other neurons, One can imagine this type of connection as a competitive gang of neurons. Each gang excites it self and its gang members and inhibits all members of other gangs. After a few rounds of signal interchange, the neurons with an active output value will win, and is allowed to update its and its gang member's weights. (There are two types of connections between two neurons, excitatory or inhibitory. In the excitatory connection, the output of one neuron increases the action potential of the neuron to which it is connected. When the connection type between two neurons is inhibitory, then the output of the neuron sending a message would reduce the activity or action potential of the receiving neuron. One causes the summing mechanism of the next neuron to add while the other causes it to subtract. One excites while the other inhibits.)

2.5.3 Learning

The brain basically learns from experience. Neural networks are sometimes called machine learning algorithms, because changing of its connection weights (training) causes the network to learn the solution to a problem. The strength of connection between the neurons is stored as a weight-value for the specific connection. The system learns new knowledge by adjusting these connection weights. The learning ability of a neural network is determined by its architecture and by the algorithmic method chosen for training.

The training method usually consists of one of three schemes:

1. **Unsupervised learning**

The hidden neurons must find a way to organize themselves without help from the outside. In this approach, no sample outputs are provided to the network against which it can measure its predictive performance for a given vector of inputs. This is learning by doing.

2. **Reinforcement learning**

This method works on reinforcement from the outside. The connections among the neurons in the hidden layer are randomly arranged, then reshuffled as the network is told how close it is to solving the problem. Reinforcement learning is also called supervised learning, because it requires a teacher. The teacher may be a training set of data or an observer who grades the performance of the network results. Both unsupervised and reinforcement suffers from relative slowness and inefficiency relying on a random shuffling to find the proper connection weights.

3. **Back propagation**

This method is proven highly successful in training of multilayered neural nets. The network is not just given reinforcement for how it is doing on a task. Information about errors is also filtered back through the system and is used to adjust the connections between the layers, thus improving performance.

2.5.3.1 Off-line or On-line

One can categorize the learning methods into yet another group, off-line or on-line. When the system uses input data to change its weights to learn the domain knowledge, the system could be in training mode or learning mode. When the system is being used as a decision aid to make recommendations, it is in the operation mode, this is also sometimes called recall.

Off-line: In the off-line learning methods, once the systems enters into the operation *mode, its weights are fixed and do not change any more. Most of the networks are of the* off-line learning type.

On-line: In on-line or real time learning, when the system is in operating mode (recall), it continues to learn while being used as a decision tool. This type of learning has a more complex design structure.

2.5.3.2 Learning Laws

There are a variety of learning laws which are in common use. These laws are mathematical algorithms used to update the connection weights. Most of these laws are some sort of variation of the best known and oldest learning law, Hebb's Rule. Man's understanding of how neural processing actually works is very limited. Learning is certainly more complex than the simplification represented by the learning laws currently developed. Research into different learning functions continues as new ideas routinely show up in trade publications etc. A few of the major laws are given as an example below.

Hebb's Rule: The first and the best known learning rule were introduced by Donald Hebb. The description appeared in his book the organization of Behavior in 1949[7]. This basic rule is: If a neuron receives an input from another neuron and if both are highly active (mathematically have the same sign), the weight between the neurons should be strengthened.

Hopfield Law: This law is similar to Hebb's Rule with the exception that it specifies the magnitude of the strengthening or weakening. It states, "If the desired output and the input are either active or both inactive, increment the connection weight by the learning rate, otherwise decrement the weight by the learning rate." (Most learning functions have some provision for a learning rate, or learning constant. Usually this term is positive and between zero and one.)

The Delta Rule: the data rule is a further variation of Hebb's Rule, and it is one of the most commonly used. This rule is based on the idea of continuously modifying the strengths of the input connections to reduce the difference (the delta) between the desired output value and the actual output of a neuron. This rule changes the connection weights in the way that minimizes the mean squared error of the network. The error is back propagated into previous layers one layer at a time. The process of back-propagating the network errors continues until the first layer is reached. The network type called Feed forward, Back-propagation derives its name from this method of computing the error term. This rule is also referred to as the Windrow-Hoff Learning Rule and the Least Mean Square Learning Rule.

Kohonen's Law: This procedure, developed by Teuvo Kohonen, was inspired by learning in biological systems. In this procedure, the neurons compete for the opportunity to learn, or to update their weights. The processing neuron with the largest output is declared the winner and has the capability of inhibiting its competitors as well as exciting its neighbors. Only the winner is permitted output, and only the winner plus its neighbors are allowed to update their connection weights.

The Kohonen rule does not require desired output. Therefore it is implemented in the unsupervised methods of learning. Kohonen has used this rule combined with the on-center/off-surround intra- layer connection, to create the self-organizing neural network, which has an unsupervised learning method.

2.6 The Back Propagation Algorithm

The best-known example of a neural network training algorithm is back propagation. Modern second-order algorithms such as conjugate gradient descent and Levenberg-Marquardt are substantially faster (e.g., an order of magnitude faster) for many problems, but back propagation still has advantages in some circumstances, and is the easiest algorithm to understand. We will introduce this now, and discuss the more advanced algorithms later. There are also heuristic modifications of back propagation which work well for some problem domains, such as quick propagation and Delta-Bar-Delta and are also included in ST Neural Networks.

In back propagation, the gradient vector of the error surface is calculated. This vector points along the line of steepest descent from the current point, so we know that if we move along it a "short" distance, we will decrease the error. A sequence of such moves (slowing as we near the bottom) will eventually find a minimum of some sort. The difficult part is to decide how large the steps should be.

Large steps may converge more quickly, but may also overstep the solution or (if the error surface is very eccentric) go off in the wrong direction. A classic example of this in neural network training is where the algorithm progresses very slowly along a steep, narrow, valley, bouncing from one side across to the other. In contrast, very small steps may go in the correct direction, but they also require a large number of iterations. In practice, the step size is proportional to the slope (so that the algorithm settles down in a minimum) and to a special constant: the learning rate. The correct setting for the learning rate is application-dependent, and is typically chosen by experiment; it may also be time-varying, getting smaller as the algorithm progresses.

The algorithm is also usually modified by inclusion of a momentum term: this encourages movement in a fixed direction, so that if several steps are taken in the same direction, the algorithm "picks up speed"[8], which gives it the ability to (sometimes) escape local minimum, and also to move rapidly over flat spots and plateaus.The algorithm therefore progresses iteratively, through a number of epochs. On each epoch, the training cases are each submitted in turn to the network, and target and actual outputs compared and the error calculated. This error, together with the error surface gradient, is used to adjust the weights, and then the process repeats. The initial network configuration is random and training stops when a given number of epochs elapse, or when the error reaches an acceptable level, or when the error stops improving.

2.7 What can you do with an N.N. and What not?

In principle, N.Ns can compute any computable function, i.e., they can do everything a normal digital computer can do (Valiant, 1988; Siegelmann and Sontag, 1999; Orponen, 2000; Sima and Orponen, 2001) [9], or perhaps even more, under some assumptions of doubtful practicality.

Practical applications of N.Ns most often employ supervised learning. For supervised learning, you must provide training data that includes both the input and the desired result (the target value). After successful training, you can present input data alone to the NN (that is, input data without the desired result), and the NN will compute an output value that approximates the desired result. However, for training to be successful, you may need lots of training data and lots of computer time to do the training. In many applications, such as image and text processing, you will have to do a lot of work to select appropriate input data and to code the data as numeric values.

In practice, N.Ns are especially useful for classification and function approximation/mapping problems which are tolerant of some imprecision, which have lots of training data available, but to which hard and fast rules (such as those that might be used in an expert system) cannot easily be applied. Almost any finite-dimensional vector function on a compact set can be approximated to arbitrary precision by feed-forward N.Ns (which are the type most often used in practical applications) if you have enough data and enough computing resources.

To be somewhat more precise, feed-forward networks with a single hidden layer and trained by least-squares are statistically consistent estimators of arbitrary square-inerrable regression functions under certain practically-satisfiable assumptions regarding sampling, target noise, number of hidden units, size of weights, and form of hidden-unit activation function. Such networks can also be trained as statistically consistent estimators of derivatives of regression functions and quintiles of the conditional noise distribution .Feed-forward networks with a single hidden layer using threshold or sigmoid activation functions are universally consistent estimators of binary classifications, under similar assumptions. Note that these results are stronger than the universal approximation theorems that merely show the existence of weights for arbitrarily accurate approximations, without demonstrating that such weights can be obtained by learning.

Unfortunately, the above consistency results depend on one impractical assumption: that the networks are trained by an error (or misclassification rate) minimization technique that comes arbitrarily close to the global minimum. Such minimization is computationally intractable except in small or simple problems .In practice, however, you can usually get good results without doing a full-blown global optimization; e.g., using multiple (say, 10 to 1000) random weight initializations is usually sufficient.

One example of a function that a typical neural net cannot learn is `Y=1/X` on the open interval (0, 1). An open interval is not a compact set. With any bounded output activation function, the error will get arbitrarily large as the input approaches zero. Of course, you could make the output activation function a reciprocal function and easily get a perfect fit, but neural networks are most often used in situations where you do not have enough prior knowledge to set the activation function in such a clever way. There are also many other important problems that are so difficult that a neural network will be unable to learn them without memorizing the entire training set, such as:

- Predicting random or pseudo-random numbers.
- Factoring large integers.
- Determine whether a large integer is prime or composite.
- Decrypting anything encrypted by a good algorithm.

And it is important to understand that there are no methods for training N.Ns that can magically create information that is not contained in the training data.

2.8 Where are Neural Networks Being Used?

Neural networks are performing successfully where other methods do not, recognize and matching complicated, vague, or incomplete patterns. Neural networks have been applied in solving a wide variety of problems.

The most common use for neural networks is to project what will most likely happen. There are many areas where prediction can help in setting priorities. For example, the emergency room at a hospital can be a hectic place; to know who the most critical needs help can enable a more successful operation. Basically, all organizations must establish priorities, which govern the allocation of their resources. Neural networks have been used as a mechanism of knowledge acquisition for expert system in stock market forecasting with astonishingly accurate results. Neural networks have also been used for bankruptcy prediction for credit card institutions.

Although one may apply neural network systems for interpretation, prediction, diagnosis, planning, monitoring, debugging, repair, instruction, and control, the most successful applications of neural networks are in categorization and pattern recognition. Such a system classifies the object under investigation (e.g. an illness, a pattern, a picture, a chemical compound, a word, the financial profile of a customer) as one of numerous possible categories that, in return, may trigger the recommendation of an action (such as a treatment plan or a financial plan.

2.9 Summary

In this chapter I tried to concentrate on the algorithms and the structure of the neural network, here there are two categories, which are Feed-forward Network and Feedback Network. And after that I have to moved ahead to give a hint about how the neural network can be trained. The neural network able to be trained in two methods the first one is Supervised Learning and the second one is Unsupervised Learning.

CHAPTER THREE

NEURAL NETWORK IN MEDICINE

3.1 Overview

The aim of this chapter is to demonstrate some basic applications of neural networks in medicine. It presents understanding the concept of a diagnosis prediction via an artificial neural network knowledge base. One of the major applications of medical information has been the metabolic disease diagnosis, and the general theme of this chapter is the ethical issues in the use of neural network, based methodologies for image interpretation in medicine, I will discuss a clinical application also, finally the chapter ends with artificial neural network in immunology.

3.2 Integration of Neural Networks and Knowledge-Based Systems in Medicine

Knowledge-Based Systems are used in medical diagnoses. They have the advantage to give an explanation of a diagnosis. But a main problem when dealing with Knowledge-Based Systems is the acquisition of knowledge. Artificial Neural Networks deal with knowledge in a sub symbolic form. Incomplete and imprecise data can be processed by approximating not linear relations in data. In a laboratory or medical system the integration of the neural network system into the decision making process may be required. We realized this by building a hybrid system consisting, first, of graphical visualizations methods and second, a machine learning module generating rules out of the neural network. The rules are presented in a form, which can be understood by humans and used in Knowledge-Based Systems. Keywords: Knowledge-Based System, Neural Network, decision making, visualization, machine learning. Integration of Neural Networks and Knowledge-Based Systems

Knowledge-Based Systems are used in medical diagnoses. They have the advantage to give an explanation of a diagnosis. This is very important especially in the domain of medicine where the user wants to have the diagnosis proved. But a main difficulty when dealing with Knowledge-Based Systems is the acquisition of the domain knowledge. There are several problems with it. It is difficult to transform the explicit and implicit knowledge of the expert's domain, which also partly consists of own experience, in a form which is suitable for a knowledge base. The knowledge can also be inconsistent or incomplete. A second problem is that Knowledge-Based Systems are not able to learn from experience or to operate with cases not represented in the knowledge base.

Artificial Neural Networks deal with knowledge in a sub symbolic form. They can solve non-linear problems often better than conventional methods and are capable to approximate non linear relations in data. In addition, incomplete and imprecise data can be processed. Neural networks learn in a massively parallel and self-organizing way. Unsupervised learning neural networks, like Kohonen's self organizing feature maps [10], learn the structure of high-dimensional data by mapping it on low dimensional topologies, preserving the distribution and topology of the data. But large neural networks can only be interpreted with analyzing tools. We developed a visualization method, the so called U-Matrix methods, to detect the structure of large two-dimensional Kohonen maps. It generates a three-dimensional landscape on the map, whereby valleys indicate data which belongs together and walls separate subcategories [11].

In a laboratory or medical system the integration of the neural network system into the decision making process may be useful. The knowledge of neural networks, however, is in this form not communicable; i.e. it is necessary to transform the knowledge into a form, which, first, can be understood by humans and second, can be processed by knowledge based systems. Knowledge based systems have the advantage that they can give an explanation of a diagnosis. By integrating both paradigms, knowledge based systems and neural networks, the disadvantages of both approaches can be redressed.

We are developing a hybrid system REGINA (rule extraction and generation in neural architecture) which consists of several parts. An unsupervised learning neural network maps the (preprocessed) data space onto a two-dimensional grid of neurons, whereby it preserves the distribution and topology of the input space. But only together with a visualization module, called U-Matrix methods, we are able to detect structure in the data and classify it. A three-dimensional colored landscape will be generated in which walls separate distinct subclasses and subcategories are represented by valleys. A machine learning algorithm extracts rules out of the learned neural network [12]. In distinction to other machine learning algorithms like ID3 our algorithm considers the attributes by selecting those which are relevant for the classification. This corresponds to the proceeding of a medical expert. The rules can be used as a knowledge base for an expert system. Also fuzzy rules can be extracted out of the neural network.

3.3 Artificial Neural Network Based Cardiovascular Modeling

One approach to cardiovascular modeling is to build a model representative of a group of individuals with similar characteristics (i.e., sex, age, physical condition, medical condition, etc.). However, cardiovascular behavior is unique to each individual [13], thus a generic cardiovascular model used in a medical diagnostic system would not be as sensitive as a system based on a model that is adapted to the patient being diagnosed. To develop these models without a cardiovascular expert, the modeling must be based on an adaptive technology that can be automated. The A.N.N. technology fits this category.

The ANN technology was selected for the cardiovascular modeling because of its many capabilities including sensor fusion, which is the combining of values from several different sensors. Sensor fusion enables the A.N.Ns to learn complex relationships among the individual sensor values, which would otherwise be lost if the values were individually analyzed. In medical modeling and diagnosis, this implies that even though each sensor in a set may be sensitive only to a specific physiological variable, ANNs are capable of detecting complex medical conditions by fusing the data from the individual biomedical sensors.

Recurrent ANNs were selected for the cardiovascular modeling application to capture the temporal information in physiological variables. These variables are time-series data from which both the absolute values and the rates of change need to be modeled. Recurrent ANNs recycle a small portion of information from time t-1 at time t. indirectly; decreasing portions of information from time t-2, t-3, t-4, etc. are also captured, thus enabling recurrent A.N.Ns to model the temporal dynamics in data.

Figure 3.1 illustrates a prototype tool that generates an ANN model of the cardiovascular system from physiological variables received from biomedical sensors attached to an individual.

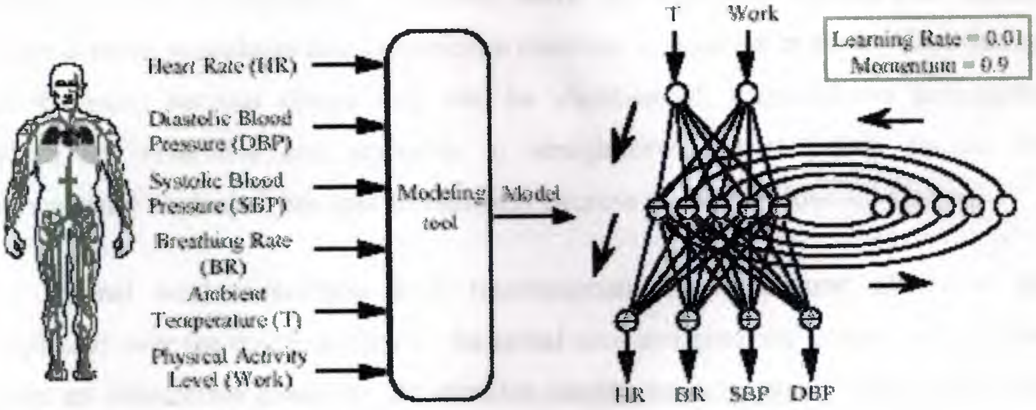

Figure 3.1 Illustrates a prototype tool that generates an ANN model

Figure 3.1, this figure illustrates a modeling tool that takes a sequence of physiological variables from biomedical sensors and learns the temporal dynamics of these variables to produce an ANN-based cardiovascular model. On the right, this figure illustrates the configuration of the ANN produced by the modeling tool. The ANN has two inputs, four outputs, and five hidden processing elements. The ANN takes the ambient temperature and the physical activity as input the four outputs, heart rate, breathing rate, systolic blood pressure, and diastolic blood pressure, are clamped to the "actual" values during the training phase. For the initial cardiovascular model prototypes, the "actual" values are generated by a no adaptive cardiovascular model. During the modeling phase, the temperature and the work are input to the ANN, and the values at the outputs are taken as the modeled variables. The feedback links going through the five processing elements on the right side of the ANN enable it to capture temporal information in the data.

During the adaptation phase, the training algorithm receives physiological data from an individual via biomedical sensors and automatically develops the ANN-based cardiovascular model. After development, the model can generate the appropriate physiological responses for simulations with varying levels of physical activity.

3.4 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 [14].

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 flanked by anode above and below) is preferred by our patients to a statistically significant degree. 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 [15].

3.5 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.

3.6 A Neural Network Model for Metabolic Disease Diagnosis

We have developed a prototype computer program, MetaNet, that uses combination of artificial neural networks and knowledge-based expert systems to assis in the diagnosis of inborn errors of metabolism in children Results of amino acid analysis data of normal children, and of patients diagnosed with a number of amino acid and organic acid abnormalities were used as inputs to train the neural network component of the program. To diagnose new cases, plasma or urinary amino acid results are entered. The knowledge-based expert system then asks questions of the user regarding the presence or absence of common clinical and/or biochemical abnormalities. Using both the amino acid data and the answers to the questions; the MetaNet program integrates the output of the neural network and the results of the knowledge-based expert system to yield a provisional diagnosis.

The diagnostic output is accompanied by a numerical belief vector, which indicates the degree of confidence of the program in the diagnosis. Altering any of the input variables followed by reprocessing of the data generates a new diagnostic output and a revised belief vector. This allows analysis of the importance of any input variable to the proposed diagnosis. The knowledge-based expert system also includes a section entitled Independent Metabolic Disease Reference Documents, which provides additional information about a suspected metabolic disease when requested by the user.

The neural network component consists of eight, three-layer neural networks that are trained using a back-propagation approach. Analysis of the hidden layers following training of the neural network revealed both expected and novel, unexpected connections between specific diagnoses and clusters of amino acids. Such data may be used as a guide for future investigation of the contribution of the metabolism of specific amino acids to amino acid disorders.

3.7 Neural Network for Plasma Torch Operation and Process Control

The prediction and classification features of ANN have been used to support torch control at DIAL, and other potential applications are being identified. ANN currently has a modular design to use it as a component in a larger intelligent support system. A.N.N. derives its topology from the interconnections of neurons in the human brain. It learns the characteristics of each process by observation, and adapts its strategy for a particular situation. Neural networks automatically take advantage of process upgrades and compensate for sys-tem degradation. Nonlinear and multivariate capabilities of neural networks make it ideal for even direct control. The network's neurons (represented by the filled circles) work together to make decisions based on their individual inputs and weights. The network is trained by sampling thousands of examples of actual data where input patterns were received and correct decisions were made the network learns to recognize these correct patterns using a neural mathematical model. After being trained to recognize input patterns, the neural network can make good decisions even for new and unfamiliar patterns.

Neural networks handle time series prediction in a unique way as they are versatile and accommodate nonlinear data with noise. It adjusts its own model based on the behavior of historical data, to predict future behavior fairly accurately. As many of the time series have a significant chaotic component, neural networks do a far better job in handling them than other models of time series prediction.

Time series prediction techniques have been applied at DIAL to plasma torch voltages to stabilize output power using the back propagation model of an artificial neural net-work. The torch power fluctuation is caused by the variation of gas pressure inside the electrode to move the arc attachment point. Operation with small power fluctuations is not usually detrimental. However, power stabilization allows more freedom for gas pressure variation, and this can be very useful. The application of ANN to solve this problem has eliminated the power variation due to gas pressure for the trained operating conditions.

During classification, a fully trained ANN recognizes a familiar pattern or makes an approximate guess on the unfamiliar patterns. DIAL has also applied ANN to characterize a simulated waste feed stream. Physical properties of waste materials have been used as inputs to ANN to identify the materials. In the present model, a fuzzy ARTMAP network uses four parameters to identify ten materials. The system classified 94% of the test data correctly.

3.8 Application of Neural Network in Immunology

The immune system recognizes agents foreign to the host organism and raises appropriate responses. Foreign includes viruses, bacteria, parasites, fungi, tumors, and transplants. The immune recognition process in vertebrates involves major histocompatibility complex (MHC) molecules which bind short peptides and display them on the cell surface for recognition by T-cells of the immune system [16]. Binding of peptides to MHC molecules is necessary for immune recognition, but only a limited number of peptides can bind to a specific MHC molecule. Determining which peptides can bind to specific MHC molecules is crucial for understanding the basis of immunity and is important for identifying of candidates for the design of vaccines and immunotherapeutic drugs.

Prediction of peptide binding to MHC molecules is a combinatorial pattern recognition problem. A peptide is a chain composed of amino acids. The majority of the peptide cores that are responsible for binding to MHC molecules are 9 amino acids long, although lengths of 8- 12 amino acids are also common. There are 5.12′10 11 peptides of length nine that can be composed of twenty naturally occurring amino acids. Peptide sets specific for a particular MHC variant in some cases overlap sets specific for other MHC variants. Often, these sets are exclusive. Human MHC is known as human leukocyte antigen (HLA); more than 700 variants have been characterized to date. An individual may express up to 20 different HLA molecules. An immune system has to discriminate between self and foreign peptides, tolerate more than 10 7 self peptides and respond appropriately to a relatively small subset of targets from more than 10 11 potential foreign peptides. For a given protein the peptide targets of immune recognition often differ between individuals even when some of their HLA molecules are identical [17].

The importance of the computational analysis in immunology is increasing with recent advances in molecular and clinical immunology. These advances have resulted in accumulation of experimental data and in improved understanding of immunological processes. Immunological databases are growing in both size and complexity [18]. This complexity growth has roots in the combinatorial aspects of immunology: genes that encode products of the immune system are the most variable (polymorphic) gene super family of the organism. The immune system products interact with, and facilitate appropriate responses against, practically unlimited number of targets. The immune response targets include peptides derived from environmental-, microbial- or self-antigens. The discovery of specific targets of immune responses is important, but is becoming increasingly difficult as we learn more about the complexity of the domain. Computational models of immune interactions include those that utilize binding motifs quantitative matrices, artificial neural networks (Ann's), hidden Markov models (HMMs), and molecular modeling. Computational models have been used in immunology to a) minimize the number of experiments required for the determination of targets of immune responses, or b) to conduct large-scale computational simulations facilitating knowledge discovery when the experimental approach is not possible. Computational methods for data analysis and modeling in immunology must provide the ability to a) deal with fuzzy data, b) deal with incomplete data, c) tolerate noise and errors, and d) easily incorporate new data.

Models based on ANNs have played a prominent role in computational immunology applications. ANN-based models have proven superior in accuracy and ease of both application and model refinement. The aims of this article are a) to discuss the use of computational models in immunology, and b) to describe successful applications of ANN models in immunology.

3.9 Summary

This chapter presented applications of neural networks in medicine. The chapter described the integration of neural networks and knowledge based system in medicine, and presented the clinical application, I then introduce artificial neural network based cardiovascular modeling, I discuss model neural network for metabolic Disease Diagnosis, and the artificial neural network application for plasma torch operation and process control.

CHAPTER FOUR

MEDICAL DIAGNOSTIC AIDES

4.1 Overview

Medical information systems provide an inherent instrument for knowledge discovery. The purpose of this chapter is to create a medical diagnostic that will serve as a diagnostic tool for medical analysis and diagnosis by neural networks. This tool can help guide health care regarding preventive and predictive action, and modeling diagnosing the cardiovascular system. This chapter also provides neural networks in diagnostic decision support systems the results generated by this chapter neural network model for disease diagnosis.

4.2 Medical Analysis and Diagnosis by Neural Networks

Almost all the physicians are confronted during their formation by the task of learning to diagnose. Here, they have to solve the problem of deducing certain diseases or formulating a treatment based on more or less specified observations and knowledge. Certainly, there is the standard knowledge of seminars, courses and books, but on one hand medical knowledge outdates quickly and on the other hand this does not replace own experience. For this task, certain basic difficulties have to be taken into account:

- The basis for a valid diagnosis, a sufficient number of experienced cases, is reached only in the middle of a physician's career and is therefore not yet present at the end of the academic formation.
- This is especially true for rare or new diseases where also experienced physicians are in the same situation as newcomers.
- Principally, humans do not resemble statistic computers but pattern recognition systems. Humans can recognize patterns or objects very easily but fail when probabilities have to be assigned to observations.

These principal difficulties are not widely known by physicians. Also studies who revealed that about 50% of the diagnoses are wrong do not impede the selfconscience of some physicians. It is not by chance that the disease AIDS which manifests by a myriad of infections and cancer states was not discovered directly by treating physicians but by statistical people observing the improbable density of rare cancer cases at the U.S. west coast. An important solution for the described problem lies in the systematic application of statistical instruments. The good availability of computers ameliorates the possibilities of statistically inexperienced physicians to apply the benefits of such a kind of diagnosis:

- Also physicians in the learning phase with less experience can obtain a reliable diagnosis using the collected data of experienced colleagues.
- Even in the case of rare diseases, e.g. septic shock, it is possible to get a good diagnosis if they use the experience of world-wide networked colleagues.
- New, unknown diseases can be systematically documented even if this induces complex computations which are not known to the treating physician.
- Also in the treatment of standard diseases a critical statistical discussion for the use of operation methods or medical therapies may introduce doubts in the physicians own, preferred methods as it is propagated by the ideas of Evidence Based Medicine EBM [19].

A classical, early study [20] in the year 1971 showed these basic facts in the medical area. At the university clinic of Leeds (UK) 472 patients with acute abdominal pain where examined and diagnosed. With simple, probability-based methods (Bayes classification) the diagnostic decision probabilities were computed based on a data base of 600 patients. Additionally, a second set of probabilities were computed by using a synthetic data base of patients build on the interviews of experts and questionnaire sheets about 'typical' symptoms. Then, the 472 cases were diagnosed by an expert round of 3 experienced and 3 young physicians. The result of this experiment was as follows:

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

4.3 Diagnosis by Growing Neural Networks

The neural network chosen for our classification task is a modified version of the supervised growing neural. Compared to the classical multilayer perceptron trained with back propagation which has reached a wide public, this network achieved similar results on classification tasks but converges faster. It is based on the idea of radial basis functions. Its additional advantage is the ability to insert neurons within the learning process to adapt its structure to the data. In our case we had only 70 patients with the diagnosis "septic shock". Our classification is based on 2068 measurement vectors (16-dimensional samples) from variable set V taken from the 70 septic shock patients. 348 samples were deleted because of too many missing values within the sample. With 75% of the 1720 remaining samples the SGNG was trained and with 25% samples from completely other patients than in the training set it was tested. The variables were normalized (mean 0, standard deviation 1) for analysis. The network chosen was the one with the lowest error on the smoothed test error function. Three repetitions of the complete learning process with different, randomly selected divisions of the data were made. The results are presented in Table 4.1.

Table 4.1 Correct classifications, sensitivity, and specificity with standard deviation; minimum and maximum in % from three repetitions.

Measure	Mean value	Standard deviation	minimum	maximum
Correct classification	67.84	6.96	61.17	75.05
Sensitivity	24.94	4.83	19.38	28.30

To achieve a generally applicable result ten repetitions would be better, but here it is already clear: with the low number of data samples the results can only have prototypical character, even with more cleverly devised benchmark strategies. On average we have an alarm rate (1- specificity) of 8.39% for survived patients showing also a critical state and a detection of about 1 out of 4 critical illness States. For such a complex problem it is a not too bad, but clearly no excellent result. An explanation for this low number is grounded in the different, individual measurements of each patient.

4.4 Diagnosis by Rule Based Networks

Results of classification procedures could provide a helpful tool for medical diagnosis. Nevertheless, in practice physicians are highly trained and skilled people who do not accept the diagnosis of an unknown machine (black box) in their routine. For real applications, the diagnosis machine should be become transparent, i.e. the diagnosis should explain the reasons for classification. Whereas the explanation component is obvious in classical symbolic expert system tools, neural network tools hardly explain their decisions. This is also true for the SGNG network used in the previous section. Therefore, as important alternative in this section we consider a classification by learning classification rules which can be inspected by the physician. The details of the network structure and the learning algorithm can be found in [21].

Now we present the results of the rule generation process with our previously introduced septic shock data set. The data set is 16-dimensional. A maximum of 6 variables for every sample was allowed to be missing. The missing values were replaced by random data from normal distributions similar to the original distributions of the variables. So it was assured that the algorithm can not learn a biased .Result due to biased replacements, e.g. Class-dependent means. We demand a minimum of 10 out of 17 variables measured for each sample, so there remained 1677 samples out of 2068 for analysis. The data we used in 5 complete training sessions – each one with a different randomly chosen training data set – was in mean coming from class 1 with a percentage of 72.10% and from class 2 with a percentage of 27.91%. In the mean 4.00 epochs were needed (with standard deviation 1.73, minimum 3 and maximum 7). Test data was taken from 35 randomly chosen patients for every training session, containing no data sample of the 35 patients in the training data set. In general, results of a patient classification or prediction task are true only with certain probability. Therefore, any prognostic system can not predict always the correct future state but may just give early warnings for the treating physician. In Table 4.2 the classification results are presented.

Table 4.2 Mean, standard deviation, minimum and maximum of correct classifications and not classifiable data samples of the test data set. In %.

	Mean	Standard deviation	Minimum	Maximum
Correct classification	68.42	8.79	52.92	74.74
Not classified	0.10	0.22	0.08	0.84

Average specificity ("deceased classified / all deceased") was 88 % and average sensitivity ("survived classified / all survived") was 18.15 %. The classification result is not satisfying. Deceased patients were not detected very well. Reasons for this can be the very individual behavior of the Patients and the data quality (irregularity of measurements, missing values). In this way it seems not possible to classify all the patients correctly, but it could be that in some areas of the data space the results are better (local rules). In the mean 22.80 rules were generated for the class survived and 17.80 rules were generated for class deceased.

4.5 Modeling and Diagnosing the Cardiovascular System

Neural Networks are used experimentally to model the human cardiovascular system. Diagnosis can be achieved by building a model of the cardiovascular system of an individual and comparing it with the real time physiological measurements taken from the patient. If this routine is carried out regularly, potential harmful medical conditions can be detected at an early stage and thus make the process of combating the disease much easier.

A model of an individual's cardiovascular system must mimic the relationship among physiological variables (i.e., heart rate, systolic and diastolic blood pressures, and breathing rate) at different physical activity levels. If a model is adapted to an individual, then it becomes a model of the physical condition of that individual. The simulator will have to be able to adapt to the features of any individual without the supervision of an expert. This calls for a neural network.

Another reason that justifies the use of ANN technology is the ability of ANNs to provide sensor fusion which is the combining of values from several different sensors. Sensor fusion enables the ANNs to learn complex relationships among the individual sensor values, which would otherwise be lost if the values were individually analyzed. In medical modeling and diagnosis, this implies that even though each sensor in a set may be sensitive only to a specific physiological variable, ANNs are capable of detecting complex medical conditions by fusing the data from the individual biomedical sensors.

This model could be used to monitor employees in hazardous environments like fire-fighters. The system could be used to determine whether firemen have recovered sufficiently from the last inhalations of smoke to be allowed to enter smoke-filled environments again. The advantages that such a system can offer are obvious. People can be checked for heart diseases quickly and painlessly and thus detecting any disease at an early stage. Of course, the system doesn't eliminate the need for doctors since a human expert is more reliable.

4.6 Neural Networks in Diagnostic Decision Support Systems

Recent advances in mathematical theories and information technology have increased the reliability of artificial decision support systems and allowed them to gain greater acceptance among skeptical health care professionals. One application of computer-aided decision making is in prediction of medical diagnosis. An artificially intelligent decision maker can identify the ailment of a patient from symptoms such as blood pressure, pulse, signals from an electrocardiogram, images from MRI and PET scans, etc. A physician may either consult the support system for guidance or rely on it completely to seek the most likely diagnosis under uncertain conditions.

Any decision support system requires a knowledge base, a body of knowledge from which the system retrieves specific information needed for decision making. Designing an effective decision support system requires a vast reservoir of past patient data. This knowledge base can be as specific as to capture characteristics of only two types of disease, or so broad as to represent a multitude of ailments in a large population. Given such a knowledge base, artificial neural networks can accurately represent the functional relation between symptoms and diagnoses.

Patterned after biological neurons, artificial neural networks are often used to represent knowledge and process information, especially when a complex relationship exists among sets of data. These highly distributed models are easily implemented on computers. Building a network model, referred to as training, generally requires a set of input/output data that demonstrates 'proper' network behavior. Once trained, a good neural network model should be able to generate correct outputs (i.e., diagnoses) given a set of inputs (i.e., symptoms) it has not encountered previously. Neural networks can be especially powerful in estimating dynamics of systems whose physics or internal components are poorly understood and which can not be modeled via classical methods. Applications of the technology include approximation of nonlinear functions, image processing, signal processing, adaptive control, and robotics1- 3. Neural networks have also proven their value in more nontraditional areas, such as economic forecasting and financial advising [22].

Recent investigations have repeatedly shown that neural networks outperform conventional statistical techniques in diagnosis prediction. Armoni and colleagues used neural networks to estimate the diagnostic probabilities of insulin dependent diabetes mellitus and compared their results with a multiple variable regression approach 6. Diagnostic probabilities refer to the frequency of a disease in a group of people with specific symptoms. In this study, an example involves the probability of diagnosing a patient having high blood glucose levels with insulin-dependent diabetes mellitus. The experiment was based on a data set consisting of 200 records. Of the 200 records, 150 were used to train the neural network and build, the multiple variable regression models and the remaining 50 were used to test each approach. The final neural network had a prediction accuracy that exceeded the multiple regression variable models by 9%. Although this appears to be a small margin of superiority, their results are consistent with other research: for some applications, conventional statistical techniques are not as accurate as neural networks7, 8.

In another study, Smolek et al. introduce a neural network model that distinguishes between patients who have the eye disease keratinous and those who are suspected to have it 9. Keratoconus is clinically characterized by certain biochemical, structural, topographical changes on the corneal surface10, 11. Keratoconus suspects, on the other hand, are corneas that exhibit only some of the symptoms seen in keratoconus.

In training the network, corneal topographic patterns were used as input. The optimal network model was trained with a data set consisting of 150 records and had 100% accuracy when tested against a separate set of 150 records. Overall, the network had superior performance over four conventional videokerato graphic methods used to detect the two types of disease. Keratoconus suspects often pose problems during clinical/preoperative screening, perhaps resulting in unnecessary surgical procedures. It is therefore essential to have a robust classification tool, such as neural networks, to prevent unneeded or improper treatments.

While these studies demonstrate the power of neural networks, outstanding performance in diagnosis prediction is not guaranteed. The predictive accuracy of the neural network depends largely on the amount of data and the classes of data that are initially available when building the model. For example, Arle and colleagues used various combinations of data sets to train neural networks for identifying three types of brain tumors in children12. Their network model had an accuracy of only 58% when three types of spectroscopy data served as input. However, when data on MRI characteristics, age, sex, and tumor size were added to the network's set of input variables, accuracy improved to 72%. It is interesting to note that a neuroradiologist had similar classification accuracy (73%) using the same information. Accuracy of prediction peaked at an impressive value of 95% when the network used all available types of data. One would expect that an initially larger database, coupled with more ways to characterize a disease, leads to an increased accuracy in the network's predictive capability.

The widespread use of computer-based diagnosis is still much debated and restricted 13. Is it conceivable that this valuable tool could replace physicians completely in the near future? Most patients would probably prefer that a human being, not a computer, decide on their state of health. It is argued that an artificial decision support system is unable to capture the more qualitative symptoms, such as the patient's facial or vocal expression. Only a doctor has the insight to interpret these signs. However, diagnosis greatly depends on subjective evaluation by the physician, which can be erroneous. Robust computer based diagnostic tools, such as neural networks; provide objective methods that supplement other types of information available to the physician. They allow the health care professional to diagnose a condition accurately and quickly, with increased confidence. In the long run such implementations might reduce hospital costs passed down to patients.

4.7 A Neural Network Model for Metabolic Disease Diagnosis

We have developed a prototype computer program, MetaNet, that uses a combination of artificial neural networks and knowledge-based expert systems to assist in the diagnosis of inborn errors of metabolism in children.

Results of amino acid analysis data of normal children, and of patients diagnosed with a number of amino acid and organic acid abnormalities were used as inputs to train the neural network component of the program. To diagnose new cases, plasma or urinary amino acid results are entered. The knowledge-based expert system then asks questions of the user regarding the presence or absence of common clinical and/or biochemical abnormalitie using both the amino acid data and the answers to the questions, the metaNet program integrates the output of the neural network and the results of the knowledge-based expert system to yield a provisional diagnosis.

The diagnostic output is accompanied by a numerical belief vector, which indicates the degree of confidence of the program in the diagnosis. Altering any of the input variables followed by reprocessing of the data generates a new diagnostic output and a revised belief vector. This allows analysis of the importance of any input variable to the proposed diagnosis. The knowledge-based expert system also includes a section entitled Independent Metabolic Disease Reference Documents, which provides additional information about a suspected metabolic disease when requested by the user.

The neural network component consists of eight, three-layer neural networks that are trained using a back-propagation approach. Analysis of the hidden layers following training of the neural network revealed both expected and novel, unexpected connections between specific diagnoses and clusters of amino acids. Such data may be used as a guide for future investigation of the contribution of the metabolism of specific amino acids to amino acid disorders.

4.8 Summary

This chapter presented medical diagnostic aides and medical analysis by neural networks, and I have described the diagnosis by growing neural networks and by rule based networks, provided neural network are used experimentally to model the human cardiovascular system.

CONCLUSION

Artificial neural networks are one of the promises for the future in computing. They offer an ability to perform tasks outside the scope of traditional processors. They can recognize patterns within vast data sets and then generalize those patterns into recommended courses of action, even though they are not traditionally programmed, the designing of neural networks does require a skill. It requires an "art." This art involves the understanding of the various network topologies, current hardware, current software tools, the application to be solved, and a strategy to acquire the necessary data to train the network.

Chapter one described the definition of artificial neural network and provided brief history of neural networks. After that I have moved to give a hint about how artificial neurons they work.

Chapter two is tried to concentrate on the algorithms and the structure of the neural network which are feed-forward network and feedback network.

Chapter three is presented application of neural networks in medicine and described the integration of neural networks and knowledge based systems in medicine, and presented the clinical application.

Chapter four presented medical diagnostic aides and medical analysis by neural networks and described the diagnosis by growing neural networks and by based networks.

Artificial neural network is being developed all the time especially in medicine it will be important to investigate and learn about neural network, then emergence, structure and application in real life, and to form on medical application of neural network, and to domesticate a medical neural network application in real life, namely medical diagnostic aids, these were the objectives of this project, which have all been accomplished.

REFERENCES

[1] Academic writing from the World Wide Web:
http://www.dacs.dtic.mil/techs/neural/neural_toc.html

[2] Academic writing. From the World Wide Web:
http://www.dacs.dtic.mil/tech/neural/neural_toc.html

[3] Academic writing from the World Wide Web:
http://www.usps.gov/websites/

[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] D. Jones. "Neural Networks for Medical Diagnosis." *Handbook of Neural Computing Applications*", Academic Press, pp. 309-318, 1990.

[14]Fox PT. Basic Principles and Neurosurgical Applications of Positron Emission Tomography. Neurosurg Clin N Am 1997; 8: 293–306.

[15]Kohonen T: Self-organised formation of topologically correct feature maps. Biological Cybernetics, 1982; 43: 59-69

[16] Cresswell P. Assembly, transport, and function of MHC class II molecules. Annual Review in Immunology, 12, 259-293, (1994).

[17] Wang R., Doolan D.L., Le T.P., Hedstrom R.C., Coonan K.M., Charoenvit Y., Jones J.A. and Hoffman S.L. Induction of antigen-specific cytotoxic T lymphocytes in humans by a malaria DNA vaccine. Science, 282, 476-480, (1998).

[18] Brusic V., Rudy G. and Harrison L.C. MHCPEP – a database of MHC-binding peptides: update 1997. Nucleic Acids Research, 26, 368-371(1998).

[19] Mamitsuka H. Predicting peptides that bind to MHC molecules using supervised learning of hidden Markov models. Proteins 33, 460-474, (1998).

[20] Zhang C., Anderson A. and DeLisi C. Structural principles that govern the peptide-binding motifs of class I MHC molecules. *Journal of Molecular Biology*, 281, 929-947, (1998).

[21] Academic writing from the World Wide Web:

http://hiru.mcmaster.ca/

[22] Brause R., Hamker F., Paetz J.: Septic Shock Diagnosis By Neural Networks And Rule Based Systems; in: L.C. Jain: Computational Intelligence Techniques In Medical Diagnosis And Prognosis, Springer Verlag 2001, in press