3   Artificial Neural Networks

3.1   Introduction

The mystery of the human brain's capability to solve complex problems has fascinated scientists for many centuries. Studies revealed that the human brain uses a web of highly interconnected neurons or processing elements to process complex data. Each neuron is independent and can function without any synchronization to other events taking place. (Rao & Rao, 1995). Some scientists tried to reconstruct or simulate the way the brain works by using binary valued information processing units, which are abstracted versions of their biological counterparts. "Much of this surge of attention results, not from interest in neural networks as models of the brain, but rather from their promise to provide solutions to technical problems of "artificial intelligence" that the traditional, logic based approach did not yield." (Muller & Reinhardt, 1990:preface.)

Artificial neural networks are a promising new field. Not only does it yield a better understanding of how the brain's complex information processing abilities work, but it also solves difficult problems too complex for conventional information processing techniques such as statistics.

3.2   Overview of History and Development of Artificial Neural Networks

The study of the human brain goes back a long way. Nelson and Illingworth (1991) mention the work described in the Edward Smith Papyrus. It is a medical paper about the sensory and motor locations in the brain, written around 3000 B.C, almost five millennia back. It was only until this century that researchers tried to simulate the actual functioning of a human brain.

The first person to use the human brain as a computing paradigm was Alan Turing in 1936 (Nelson & Illingworth, 1991). In 1943, McCulloch and Pitts (cited in Nelson & Illingworth, 1991) wrote the first paper about the theory of how the nervous system might work and also simulated a simple neural network with electrical circuits. Researchers began to imitate the biological model to create intelligent machines. Donald Hebb made the connection between psychology and physiology and pointed out that a neural pathway is reinforced every time it is used. He formulated his learning rule, still referred to as the Hebb rule of learning in 1949 (Nelson & Illingworth, 1991). This rule states that changes in connections between neurons are proportional to the activation of the neurons. This was the formal basis for the creation of neural networks that have the ability to learn.

Research expanded and neural network terminology started to appear in the 1950s. In 1957 Rosenblatt expanded on the theory of Hebbian learning and incorporated it into a two layer network, calling the result a "perceptron" (Blum, 1992). Rosenblatt formulated his own learning rule "the perceptron convergence theorem". This rule describes the weights adjusted in proportion to the error between output neurons and

target outputs. Many neural networks still use this method of adjusting weights until the desired set of weights is achieved to learn and predict outcomes (Blum, 1992).

A very important turning point in the development of ANNs was at the Dartmouth Summer Research Project of Artificial Intelligence (AI) in 1956. This project provided the momentum for many different projects in the 1950s and 1960s such as MADELINE (Multiple ADAptive LInear Element), the first neural network to be applied to a real word problem. This application consists of adaptive filters to eliminate echoes on telephone lines. MADELINE has been in commercial use for several decades (Blum, 1992). The 1960s were also a period where the potential of neural networks were blown out of proportion. "Some observers were disappointed as promises were left unfulfilled. Others felt threatened by the thought of "intelligent machines"" (Nelson & Illingworth 1991:29).

Neural network development came almost to a standstill in the 1970s but research continued in Japan and Europe. Interest in neural networks renewed in 1982 when John Hopfield presented his neural network paper to the National Academy of Sciences. The emphasis was on practicality. He showed how these networks worked and what it could do (Nelson & Illingworth, 1991).

Various disciplines became interested in the use of ANNs to address complex problems in the last two decades, ranging from cognitive psychology, physiology, medicine, computer science, electrical engineering, economy and even philosophy. ANNs have barely reached its late infancy stage. "Hopefully the rich blend of

intellects and backgrounds and divergent objectives will continue the quest" (Nelson & Illingworth, 1991:34).

3.3 Definition of Artificial Neural Networks

Artificial neural networks (ANNs) are a new information processing technique that attempts to simulate or mimic the processing characteristics of the human brain (Medsker, Turban & Trippi, 1993). An artificial neural network is an algorithm for a cognitive task, such as leaning or optimization, recognition of a pattern or retrieval of large amounts of data (Muller & Reinhardt, 1990). ANNs were inspired by studies of the central nervous system and the brain (Medsker et al., 1993; Klimasauskas, 1993) and therefore share much of the terminology and concepts with its biological counterpart. This biological analogy will be discussed in the next section.

3.4 "Anatomy" and "Physiology" of Artificial Neural Networks: A Discussion of Concepts and Terms

Neural networks were initially developed to gain a better understanding of how the brain works. It resulted in computational units, called neural networks, that work in ways similar to how we think the neurons in the human brain work. Several human characteristics such as "learning, forgetting, reacting or generalizing" and also the biological aspects of networks consisting of neurons, dendrites, axons and synapses were ascribed to these artificial neural networks in order to promote understanding of these abstract terms (Nelson & Illingworth, 1991). Some of the terminology of neural networks will be reviewed briefly.

3.4.1   Biological Neural Networks

The human brain is composed of cells called neurons. Estimates of the number of neurons in the human brain range up to 100 billion (Medsker, et al., 1993). Neurons function in groups called networks. Each network contains several thousand highly interconnected neurons where each neuron can interact directly with up to 20 000 other neurons (Nelson & Illingworth, 1991). This architecture can be described as parallel distributed processing, where the neurons can function simultaneously (Muller & Reinhardt, 1990). In contrast with conventional computers which process information serially, or one thing at a time, the human brain's parallel processing ability enables it to outperform supercomputers in some areas regarding complexity and speed of problem solving such as pattern recognition (Blum, 1992).

A typical biological neuron (Figure 3.1) consists of a cell body containing a **nucleus**, **dendrites** which provides input to the cell and an **axon**, which carries the output signal from the nucleus (Hawley, Johnson & Raina, 1993). Very often, the axon of one neuron merges with the dendrites of a second neuron. Signals are transmitted through synapses. A synapse is able to increase or decrease the strength of the connection and causes inhibition or excitation of a subsequent neuron (Nelson & Illingworth, 1991). Although there are many different neurons, this typical neuron serves as a functional basis to make further analogies to artificial neural networks.
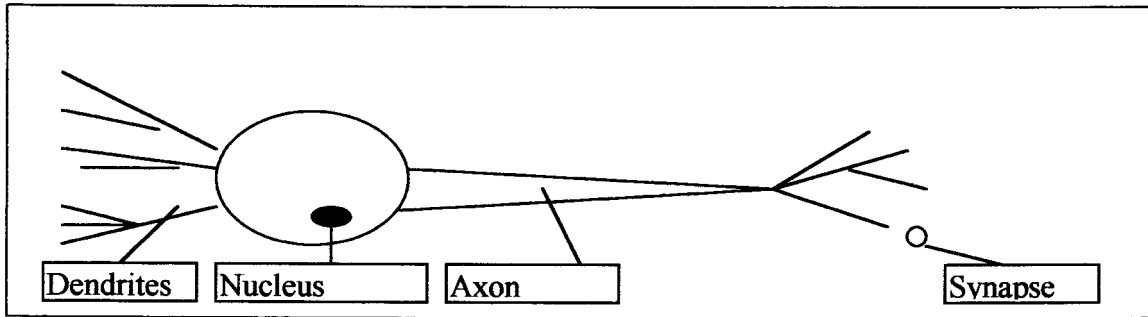
Figure 3.1: A biological neuron (Medsker, et al., 1993:5)

"Artificial neural networks are based only loosely on biology. We don't understand how the brain works or what intelligence really is." (Nelson & Illingworth, 1991: 41) ANNs are an attempt to create a technique to process information in the same fascinating and highly successful way the human mind does. It borrows features from the biological system, to enable scientists to form compact representations of complex problems such as image recognition and forecasting or prediction.

3.4.2 Artificial Neural Networks (ANNs):

The building blocks of an artificial neural network is also referred to as a neuron, or sometimes as a node, a perceptron or processing element. These artificial neurons (Figure 3.2) bear only a modest resemblance to biological neurons in the sense that they can perform approximately three of the processes we know neurons perform. (Biological neurons perform about 150 processes in the human brain) (Nelson & Illingworth, 1991). An artificial neuron receives an incoming stimulus **(input)** analogous to the impulses that the dendrites of biological neurons receive. The second function is to **calculate** a total for the combined input signals and compare it to some

threshold level. Finally, it determines the **output** and sends it out just like a biological

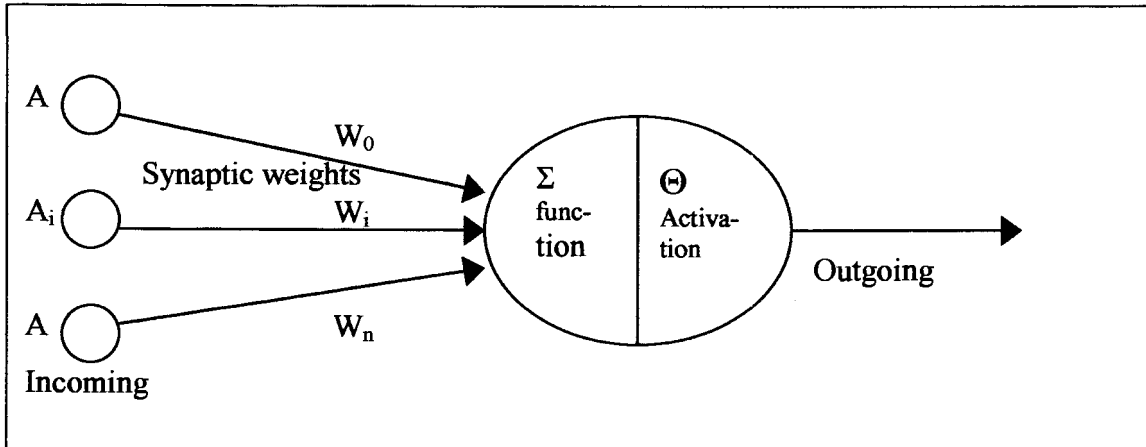neuron sends out an output through its axon (Muller & Reinhardt, 1990).

Figure 3.2: An artificial neuron (Blum, 1992: 37).

Several of these artificial neurons or nodes can be combined to make a **layer** of nodes
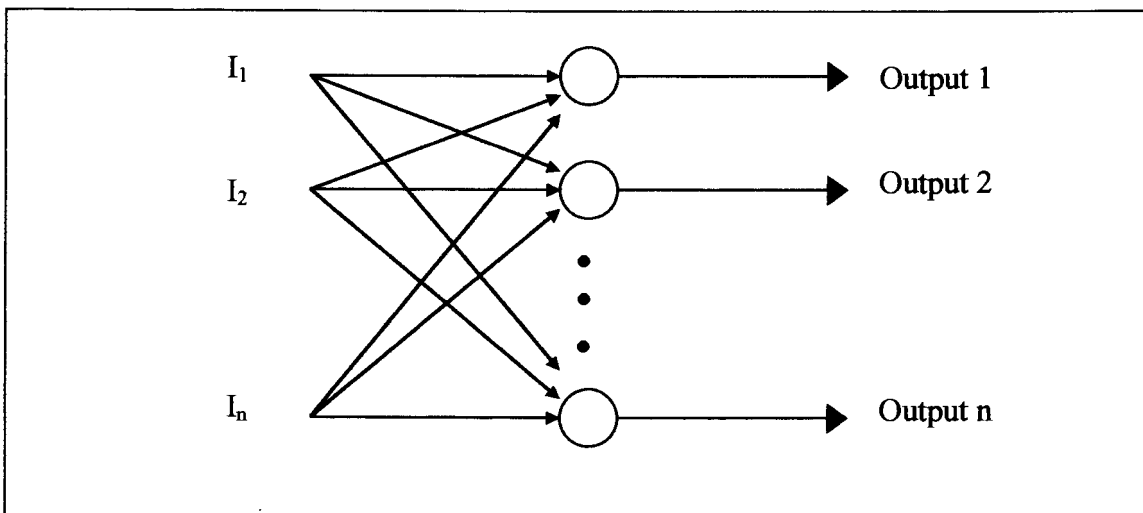
as illustrated in Figure 3.3.

Figure 3.3: Inputs to several nodes to form a layer (Nelson & Illingworth, 1991: 49).

To form an artificial neural network (Figure 3.4), several layers are connected to each other.
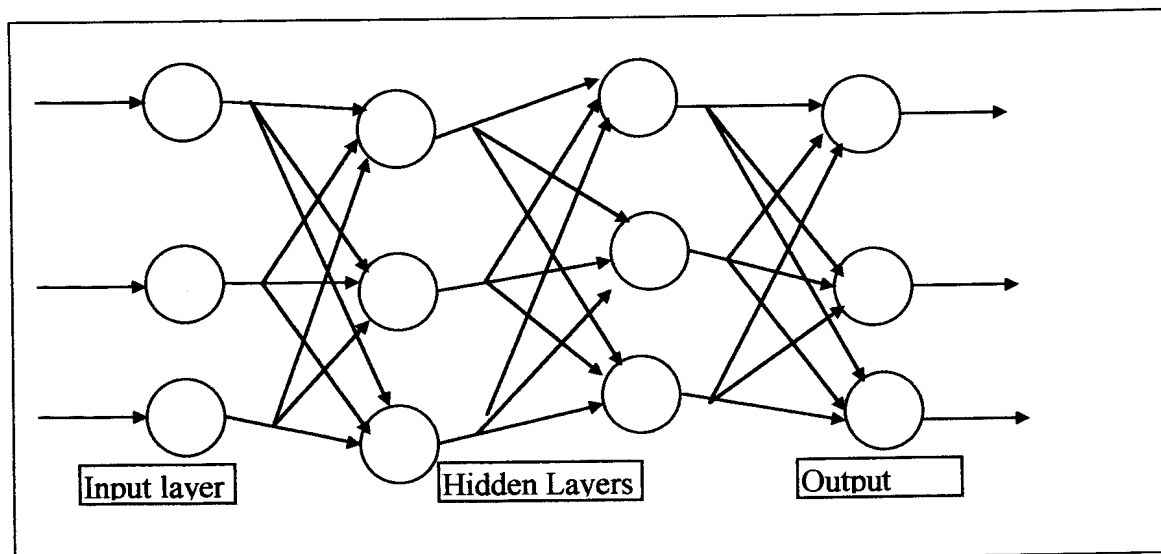

Figure 3.4: Connection of several layers to form a network (Nelson & Illingworth, 1991:50).

The first layer that receives the incoming stimuli is referred to as the **input layer.** The network's outputs are generated from the **output layer** and all the layers in between are called the **hidden layers** or **middle layers.**

The "anatomy" of artificial neural networks has just been reviewed. The terminology used in the "physiology" of an artificial neural network will be discussed next.

The first layer of neurons, called the input layer, receives the incoming stimulus. The next step is to calculate a total for the combined incoming stimuli. In the calculation of the total of the input signals, there are certain weighting factors: Every input is given a relative **weight** (or mathematical value) which affects the impact or importance of that input. This can be compared to the varying synaptic strengths of

the biological neurons. Each input value is multiplied with its weight value and then all the products are added up for a weighted sum. If the sum of all the inputs is greater than the threshold, the neuron generates a signal (output). If the sum of the inputs is less than the threshold, no signal (or some inhibitory signal) is generated. Both types of signals are significant (Blum, 1992; Nelson & Illingworth, 1991). These weights can change in response to various inputs and according to the network's own rules for modification. This is a very important concept because it is through repeated adjustments of weights that the network "learns" (Medsker, et al., 1993).

Medsker, Turban and Trippi (1993:10) summarized the crucial steps of the **learning** process of an artificial neural network very effectively:

"An artificial neural network learns from its mistakes. The usual process of learning or training involves three tasks:

1) Compute outputs.

2) Compare outputs with desired answers.

3) Adjust the weight and repeat the process."

The learning process usually starts by setting the weights randomly. The difference between the actual output and the desired output is called Δ. The objective is to minimize Δ, or even better, eliminate Δ to zero. The reduction of Δ is done by comparing the actual output with the desired output and by incrementally changing the weights every time the process is repeated until the desired output is obtained.

59

Hawley, et al. (1993) compared the learning process of an artificial neural system (ANS) with the training of a pet: "An animal can be trained by rewarding desired responses and punishing undesired responses. The ANS training process can also be thought of as involving rewards and punishments. When the system responds correctly to an input, the "reward" consists of a strengthening of the current matrix of nodal weights. This makes it more likely that a similar response will be produced by similar inputs in the future. When the system responds incorrectly, the "punishment" calls for the adjustment of the nodal weights based on the particular learning algorithm employed, so that the system will respond differently when it encounters similar inputs again. Desirable actions are thus progressively reinforced, while undesirable actions are progressively inhibited." (Hawley, et al. (1993:33)

The learning of a neural network takes place in its **training** process. Every neural net has two sets of data, a **training set** and a **test set**. The training phase of a neural network consists of presenting the training data set to the neural network. It is in this training process, that the network adjusts the weights to produce the desired output for every input. The process is repeated until a *consistent* set of weights is established, that work for all the training data. The weights are then "frozen" and no further learning will occur. After the training is complete, the data in the test set is presented to the neural network. The set of weights as calculated by the training set is then applied to the test set. The presentation of the test set is the final stage in the neural network where the answer is given whether it is to predict an outcome, find a correlation, or recognize a pattern (Blum, 1992; Medsker, et al., 1993; Nelson & Illingworth, 1991). This type of learning, where a training set of actual data is used to train the neural net, is also referred to as **supervised learning** (Nelson & Illingworth,

1991). Some neural nets learn through **unsupervised learning** where there are no data available to train on. Such a network looks for regularities or trends in the input signals and makes adaptations according to the function of the network. "At the present state of the art, unsupervised learning is not well understood and is still the subject of much research." (Nelson & Illingworth, 1991:133).

Another term that justifies some explaining is the **programming of a neural network.** "Artificial neural networks are basically software applications that need to be programmed" (Medsker, et al., 1993:22). A great deal of the programming is about the training algorithms, transfer functions and summation functions. According to Medsker, et al. (1993) it makes sense to use standard neural network software where computations are preprogrammed. Several of these preprogrammed neural networks are available on the market. Every person using an artificial neural network however, has certain additional programming that needs to be done. It might be necessary to program the layout of the database, to separate the data into two sets, namely, a training set and a test set, and lastly to transfer the data to files suitable for input into the standard artificial neural network.

The basic components of a general neural network have been discussed. The next section will review different types of neural networks.

3.5   Different Types of Artificial Neural Networks

There are different types of neural networks, categorized by their topology (the number of layers in the network). To provide just a limited overview of the basic

61

types of neural networks, the single layer network, the two layer network and multi layer networks will be discussed briefly (Rao & Rao, 1995).

3.5.1   Single Layer Networks

The single layer network has only one layer of neurons and can be used for pattern recognition. The specific type of pattern recognition in this case is called autoassociation, where a pattern is associated with itself. When there is some slight deformation of the pattern, the network is able to relate it to the correct pattern.

3.5.2   Two Layer Networks

Some models have only two layers of neurons, directly mapping the input patterns to the outputs. Two layer models can be used when there is good similarity of input to output patterns. When the two patterns are too different, hidden layers are necessary to create further internal representation of the input signals. Two layer networks are capable of heteroassociation where the network can make associations between two slightly different patterns (Blum, 1992; Nelson & Illingworth, 1991).

3.5.3   Multi Layer Networks

Several types of multi layer networks exist. The most common multi layer network is the back propagation network. According to Rao & Rao (1995), over 80% of all neural network projects in development use back propagation. "Back propagation is

the most popular, effective, and easy-to-learn model for complex, multi layered networks." (Nelson & Illingworth, 1991:121). Most back propagation networks consist of three layers, an input layer, an output layer and a hidden or middle layer (Figure 3.5). The connections between the layers are forward and are from each neuron in one layer to every neuron in the next layer.
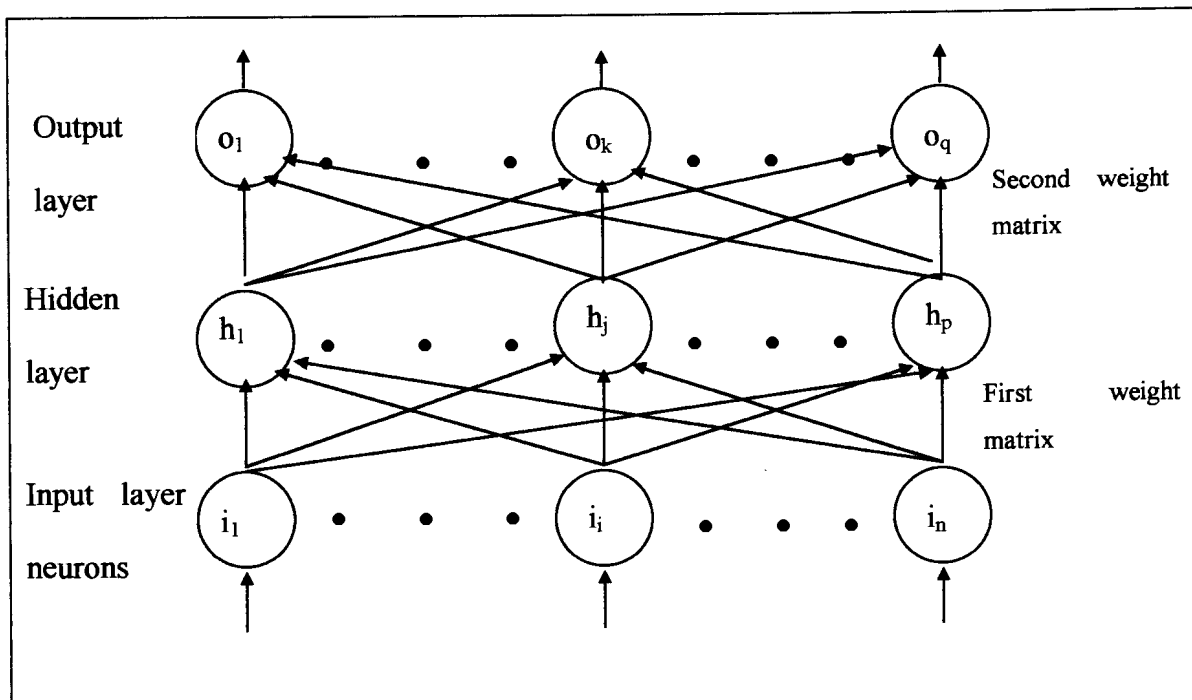
Figure 3.5: Diagram of a back propagation neural network (Blum, 1992: 56).

There are two phases in the learning process of a back propagation neural network, and it works the same as the learning process described earlier. The first stage propagates the input signal through the network in a forward manner (from the input layer, to the hidden layer, to the output layer). The second stage is to calculate the difference between the actual output and the desired output and to adapt the output, by changing the weights in the network in a backward manner (from the output layer, to

63

the hidden layer, to the input layer). The error signals of the output are propagated back into the network for each cycle. At each back propagation, the hidden layer neurons adjust the weights of connections and reduce the error in each cycle until it is finally minimized (Blum, 1992). This process was summarized by Nelson and Illingworth, (1991: 122): "The whole sequence involves two passes: a forward pass to estimate the error, then a backward pass to modify weights so that the error is decreased." Back propagation networks require supervised learning where the network is trained with a set of data (training set) similar to the test set.

3.6   Current Applications of Artificial Neural Networks

Current applications of artificial neural networks include forecasting, image recognition, text processing and optimization (Blum, 1992).

3.6.1   Forecasting or Prediction

Intelligent forecasting is predicting future events based on historical data. A set of "historical" data can be chosen for a neural net to form a set of pattern associations. Once a neural network is trained with the pattern associations of input and output factors of the historical data, the net will "recall" output patterns when presented with input patterns. When a new set of data is presented to the trained neural net, the network can predict future events by applying the trained pattern associations to the new set of inputs (Blum, 1992).

An example of the prediction ability of neural networks is the "Airline Marketing Tactician" (AMT) from a company called BehavHeuristicts, Inc. in Silversprings, Maryland. This system is trained to monitor patterns on seat bookings on airplanes, pricing, no-show rates of passengers, etcetera, to maximize profit and minimize overbooking. The system predicts demand and no-show rates and advises a user to raise or lower the number of seats for each fare (Nelson & Illingworth, 1991). The prediction ability of neural networks is also very commonly used in the financial markets. "Financial applications that require pattern matching, classification, and prediction such as corporate bond rating, credit evaluation, and underwriting have been proven to be excellent candidates for this new technology "(Salchenberger, Cinar & Lash, 1993:230).

Blum (1992) specifically referred to the excellent forecasting and prediction abilities of back propagation neural networks. Several other investigators also proved back propagation neural networks to be highly applicable in the prediction of bankruptcy (Odom & Shara, 1993; Raghupathi, Schkade & Raju, 1993; Rahimian, Singh, Thammachote & Virmani, 1993). Odom and Shara (1993) specifically compared the predictive ability of a neural network and multivariate discriminant analysis model in bankruptcy prediction. The authors concluded that the neural network performed better on both the original set of data and the holdout sample (training set and test set). Salchenberger et al. (1993) confirmed the findings that back propagation neural networks predicted more accurately than any other method originally used. These research findings show promise in using back propagation neural networks for prediction purposes.

3.6.2 Image Recognition

An example of the image recognition ability of neural networks is the project of Paul Gorman of Bendix Aerospace (cited in Nelson & Illingworth, 1991). He trained a neural network to recognize underwater targets by sonar and to tell the difference between a mine and a rock shaped like a mine. The neural network performed better than trained human listeners or the traditional technique called nearest neighbor classifier and could recognize 90% of the mines correctly.

The area of image recognition also include recognition of handwriting, recognition of human speech (Blum, 1992) and even to estimate speech intelligibility of hearing impaired speakers (Metz, Schiavetti & Knight, 1992). In this last study, a back propagation neural network was used to predict the intelligibility of hearing-impaired speakers from acoustic speech parameters. The study attempted to classify hearing impaired persons into 4 groups of varying speech intelligibility. The network very successfully classified hearing impaired persons into the first and last group (most and least intelligible) but the neural network experienced difficulty classifying middle categories probably due to the variable chosen to separate the different classes. This experiment is currently being expanded to improve network performance.

3.6.3 Text Processing

An example of a neural network's text processing abilities is a simple spell checker, designed by Jagota and Jung of SUNY, Buffalo (cited in Blum, 1992). Text processors can also be combined with speech recognition systems. Some types of

neural networks are bi-directional and can perform both functions where inputs and outputs can be reversed to achieve the desired function. If such a bi-directional system is given a word, it can return the pronunciation or the corrected spelling or both, depending on the specified output and input (Blum, 1992).

3.6.4 Optimization

Neural networks can also be used to solve difficult optimization problems such as cost minimization where numerous factors can influence a manufacturing process (Blum, 1992). An example of such an application is used in the GTE Laboratories fluorescent bulb manufacturing plant (cited in Nelson & Illingworth, 1991). A neural network was trained to monitor the production line and keep track of all the variables that influence production such as heat, pressure and the chemicals used to make the bulbs. The neural network determines and monitors optimum manufacturing conditions and can shut down the plant in emergency situations.

3.7 Advantages of Artificial Neural Networks over Conventional Statistical Methods

"One could argue that in many cases it would be *possible* to formulate a statistical approach to the same problem. For example in the image recognition applications, the program could make probabilistic guesses about what character is being viewed based on the results of a statistical model. There are several problems in this approach, however, which is why progress in the fields of pattern recognition and handwriting recognition was so slow prior to the advent of applied neural networks" (Blum,

1992:7). Some of the advantages of artificial neural networks as described by Blum, (1992) will be reviewed briefly.

3.7.1 Less Need to Determine Relevant Factors A Priori

To formulate a statistical model, one should know what factors one wish to correlate. With neural networks, irrelevant data has such low connection strength that it has no effect on the outcome. Neural networks excel at determining what data is relevant.

3.7.2 Sophistication of the Model

When hundreds of factors are at play, even if some only have a very small effect, neural network models are much more likely to be more accurate for difficult problems than any statistical model.

3.7.3 Directness of the Model

A statistical method is a more indirect way of learning correlations, where artificial neural networks model a problem directly. The example the Blum (1992) describes is to map pixelated images to alphabet letters. A neural network would simply connect the objects (all pixels of the image are neurons and are connected through a hidden layer to the output neurons that guess the letter). If a statistical method were used, the first step would have been to determine factors that are likely to influence the guess of the character. The next step to formulate a statistical model, run the model, analyze

the results and then to build a system that incorporates the results. If the character can still not be identified correctly, the whole process should be repeated with other factors that are likely to influence the guess. Although it is possible to solve a problem like this with a statistical model, it requires much more time, planning and trial and error guesses.

3.7.4 Fault Tolerance

Neural networks are extremely fault tolerant and can learn from and make decisions based on incomplete data (Nelson & Illingworth, 1991). Even if some of the hardware fails, the neural network system will not be considerably changed. Blum (1992) even suggests to train on noisy data to possibly enhance post training performance.

3.7.5 Inherent Parallelism

ANNs simulate the human brain's parallelism where neurons are highly interconnected and function independently and in parallel. There are no time dependencies among synapses of the same layer all of them can work in parallel and simultaneously. Although digital computers have to simulate this parallelism, true neural network hardware really perform operations in parallel. This feature makes very fast decisions possible and the solving of very complex problems (Blum, 1992, Nelson & Illingworth, 1991).

3.8   Limitations of Neural Networks

"There is still a tendency to portray neural networks as magical, a sort of black box that does magical things" (Nelson & Illingworth, 1991:263). ANNs however, have a number of limitations that should be reviewed (Nelson & Illingworth, 1991).

Neural networks do not excel in precise exact answers. It can for example, not be used to do finances. Neural networks have the tendency to generalize.

Neural networks can not count. Counting has to take place in a sequential mode and neural networks function in parallel.

Designing a neural network is somewhat of a mysterious process. The learning process of a neural network is a tedious and painstaking trial-and-error effort. There are no standards for learning algorithms for ANNs. Another factor of importance influencing the learning process is the quality of the material that is used to train on.

Scaling is another problem. The networks may perform very well on the training and test set in the laboratory but less well as soon as it is implemented as a commercial model.

Another limitation is that ANNs can sometimes generalize or guess incorrectly. These mistakes are hard to undo since it spreads out through the network. Back propagation algorithms address this issue by extensive training on a set of data before any generalizations or guesses are made.

"In general, a neural network can not justify its answers. There is no facility to match the "how" or "why" found in expert systems. There is no way to stop it and say, "What are you doing now?" It is as if the network were instead saying, "Trust me, trust me." (Nelson & Illingworth, 1991:75). There are current efforts to build "knowledge extraction tools" for neural networks also called "justification systems" to verify the learned relationships directly (Blum, 1992).

3.9   Summary

DPOAE measurements are potentially a fantastic new objective, rapid, non-invasive, inexpensive and accurate test of auditory sensitivity. Conventional statistical methods however, could not yet provide a general rule to predict pure tone thresholds given DPOAE results.

Artificial neural networks are a new information processing technique proved to be highly applicable in the areas of prediction and correlation finding. The application of neural networks to the field of audiology, specifically, DPOAEs to predict pure tone thresholds, could result in an ideal objective testing procedure for special populations. It would have a profound positive effect on current screening procedures, as well as the differential diagnosis of sensorineural hearing losses, in the assessment of the peripheral ear.

4 Research Methodology

4.1 Introduction

Leedy (1993) gave one very interesting viewpoint on the essence of research methodology. "The process of research, then, is largely circular in configuration: It begins with a problem; it ends with that problem solved. Between crude prehistoric attempts to resolve problems and the refinements of modern research methodology the road has not always been smooth, nor has the researcher's zeal remained unimpeded." (Leedy, 1993:9).

The problem inspiring this research project has already been elaborately stated in Chapter 1. In short, the need for an objective, non-invasive and rapid test of auditory functioning has led to numerous previous studies attempting to develop such a procedure. Shortcomings in conventional statistical methods prevented accurate predictions of hearing ability with distortion product otoacoustic emissions. A new form of information processing called artificial neural networks might prove useful in the solving of this problem.

4.2 Aims of Research

The aims of this research project are:

4.2.1   Main Aim

The main aim is to predict hearing ability at 500 Hz, 1000 Hz, 2000 Hz, and 4000 Hz with distortion product otoacoustic emission (DPOAE) responses in normal and hearing-impaired ears with the use of artificial neural networks.

4.2.2   Sub Aims

The first sub aim is to determine optimal neural network topology to ensure accurate predictions of hearing ability at 500 Hz, 1000 Hz, 2000 Hz and 4000 Hz. The number of input nodes and number of output neurons are determined by the number of input- and output data. The number of middle layer neurons however, should be determined by trial and error until the required accuracy of prediction in the training stage is reached.

The second sub aim is to train a neural network with sufficient data to predict pure tone thresholds with DPOAE results. Sufficient data implies enough data from different categories of hearing loss to ensure accurate training and prediction of various hearing abilities.

The third sub aim is to determine the possible effects of age and gender on the distortion product.

4.3 Research Design

For this research project, the chosen research design was a multivariable correlational study (Leedy, 1993). The correlation between selected variables of DPOAE and selected variables of pure tone thresholds was studied by the use of artificial neural networks.

The selected variables for DPOAE measurements are:

1. The frequency of f1.

2. The frequency of f2.

3. The loudness level of f1 (L1).

4. The loudness level of f2 (L2).

5. The pattern of present and absent DPOAE responses of 8 DP Grams.

6. The age and gender variables.

The selected variables for pure tones are:

1. The frequency of the pure tone.

2. The lowest dB level where a response can be measured 50% of the time.

4.4 Subjects

For this study, data obtained from 70 subjects (120 ears, in some cases only one ear fell within subject selection specification) were used to train a neural network to

predict pure tone thresholds given only the distortion product responses. Subjects were recruited from a private audiology practice as well as a school for hard of hearing children. The subjects included 28 males and 42 females, ranging from 8 to 82 years old.

4.4.1 Criteria for the Selection of Subjects

The following criteria were chosen for the selection of subjects.

4.4.1.1 Hearing Ability

In order to train a neural network with sufficient data to make an accurate prediction of hearing ability, data across all groups of hearing impairment were needed. For this study, subjects were chosen that had varying hearing ability, ranging from normal to moderate severely sensorineural hearing impaired. To obtain an equal amount of data in different areas of hearing impairment, data in three different categories of hearing impairment were included, namely normal hearing ability, mild hearing losses and moderately severe hearing losses.

There are two general classification systems to classify hearing level as being normal or impaired (Yantis, 1994). The first method converts hearing levels into a rating scale based on percentage. A Pure tone threshold average (PTA) for the frequencies 500 Hz, 1000 Hz, 2000 Hz and 3000 Hz is calculated, 25dB is subtracted (which is assumed to be the normal range) and the answer is multiplied by 1.5 to find percentage of impairment for each ear.

The second approach to describe normal ranges and hearing impairment also uses monaural PTA in the speech frequencies but adds additional descriptors to the different levels. Clark 1981 (cited in Yantis, 1994) modified Goodman's recommendations from 1965 into the following categories:

-10 to 15dB	Normal hearing
16 to 25dB	Slight hearing loss
26 to 40dB	Mild hearing loss
41 to 55dB	Moderately severe hearing loss
56 to 70dB	Severe hearing loss
91dB plus	Profound hearing loss

For this study, the second approach to classification of hearing impairment (as used by Clark, 1981 in Yantis, 1994) was used. Subjects with normal hearing, slight hearing loss, mild hearing loss and moderately severe sensorineural hearing loss were included in the study. To divide the subjects into three groups of 40 ears each, the group with normal hearing ranged from 0 dB to 15 dB. The group with slight and mild hearing loss ranged from 16 to 35dB and the moderately severe hearing-impaired group had PTAs in the range of 36 - 65dB. It should be noted that according to Clark's (1981) (cited in Yantis, 1994) specification the moderate hearing loss group only includes hearing losses of up to 55 dB, whereas the severely hearing impaired group extends to 70 dB. DPOAEs has been reported in ears that have a hearing threshold as high as 65dB HL (Moulin, et al., 1994) at the frequencies close to the primaries. It was therefore decided to combine the category of moderate and severe

76

hearing impairment to form the category moderately severe hearing impairment ranging from 36 to 65 dB HL. The data was divided into three groups merely to ensure that an equal amount of data was obtained in each category. Another modification to Clark's classification system has been made. In addition to the frequencies used by Clark (1981) (cited in Yantis, 1994) to determine the PTA, namely 500 Hz, 1000 Hz, 2000 Hz and 3000 Hz, for this study 4000 Hz was also taken in consideration in the classification of hearing impairment. The reason for this modification is that DPOAE measurements are required at 4 kHz to predict the pure tone threshold at 4 kHz.

4.4.1.2 Middle Ear Functioning

The second selection criterion was normal middle ear functioning. Otoacoustic emissions can only be recorded in subjects with normal middle ear function. Only a very small amount of energy is released by the cochlea and is transmitted back through the oval window and ossicular chain to vibrate the tympanic membrane. Normal middle ear function is crucial to this transmission process (Norton, 1993; Osterhammel, Nielsen & Rasmussen, 1993; Zhang & Abbas, 1997).

Normal middle ear functioning was determined by otoscopic examination and tympanometry.

4.4.1.2.1 Otoscopic Examination

Otoscopic examination was performed to determine the amount of wax in the ear canal, for excessive wax may block the otoacoustic emission microphone and prevent the reading of a response. The second aspect that was investigated was the light reflection on the tympanic membrane, indicative of a healthy tympanic membrane (Hall III & Chandler, 1994).

4.4.1.2.2 Tympanometry

A subject's tympanometry results must have been within the following specifications to be included in the study:

A normal type A tympanogram was one of the criteria for normal middle ear functioning. A type A tympanogram has a peak (or point of maximum admittance) of 0 to -100 daPa. The peak may even be slightly positive, for example +25daPa (Block & Wiley, 1994). A type A tympanogram's static immittance when measured at 226 Hz ranges from about 0.3 cm^3 to 1.6 cm^3 (Block & Wiley, 1994). Subjects demonstrating type A tympanograms within these specifications were accepted for the study.

4.4.1.3   Attention Span

Only persons that were able to cooperate for approximately an hour were included in the study. Subjects had to be able to follow instructions and sit quietly and still in one position for about forty minutes for DPOAE testing. Subjects demonstrating inadequate ability to follow instructions or cooperate during pure tone audiometry, tympanometry or DPOAE testing were not included in the study. Some of the reasons subjects were excluded from the study in this regard include very young age, ill health and hyperactivity.

4.4.1.4   Criteria Regarding subject Age and Gender

There is some debate regarding the effect of age on distortion product otoacoustic emissions. In a study by Lonsbury-Martinet al. (1991), a negative correlation between DPOAE measurements and age for subjects 20-60 years was reported. In their report however, it is suggested that this negative correlation is due to changes in hearing threshold associated with aging. A study by Stover and Norton (1993) (cited in He & Schmiedt, 1996) also indicated that the difference in DPOAEs between younger and older subjects can be attributed to the sensitivity changes, rather than the aging itself. According to He and Schmiedt (1996) a 60-year-old person with normal hearing (PTA < 15dB) will therefore have the same DPOAEs as a 12-year-old with the same pure tone threshold levels.

There was therefore no selection criteria regarding age. The only population that was excluded in this study is the pediatric population, due to differences in middle ear

properties such as canal length, canal volume and middle ear reverse transmission efficiency that may cause differences in DPOAE amplitudes (Lasky, 1998a; Lasky, 1998b; Lee, Kimberley & Brown, 1993).

There was also no selection criteria regarding gender. Gaskill and Brown (1990) and Cacace et al. (1996) reported that DPOAEs were significantly larger in female than male subjects tested in the frequency range of 1000- 5000Hz. Both studies however, indicated that the female subjects in their studies had more sensitive auditory thresholds than the males (an average of 2.4 dB better). The differences found between the two groups could therefore not be explained by gender only.

Lonsbury-Martin et al. (1990) conducted a study to investigate basic properties of the distortion product including the effect of gender on the prevalence of DPOAEs. A comparison of DPOAE amplitudes and thresholds failed to reveal any significant differences except a minor difference at 4 kHz.

Gender effects on DPOAEs are apparently limited to minor differences in DPOAE amplitudes and thresholds and therefore gender was not one of the selection criteria for this study.

4.4.2   Subject Selection Procedures

The procedure in which subjects were selected started with a brief interview, following an otoscopic examination of the external meatus, tympanometry and pure tone audiometry.

4.4.2.1   Case History and Personal Information

A short interview was performed to obtain a limited case history and some personal information. The research project was also discussed with the subject in a very brief manner and any questions answered. The purpose of the case history was to obtain enough personal information to open a new subject file and obtain the subject's age and gender for later studies of these effects on DPOAEs. Second, information regarding hearing status such as any complaints of tinnitus and vertigo, the amount of noise exposure and complaints of middle ear problems was obtained. In the analysis of data, some subjects may exhibit abnormal DPOAEs in conjunction with normal pure tone thresholds. In a study by Attias, Furst, Furman, Horowitz and Bresloff (1995), it was found that in some cases, subjects with normal pure tone thresholds of 0 dB exhibited abnormal otoacoustic emissions, due to noise exposure. The effects of noise exposure can clearly be seen long before the actual hearing loss occurs. This is also true for ototoxic medication (Danhauer, 1997). By obtaining this background information, possible reasons for abnormal DPOAEs may be formulated. Appendix A reviews the aspects which were addressed in the short interview. This short interview lasted approximately 10 minutes.

4.4.2.2   Otoscopic Examination

Otoscopic examination of both ears was performed, with special attention to the light reflection of the eardrum and excessive wax in the ear canal. The duration of the otoscopic examination was about 3-5 minutes.

4.4.2.3   Tympanometry

The next step in the subject selection procedure was to obtain a tympanogram to determine middle ear functioning. The subject was instructed to sit in front of the tympanometer and not to speak or swallow. Tympanometry was performed in both ears and the duration of the procedure was about 5 minutes.

4.4.2.4   Traditional Audiogram

If the subject had normal middle ear functioning, the subject selection procedure continued. A traditional audiogram was obtained from the subject. The frequencies that were tested during pure tone air conduction was 125 Hz, 500 Hz, 1000 Hz, 2000 Hz, 4000 Hz and 8000 Hz. If a hearing loss was present, or if any of the frequencies except 8000 Hz had a threshold >15 dB, then pure tone bone conduction was also performed. If sensorineural hearing losses varied with more than 15 dB between adjacent frequencies, in between frequencies such as 3000 Hz or 750 Hz were also tested. Only subjects with sensorineural hearing losses (no gap between air conduction and bone conduction) were accepted for the study. Threshold determination was in 5dB steps and a threshold was defined as 50% accurate responses at a specific dB level (Yantis, 1994).

Audiograms from subjects were then analyzed. All audiograms indicating normal hearing (500 Hz, 1000 Hz, 2000 Hz, 3000 Hz and 4000 Hz below 15 dB) were included in the first group. Audiograms indicating hearing loss were analyzed in terms of the degree and configuration of the hearing loss. Mild hearing loss,

indicating a hearing loss between 16-35 dB in the frequency region 500-4000 Hz were categorized in the second group, namely mild hearing loss. Audiograms indicating hearing loss of 36-65 dB in the frequency region of 500-4000 Hz were categorized in the third group, namely moderately severe hearing loss. In each category, 40 audiograms were included.

The duration of pure tone audiometry was approximately 15-20 minutes.

If a subject demonstrated normal middle ear functioning and a pure tone audiogram that could be categorized into one of the three groups, DPOAE measurements were performed within the next hour. This procedure will be discussed in data collection procedures.

4.5   Apparatus

The apparatus for the different sections of research were as follows:

4.5.1   Subject Selection Apparatus

- For otoscopic examination of the external meatus and tympanic membrane an otoscope was used, specifically the Welch Allyn pocketscope model 211.

- For tympanometric measurements the GSI 28 A middle ear analyzer, calibrated April 1997 was used (Testing was performed in January 1998).

- For determination of auditory pure tone thresholds, the GSI 60 Audiometer, calibrated April 1997 was used. The model of the earphones on the audiometer was 296 D 200-2. Pure tone thresholds were measured in a sound proof booth.

4.5.2   Data Collection Apparatus

- The measurement of Distortion Product Otoacoustic Emissions were conducted with a Welch Allyn GSI 60 DPOAE system and the probe was calibrated for a quiet room in January, 1998. All measurements were made in a quiet room.

4.5.3   Data Preparation Apparatus

- For the preparation of data files, a Pentium 200 MMX computer was used. The software included Excel for Windows 1998.

4.5.4   Data Analysis Apparatus

- For the training of the neural network, the back propagation neural network from the software by Rao and Rao, 1995 (in addition to the book) was be used. The neural network was trained on a Pentium 200 MMX.

- Further analysis of data was performed in Excel for Windows 1998 and with custom software.

4.5.5   Preliminary Study

The reason for the preliminary study was twofold: First, to determine which persons may participate as subjects and second, which stimulus parameters to use in the measurements of DPOAEs.

4.5.5.1   Determination of Subject Selection Criteria

A very large part of the determination of subject selection criteria was based on an extensive overview of related literature. The researcher did however conduct a series of DPOAE measurements on subjects with various categories of hearing ability to confirm current subject selection criteria. Just a few of the interesting finds during DPOAE measurement of the preliminary study will be discussed briefly.

To confirm the studies of the importance of normal middle ear functioning by researchers such as Zhang and Abbas, (1997); Osterhammel et al., (1993); Hall III et al., (1993) and Kemp et al., (1990), a few DPOAE measurements were performed on subjects that displayed acceptable hearing ability for this study but small variations in tympanometric results. One subject had perfect hearing (pure tone hearing thresholds of 0 dB HL at all frequencies) but no airtight seal could be obtained as a result of grommets in the tympanic membrane. This subject displayed very high levels of low frequency background noise during DPOAE testing and it was difficult to distinguish the DPOAE responses from the noise floor at most of the low and mid frequencies. Another subject had a mild sensorineural hearing loss but the tympanogram's compliance was just below 0.3cc. This subject also demonstrated very high levels of

low and mid frequency noise with indistinguishable DPOAE responses above the noise floor. A normal type A tympanogram with static compliance of 0.3-1.75cc was therefore set as one of the subject selection criteria.

A few measurements were also made in the ears of severely hearing impaired subjects and varying levels of stimuli was used. Another aspect that became apparent after a few tests were conducted was the absence of DPOAEs in persons with hearing losses greater than 65dB HL. This confirmed studies by Moulin et al., (1994) and Spektor et al., (1991) which found that when stimuli lower than 65dB SPL are used, DPOAEs can not be measured in ears with a hearing loss exceeding 65dB HL. Therefore, for this study, only subjects were included with sensorineural hearing losses of up to 65 dB HL.

These same tests also revealed that when very high intensity primaries were used (such as 70- 80dB SPL), in some instances one could observe "passive" emissions from the ears of these severely hearing impaired subjects. The reason for passive emissions, according to Mills, (1997) is that very high level stimuli can stimulate broad areas of the basilar membrane and phase relations between travelling waves can cause these "passive" emissions that do not correspond well to hearing sensitivity or frequency specificity. In this preliminary study, passive emissions were only observed when stimuli levels were higher than 70dB. It was therefore decided not to use stimuli levels higher than 70dB.

A second series of tests were conducted to determine optimal stimulus parameters.

4.5.5.2 Determination of Optimal Stimulus Parameters

Most of the stimulus parameters for this study were derived from an in depth literature study. Parameters such as the frequency ratios between the primaries, the loudness levels of L1 and L2 and whether to measure DP Grams or I/O functions were selected on recommendation of other previous studies. There are however a few stimulus parameters that requires some experimenting in order to determine applicability and practicality for a certain research project. One such example is the configuration setup, or specifically, the number of frames of data that will be collected in each measurement. The GSI-60 DPOAE system offers two possibilities, a screening option and a diagnostic option.

The screening option collects a maximum of 400 frames before stopping each primary tone presentation. Not every test runs up to 400 frames, if a very clear response is measured, the measurement can be made in as little as 10 frames. Test acceptance conditions for the screening configuration are a cumulative noise level of at least – 6dB SPL and either a DPOAE response amplitude that is 10 dB above the noise floor or a cumulative noise level of at least –18 dB SPL (GSI-60 manual, p2-44). A maximum of 400 frames are measured, and if no clear response was present, the results are labeled "timed out."

The diagnostic option runs up to 2000 frames for each primary tone presentation. The minimum number of accepted frames is 128. Test acceptance conditions are that the distortion product minus the average noise floor should be at least 17 dB.

After a few measurements in both configurations it became clear that the diagnostic option requires much more testing time. Testing time of one single DP Gram measured at low level stimuli in the diagnostic configuration could increase testing time up to 12 minutes. Even though the general noise floor was slightly lower during the diagnostic option, it was not practical to conduct 8 DP Grams in each ear with tests lasting 6-12 minutes each. It would take between an hour and one and three quarters of an hour to measure one ear alone with DPOAEs. It was therefore not practical to evaluate 120 ears with the diagnostic option. The screening option with a testing time of up to 2 minutes per DP Gram was selected for this study. One ear could be evaluated in about 15 minutes with DPOAEs and the screening procedure yielded very much the same information.

Lastly, the stimulus parameter that required some experimenting was the selection of the frequencies of the primary tone pairs. The GSI-60 DPOAE system has a "Custom DP" function where the examiner can choose any primary frequencies for DPOAE measurement. After a few tests it became clear that care should be taken when selecting primary tones. Not only should the frequency ratio of the primaries preferably be 1.2, but the frequency values from one tone pair to the next should be at least one octave apart to avoid interaction between stimuli (GSI-60 manual, p2-39). The GSI-60 measures the noise floor from the first primary tone pair per group, and if frequency pairs are selected too close to each other, very high levels of noise are being measured. So after a lot of changes in primary tone pairs were made to avoid interaction between stimuli, the researcher ended up with stimuli very similar to the default stimuli of the GSI-60. It was therefore decided to use the default primary

frequencies of the GSI-60 for this study by activating all four octaves. (It seems that those stimuli are set as default for a very obvious reason.)

Just for practicality, a few test runs that incorporated the whole data collection procedure were conducted to determine the amount of time required testing each subject. This was determined in order to schedule appointments. As seen in Table I, the whole data collection procedure lasted about an hour. In some cases, especially in the case of subjects with a hearing loss, more time was required for bone conduction but on the average, one hour was sufficient to test one subject.

Table I: Time required testing one subject

Subject history	5 minutes
Audiometry	15 minutes
Otoscopic examination	5 minutes
Tympanometry	5 minutes
DPOAE measurements left ear	15 minutes
DPOAE measurements right ear	15 minutes
Total testing time	**60 minutes**

4.6 Data Collection Procedures

The following procedure was conducted in the various stages of the research project.

4.6.1 Data Collection Procedures in the Selection of Subjects

In the selection of subjects, the procedure included a short interview, an otoscopic examination, tympanometry and pure tone audiometry. Data that was collected during the interview, the otoscopic examination and tympanometry was used for subject selection only. Data that was collected during pure tone audiometry was not only used in the selection of subjects, but also in the main purpose of the study, namely to train a neural network to predict pure tone thresholds given the distortion product responses. These procedures were discussed in **4.4.2 Subject Selection Procedures**.

4.6.2 Data Collection Procedures Conducted During Research

In order to train a neural network to predict pure tone thresholds given only the distortion product responses, two sets of data should be collected namely each subject's pure tone thresholds and each subject's DPOAEs.

4.6.2.1 Data Obtained From Pure Tone Audiometry

The necessary pure tone audiometry data has already been obtained during subject selection and the collection procedure for this set of data has been described in the section **4.4.2.4. Traditional Audiogram**.

4.6.2.2   Data Obtained from DPOAE Measurements

The second set of data that was collected was each subject's DPOAE responses. The procedure for the collection of this set of data is quite complex, due to the number of stimulus parameters that should be specified. There is a four dimensional space in which the stimulus parameters for DPOAE measurement should be specified (Mills, 1997). The frequencies of the two primary stimulus tones f1 and f2 (f1>f2), the frequency ratio of f2/f1 (how many octaves apart the two frequencies are), the loudness level of f1 (which is L1) and the loudness level of f2 (which is L2). Furthermore, the difference in loudness level between L1 and L2 should also be specified.

4.6.2.2.1   Specification of Stimulus Parameters for DPOAE Measurements

The specification of stimulus parameters will be reviewed shortly.

4.6.2.2.1.1   The Selection of the Frequencies

In the case of the GSI-60 Distortion Product otoacoustic emissions system, the number of octaves that should be tested can be specified as well as the amount of data points to plot between octaves. The octaves available are 0.5 - 1 kHz; 1-2 kHz; 2-4 kHz and 4 -8 kHz. All of these octaves was selected for DPOAE testing because information regarding all these frequencies was required to make comparisons with the audiogram in the frequency range 500 - 4000 Hz. The amount of data points between frequencies could be any number between 1 and 20. The more data points

per octave, the longer the required test time since more frequency pairs are tested between frequencies. The GSI-60 manual suggests three data points per octave to be adequate, not increasing the test time too much but yielding enough information regarding DPOAE prevalence between frequencies. In the case of the pure tone audiogram, in-between frequencies were only tested when hearing losses between frequencies varied more than 15 dB (to measure the slope of the hearing loss) and only one or in extreme cases two in-between frequencies were evaluated. The selection of three data points between octaves in the case of DPOAE measurement should therefore be adequate.

The frequencies tested by the GSI-60 when all four octaves are activated and three data points per octave is specified amount to 11 frequency pairs. The 11 frequency pairs are presented in Table II.

f1 and f2 are always presented simultaneously as a pair, f1>f2 and the f2/f1 ratio is ~1.2.

Table II: The 11 frequency pairs tested by the GSI-60 DPOAE system when all four octaves are activated.

PAIR	1	2	3	4	5	6	7	8	9	10	11
f1 Hz	500	625	781	1000	1250	1593	2000	2531	3187	4000	5031
f2 Hz	593	750	937	1187	1500	1906	2406	3031	3812	4812	6031

4.6.2.2.1.2 The Selection of the Frequency Ratio of the Primary Frequencies (f2/f1)

Several studies investigated the effect of the frequency ratio on the occurrence of DPOAEs (Cacace et al., 1996; Popelka, Karzon & Arjmand, 1995; Avan & Bonfils, 1993; He & Schmiedt, 1997).

It appears that the frequency ratio of 1.2 - 1.22 is most applicable to a wide range of clinical test frequencies (0.5-8kHz) and a wide range of stimulus loudness levels. A stimulus ratio of f2/f1 = 1.2 was therefore selected for this study.

4.6.2.2.1.3 The Selection of the Loudness Levels of the Primaries, L1 and L2.

As mentioned in the introduction, there are two ways of eliciting a DPOAE response. Either the frequencies are changed and the loudness level kept constant, this is sometimes referred to as a "distortion product audiogram" (DP Gram), or the frequencies are being kept constant while the loudness level is changed (an input/output function (I/O) is obtained). In this case, several DP audiograms were obtained. All the frequencies selected for all four octaves were presented to the subjects at different loudness levels, starting with maximum loudness levels at L1= 70 dB; L2 =60 dB. Loudness levels were decreased in 5 dB steps until DP "thresholds" (lowest intensities where DP responses can be distinguished from the noise floor) for all the frequencies were obtained. The lowest loudness level for the primaries that was tested was L1= 35 dB; L2= 25dB. Eight loudness levels were therefore evaluated resulting in eight DP "audiograms" for each ear.

4.6.2.2.1.4 The Relative Loudness Levels of the Primaries (L1 and L2)

An overview of several studies indicated the following loudness level ratios to be most suitable for the detection of DPOAEs: L1>L2 by 10dB (Stover et al., 1996a), L1>L2 by 15 dB (Gorga et al., 1993) and L1>L2 by 10-15 dB (Norton & Stover, 1994). A study by Mills (1997) indicated that more DPOAEs were recorded when L1>L2 than L1 = L2.

A loudness level ratio of L1>L2 by 10 dB was chosen for this study.

4.6.2.2.1.5 The Criteria for DPOAE Threshold

The detection threshold for a distortion product otoacoustic emission depends almost entirely on the noise floor and the sensitivity of the measuring equipment (Martin et al., 1990b). A distortion product with an amplitude less than the noise floor can not be detected (Kimberley & Nelson, 1989; Lonsbury-Martin et al., 1990). Most researchers specify a DP response to be present if the DP response is 3-5 dB above the noise floor. Harris and Probst (1991:402) specified a DP response as "the first response curve where the amplitude of 2f1-f2 is $\geq$ 5 dB above the level of the noise floor." Lonsbury-Martin et al., (1990) reported detection thresholds for DPOAE measurements 3 dB above the noise floor. Lonsbury-Martin (1994) set the criterion level for a DPOAE threshold at $\geq$ 3 dB.

For this study, a detection threshold for a DPOAE response will be defined as the first response where the distortion product (2f1-f2) is 3 dB above the noise floor.

4.6.2.2.2 DPOAE Testing Procedure

DPOAE measurements were performed directly after the subject selection procedure. Subjects were instructed to sit next to the GSI 60 DPOAE system, not to talk and to remain as still as possible. Subjects were allowed to read as long as they kept their heads as still as possible. First, a new file was opened for the subject. Then the DPOAE probe tip was inserted into the external meatus in such a manner that an airtight seal was obtained.

Eight tests or DP Grams were performed in each ear. Every DP Gram consisted of eleven frequency pairs. Every frequency pair consisted of two pure tones, f1 and f2 presented to the ear simultaneously (see Table II for the 11 frequency pairs). The eleven frequency pairs were presented to the ear in a sweep, one at a time starting with the low frequencies, ending with the high frequencies.

The first DP Gram was conducted on the loudness levels F1 = 70dB SPL, F2 = 60dB SPL. The second DP Gram was conducted 5 dB lower at F1 = 65 dB SPL, F2 = 55 dB SPL. The third DP Gram was conducted 5 dB lower than the second, namely F1 = 60 dB SPL, F2 = 50 dB SPL. A total of eight DP Grams were conducted, each one 5 dB lower than the previous one. The lowest intensity DP Gram that was performed was F1 = 35 dB SPL, F2 = 25 dB SPL.

The procedure was repeated for both ears if both ears fell within selection criteria. The duration of DPOAE testing of eight DP Grams for one ear was between 15-20

minutes. If a subject was tested binaurally, the duration of DPOAE testing was approximately 30-40 minutes.

4.7   Data Preparation Procedures

The first step in the data preparation procedures is to organize all raw data in files.

4.7.1   Creation of a Data File for Each Ear

Each ear has its own file. A file is merely a row of numbers, depicting the test results in a certain order. The first column represents the subject or file number, the second number the DP Gram number, then the ear that has been tested (left or right) and so the numbers continue until all data relating to the DPOAE testing procedure and pure tone testing results have been depicted. Table III represents a data file for one DP Gram. 8 DP Grams for each ear were conducted. The complete data file for one ear would therefore have 88 rows of data under each column number. The column numbers in the top row is explained to indicate which measurement that column represents in the section following the Table III.

Table III: Example of a data file for one DP Gram

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1	1	R	500	70	593	60	406	8	0	N	0	0	5	0	0	0	24	F
1	1	R	625	70	750	60	500	9	-1	T/O	0	0	5	0	0	0	24	F
1	1	R	781	70	937	60	625	14	-6	A	0	0	5	0	0	0	24	F
1	1	R	1000	70	1187	60	812	3	-2	N	0	0	5	0	0	0	24	F
1	1	R	1250	70	1500	60	1000	12	-6	A	0	0	5	0	0	0	24	F
1	1	R	1593	70	1906	60	1281	-1	-9	A	0	0	5	0	0	0	24	F
1	1	R	2000	70	2406	60	1593	13	-7	A	0	0	5	0	0	0	24	F
1	1	R	2531	70	3031	60	2031	5	-8	A	0	0	5	0	0	0	24	F
1	1	R	3187	70	3812	60	2562	7	-9	A	0	0	5	0	0	0	24	F
1	1	R	4000	70	4812	60	3187	8	-6	A	0	0	5	0	0	0	24	F
1	1	R	5031	70	6031	60	4031	5	-6	A	0	0	5	0	0	0	24	F

Explanation of column numbers for Table III:

1   Subject number.

2   Number of DP Gram.

3   Ear that is being tested (right or left).

4   Frequency of f1 in Hz.

5   Loudness level of L1 in dB SPL.

6   Frequency of f2 in Hz.

7   Loudness level of L2 in dB SPL.

8   Distortion product frequency in Hz.

9   Distortion product amplitude in dB SPL.

10   Loudness level of noise floor in dB SPL.

11   Test status (A= accepted, N= noisy, T/O= timed out response).

12   Pure tone threshold of 250 Hz in dB HL.

13   Pure tone threshold of 500 Hz in dB HL.

14   Pure tone threshold of 1000 Hz in dB HL.

15   Pure tone threshold of 2000 Hz in dB HL.

16   Pure tone threshold of 4000 Hz in dB HL.

17   Pure tone threshold of 8000 Hz in dB HL.

18   Subject age.

19   Subject gender.

The next step in the preparation of data was to select the type of neural network needed for this study and also the topology of the neural network.

4.7.2   Selection of the Type of Neural Network

A back propagation network was chosen for this study for two reasons: **1)** A possible nonlinear correlation is suspected between DPOAE thresholds and traditional pure tone thresholds. Metz, et al., 1992 reported the back propagation neural network to be very successful in dealing with nonlinearities that potentially occur in complex data sets. According to Blum, 1992, the back propagation neural network is capable of nonlinear mappings and able to generalize well. **2)** The purpose of this study is to predict pure tone thresholds with distortion product thresholds with the use of neural networks. According to Blum, (1992), the back propagation neural network is highly applicable in the areas of forecasting and prediction. Tam and Kiang, (1993) indicated a back propagation neural network to be very effective in the prediction of bank failure. Salchenberger, et al. (1993) also chose a back propagation neural network for their prediction study where thrift institution failures were predicted and obtained predictions better than any other method originally used.

To summarize, back propagation networks are applicable in the areas of *prediction* and can be used where a possible *nonlinear correlation* is sought between two sets of data.

4.7.2.1 Selection of the Topology of the Neural Network

Aspects that determine the topology of a neural network are:

4.7.2.1.1 Number of Layers in the Neural Network

"A neural network has its neurons divided into subgroups, or fields, and elements in each subgroup are placed in a row, or column, in the diagram depicting the network." (Rao & Rao, 1995:81).

For this back propagation neural network a three-layer structure was chosen: The first layer is an input layer only. The third layer is the output layer and the second layer, also referred to as the hidden layer, categorizes the input pattern and serves as a connection between the first and third layer.

4.7.2.1.2 Number of Nodes in the Input Layer

The number of input data sets that the neural network is trained with determines the number of nodes in the input layer. For example, if one threshold value at each of the 11 distortion product frequencies is used to train the neural network, the input layer will consist of 11 nodes. If two values at each of the 11 distortion product frequencies are used, such as the threshold value and the amplitude value, then the number of nodes in the input layer will be 22.

Several experiments were conducted to find the optimal number of input nodes for this study. These "trial runs" to determine the optimal topology of the neural network are described in **4.7.2.1.5: Trial Runs to Determine Neural Network Topology**.

4.7.2.1.3 Number of Neurons in the Output Layer

In the case of the output and hidden layers, the components are being referred to as neurons because of the two layers of connectivity (an input and an output) which gives it the similar structure as a neuron with a synapse on each side.

The number of aspects that is being predicted determines the number of neurons in the output layer. For example, if the neural network has to predict hearing thresholds at 500 Hz, 1000 Hz, 2000 Hz and 4000 Hz, then the number of output neurons will be four. If the neural network has to predict only one frequency, then only one output neuron in needed. Even though the aim of this study was to predict hearing ability at all four these frequencies, one network does not necessarily have to do it simultaneously, the same results can be achieved by four different networks, trained to predict only one of the frequencies. The trial runs that were conducted to determine the optimal number of output neurons for this study are also discussed in **4.7.2.1.5: Trial Runs to Determine Neural Network Topology**.

4.7.2.1.4 Number of Neurons in the Middle- or Hidden Layer

The number of neurons in the hidden or middle layer cannot be determined merely by the amount of input or output data but is a function of the diversity of the data (Blum,

1992). The number of middle layer neurons determines the accuracy of prediction during the training period. With an insufficient number of middle neurons, the network is unable to form adequate midway representations or to extract significant features of the input data (Nelson & Illingworth, 1991). With too many middle neurons the network has difficulty to make generalizations (Rao & Rao, 1995; Nelson & Illingworth, 1991). The number of middle layer neurons was determined by trial and error, based on the accuracy of the prediction during the training period. All these trial runs are discussed in the following section.

4.7.2.1.5 Trial Runs to Determine Neural Network Topology

The **first scenario** encompassed all the data from all 120 ears. DPOAE thresholds were determined for all 120 ears at all 11 DPOAE frequencies (in other words, the lowest L1 value that still yielded a DPOAE response). The criteria for a DPOAE threshold was that the lowest L1 DPOAE response had to be 3 dB above the noise floor and that the test status had to be "accepted".

All the lowest L1 values where a DPOAE response was measured were used as input data for the neural network. There were however some of the hearing impaired subjects that did not have any DPOAE responses at certain frequencies, and no DPOAE threshold values were available to use as input data. All these absent DPOAE thresholds were depicted with a "zero".

The input level of this neural network therefore had 11 nodes and each represented the L1 dB SPL value where the DPOAE threshold was measured. The neural network had

to predict hearing ability at 500 Hz, 1000 Hz, 2000 Hz and 4000 Hz in dB SPL. There were therefore 4 output neurons in this network. The number of middle level neurons were set at 20 and the acceptable prediction error during the training period at 5 dB for this test run. After a few hours it became clear that the neural network was unable to converge during the training period and that no accurate predictions could be made. For the next few trial runs, middle level neurons were increased up to 100 or the acceptable prediction error during the training period were decreased to 1 dB. All these changes did not improve convergence or prediction ability. It became clear that the absence of DPOAE thresholds in the hearing impaired population (about 66% of the subjects) called for a different data preparation method.

In the **second scenario**, an attempt was made to determine necessary neural network topology and acceptable prediction error with only those subjects that had DPOAE responses at all 11 frequencies. There were 20 ears with DPOAE thresholds at all 11 DPOAE frequencies, and naturally, almost all (19) had normal hearing (pure tone thresholds < 15 dB HL). Many different trial runs were conducted to determine the effects of the number of middle level neurons and acceptable error during training on the prediction abilities of the neural network. The general tendency revealed that more accurate predictions were made with higher numbers of middle level neurons (around 100) but that the acceptable error during training did not have a great influence on the accuracy of the prediction. An acceptable error of 5 dB in the training stage did not worsen prediction abilities compared to a training error of 1dB. It was actually found in some instances that the network had a better ability to generalize with the larger training error of 5dB. Just for general interest, one example of the second scenario trail runs will be discussed briefly.

All ears with DPOAE responses at all 11 frequencies were selected. (There were 20 ears, 19 had normal hearing (0-15 dB HL) and one had a mild hearing loss (25dB HL). For input data, only the eight highest DPOAE frequencies were used. The 3 low DPOAE frequencies were omitted because of high levels of low frequency noise. This time DPOAE amplitudes were used instead of DPOAE thresholds. The DPOAE amplitudes at $L1 = 65$, $L2 = 55$ were used as input values for the eight high frequencies. The neural network was programmed to predict only one high frequency, namely 2000 Hz. The number of middle neurons was set at 20 and the acceptable error during training at 0.5dB. The network converged fairly quickly and predictions turned out to be extremely accurate. 2000 Hz could be accurately predicted within 10 dB 100% of the time and within 5 dB 83% of the time. Although this seems like a cause for celebration, one should ask oneself what the relevance of such a prediction is. If all the ears in the training set are normal ears, and the network predicts all the ears as normal, would it necessarily know an ear with a hearing loss if it encountered one? All that could be derived from this trial run was that it was time to try a new data preparation method to incorporate all data from hearing impaired subjects as well. Accurate predictions of hearing ability across different categories of hearing impairment can only be made if a neural network is trained with sufficient data to recognize all the different categories.

Scenario three required drastic changes in the way the data is presented to the neural network. Up to now, input data consisted of decibel sound pressure level (SPL) quantities, depicting either a DPOAE threshold at a certain L1 value or DPOAE amplitude. Output data also predicted hearing thresholds in decibel sound pressure

level (dB SPL) values. For scenario three, a whole new approach was used. All data was rewritten in a binary format. The presence of a DPOAE response was depicted with a "1" whereas the absence of a DPOAE response was depicted with a "0".

The criteria for the presence of a DPOAE response was that the DPOAE response had to be 3 dB above the noise floor and that the test status had to be "accepted". All responses less than 3 dB above the noise floor or with a test status that was "noisy" or "timed out" were regarded as absent responses. (It should be noted that Kemp (1990) warned that in order to determine if a response is 3 dB above the noise floor, one could not merely subtract the noise floor from the DPOAE amplitude in its decibel form. The two values should be converted back to their pressure value (Watt/m^2), then subtracted.)

Responses from each of the eight DP Grams in each of the 120 ears were rewritten in this binary format. In the end, each ear had a row of 88 numbers ("ones" and "zeros") and every number depicted the presence or absence of a DPOAE response at one of the 11 DPOAE frequencies and one of the 8 loudness levels. These 88 numbers served at input information in the neural network (the network therefore had 88 input nodes). The only information available to the neural network in this trail run was therefore the pattern of absent and present responses at all eight loudness levels.

Another drastic change was made in the way the pure tone audiogram was depicted. As a first level approach every audiogram was graded into seven categories of average hearing ability. Each category spanned 10 dB, category one ranged from 0-10dB, category two from 11-20 dB, three from 21-30 dB and so forth. The seven

categories can be seen in Table IV. Each category of hearing ability was determined by taking the average of 500 Hz, 1000 Hz, 2000 Hz and 4000 Hz. Each ear had one number in the end, depicting its average hearing ability according to one of the seven categories. The network had only binary input information, not dB SPL values, and only had to guess a category, not a decibel value. The decibel hearing level categories of the audiogram were therefore used in its hearing level (HL) form.

To present average hearing ability according to one of the seven categories in a binary fashion, each ear had seven number places (or columns). Column one represented hearing ability in category one, column two represented hearing ability in category two and so forth. To indicate average hearing ability, the column that represented that specific hearing ability was given a "one" and the rest "zeros". For example, an ear with an average hearing ability of 29 dB HL would fall in category 3. This ear would be written: [0 0 1 0 0 0 0]. An ear with an average hearing ability of 5 dB HL would be written as [1 0 0 0 0 0 0], therefore depicting category one.

Table IV : Seven categories of hearing ability for scenario three and four

Category 1	0 -10 dB
Category 2	11-20 dB
Category 3	21-30 dB
Category 4	31-40 dB
Category 5	41-50 dB
Category 6	51-60 dB
Category 7	61-70 dB

The neural network was trained with the 88 input nodes depicting the pattern of present and absent DPOAE responses at all 11 DPOAE frequencies and all 8 loudness levels as well as the average hearing ability in one of the seven categories. The number of middle level neurons was set at 140 and the prediction error at 5%.

The results of scenario three is presented in Chapter 5.

This binary approach offered the first solution to the problem of absent DPOAE results. For the first time all the data could be used and the neural network could be trained with data across all categories of hearing impairment. Scenario three however, predicted only average hearing abilities across the whole audiogram. The main aim of this study is to predict hearing ability at the frequencies 500 Hz, 1000 Hz, 2000 Hz and 4000 Hz. It was decided to take the binary approach one step further, by predicting hearing ability at a specific frequency, one at a time.

4.7.2.1.6 Prediction of Hearing Ability at 500 Hz, 1000 Hz, 2000 Hz and 4000 Hz

Prediction of hearing abilities at specific frequencies was conducted in scenario four.

Scenario four used the same DPOAE input information as scenario three, which was the 88 columns of binary information, depicting present and absent DPOAE responses at all the DP Grams and DPOAE frequencies. Scenario four also used the seven categories of hearing ability to write output information in a binary format. Instead of using the average hearing ability of a subject as output information, only the pure tone frequency to be predicted was used. Four different neural networks were used to

predict 500 Hz, 1000 Hz, 2000 Hz and 4000 Hz, one at a time. For each neural network, the number of middle neurons was set at 140 and the acceptable error during training at 5%. The neural network took about 4 days non-stop to predict one frequency of all 120 subjects.

After completion of neural network prediction it became clear that in some instances, certain categories had very little hearing-impaired data. In the case of 500 Hz for example, many of the subjects with hearing losses had normal hearing at 500 Hz (such as subjects demonstrating ski slopes). Category 7 in the case of the 500 Hz prediction had only data for one ear. Category 6 had only data for six ears and category 5 only data for five ears. It could be possible that the neural network did not have sufficient data in every category to train on and this aspect might influence the accuracy of the prediction. It was decided to enlarge the categories depicting hearing impairment to 15 dB, in order to attempt to include more hearing-impaired data in every category. In **scenario five**, hearing ability was divided in five categories. Categories that depicted normal hearing spanned 10 dB whereas categories that depicted hearing impairment spanned 15 dB. The five categories are presented in Table V.

Table V. The five categories of hearing ability for scenario five.

Category 1	0 – 10 dB HL
Category 2	11 – 20 dB HL
Category 3	21 – 35 dB HL
Category 4	36 – 50 dB HL
Category 5	51 – 65 dB HL

The network was trained with the binary written DPOAE responses and hearing abilities in the five categories. The number of middle level neurons was set at 140 and the acceptable training error at 5%. The network was trained with the data of 119 ears and predicted one ear. This process was repeated 120 times to predict every ear once.

The prediction of hearing ability at 500 Hz, 1000 Hz, 2000 Hz and 4000 Hz as well as the prediction of average hearing ability were performed in both seven categories (as in scenario 4) and in five categories (as in scenario 5). The differences in results between these two scenarios will be discussed in Chapter 5 and Chapter 6.

4.7.2.1.7   The Inclusion of the Age and Gender Variables

To determine the effects of age and gender on the distortion product, it was decided to include these variables into the neural network as input information. The variables age and gender were included in the network run where the network had to predict average hearing ability.

The variables age and gender also had to be presented to the neural network in a binary format. For the network run that included the gender variable, it was very easy to depict the new variable in a binary mode. The one gender was given a zero and the other gender a one. The one extra input did not influence the complexity of the neural network topology to such an extent that it was necessary to include more middle neurons in the hidden layer. This neural network therefore had 89 input nodes, 140 middle layer neurons and seven output neurons, one for every 10dB category. The prediction error during training was set at 5%. The neural network was exactly the

same as for the prediction of average hearing ability, except for the extra input variable, gender. The neural network had to predict average hearing as being in one of the seven 10dB categories of scenario four.

The age variable was also incorporated in a neural network run to predict average hearing ability in the seven 10dB categories of scenario four to determine its effect on the distortion product. To represent the age variable to the neural network in a binary format required much more input neurons. Subject age ranges from 8 to 82 years old. To present this to the neural network in a binary mode, nine categories of different ages were created, every category spanning 10 years. It was written in a binary format in the same way that hearing ability categories were. For example, a subject with an age of 12 would fall in the second 10-year category and would be written binary as [0 1 0 0 0 0 0 0]. A subject with an age of 82 would fall in the ninth 10-year category and would be binary written as [0 0 0 0 0 0 0 0 1]. The network that was presented with subject age had therefore nine more input neurons, amounting to a total of 97 input neurons. (The network had 88 regular input nodes to represent all absent and present DPOAE responses at the 8 DP Grams of all 11 DPOAE frequencies plus 9 input nodes to represent the age category). The middle level neurons were kept at 140 and the network had seven output neurons, one for every 10dB category.

To determine the combined effects of gender and age, one neural network was run to include both variables at the same time. The network therefore had 98 input nodes, 140 middle level neurons and seven output neurons for the seven 10dB categories of scenario four. Prediction accuracy during training was set at 5%.

4.8 Data analysis procedures

After the completion of a neural network run, the results were given in a table format, with 120 rows (each ear had one row) and 15 columns of numbers (as in the case of scenario four). The first column number depicted the ear number, the other 14 the actual hearing category and predicted hearing category, written in a binary format. To illustrate this concept, an example of a neural network's output for the data of 10 ears is presented in Table VI. The predicted frequency was 1000 Hz.

Ear 1 had an actual hearing threshold of 5 dB at 1000 Hz, therefore a category one. The category was depicted binary by the "1" in the "actual" (A) column of Category 1. All the other "actual" (A) columns of the other categories for ear one is therefore "0". The neural network investigated the pattern of the input information and made more than one prediction for possible categories of hearing ability for this ear.

The category where the most energy is concentrated, is taken as the prediction of the neural network, and in the case of ear 1, it is in category 1. This ear's hearing ability was therefore correctly predicted as a category 1.

110

Table VI: Example of the results of the neural network's prediction of 1000 Hz for 10 ears, (scenario four). A= Actual hearing category, P= Predicted hearing category.

EAR #	CATEGORY 1		CATEGORY 2		CATEGORY 3		CATEGORY 4		CATEGORY 5		CATEGORY 6		CATEGORY 7	
	A	P	A	P	A	P	A	P	A	P	A	P	A	P
1	1	1	0	0.4	0	0	0	0	0	0.03	0	0.1	0	0
2	0	0.3	1	0.9	0	0.3	0	0.01	0	0.01	0	0.1	0	0.01
3	0	0.2	0	0	0	0.01	0	0.3	0	0.01	1	0.4	0	0.9
4	0	0	0	0	1	1	0	0	0	0	0	0	0	0
5	1	1	0	0	0	0.3	0	0.02	0	0.03	0	0.4	0	0.02
6	0	0.05	1	1	0	0.44	0	0.14	0	0.05	0	0.01	0	0
7	0	0.01	0	0.03	0	1	0	0.8	1	0.6	0	0.4	0	0.02
8	0	0.01	1	1	0	0.3	0	0.21	0	0.12	0	0.02	0	0.01
9	1	1	0	0.2	0	0.15	0	0.11	0	0.02	0	0.01	0	0
10	0	0	0	0	0	0	0	0	0	0.02	0	0.3	1	0.7

In the case of ear 3, the actual hearing threshold of the specific frequency fell in Category 6. After analysis of this ear's pattern of present and absent DPOAE responses the neural network predicted hearing ability to be most probably in Category 7. It should be noted that some energy was also predicted in other categories. The category where the most energy is predicted is taken as the predicted category. Another ear in this example that was wrongly predicted is ear 7, predicted as a Category 3, instead of a Category 5. All the other ears were predicted correctly.

The data analysis therefore consisted of analyzing the actual and predicted values of all 120 ears and to determine how many were predicted accurately, how many within one class and how many were predicted incorrectly.

Another aspect that was determined for every frequency was the percentage accurate prediction of normal hearing for every frequency. This was determined in terms of false positive responses (how many subjects with normal hearing were predicted as hearing impaired) and false negative responses (how many subjects with hearing impairment were predicted as having normal hearing ability) at every frequency.

4.9   Summary

The need for an objective non-invasive and accurate test of auditory functioning inspired this research project. The aim of this research project was to predict hearing ability at 500 Hz, 1000 Hz, 2000 Hz and 4000 Hz, with DPOAEs and artificial neural networks. Data obtained from DPOAE results and pure tone thresholds of 120 ears were used to train the neural network. Subject selection criteria included varying degrees of sensorineural hearing loss and normal middle ear functioning. Subjects ranged from 8 to 82 years old and included 28 males and 42 females.

The distortion product otoacoustic emission has numerous variables that influence the effectivity in which measurements can be made. For this research project, eight DP Grams at 5dB intervals ranging from L1=70dB SPL to L1=35dB SPL were measured. A frequency ratio of 1.2 was selected for the two primaries and the loudness level ratio of the two primaries was L1>L2 by 10dB. The frequency range of F1= 500 to F1= 5031 was tested.

The neural network that was chosen for the prediction of 500 Hz, 1000 Hz, 2000 Hz and 4000 Hz was a back propagation neural network. The network had 140 middle

neurons, 88 input nodes and seven output neurons in scenario four, five output neurons in scenario five. The network's acceptable prediction error during training was set at 5%. All data that was used for neural network training was rewritten in a binary format.

Hearing ability was predicted in two scenarios. In scenario four, hearing ability was predicted into one of seven 10dB categories (Table V). In scenario five, the network had to predict hearing ability into one of five categories, the first two spanned 10dB and the rest 15dB. The neural network was not trained with the precise decibel values of a hearing threshold but with the categorical value. Four different networks were trained for the four prediction frequencies 500 Hz, 1000 Hz, 2000 Hz and 4000 Hz. Data analysis consisted of analyzing the actual and predicted values of all 120 ears and to determine how many were predicted accurately, how many within one class and how many were predicted incorrectly.

There are numerous variables that influenced the outcome of this research project. It is quite possible that different DPOAE settings such as other frequency ratios or different loudness levels could yield different results (Cacace et al., 1996). It is also possible that a different type of neural network or a network with a different topology could affect the results significantly (Nelson & Illingworth, 1991). It was attempted to specify all the stimulus variables that could have an effect on the outcome of this research project in great detail in the preceding Chapters.