NEURAL NETWORKS APPROACH IN DIAGNOSING CLASSES OF ANAEMIA

SHUZLINA BINTI ABDUL RAHMAN

UNIVERSITI UTARA MALAYSIA

2000

Sekolah Siswazah
(Graduate School)
Universiti Utara Malaysia

PERAKUAN KERJA KERTAS PROJEK
(Certification of Project Paper)

Saya, yang bertandatangan, memperakukan bahawa
(I, the undersigned, certify that)

SHUZLINA BINTI ABDUL RAHMAN

calon untuk Ijazah
(candidate for the degree of) Master Science Of (Information Technology)

telah mengemukakan kertas projek yang bertajuk
(has presented his/her project paper of the following title)

NEURAL NETWORKS APPROACH IN DIAGNOSING CLASSES OF ANAEMIA

seperti yang tercatat di muka surat tajuk dan kulit kertas projek
(as it appears on the title page and front cover of project paper)

bahawa kertas projek tersebut boleh diterima dari segi bentuk serta kandungan, dan meliputi bidang ilmu dengan memuaskan.
(that the project paper acceptable in form and content, and that a satisfactory knowledge of the field is covered by the project paper).

Nama Penyelia
(Name of Supervisor) : Dr. Fadzilah Siraj

Tandatangan
(Signature) : Fadzilah

Tarikh
(Date) : 5/10/2000

NEURAL NETWORKS APPROACH IN DIAGNOSING CLASSES OF ANAEMIA

A thesis submitted to the Graduate School in partial fulfilment of the requirements for the degree Master of Science (Information Technology), Universiti Utara Malaysia

by

Shuzlina binti Abdul Rahman

GRADUATE SCHOOL
UNIVERSITI UTARA MALAYSIA

PERMISSION TO USE

In presenting this thesis in partial fulfilment of the requirements for a post graduate degree from Universiti Utara Malaysia, I agree that the University Library may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or part, for scholarly purposes may be granted by my supervisor(s) or, in their absence, by the Dean of the Graduate School. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without any written permission. It is also understood that due recognition shall be given to me and to Universiti Utara Malaysia for any scholarly use which may be made of any material from my thesis.

Request for permission to copy or to make other use of materials in this thesis, in whole or part, should be addressed to:

Dean of Graduate School

Universiti Utara Malaysia

06010 Sintok

Kedah Darul Aman

ABSTRAK

Perkhidmatan unit hematologi amat diperlukan dalam mengenalpasti penyakit pesakit melalui kajian dari sampel darah. Pelbagai jenis penyakit dapat dikenalpasti oleh seorang pakar hematologi melalui beberapa analisa terhadap kandungan darah pesakit. Namun ratusan kes yang diterima dan pelbagai faktor yang perlu diambilkira telah melambatkan proses sesuatu keputusan dibuat. Hal ini boleh diatasi dengan menggunakan rangkaian neural sekiranya dilatih dengan sejumlah data yang mencukupi, merangkumi semua faktor yang diperlukan untuk mengkelaskan sesuatu penyakit melalui pengecaman corak. Kajian tesis ini telah menggunakan model "multilayer perceptron" dengan pembelaj aran rambatan-balik untuk pengkelasan anemia. Di samping itu, beberapa pembolehubah yang mempengaruhi prestasi model juga telah dikenalpasti. Model yang dihasilkan dinilai prestasinya dan telah berjaya mengkelaskan anemia dengan 72.78% bagi data latihan dan 71.56% bagi data ujian. Model yang dihasilkan seterusnya dibandingkan dengan model "Radial Basis Function" dan "Regression" dan telah menunjukkan prestasi yang terbaik.

ABSTRACT

Hundreds of haematology forms are directed to Haematology unit every day from various departments from physicians that need the right diagnosis in patient's blood. The processing may take several days depending on the workload and available resources. A combination of various factors has to be considered before a haematologist can diagnose classes of anaemia and is normally performed in several stages. The process can actually be performed using neural network approach, as it is capable in pattern recognition. Knowing the relevant factors that influence anaemia classification, a model of neural network can be produced if it is trained with sufficient data sets. Hence, this thesis presents the neural network model for anaemia classification and identifies parameter that affects its performance using backpropagation. The model is then implemented and the performance of the neural network is assessed. The model was able to diagnose classes of anaemia with 7 1.5 6% generalization. Finally, the model was compared with Radial Basis Function and Regression model to show that Multilayer Perceptron outperforms the other two models.

ACKNOWLEDGEMENTS

In the name of Allah, Most Gracious, Most Merciful.

I would especially like to take this opportunity to thank Dr Fadzilah Siraj, School of Information Technology, Universiti Utara Malaysia (UUM) who has supervised me in this study and helped in shaping the final form of this paper.

I would like to acknowledge the assistance received from Dr Abdul Rashid Mohd Ibrahim, a hematologist at Bank Darah of Hospital Alor Setar, Kedah who has allowed me to assess the information from the department. His time and assistance is really appreciated. I would also like to thank the staff at Haematology Lab who have directly involved and gave their full cooperation.

The love and understanding of my husband, Dr Muhd. Helmi, and my beloved mother, Hajjah Bashah that made this works possible and worthwhile. Last but not least, all my friends who have given their assistance throughout the study. There is a tremendous sense of achievement in completing this study.

SHUZLINA BINTI ABDUL RAHMAN
SEKOLAH SISWAZAH
UNIVERSITI UTARA MALAYSIA
OCTOBER 2000

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDIXES

CHAPTER 1
INTRODUCTION

In this chapter, the first section describes the context of the study that gives an introduction to neural networks and its application. The second section presents statement of purpose, while the third section presents the objectives of the study followed by study significance. Finally, the scope of the study that includes the limitations of the study is presented.

1.1 The Context of The Study

The development of computers has been very fast and computers have become important tools in this Information Communication Technology's (ICT) world. Nevertheless, it still lack the flexibility of processing in some areas as what the human brain does especially in the area of pattern recognition, prediction or forecasting in business, modelling and diagnosing in medical, and others.

Artificial Neural Network or neural network is relatively recent development in the information science that has the ability to model human like computing strategies to improve the performance of computers. They differ from the usual computer programs in that they "learn" from a set of examples rather than being programmed to get the right answer. Neural networks have been applied in many areas, ranging from business, engineering, medical and others. This study will focus the application of neural networks in haematology, an area in medicine.

1.2 Statement of Purpose

Hundreds of haematology forms are directed to Haematology unit every day from various departments in Hospital Alor Setar, mostly from physicians that need the right diagnosis in patient's blood. The processing at Haematology unit may take several days depending on the workload and available resources. Several procedures involved before the right diagnosis can be made. And the diagnosis must be made with the existence of Hematologist who is the key person in giving the final diagnosis.

Hematologist needs information of Full Blood Picture (FBP) and Full Blood Count (FBC) as well as the clinical features written by medical officers. After considering those factors, only then the haematology will come out with a diagnosis of the patient. All these information is recorded in a haematology form. The complex underlying relationships between the information can be applied in neural network model that has the ability like a human brain if it is trained with sufficient data. Therefore, it is likely that it can also be trained of what hematologist does in diagnosing classes of anaemia.

1.3 The Objectives of The Study

The goal of the study is to produce a neural network model or precisely a multilayer perceptron model with the optimum values in its parameters, thus making a diagnosis of anaemia. It is important to have the optimal value of number of hidden units, learning rate and momentum in the development model. Without the right value of

those parameters, the model is unable to make good generalization for the domain area.

The study aims to compare the performance of multilayer perceptron with radial basis function and regression model. The specific objective is to produce a model that can help hematologist in diagnosing classes of anaemia using a set of data sets that will be determined in the study.

1.4 Study Significance

This study investigates the importance of having the optimum values in number of hidden units, learning rate and momentum that constructs the suitable model for predicting the classification of anaemia. The study shows that performance of multilayer is much better than the alternative models such as radial basis function and the regression model in this problem domain.

The suitable model that was attained in this study can be used by haematologist to diagnose classes of anaemia by providing the input variables. The findings in this study indeed a milestone for further enhancement in the field of medical, particularly haematology field for Hospital Alor Setar, Kedah.

1.5 Scope of Study

The study data was obtained from Haematology Unit of Hospital Alor Setar, Kedah. For the purpose of this study, patients that suffered from anaemia were chosen to be the inputs for obtaining the neural network model. In order to determine the relevant inputs for the model, a questionnaire is developed in addition to the haematology form that act as the main sources of data collection. However, clinical features were not considered although it is known as one of the important factors needed in any diagnosis. In most cases, the clinical features written by the medical officers were not completed and sometimes were not being filled.

The data collected comprises of anaemia patients in 1999 that total of 732 cases. This study is solely concerned with anaemia although there are many diagnoses that can be extracted from the haematology form such as leukemia, AIDS (Acquired Immune Deficiency Syndrome), lead poisoning and others. From seventeen classes of anaemia identified in the early stages, only classes of more than 5% of total distribution was considered resulting eight classes for the final output.

CHAPTER 2
LITERATURE REVIEW

This chapter presents a literature review on neural networks applications. The first section discusses applications of neural network in general, while the second section focuses on neural network applications in medical. In the final section, the application of neural networks in haematology is presented.

2.1 Application in General

The pioneering work of neural network in the modern era has started since 1943 by McCulloch and Pitts (McCulloch *et al.*, 1943). To date, there has been an explosive growth research in this field and has attracted many investigators, including academician, physicians, psychologist, and neurobiologist beginning early 1980s. An approach to the pattern recognition problem was introduced by Rosenblatt (Rosenblatt, 1958) in his work on the perceptron and they are now many successful projects and on-going projects that utilized the ability of neural networks in their applications.

The applications of neural networks are almost limitless but they fall into several main categories like classification, modelling, forecasting and novelty detection. Some examples of successful applications include; credit card fault detection (Alaskerov *et al.*, 1997), pattern recognition (Ramli *et al.*, 1996, Rietveld *et al.*, 1999), handwritten character recognition (Le Chun *et al,* 1990; Tay, 1991; Karim *et*

al., 1998; Kamaruzzaman *et al.*, 1998), colour recognition (Yaakob *et al.*, 1999), and share price prediction system (Sanugi *et al.*, 1996; Lim *et al.*, 1996) and others.

Many researchers have compared Artificial Neural Netwoks (ANNs) and Logistic Regression (LR) models. They have shown that neural networks are able to make a better generalization over the traditional statistical methods such as regression techniques (Lapuerta *et al.,* 1995; Erler *et al.;* 1995; Shanker, 1996; Lapuerta *et al.*; 1997, Armoni, 1998).

2.2 Application in Medical

One of the major goals of observational studies in medicine is to identify patterns in complex data sets. Literatures have shown that medical has benefited much from this technology. It has been successfully applied to various areas of medicine to solve non-linear problems. The applications include prediction of diagnosis such as myocardial infarction (Baxt, 1991), several type of cancers (Astion *et al.*, 1992; Wilding *et al*, 1994; Wu *et al.*, 1993), the onset of diabetes melitus (Shanker, 1996), survival prediction in AIDs (Ohno-Machado, 1996b), eating disorders (Buscema *et al.*, 1998) and others. Applications in signal processing and interpretation involve EEGs or electroencephalogram analysis (Makeigh *et al.*, 1996), ECGs or electrocardiograms (Bortolan *et al.*, 1993), EMGs or electromyelogram (Chiou *et al.*, 1994), and EGGs or electrogastrograms classifications (Lin *et al.,* 1997). The use of neural networks in biochemical analysis includes identification of haeterotrophic marine bacteria (Giacomini *et al.*, 1997), and analysis of biochemical markers for early assessment of acute myocardial infarction (Ellenius *et al.*, 1997).

Some applications in radiology includes classification and detection of interstitial lung disease (Katsuragawa *et al.*, 1997; Ishida *et al.*, 1997), detection of lung nodules (Wu *et al.*, 1995), several types of classification in mammography (Huo *et al.*, 1996; Doi *et al.*, 1997).

In 1994, a group of researchers employed a hierarchical artificial neural network system to classify cervical cells (Bazoon *et al.*, 1994). The system consists of a hierarchical structure of BP ANNs that "filters" out the abnormal from the normal and then makes a finer discrimination as to degree of normalcy or abnormalcy. The total system accuracy rate is 93.7% for the classification of cells into four categories.

In another study conducted by Bortolan (Bortolan *et al.*, 1991), a combination of two techniques of pattern recognition i.e., cluster analysis and neural networks are used to investigate the specific problem of the diagnostic classification of 12-lead electrocardiograms (ECGs). For this study, a previously used database established at the University of Leuven, Belgium was applied. Sensitivity, specificity, and total and partial accuracy were the indices used for the assessment of the performance. Several neural networks have been obtained by either varying the training set (considering clusters of the original learning set) or adjusting some components of the architecture of the networks. The combination of different neural networks has shown satisfactory performances in the diagnostic classification task.

Performance of the neural network strategy has shown higher performance than Cox regression models in predicting clinical outcomes of the risk of coronary artery disease (Lapuerta *et al.*, 1995). In addition to this study, Lapuerta *et al.* compared the

prediction of survival of neural networks and logistic regression models on alcoholic patients with severe liver disease. The study reveals that neural networks were more successful in classifying patients into low and high-risk group.

A similar study carried out by Armoni shows that neural network prediction was more accurate than linear regressions (Armoni, 1998) for prediction the diagnostic probabilities of insulin-dependent diabetes mellitus. The results suggest the use of a neural network should be considered whenever prediction of diagnosis is required.

The modelling capabilities of neural networks are compared to traditional methods like logistic regression to predict the onset of diabetes mellitus in Pima Indian Women (Shanker, 1996). The results indicate that neural networks are indeed a viable approach to classification. The chance probabilities of classification in training and test samples are around 55% while the classification rates achieved by neural networks are around 78% in training and 81% in the test sample.

In addition to the above studies, the performance of six different neural network models and a simple auto-regressive model were tested empirically to forecast the fetal heartbeats. The outcome can be regarded as an example demonstrating that good results can be obtained when various methods are combined in a single neural network. The best performance was obtained with a network that used layer delay, layer feedback, and unit feedback (Ulbricht *et al.*, 1996).

In the area of medical image processing, Doffner *et al.,* demonstrated that neural network can be effectively be used as a tool in medical decision-making (Doffner *et*

al., 1996). They applied neural network in the interpretation of planar thallium-201 scintigrams for the assessment of coronary artery disease.

2.3 Application in Haematology

An attempt was made to create an expert system with sufficient accuracy to diagnose classes of anaemia and report presumptive diagnoses directly on the haematology form (Birndort *et al.*, 1996). The purpose is to simulate the processes of human experts that can reliably achieve diagnostic separability by pattern analysis. In doing this, they constructed a hybrid expert system combining rule-based and artificial neural network (ANN) models to evaluate microcytic anaemia in a 3-layered program using haematocrit (HCT), mean corpuscular volume (MCV), and coefficient of variation of cell distribution width (RDWcv) as inputs. These measurements are available as standard output on most haematology analyzers. Three categories of microcytic anaemia were considered, iron deficiency (IDA), haemoglobinopathy (HEM), and anaemia of chronic disease (ACD). The performance of the model was evaluated with actual case data. The results show that the model was successful in correctly classifying 96.5% of 473 documented cases of microcytic anaemia and anaemia of chronic disease. This result exhibits sufficient accuracy to be considered for use in reporting microcytic anaemia diagnoses on haematology forms.

The leukocyte-vessel wall interactions are studied in post capillary vessels by intravital video microscopy during in vivo animal experiments (Egmont-Petersen *et al.*, 2000). Sequences of video images are obtained and digitized with a frame grabber. A method for automatic detection and characterization of leukocytes in the

video images is developed. Individual leukocytes are detected using a neural network that is trained with synthetic leukocyte images generated using a novel stochastic model. This model makes it feasible to generate images of leukocytes with different shapes and sizes under various lighting conditions. Experiments indicate that neural networks trained with the synthetic leukocyte images perform better than networks trained with images of manually detected leukocytes. The best performing neural network trained with synthetic leukocyte images resulted in an 18% larger area under the ROC curve than the best performing neural network trained with manually detected leukocytes.

CHAPTER 3
NEURAL NETWORKS

This chapter gives a brief overview of networks. The first section presents an introduction to neural networks. The second section discusses a historical account of neural networks, while section three discusses a perceptron model. In the following sections, the architecture, the types of learning and how it works are discussed. The multilayer perceptron model and back propagation concepts are mentioned in section six and seven. Finally, a discussion of neural networks as tools in medical is presented.

3.1 Introduction to Neural Networks

The field of Artificial neural networks (ANN) or commonly known as neural network has received a great deals of interest due to its capability in findings the relationships from complex data. It is also referred to as neurocomputers, connectionist systems, parallel distributed processor or self-adaptive models (Haykin, 1999; Ohno-Machado, 1996a).

Tsoukalas and Uhrig define neural networks as, *"a data processing system consisting of a large number of simple, highly interconnected processing elements (artificial neurons) in an architecture inspired by the way structure of the cerebral cortex of the brain"* (Tsoukalas and Uhrig, 1997).

Due to complexity that governs its training, neural network is still viewed as a "black box" (Turner, 1994) that can learn, generalize, and cluster data. They have the ability to derive meaning from complicated or imprecise data that can be used to extract patterns, to organize the information and detect trends into a useful form (Turner, 1994; Tsoukalas and Uhrig, 1997; Haykin, 1999).

Neural network models have three primary components:

- ***The input data layer***

  This layer represents the raw information that is fed into the model.

- ***The hidden layer(s)*** which is commonly referred to as the black box.

  This layer contains two main processes: the weighted summation function and the transformation functions or known as activation function. Both of these functions relate the values from the input data to the output.

- ***The output layer***, the estimated property value(s)

  This layer's behavior depends on the hidden units and the weights between the hidden and output unit.

The models have several features that set them apart from other models. These features are: (Krose & van der Smagt, 1991)

i. *a set of processing units, sometimes called neurons, nodes, or cells;*

ii. *a state of activation for every unit;*

iii. *connections between the units. Generally each connection has a weight value associated by it;*

iv. *a propagation rule, which determines the effective input of the unit from its external inputs;*

v. *an activation function that determines the activation of the unit from its effective input;*

vi. *an external input or offset for each unit;*

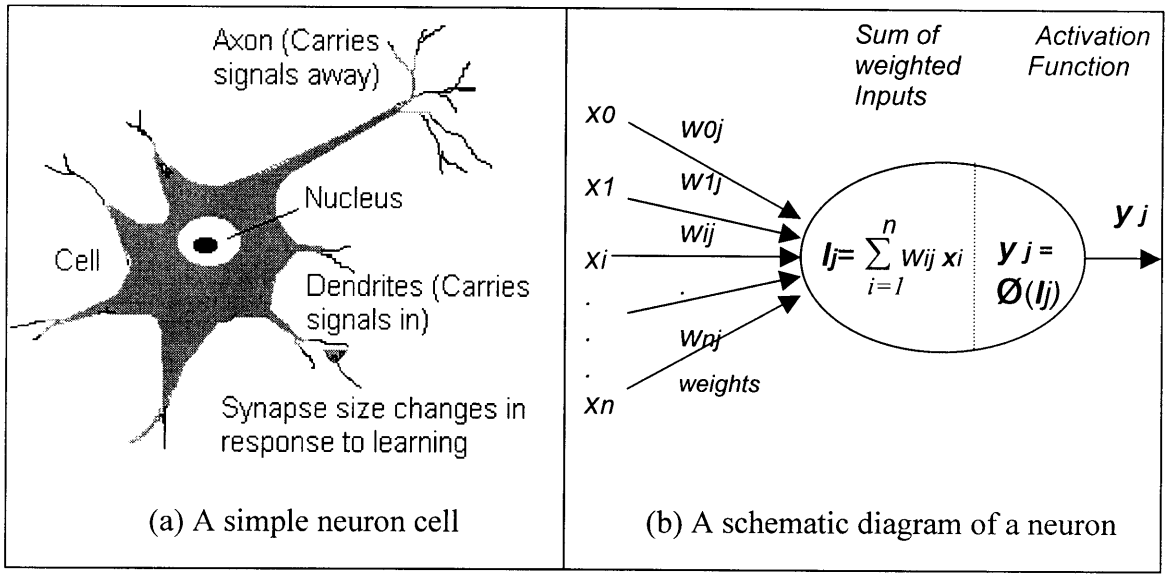
vii. *an environment for the system to operate in, providing input signals and, if necessary, error signals.*

3.2 History of Neural Network

Neural network research has its origins in the work developed by McCulloch and Pitts in 1943, who developed mathematical models based on observational studies of real neurons (Ohno-Machado, 1996a). They introduced the idea of a step threshold, but it did not have the ability to learn (Haykin, 1999). In 1962, Rosenblatt developed the *perceptron* model, which generates much interest because of its ability to solve some simple pattern classification problems. However, the interest started to fade in 1969 when Minsky and Papert provided mathematical proofs of the limitations of the perceptron and pointed out its weaknesses in computation. The drawbacks led to the temporary decline of the field of neural networks. However, the popularity is back in mid-1980s when Rumelhart (Rumelhart, 1986) developed the backpropagation algorithm.

Figure 3.1 compares a neuron cell and a schematic diagram of neuron. In figure 3.1(a), one can see a simple processing element that receives and combines signals from other neurons through input paths called dendrites. Each signal coming into a neuron along a dendrite passes through a synapse or synaptic junction. As the synaptic strengths of the neurons adjusted, the brain "learns" and stores information.

The processing units of neural networks are very similar to their biological counterpart, the neuron (see Figure 3.1(b)). The input signals are represented by x_1, x_2, ..., x_n. Each of these inputs is modified by weights (sometimes called synaptic weights) represented by w_1, w_2, ..., w_n. The processing element consists of two parts. The first part simply sums the weighted inputs resulting in quantity **I** and the second part is a nonlinear filter which usually called the activation function, through which the combined signal flows.



(a) A simple neuron cell

(b) A schematic diagram of a neuron

Figure 3.1: Comparison between neuron cell and artificial neuron

3.3 Perceptron

A perceptron is the simplest form of a neural network model, a term coined by Rosenblatt in 1958 that used for the classification of patterns said to be linearly separable. The perceptron is a neural network that has only two layers; the input layer and the output layer. In the perceptron, the output layer consists of exactly one node but the input layer can have many nodes. It is a feedforward network, which means that the inputted data is passed from the first layer directly to the second layer, or output layer. Figure 3.2 shows signal-flow graph of the perceptron.

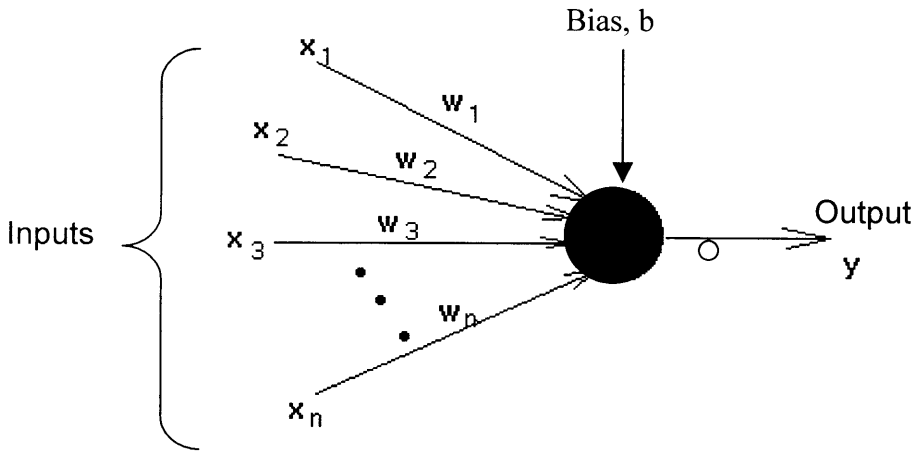


Figure 3.2: Signal flow graph of the perceptron

Input nodes are presented by $x_1, x_2, ..., x_n$ while weights are denoted by $w_1, w_2, ..., w_n$ and bias is denoted by ***b.*** Since the perceptron has only one output node that is fed from all input nodes, the output from the network is given by equation e.q. 3.1.

$$0 = \begin{cases} 1 \text{ if a is nonnegative} \\ 0 \text{ if a is negative} \end{cases}$$

where

$$y = \sum_{i=1}^{m} W_i x_i + b \qquad \text{Eq. 3.1}$$

3.4 Network Architecture

There are two types of network architectures: Feedforward Networks and Recurrent Networks. Each of this architecture represents a different way of connecting the nodes within each layer and of interconnecting the layers with each other.

a) FeedForward Networks

In this type of network, the nodes are connected to each other in such a way that the data flows from the input layer to the output layer without passing through any layer more than once. They are extensively used in pattern recognition. This type of organisation is also referred to as bottom-up or top-down.

b) Recurrent network

In this type of network, there are connections from nodes in one layer to nodes in a previous layer, or connections between nodes in the same layer, or possibly both.

The presence of feedback loops has a profound impact on the learning capability of the network and its performance (Haykin, 1999).

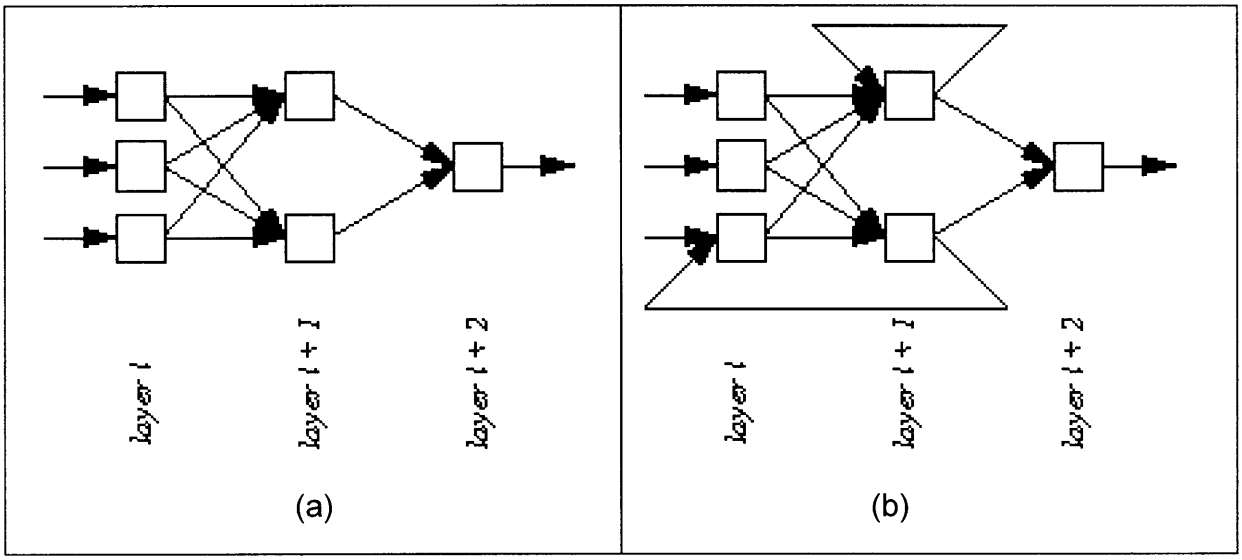


Figure 3.3: Network Architecture: a) Feedforward Network and b) Recurrent Network

3.5 Types of Learning

In order for a neural network to be used in any form, it must be able to "learn". Haykin (Haykin, 1999) defines learning as a process by which the free parameters of a neural network are adapted through a process of stimulation by the environment in which the network is embedded. Learning in neural networks means finding a set of weights that minimizes the overall error (Ohno-Machado, 1996a). The type of learning is determined by the manner in which the parameter changes take place.

There are two types of learning: supervised learning and unsupervised learning. Supervised learning incorporates an external teacher, so that each output unit is told what its desired response to input signals ought to be. During the learning process global information may be required. Paradigms of supervised learning include error-

correction learning, reinforcement learning and stochastic learning.

Unlike supervised, unsupervised uses no external teacher and is based upon only local information. It is also referred to as self-organisation, in the sense that it self-organises data presented to the network and detects their emergent collective properties. Paradigms of unsupervised learning are Hebbian learning and competitive learning.

Lawrence (Lawrence, 1991) points out that there is a big difference in how data get organized between supervised and unsupervised networks. Supervised neural networks are generally used for prediction, evaluation, or generalization. Unsupervised networks, such as Kohonen networks are best applied to clustering or recognition types of problems. Generally, the more example sets that are presented to a network for training and testing, the better the training will be able to make valid correlations and generalizations (Tsoukalas and Uhrig, 1997). However, there must be good distribution of possible inputs and outputs.

3.6 How Neural Networks Work

The behaviour of an ANN (Artificial Neural Network) depends on both the weights and the activation function (also known transfer, squashing, input-output, gain function) that is specified for the units. This function typically falls into one of three categories: linear threshold (usually used in perceptron), logistic or sigmoid and hyperbolic tangent (bipolar sigmoid) as shown in Figure 3.4. The sigmoid function is by far the most common form of activation function used in the construction of neural networks.

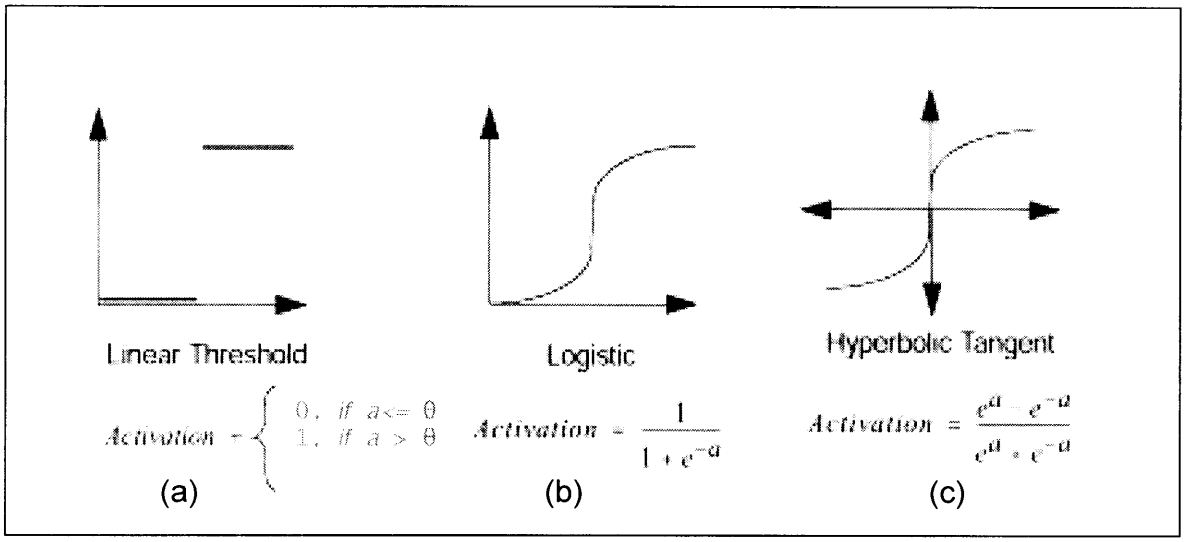


Figure 3.4: Categories of Activation Functions

3.7 Multilayer Perceptron (MLP)

MLP is a neural network model that consists of a set of sensory units (sensor nodes) that constitute the input layer, one or more hidden layers of computation nodes, and an output layer of computation nodes (Haykin, 1999). Since its development in the mid-1980's, MLP's have been applied to a diverse range of practical problems such as signal processing, forecasting, pattern classification, speaker identification and speech recognition, handwriting and character recognition, control and telecommunications problems (Gallagher, 1999).

MLP is trained in a supervised manner with a highly popular algorithm known as error back-propagation learning rule. According to Tsoukalas, about 80% of all applications utilize this backpropagation algorithm in one form or another (Tsoukalas and Uhrig, 1997). Haykin describes MLP of having two kinds of signals: Function signals and Error signals (Haykin, 1999). (See Figure 3.5)

Function signal is an input signal that comes in at the input end of the network, propagates forward through the network, and emerges at the output end of the network as an output signal. It is referred to as a "function signal" for two reasons. First, it is presumed to perform a useful function at the output of the network. Second, at each neuron of the network through which a function signal passes, the signal is calculated as a function of the inputs and associated weights applied to that neuron.

Error signals originates at an output neuron of the network, and propagates backward (layer by layer) through the network. It is referred to as an "error signal" because its computation by every neuron of the network involves an error-dependent function in one form or another.

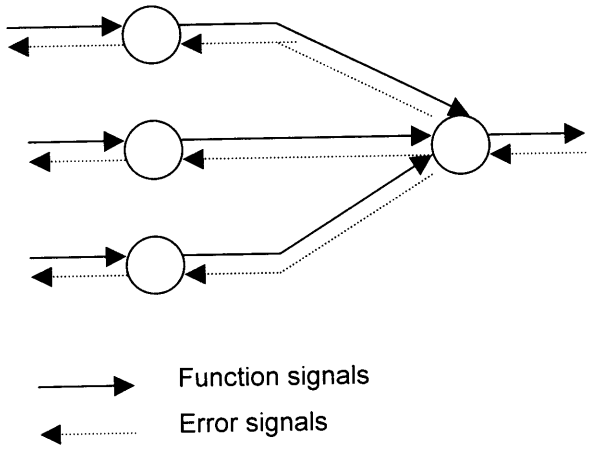


Figure 3.5: Two basic signal flows in multilayer perceptron

Each hidden or output neuron of a multi layer perceptron is designed to perform two computations (Haykin, 1999):

1. The computation of the function signal appearing at the output of a neuron, which is expressed as a continuous nonlinear function of the input signal and synaptic weights associated with that neuron.
2. The computation of an estimate of the gradient vector (i.e. the gradients of the error surface with respect to the weights connected to the inputs of a neuron), which is needed for the backward pass through the network.

3.8 Backpropagation

Backpropagation algorithm is the first widely used method for training an MLP to perform a supervised learning (Rumelhart, 1986). It is the most frequently used algorithm for neural network learning in medical applications (Ohno-Machado, 1996a). The algorithm provides a means of adjusting the weights in an MLP, given a set of training data. It allows the gradient of the error function to be calculated (composed of partial gradient information with respect to different groups of weights in the network), and implements gradient descent on the error function in weight space.

The basic backpropagation algorithm consists of three steps (see Figure 3.6). The input pattern is presented to the input layer of the network. These inputs are propagated through the network until they reach the output units. This forward pass produces the actual or predicted output pattern. Since it is a supervised learning algorithm, the desired outputs are given as part of the training vector. The actual network outputs are subtracted from the desired outputs and an error signal is produced. This error signal is then the basis for the backpropagation step, whereby the errors are passed back through the neural network. The connection weights are then adjusted and the neural network has just "learned" from an experience.

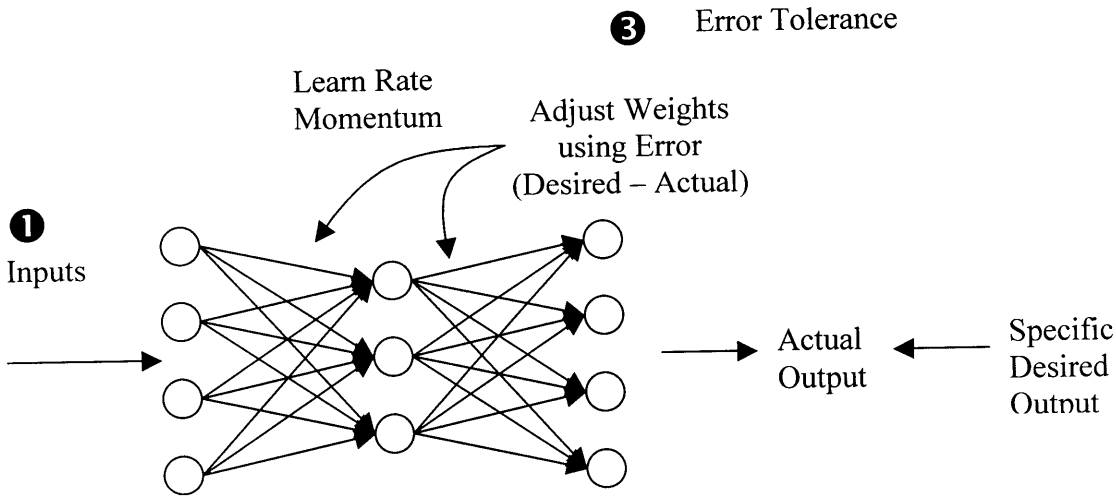


Figure 3.6: Backpropagation networks

3.9 Neural Networks as Tools in Medical

In medical literature, the most popular algorithm for estimating parameters in neural networks is backpropagation (Reggia, 1993) which mostly used for classification, or pattern recognition problems (Ohno-Machado, 1996a). A search in National Library of Medicine (PubMed) found 2844 citations that consist of past and ongoing work up to October 2000 involves neural networks. Eight percent (8%) of total citations is from year 1981 while ninety two percent (92%) is from year 1991 to 2000. This shows a tremendous increase in the work involves with neural network for the past decade.

An artificial neural network was deployed to predict the intracranial hemorrhage in preterm neonates (Zernikow *et al.,* 1998). A total of 865 preterm (<32 weeks gestation) and/or low birth weight (<1500 g) babies admitted between April 1990 and October 1996, 79 of which had an intracranial hemorrhage. They compared the performance of statistical model namely logistic regression models (LRs) and artificial neural networks (ANNs) to identify patterns in complex data sets.
After randomly dividing the data set into 2 equal groups, they created a step-wise LR model and trained an ANN using one of the data sets. Each model was then validated against the other data set, to which it had not been exposed.

The ANN devised for this study consisted of a 3-layer, perceptron-type system with 13 input nodes, each connected to 4 hidden nodes, which were all connected to 1 output node. The ANN used an adaptive gradient learning algorithm with a tangential transfer function in the hidden nodes.

The sensitivity (ability to correctly identify true cases) of the ANN was higher than the LR model at the 75%, 80%, 85%, and 90% specificity (ability to correctly identify patients without intracranial hemorrhage) levels. In addition, the area under the receiver operating characteristic curve was greater (0.935 vs. 0.884; $p < 0.02$) for the ANN than for the LR model. Of the 13 input variables to the ANN, all seemed to contribute significantly to the prediction except "capillary PCO2 on admission" and "gender".

3.9.1 Neural Networks Versus Regression Models

Most studies in the medical sciences are familiar with conventional statistical methods for classification, such as multiple nonlinear regression, and linear discriminant analysis or logistic regression. The complexity of the diagnostic task is thought to be one of the prime determinants of model selection (West, et. al., 2000).

Since the networks start processing the data without any preconceives hypothesis, they start with random weight assignments to various input variables and adjustments are made based on the difference between predicted and actual output. This allows for unbiased and better understanding of data. In addition, neural network helps in explaining subtle relationship between various input variables and output. Neural networks can be retrained using additional input variables and number of individuals and once trained, they can be called on to predict in a new data sets. There are several models of neural networks available to choose from according to a particular problem. Once trained, they are very fast and are cost saving since the accuracy of generalization increases.

Nevertheless, they are some disadvantages since there is no set rules in selecting the networks models. Practitioners are left to choose the model that suits their problem domain. In order to utilize this technology, one must have an in-depth understanding about the technology and the nature of the problem. Furthermore, the training is quite time consuming and requires patience.

CHAPTER 4

ANAEMIA

This chapter describes anaemia in general. However, this chapter focuses on the symptons of anaemia. Furthermore, it also discusses the procedure taken by physicians in classifying anaemia patient.

4.1 Descriptions

"Anaemia" is a common medical problem. The word anaemia is composed of two Greek roots that together mean "without blood" (Ed-Uthman, 1998). Signs and symptoms of anaemia include weakness fatigue, palpitation, light-headedness, difficulty in swallowing, loss of appetite, nausea, constipation, diarrhea, stomatitis and others. The patient looks pale, the nail may be dry and brittle, and tongue may be inflamed. In severe anaemia, heart failure and swelling of both limbs can occur. In mild anaemia, none of the above signs and symptoms may appear (Orkin, 1992).

Patients with anaemia have a significant reduction in red cell mass and a corresponding decrease in the oxygen-carrying capacity of the blood (Orkin, 1992). In General Medical Officer's manual (Luiken *et al.*, 1999), anaemia is define as a decreased level of haemoglobin more than two standard deviations below the expected mean for age and sex. Anaemia itself is not a disease but a sign of disease

(Rapaport, 1987; DeLoughery, 1999). This means underlying disease is presents that demand an explanation.

4.2 Procedures in Classifying Anaemia

In evaluating the anaemic patient, the physician will proceed in orderly fashion so that the correct diagnosis can be established with a minimum number of laboratory tests and procedures. After comprehensive history and meticulous physical examination, the physician will first perform 'Full Blood Count' (FBC). Nowadays, FBC can be done through electronic blood counters. The information given by this machine includes red cell indices such as Haemoglobin level, Packed Cell Volume, Mean Red Cell Volume and the non-red cell indices such as platelet and total white cell counts. If the FBC has suggested that the patient is anemic, further investigation will be carried out. The first further investigation that is normally carried out is Full Blood Picture (FBP), as this investigation provides most of information needed to classify anaemia. To do FBP, a drop of blood is smeared on a piece of slide and staining is done. The blood smear is then studied under microscope and important information can be derived from it (Rubin, 1997).

There are many ways to classify anaemia such as classification based on its causes e.g. failure of red cell production or excess loss or destruction, but one of the most commonly used classification is based on red cell indices (Hoffbrand *et al.*, 1993). After examining the red cell indices and combining them with non-red cell indices, the haematologist will further proceed to study the blood smear or full blood picture

that can help he or she in classifying anaemia. The classification of anaemia will guide the haematologist or physician in choosing further investigation needed to find the cause of anaemia such as Bone Marrow Examination or Iron Study. This is essentially important as the numbers of further tests are numerous and some investigations are very expensive. Factors being considered in classifying anaemia including patients background, red cell indices, non-red cell indices (white blood cell, platelet) and peripheral blood smear (Rubin, 1997).

4.2.1 Patient Background

Age and gender are two vital parameters to determine whether the patient is anemic or not as different age group and gender have different normal value. For example, a 20 years old man with haemoglobin of 11.5 can be considered to be anemic but not in a 20 years old woman. Table 4.1 shows the normal value of haemoglobin and other parameters in male and female (Hoffbrand *et al.*, 1993).

Table 4.1 : Normal adult red cell values

	Male	**Female**
Haemoglobin (Hb)* (g/dl)	13.5-17.5	11.5-15.5
Haematocrit (PCV) (%)	40-52	36-48
Red cell count (x 10^{12}/l)	4.5-6.5	3.9-5.6

4.2.2 Haemoglobin and Red Cell Indices

Red cell indices measured includes Haemoglobin level, Mean Corpuscular Volume (MCV), and Pack Cell Volume or haematocrit. MCV is a calculated value by dividing the haematocrit from the RBC count. The measurements of these red cell indices help the clinician or haematologist classify the type of anaemia (Rapaport, 1987).

- Decreased MCV is known as Microcytic Anaemia (small RBC, less haemoglobin) e.g. iron deficiency.
- Elevated MCV is known as Macrocytic Anaemia: The RBC are macrocytic (bigger than normal) because of incomplete maturation in the bone marrow due to the deficiency of vitamin B12 and Folic acid.
- Dimorphic anaemia: In this anaemia both small and large RBCs appear in the circulation due to combined Iron, Folic acid, and Vitamin B12 deficiency.

4.2.3 Reticulocyte Counts

New red blood cells which has just been produced by our body can be distinguished from 'old' red cell as these new cells contain a residual component known as RNA that can be stained by using special dye. The normal reticulocyte counts is between 0.5 to 2 %. In disease state, the reticulocyte count can be increased such as in haemolytic anaemia or decreases such as in marrow failure (Orkin, 1992).

4.2.4 White Blood Cell

White blood cell or leukocyte is also a component of blood that has to be considered in classifying anaemia. They involve in body defense system. Leukocyte can be

divided into two groups phagocyte and immunocytes. Granulocyte which includes three types of cell: neutrophils, eosinophils and basophils. Together with monocytes it comprises phagocyte. The function of phagocytes and immunocytes mainly in body defense system against infection. Changes in quantity or morphology may occur in certain medical condition and close examination of these cell may reveal the underlying pathology and help to establish type of anaemia such as elevation of neutrophils in haemolysis or haemorhage or elevation of eosinophils in parasite infestation or allergic reaction (Orkin, 1992).

4.2.5 Full Blood Picture

The shapes and characteristics examined during Full Blood Picture studies includes Elliptocyte / Pencil cells, Ovalocytes, Tear drop cells, Schistocytes / Fragmented cells, Polychromasia, Sperocytes, Targets Burr cell, Acanthocyte, Sickle cell, Rouleoux Aggregates Neclueted (NRBC), Bite cells, Stomatocytes, and Basophilic Stippling. Abnormal red cell morphology or red cell inclusions may suggest particular diagnosis (Rapaport, 1987). The color intensity and size of the red cell is examined. The red cell will be classified into hypochromic, normochromic , microcytic, normocytic or macrocytic. When causes of both microcytosis and macrocytosis is present, e.g. mixed iron and folate/B12 deficiency, the red cell indices obtained from FBC may be normal but the FBP will be abnormal as red cells studied under microscope may reveal 'dimorphic' appearance (a dual population of large, well haemoglobinized cells and and small, hypochromic cells). During the blood film examination, the white cell differential count is also performed, platelet number and

morphology are assessed and the presence of abnormal cell, e.g. normoblasts, granulocyte precursors or blast cell is noted.

As been mentioned previously, the combination of various parameters is needed in classifying the type of anaemia. As the causes of anaemia are numerous, they are normally classified into several main groups as shown in figure 5.1.

has been used by the haematologist that influence anaemia classification. The questionnaire mainly has three sections i.e. Full Blood Count (FBC), Full Blood Picture (FBP) and Result (Target). The target is the diagnosis provided by the haematologist who studies the above investigations. A questionnaire sample is attached in Appendix B.

5.2 Data Set Description

The attributes and its domain that have been simplified from the questionnaire is presented in Table 5.1. Seventeen attributes or variables were used in the training model. The last attributes represent the target or output value.

5.3 Data Preparation

One goal of data preparation is to reduce nonlinearity when its character is known and let the network resolve the hidden nonlinearities that are not understood (Tsoukalas and Uhrig, 1997).

5.3.1 Data Cleansing

Data cleansing deals with incomplete patterns, unreliable and uneconomical of patterns. Although neural network can work with incomplete data sets, missing data can create serious problems (Tsoukalas and Uhrig, 1997). Therefore, data sets with missing values and unreliable are removed at this stage which resulted in 600 data sets.

CHAPTER 5
METHODOLOGY

This chapter describes the methodology used in producing the most suitable model of neural network. This is accomplished by determining the optimum size of hidden units, the optimum value of learning rate and momentum, the suitable weight distribution and update, the suitable activation function and finally validate the results with other models like Radial Basis Function (RBF) and Regression. The chapter starts with data source, data sets descriptions and data preparation.

5.1 Data Source

Data was obtained from Pathology Lab of Hospital Alor Setar, Kedah with assistance from Dr Abdul Rashid Mohd Ibrahim, a haematologist at Unit Bank Darah. Much information was provided regarding the concept of haematology particularly blood's composition and factors that influence anaemia classification.

Data collections were mainly based on the information written in the haematology form. Clinicians that requested the blood investigation are required to fill in the haematology form properly. A total of 732 data of the year 1999 sets were collected. Several examples of Haematology form that were used during the study are attached in Appendix A. Apart from the Haematology form, a questionnaire was also developed. The objective of the questionnaire is to consider all the available data that

Table 5.1: Attributes Information

Attributes	**Domain**
1. Age	Age
2. Sex	0,1
(Full Blood Count, FBC)	
3. FBC-White Blood Cell (WBC)	(exact value) X10^3/*u*L
4. FBC-Red Blood Cell (RBC)	(exact value) X10^6/*u*L
5. FBC-Haemaglobin (HGB)	(exact value) G/dL
6. FBC-Haematocrit (HCT)	(exact value) %
7. FBC-Platelet (PLT)	(exact value) X10^3/*u*L
8. FBC-Reticulocyte (RTC)	(exact value) %
9. FBC-Eosinophil (EOS)	(exact value) %
(Full Blood Picture, FBP)	
10. FBP- Red Blood Cells (RBC)	1-5
11. FBP-Shape1	1-16
12. FBP-Shape2	1-16
13. FBP-Shape3	1-16
14. FBP-Shape4	1-16
15. FBP-WhiteCount1	1-9
16. FBP-WhiteCount2	1-9
17. FBP-Platelet	1-9
Classes of Anaemia	
18. Target	1-17

5.3.2 Data Selection

Tsoukalas and Uhrig stated that there must be enough examples of sufficient varieties for training so that the network will be able to make valid correlation and generalizations for unfamiliar cases. They added that the varieties must include good distribution of possible inputs and outputs. The distribution of output patterns before selection is shown in Figure 5.1 with seventeen classes of anaemia.

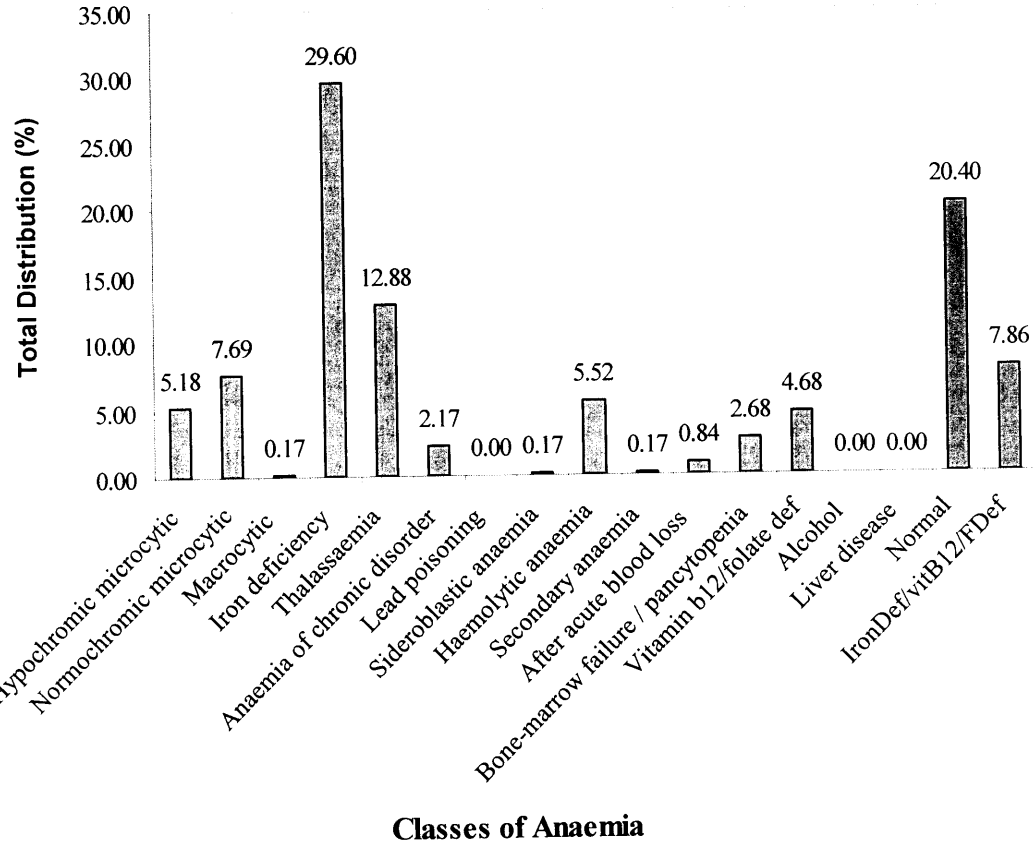


Figure 5.1: Pattern distribution of anaemia classes (before selecting)

Classes that have more than 5% of total distribution ware selected. This is to avoid of what Lowe (Lowe, 1990) states, "Adaptive networks trained on a 1-from-*c* 1 classifier problem exhibit a strong bias in favor of those classes which have the largest membership in the training data."

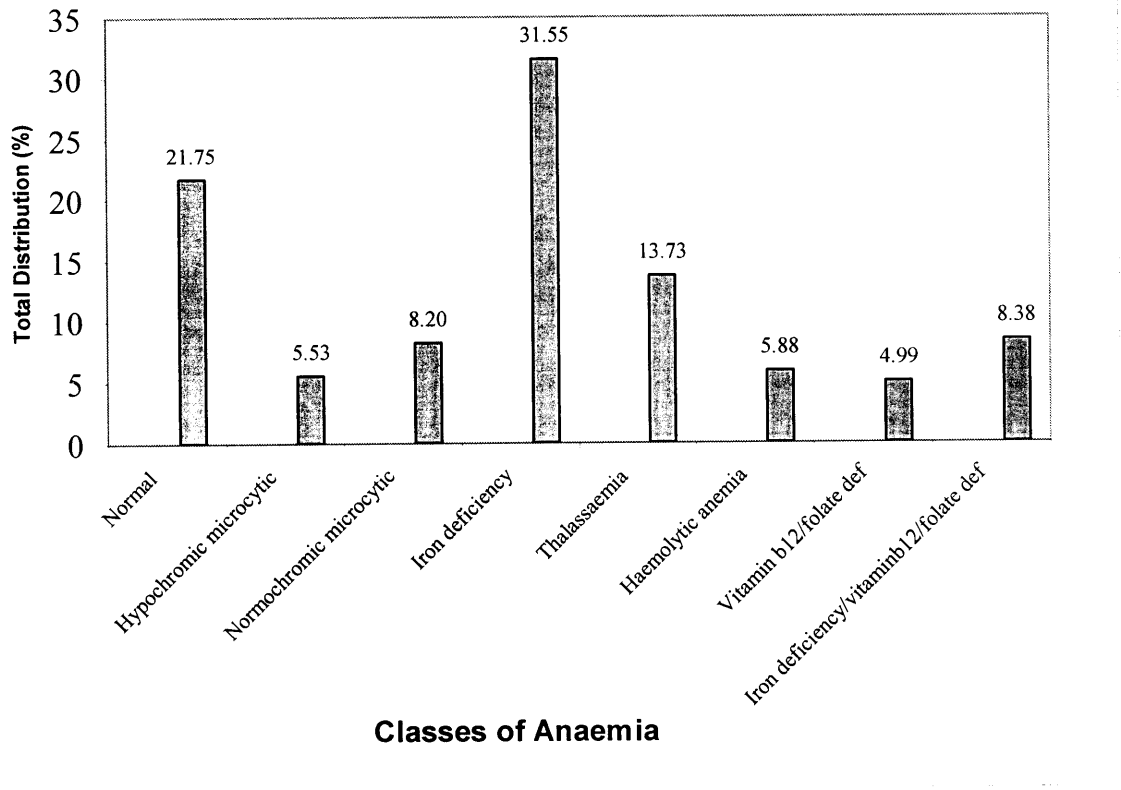


Figure 5.2: Pattern distribution of anaemia classes (after selecting)

Figure 5.2 shows the distribution of data sets after removing the rare categories that resulted in 561 data sets belong to eight classes of anaemia. However, distribution of anaemia classes was still not well distributed. Thus, a portion of these data sets was randomly removed using SPSS software as suggested by Tsoukalas and Uhrig, that the variety must include a good distribution of possible inputs and outputs distribution of data sets. The final distribution of eight anaemia classes consists of 400 data sets (see Figure 5.3).

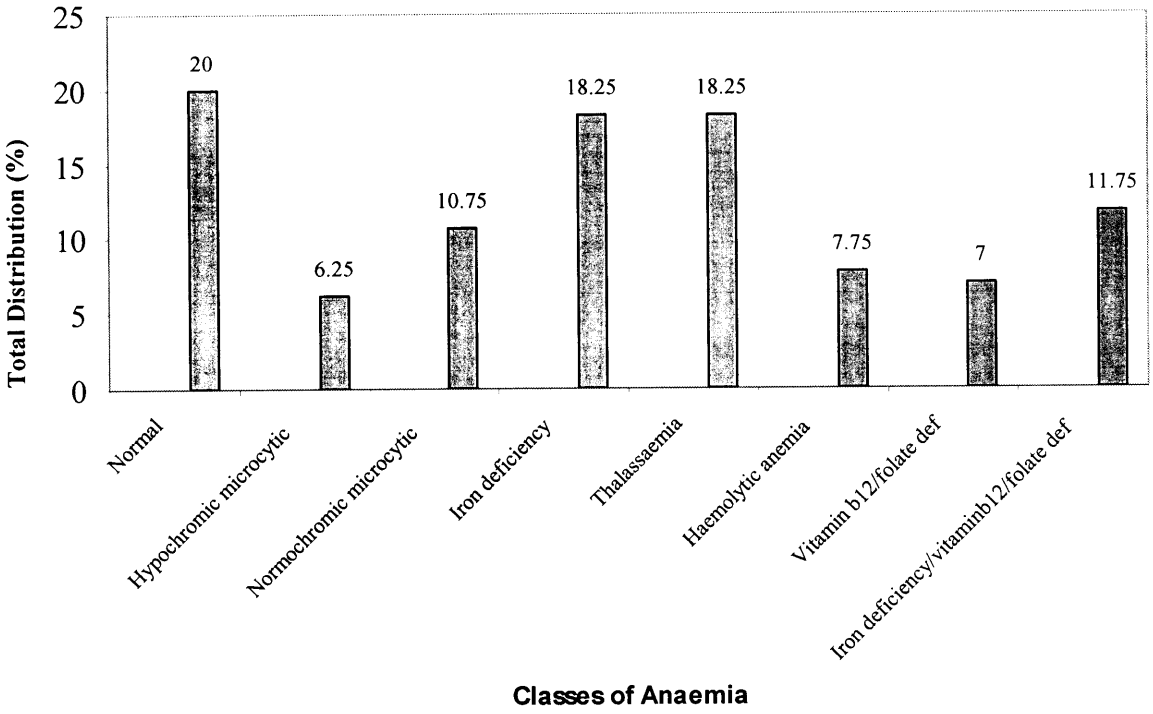


Figure 5.3: Distribution of data sets from eight classes of anaemia

5.3.3 Data Preprocessing

Before training can take place, data must be processed in a form that is meaningful to neural network. Each input variable should be preprocessed so that its mean value, averaged over the entire training set, is close to zero, or else it is small compared to its standard deviation (LeCun, 1993). Neural networks are very sensitive to absolute magnitudes. If one input ranges from 1000 to 1,000,000 and a second input ranges from 0 to 1, fluctuations in the first input will tend to swamp any importance given to the second (Tsoukalas and Uhrig, 1997).

In this study, all the input patterns have been normalized and scaled down between 0 and 1 so that, they correspond roughly to the same range of values. Normalization as been defined by Tsoukalas is simply by dividing all values of a set by an arbitrary reference value, usually the maximum value. However, this process carries with it the potential for loss of information as it can distort the data if one or a few values are much larger than the rest of the data.

$$\mathbf{X_{new}} = \frac{\mathbf{X_{max}} - \mathbf{X_{old}}}{\mathbf{X_{max}}} \qquad (5.1)$$

Only one neuron is needed when the choice is between two categories, but one neuron is needed for each category when there are more than two alternatives (Tsoukalas and Uhrig, 1997). This situation is described as a non-distributed representation. Since the final distribution of data sets in this study has eight different classes, thus eight neurons were used to represent the output (see Figure 5.4).

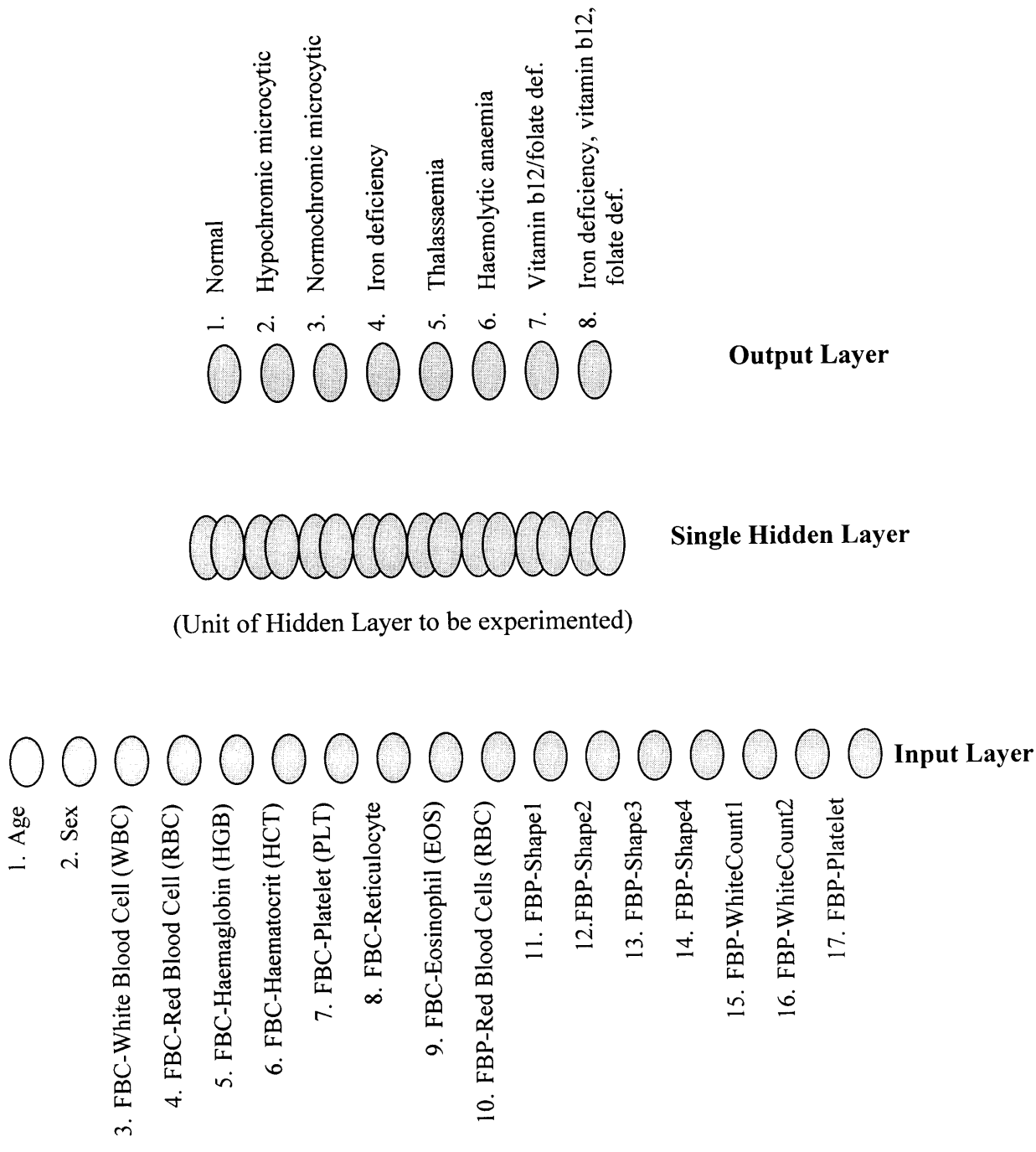


Figure 5.4: Schematic Representation of Model used in Training
(17 input nodes and 8 output nodes)

5.4 Data Training and Testing

The aim of this section is to get the suitable model by determining the optimum value of hidden unit, learning rate and momentum. This was achieved by using Neural Connection (v 2.0) software to train the available data sets. There is a relationship between gain, learning rate, and weights in backpropagation neural network (Thimm *et al.*, 1996).

The essence of back-propagation learning is to encode an input-output mapping into the synaptic weights and the thresholds of a multilayer perceptron (Haykin, 1999). The hope is that the network becomes well trained so that it learns enough about the past to generalize the future.

In this study, a *cross-validation's* rule was used that provides an appealing guiding principle (Stone, 1974) where the available data set was randomly partitioned into a training set and a test set. The training set was further partitioned into two disjoint subsets: the estimation subset (training set) used to select the model and validation subset, used to test or validate the model. The division of data sets is shown below:

Training set – 80%; Test set –10%; Validation set – 10%

The training data is used to train the model while the validation data is used to monitor neural network performance during training. The test data is used to measure the performance of a trained model. The motivation here is to validate the model on a

data set different from the one used for parameter estimation. In this way, we may use the training set to assess the performance of various candidate models, and thereby choose the "highest" one. However, there is a distinct possibility that the model with the highest performing parameter values may end up overfitting the validation set (Haykin, 1999). In guiding against this possibility, the generalization performance was measured against the data set.

5.4.1 Early Stopping Method of Training

A multi layer perceptron trained with the backpropagation algorithm learns in stages, moving from the realization of fairly simple to more complex mapping functions as the training session progresses (Haykin, 1999). Haykin added that this is exemplified by the fact that in a typical situation the mean-square error decreases with an increasing number of epochs during training. It starts-off at a large value, decreases rapidly, and then continues to decrease slowly as the network makes its way to a local minimum on the error surface. Therefore, it is very difficult to figure out when it is suitable to stop training if we were to look at the learning curve for training by itself. Furthermore, it is possible for the network to end up overfitting the training data if the training session did not stop at the right point.

In this exploratory study, the onset of overfitting is identified by cross-validation method that has been discussed in the previous section. The procedure is referred to as the early stopping method of training (Haykin, 1999). The following figure (see Figure 5.5) shows the conceptualized forms of two learning curves. The estimation

learning curve decreases monotonically for an increasing number of epochs. In contrast, the validation learning curve decreases monotonically to a minimum, it then starts to increase as the training continues.

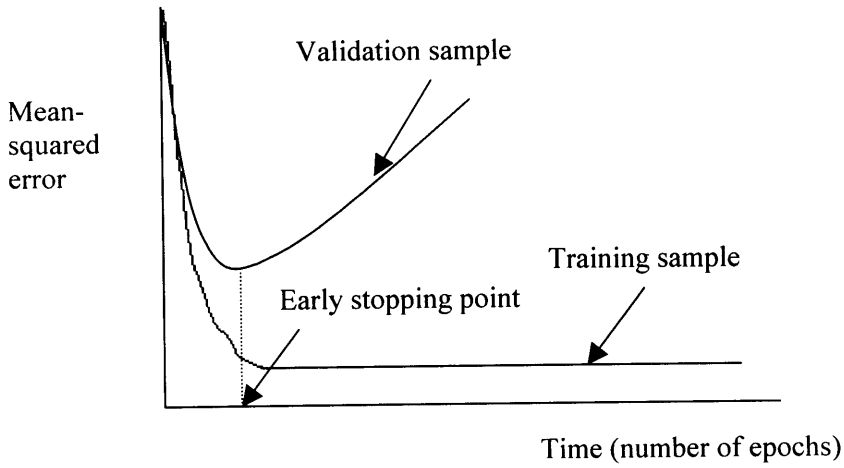


Figure 5.5: Illustration of the early-stopping rule based on cross validation

5.4.2 Training Initialization

- Weight seed, learning rate and momentum were initialized to 0.1.
- Each configuration of Multi Layer Perceptron (MLP) was trained with ten simulations each with a different condition.

- ***Weight Distribution*** – Uniform where the initial weights were randomly selected and have values that cannot exceed the range set.

- ***Learning rule***

  a) Algorithm - Steepest descent method was used rather than conjugate gradient because it measures the gradient of the error surface after each iteration and changes the weights in the direction of the steepest descent.

  b) Weight Update – Batch Mode (epoch) where weight updating is performed after the presentation of all the training examples that constitute an epoch.

- ***Stopping Criteria*** – In avoiding overfitting during training, the training was stopped once the validation error rate "starts to go up", as discussed in section 5.5.

- ***Activation Function*** - Sigmoid was used for two main reasons:

  (1) The data sets used in this study were real values and positive. (2) Sigmoid is the most commonly used compared to other activation function. It is defined as a strictly increasing function that exhibits a graceful balance between linear and nonlinear behavior (Haykin, 1999).

- ***Number of hidden layers*** – Single layer. The Universal Approximation Theorem states that *"a single hidden layer is sufficient for a multilayer perceptron to compute a uniform ε approximation to a given training set represented by the set of inputs $x_1, \ldots, x_{m0}$ and a desired (target) output $f(x_1, \ldots, x_{m0})$.* (Haykin, 1999)

Table 5.2 presents the variable parameters that were used in MLP.

Table 5.2: Variable Parameters of Multilayer Perceptron

Parameter	**Symbol**	**Typical Range**
Number of hidden units	m_1	$(2, \infty)$
Learning-rate parameter	α	(0,1)
Momentum constant	μ	(0,1)

The nature of this study is well suited to multilayer perceptron (MLP) to train the data of type supervised manner (data with target value) with a highly popular algorithm, *error backpropagation algorithm.* In this study, the following parameters are explored in order to find the most suitable one.

(i) Size of hidden unit, m_1

(ii) Value of learning rate, α

(iii) Value of momentum, μ

(iv) Activation functions between Sigmoid, Tangen and Linear

(v) Weight distributions between Uniform and Gaussian mode

(vi) Weights update between Epoch or Pattern mode

(vii) Comparison of MLP model with other alternative models like Radial Basis Function (RBF) and Regression (REG) model

5.4.3 Exploratory on Different Parameters

The aim of this section is to obtain the network with the highest performance on the test set.

(i) Number of Hidden Units

Number of hidden units has great influence on performance of the networks. The presence of hidden layers allow neural networks to approximate a variety of non-linear functions (Mitchell, 1997). Too few neurons in the hidden layer prevent it from correctly mapping inputs to outputs, while too many impede generalization and increase training time. Too many neurons may allow the network to "memorize" the patterns presented to it without extracting the pertinent features for generalization (Tsoukalas and Uhrig, 1997). Hence, an optimum size has to be attained and is often accomplished through experimentation.

(ii) Learning rate

Learning rate α is another important parameter that may effect the generalization of networks. Tsoukalas and Uhrig pointed out that α cannot be negative because this would cause the change of the weight vector to move away from the ideal weight vector position. Therefore, α must always be positive. Both analytically and experimentally have shown that if α is greater than 2, then the network is unstable, and if α is greater than 1, the weight vector will overshoot the ideal position (Tsoukalas and Uhrig, 1997). Hence, α should be in the range between 0 and 1 Until

now, there is no specific rule about the value that the learning rate should have. In this exploratory study, α is changed from 0.1 to 1.0 as the networks learn.

(iii) ***Momentum coefficient***

Momentum μ can enhance the stability of training process by reducing the training time. Haykin (Haykin, 1999) says that momentum is used to keep the training process going in the same general direction analogous to the way that momentum of a moving object behaves. This prevents the learning algorithm from stopping in a local minimum rather than the global (ideal) minimum.

(iv) ***Weight Distributions***

This section aims to see the difference of training and testing performance between Uniform and Gaussian distribution. Starting form this section onwards, the optimal value gained from (*i*) to (*v*) is applied in the networks. Neural Connection allows weight distribution in two forms, Uniform and Gaussian distributions. It also allows the change of the weight values range. With a Uniform distribution, the initial weights are randomly selected and have values that cannot exceed the range set. With Gaussian distribution, the initial weights are randomly selected based on a Gaussian distribution with a variance equal to the selected range value.

(v) ***Weights Update***

While different weight distributions are observed in the previous section *(iv)*, this section aims to investigate the performance of networks with different weight update,

namely Epoch and Pattern mode. In the epoch mode or also known as batch or off-line mode, weight updating is performed after the presentation of all the training examples that constitute an epoch. On contrary, weight updating is performed after the presentation of each training example in the pattern mode. Pattern mode is also referred to as on-line, sequential or stochastic mode.

(vi) Activation Functions

Although Sigmoid is chosen as the preferred activation function, other activation functions like Tangen and Linear are also interesting to be observed. The networks still had the same architecture with the previous stage (17: m_h:8). This is to ensure that the suitable activation function is used in producing the most suitable model.

(vii) Validating MLP against alternative models

Finally, the MLP model with the parameter gained from *(i)* to *(v)* is compared with Radial Basis Function (RBF) and Regression model. This is to ensure that MLP is the most suitable model in this study domain. RBF is another modelling and forecasting tool to classify or predict the patterns. Training in RBF uses one stage process rather than iterative process as in MLP. Regression on the other hand, makes two critical assumptions. The model assumed that the outputs and the inputs are linearly related and there is no interaction between the input variables.

CHAPTER 6
RESULTS AND DISCUSSION

The questions that have been posed in the previous chapter are explored in this chapter. Section one to six presents the results of varying the parameters and the effects of those parameters are discussed and presented graphically. The following section compared the most suitable model with other neural network models namely, Radial Basis function and Statistical models.

6.1 Effect of Hidden Units

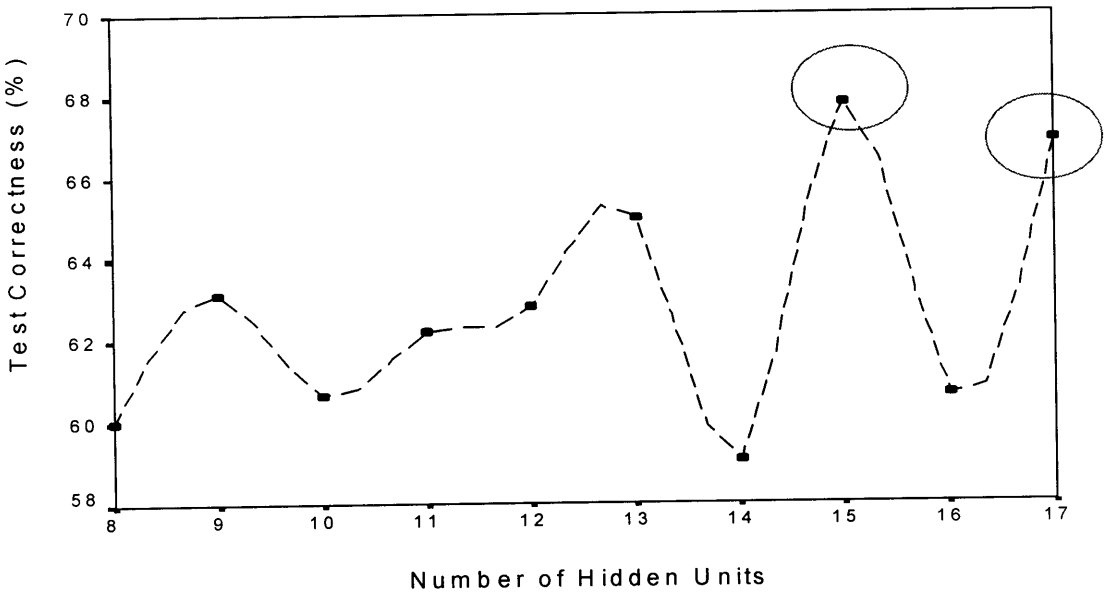
This section investigates the effect on network generalization when the input size is fixed but the hidden units size are varied. The criterion used in selecting the hidden unit is the highest average test and lowest average training correctness. The networks have the following architecture: 17: m_h: 8. For a single hidden layer, it is common practice to initially make the number of neurons equal to about two-third of the number in the input layer (Bailey and Thompson, 1990). In this study, the experiment was simulated with 10 different hidden units, starting from 8 hidden units up to 17 hidden units.

The purpose of simulations set is merely to ascertain the best two hidden units by seeing the correctness gained in the testing set. Therefore, the learning rate parameter α and momentum constant μ are arbitrarily sets to some nominal values, and in this case, they were set to 0.1. The best two hidden units are then retrained with different weight seeds that value from 1 to 10. This is to ensure that the selected value does

give a better generalization than the other one. The same criterion in selecting the hidden units was applied. Figure 6.1 shows the training and test correctness as the number of hidden unit increased. In general, the graph shows that the accuracy of networks for both data sets (data training and testing) was found in the larger size of hidden units when compared with hidden units 8. The best two results for training and generalization occurs for networks with 15 and 17 hidden units. The detail of the results is shown in Appendix C (Table 1).

Table 6.1: Train and test correctness with different size of Hidden units, m_h

m_h	8	9	10	11	12	13	14	15	16	17
Train	61.13	63.36	62.31	62.78	62.78	62.66	60.59	67.78	63.09	66.49
Test	60.00	63.13	60.63	62.19	62.81	65.00	59.06	67.81	60.63	66.88



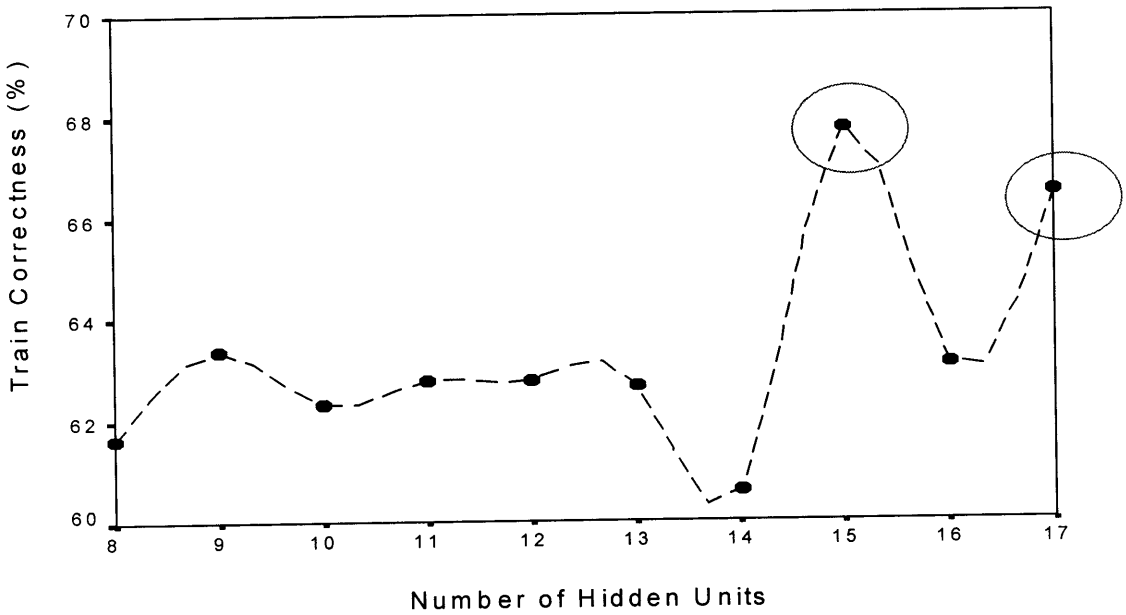


Figure 6.1: The correctness of networks with a topology of 17: m_h: 8 with different size of hidden units. The graph on the top is the training accuracy. The graph on the bottom is the test accuracy.

The experiment is further trained with different sizes of weight seeds. Each hidden unit is retrained with weight seed that was varied from 1 to 10. The result is shown in Figure 6.2.

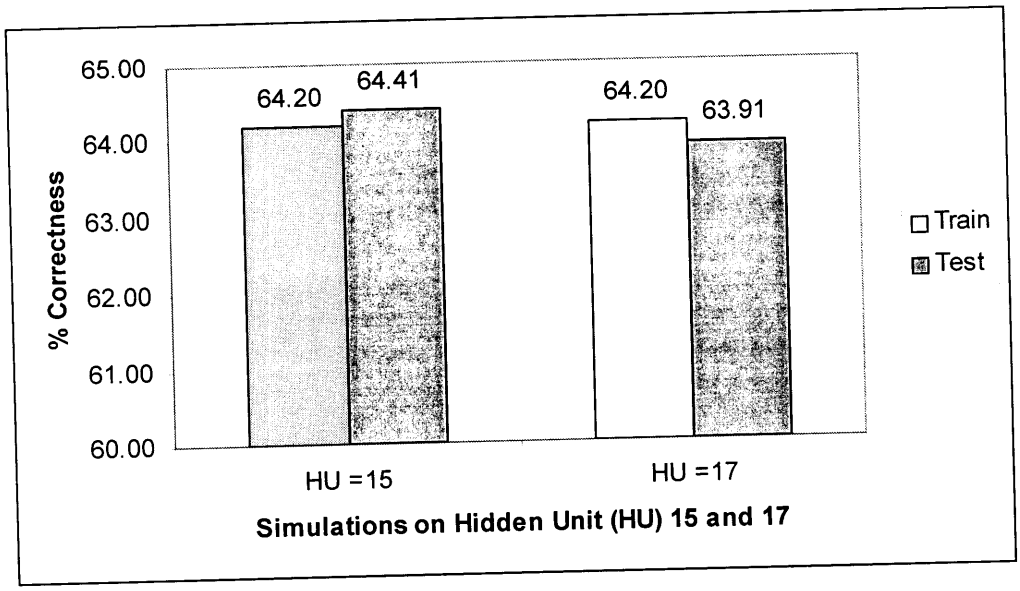


Figure 6.2: Simulations on Hidden unit 15 and 17. The networks generalize better with hidden unit 15.

The results show that hidden unit 15 is the optimum size for the topology of 17: m_h:8. The generalization performance is very poor for hidden units 8 to 14. This indicates that networks require larger number of hidden units in order to achieve good generalization. However, this does not necessarily means that a smaller size of hidden units would not generalize better. The trend may vary according to the generating network size (number of inputs, hidden units and outputs), and the nature of the target function. The detail of the results is shown in Appendix C (Table 2).

6.2 Effect of Learning Rate

This section investigates the network generalization of the multilayer perceptron with the hidden unit size fixed, m_{opt} =15 but the learning rate, α (also known as learning coefficient) is varied from 0.1 to 1.0. The momentum constant μ is arbitrarily sets to 0.1. The same criterion is used as in the above section, in which the learning rate that has the average highest test correctness is selected. The same neural networks architecture is used (17: m_h:8) but the value of hidden units (m_h) is fixed to 15.

Table 6.2: Train and test correctness with different size of Learning Rate α

α	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
Train	67.78	67.85	67.82	67.89	67.89	71.68	72.78	69.42	71.61	72.11
Test	67.81	67.50	67.19	67.19	66.88	61.56	71.56	68.44	71.25	70.31

Table 6.2 shows the training and test correctness as the value of α increases. No large differences is observed from α = 0.1 to 0.5. However, the performance starts to improve when α = 0.6. It is of interest to observe that both data sets achieved their best correctness when α = 0.7. Figure 6.3 shows the trend of training and testing correctness as the value of α increase. The detail of the results is shown in Appendix C (Table 3).

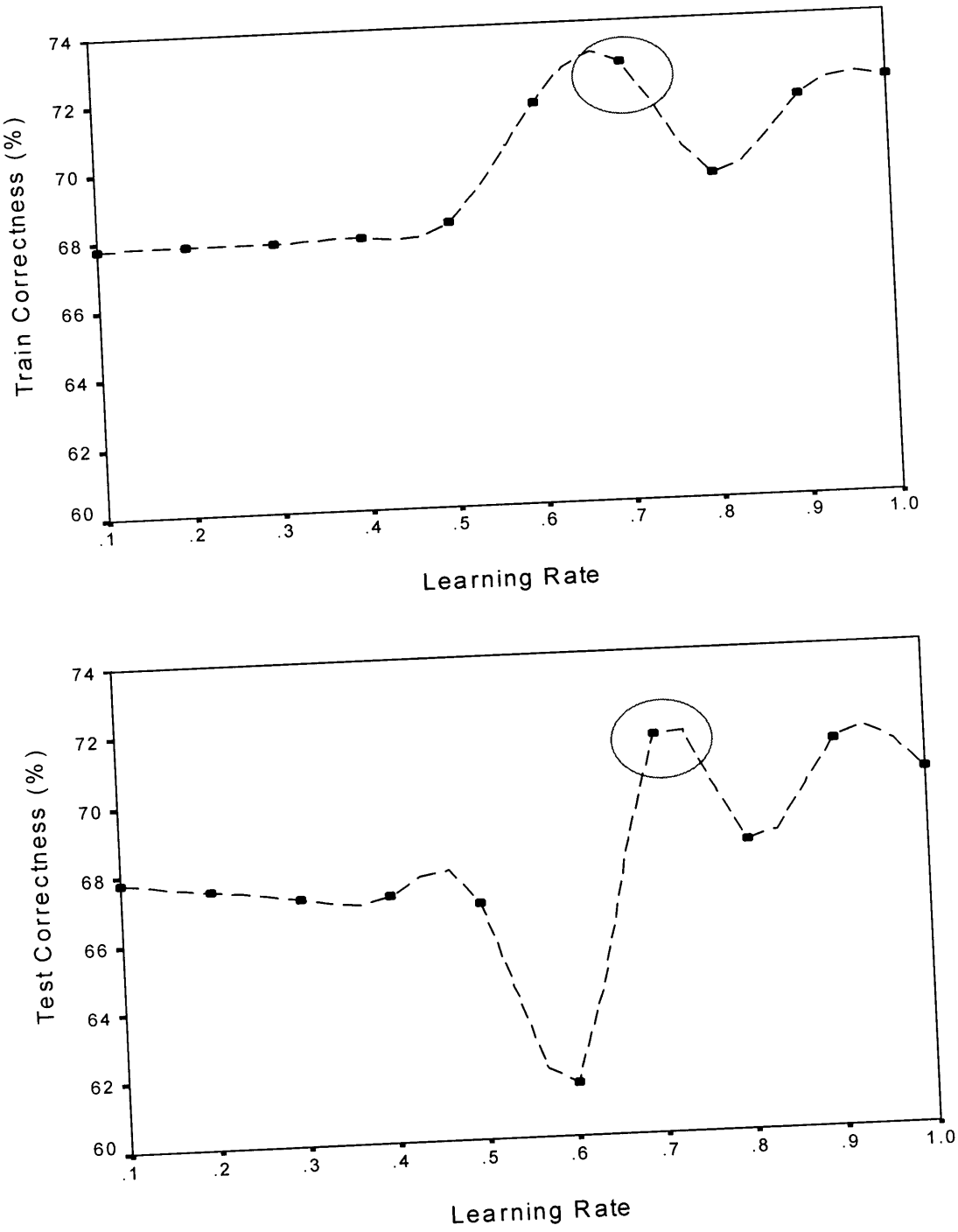


Figure 6.3: The correctness for networks with a topology 17:15:8 for different values of learning rate α. The graph on the top is the training correctness. The graph on the bottom is the test correctness.

6.3 Effect of Momentum

The section investigates the training behaviour of the networks with the hidden units size (m_{opt} =15) and learning rate fixed but the momentum μ is increased. Momentum with highest test correctness is selected. The networks still used the same architecture as in the previous section 17:15:8.

Table 6.3: Train and test correctness with different size of momentum μ

μ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
Train	67.78	67.82	67.89	67.82	67.89	67.73	67.85	68.32	68.24	66.88
Test	67.81	67.81	67.50	67.19	67.19	67.19	67.19	67.50	67.50	66.88

Table 6.3 presents the training and test correctness as the value of momentum μ is increased. On average, the results show that the generalization decreases as the value of μ increases. Although μ = 0.1 and 0.2 in the testing set gave the most highest correctness, μ = 0.1 is selected with reference to the setting criterion; in which the best generalization is obtained from the highest test and lowest training. Figure 6.4 shows the trend of training and testing correctness as the value of μ is increased. The detail of the results is shown in Appendix C (Table 4).

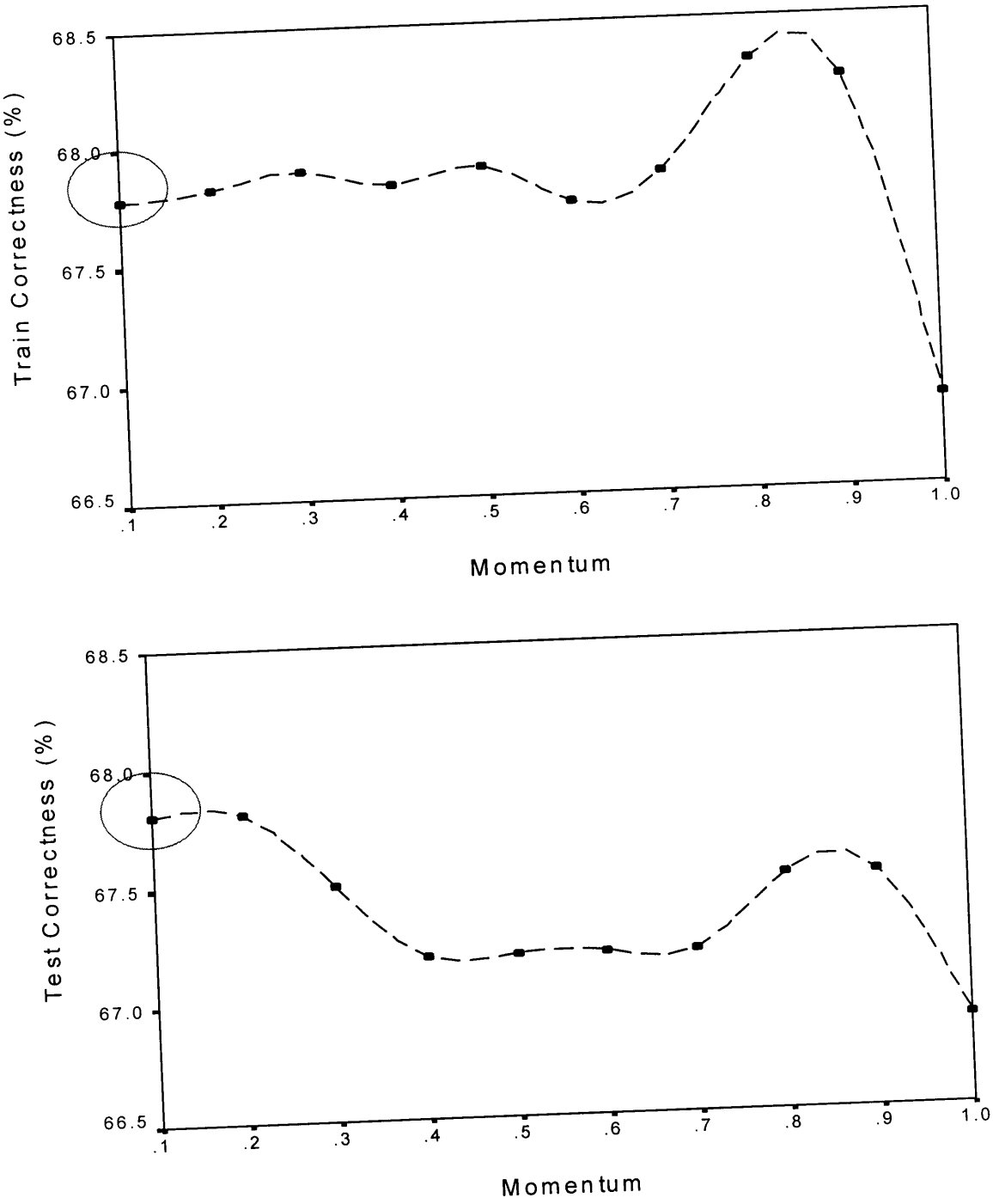


Figure 6.4: The correctness for networks with a network 17:15:8 for different values of momentum μ. The graph on the top is the training correctness. The graph on the bottom is the test correctness.

6.4 Effect of Different Weight Distributions

This section examines Uniform and Gaussian distribution functions in Multilayer Perceptron training. With the same networks architecture 17:15:8, the behaviour of the networks is observed. In this section, the optimal number of hidden unit, values of learning rate and momentum are applied to the network. From the chart below, these two functions produced different performance correctness. As seen in Figure 6.5, best error of Uniform is better than that of Gaussian function. The results are plotted in Figure 6.6 show the correctness gained between these two distribution functions. The detail of the results is shown in Appendix C (Table 5).

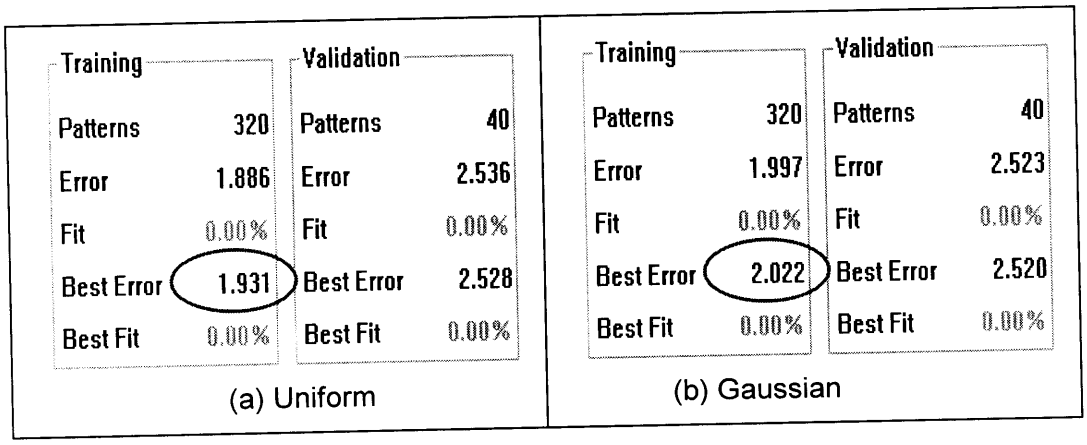


Figure 6.5: Best Error of Uniform and Gaussion Distribution functions. The best error is achieved with Uniform weight distributions.

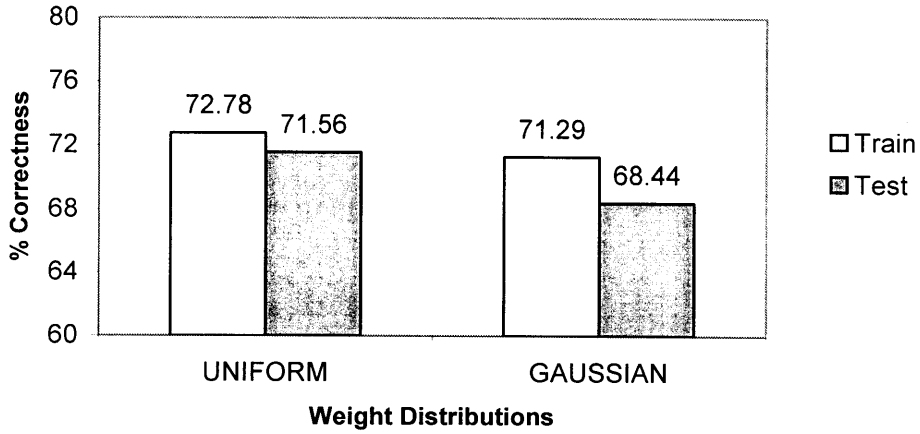


Figure 6.6: Correctness performance of the two functions. The performance of network with Uniform distribution is higher than that of Gaussian with a difference of 1.49%.

6.5 Effect of Different Weights Update

This section examines the network performance when applied with different mode of weight update. Epoch and Pattern weight updates are observed. The same network architecture is applied for both 17:15:8. In this section, the optimal number of hidden units, learning rate and momentum are applied to the networks and their performance is observed. The networks give best error with the epoch mode (see Figure 6.7). The results are also plotted in Figure 6.8 to show the correctness gained between these two distributions function. The detail of the results is shown in Appendix C (Table 6).

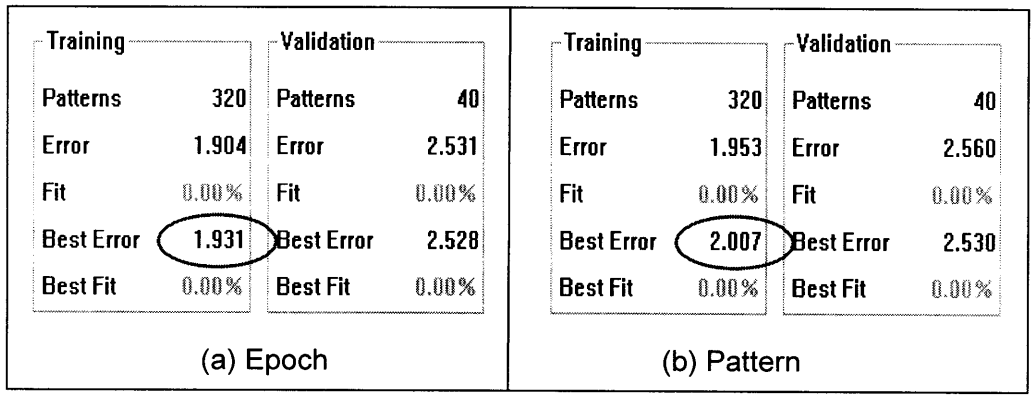


Figure 6.7: Best Error of Epoch and Pattern weights update. The Best error is achieved with Epoch distribution functions.

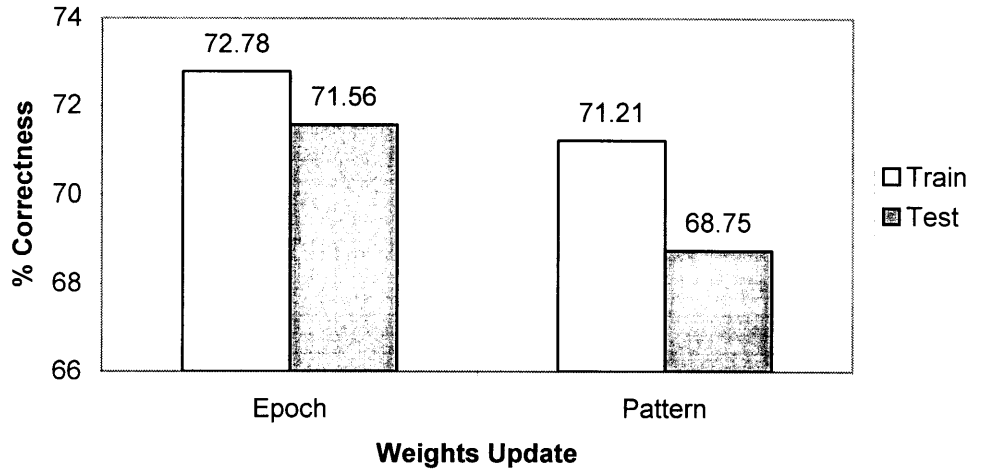


Figure 6.8: Correctness performance of the two functions. Networks performance with Epoch update gives better generalization behaviour with correctness of 71.56%compared to Pattern update of 68.44%.

6.6 Effect of Activation Functions

This section examines the networks performance on different type of activation functions namely, Sigmoid, Tanh and Linear activation function. Sigmoid activation function is preferred in the beginning of the training study as it meets the nature of the problem that has data sets of real values. Using the same network architectures of 17:15:8, the behaviour of the three activation functions was observed. The results are plotted in Figure 6.9.

The performance of networks in training and testing data sets of sigmoid activation functions is better than hyperbolic Tangent activation functions. Thus supporting the statement made by Haykin (Haykin, 1999) and other researchers why sigmoid is the most preferred form of activation function. However, hyperbolic Tangent activation function (also known as Bipolar sigmoid) would be more suitable if the data sets are in the range of -1 and 1. Linear activation function does not give any performance value. The results indicate that the non-linear neural networks are indeed suitable for the research domain. The detail of the results is shown in Appendix C (Table 7).

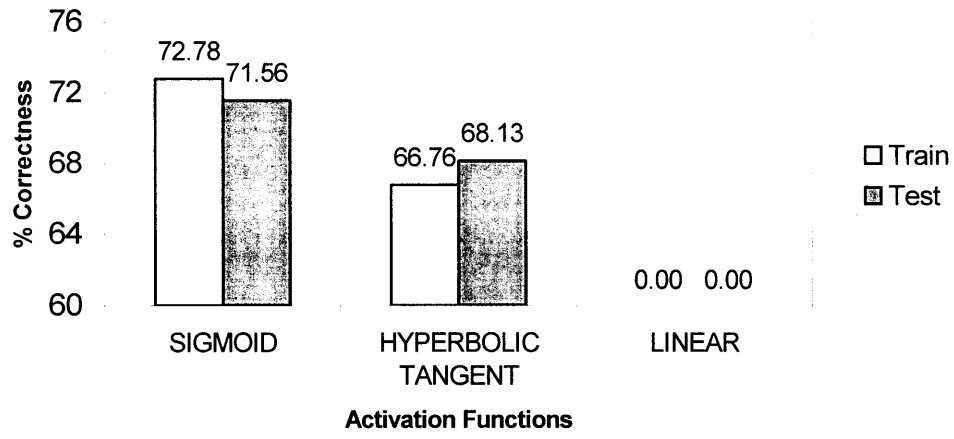


Figure 6.9: Correctness performance between Sigmoid, Tan and Linear activation functions. Sigmoid produces the highest generalization of the three activation functions.

6.7 Comparison Between Alternative Models

This section evaluates the performance of Multilayer perceptron (MLP) and other neural network models namely, Radial Basis Function (RBF) as well as statistical model such as Regression (REG) model. With architecture of 17:15:8, the network is set with the optimum parameters gained from the previous sections (section one to six). The detail of the results is shown in Appendix C (Table 8).

Figure 6.10 shows the performance of MLP, RBF and REG models. The result of MLP is indeed superior to the other models for both data sets. When compared with RBF and REG models, a difference of 7.81% and 10.62% were observed respectively.

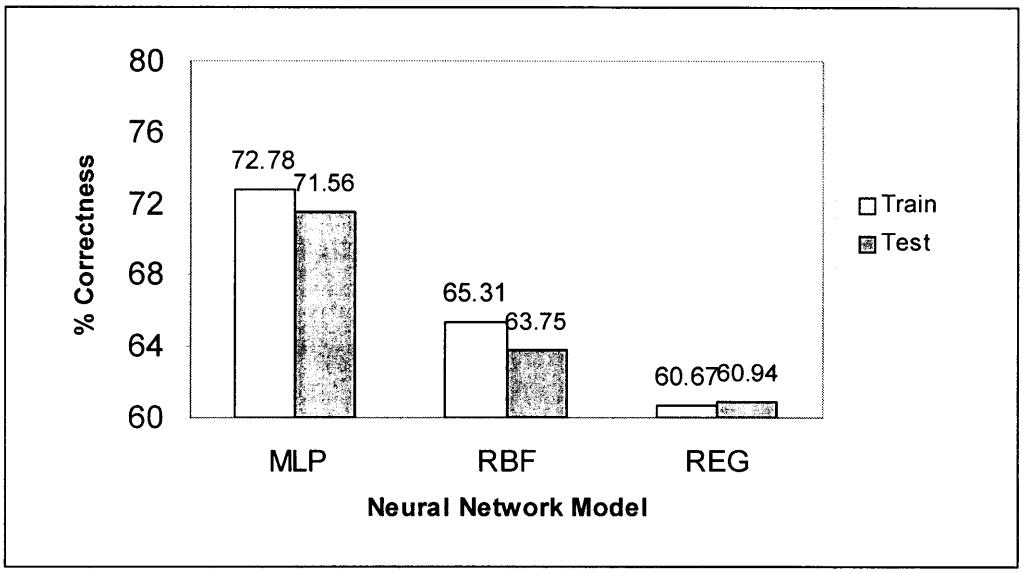


Figure 6.10: Correctness performance between Multilayer perceptron (MLP), Radial Basis Function (RBF), and Regression (REG).

Figure 6.10 shows that the Regression model performance is the lowest compared to neural networks model (Multilayer Perceptron and Radial Basis Function). Similar results are obtained when multilayer perceptron model used linear activation functions (see Figure 6.9). These two results indicate that non-linear model such as multilayer perceptron and radial basis function are able to generalize the problem domain and multilayer perceptron has shown to be the most suitable model in the context of this study.

Figure 6.11 shows the schematic diagram of multilayer perceptron with 15 numbers of hidden units while Table 6.6 shows the most suitable parameters that were gain in this study.

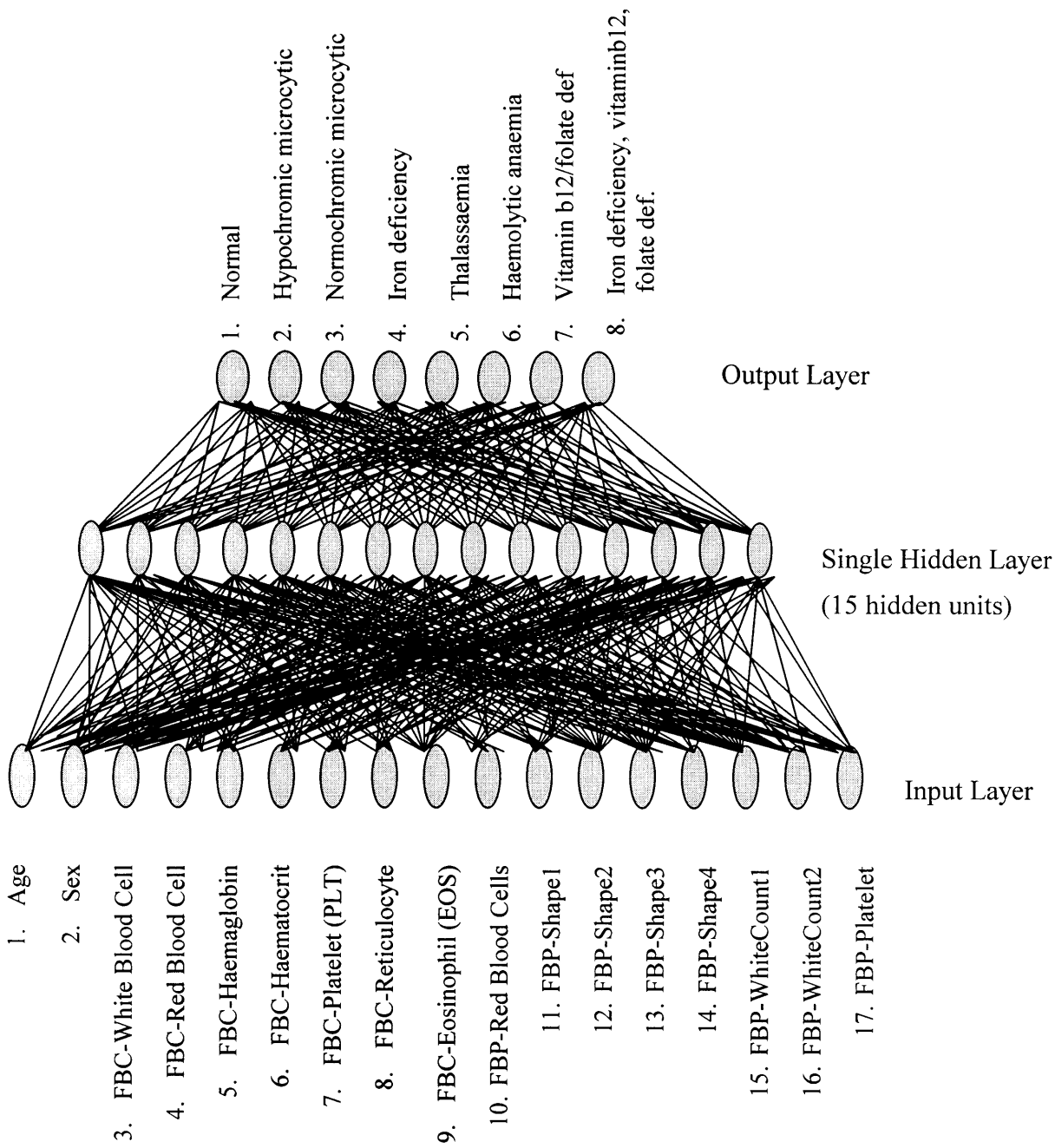


Figure 6.11: Schematic Representation of the most suitable model attained with 17 units of input layers, 15 units of hidden units and 8 units of outputs.

Table 6.4: Configuration of Optimized Multilayer Perceptron	
Parameter	**Value**
Optimum number of hidden units, m_{opt}	15
Optimum learning rate parameter, m_{opt}	0.7
Optimum momentum constant, μ_{opt}	0.1
Activation functions	Sigmoid
Weight Distributions	Uniform
Weights Update	Epoch

Summary

The most suitable model was attained with 15 numbers of hidden units, learning rate of 0.7 and momentum of 0.1. This shows that the problem domain needs a larger number of hidden units to increase the generalization performance. The performance is much better when the value of learning rate and momentum are at optimal values. The results also indicate that uniform weight distribution and epoch weight update give better generalization performance than their counterparts. When comparing with other alternative models, MLP obtained higher generalization than RBF and REG models with 72.78% and 71.56% correctness for both training and testing data respectively.

CHAPTER 7

CONCLUSION AND RECOMMENDATION

The results of the study show that the multilayer perceptron (MLP) model used for predicting the classification for anaemia with backpropagation learning rule can be achieved by applying the optimal value of number of hidden units, learning rate, and momentum to the selected patient. In addition, the model achieved higher generalization compared to Radial Basis Function (RBF) and Regression model when sigmoid activation function was employed.

The highest performance was obtained when the number of hidden units is 15, learning rate is 0.7 and momentum is 0.1. The testing and generalization correctness is 71.56 and 72.78 respectively. This result has demonstrated the ability of multilayer perceptron for predicting classes of anemia and can be used by haematologist and other medical staff.

This study is a retrospective study looking at data collected in 1999. Therefore, there are few limitations and weaknesses. Many of the laboratory forms, particularly in the clinical features section forms were not thoroughly filled by the physician as they were not aware that this form will be used in this study. Hence, these essential data were not being considered. These data however, are very important in improving the accuracy in diagnosing classes of anemia. If this study was done prospectively, the

physician that requested the related investigation can be reminded to fill in the above section beforehand.

For future work, an improvement can be tackled in the following aspects to intensify the usefulness and the generalization of the model. In the pre-processing phase for example, the proper method was not used to remove the unequal distribution of the data set. The data was randomly removed using the ordinary statistical software. The data that has been removed may contain important information. A better method such as hierarchical neural networks (HNNs) as suggested by Ohno-Machado (Ohno-Mahado, 1996a) can provide a way of enhancing the sensitivity to rare categories without decreasing its specificity.

A higher generalization may be accomplished if several layers of hidden units are added. The ability of hidden neurons to extract higher-order statistics is particularly valuable when the size of the input layer is large (Haykin, 1999). As in this study domain, the size of input neuron of seventeen is considerably large. Hence, two layers of hidden units can be employed to increase the generalization performance.

Since training neural networks with backpropagation is time consuming (Ohno-Machadoc *et al.*, 1998), it is desirable that a minimum number of representative cases be kept in the training set in which redundant cases should be removed. However, the removal of redundant cases should be carefully monitored so that classification performance is not affected.

The Neural Connection's software that was used for the training particularly, in the data input tools could be enhanced to increase the sensitivity of the program. Although the software allows the segregation of data sets in a form of percentage or number of cases, it does not take care the distribution of output (target) classes to be equally distributed among the data sets. Consequently, some classes may have a little number of output cases or null to be worst in the data sets that may influence the overall performance.

Providing a user interface is another way to embellish this study. In this way, a user does not need to know the underlying process that may troublesome them. They simply enter a set of input patterns and get the prediction result. In addition to that, the range of output between classes may change in the future. Therefore, one who intent to pursue this study can just make an amendment in the interface program without touching the model.

Recurrent Neural Networks (RNNs) is another alternative models that can be employed in this study to do classification task in which supervised learning can take place. RNN can have the same architecture as standard feed-forward neural networks, except that they allow feedback connections (Haykin, 1999). An alternative for building recurrent neural networks is the use of backpropagation through time (Rumelhart, 1986).

Besides the above suggestions, integrating with other intelligent database techniques such as data mining is another way to expand this classification model. In spite of the fact that neural network is capable to make decision like human, it has a drawback of

explaining why such a result is concluded. Thus utilizing neural network alone can only materialize decision making without decision support. On the other hand, data mining is capable to extract the rules of decision making. Combination of these two techniques would produce an expert system that can be most beneficial for future applications.

As a conclusion, this study has achieved its objectives to produce the multilayer perceptron with backpropagation in predicting the classification for anaemia. This study is hopefully can inspire more works in the field of medical particularly in the haematology studies.

BIBLIOGRAPHY

Alaskerov, E., Freisleben, B. and Rao, B., 1997. CARDWATCH: A Neural Network Based Data Mining System for Credit Card Fraud Detection, *Proceedings of the IEEE/IAFE*, 220-226.

Armoni, A., 1998. Use of neural networks in medical diagnosis, *MD Computing*, Mac-Apr;15(2):100-4

Astion, M.L., and Wilding, P. 1992. Application of neural networks to the interpretation of laboratory data in cancer diagnosis, *Clin Cehm* Jan;38(1):34-8

Bailey, D., and Thompson, D. 1990. Developing Neural Network Applications, *AI Expert*, Vol. 5, No.9, 38-47.

Bazoon, M., A.S. Deborah, and Cui, C. 1994. A Hierarchical Artificial Neural Network System for the Classification of Cervical Cells, *Conference Paper in ICNN'94,* University of Guelph, Ontario, Canada.

Birndorf, N.I., Pentecost, J.O., Coakley, J.R., and Spackman, K.A. 1996. An expert system to diagnose anemia and report results directly on hematology forms, *Comput Biomed Res* Feb;29(1):16-26.

Bortolan, G., *et al.* 1991. "ECG classification with neural networks and cluster analysis." *Proceedings Computers in Cardiology.* Held: Venice, Italy, 23-26 Sept.

Bortolan, G., and Willems, J.L. 1993. Diagnostic ECG classification based on neural networks. *Journal of Electrocardiology*, 1993, 26 Suppl:75–9.

Buscema, M., Mazzeti, di P.M., Salvemini, V., Inraligi, M., and Indrimi, M. 1998. Application of artificial neural networks to eating disorders. *Subst Use Misuse* Feb:33(3):765-91.

Chee, P.L., Hoon, H.T., Thien, S.L., Robert, Harrison, F., and Kennedy, R.L., 1999. Application of and Adaptive Neural Network to Medical Decision Support, URL: http://www/journal.au.edu/ijcim/jan99/ijcim_ar2.html

Chiou, Y.S., and Lure, Y.M. 1994. Hybrid lung nodule detection (HLND) system. *Cancer Letter*s, Mar 15, 77(2-3):119–26.

Dean, F.S. 1999. An artificial neural networks predicts intracranial hemorrhage in preterm neonates better than a logistic regression model. URL: http://www.informatics-review.com/subscribers/Volume_2/num8/neural.html

DeLoughery G.T., 1999, Anemia: An Approach to Diagnosis, Division of Hematology & Medical Oncology, March 15. URL: http://www.ohsu.edu/som-hemonc/handouts/deloughery/anemia.shtml

Dorffner, G., Erich, P., Markus, M., Stefan, K., Paolo, P., Gerold, P., Heinz, S., 1996. Experiences with Neural Networks as a Diagnostic Tool in Medical Image Processing.

Doi, K., Giger, M.L., Nishikawa, R.M., and Schmidt, R.A. 1997. Computer vision and artificial intelligence in mammography: *Proceedings of the International Symposium on"Diagnosis and Therapy of Breast Cancer"*, February, Germany (in press).

Ed Uthman, MD, 1998, Understanding Anemia,University Press of Mississippi URL:http://www.neosoft.com/~uthman/unanemia/unanemia_outline.html

Egmont-Petersen, M., Schreiner, U., Tromp, S.C., Lehmann, T.M., Slaaf, D.W. 2000. Detection of leukocytes in contact with the vessel wall from in vivo microscope recordings using a neural network, Arts T Department of

Biophysics, Maastricht University, The Netherlands. *IEEE Trans Biomed Eng,* Jul;47(7):941-951

Ellenius, J., Groth, T., and Lindahl, B. 1997. Neural network analysis of biochemical markers for early assessment of acute myocardial infarction, *Stud Health Technology Inform*;43 Pt A:382-5

Erler, B.S., Vitagliano, P., and Lee, S. 1995. Superiority of neural networks over discriminant functions for thalassemia minor screening of blood cell microcytosis, ArchPathol Lab Med, Apr; 119(4):350-4.

Fausett, Laurence V., *Fundamentals of Neural Networks: Architecture, Algorithms, and Applications*, New Jersey: New Jersey: Prentice Hall PTR, 1994

Gallagher, M.R., 1999. Multi-layer Perceptron Error Surfaces: Visualization, Structure and Modelling, Department of Computer Science and Electrical, PhD Thesis, Engineering University of Queensland, St Lucia 4072, Australia, June 30 (Revised January, 2000)

Giacomini, M., Ruggiero, C., Bertone, S., and Calegari, L., 1997. Artificial neural network identification of heterotrophic marine bacteria based on their fatty-acid composition. *IEEE Trans Biomed Eng,* Dec;44(12):1185-91

Haykin, S., 1999. *Neural Networks: A Comprehensive Foundation*, Intl Edition 2nd Edition, Prentice Hall Intl. Inc.

Hoffbrand, A.V., and Pettit, J.E., 1993, Essential Hematology, 3rd Edition, Blackwell Scientific Publications.

Huo Z, and Giger ML. 1996. Integrating rules and artificial neural networks in the classification of mass lesions in digital mammograms. *Proc. World Congress on Neural Networks '96*, Vol. 1, 1166-1169

Hussain, A., Abdul Samad, S., and Soon, K.T. 1999. Theory, Methodology and Implementation of The Malay Text-To-Speech System, *MJCS*, Vol.12, No. 1.4.

Ishida, T., Katsuragawa, S., Ashizawa, K., MacMahon, H., and Doi, K. 1997. Application of artificial neural networks for quantitative analysis of image data in chest radiographs for detection of interstitial lung disease. *Med Phys* (submitted).

Isselbacher, Braunwald, Wilson, Martin, Fauci, Kasper. 1994, Harrison's Principle of Internal Medicine, Mc-Graw Hill, 13th Edition: Vol1, 313

Kamruzzaman, J., and Aziz, S.M., 1998. A Neural Network Character Recognition System Using Double Backpropagation, *MJSC* Vol.11, No.1.7.

Karim, M.E., Hossain, M., Mottalib, M.A. , and Zaman, T. , 1998. A modified Neural Network Learning Approach and Its Application to Bengali Character Recognition, *MJCS* Volume 11, No. 2.9.

Katsuragawa, S., Doi, K., MacMahon, H., Monnier-Cholley, L., Ishida, T., and Kobayashi, T. 1997. Classification of normal and abnormal lungs with interstitial diseases by rule-based method and artificial neural networks. *J of Digital Imaging 19*: 108-114.

Krose, Ben J. A., and van der Smagt., P.P. An Introduction to Neural Networks. University of Amsterdam, 1991

Lapuerta, P., Azen, S.P., and La, B. 1995. Use of Neural Networks in predicting the risk of coronary artery disease, *Comput Biomed Res*, Feb;28(!):38-52

Lapuerta, P., Rajan, S., and Bonacini, M. 1997. Neural networks as predictors of outcomes in alcoholic patients with severe liver disease, Hepatology. Feb;25(2):302-6

Lawrence, J., 1991. Data Preparation for a Neural Network. *AI Expert*, Vol. 6, No.11, 34-41.

LeCun, Y., 1993. *Efficient Learning and Second-order Methods, A Tutorial at NIPS 93*, Denver.

LeCun, Y., Boser, B., Denker, J.S., Henderson, D., Howard, R.E. Hubbard, W., and Jackerl, L.D. 1990, Handwritten digit recognition with a backpropagation network. Advances in Neural Information Processing Systems, 2, 248-257.

Lim, S.H., Liang, W.T., and Subramanian, R.K. 1996. Neural Network Based Share Price Prediction System, *ATIP96.067 : Malaysian Computer Science & Applications (REDECS '96)*, Universiti Putra Malaysia, 26-27 June.

Lin, Z., Maris, J., Hermans, L., Vandewalle, J., and Chen, J.D. 1997. Classification of normal and abnormal electrogastrograms using multilayer feedfordward neural networks, *Med Biol Eng Comput,* May; 35(3):199-206

Lowe, D., Webb, A.R. 1990. Exploiting prior knowledge in network optimization: An illustration from medical prognosis. *Network*, 1:299-323

Luiken, G., Phares, J.C., and Millard, F. 1999. General Medical Officer (GMO) Manual: Clinical Section: Hematology. URL: http://www.vnh.org/GMO/ClinicalSection/26Aenmia.html

Makeigh, S., Jung, Tzyy-Ping., and Sejnowski, T.J. 1996. Using Feedforward Neural Networks to Monitor Alertness from Changes in EEG Correlation and Coherence, Advances in Neural Information Processing Systems 8, MIT Press, Cambridge MA, 931-937.

Mat Shaari, A.R., Ramli, A.R., Mahmod, R. and Bidin, A.R. 1996. Automatic Recognition of Malaysian Car Registration Numbers, *ATIP96.067 :*

Malaysian Computer Science & Applications (REDECS '96), Universiti Putra Malaysia, 26-27 June.

McCulloch, W.S., and Pitts, W., 1943. A logical calculus of ideas immanent in nervous activity, *Bulletin of Mathematical Biopyhsics,* Vol. 5, 115 –133.

Mehdi, B., Deborah, A.S., and Chen, C., 1994. Computing and Information Science, H George, Genetics and Molecular Biology, University of Guelph, Guelph, Ontarion, CANADA NIG 2W1, *Conference Paper in ICNN '94* (Orlando, Florida, USA).

Minsky M; Papert S. 1969. Perceptrons; An Introduction to Computational Geometry. MA:MIT Press, Cambridge

Mitchell, T.M., 1997. *Machine Learning*, Boston: WCB/McGraw Hill.

Ohno-Machado, L., 1996a. Medical Applications of Artificial Neural Network: Connectionist Model of Survival, Dissertation, March.

Ohno-Machado, L., 1996b. Sequential use of neural networks for survival prediction in AIDs, *Proc AMIA Fall Symp*, 170-4.

Ohno-Machado, L., Fraser, H.L., Ohrn, A. 1998. Improving machine learning performance by removing redundant cases in medical data sets. *Proc AMIA Symp* 523-7.

Orkin, 1992. *Harrison's Principle of Internal Medicine*, Mc-Graw Hill, 13th Edition: Vol1, 313

Ning L.W, Khalid M, and Yusof R. 1999. Design of an Automated Data Entry System for Handwritten Forms.

Partridge, D., Abidi, S.S.R., and Alwyn, G., 1996. Neural Networks Applications in Medicine, ATIP96.067 : Malaysian Computer Science & Applications (REDECS '96), Universiti Putra Malaysia, 26-27 June.

Ramli, A.R., and Kumar S.V., 1996. Pattern Recognition for IC Mark Inspection, ATIP96.067 : Malaysian Computer Science & Applications (REDECS '96), Universiti Putra Malaysia, 26-27 June.

Rapaport, S.I., 1987. Introduction to Hematology, 2nd Edition, J.B. Lippincott Company, Philadelphia,10.

Reggia, J.A., 1993. Neural Computation in medicine. *Artificial Intelligence in Medicine*, Apr, 5(2):143-57.

Rietveld, S., Oud,M., and Dooijes, E.H. 1999. Classification of asthmatic breath sounds: preliminary results of the classifying capacity of human examiners versus artificial neural networks. Computers and biomedical research, Faculty of Psychology, University of Amsterdam, The Netherlands, *Comput-Biomed Res,* Oct, Vol. 32(5) 440-8.

Ritschel, W.A.., Ajukeswaran, R., and Hussain, A.S. 1995. Application of neural networks for the prediction of human pharmacokinetic parameters, *Methods Find Exp Clin* Pharmacol Nov; 17(9):629-43.

Rosenblatt, F., 1958. The Perceptron: A probabilistic model for information storage and organization in the brain, *Physiological Review*, Vol. 65, 386-408.

Rubin, R.N., 1997. Hematologic Disease: *Medicine*, 3rd Edition, Williams & Wilkins, 109.

Rumelhart, D.E., Hinton, G.E., and Williams, R.J., 1986. Learning internal representations by error propagation in Rumelhart, D.E., and McClelland, J.L. eds., Vol. 1, Chapter 8, Cambridge, MA:MIT Press.

Russell, I.F., Neural Networks. Journal of Undergraduate Mathematics and its Applications, vol. 14, No 1, URL http://uhavax.hartford.edu/disk$userdata/faculty/compsci/www/neural-networks-tutorial.html

Sanugi, B., and San, W.S., 1996. Neural Network Approach In Predicting KLSE Composite Index, LT/M.Bil. 7.

Shanker, M.S., 1996. Using Neural Networks to Predict the Onset of Diabetes Mellitus, Department of Administrative Sciences, Kent State University, Kent, OH 44242, In *Journal of Chemical Information and Computer Sciences*, 36

Siraj, F., Zakaria, A., Yassin, A., Wan Ishak, W.H, 1999. Neural Network Approach On-Online Handwritten Signature Verification System.

Stone, M., 1978. "Cross-validatory choice and assessment of statistical predictions, " Journal of the Royal Statistical Society, Vol. B36, 111-133.

Tay, Y.H., 1997. Handwritten Character Recognition by Fuzzy Artmap Neural Network with Application to Postcode Recognition, Thesis of University of Technology Malaysia.

Tay, Y.H., and Khalid, M. 1997. Comparison of Fuzzy ARTMAP And MLP Neural Networks For Hand-written Character Recognition, IFAC Symposium on AI in Real-Time Control 1997 (AIRTC'97), Kuala Lumpur, Malaysia, 363-371.

Thimm, G., Moerland, P., and Fiesler, E., 1996. The Interchangeability of Learning Rate and Gain in Backpropagation Neural Network IDIP, CII-1920 Martigny, Switzerland, Appeared in *Neural Computation* 8(2), Feb.

Toronen, P., Kolehmainen, M., Wong, G., and Castren, E. 1999. Analysis of gene expression data using self-organizing maps. *FEBS Lett*, May 21;451(2):142-6

Tsoukalas, L. H., and Uhrig, R. E., 1997. *Fuzzy and Neural Approaches in Engineering*, John Wiley & Sons, Inc.

Ulbricht, C., and Dorffner, G. 1996. Forecasting Fetal Heartbeats with Neural Networks, ext. version, Solving Engineering Problems with Neural Networks, 403-406.

Yaakob, R., Sulaiman, M.N., Mahmood, R., Mohamed, M.T.M., and Ramli, A.R, 1999. Multi-Counterpropagation Network Model For Colour Recognition, *MJCS,* Vol. 12:1.5.

Wilding, P., Morgan, M.A., Grygotis, A.e. Shoffner, M.A., and Rosato, E.F. 1994. Diagnosis of backpropagation neural networks to diagnosis of breast and ovarian cancer, *Cancer Lett Mar* 15:77(2-3):145-53

West, D. and West, V. 2000. Model selection for a medical diagnostic decision support system: A breast cancer detection case. *Artif Intell Med,* Nov 1;20(3):183-204

Wu, Y., Doi, K., and Giger, M.L. 1995. Detection of lung nodules in digital chest radiographs using artificial neural networks: A pilot study. J Digital Imaging 8: 88-94

Wu, Y., Giger, M.L., Doi, K., Vyborny, C.J., Schmidt, R.A., and Metx, C.E. 1993. Artificial Neural Netwrok in mammography: Application to decision making in the diagnosis of breast cancer. Radiology, 187 (1):81-87

Zernikow, B., Holtmannspoetter, K., Michel, E., Theilhaber, M., Pielemeier, W., and Hennecke, K.H. 1998. "Artificial neural network for predicting intracranial haemorrhage in preterm neonates". *Acta Paediatric* Sep;87(9):969-75

APPENDIX A

HAEMATOLOGY FORM

APPENDIX A : HAEMATOLOGY FORM

PER-TAT 301

KEMENTERIAN KESIHATAN MALAYSIA
PERKHIDMATAN PATOLOGI

HOSPITAL ALOR SETAR

1184/99

UNTUK KEGUNAAN MAKMAL

LAB NO

1. Nama : [redacted]		2. No. Pendaftaran :	
3. No. K/P : [redacted]		4. Jantina : ☒ Lelaki	☐ Perempuan
5. Umur : 6 thn	6. Keturunan : Melayu	7. Wad / Klinik :	
8. Tarikh Masuk Wad : 15/8/99	9. Pekerjaan :	10. Taraf Perkahwinan :	11. ☐ Bayar ☐ Percuma

12. No. Laporan Dahulu :

13. Butiran Penting

	Ya	Tidak
Jaundice	☐	☐
Lymphadenopathy	☐	☐
Hepatomegaly	☐	☐
Splenomegaly	☐	☐
Bleeding Tendency	☐	☐
H/O Transfusion	☐	☐

14. Ringkasan Klinikal, Penemuan Pembedahan dan Riwayat Keluarga :

6/12/99
Renal calculi & renal impairment
Hb. 6.2
° Active bleeding ° Melaena
TRO Septic Cause
& Anemia

Haematinics :

Drug / Chemical History

Data Makmal Terdahulu

Hb

Platelet

TWDC

15. Diagnosis : Renal calculi & renal impairment

16. Kategori Permohonan/Jenis Ujian :

Patologi Kimia	Klinikal	Hematologi	Histo/Saitologi	Mikro/Immunologi
B.Sugar ☐	Bld. Count ☐	PBP ☐	Specimen	Specimen / Ujian
B.Urea ☐	ESR ☐	BM Asp. ☐		
S.Elec ☐	BFMP ☐	Hb Analysis ☐		
B.Gases ☐	U.Sugar ☐	Coagulation ☐		
S.Billirubin ☐	U.Alb. ☐			
LFT ☐	U.ME ☐			
Se.Creatinine ☐	Stool ME ☐			

Lain-Lain FBP

17. Pengambilan Specimen : Tarikh : 16/12/99 Masa

18. Nama Doktor :

19. Tarikh :

PEGAWAI PERUBATAN
Y/M WAD X2
HOSPITAL ALOR SETAR

Tandatangan dan Cop Doktor

Patologi/95

LAPORAN : SILA LIHAT SEBELAH

KEMENTERIAN KESIHATAN MALAYSIA
PERKHIDMATAN PATOLOGI
HOSPITAL A/S

1179/99 PER-PAT 301

UNTUK KEGUNAAN MAKMAL
LAB NO

1. Nama : [redacted]			2. No. Pendaftaran :
3. No. K/P : [redacted]			4. Jantina : ☑ Lelaki ☐ Perempuan
5. Umur : 63	6. Keturunan :	7. Wad / Klinik :	X/- 24
8. Tarikh Masuk Wad :	9. Pekerjaan :	10. Taraf Perkahwinan :	11. ☐ Bayar ☐ Percuma

12. No. Laporan Dahulu :

13. Butiran Penting

	Ya	Tidak
Jaundice	☐	☐
Lymphadenopathy	☐	☐
Hepatomegaly	☐	☐
Splenomegaly	☐	☐
Bleeding Tendency	☐	☐
H/O Transfusion	☐	☐

14. Ringkasan Klinikal, Penemuan Pembedahan dan Riwayat Keluarga :

c/o back pain (vertebral ego)
since 4/ [illegible] 98
since 3/12 ago increasing
unable to walk Fd/t pain

X-ray: lytic lesion pelvic
- both head of humerus

uric acid 99
Sr Ca - 2.68

Haematinics

Drug / Chemical History :

Data Makmal Terdahulu

Hb :

Platelet

TWDC

Diagnosis : ? Multiple Myeloma

16. Kategori Permohonan/Jenis Ujian :

Patologi Kimia		Klinikal		Hematologi		Histo/Saitologi	Mikro/Immunologi	
						Specimen	Specimen	Ujian
B.Sugar	☐	Bld. Count	☐	PBP	☐			
B.Urea	☐	ESR	☐	BM Asp.	☐			
S.Elec	☐	BFMP	☐	Hb Analysis	☐			
B.Gases	☐	U.Sugar	☐	Coagulation	☐			
S.Billirubin	☐	U.Alb.	☐					
LFT	☐	U.ME	☐					
Se.Creatinine	☐	Stool ME	☐					

Lain-Lain

17. Pengambilan Specimen : Tarikh : 15/8/99 Masa

18. Nama Doktor : Dr. Sythan Lah

19. Tarikh :

Pegawai Perubatan
Hospital Alor Setar
Tandatangan dan Cop Doktor

Patologi/95

LAPORAN "SILA LIHAT SEBELAH"

L—PNMB_AS

KEMENTERIAN KESIHATAN MALAYSIA
PERKHIDMATAN PATOLOGI

HOSPITAL

PER-PAT 301

UNTUK KEGUNAAN MAKMAL

LAB NO :

1. Nama : [redacted]
2. No. Pendaftaran :
3. No. K/P : [redacted]
4. Jantina : ☐ Lelaki ☐ Perempuan
5. Umur :
6. Keturunan :
7. Wad / Klinik :
8. Tarikh Masuk Wad : 17/8/99
9. Pekerjaan :
10. Taraf Perkahwinan :
11. ☐ Bayar ☐ Percuma
12. No. Laporan Dahulu :
13. Butiran Penting

	Ya	Tidak
Jaundice	☐	☐
Lymphadenopathy	☐	☐
Hepatomegaly	☐	☐
Splenomegaly	☐	☐
Bleeding Tendency	☐	☐
H/O Transfusion	☐	☐

Haematinics

Drug / Chemical History

Data Makmal Terdahulu

Hb :

Platelet :

TWDC :

14. Ringkasan Klinikal, Penemuan Pembedahan dan Riwayat Keluarga :

G1 P0 A0 @ 34/52
Hb 4.9
Asymptomatic
transfused 4U PC

15. Diagnosis : Severe Anemia

16. Kategori Permohonan/Jenis Ujian :

Patologi Kimia		Klinikal		Hematologi		Histo/Saitologi	Mikro/Immunologi	
						Specimen	Specimen	Ujian
B.Sugar	☐	Bld. Count	☐	PBP	☐			
B.Urea	☐	ESR	☐	BM Asp.	☐			
S.Elec	☐	BFMP	☐	Hb Analysis	☐			
B.Gases	☐	U.Sugar	☐	Coagulation	☐			
S.Billirubin	☐	U.Alb.	☐					
LFT	☐	U.ME	☐					
Se.Creatinine	☐	Stool ME	☐					

Lain-Lain : FBP

18 AUG 1999

17. Pengambilan Specimen : Tarikh : 17/8/99 Masa : 11 pm

18. Nama Doktor : TEO

19. Tarikh :

PEGAWAI PERUBATAN
YANG BERTANGGUNGJAWAB
HOSPITAL ALOR SETAR
Tandatangan dan Cop Doktor

81

Patologi/95

LAPORAN "SILA LIHAT SEBELAH"

L—PNMB., AS.

KEMENTERIAN KESIHATAN MALAYSIA
PERKHIDMATAN PATOLOGI

HOSPITAL ALOR SETAR

1184/99

PER-PAT 301

APPENDIX A : HAEMATOLOGY FORM

UNTUK KEGUNAAN MAKMAL

LAB NO

1. Nama : [redacted]		2. No. Pendaftaran :
3. No. K/P : [redacted]		4. Jantina : ☒ Lelaki ☐ Perempuan
5. Umur : 61 thn	6. Keturunan : Melayu	7. Wad / Klinik : V2 / 27
8. Tarikh Masuk Wad : 15/8/99	9. Pekerjaan :	10. Taraf Perkahwinan : 11. ☐ Bayar ☐ Percuma

12. No. Laporan Dahulu :

13. Butiran Penting

	Ya	Tidak
Jaundice	☐	☐
Lymphadenopathy	☐	☐
Hepatomegaly	☐	☐
Splenomegaly	☐	☐
Bleeding Tendency	☐	☐
H/O Transfusion	☐	☐

14. Ringkasan Klinikal, Penemuan Pembedahan dan Riwayat Keluarga :

61/M/Malay

Renal calculi & renal impairment

Hb. 6.2

- Active bleeding
- Melaena

TRO [redacted]/Suspective Cause

- Anemia

Haematinics :

Drug / Chemical History

Data Makmal Terdahulu

Hb

Platelet

TWDC

15. Diagnosis : Renal calculi & renal impairment

16. Kategori Permohonan/Jenis Ujian:

Patologi Kimia		Klinikal		Hematologi		Histo/Saitologi		Mikro/Immunologi	
						Specimen		Specimen	Ujian
B.Sugar	☐	Bld. Count	☐	PBP	☐				
B.Urea	☐	ESR	☐	BM Asp.	☐				
S.Elec	☐	BFMP	☐	Hb Analysis	☐				
B.Gases	☐	U.Sugar	☐	Coagulation	☐				
S.Billirubin	☐	U.Alb.	☐						
LFT	☐	U.ME	☐						
Se.Creatinine	☐	Stool ME	☐						

Lain-Lain FBP

17. Pengambilan Specimen : Tarikh : 16/08/99 Masa

18. Nama Doktor :

19. Tarikh :

PEGAWAI PERUBATAN
Y/M WAD X2
HOSPITAL ALOR SETAR

Tandatangan dan Cop Doktor

Patologi/95

LAPORAN "SILA LIHAT SEBELAH"

APPENDIX B

QUESTIONNAIRE

Age: _____ Sex: M ☐ F ☐

Objective: To determine factors that influence classification of anemia.

(A) FULL BLOOD COUNT (FBC)

1.	**WBC**		X10^3/uL
2.	**RBC**		X10^6/uL
3.	**HGB**		g/dL
4.	**HCT**		%
5.	**PLT**		X10^3/uL
6.	**Reticulocyte**		%
Differential Count			
7.	**Eosinophil**		%

(B) FULL BLOOD PICTURE (FBP) – Peripheral Blood film

8) Red Blood Cell (RBC)		**9) Shapes / characteristics**	
(Color, size)		1. Elliptocyte / Pencil cells	
1. Hypochromic, microcytic		2. Ovalocytes	
2. Hypochromic, normocytic		3. Tear drop cells	
3. Normochromic, normocytic		4. Schistocytes / Fragmented cells	
4. Normochromic, microcytic		5. Polychromasia	
5. Normochromic, macrocytic		6. Sperocytes	
		7. Targets	
		8. Burr cell	
		9. Acanthocyte	
		10. Sickle cell	
		11. Rouleoux	
		12. Aggregates	
		13. Neclueted (NRBC)	
		14. Bite cells	
		15. Stomatocytes	
		16. Basophilic stippling	

10) White Count (WC)		**11) Platelets**	
1. Normal		1. Normal	
2. Leucocytosis		2. Increased	
3. Left shift		3. Reduced	
4. Plasma cell			
5. Leucopenia			
6. Monocytosis			
7. Toxic granulation			
8. Hypersegmented neutrophilia			
9. Vacuolation			

(C) CLASSIFICATION (Target)

1. Normal
2. Microcytic, hypochromic
3. Normocytic, normochromic
4. Macrocytic
5. Iron deficiency
6. Thalassaemia
7. Anemia of chronic disease
8. Lead poisoning
9. Sideroblastic anaemia
10. Haemolytic anaemia
11. Secondary anaemia
12. After acute blood loss
13. Bone-marrow failure
14. Vitamin B12 deficiency / Folate deficiency
15. Alcohol
16. Liver diseases
17. Infection

APPENDIX C

EXPERIMENT RESULTS

TABLE 1: Neural Networks Results with different size of Hidden Units (HU)

HU	O1	O2	O3	O4	O5	O6	O7	O8	AVG
8	54.69	82.50	61.56	35.63	43.75	71.56	75.94	63.44	61.13
9	61.88	82.81	64.06	40.31	46.88	72.50	75.63	62.81	63.36
10	55.63	84.06	62.19	38.13	47.50	72.81	74.38	63.75	62.31
11	59.69	82.50	62.81	37.19	49.06	72.19	74.38	64.38	62.78
12	57.50	83.13	63.44	35.31	49.38	72.50	75.00	65.94	62.78
13	56.88	81.88	63.75	35.94	49.38	71.56	75.63	66.26	62.66
14	50.31	85.63	60.94	31.56	42.81	72.19	75.63	65.63	60.59
15	**74.06**	**81.56**	**64.38**	**43.13**	**52.19**	**79.38**	**79.38**	**68.13**	**67.78**
16	56.25	81.88	64.69	37.50	48.44	71.25	77.19	67.50	63.09
17	**68.75**	**80.00**	**66.56**	**44.06**	**52.19**	**78.13**	**77.50**	**64.69**	**66.49**

(a) Training Results

HU	O1	O2	O3	O4	O5	O6	O7	O8	AVG
8	60.00	80.00	47.50	30.00	47.50	80.00	70.00	65.00	60.00
9	60.00	77.50	50.00	45.00	55.00	82.50	70.00	65.00	63.13
10	55.00	80.00	50.00	40.00	47.50	82.50	70.00	60.00	60.63
11	60.00	82.50	50.00	30.00	60.00	80.00	70.00	65.00	62.19
12	57.50	77.50	50.00	42.50	57.50	80.00	70.00	67.50	62.81
13	60.00	82.50	52.50	47.50	60.00	82.50	70.00	65.00	65.00
14	60.00	82.50	45.00	25.00	45.00	80.00	70.00	65.00	59.06
15	**72.50**	**77.50**	**62.50**	**50.00**	**57.50**	**82.50**	**77.50**	**62.50**	**67.81**
16	57.50	77.50	50.00	37.50	50.00	77.50	70.00	65.00	60.63
17	**62.50**	**80.00**	**57.50**	**52.50**	**62.50**	**82.50**	**70.00**	**67.50**	**66.88**

(b) Testing Results

TABLE 2: Neural Networks results with different size of Weight Seeds

WS	O1	O2	O3	O4	O5	O6	O7	O8	AVG
1	74.06	81.56	64.38	43.13	52.19	79.38	79.38	68.13	67.78
2	60.94	81.25	64.38	43.13	51.56	78.75	77.50	69.06	65.82
3	61.56	81.88	63.13	39.06	50.00	73.44	77.81	65.63	64.06
4	60.31	82.50	62.81	40.00	50.00	73.75	76.25	66.88	64.06
5	60.94	81.56	60.94	40.94	53.75	76.56	76.25	67.50	64.81
6	54.38	83.13	63.13	33.75	48.44	72.19	74.69	65.00	61.84
7	60.31	80.63	65.00	44.06	51.56	74.69	76.56	66.25	64.88
8	56.56	82.81	62.81	37.19	47.81	71.88	75.63	65.31	62.50
9	57.50	80.94	61.88	38.75	50.94	72.50	75.94	65.94	63.05
10	58.75	82.19	65.63	38.13	48.44	70.94	75.00	66.56	63.21
									64.20

2(ai) Training Results (Hidden Unit = 15)

WS	O1	O2	O3	O4	O5	O6	O7	O8	AVG
1	72.50	77.50	62.50	50.00	57.50	82.50	77.50	62.50	67.81
2	57.50	82.50	52.50	47.50	57.50	80.00	70.00	65.00	64.06
3	57.50	77.50	60.00	42.50	57.50	80.00	70.00	67.50	64.06
4	55.00	77.50	57.50	40.00	57.50	85.00	70.00	65.00	63.44
5	55.00	75.00	57.50	52.50	62.50	80.00	75.00	60.00	64.69
6	60.00	77.50	50.00	30.00	52.50	82.50	67.50	67.50	60.94
7	60.00	77.50	57.50	57.50	55.00	82.50	75.00	65.00	66.25
8	57.50	77.50	52.50	32.50	52.50	82.50	70.00	60.00	60.63
9	60.00	77.50	55.00	37.50	60.00	80.00	70.00	60.00	62.50
10	52.50	82.50	55.00	42.50	60.00	85.00	72.50	67.50	64.69
									63.91

2(aii) Testing Results (Hidden Unit = 15)

WS	O1	O2	O3	O4	O5	O6	O7	O8	AVG
1	66.56	78.13	64.69	44.38	54.06	76.88	75.31	69.69	66.21
2	61.25	80.31	61.56	43.75	54.38	77.19	78.44	67.19	67.19
3	59.06	80.94	65.94	38.13	49.38	73.44	76.25	66.56	63.71
4	70.63	76.88	64.69	48.13	55.63	82.19	78.44	70.94	68.44
5	56.88	82.19	65.63	39.38	50.94	72.19	76.88	66.88	63.87
6	62.19	80.63	63.44	41.88	52.81	79.38	77.81	69.06	65.90
7	54.69	82.81	63.44	35.00	48.75	72.50	76.25	65.94	62.42
8	55.31	80.94	65.63	35.63	48.44	71.88	76.25	65.63	62.46
9	63.13	80.63	64.38	43.75	52.19	80.63	79.06	67.19	66.37
10	57.50	82.81	63.44	39.69	51.88	74.06	76.88	67.81	64.26
									65.08

2(bi) Testing Results (Hidden Unit = 17)

WS	O1	O2	O3	O4	O5	O6	O7	O8	AVG
1	57.50	85.00	65.00	50.00	62.50	82.50	75.00	70.00	68.44
2	57.50	75.00	55.00	47.50	55.00	80.00	75.00	67.50	64.06
3	57.50	80.00	57.50	40.00	57.50	80.00	70.00	62.50	63.13
4	70.00	72.50	57.50	55.00	62.50	85.00	75.00	65.00	67.81
5	60.00	80.00	47.50	40.00	62.50	80.00	70.00	67.50	63.44
6	57.50	82.50	57.50	45.00	67.50	85.00	67.50	62.50	65.63
7	60.00	80.00	55.00	30.00	60.00	77.50	70.00	60.00	61.56
8	57.50	75.00	52.50	30.00	52.50	82.50	70.00	60.00	60.00
9	55.00	80.00	60.00	50.00	65.00	80.00	72.50	65.00	65.94
10	57.50	80.00	52.50	40.00	60.00	85.00	70.00	67.50	64.06
									64.41

2(bii) Testing Results (Hidden Unit = 17)

TABLE 3: Neural Networks Results with different size of Learning Rate (LR)

LR	O1	O2	O3	O4	O5	O6	O7	O8	AVG
0.1	74.06	81.56	64.38	43.13	52.19	79.38	79.38	68.13	67.78
0.2	74.06	81.56	64.38	43.13	52.50	79.38	79.69	68.13	67.85
0.3	74.06	81.56	64.38	43.13	52.50	79.38	79.38	68.13	67.82
0.4	74.06	81.56	64.69	43.13	52.19	79.38	79.69	68.44	67.89
0.5	77.19	80.94	66.88	43.75	52.81	78.13	78.75	67.81	68.28
0.6	82.19	80.94	68.75	46.56	57.81	81.56	81.88	73.75	71.68
0.7	**84.38**	**80.00**	**70.31**	**44.38**	**60.94**	**82.81**	**83.13**	**76.25**	**72.78**
0.8	80.63	83.44	60.00	43.13	59.38	85.00	74.38	69.38	69.42
0.9	**88.13**	**87.50**	**67.50**	**39.38**	**48.13**	**86.88**	**79.69**	**75.63**	**71.61**
1	88.13	84.69	67.81	41.25	45.63	83.13	84.06	82.19	72.11

3(a) Training Results

LR	O1	O2	O3	O4	O5	O6	O7	O8	AVG
0.1	72.50	77.50	62.50	50.00	57.50	82.50	77.50	62.50	67.81
0.2	72.50	77.50	62.50	50.00	57.50	82.50	77.50	60.00	67.50
0.3	72.50	77.50	62.50	50.00	57.50	80.00	77.50	60.00	67.19
0.4	72.50	77.50	62.50	50.00	57.50	80.00	77.50	60.00	67.19
0.5	72.50	77.50	62.50	52.50	57.50	77.50	75.00	60.00	66.88
0.6	60.00	85.00	55.00	27.50	52.50	80.00	70.00	62.50	61.56
0.7	**72.50**	**75.00**	**67.50**	**42.50**	**77.50**	**90.00**	**80.00**	**67.50**	**71.56**
0.8	70.00	85.00	62.50	37.50	70.00	87.50	67.50	67.50	68.44
0.9	**75.00**	**82.50**	**67.50**	**47.50**	**62.50**	**90.00**	**75.00**	**70.00**	**71.25**
1	75.00	85.00	62.50	45.00	57.50	85.00	77.50	75.00	70.31

3(b) Testing Results

TABLE 4: Neural Networks results with different size of momentum (MV)

MV	O1	O2	O3	O4	O5	O6	O7	O8	AVG
0.1	**74.06**	**81.56**	**64.38**	**43.13**	**52.19**	**79.38**	**79.38**	**68.13**	**67.78**
0.2	**74.06**	**81.56**	**64.38**	**43.13**	**52.50**	**79.38**	**79.38**	**68.13**	**67.82**
0.3	74.06	81.56	64.69	43.13	52.50	79.38	79.69	68.13	67.89
0.4	74.06	81.56	64.38	43.13	52.50	79.38	79.38	68.13	67.82
0.5	74.06	81.56	64.69	43.13	52.19	79.38	79.69	68.44	67.89
0.6	74.06	81.56	65.00	42.81	52.19	79.06	79.06	68.13	67.73
0.7	74.38	81.25	65.31	42.81	52.81	79.38	79.38	67.50	67.85
0.8	75.63	81.25	65.94	42.81	53.13	79.69	80.00	68.13	68.32
0.9	76.88	81.56	65.31	43.13	52.50	79.06	80.31	67.19	68.24
1	70.00	80.00	62.50	50.00	57.50	80.00	75.00	60.00	66.88

4(a) Training Results

MV	O1	O2	O3	O4	O5	O6	O7	O8	AVG
0.1	**72.50**	**77.50**	**62.50**	**50.00**	**57.50**	**82.50**	**77.50**	**62.50**	**67.81**
0.2	**72.50**	**77.50**	**62.50**	**50.00**	**57.50**	**82.50**	**77.50**	**62.50**	**67.81**
0.3	72.50	77.50	62.50	50.00	57.50	82.50	77.50	60.00	67.50
0.4	72.50	77.50	62.50	50.00	57.50	80.00	77.50	60.00	67.19
0.5	72.50	77.50	62.50	50.00	57.50	80.00	77.50	60.00	67.19
0.6	72.50	77.50	62.50	50.00	57.50	80.00	77.50	60.00	67.19
0.7	72.50	77.50	62.50	50.00	57.50	80.00	77.50	60.00	67.19
0.8	72.50	77.50	62.50	50.00	60.00	80.00	77.50	60.00	67.50
0.9	72.50	80.00	60.00	50.00	57.50	80.00	77.50	62.50	67.50
1	70.00	80.00	62.50	50.00	57.50	80.00	75.00	60.00	66.88

4(b) Testing Results

TABLE 5: Neural Networks Results with different Weight Distributions

	O1	O2	O3	O4	O5	O6	O7	O8	AVG
Uniform	**84.38**	**80.00**	**70.31**	**44.38**	**60.94**	**82.81**	**83.13**	**76.25**	**72.78**
Gaussian	77.81	80.63	66.25	45.00	61.88	83.44	79.69	75.63	71.29

(a) Training Results

	O1	O2	O3	O4	O5	O6	O7	O8	AVG
Uniform	**72.50**	**75.00**	**67.50**	**42.50**	**77.50**	**90.00**	**80.00**	**67.50**	**71.56**
Gaussian	70.00	80.00	62.50	42.50	70.00	82.50	72.50	67.50	68.44

(b) Testing Results

TABLE 6: Neural Networks Results with different Weight Updates

Epoch	**84.38**	**80.00**	**70.31**	**44.38**	**60.94**	**82.81**	**83.13**	**76.25**	**72.78**
Pattern	82.81	74.06	67.19	45.31	60.31	83.44	81.88	74.69	71.21

(a) Training Results

Epoch	**72.50**	**75.00**	**67.50**	**42.50**	**77.50**	**90.00**	**80.00**	**67.50**	**71.56**
Pattern	70.00	72.50	65.00	47.50	70.00	87.50	75.00	62.50	68.75

(b) Testing Results

TABLE 7: Neural Networks Results with different Activation Functions

Sigma	**84.38**	**80.00**	**70.31**	**44.38**	**60.94**	**82.81**	**83.13**	**76.25**	**72.78**
Tan	83.13	76.56	62.81	40.31	51.25	75.94	82.19	61.88	66.76
Linear									0.00

(a) Training Results

Sigma	**72.50**	**75.00**	**67.50**	**42.50**	**77.50**	**90.00**	**80.00**	**67.50**	**71.56**
Tan	72.50	80.00	60.00	50.00	60.00	82.50	77.50	62.50	68.13
Linear									0.00

(b) Testing Results

TABLE 8: Neural Networks Results with different models

MLP	**84.38**	**80.00**	**70.31**	**44.38**	**60.94**	**82.81**	**83.13**	**76.25**	**72.78**
RBF	68.44	80.31	66.88	42.50	48.75	79.69	72.19	63.75	65.31
REG	49.38	81.56	61.88	31.56	45.94	70.94	76.25	67.81	60.67

(a) Training Results

MLP	**72.50**	**75.00**	**67.50**	**42.50**	**77.50**	**90.00**	**80.00**	**67.50**	**71.56**
RBF	60.00	80.00	60.00	47.50	55.00	85.00	60.00	62.50	63.75
REG	57.50	77.50	52.50	32.50	50.00	80.00	70.00	67.50	60.94

(b) Testing Results

APPENDIX D

DATA SOURCES

Citation Request :

If you publish results when using this database, then please include the following information in your acknowledgements.

1) Title : Haematology Databases (1999)

2) Sources : Haematology Unit, Hospital Alor Setar, Kedah

Dr. Abdul Rashid Mohd Ibrahim
Haematologist
Unit Bank Darah
Hospital Besar Alor Setar,
Kedah, Malaysia

3) Collected by : Shuzlina Abdul Rahman

Master in Information Technology; 1999/2000
Date of Collection : July - August, 2000

4) Usage : The database was used to predict classes of anaemia for Master Thesis

5) Relevant Information:

Classes	Descriptions	# instances
1	hypochromic microcytic	31
2	normochromic microcytic	46
3	macrocytic	1
4	iron deficiency	177
5	thalassaemia	77
6	anemia of chronic disorder	13
7	lead poisoning	0
8	sideroblastic anemia	1
9	haemolytic anemia	33
10	secondary anemia	1
11	after acute blood loss	5
12	bone-marrow failure	14
13	vitamin b12/folate def	28
14	alcohol	0
15	liver disease	0
16	normal	122
17	classes 4 & 13	46
Total		**595**

APPENDIX E

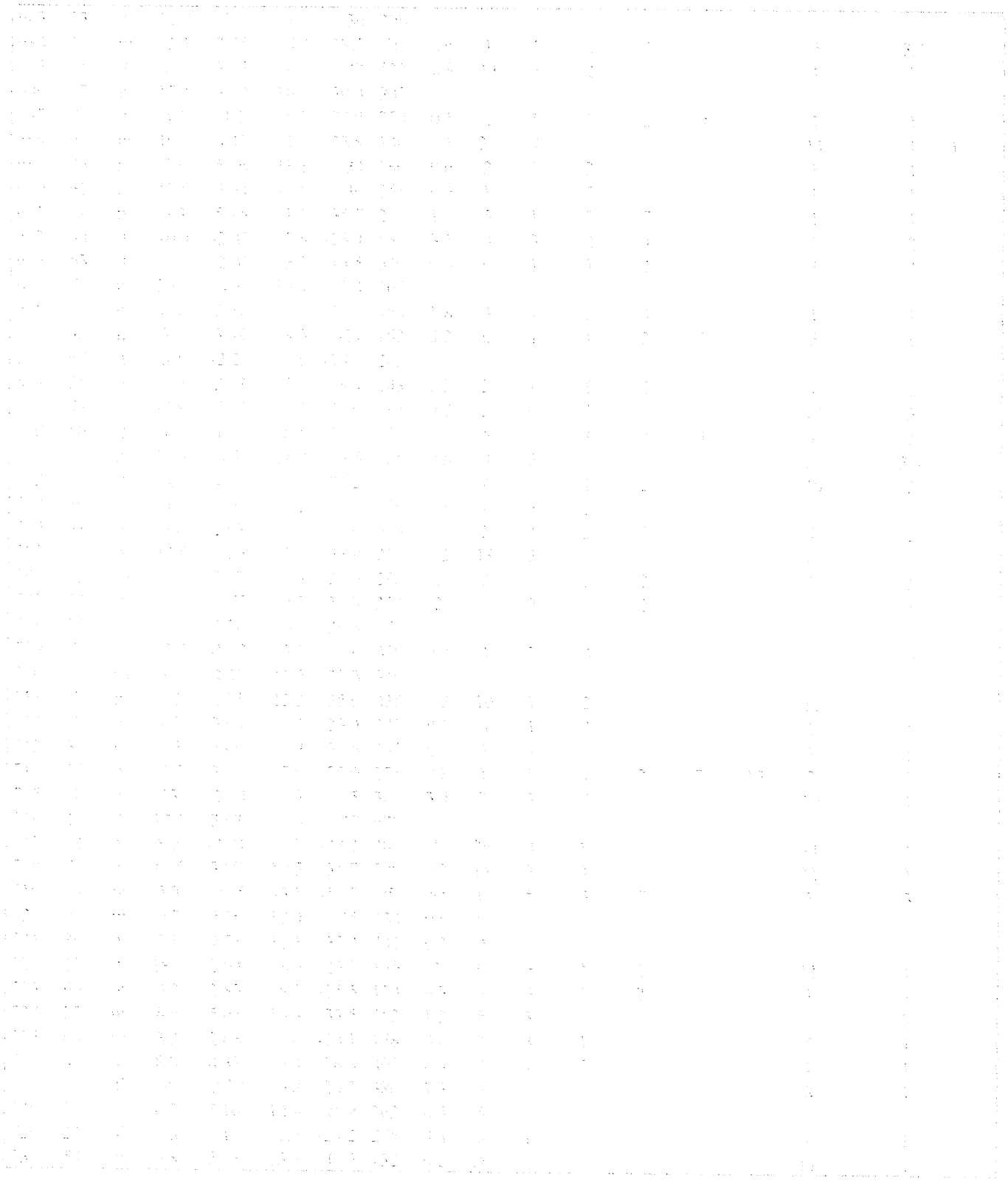
RAW DATA

* The total number does not equal to initial databases (730) since classes 17 is grouped of classes 4 & 13.

6) Number of Attributes : 18 plus the class attribute

7) Attributes Information :

#	Symbol	Description	
1	AG	Age	
2	SX	Sex	
3	WBC	FBC - White Blood Cell	**10^3/uL**
4	RBC	FBC - Red Blood Cell	**10^6/uL**
5	HGB	FBC - Haemaglobin	**g/dL**
6	HCT	FBC - Haematocrit	**%**
7	PLT	FBC - Platlet	**10^3/uL**
8	Rtc	FBC - Reticulocyte	
9	Eos	FBC - Eosinophil	
10	Rbc	FBP - Reticulocyte	
11	S1	FBP - Shape1	
12	S2	FBP - Shape2	
13	S3	FBP - Shape3	
14	S4	FBP - Shape4	
15	Wc	FBP - White Cell1	
16	Wc	FBP - White Cell2	
17	Pt	FBP - Platlet	
18	Tg	Target (Classes)	

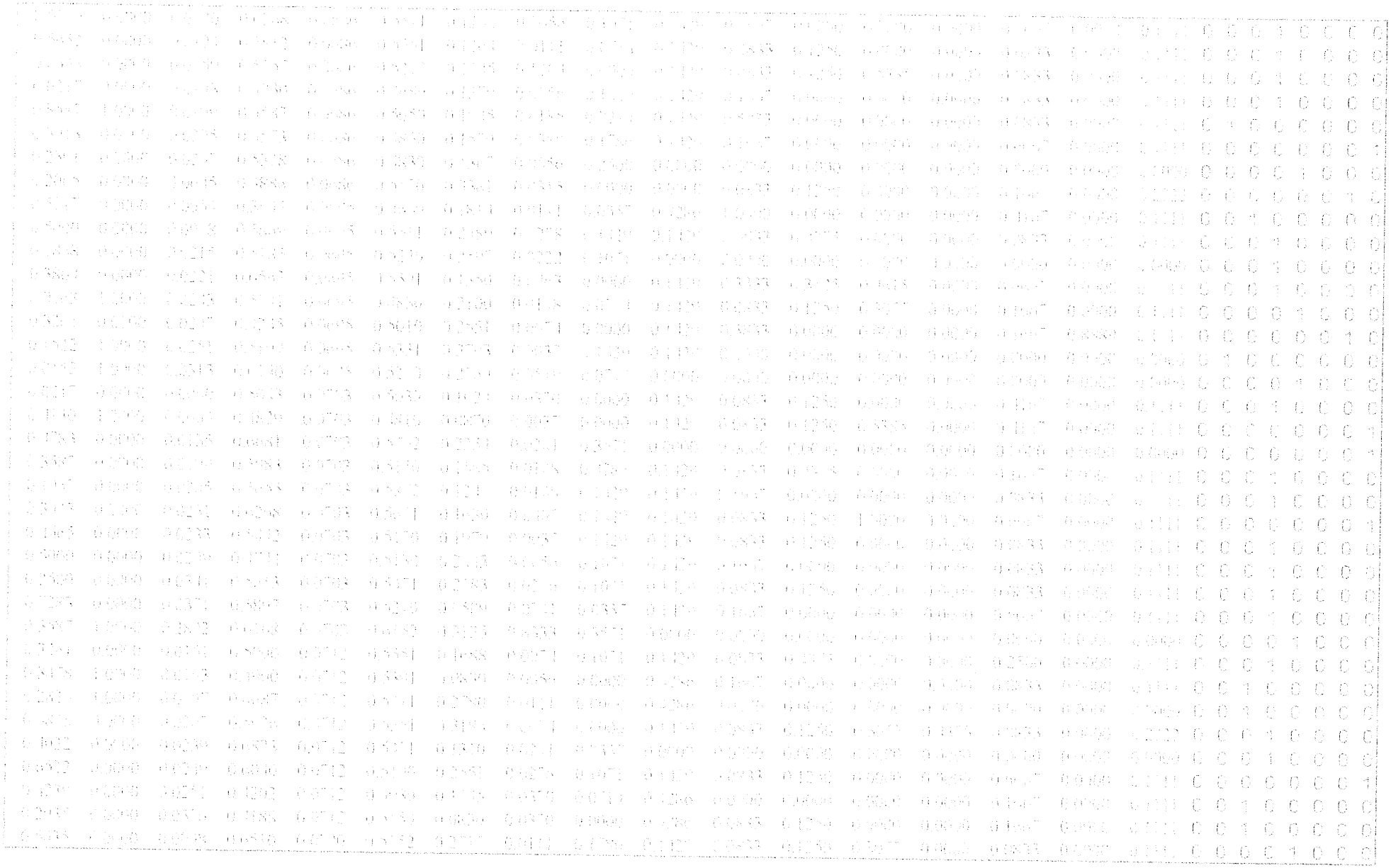


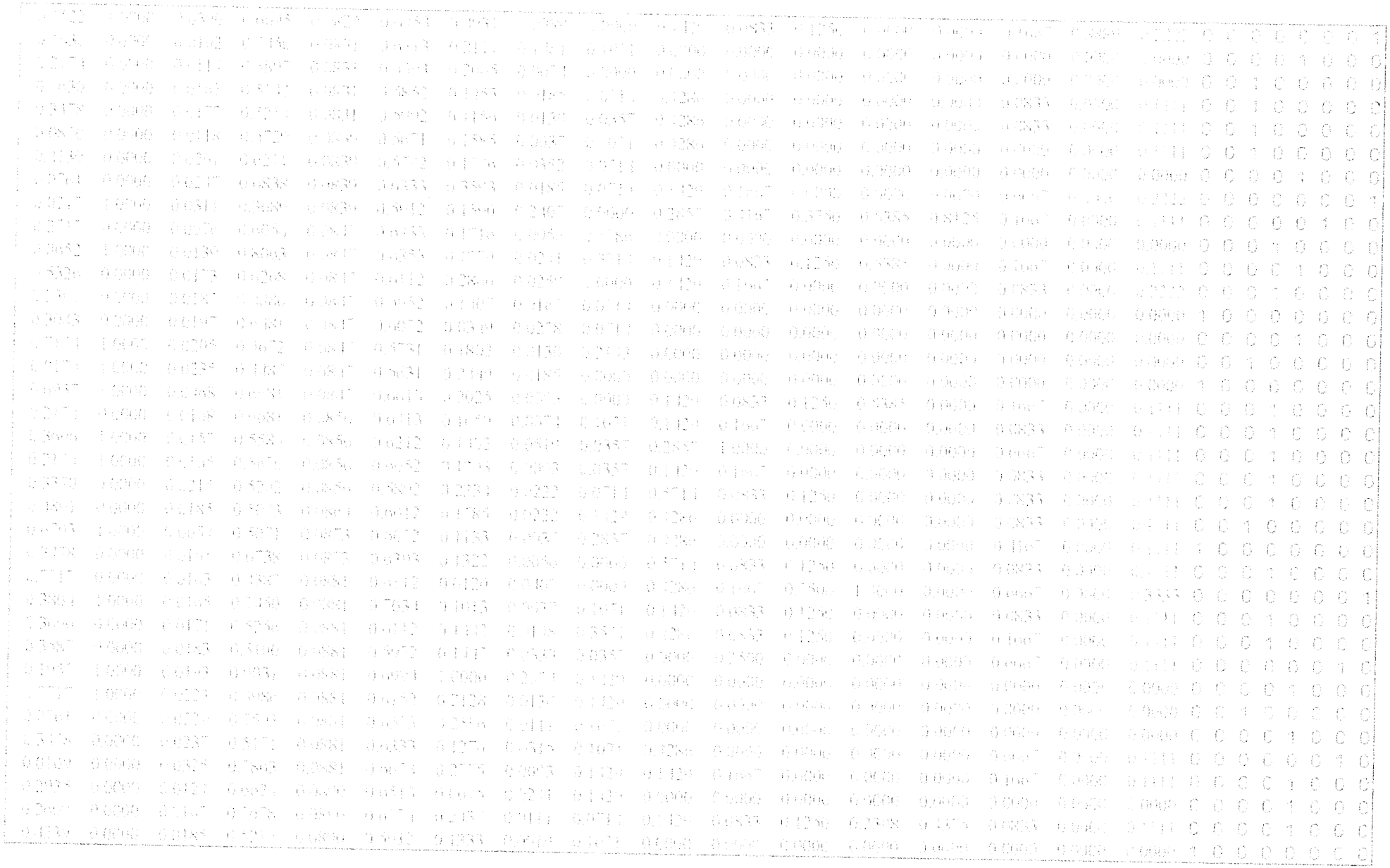
APPENDIX F

NORMALIZED DATA

APPENDIX F: NORMALIZED DATA

AG	SN	WBC	RBC	HGB	HCT	PLT	Rtc	Eos	RBC	SHP	SHP	SHP	SHP	WC	WC	Pt	Target							
0.0326	0.0000	0.0273	0.2023	0.0195	0.1713	0.3881	0.0944	0.0714	0.1429	0.1667	0.3125	0.5385	0.8125	0.0833	0.0000	0.1111	0	0	0	0	1	0	0	0
0.0543	1.0000	0.0617	0.1897	0.0195	0.1775	0.3363	0.3667	0.0000	0.2286	0.1667	0.3750	0.0000	0.0000	0.1667	0.0000	0.2222	0	0	0	0	0	1	0	0
0.1667	0.0000	0.0073	0.2667	0.0229	0.1964	0.2654	0.0407	0.0714	0.1429	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.1000	0	0	0	1	0	0	0	0
0.3696	0.0000	0.0183	0.2279	0.0229	0.2164	0.4565	0.0333	0.1071	0.1429	0.1667	0.2500	0.4615	0.5625	0.1667	1.0000	0.2222	0	0	0	0	1	0	0	0
0.0435	0.0000	0.0339	0.2667	0.0229	0.2184	0.3387	0.2000	0.1429	0.1429	0.1667	0.1875	0.3846	0.0000	0.2500	0.0000	0.1111	0	0	0	0	1	0	0	0
0.5652	0.0000	0.0068	0.1153	0.0237	0.1994	0.1419	0.1074	0.0357	0.2286	1.0000	0.0250	0.0000	0.0000	0.0833	0.0000	0.1111	0	0	0	0	0	1	0	0
0.5043	0.0000	0.0245	0.1166	0.0254	0.1703	0.1130	0.0037	0.0000	0.2286	1.0000	0.0000	0.0000	0.0000	0.2500	0.0000	0.1111	0	0	0	0	0	1	0	0
0.6435	0.0000	0.0250	0.1624	0.0254	0.1901	0.2059	0.0630	0.0714	0.2286	0.3333	0.0000	0.0000	0.0000	0.1667	0.0000	0.1111	0	0	0	0	0	1	0	0
0.2391	0.0000	0.0100	0.1980	0.0268	0.2064	0.0978	0.0630	0.0000	0.1429	0.1667	0.3125	0.0000	0.8125	0.1667	0.0000	0.1111	0	1	0	0	0	0	0	0
0.1130	1.0000	0.0166	0.1068	0.0263	0.2111	0.0513	1.0000	0.0000	0.7143	0.2500	0.1875	0.3846	0.0000	0.0833	0.1250	0.3333	0	0	0	0	0	1	0	0
0.0217	1.0000	0.0549	0.1505	0.0263	0.1924	0.2451	0.2000	0.1429	0.2286	0.3333	0.0000	0.0000	0.0000	0.1667	0.0000	1.0000	0	0	0	0	0	1	0	0
0.2391	1.0000	0.0143	0.1738	0.0271	0.2545	0.0252	0.0014	0.0000	0.0000	0.1667	0.2500	0.4615	0.7500	0.0667	0.0000	0.3333	0	0	0	0	0	1	0	0
0.5217	0.0000	0.0153	0.3832	0.0271	0.2966	0.3318	0.0185	0.1786	0.1429	0.0833	0.1250	0.3077	0.0000	0.0833	0.0000	0.2222	0	0	0	1	0	0	0	0
0.6630	0.0000	0.0171	0.2450	0.0271	0.2101	0.3667	0.0211	0.0357	0.1429	0.0833	0.1250	0.5385	0.0000	0.0833	0.0000	0.2222	0	0	0	1	0	0	0	0
0.0163	0.0000	0.0287	0.3435	0.0271	0.2635	0.3381	0.0463	0.1786	0.2857	0.2500	0.8125	0.0000	0.0000	0.0833	0.0000	0.2222	0	0	0	0	1	0	0	0
0.0087	1.0000	0.0541	0.1453	0.0271	0.2044	0.1173	0.1944	0.0714	0.7143	0.4167	0.3750	0.0000	0.0000	0.1667	0.0000	0.1111	0	0	0	0	0	1	0	0
0.1413	0.0000	0.0137	0.1368	0.0280	0.2001	0.1061	0.0556	0.1071	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0	0	0	0	0	1	0	0
0.5519	0.0000	0.0137	0.1721	0.0288	0.2305	0.0070	0.3556	0.0000	0.2286	1.0000	0.8125	0.1538	0.0000	0.2500	0.0000	0.3333	0	0	0	0	0	1	0	0
0.1130	1.0000	0.0221	0.3775	0.0288	0.2966	0.0761	0.0241	0.0357	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0	0	0	0	0	0	1	0
0.5217	0.0000	0.0070	0.1695	0.0297	0.2181	0.1217	0.0222	0.1429	0.2286	0.0000	0.0000	0.0000	0.0000	0.0833	0.0000	0.1111	0	0	1	0	0	0	0	0
0.1565	0.0000	0.0124	0.2677	0.0305	0.2315	0.1390	0.0185	0.0357	0.5714	0.0833	0.1250	0.2308	0.0000	0.0833	0.0000	0.1111	0	0	0	0	1	0	0	0
0.5326	1.0000	0.0197	0.3489	0.0305	0.3126	0.2177	0.0148	0.1071	0.1429	0.0833	1.0000	0.0000	0.0000	0.1667	0.0000	0.1111	0	1	0	0	0	0	0	0
0.8478	0.0000	0.0088	0.3547	0.0314	0.2775	0.2071	0.0111	0.0714	0.1429	0.0833	0.1250	0.0000	0.0000	0.1167	0.0000	0.1111	0	0	0	0	0	0	0	1
0.5761	0.0000	0.0157	0.3519	0.0314	0.2966	0.2975	0.0111	0.0000	0.1429	0.0833	0.1250	0.0000	0.0000	0.0833	0.0000	0.2222	0	0	0	1	0	0	0	0
0.1141	0.0000	0.0257	0.1715	0.0314	0.3186	0.3062	0.0542	0.1786	0.1429	0.0833	0.1250	0.3077	0.0000	0.0833	0.0000	0.2222	0	0	0	1	0	0	0	0
0.4891	1.0000	0.0249	0.1823	0.0322	0.2265	0.1928	0.0407	0.0000	0.2286	0.3333	0.3750	0.0000	0.0000	0.0833	0.0000	0.1111	0	0	0	0	0	1	0	0
0.0543	1.0000	0.0189	0.1631	0.0331	0.3026	0.2517	0.0667	0.2143	0.1429	0.0833	0.1250	0.3077	0.0000	0.1000	0.0000	0.1111	0	0	0	0	1	0	0	0
0.1630	1.0000	0.0227	0.3031	0.0331	0.2665	0.3066	0.1444	0.1071	0.1429	0.5000	0.0000	0.3846	0.0000	0.1000	0.0000	0.0000	1	0	0	0	0	0	0	0
0.1413	1.0000	0.0130	0.1181	0.0331	0.2325	0.0653	0.3704	0.0000	0.2286	0.4167	0.7500	1.0000	0.0000	0.1667	0.0000	0.1111	0	0	0	0	0	1	0	0
0.5543	0.0000	0.0092	0.3402	0.0339	0.3317	0.2644	0.0926	0.0357	0.1429	0.0833	0.1250	0.3077	0.0000	0.1667	0.0000	0.2222	0	0	0	0	0	0	0	1
0.7174	0.0000	0.0173	0.3761	0.0339	0.3387	0.1899	0.0056	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0	0	0	1	0	0	0	0
0.5652	1.0000	0.0173	0.3818	0.0339	0.3166	0.3244	0.0241	0.1429	0.1429	0.0833	0.1250	0.0000	0.0000	0.0833	0.0000	0.2222	0	0	0	1	0	0	0	0
0.2391	0.0000	0.0096	0.3348	0.0347	0.3126	0.1522	0.0148	0.2143	0.1429	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0	1	0	0	0	0	0	0
0.0087	0.0000	0.0080	0.2365	0.0356	0.2615	0.1205	0.1111	0.1429	0.1429	0.4167	0.3750	0.0000	0.0000	0.0833	0.0000	0.1111	0	1	0	0	0	0	0	0





APPENDIX G

USER MANUAL

NEURAL CONNECTION ver. 2.0

Neural Connection is a software system that allows complex applications for solving business problems using neural computing and other techniques.

Neural networks

In general, neural networks are simply a new way of data analyzing. What makes them differ and useful is their ability to learn complex patterns and trends in the data sets rather than being programmed to get the right answer. Neural networks are very good in face recognition, speech recognition, medical analysis and others.

By producing systems that learn the relationships between data and results, neural networks avoid many of the problems of conventional computing. Given new, unseen data, a neural network can make a decision or prediction based upon its experience just as a human can.

Installing Neural Connection

Neural Connection is provided with a setup program that automatically installs Neural Connection and its associated files onto the hard disk. The setup program must be run before it can be used. The next section explains how to setup Neural Connection in the personal computer.

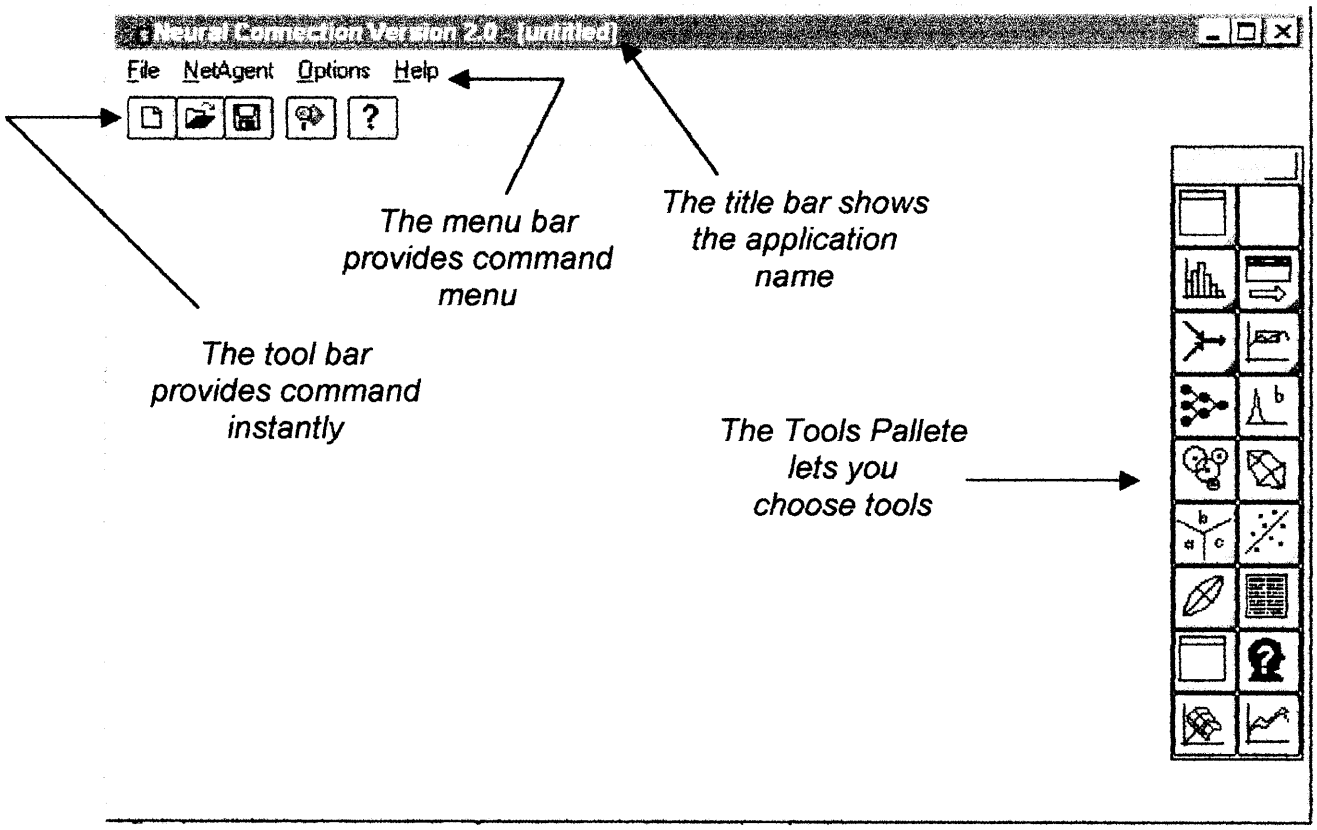


Figure 2.0: Neural Connection workspace

The workspace is the blank canvas on which you build the data analysis applications. It has its own set of commands that are called from the *menu bar*. The workspace allows you to build, train, and run analysis applications. An application is a method for modeling your problem, which may include inputting the data, processing them in some way, and producing the output in a useful format. Tools are selected as icons from a palette, and moved onto the workspace, where they can be connected to other tools. These connections determine an application's topology, and the path along which data flows.

If you do not see the name of this application, select a new drive or directory where the file may reside.

2. Click OK. The workspace now looks like Figure 4.0

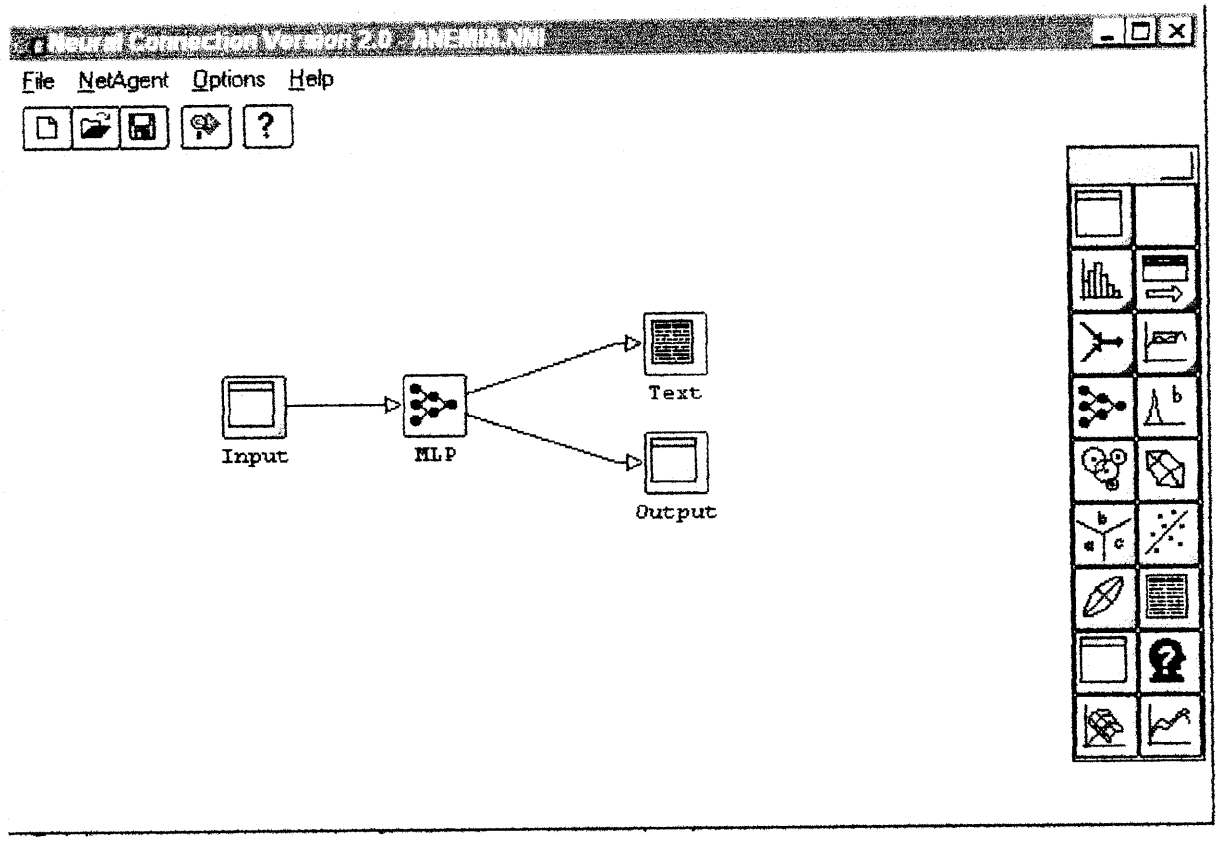


Figure 4.0

Viewing Results in the Data Output Tool

To view the prediction result, point to the icon "Output" and click the right mouse button and menu as Figure 5.0 appears. Choose View.

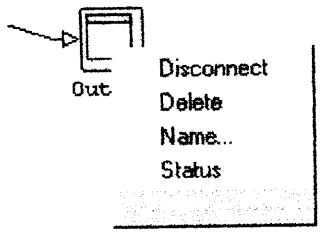


Figure 5.0

From the View menu, choose the data set that you want to view. There are four type of data set and each is identified by the following colored codes in the left hand column on the screen:

Type of Data Sets

T = Training data, in cyan

V = Validation data, in bright green

X = Test data, in yellow

R = Run data, in pale green

To view the results of the prediction data, choose R (run data). The prediction results window is shown in Figure 6.0. Scroll bars are provided to move around the text display if the results do not all fit in the window. The fields within data sets are identified by the following colored codes in the top row on the screen:

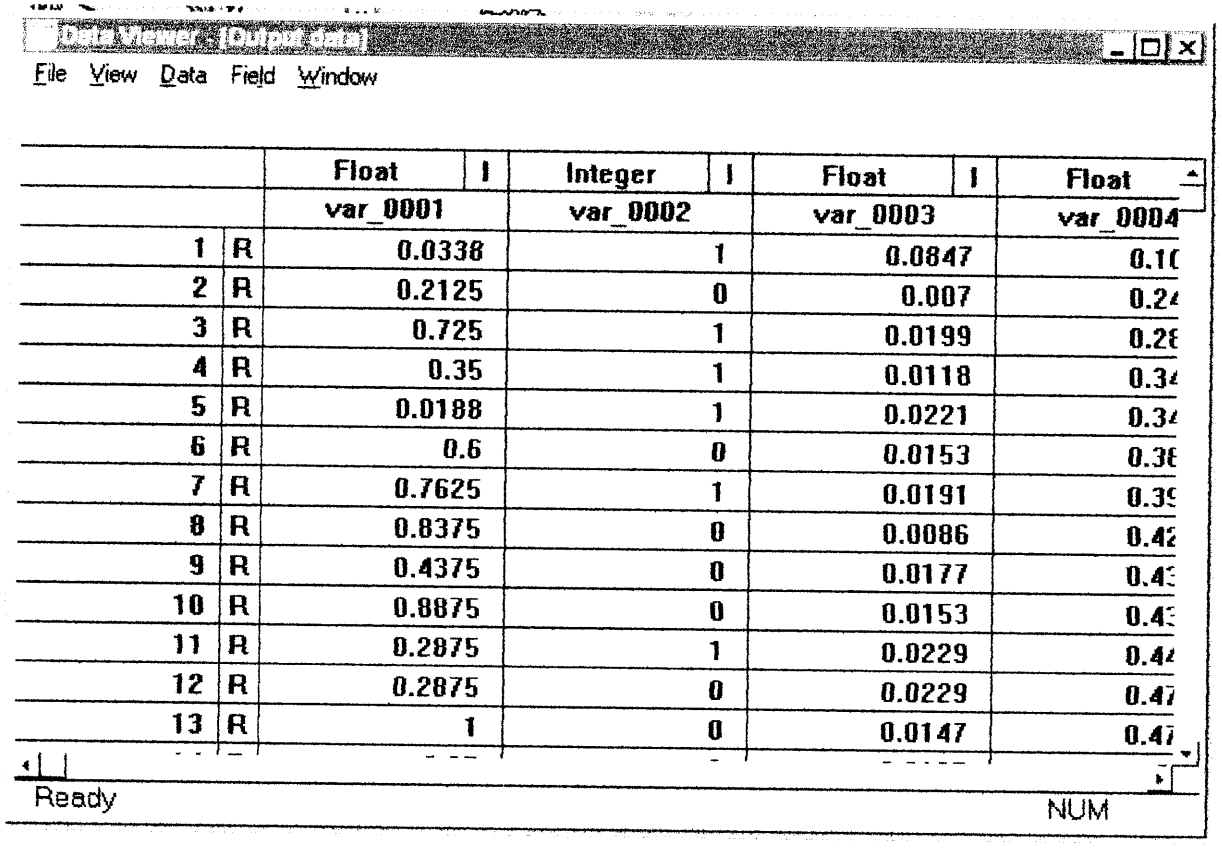

File View Data Field Window
Float
Integer
Float
Float
var_0001
var_0002
var_0003
var_0004
1 R 0.0338 1 0.0847
2 R 0.2125 0 0.007
3 R 0.725 1 0.0199
4 R 0.35 1 0.0118
5 R 0.0188 1 0.0221
6 R 0.6 0 0.0153
7 R 0.7625 1 0.0191
8 R 0.8375 0 0.0086
9 R 0.4375 0 0.0177
10 R 0.8875 0 0.0153
11 R 0.2875 1 0.0229
12 R 0.2875 0 0.0229
13 R 1 0 0.0147
Ready
NUM

Figure 6.0

Field Usage

I = Input fields, in blue

T = Target fields, in yellow

R = Reference fields, in green

* = Unused fields

M = Network Target fields, in pale blue, are named MTARGET1, MTARGET2, etc.

O = Network Output fields, in pale green, are named OUTPUT1, OUTPUT2, etc.

For anaemia classification:

var_0001 – var_0017 is the Input fields

var_0018 – var_0025 is the Target fields

MTarget1 – MTarget8 is the Network Target fields

Output1 – Output8 is the Output fields

The following table, Table 1.0 shows how to read the prediction result that is labeled as Output data.

Output1	Output2	Output3	Output4	Output5	Output6	Output7	Output8	Result / Prediction
1	0	0	0	0	0	0	0	Normal
0	1	0	0	0	0	0	0	Hypochromic microcytic
0	0	1	0	0	0	0	0	Normocromic microcytic
0	0	0	1	0	0	0	0	Iron deficiency
0	0	0	0	1	0	0	0	Thalassemia
0	0	0	0	0	1	0	0	Haemolytic anemia
0	0	0	0	0	0	1	0	B12 def./Folate def.
0	0	0	0	0	0	0	1	Iron/B12 def./Folate def.

Table 1.0

Quitting Neural Connection

To quit Neural Connection from the workspace:

1. From the File menu, choose Exit (Alt, F, X)
2. If an application has unsaved changes, Neural Connection asks if you want to save the application. Either click **Yes** to save your work or press **Enter** to quit without saving.
3. If the application hasn't been named, Neural Connection asks you to provide a name for the application. Type a name; your application is automatically store with an *.nni* filename extension.