Chapter One

Introduction

1.1 General View

Medical information systems in modern hospitals and medical institutions have become larger and this trend causes big difficulties for extracting useful information for decision support systems. Increased number of clinical databases, increases manual processing which increases the cost. Most of the medical data sets are non-linear patterns and difficult to classify. Therefore, there is a need to enhance the classification by means of machine learning methods[1].

Artificial neural networks seemed to be one of the best approaches in machine learning methods. The use of neural networks for diabetic Diagnosis has also attracted the interest of the medical informatics community because of their ability to model the nonlinear nature. Gradient based methods are one of the most widely used error minimization methods used to train back propagation networks. Back propagation algorithm is a classical domain dependent technique for supervised training. It works by measuring the output error calculating the gradient of this error, and adjusting the ANN weights and biases in the descending gradient direction. Back propagation is the most commonly used and the simplest feed forward algorithm used for classification[2][3].

Diabetes is disease in which the body does not properly produce insulin. It is one of the most common chronic disease, which can lead to serious long term complications. Type I diabetes, the disease is caused by the failure of pancreas, to produce sufficient insulin, and type II the body is resistant to the insulin it makes, which leads to an uncontrolled increase of blood glucose unless the patient uses insulin or drug. The blood glucose level that is elevated for a long period can result in metabolic complications such as kidney failure, blindness, and an increased chance of heart attacks. To prevent or postpone such complications strict control over the diabetic blood glucose level is needed[3][1] .

The goal of this study is to propose the use of ANNs for efficient computer-based analysis are essential for diagnosing of Diabetes Mellitus because provides an improvement in accuracy, sensitivity, and specificity in comparison with other methods there manual data analysis of the diabetes mullitus has become inefficient.

1.2 The problem statement

Increased number of diabetic clinical databases, increases traditional manual data analysis and processing has become inefficient ,which increases the cost. The diabetic data sets are non-linear patterns and difficult to classify and diagnosis.

1.3 Objective

The goal of this study is to Design a novel approach for diagnosing Diabetes using Artificial neural networks.

1-4 Methodology

In this study, three-layered Multilayer Perceptron (MLP) in the three network : feed-forward neural network architecture ,Elman network architecture , and recurrent network architecture ,Using nntool in MATLAB BPF was designed and trained ,first one hidden layer between input and output layer with 3,4,5,6,7,8,9 neuron in hidden layer . type of activation function which chooses firstly is log sigmoid[7].

Then for the same network, activation function was replaced first by tan sigmoid function.

was used and trained with the error back propagation algorithm. The back propagation training with generalized delta learning rule is an iterative gradient algorithm designed to minimize the root mean square error between the actual output of a multilayered in the three type of neural network and a desired output. Each layer is fully connected to the previous layer, and has no other connection .

Chapter Two

Literature Review

B.Y.Baha ,B.Yola and etc .., Artifical Neural Networks To Detectect Risk Of Type 2 Diabetes , June, 2012 , In this research, 7 risk factors and their strength of association to the development of Type 2 diabetes was used as relative weight of input variables. A multilayerfeedforward architecture with backpropagation algorithm was designed using Neural Network Toolbox of Matlab. The network was trained using batch mode backpropagation with gradient descent and momentum. Best performed network identified during the training was 2 hidden layers of 6 and 3 neurons, an output layer of 1 neuron, logsigmoid transfer function at the hidden layers and a linear transfer function at the output layer [1].

F.Amato, A.López andetc .. , Artificial neural networks in medical diagnosis , 7th January 2013, To streamlined the diagnostic process in daily routine and avoided misdiagnosis, artificial intelligence methods (especially computer aided diagnosis and artificial neural networks) could be employed , These adaptive learning algorithms could handle diverse types of medical data and integrate them into categorized outputs. In this paper, them briefly review and discuss the philosophy, capabilities, and limitations of artificial neural networks in medical diagnosis [2].

B.Adeyemo and A.E.Akinwonmi , On the Diagnosis of Diabetes Mellitus Using Artificial Neural Network Models , September 2011 , In this work Artificial Neural Network models were developed using both classification and predictive neural networks for the rapid diagnosis of diabetes mellitus. Both neural network models were able to learn the problem with the predictive network giving a better performance of 84% correctly classified records as opposed to 76% achieved by the classifier network on the same data set, A combined Diagnosis and Treatment neural network was also modeled using various neural network architectures. The GRNN/PNN network gave the best result out of the three architectures used. The other networks were unable to model the problem [3].

Og˘uz.Karan ,C.Bayraktar and etc .. , Diagnosing diabetes using neural networks on small mobile devices , 2012 , This paper presented a novel approach for diagnosing diabetes using neural networks and pervasive healthcare computing technologies , A distributed end-to-end pervasive healthcare system utilizing neural network computations for diagnosing illnesses was developed [8].

M.Pradhan. and Dr. R.K.Sahu , Predict the onset of diabetes disease using Artificial Neural Network (ANN) , April 2011, In this paper we have experimented and suggested an Artificial Neural Network (ANN) , the model is trained with Back Propagation (BP) algorithm and GA (Genetic Algorithm) and classification accuracies are compared. The designed models are also compared with the Functional Link ANN (FLANN) and several classification systems like NN (nearest neighbor) , the model and it can be a very good candidate for many real time domain applications as these are simple with good performances [9].

P.Venkatesan and S.Anitha , Application of a radial basis function neural network for diagnosis of diabetes mellitus , 10 NOVEMBER 2006 , In this article an attempt is made from them to studed the applicability of a general purpose, supervised feed forward neural network with one hidden layer, namely. Radial Basis Function (RBF) neural network .Diabetes database used for empirical comparisons and the results show that RBF network performs better than other models [10].

S.Kabilan,S.Kannan, and etc .. , Diagnosis of Diabetes Mellitus Type 2 , February 2014 , The proposed system could improve the strategy to a better level where artificial metaplasticity on perceptrons is implemented on neural network , this network container 13input nodes, 2 hidden layers and one output node. and this system can increase the efficiency of the system which is in existence. The design, training and testing of neural network is done in MATLAB [11].

O. Wahyunggoro, A.E. Permanasari, andetc .., Utilization of Neural Network for Disease Forecasting , This paper will presented the use of neural network to learn the historical patterns of disease incidence to forecast future incidence , The results show the advantages of neural network for supporting policy/decision makers in developing long term strategies regarding the number of disease incidence [12].

K. Rajeswariand V. Vaithiyanathan , Fuzzy based modeling for diabetic diagnostic decision support using Artificial Neural Network , April 2011 , The study is on people approaching diabetician with either past history of Diabetics or new case with symptoms of diabetics, Some cases are Normal patients without diabetics. The required parameters are estimated by interviewing patients. Later the parameters are modeled using a fuzzy approach and after normalization classified by Artificial neural networks as ‘Close to Type 2 diabetic’ or not. This result may indicate the effectiveness of proposed algorithm to optimally model the diagnostic process for small or large datasets; especially, due to its computational simplicity [13].

Chapter Three

Theoretical Background

3.1 Artificial neural networks (ANNs)

neural networks can be viewed as massively parallel computing systems consisting of an extremely large number of simple processors with many interconnections . neural network models attempt some organizational principles (such as learning , generalization , adaptively , fault tolerance and distributed representation and computation) in network weighted directed graphs in which the nodes are artificial neurons and directed edges (with weights) are connections between neuron outputs and inputs[2][3][4][5].

the main differences between neural network and other approaches to pattern recognition are that these network have the ability to learn complex non-linear input-output relationship , and use sequential training procedures. Moreover they have the general characteristic of adapting themselves to the data .

ANN is a massively parallel distributed processor called artificial neurons(ANs) that has the ability for storing experiential knowledge and making it available for use . It resembles the brain in :

1- Knowledge is acquired by the network through a learning process .

2- Inter-neuron connection strengths known as synaptic weights are used to store knowledge.

3.2ANNs benefits

The advantages of neural network information processing arise from its ability to recognize and model nonlinear relationships between data . in biologic systems, clustering of data and nonlinear relationships are more common than strict linear relationships[5].

Conventional statistical methods can be used to try to model nonlinear relationships , but they require complex and extensive mathematical modeling .

neural networks provide a comparatively easier way to do the same type of analysis .

neural network have evolved over the the past years , and used in making financial and weather predictions . neural network have been used in a variety of medical applications , mainly in image interpretation , laboratory studies , and clinical diagnosis . one of their most valuable assets is their ability to predict prognosis . ANNs are more accurate at predicting prognosis than conventional TNM staging[3][4].

3.3 Information processing methodology in ANNs

Neural networks can do perform complex functions in various fields if there are enough resources. The ANs in the network can be trained by adjusting the values of the connections (weights) between elements. to get the final weights for a particular function that make the output of the neural match the target[4][5].

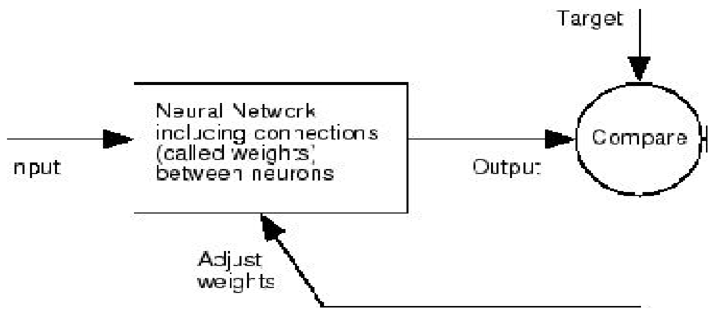


Figure 3.1 Neural network adjusting system

An artificial neuron in shown in figure3.2

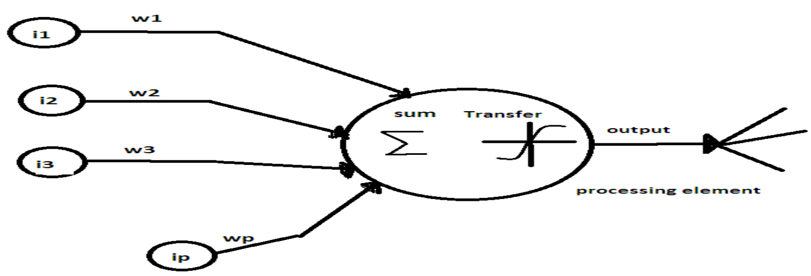


Figure3.2: Mathematical model of an artificial neuron.

The neuron receives coming input and multiplies each of them by weight and then summated all , this is called the summation function . the network output is as in equation (1).

$$net = \sum_{i=0}^{d} w_{ji x_i} \quad -(1)$$

then output was normalized using a non linear transformation by transfer (activation) function as explained in equation(2).

$$y = f(net) \quad -(2)$$

Popular choices for the activation function f include

(1) the thresholding function

$$f(net) = \begin{cases} 1 \; if \; net \; \geq 0 \; or \\ 0 \; otherwise \end{cases} \text{-(3)}$$

$$f(net) = \begin{cases} 1 \; if \; \geq 0 \; or \\ -1 \; othewise \end{cases} \quad -(4)$$

(2) the identity (linear) function

$$f(net) = net \quad -(5)$$

(3) the logistic (sigmoid) function

$$f(net) = \frac{1}{1+e^{-\beta net}} \quad -(6)$$

Where β is slop of the nonlinear activation

(4) the hyperbolic tangent (sigmoid) function

$$f(net) = \tanh(\beta net) = \frac{2}{1 + e^{-\beta net}} - 1 \quad -(7)$$

The type of activation function plays an important role in information understanding operation and it is identified according to the application[10].

3.4.ANNs learning

After ANN designed it must be learned on information understanding , learn simply means adjusting ANN weights unit the network understand the information . there are two way for ANN learning supervised and unsupervised learning[3] .

3.4.1 Supervised learning

In the initial phase of operation the neural network “learns” .this is accomplished in the supervised learning method , by showing the network input data and the known outcome . the interwoven weights are adjusted by the training algorithm to reproduce the desired answer.

Initially when a neural network is presented with a pattern , it makes a random guess as to what the answer might be . it then sees how far its answer was from the actual one and makes an appropriate adjustment to its connection weights. This process , known as the backward propagation of error , is repeated unit the evolved set of weights best reproduces the entire training and testing database results in an answer that is most accurate. After the learning phase, the network is given the input data , and it gives its best answer based on prior learning[5] .

3.4.2. Unsupervised learning

in unsupervised learning, the desired response is not known a priori . thus , explicit error information is unavailable to improve network behavior . since no information is available as to the correctness or incorrectness of response , learning must somehow be accomplished based on observation of responses to inputs that yield marginal or no knowledge about them .

the network is therefore training towards some optimum output where optimum is usually some clustering of the data[5] .

3.5types of ANNs

there are many different kinds of neural networks . this thesis was interested in three of them which are feed-forward back propagation network , Recurrent neural network , and Elman network .

3.5.1feed-forward back propagation network (BPNN)

The propagation net was first introduced by G.E. Hinton ,E. Rumelhart and R.J. Williams in 1986 and is one of the most powerful neural net types. It has the same structure as the Multi-Layer-Perceptron and uses thebackpropagation learning algorithm . figure3-3 – show BP NN architecture[6].

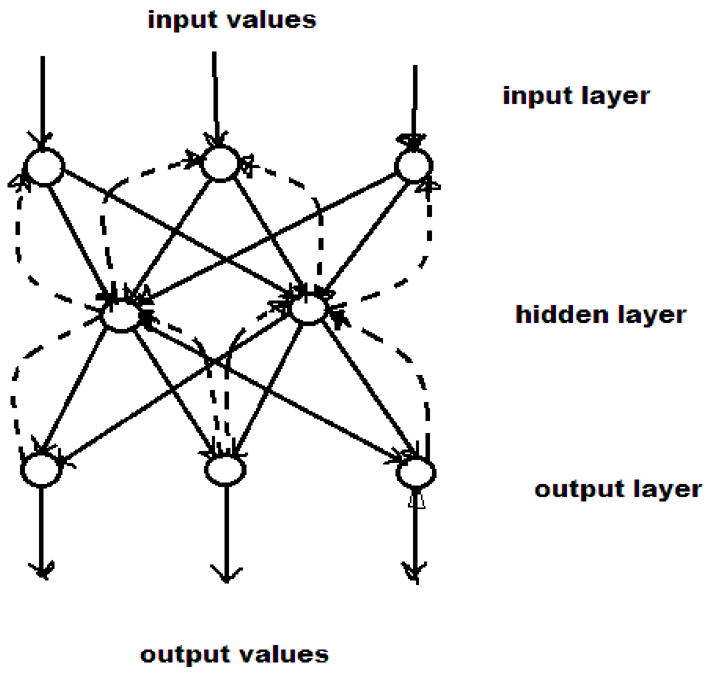


Figure 3.3: three layer BP NN architecture.

In the employment of the backpropagation algorithm , each iteration of training involves the following step: 1) a particular case of training data is fed through the network in a forward direction, producing results at the output layer ,2) error is calculated at output nodes based on known target information , and the necessary changes to the weights that lead into the output layer are determined based upon this error calculation ,3) the changes to the weights that lead to the preceding network layers are determined as a function of the properties of the neurons to which they directly connect (weight changes are calculated , layer by layer , as a function of the errors determined for all subsequent layers , working backward toward the input layer) unit all necessary weight changes are calculated for the entire network . the next iteration begins ,and the entire procedure is repeated using the next training pattern[6].

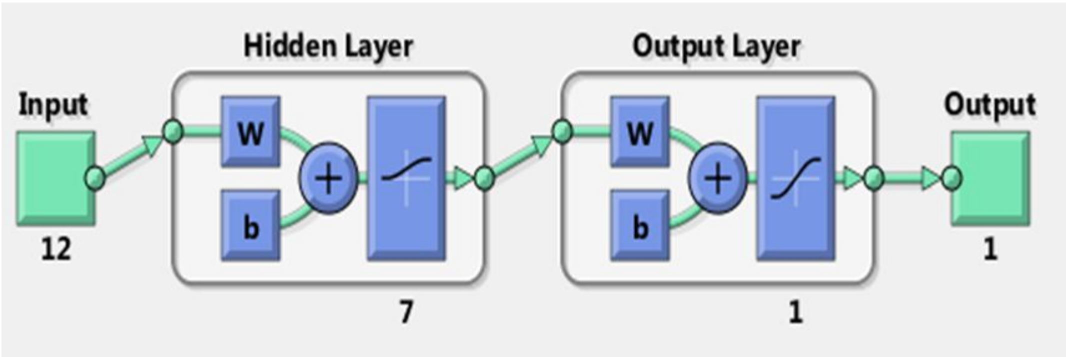


Figure 3.4: the ideal BP NN architecture.

3.5.1.1Applications of BP NN

BN NN might be a good idea in the following situations :

1- A large amount of input/output data is available , but you're not sure how to relate it to the output .
2- The problem appears to have overwhelming complexity , but there is clearly a solution .
3- It is easy to create a number of examples of correct behavior.
4- The solution to the problem may change over time , within the bounds of the given input and output parameters.
5- Outputs can be "fuzzy",or non-numeric[6].

3.5.2Recurrent neural network (RNN)

Layer recurrent network are similar to feed_forwardnetworks , except that each layer has a recurrent connection with a tap delay associated with it .
This allows the network to have an infinite dynamic response to time series input data .in feed forward networks employing hidden unit and a learning algorithm , the hidden units develop internal representations for the input patterns that record those patterns in a way which enables the network to produce the correct output for a given input .
In RNN architecture , the context units remember the previous internal state . thus, the hidden units have the task of mapping both an external input , and also the previous internal state of some desired output. Because the patterns on the hidden units are saved as context, the hidden units must accomplish

this mapping and at the same time develop representations which are useful encoding of the temporal properties of the sequential input.

Thus, the internal representations that develop are sensitive to temporal context; the effect of time is implicit these internal states .note ,however , that these representations of temporal context need not be literal .they represent a memory which is highly task- and stimulus-dependent[6].

Figure3-4 – show simple recurrent network.

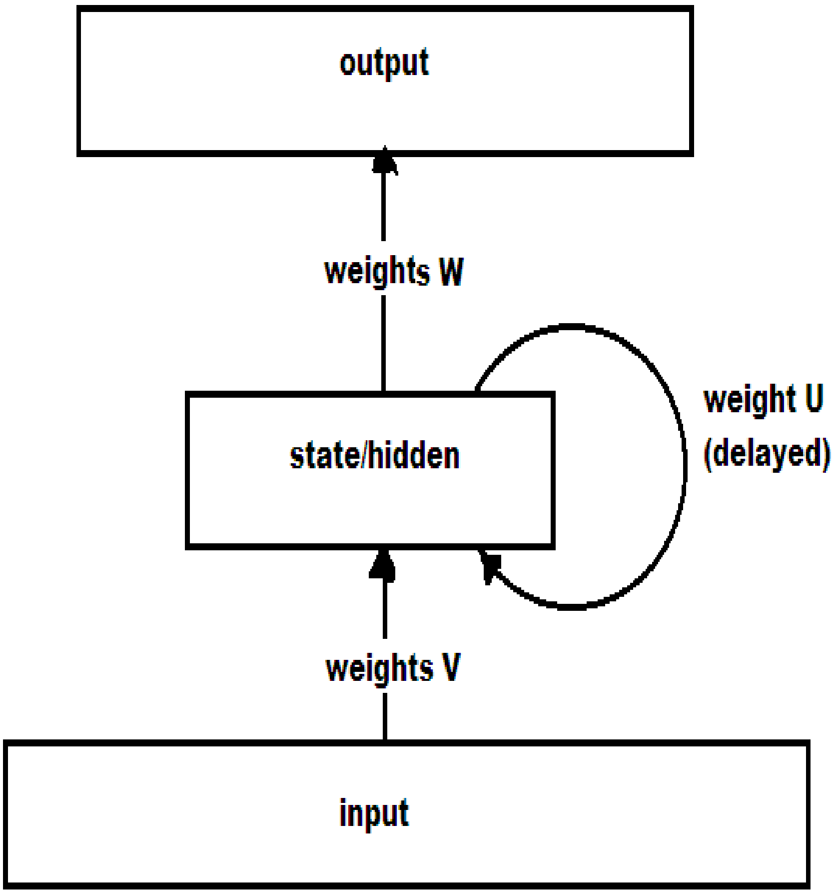


Figure3.5 : simple RNN

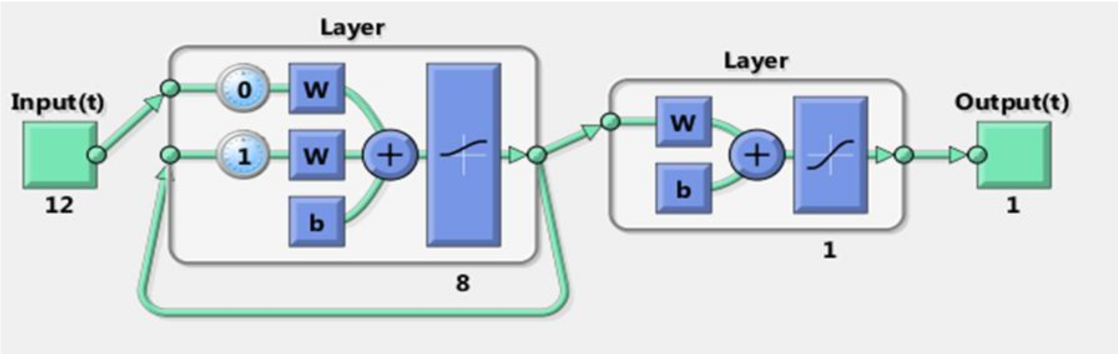


Figure3.6 : the ideal RNN architecture

3.5.2.1Applications of RNN

RNN was used in many field such as complex time series prediction, adaptive robotice and control , connected handwriting recognition ,image classification , aspects of speech recognition , protein analysis , and other sequence learning problem ,RNN has many characteristics which make it able to handle of all this applications sutch as ability to :

1- Recognition of temporally extended patterns in noisy input sequences
2- Recognition of the temporal of widely separated events in noisy input streams.
3- Extraction of information conveyed by the temporal distance between events.
4- Stable generation of precisely time rhytms , smooth and non-smooth periodic trajectories.
5- Robust storage of high-precision real numbers across extended time intervals[6].

3.5.3Elman Network

Elman Network is special case of basic architecture for recurrent network ; it was employed by Jeff Elman . a three-layer netwok is used, with the addition of a set of “context units” in the input layer . there are connections from the hidden layer to these context units fixed with a weight of one . at each time step, the input is propagated in a standard feed forward fashion , and then a simple back prop-like learning rule is applied (this rule is not performing proper gradient descent, however). The fixed back connections result in the context units always maintaining a copy of the previous value of hidden units (since they propagate over the connections before the learning rule iaapplied)[6]. Figure3-5 – show simple one cycle of state vector (context unit).

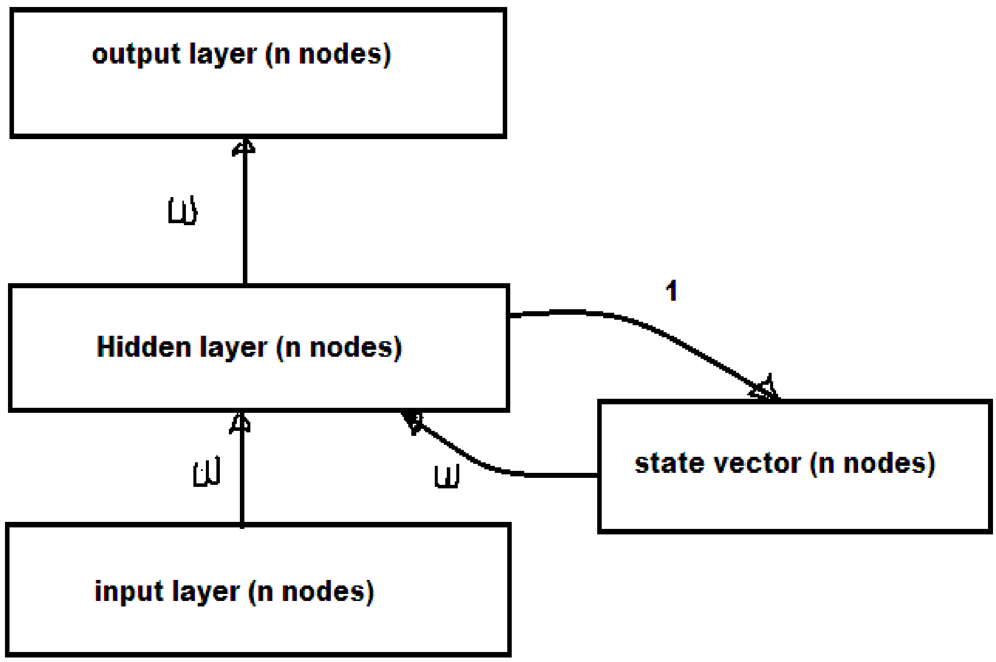


Figure3.7 : architecture of elman network .

In figure – ω signifies total interconnection with trainable weights ; 1 signifies that the activations at destination are a copy of the activations at the source in the previous processing cycle .

The state vector in elman network provides the potential for such networks to store information about previous input .an ordinary back propagation network without some from of feedback loop would be unable to perform tasks which require it to know what the previous input was : such networks could not recognize a tune or any other temporal structure[6] . figure3-6 – show elman network with many state vectors.

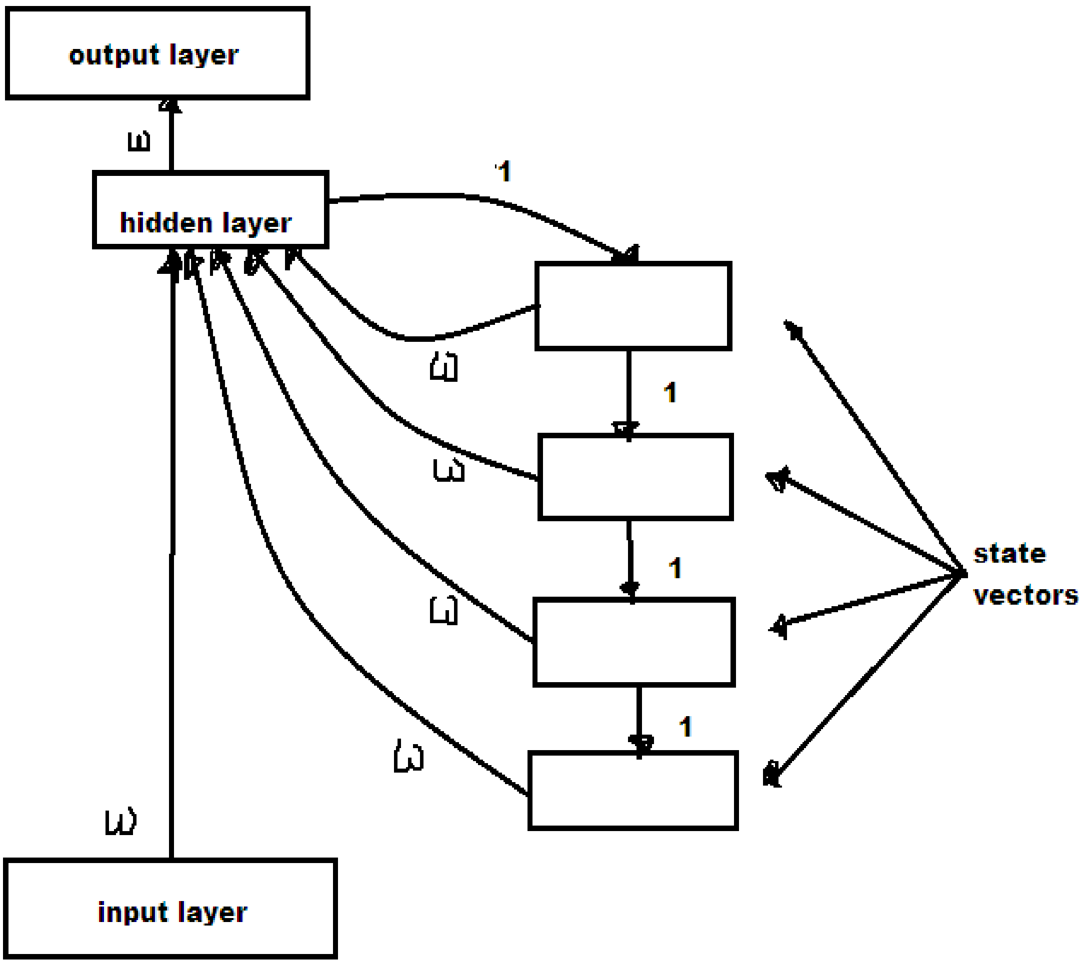


Figure3.8 :elman network with 4 state vectors.

Note that in figure – the connections from the state vectors to hidden layer feedback hidden layer activations from the preceding four processing cycle .

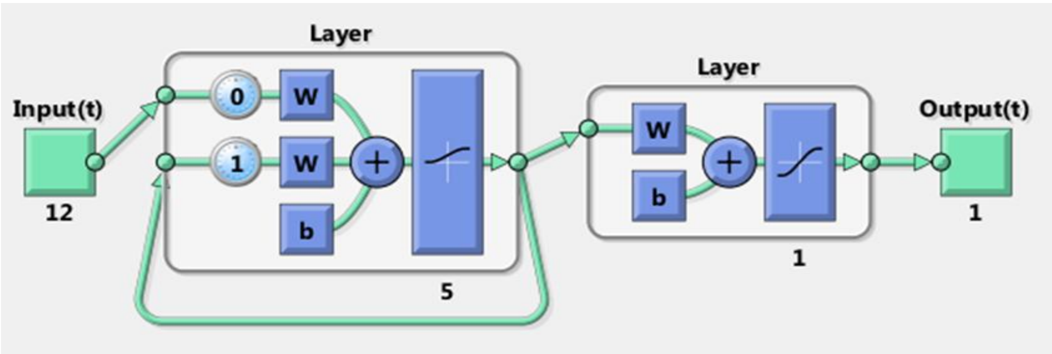


Figure3.9 : the ideal elman network architecture.

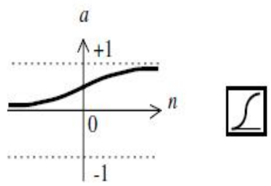
3.5.3.1Applications of elman network

Elman introduced a particular class of recurrent network in which the feedback connections are from the state vector to the hidden layer . as illustrated in previous figure , so it used in all RNN applications and also it used in biomedical engineering and nuclear engineering problems[6].

3.6Neuron Model (logsig, tansig, purelin)Activation function

An elementary neuron with R inputs is shown below. Each input is weighted with an appropriate w. The sum of the weighted inputs and the bias forms the input to the transfer function f. Neurons can use any differentiable transfer function f to generate their output.

Multilayer networks often use the log-sigmoid transfer function logsig[4].

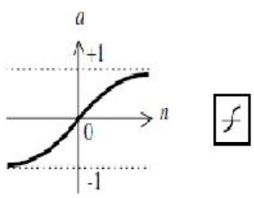


a = logsig(n)

Log-Sigmoid Transfer Function

The function logsig generates outputs between 0 and 1 as the neuron’s net input goes from negative to positive infinity.

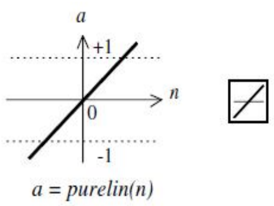
Alternatively, multilayer networks can use the tan-sigmoid transfer function tansig.



a = tansig(n)

Tan-Sigmoid Transfer Function

Occasionally, the linear transfer function purelin is used in backpropagation Networks[4]



Linear Transfer Function

If the last layer of a multilayer network has sigmoid neurons, then the outputs of the network are limited to a small range. If linear output neurons are used the network outputs can take on any value[4].

3.7 Neural network for Diabetes Mellitus classification

Developing a classifier involves choosing an appropriate classifier model and then using the training algorithm to train and then test the input data to classify them into different categories.

Backpropagation algorithm will be used in this project as a training function to train multilayer feedforward , Elman , Recurrent neural network to solve Diagnosis Of Diabetes Mellitus problems.

There are generally four steps in the training process:

1-Assemble the training data .

2-Create the network object .

3-Train the net work .

4-Simulate the network response to new inputs .

The backpropagation training algorithm is an iterative gradient algorithm designed to minimize the mean square error between the actual output to a multi-layer feed-forward, Elman , Recurrent perceptron and the desired output.

3.8 Artificial neural network design

One of important gols ANNs is processing information just as people do , well design and train of neural network make it qualified for decision maker operation when it faced new data outside training data ; this well provide ANNs high reliability exactly like experiment humans.

There are two problem faced designer of ANNs in any application :

1- Network structure. 2- Network generalization.

3.8.1 Network structure

When design ANNs a suitable Architecture for the specific application must be well choose , this involve:

1- Choosing suitable network type for application .

2- Number of layers.

3- Number of nodes in hidden layers.

4- Activation function between layers .

Number of layer , nodes and activation functions determined according to the application and there is not specific rule for choose. This study suggest new approach use trail and error but with specific strategy.

First initial numbers of layers , nodes and activation function will determined and value of performance error will record, then for the same architecture; activation function will replaced and also performance error record and so on . finally activation function which provides the least performance error was choosing .

Now number of hidden nodes will increased and then decreased and performance error also record the purpose of these is to study is the application need more or less of nodes ; if performance error value is still high a new layer must be added.

Finally a new ANN which gives the least performance error was choosing.

3.8.2 Network generalization

Designer of ANN always faced about degree of his network generalization i.e. although ANN well designer and trained and performance error decrease to the least value; ANN fail when handle with new input data and give worst performance.

To handle with problem and make ANN more general there are two ways:

1- Regularization:

Here ANN was trained until it's performance become un changeable then sum squared error and sum squared weights for network were fixed at last values of it.

2- Early stopping :

Here training data will divided into three groups ,first group for training , the second for validation test, and the third for testing . then error in performance for second group was observed , if ANN loose it validation to handle with new data ; performance error was increase then the training was stopped and weights were fixed in last values before network loose there generalization. The third group is just for testing network in last values of weights for it .

the input vectors and target vectors will be randomly divided, with 70% used for training, 15% for validation and 15% for testing.

To achieve classification of the Diabetes is disease into tow type the number of neurons in the output should be one .logsig which restrict the result between 0 and 1 is used as a transfer function for the first layer

3.9Diabetes Mellitus (DM)

Diabetes Mellitus (DM), literally translated "Urine flow promoted by sweetness" is a long standing or chronic disease which occurs when the pancreas does not produce sufficient insulin, or when the body cannot effectively use the insulin it produces .It is an important and relatively common medical condition and is a risk factor for many other medical conditions likeDevelopment of renal failure, blindness, kidney disease and coronary artery disease are types of the severe damage which are resulted by improper management and late diagnosis of diabetes. stroke, peripheral vascular disease and coronary artery disease (heart attack)[3].

Diabetes Mellitus is according to the World Health Organization there are approximately million people in this world suffering from diabetes. The number of diabetic patients is expected to increase by more than 100% by the year 2030 Diabetes Mellitus patients in this part of the world have been at serious risk and have lost more money in the process of finding out what ailment they are living with, this is due to the enormous processes or procedures involved in diagnosing a diabetes mellitus patient[7].

Common manifestations of diabetes are characterized by insufficient insulin production by pancreas, ineffective use of the insulin produced by the pancreas or hyperglycemia. Causes like obesity, hypertension, elevated cholesterol level, high fat diet and sedentary lifestyle are the common factors that contribute to the prevalence of diabetes.

Diabetes Mellitus is a chronic disease and a major public health challenge worldwide.

According to the International Diabetes Federation (IDF), there were 285 milliondiabetic people in 2010 [15].representing around 7% of the adult populationworldwide. This high prevalence reflects the epidemic nature of diabetes, where thisnumber is expected to rise to 439 million by 2030, with much higher prevalence indeveloping countries[7].

Diabetes is a disease primarily associated with an increase in the level of plasmaglucose (hyperglycemia) caused by insulin deficiency, (type 1 diabetes) formerly known as insulin dependent or childhood onset diabetes,

insulin resistance (type 2 diabetes)known as non-insulin dependent or adult onset diabetes. The wide spread of type 2 diabetes can beattributed to the change of lifestyle, and decreased physical activity levels. Thischange in life style has resulted in higher rates of overweight and obesity, which arekey risk factors for diabetes[7].

Gestational DM which is diabetes mellitus occurring during pregnancy to a woman who was free of DM prior to getting pregnant and some other specific disease conditions which are known to cause DM[6][7].

103Medical practice and literature have shown that Impaired Fasting Glucose (IFG) andImpaired Glucose Tolerance (IGT) represent intermediate states which exist betweennormal glucose homeostasis and diabetes [15].IFG is defined by an elevated FastingPlasma Glucose (FPG) concentration (>=100 and <126 mg/dl), while IGT is definedby an elevated 2-hour plasma glucose concentration (>=140 and <200 mg/dl) after a75g glucose load on the oral glucose tolerance test (OGTT). However, IFG was notfound to be equivalent to IGT. People with IFG not necessarily would also be IGTand the opposite is also true. Therefore, the WHO has decided to use the OGTT test todiagnose the pre diabetes stage. Several studies have also shown that IFG can be arisk factor for IGT, where people with IFG need to have the OGTT test to confirmthat diagnosis of pre diabetes. Current estimates indicate that up to 70% of individualswith these pre-diabetic states (IFG and/or IGT) eventually develop diabetes [7] withinfew years.

Even though there is no established cure for diabetes, indeed, the blood glucose level of diabetic patients can be controlled by well-established treatments, proper nutrition and regular exercise

Chapter Four

The Proposed System

4.1Data Set

When taking dataset of patients, Questionnaire and Interview methods are followed to speed up the process of diagnosis. Moreover, when taking a fairly long term measurement, a reasonable estimation method from the roughly observed level data is very effective

Data collected from 500 individuals (250 diabetic, 250 non-diabetic) attending a private Abuaaqlh Diabetes Center in Wad Medani during the period from December 2014 to January 2015 were used in this work for empirical comparison of the network models.Abuaaglh is a general health center , but more attention and follow-up for patients with diabetes.

Data from 500 subjects of age 0 years and above was collected using aquestionnaire regarding demographic data, history and anthropometric measures.Furthermore, blood pressure and blood samples were analyzed to measure fastingvenous and 2-hour post-glucose load (oral glucose tolerance test (OGTT)) for theparticipants of the survey in health care centers. Other attributes were also collected,but omitted here as they are not relevant to the current study.The data collected about each subject include Age in years (min 0, max 95), Sexcode ((1) male/ (0) female), Family history of diabetes (yes/no), Body mass index(BMI) (min 1, max 120), Systolic blood pressure (min 90,max 220) (BPSYS), Diastolic blood pressure (BPDIAS) (min 50, max 120),Fasting Plasma Glucose (FPG), (min 55,max 400 mg/dl) physical activity(yes or no),pregnancy in females (yes or no), diabetes in the family (yes or no),pedigree of diabetes(type1 , type2),urine sugar (1,2,3)[7] .

The dataset contains 500 samples and two classes.

Class 1 : normal

Class 2 : Diabetes

All samples have 12 features

Feature 1 :

age

Feature 2 :

sex

Feature 3 :

physical activity

Feature 4:

pregnancy

Feature 5 :

diabetes in the family

Feature 6:

body mass index(weight)

Feature 7 :

body mass index(height)

Feature 8 :

diastolic blood pressure

Feature 9 :

systolic blood pressure

Feature 10 :

pedigree of diabetes

Feature 11 :

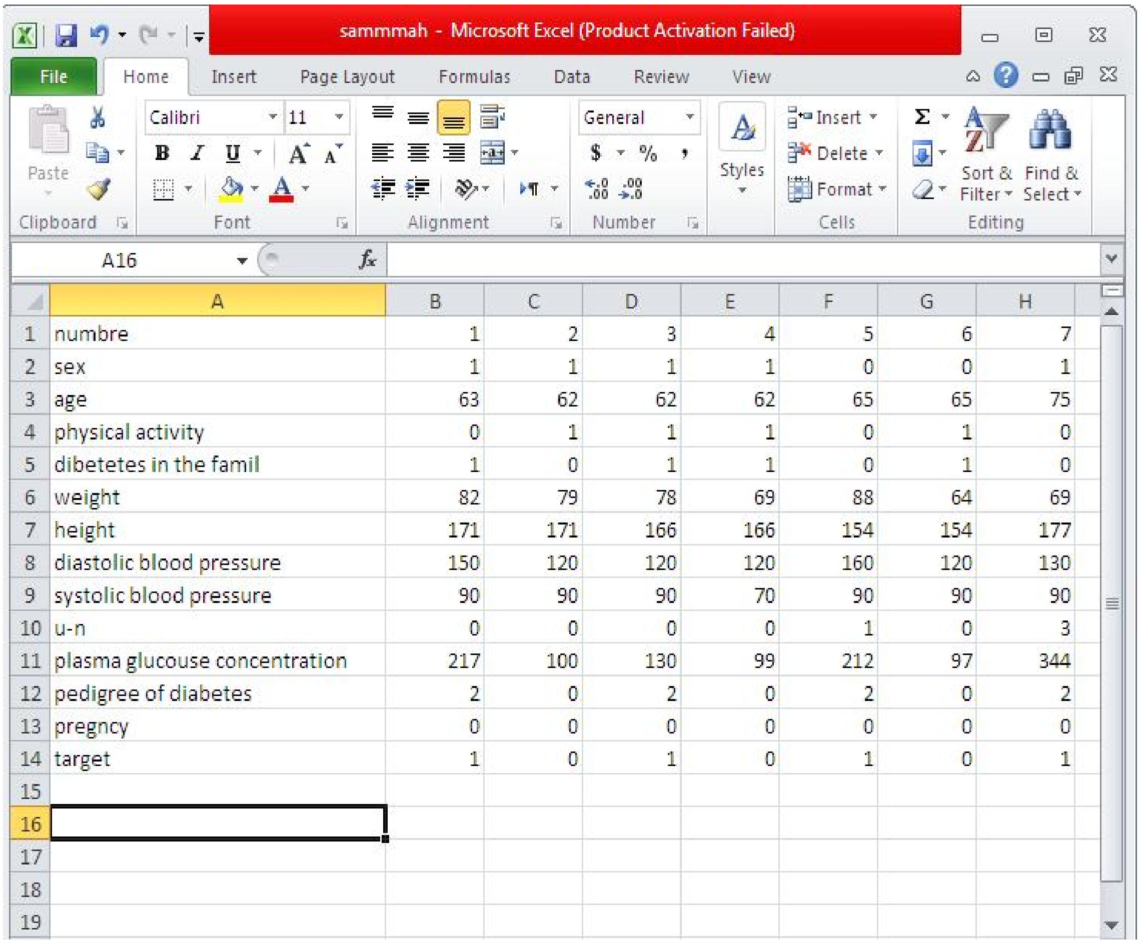
plasma glucose concentration

Feature 12 :

urine sugar.

Include the final research copy from Questionnaire used in studye.

The introduction of data in Microsoft Excel program and so are organized in the form of matrices to facilitate the entry of the program MATLAB



	A	B	C	D	E	F	G	H
1	numbre	1	2	3	4	5	6	7
2	sex	1	1	1	1	0	0	1
3	age	63	62	62	62	65	65	75
4	physical activity	0	1	1	1	0	1	0
5	dibetetes in the famil	1	0	1	1	0	1	0
6	weight	82	79	78	69	88	64	69
7	height	171	171	166	166	154	154	177
8	diastolic blood pressure	150	120	120	120	160	120	130
9	systolic blood pressure	90	90	90	70	90	90	90
10	u-n	0	0	0	0	1	0	3
11	plasma glucouse concentration	217	100	130	99	212	97	344
12	pedigree of diabetes	2	0	2	0	2	0	2
13	pregncy	0	0	0	0	0	0	0
14	target	1	0	1	0	1	0	1

Figure4.1 : get in the data set in Microsoft Excel

The outputs in the ANN model are NORMAL and ABNORMAL. The dataset has 500 observations used as training data. The first 250 data are from normal people and others belong to abnormal people. These samples were taken from real patients of a hospital.

4.2Diagnosing Diabetes mellitus using ANNs

The goal of this study is to design a novel approach for diagnosing diabetes using Artificial Neural Networks. The study aims also identify the best ANNs type which is more suitable for diagnosesdiabetes .so three ANNs was designed which are feed forward back propagation ,Recurrent and Elman network to designed which one has the best performance .

The sixth benchmark problem is a pattern recognition problem. The objective of the network is to decide whether an individual has diabetes, based on personal data (age, number of times pregnant) and the results of medical examinations (e.g., blood pressure, body mass index, result of glucose tolerance test, etc.)[7].

Each entry in the table represents 10 different trials, where different random initial weights are used in each trial. In each case, the network is trained until the squared error is less than 0.05[7].

The conjugate gradient algorithms and resilient backpropagation all provide fast convergence. The results on this problem are consistent with the other pattern recognition problems considered. The RP algorithm works well on all the pattern recognition problems. This is reasonable, because that algorithm was designed to overcome the difficulties caused by training with sigmoid functions, which have very small slopes when operating far from the center point. For pattern recognition problems, you use sigmoid transfer functions in the output layer, and you want the network to operate at the tails of the sigmoid function[7].

The Flowchart used for the design of the three Neural Networks in the study shown in figure 4-1.

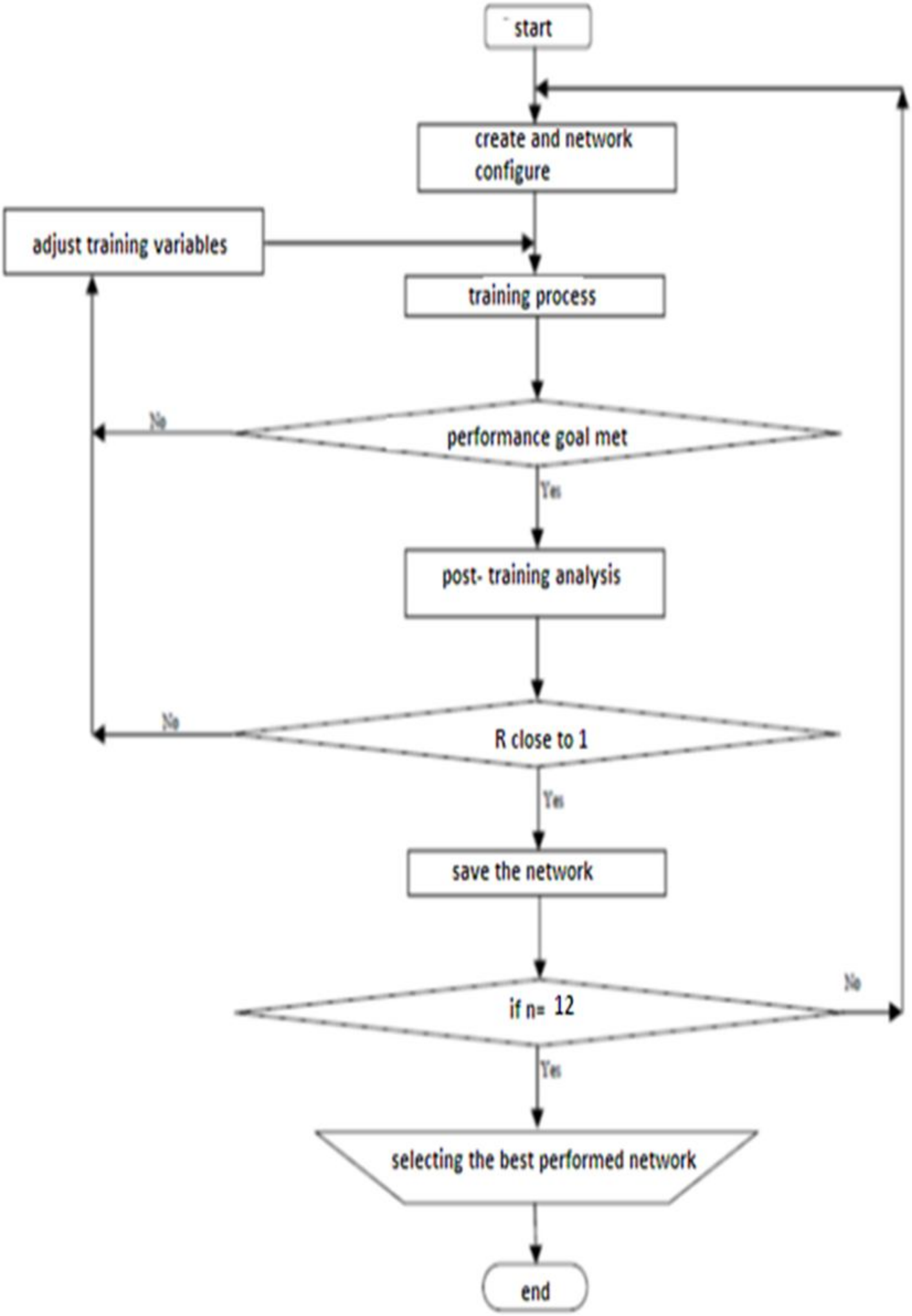


Figure. 4.1: Flowchart used for the design of Neural Network.

4.2.1 Feed-Forward Back Propagation (BP NN) Design

Using nntool in MATLAB BPF was designed and trained ,first one hidden layer between input and output layer with 3,4,5,6,7,8,9 neuron in hidden layer . type of activation function which chooses firstly is log sigmoid[7].

Then for the same network, activation function was replaced first by tan sigmoid function.

4.2.2 Recurrent Network (RNN) Design

As in the case of (BP NN) .

4.2.3 Elman Network Design

As in the case of (BP NN) and (RNN).

Chapter Five

The Results and Desiccation

5.1 The Results

the networks was trained in three ANNs was designed which are feed forward back propagation ,Recurrent and Elman network to designed which is with different number of hidden layer , different training function and different performance as shown in tables5.1 , 5.2 , 5.3 , 5.4 , 5.6 .

Table 5-1 :shown in the Results of elmanNetworks Training with log training function and different number of hidden layer neuron

		Training function	**epoch**	**performance**
Elman	12-3-1	Log	12	5.8503e-09
Elman	12-4-1	Log	10	7.8858e-08
Elman	12-5-1	Log	6	2.0109e-09
Elman	12-6-1	Log	10	6.0592e-09
Elman	12-7-1	Log	14	1.9874e-09
Elman	12-8-1	Log	11	1.0109e-05
Elman	12-9-1	Log	9	0.00078137

Table 5.2 : shown in the Results of elman Networks Training with Tan training function and different number of hidden layer neuron

		Training function	**epoch**	**performance**
Elman	12-3-1	Tan	10	1.2469e-09
Elman	12-4-1	Tan	8	1.18e-08
Elman	12-5-1	Tan	10	0.00083289
Elman	12-6-1	Tan	12	2.4745e-09
Elman	12-7-1	Tan	10	2.524e-09
Elman	12-8-1	Tan	10	0.000997847
Elman	12-9-1	Tan	9	0.011887

Table 5.3 : shown in the Results of RRR Networks Training with Log training function and different number of hidden layer neuron

		Training function	**epoch**	**performance**
RNN	12-3-1	Log	11	3.9547e-09

RNN	12-4-1	Log	15	8.385e-07
RNN	12-5-1	Log	9	9.2635e-09
RNN	12-6-1	Log	10	0.0017471
RNN	12-7-1	Log	10	2.7758e-06
RNN	12-8-1	Log	12	0.00019097
RNN	12-9-1	Log	8	0.0047844

Table 5.4 : shown in the Results of RRR Networks Training with Tan training function and different number of hidden layer neuron

		Training function	epoch	performance
RNN	12-3-1	Tan	14	0.013214
RNN	12-4-1	Tan	9	5.6475e-09
RNN	12-5-1	Tan	11	0.01514
RNN	12-6-1	Tan	13	1.8845e-10
RNN	12-7-1	Tan	10	1.1816e-09
RNN	12-8-1	Tan	8	0.013022
RNN	12-9-1	Tan	11	1.8242e-09

Table 5.5 : shown in the Results of BP NN Networks Training with Log training function and different number of hidden layer neuron

		Training function	epoch	performance
BP NN	12-3-1	Log	12	5.9346e-09
BP NN	12-4-1	Log	13	4.8434e-07
BP NN	12-5-1	Log	7	0.0028903
BP NN	12-6-1	Log	10	5.6707e-09
BP NN	12-7-1	Log	8	7.6156e-08
BP NN	12-8-1	Log	10	5.5391e-09
BP NN	12-9-1	Log	10	0.001472

Table 5.6 : shown in the Results of BP NN Networks Training with Tan training function and different number of hidden layer neuron

		Training function	epoch	performance
BP NN	12-3-1	Tan	14	0.012954
BP NN	12-4-1	Tan	13	7.8861e-07
BP NN	12-5-1	Tan	10	3.0584e-09
BP NN	12-6-1	Tan	9	2.4274e-09
BP NN	12-7-1	Tan	13	0.00011178

BP NN	12-8-1	Tan	11	0.0011952
BP NN	12-9-1	Tan	9	0.8056e-09

5.2 TheDesiccation

The property training record best_epoch indicates the iteration at which the validation performance reached a minimum.

The validation and test curves are very similar. If the test curve had increased significantly before the validation curve increased, then it is possible that some over fitting might have occurred[4].

The validating the network is to create a regression plot, which shows the relationship between the outputs of the network and the targets.

If the training were perfect, the network outputs and the targets would be exactly equal, but the relationship is rarely perfect in practice.

The best of training algorithm it is the Lavemberg-Marquardt because The results indicate Quick Algorithm as the most accurate algorithm since the value of correlation and R-squared close to 1[4].

The following figure shows the result. The result here is reasonable, because the test set error and the validation set error have similar characteristics, and it doesn't appear that any significant overfitting has occurred. The next step is to perform some analysis of the network response. Put the entire data set through the network (training, validation, and test) and perform a linear regression between the network outputs and the corresponding targets. First,and calculate the network outputs[4].

5.3Theresults obtainedin the study

The result obtained that amount the three neuralnetwork.

1-elman network

2- BPNN network

3-RNN network

1- theelman network with 5neuron in the hidden layer with TANsigmoid activation function has performance 0.00083289 .

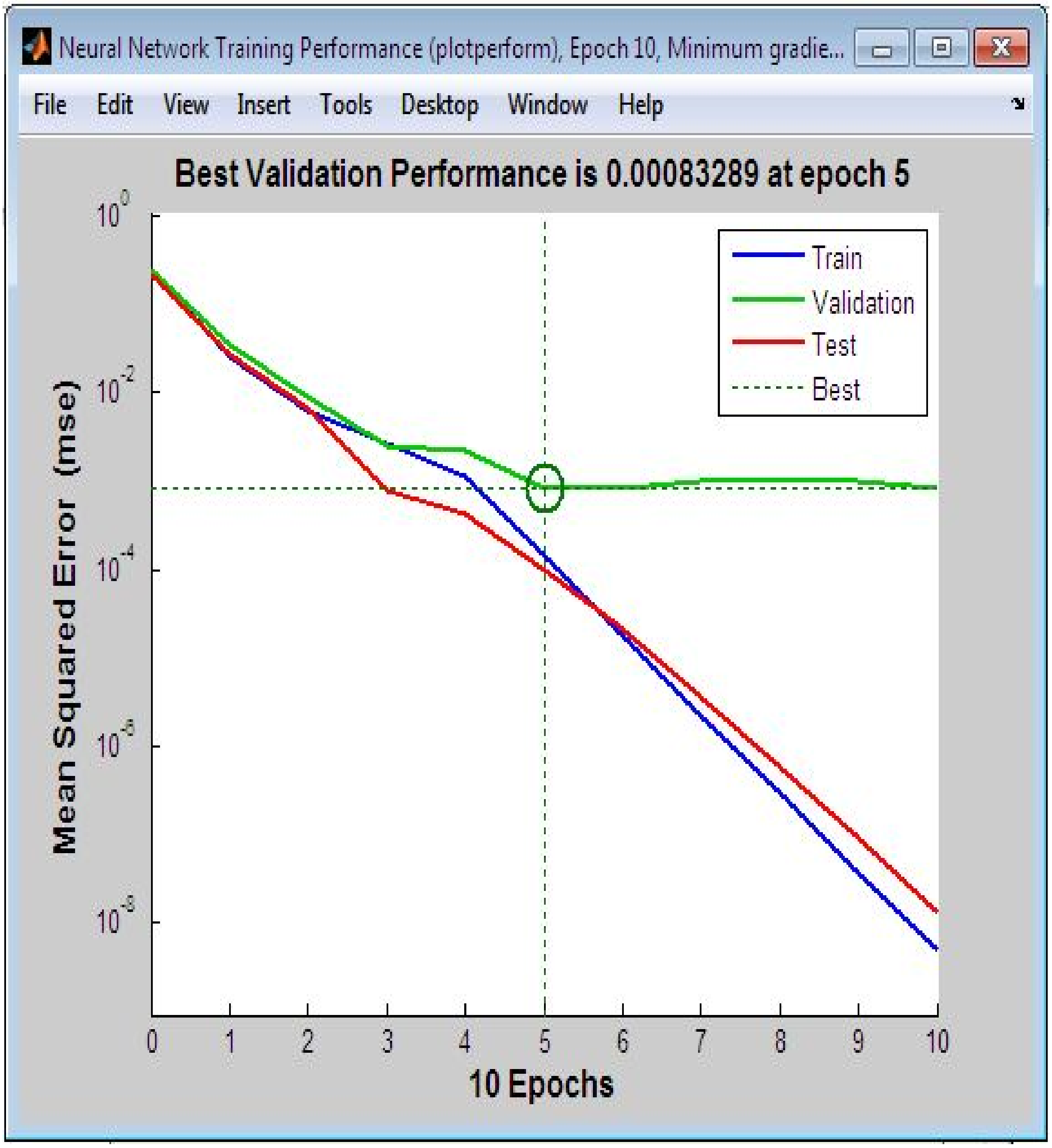


Figure5.2 : training result for elman with 5 neuron and TANsigmoid activation function.

2- For the BPNN with 7 neuron in the hidden layer with tan sigmoid activation function has performance 0.00011178

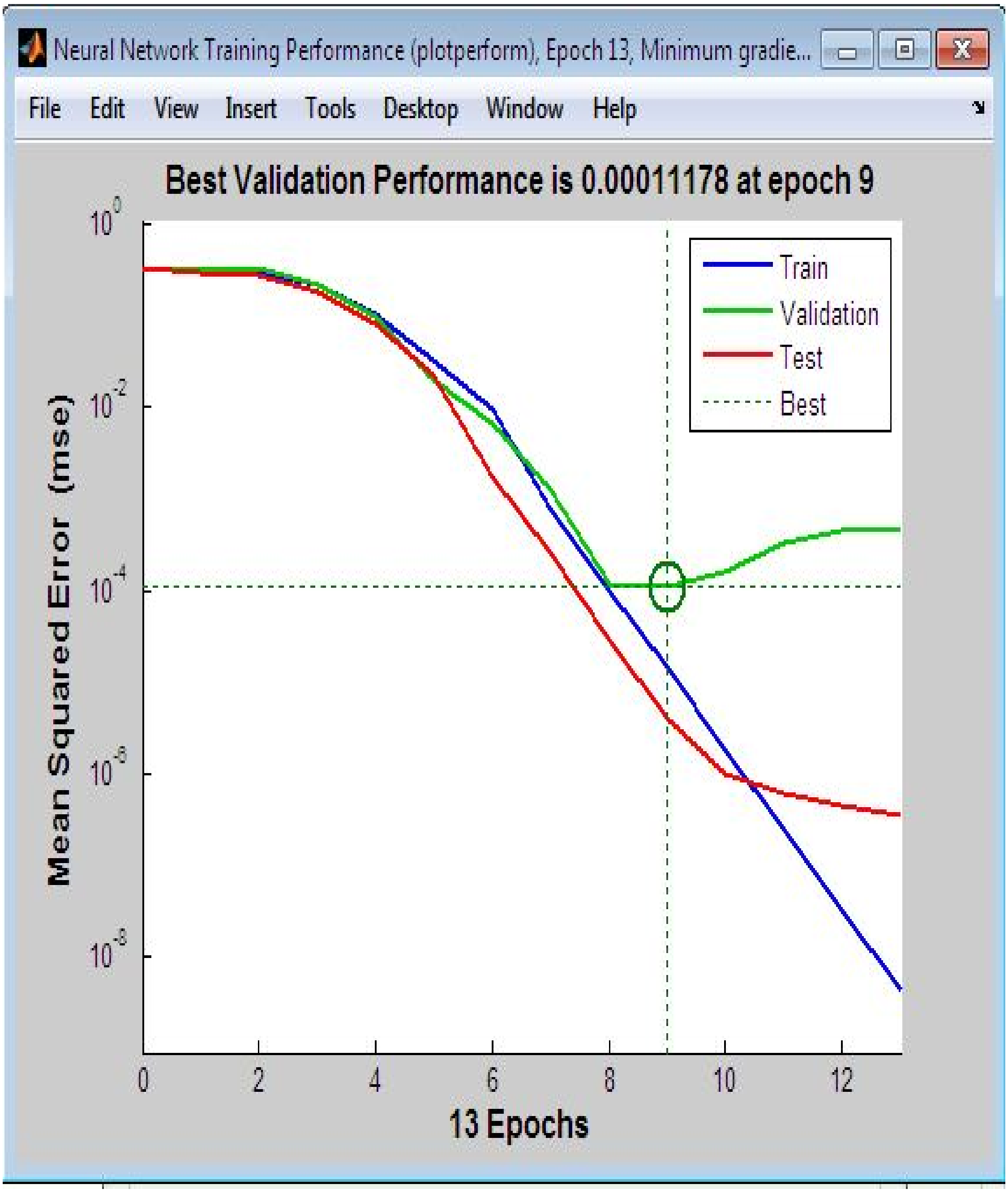


Figure5-4 : training result for BP NN with 5 neuron and TAN sigmoid activation function.

3- For the RNN with 8 neuron in the hidden layer with LOG sigmoid activation functionhas performance 0.00019097.

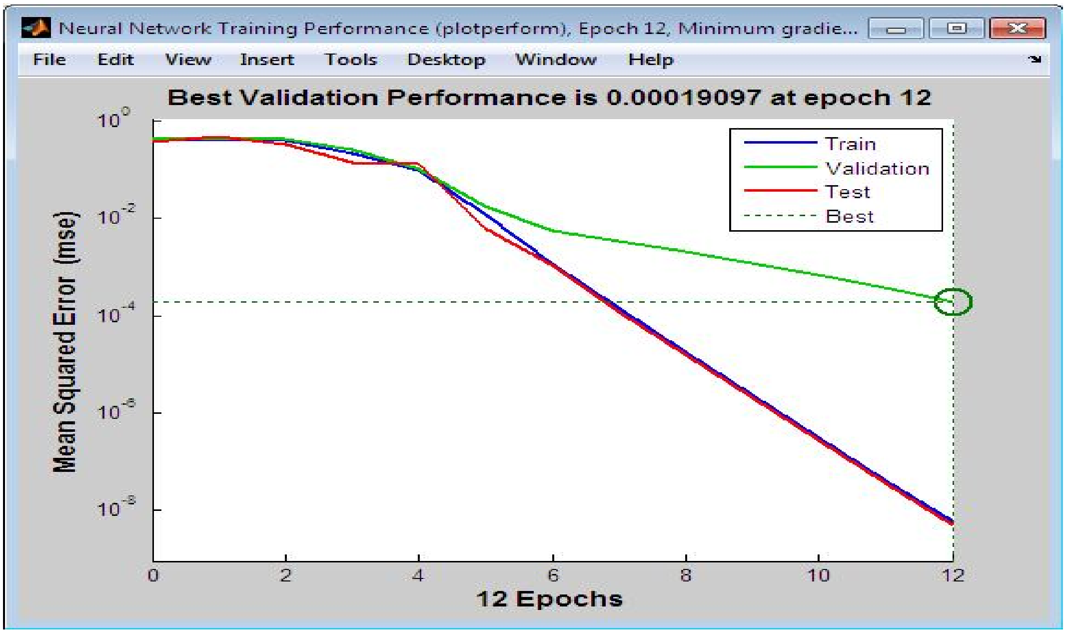


Figure5-6 : training result for RNN with 8 neuron and LOG sigmoid activation function.

This table shown in The result obtained that amount the three neural network.

Table5.7: the result obtained that amount the three neural network.

	BPNN	RNN	ELMAN
Best of Performance	0.00011178	0.00019097	0.00083289
Number of neuron in hidden layer	7	8	5
Activation function	TAN	LOG	TAN

For the BPNN with 7 neuron in the hidden layer with TAN sigmoid activation function has performance 0.00011178 that is best performance network in this study

Chapter six

Conclusion and Recommendation

6.1 Conclusion

The goal of this study is to Design a novel approach for diagnosing Diabetes using Artificial neural networks. The study aims also identify the best ANNs type which is more suitable for Diagnoses Diabetes .so three ANNs was designed which are feed forward back propagation ,Recurrent and Elman network to designed which one has the best performance .

Using MATLAB BPF was designed and trained in the BB NN , RNN, and elman network ,first one hidden layer between input and output layer with changed about (3,4,5,6,7,8,9) neuron in hidden layer . type of activation function which chooses firstly is log sigmoid and tan sigmoid function.

The result obtained that amount the three neural network.

1- theelman network with 5 neuron in the hidden layer with TAN sigmoid activation function has performance 0.000832892- For the BPNN with 7 neuron in the hidden layer with TAN sigmoid activation function has performance 0.00011178 that is best performance network in this study. 3- For the RNN with 8 neuron in the hidden layer with LOG sigmoid activation function has performance 0.00019097.

6.2 Recommendation

The recummention of this work are :

1- Competition neurofuzzy will be recommended
2- The receiving of the aspirations of this work linked to telemedicine or other communication techniques.
3- after the relative success in the diagnosis of diabetes and early detection to him, according to his competence by high it is recommending its use as a system are reliable in all health facilities, center professional diagnosis of diabetes in Sudan
4- Finally, due to difficulty of access the special databased of patient diabetes, so I recommended that full documentation and archive of the all patient data, whose re-visit hospitals and utilities health centers frequently, to assist future research.

References

[1] - B.Y.Baha ,B.Yola and etc... , Artifical Neural Networks To DetectecRisk Of Type 2 Diabetes , June, 2012.

[2] – F.Amato, A.López and etc .. , Artificial neural networks in medical diagnosis. Received 17th December 2012.Published online 7th January 2013.

[3] - A.B.Adeyemo and A.E.Akinwonmi. On the Diagnosis of Diabetes Mellitus Using Artificial Neural Network Models. African Journal of Computing & ICT September 2011.

[4] - H. Demuth, M. Beale and etc... ,Neural Network Toolbox™ 6 ,2009

[5] - T. Abdelrazig .Y.Gabersogail and etc… ,Design and implementation of a custom Arrhythmias classification system

[6] - D .M .Mahmoud , Brain tumor recognition using artificial neural network and fussy systems, 2012 .

[7] – N.Barakat and M.N.Barakat, Hybrid Intelligent Model for the Diagnosis of Impaired Glucose Tolerance

[8] - Og˘uz.Karan ,C.Bayraktar and etc .. , Diagnosing diabetes using neural networks on small mobile devices , 2011.

[9] – M.Pradhan and Dr. Ranjit Kumar Sahu. Predict the onset of diabetes disease using Artificial Neural Network (ANN). International Journal of Computer Science & Emerging Technologies (E-ISSN: 2044-6004) 303 Volume 2, Issue 2, April 2011.

[10] - P.Venkatesan and S.Anitha Application of a radial basis function neural network for diagnosis of diabetes mellitus , CURRENT SCIENCE, VOL. 91, NO. 9, 10 NOVEMBER 2006.

[11] - S.Kabilan, S.Kannan, and etc... , Diagnosis of Diabetes Mellitus Type 2 .International Journal of Emerging Technology and Advanced Engineering. Volume 4, Issue 2, February 2014.

[12] - O.Wahyunggoro, A.E.Permanasari, and etc..., Utilization of Neural Network for Disease Forecasting ,2012.

[13] - K.Rajeswari and V.Vaithiyanathan , Fuzzy based modeling for diabetic diagnostic decision support using Artificial Neural Network , April 2011.

[14] - H. Demuth, M. Beale and etc... ,Neural Network Toolbox™ 6 ,2009

[15] م- علامزكيعيسى ،الشبكاتالعصبية (البنيهالهندسية ، الخواروميات ، التطبيقات)2000- of a custom Arrhythmias classification system , 2014.