COMPARING THE IMPACT OF ACCURATE INPUTS ON NEURAL NETWORKS

V.Vaithiyanathan[1], K.Rajeswari[2], N.Nivethitha[3], Pa.Shreeranjani[4], G.B.Venkatraman[5],
M. Ifjaz Ahmed[6].

[1]Associate Dean - Research, School of Computing, SASTRA University,Thanjavur, Tamilnadu , India
[2]Research Scholar, SASTRA University, Thanjavur, TamilNadu , India & Assistant Professor, Pimpri Chinchwad College of Engineering, University of Pune, Maharashtra, India.
[3,4] Student, School of Computing, SASTRA University, Thanjavur, Tamilnadu, India.
[5]System Administrator, School of Computing, SASTRA University,Thanjavur, Tamilnadu, India.
[6]Research Scholar, School of Computing, SASTRA University,Thanjavur, Tamilnadu, India.
vvn@it.sastra.edu, raji.pccoe@gmail.com, nivethitha.mns@gmail.com, shreeparthasarathy@gmail.com,
gbvram@gmail.com, ifjazahmed@sastra.edu.

***Abstract* - Artificial neural networks are widely used in medical diagnosis replacing most of the conventional diagnosis methods due to its accuracy and speed. This paper analyses the variation in the accuracy of diagnosis of type II diabetes using Artificial Neural Networks based on the accuracy of the inputs given to the network. It compares the efficiency of the network based on the input format. The data needed for this comparison is collected by interviewing patients who approach the diabetician with various symptoms of the disease. These symptoms can be modeled in 2 different forms. One form just specifies the presence or absence of the symptom and can be represented using Boolean values. The other form specifies the severity or frequency of occurrence of the symptom. Both these inputs are given to the system and the accuracy of the output is analyzed. This result indicates the impact of the specification of the input on the output. Comparison is done by performing regression analysis on both the outputs. Regression analysis gives the correlation between the output of the system and the target [1]. It makes use of only the most general symptoms of the disease. Further analysis can be done on other diabetes particular symptoms**.

Keywords: Diabetes, Artificial Neural Networks, Feed forward neural networks, regression.

I. INTRODUCTION

Classification is the most important part of medical diagnosis as it rightly identifies the disease and leads to providing proper treatment. The multilayer neural networks (MLNNs) have been successfully used in replacing conventional pattern recognition methods for the disease diagnosis systems [5].Many studies have been undertaken by researchers to improve the efficiency of neural networks which are widely used in diagnosis. The results obtained showed that for any system to produce the best output, the input given must be as accurate as possible. Different types of neural networks can be used, but feed forward neural networks are widely recognized as the most efficient networks in medical diagnosis. Efficient knowledge acquisition and representation are one of the central challenges for the successful construction and following use of medical-expert and knowledge-based systems in clinical practice [4].

II. RELATED WORK

A. Diabetes Mellitus:

It is the most common endocrine disorder. It is a group of metabolic disorders of carbohydrate metabolism characterized by glucose underutilization and hyperglycemia. Diabetes is a major health problem in both industrial and developing countries and its incidence is rising. It is a disease in which the body does not produce or properly use insulin, the hormone which “unlocks” the cells of the body, allowing glucose to enter the cells and fuel them [3]. This state is also called “insulin resistance” and is the reason for Diabetes type II. The most common form of Diabetes is Type 2 Diabetes [6]. Diabetes increases the risk of developing kidney disease, blindness, nerve damage, blood vessel damage and it contributes to heart disease [7]. So the correct and early diagnosis of Diabetes is of paramount importance. Proper representation and interpretation of data is important for medical classification and is the main concern of this paper.

B. Artificial Neural Networks:

Artificial Neural Networks are inspired from the biological neurons. The network consists of multiple layers of neurons connected with each other. In general there is one input layer, one or more hidden layers and one output layers. Each neuron in the input layer is connected with every neuron in the hidden layer and each neuron in the hidden layer is connected with every neuron in the output layer. The neurons process the inputs and produce the output based on the transfer function and the weights on the interconnection.

The three main parameters which determine the performance of the neural network are:

- The way in which the neurons are connected with each other.
- The learning methodology.
- The transfer function which converts the neuron input to the processed output.

The most commonly used learning methodology in medical diagnosis is supervised learning. In this, the network is provided with both the inputs and the corresponding output. The network adjusts its weight till the error is minimized and the output of the network matches with the target. Pattern recognition and regression are the most common applications of this network.

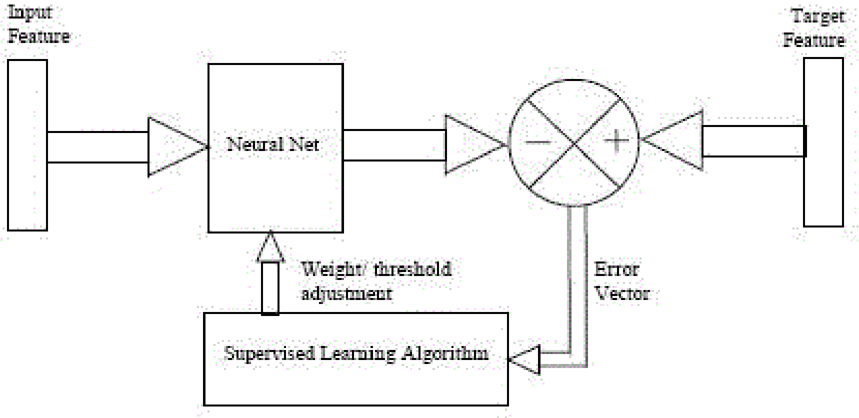


Fig 1: Artificial neural network using supervised learning.

C. Feed forward Neural network:

This network makes use of the supervised learning algorithm. The Backpropogation algorithm is the most commonly used algorithm in these networks. The network signals travel in the forward direction and the errors travel backwards. The network weights are initially assigned random values and then they are adjusted to obtain the desired output. Feed forward neural networks play a prominent role in the field of medical diagnosis compared to the other artificial neural networks.

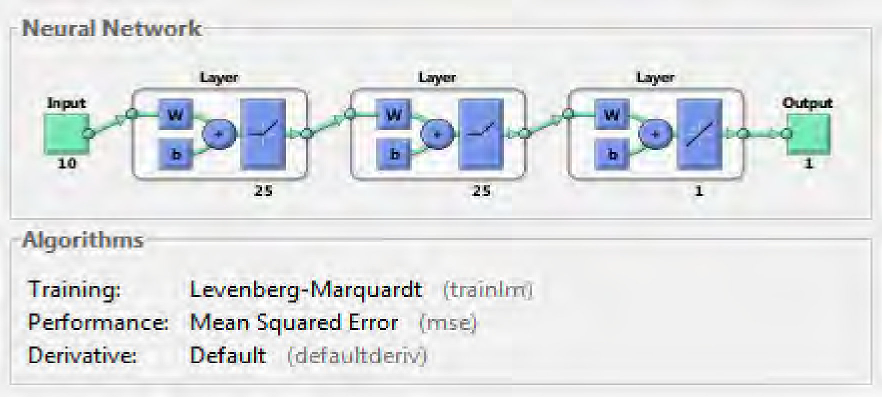
III. METHOD

A. Data preparation:

The dataset was collected from Diabetic Care and Research Centre, Sivapreethi Hospital, Tanjore, TamilNadu. The dataset also includes the records of some patients who were not diagnosed with the disease. When a doctor interrogates a patient about the symptoms of the disease it can be a simple “yes or no”, wherein the frequency of the occurrence of the symptom is not mentioned properly and may lead to a diagnosis error. When the severity of the symptoms are specified based on the frequency per day, or based on Likert scale where the frequency cannot be considered on a daily basis, the chances of a diagnosis error are reduced. A Likert scale is a psychometric scale which is used when the study involves data collection using questionnaires. It is used to specify the frequency on a scale of 10 or so for research purpose. The data was collected with extreme care so that proper analysis can be carried out.

B. Network structure: The network used in this comparative study consists of an input layer, two hidden layers and one output layer. The hidden layers makes use of “poslin” transfer function and the output layer makes use of the “purelin” transfer function.

- The poslin transfer function returns the same value if the input is greater than 0 else it returns 0.
  Poslin(n) = n, if $n > 0$.
  = 0, if $n < 0$.
- Purelin is a linear transfer function which returns the input as the output.
  Purelin(n) = n.



Training: Levenberg-Marquardt (trainlm)
Performance: Mean Squared Error (mse)
Derivative: Default (defaultderiv)

Fig 2: Neural network with two hidden layers.

C. Training algorithm:

“Trainlm” is the training algorithm used to train the network to achieve the desired output. It updates the weights and the biases based on the Levenberg-Marquardt optimization and is the most preferred training algorithm as it has the fastest convergence and high accuracy. LM algorithm can provide better generalization performance compared to the other algorithms. However, the requirement of high computer memory and longer time during training has limited the application of this algorithm for practice. Therefore, in order to apply this algorithm, a balance is always needed between the size of the ANN model and the selection of learning algorithm[3].

“trainlm” can train any network as long as its weight, net input, and transfer functions have derivative functions.

D. Algorithm:

Backpropagation is used to calculate the Jacobian jX of performance perf with respect to the weight and bias variables X. Each variable is adjusted according to Levenberg-Marquardt,

$$jj = jX * jX$$

$$je = jX * E$$

$$dX = -(jj+I*mu) \backslash je$$

where E is all errors and I is the identity matrix. In general the metric used to estimate the network performance is the mean-squared error.

IV. EXPERIMENTAL RESULTS

The table below shows the input symptom values for 10 patients. The 10 symptoms used are polyurea, polyphagia, polydipsia, nocturia, tiredness, giddiness, non-healing ulcer, sleeplessness, itching and shoulder pain.

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
P1	0	0	1	0	0	0	0	0	0	0
P2	1	1	1	1	1	0	0	0	0	1
P3	0	0	0	0	0	0	0	1	0	0
P4	1	1	1	1	1	1	0	0	0	1
P5	0	0	0	0	1	0	0	1	0	0
P6	1	0	1	1	1	1	0	1	1	0
P7	1	1	1	1	1	0	0	0	0	1
P8	0	0	0	0	0	0	0	0	0	0
P9	1	0	1	1	1	0	0	0	0	1
P10	0	0	0	0	0	0	0	0	0	0

Table I: Patient Vs Symptom matrix without specifying severity of the symptom

In the above table the value 0 indicates the absence of the symptom and 1 indicates the presence of the symptom.

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
P1	4	2	8	2	2	2	4	4	4	2
P2	9	4	5	4	6	1	2	2	2	6
P3	7	3	1	2	4	3	5	8	3	3
P4	11	4	6	5	7	6	3	3	1	7
P5	8	3	2	1	6	2	5	7	4	2
P6	15	2	7	7	6	7	4	9	6	1
P7	11	4	5	4	5	2	2	4	2	6
P8	6	3	3	2	3	1	3	2	4	1
P9	10	2	9	7	6	4	4	3	3	8
P10	3	3	4	2	4	1	5	5	2	4

Table II: Patient Vs Symptom matrix specifying the severity of the symptom.

In the above table the first 4 symptom values are based on the frequency of occurrence of the symptoms per day and the remaining symptom values are specified based on the Likert scale.

A. Regression Analysis:

Regression analysis is performed to measure the system performance. It indicates the correlation between the system output and the target of the system. The regression value (R) of 1 indicates the maximum correlation and a regression value of 0 indicates the minimum correlation between the output and the target. The regression value is also called the correlation coefficient.

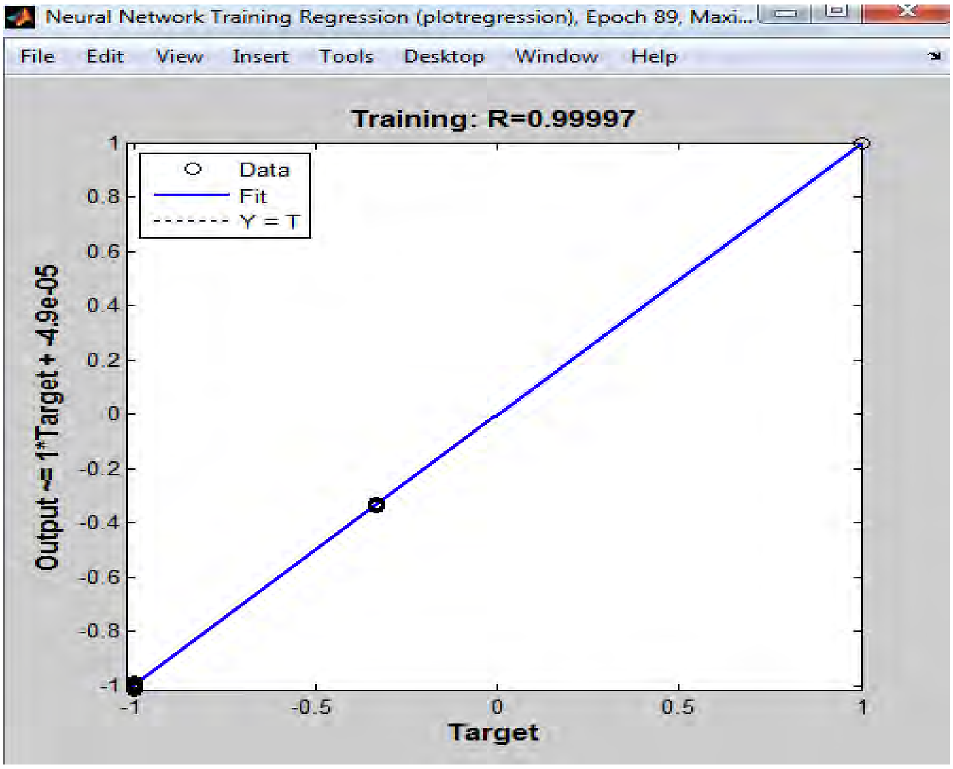


Fig 3: Regression plot for input specified as in Table 1.

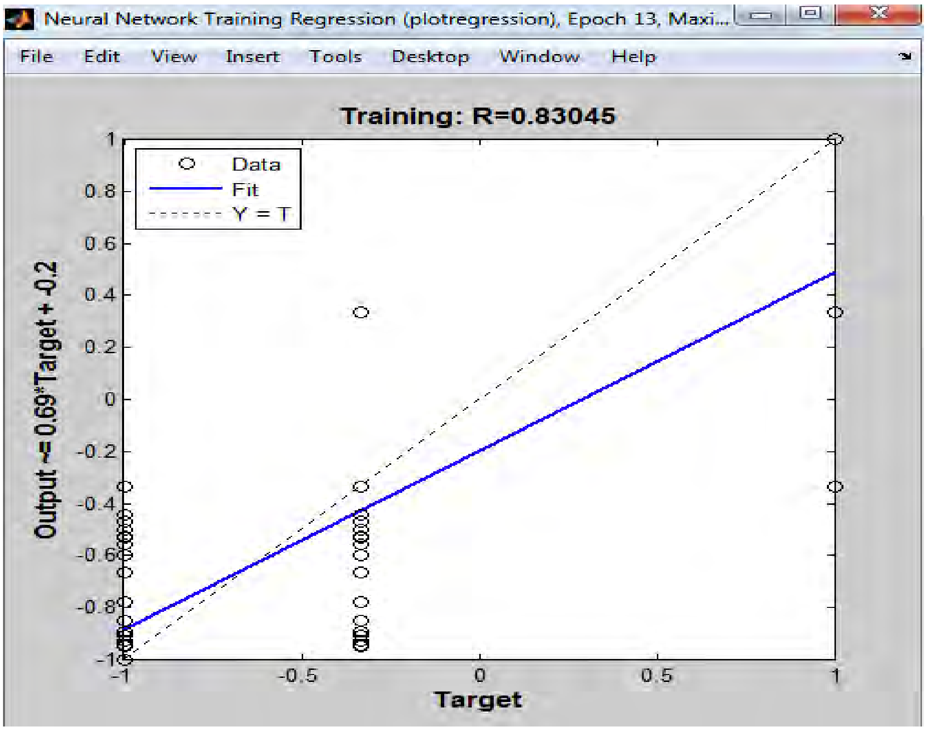


Fig 4: Regression plot for input specified as in Table 2.

In the above regression plots, the dashed line indicates the best fit produced by the algorithm and the solid line indicates the obtained output, and a perfect fit of the solid over the dashed line indicates the perfect output [1]. So from the regression plots it can be seen that the system performance improves to a great extent when the input to the network is more precise and accurate. The correlation coefficient of Figure 4 is higher than that of Figure 3 which emphasizes the same.

	Epochs	Correlation Coefficient (R)
Input not specifying severity	13	0.83046
Input specifying severity	89	0.99997

Table III : Network simulation parameters

V. CONCLUSION

This study analyses the improvement in the efficiency of the system performance based on the input accuracy. From the regression plots obtained and from Table 3 it can be seen that the first input set achieves a correlation value of 0.83046 which is less compared to the value achieved by the second set which is 0.99997. Hence the network performance improves for more accurate input. In future, analysis can be done on the symptoms used in the diagnosis of the disease.

REFERENCES

[1] Muhammad Akmal Sapon , Khadijah Ismail, Suehazlyn Zainudin and Chew Sue Ping "Diabetes Prediction with Supervised Learning Algorithms of Artificial Neural Networks" *2011 International Conference on Software and Computer Applications IPCSIT vol.9 (2011) © (2011) IACSIT Press, Singapore*.

[2] Goh Lyn Dee**,** Norhisham Bakhary, Azlan Abdul Rahman and Baderul Hisham Ahmad "A Comparison of Artificial Neural Network Learning Algorithms for Vibration-Based Damage Detection" *Advanced Materials Research Vols. 163-167 ,pp 2756-2760, (2011).*

[3] M Mohamed, E. I., Linderm, R., Perriello, G., Di Daniele, N., Poppl, S. J., and DeLorenzo, A. "Predicting type 2 diabetes using an electronic nose-base artificial neural network analysis". *Diabetes Nutrition & Metabolism, 15(4), 2002.*

[4] Michael Schuerz , Klaus-peter Adlassnig , Charles Lagor , Barbara Scheider and Georg Grabner, "Definition of Fuzzy Sets Representing Medical Concepts and Acquisition of Fuzzy Relationships Between Them by Semi-Automatic Procedures"

[5] K.Rajeswari and V.Vaithiyanathan "Fuzzy based modeling for diabetic diagnostic decision support using Artificial Neural Network", *IJCSNS International Journal of Computer S cience and Network Security, VOL.11 No.4, April 2011.*

[6] Acharya, U. R., Tan, P. H., Subramaniam, T., et al. "Automated identification of diabetic type 2 subjects with and without neuropathy using wavelet transform on pedobarograph". *Journal of Medical Systems, 32(1), pp 21–29, (2008).*

[7] Hasan Temurtas, Nejat Yumusak, Feyzullah Temurtas, "A comparative study on diabetes disease diagnosis using neural networks", *Journal Expert Systems with Applications:An International Journal Volume 36 Issue 4, May, 2009.*
http://www.learnartificialneuralnetworks.com

[8] http://www.wikipedia.com [Artificial Neural Networks].