Journal of Pharmacy and Pharmacology 11 (2023) 45-55
doi: 10.17265/2328-2150/2023.03.001

Comparison of Different Types of Artificial Neural Networks for Diagnosing Thyroid Disease

Rehab Abdulmunem Ali Alshireef
Computer Sciences Department, Libyan Academy, Libya

Abstract: Thyroid disease is one of major causes of severe medical problems for human beings. Therefore, proper diagnosis of thyroid disease is considered as an important issue to determine treatment for patients. In this paper, three methods of using neural network to achieve high precision are compared. The first one is using a multilayer feed forward architecture of artificial neural network (ANN) is adopted in the proposed design, and the back propagation is selected as learning algorithm to accomplish the training process, with 3 inputs, 5 nodes of hidden layer and 8 output but only one is active with accuracy 99.2%. Second one will be a Multi-layer Perceptron (MLP) ANN using back propagation learning algorithm to classify Thyroid disease. It consists of an input layer with 5 neurons, a hidden layer with 6 neurons and an output layer with just 1 neuron with accuracy 98.6%. Third one a multilayer feed forward network with Genetic Algorithm the input layer with 5 neurons equal to the number of the dataset features, 1 hidden layer which its neurons will be determined by the GA and it's 4 gene, and the output layer with only 1 neuron, the overall accuracy is 100% for training and in range between 96% and 98% for testing.

Key words: Artificial Neural Network, Thyroids disease.

1. Introduction

In recent years, artificial intelligence techniques are exploited for developing professional systems to diagnose different kinds of diseases with high accuracy. These systems assist staff in hospitals and medical centers to quickly diagnose patients and give them essential treatments without need for a specialist doctor. As a result, these systems decrease cost and time for diagnosis. Artificial Neural Network is the most important artificial intelligence technique that has been used to design diagnostic system for several diseases such as diabetes, heart disease, breast cancer, skin disease, and thyroid disease [1] and those are computing tools like NNs, Fuzzy Logic and Genetic algorithm. ANNs due to advantages such as self-learning, associative memory, high parallelism strength and high speed and error tolerance against noises which might be in parameters and also their cheapness in reuse of available solutions is the best option to do this [2]. The ANN called connection-oriented networks which include a set of processors act as parallel, take the sets of input in a time and produce output based on processing algorithm [3]. Artificial Neural Network (ANN) is a powerful tool to solve classification and pattern recognition problems [4]. Hence, ANN has been used to identify the type of thyroid disease. In recent years, there are many researches that have been focused in this area. Most of them used the data set from University of California, Irvine (UCI) repository of machine learning database [5]. In this data set, thyroid cases are only classified to three main types: euthyroidism (normal case), hyperthyroidism, and hypothyroidism.

The thyroid gland is one of the most important organs in the human body. It produces two active hormones that are responsible for controlling metabolism, production of proteins, regulation of body temperature, and overall energy production [6]. Therefore, proper operation of thyroid gland is essential for every organ in the human body. However,

Corresponding author: Rehab Abdulmunem Ali Alshireef. Email: rehabalshireef@gmail.com.

thyroid diseases commonly occur and lead to produce too much or too little hormones. Severe case of thyroid disorders may lead to death. Hence, correct and fast diagnosis of thyroid diseases is necessary to provide patients the required cures in early stage of diseases.

In this paper, it is used MATLAB, a software due to the flexibility, many prepared functions and proper performance to ANN simulation. This software also causes the accuracy of results and comparing the performance of the networks become increased and it is achieved the best performance to learning network in the shortest possible time. In section 2 we will show the Previous works, section 3 the overview of ANN, in section 4 describe the data set, in section 5 training and testing set, in section 6 the evaluation and discussion finally the conclusion in section 7.

2. Previous Works

By studying literature, it seems that most researchers have been used widely ANN techniques to diagnose Thyroid disease. Most used techniques are under-supervision techniques in this feature in which the researchers increase the speed and accuracy of network by applying network architecture, proper Initialized of weights and choosing proper activating function due to the type of their diagnoses.

For example, Dey et al [1] has been used ANN techniques to diagnose Diabetes disease. The applied data in this paper are taken from Manipal Sikkim Institution of Medial Science Hospital which includes 530 patients. The output includes 2 classes of 0 and 1. They suggest two feed forward architectures in which the first one includes the number of neurons in three layers as (6-10-1) and the second involves two hidden layers and the number of neurons in (6-14-14-1) layers. They indicate that log sigmoid activating function in the hidden layer neuron is considerably increase the speed and performance of networks. So, if the number of the layers increases, the abstract error will be increased. The accuracy of performance in this network for the architecture with a single hidden layer is 92.5%.

Kabari et al [7] provide a framework to diagnose skin disease using MLP ANN and can reach to an acceptable level of accuracy. In this paper, researchers have been used 4 neurons in input layer, 3 neurons in hidden layer and 2 neurons in output layer, respectively to reach the accuracy to 90%.

Gharehchopogh, F. S., et al [8] have been used ANN to diagnose heart disease. Their architecture includes MLP ANN with 60 nodes in input layer, 4 nodes in hidden layer and 2 nodes in output layer. It is back propagation learning algorithm for network learning. The value of assigned parameters for rate learning and momentum are considered 0.2 and 0.3 respectively. The achieved accuracy of performance for training and test set is 0.95% and 0.85%, respectively.

Shukla et al [9] could achieve to the acceptable results using ANN techniques to diagnose Thyroid disease. They use 3 architectures of back propagation learning algorithm, RBF network and Learning Vector Quantization Networks. The number of layers for back propagation learning algorithm is considered with a single hidden layer and the number of neuron in hidden layer is 46 and learning rate is equal to 0.8. RBF network has a single hidden layer, learning rate of 0.9 and the number of neuron in hidden layer is 100. Comparing the performance of these three architecture networks indicate that LVQ network has the best accuracy rate equal to 0.98%. Because LVQ like back propagation learning algorithm doesn’t fall into the local minimum trap as well as RBFN network which doesn’t require full cover of input space. But, among three applied ANNs in this paper, RBFN network has the least learning time.

Isa et al [10] have been used ANN to diagnose Thyroid disease. By selecting the proper activating function, they could improve the performance of ANN. The given data related to the data set of UCI site is used 215 data, MLP architecture, sigmoid activation function and a hidden layer with 7 neurons and reached to the accuracy rate of 97.6%. They also indicate that hyperbolic tangent function of MLP ANN is suitable to classify data to two classes and neural function for three classes.

3. Overview of Artificial Neural Network

Artificial neural network (ANN) is a well-known artificial intelligent technique for solving problems that are difficult to be solved by human beings or conventional computational algorithms [11].

3.1 MLF with Back Propagation Algorithm

A multilayer feed forward ANN is adopted to do the classification problem of thyroid cases because it is the most popular structure of ANN that is used for classification and pattern recognition problem [12].

In the data set, each sample is a vector of three entries (values of TSH, T4, and T3). Therefore, the designed ANN should have three neurons in the input layer. In contrast, the designed ANN should classify input samples into eight categories of thyroid cases. Thus, in the output layer, ANN should have eight neurons which are corresponding to these categories. As a result, inputs of ANN system are values of TSH, T4, and T3 for a certain sample, and output is the type of thyroid case for input sample as shown in Figure 1.

Back propagation is used as a learning algorithm to train ANN. At first, weights are initialized with random values, then at each iteration of back propagation algorithm. One input sample is applied to ANN to produce the actual output. After that, the error is computed between the actual output and desired output. Depending on this error, the synaptic weights are updated as Equation (1) to minimize error.

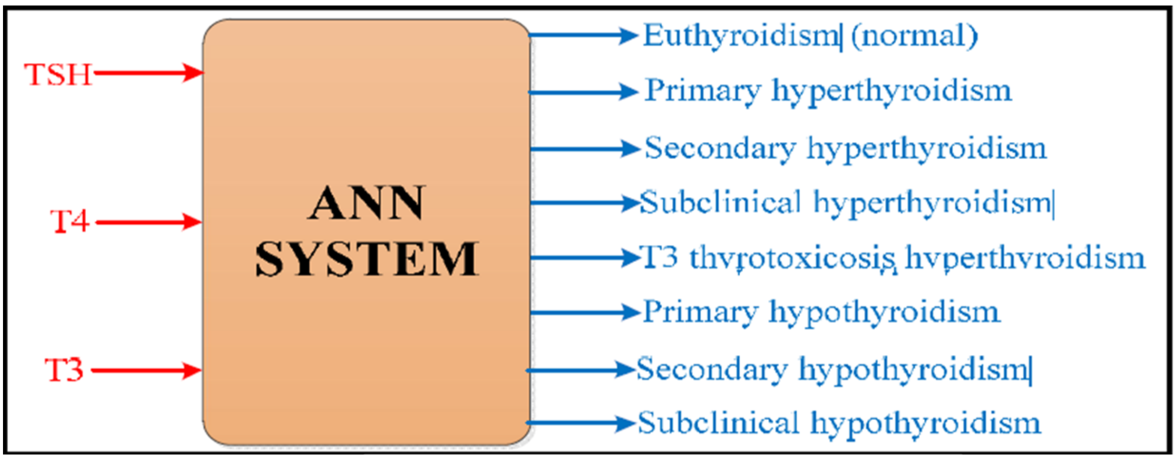

Fig. 1 The ANN system for thyroid diagnosis.

(1) $W_{i+1} = W_i + \Delta W$

Where W_{i+1} is updated value of the weights, Wi is current value of the weights, and ΔW is the updated change of weights, which is determined as Equation (2).

(2) $\Delta W = \eta \frac{\vartheta E}{\vartheta W}$

Where η is the learning rate parameter, and $\frac{\vartheta E}{\vartheta W}$ is the derivative of error with respect to value of the synaptic weights.

These processes are continued until the error reaches a very small value (approximately zero). At this time, the algorithm converges, and the training process is stopped [13]. The flowchart of back propagation algorithm is shown in Figure 2. After that, a test process is commenced to evaluate the performance of trained ANN via applying test samples that are not used in the training process. In this work, the ANN performance is computed by calculating the classification rate as Equation (3).

$$(3) \quad Classification\ rate = \frac{Number\ \ of\ test\ samples\ \ that\ are\ correctly\ \ calssified\ \ by\ ANN}{Total\ \ number\ \ of\ test\ samples} \times 100$$

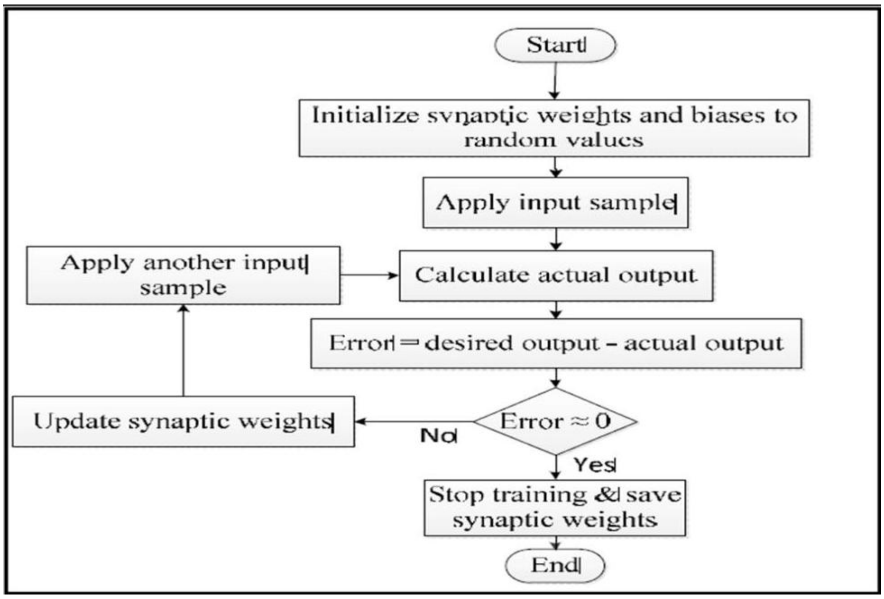

Fig. 2 Flowchart of back propagation algorithm.

3.2 MLP with Back Propagation Algorithm

MLP ANNs architecture is shown in Figure 3. The most popular architecture has three layer includes an input layer, a hidden layer and an output layer and all the connections are full in this architecture.

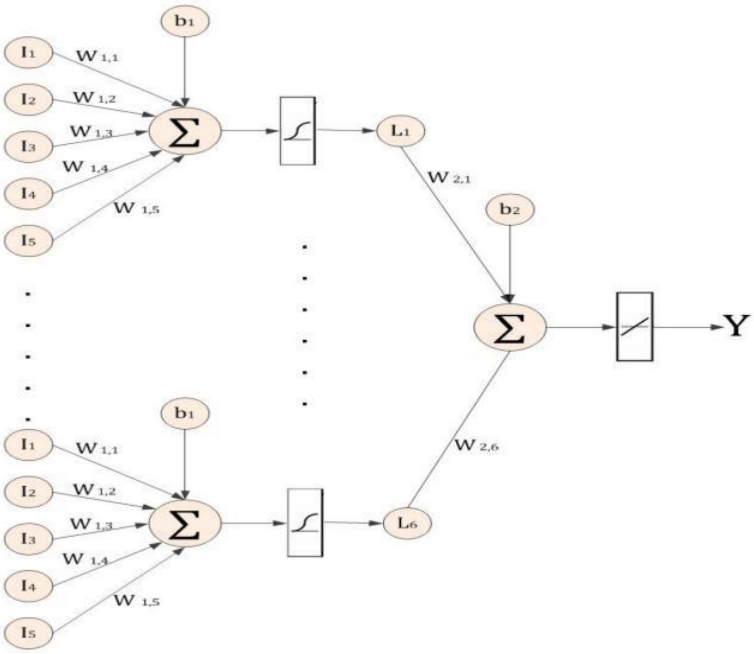

Fig. 3 MLP ANNs architecture.

3.3 MLF with GA and Back Propagation Algorithm

A multilayer neural network structure of three layers; the input layer with five neurons equal to the number of the dataset features, one hidden layer which its neurons will be determined by the GA, and the output layer with only one neuron; that using a type of second order approach as a training algorithm is used for diagnosing thyroid disease. The GA is used to find an optimum network. These two techniques will be illustrated in details at the next subsections and the general steps of this approach algorithm will be listed in the third subsection.

3.3.1 The approach algorithm

The algorithm of this approach will be listed and explained in the following steps:

1. Coding (determine the chromosomes of the GA): Each chromosome will have four genes that represent the number of neurons in the hidden layer, values of the training parameters (mu, mu_inc, and mu_dec). The first gene is integer while the others genes are real.

2. Population Initialization: The population of the individuals (chromosomes) will be initialized randomly in some pre-specified ranges, for this approach, the size of the population is set to 50 individuals.

3. Chromosomes Evaluating (Fitness Function of the GA): The whole goal is to get high classification accuracy with optimum NN that has a minimum number of neurons with lower training epochs. Thus our fitness function is as in equation (4).

(4) *Fitness = perf + hu/max_hu + epochs/max_epochs+ (1-acc) x 100*

Where *perf* is the error function of the neural network, *hu* number of neurons at the hidden layer, *max_hu* maximum number of neurons possible at the hidden layer, *epochs* number of iterations needed by network for convergence, *max_epochs* the maximum number of iterations will be proposed for convergence which here is determined by 100. Finally, Acc is the classification accuracy which is calculated as in equations (5) and (6).

(5) $ACC = \frac{\sum_{t=1}^{|N|} CalCulat\ e(n_l)}{|N|} n_t \in N$

(6) $calculate(n) = \begin{cases} 1 & \text{if classify(n) = nc} \\ 0 & \text{otherwise} \end{cases}$

Where N is the set of data items to be classified (the test set), $n \mathcal{E} N$, *nc* is the class of the item *n*, *classify (n)* returns the classification of by the neural network individual.

4. Reproduction (GA search): The time of the GA search is get after the population initialization and evaluation of each individual by the fitness function. The basic genetic operators guide these searches are:

- Selection: Selection is an important operation. A combination between two selected techniques; Ranking and Tournament will be used to select the two parents.
- Crossover: For each genetic cycle, the two selected parents will be recombined by using the uniform crossover to produce one child with probability *pc = 0.8*.
- Mutation: Each gene in the chromosome that obtained by the crossover will be muted by adding a value generated randomly in some range.
- Replacement: After evaluating the new individual produced from the crossover and mutation, a selected individual of worse fitness will be replaced by the new individual under some condition.

5. Termination Conditions (GA Convergence): The proposed GA is iterated until either the number of the genetic cycles reaches to the predetermined maximum cycle's number which in this work is set to 100 cycles or the first half of the finest part of the population is not changed for some cycles that set for 10 in this work.

4. Data set

4.1 MLF network

We collected 655 samples of real patients from certified advanced hormones laboratory in Kerbala city.

The features are 3 inputs: 1- Thyroid Stimulating Hormone (TSH); 2- Total serum thyroxin (T4); 3- Total serum triiodothyronine (T3).

Most of samples (80%) are used to train ANN, and rest of samples (20%) is used to assess the performance of trained ANN (Tables 1 & 2).

Table 1 Examples of samples in the collected data set.

No.	TSH	T4	T3	Thyroid case
1	3.29	102.35	1.67	Euthyroidism (normal)
2	0.07	142.81	3.11	Primary hyperthyroidism
3	7.05	222	3.9	Secondary hyperthyroidism
4	0.23	111.22	1.55	Subclinical hyperthyroidism
5	0.11	111	4.9	T3 thyrotoxicosis hyperthyroidism
6	29.44	23	0.55	Primary hypothyroidism
7	0.09	22	0.44	Secondary hypothyroidism
8	5.2	82.61	1.49	Subclinical hypothyroidism

Table 2 Number of samples for each thyroid case.

No.	Thyroid case	Number of sample
1	Euthyroidism (normal)	328
2	Primary hyperthyroidism	44
3	Secondary hyperthyroidism	44
4	Subclinical hyperthyroidism	46
5	T3 thyrotoxicosis hyperthyroidism	40
6	Primary hypothyroidism	48
7	Secondary hypothyroidism	40
8	Subclinical hypothyroidism	65
Total	655	

4.2 MLP network

The data is UCI University of California Iverine, The total numbers of laboratory samples are 215.

The features are: (a) T3-resin uptake test (A percentage); (b) Total Serum thyroxin as measured by the isotopic displacement method; (c) Total serum triiodothyronine as measured by radioimmuno assay; (d) Basal TSH as measured by radioimmuno assay; (e) Maximal absolute difference of TSH value after injection of 200 micro grams of thyrotropin-releasing hormone as compared to the basal value.

All features are connected and there is no loosed one. In Table 3, the numbers of laboratory samples are determined based on the type of disease.

Table 3 The number of laboratory samples is determined according to the type of disease.

Type disease	Number of samples	Type class of the UCI classification
Normal	150	1
Hyper	30	2
Hypo	35	3
Total	215	

Third neural network the same second one.

5. Training and Testing

5.1 MLF ANN

The tangent sigmoid is chosen as an activation function for hidden layer, and softmax is used for output layer. We apply input samples in the data set and their target types of thyroid cases to train ANN with various numbers of neurons in the hidden layer. However, the efficiency of ANN depends on the initial value of weights. In addition, weights are initialized with random values. Therefore, each training ANN will get different results. To solve this problem, we train ANN 10,000 times. After that, we record the maximum, minimum, and average classification rates of 10,000 trained ANNs with different number of hidden layer neurons, as shown in Table 4.

Table 4 Classification rate of trained ANNs.

Number of neurons in the hidden layer	Maximum Classification rate %	Minimum Classification rate %	Average Classification rate %
1	74.81	50.38	64.29
2	96.95	50.38	85.01
3	98.47	58.02	91.82
4	98.47	58.78	94.17
5	99.24	62.60	95.08
6	99.24	70.99	95.54
7	99.24	78.63	95.81
8	99.24	80.15	96.04
9	99.24	83.97	96.17
10	99.24	85.50	96.26

In Fig. 4, it can be found that when the number of neurons is > 5, the maximum classification rate stops increasing. In means when the number of neurons is 5, the maximum classification rate reaches the saturation value. This means that the classification rate is proportional to the number of hidden layer neurons.

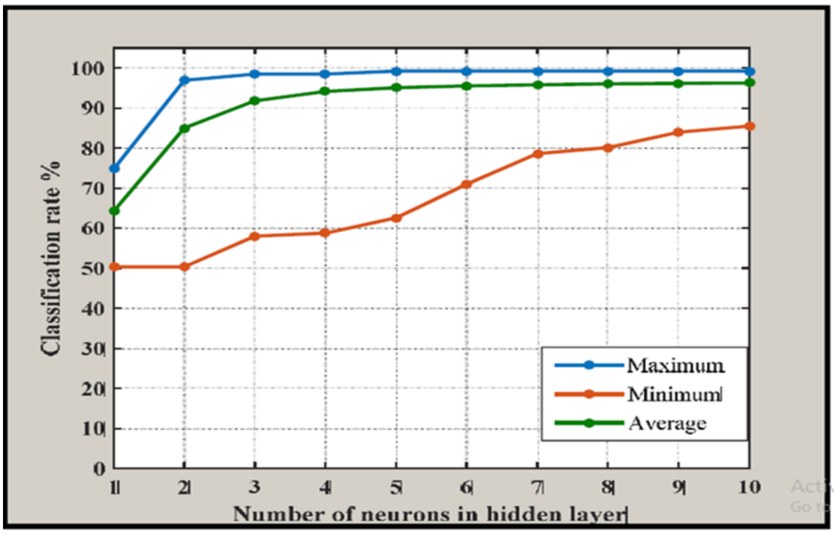

Fig. 4 Classification rate of trained ANNs.

To clarify the performance of final selected ANN, the confusion matrix is illustrated in Figure 5 to show the classification rate of the network by using 131 test samples (20% of total samples in the data set) which are not used in the training process. The green diagonal cells of the matrix demonstrate the number of samples that are correctly classified by ANN and their percentages. On the other hand, the red off-diagonal cells show the number of samples that are misclassified by ANN.

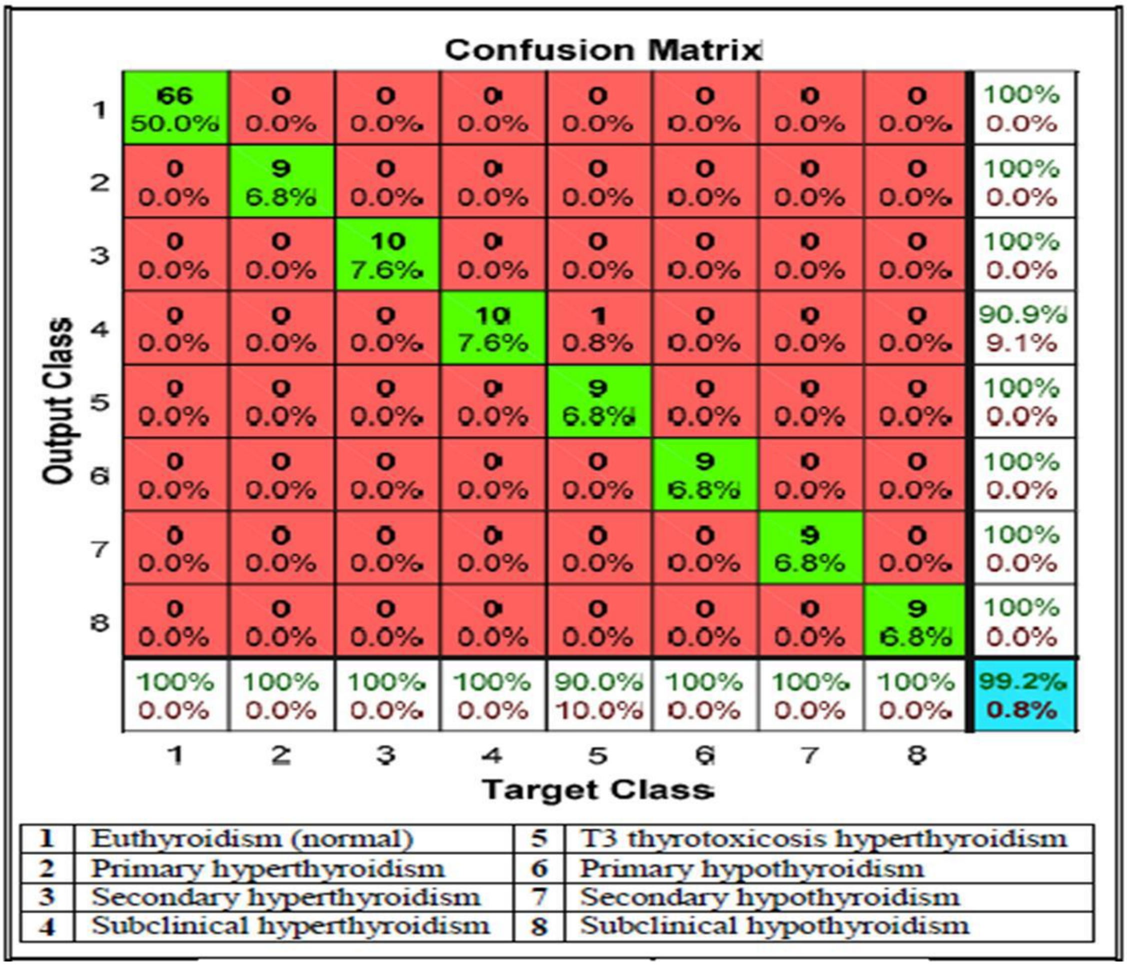

Fig. 5 Confusion matrix of the designed ANN.

5.2 MLP ANN

The techniques used to training an ANN are wide in which the most important is supervision or non-supervision learning techniques. Back propagation learning algorithm which uses under supervision learning technique could learn well based on a set of training examples and then capable of diagnosing each pattern [14]. Basic MLP constructive unit is a simple model of artificial neuron. This unit adds the inputs collection with the value of threshold. And the sum depending on type activation function (generally sigmoid) passes [15]. The network begins to train based on the random values of weights and biases [16, 17] and the training process will continue until the error is minimized [7, 16, 17]. Updating the values of all weights is performed based on error value as far as the difference between network input and output reaches to the desirable one. It gradually achieves from output layer, through hidden layer to input layer toward back [1]. Beside weights optimal adjustment, proper selection of activation function According to [10] Cause more optimal network performance. An activating function of MLP ANN has several major features such as continuous, differentiable, and monotonically non-decreasing [16, 18, 10]. The applied activation function is considered for hidden layer neurons is Logsig and for each output layer neuron is Purelin.

Due to the obtained results of Table 5, the best performance relates to the classification is 85% for Train data and 15% for Test data, respectively. The way of choosing data based on this classification is as random (using Randperm function).

Table 5 Choosing the best classification of test and train data.

Train	Test	Performance for Tansig	Performance for Logsig
70	30	0.00179	9.19e-9
75	25	0.032	1.33e-5
80	20	0.0085	0.056
85	15	0.90e-10	5.94e-10
90	10	2.6e-7	0.046
95	5	5.26e-5	0.0023

5.3 MLF ANN

In order to compare the performance of the genetic neural network techniques, firstly, data set is normalized and split into groups of training set and testing set. The splitting process will be done randomly three times with specific proportion at each time (10% to 90%, 30% to 70%, and 50% to 50%) in order to form different training/testing groups. Table 6 shows the number of instances that used as training set and testing set for each group. For each training/testing pair the GA will be applied five times to find the optimum network topology that gives a high accuracy.

Table 6 The number of instances that used as training set and testing set for each group.

Group Number	Training Set	Testing Set	Total
1	21	194	215
2	64	151	215
3	107	108	215

6. Evaluation and Discussion

First ANN is trained by back propagation algorithm. After training process, if inputs of a particular sample are applied to ANN, only one neuron of output layer will be active (its value is approximately one). In contrast, the values of the other seven will be roughly zero. The active neuron represents the category of thyroid case for input sample. Number of neurons in the hidden layer should be as smallest as possible to reduce the complexity of ANN which leads to decrease the response time of ANN. As a result, the lowest number of neurons that provides the highest maximum classification rate is selected in the proposed system. In this case, 5 neurons are chosen to be in the hidden layer. Hence, architecture of the designed ANN will be 3-5-8. Among 10,000 trained ANNs, the one that provides the maximum classification rate is chosen in the final system for thyroid diagnosis. It can be shown that classification rate of designed ANN is 99.2% which is considered very high as compared to results of previous works. with confusion matrix Only one of the 131 test samples is misclassified. This implies that ANN is almost successfully able to classify samples into eight thyroid cases.

The second ANN compares the performance of MLP ANN with the changes of activation function and the number of hidden nodes and to reach to the high accuracy. The back propagation learning algorithm performance is based on Gradient descent technique. It is considered to regulate weight connections among neurons to minimize system error between real output and target output. Although back propagation learning algorithm is the most popular algorithm to ANN training but sometimes can be inefficient. One of the main training drawback whits the algorithm is the slow convergence. It is proposed methods to improve convergence rate which includes proper selection of activating function in neurons and accurate determination of size parameter of learning rate. We are used the architectures of (5-6-1) with a single hidden layer. Figure 6 display the numbers of hidden layer nodes to the network performance.

Third ANN, see table 7, which show the optimum network by the marker rows for each group.

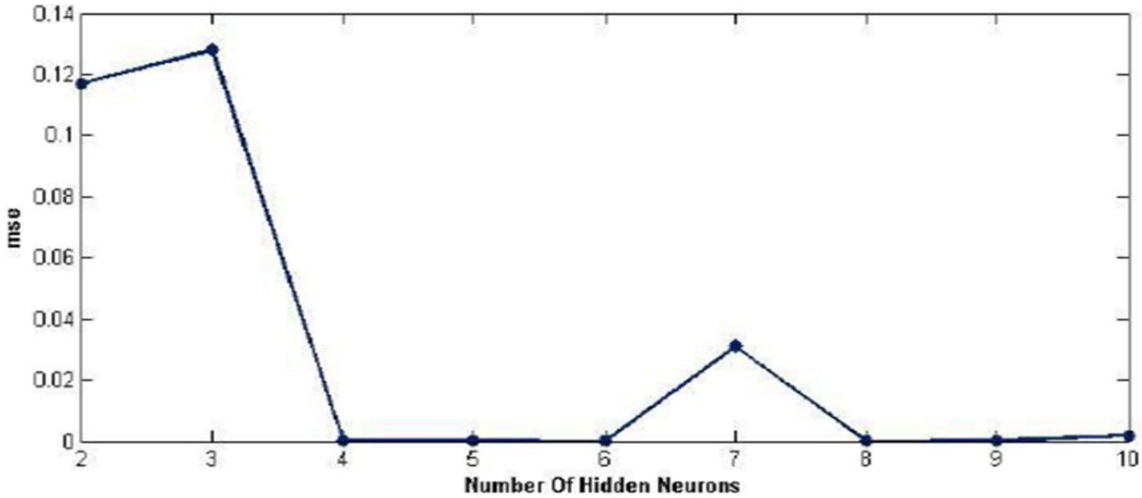

Fig. 6 The numbers of hidden layer nodes to the network performance.

Table 7 The number of instances that used as training set and testing set for each group.

Group No.	hu	mu	mu_inc	mu_dec	epochs	perf	Train Accuracy	Test Accuracy
1	6	0.067	6	0.116	100	2.74E-11	100%	96.91%
	6	0.021	2	0.205	49	2.22E-13	100%	96.91%
	6	0.006	4	0.300	35	1.91E-13	100%	96.39%
	5	0.017	7	0.101	28	8.07E-13	100%	96.91%
	6	0.084	10	0.283	67	9.57E-13	100%	97.42%
2	7	0.068	3	0.102	21	4.93E-13	100%	97.35%
	6	0.043	10	0.501	65	9.31E-13	100%	97.35%
	6	0.058615	10	0.25036	29	3.14E-13	100%	96.69%
	5	0.031	9	0.745	64	3.56E-13	100%	97.35%
	5	0.047	7	0.621	57	7.20E-13	100%	97.35%
3	8	0.021	7	0.321	73	9.05E-13	100%	97.22%
	8	0.070	8	0.702	100	2.62E-11	100%	98.15%
	12	0.359	4	0.359	48	5.35E-13	100%	97.22%
	6	0.163	5	0.163	22	1.94E-13	100%	97.22%
	14	0.259	5	0.259	63	9.29E-13	100%	97.22%

7. Conclusions

In this work, diagnosis of thyroid disease is modeled by neural network techniques.

First one A multilayer feed forward is selected as ANN architecture, and back propagation is used as training algorithm. The results of this work show that classification rate of ANN and number of neurons in the hidden layer is directly proportional. After extensive search for best network, the one with only one hidden layer that has 5 neurons is chosen to perform thyroid diagnosis system. The selected ANN has high classification rate which is about 99.2%. As a result, the proposed structure of ANN can effectively categorize the type of thyroid cases.

Second one we consider the type of appropriate activation function, correct selection of layer number and the network complexity so as to achieve the best result by comparing their performance to reach the best possible answer. By selecting a hidden layer and Logsig activation function for hidden layer and 6 neurons in the hidden layer, we can reach the classification accuracy for Thyroid disease to 98.6%.

Third one presents a study on thyroid disease diagnosis by using neural networks with second order training algorithm. The genetic algorithm was used to find the optimum network structure with high classification accuracy. Three different proportions of training/testing groups are formed. According to the results, it was seen that neural network structures could be successfully used to help diagnosis of thyroid disease. Overall accuracy of diagnosis is 100% for training and in range between 96% and 98% for testing.

References

[1] Dey, R., Bajpai, V., Gandhi, G., & Dey, B. 2008. "Application of Artificial Neural Network (ANN) technique for Diagnosing Diabetes Mellitus." *IEEE Region 10 and the Third International Conference on Industrial and Information Systems*, Kharagpur, India. pp. 1-4, doi: 10.1109/ICIINFS.2008.4798367.

[2] Liu, L., and Deng, M. 2010. "An Evolutionary Artificial Neural Network Approach for Breast Cancer Diagnosis." *Third International Conference on Knowledge Discovery and Data Mining*, Phuket, Thailand, pp. 593-596, doi: 10.1109/WKDD.2010.148.

[3] Sarasvathi, V., and Santhakumaran, A. 2011. "Towards Artificial Neural Network Model to Diagnose Thyroid Problems." *Global Journal of Computer Science and Technology* 11 (5): 53-55.

[4] Haykin, S., 2001. Neural Networks, Comprehensive Foundation. 2nd ed. Pearson Education.

[5] Lichman, M., 2013. [Online] Available at: http://archive.ics.uci.edu/ml.

[6] Nussey, S., & Whitehead, S. 2001. The thyroid gland. In: *Endocrinology: An Integrated Approac*h. Oxford: BIOS Scientific Publishers Ltd, pp. 71-111.

[7] Kabari, L. G., and Bakpo, F. S. 2009. "Diagnosing Skin Diseases Using an Artificial Neural Network." 2nd International Conference on Adaptive Science & Technology (ICAST), Accra, Ghana, pp. 187-191, doi: 10.1109/ICASTECH.2009.5409725.

[8] Gharehchopogh, F. S., and Khalifelu, Z. A. 2011. "Neural Network Application in Diagnosis of Patient: A Case Study." *International Conference on Computer Networks and Information Technology*, Abbottabad, Pakistan, pp. 245-249, doi: 10.1109/ICCNIT.2011.6020937.

[9] Shukla, A., Tiwari, R., Kaur, P., & Janghel, R. R. 2009. "Diagnosis of Thyroid Disorders using Artificial Neural Networks." *IEEE International Advance Computing Conference*, Patiala, India. pp. 1016-1020. doi: 10.1109/IADCC.2009.4809154.

[10] Isa, I. S., Saad, Z., Omar, S., et al. 2010. "Suitable MLP Network Activation Functions for Breast Cancer and Thyroid Disease Detection." Second International Conference Computational Intelligence, Modelling and Simulation, pp. 39-44, doi: 10.1109/CIMSiM.2010.93.

[11] Graupe, D. 2013. *Principles of Artificial Neural Networks*. 3rd ed. World Scientific Publishing Company.

[12] Pandya, A. S., & Macy, R. B. 1995. "Feed-Forward Networks with Supervised Learning." In: Pattern Recognition with Neural Networks in C++. CRC Press, pp. 73-145.

[13] Fausett, L., 1994. *Fundamentals of Neural Networks: Architectures, Algorithms, and Applications*. Prentice-Hall.

[14] Zhong, S., Zeng, X., Wu, S., & Han, L. 2012. "Sensitivity-Based Adaptive Lear." *IEEE Trans Neural Netw Learn Syst* 23 (3): 480-91.

[15] Ozylmaz, L., and Yildirim, T. 2003. "Artificial Neural Networks for Diagnosis of Hepatitis Disease." *Proceedings of the International Joint Conference on Neural Networks*, Portland, OR, USA. pp. 586-589. doi: 10.1109/IJCNN.2003.1223422.

[16] Belciug, S., Gorunescu, F., Gorunescu, M., & Salem, A.-B. M. 2010. "Assessing Performances of Unsupervised and Supervised Neural Networks in Breast Cancer Detection." In the 7th International Conference on Informatics and Systems (INFOS), Cairo, Egypt, pp. 1-8.

[17] Yeh, W-C. 2013. "New Parameter-Free Simplified Swarm Optimization for Artificial Neural Network Training and its Application in the Prediction of Time Series." *IEEE Transactions on Neural Networks and Learning Systems* 24 (4): 661-5.

[18] Wilamowski, B. M., and Yu, H. 2010. "Neural Network without Back Propagation." *IEEE Trans Neural Netw* 21 (11): 1793-803.