NEURAL NETWORKS OF ENGINE FAULT DIAGNOSIS BASED ON EXHAUST GAS ANALYSIS

Rafil M. Laftah
Dean Assistant of Engineering College
Basrah University
rafil_74@yahoo.com

Qusai T. Abd-Alwahab
Educational in Engineering College
Basrah University
qusaith@yahoo.com

Jaafar M. Hamzah
Lecturer in Engineering College
Thiqar University
jaafar_2011m@yahoo.com

Abstract: This work uses the Artificial Neural Networks (ANNs) for fault diagnosis of a single cylinder four stroke gasoline generator type (Astra Korea AST11700). One normal and fourteen faulty conditions are examined experimentally to produce a realistic data set, which is to be used for the training and validation of the ANNs. The resulted data was in the form of exhaust gases and engine speed records for each case separately under different loading conditions. After the learning process is completed, the ANN becomes able to make a diagnosis about the gasoline engine condition when new data is presented. The data presented to the ANN system include a subset of engine faults which were selected and executed experimentally for this topic. These include, faults in carburetor, air filter, spark plug, valves, piston rings, etc. The results showed that the multi layer training algorithm is sufficient enough in diagnose engine faults under different loading conditions. It was found that the correlation coefficient values are 0.999 and 1 for the testing and training data, respectively. The results obtained in this investigation showed that the ANN-based fault diagnosis system is capable of fault diagnosis with high reliability.

Keywords: Artificial neural network; IC engine; Exhaust gas analysis; Fault diagnosis; Prediction

الخلاصة : يَستخدِم هذا العمل الشبكاتِ العصبيةِ الاصطناعية لِتشخيص أعطالِ مولدٍ كهربائي ذو محركٍ بأسطوانةٍ واحدةً ، رباعيُ الأشواط . لقد تم إختبار أربعة عشر نوعٍ مِنَ الأعطال بالإضافة إلى الحالةِ الطبيعيةِ عملياً لغرض إنتاج مجموعة من البياناتِ الشاملة والواقعية ، إستُعمِلت هذِهِ البيانات لِتدريبِ وتحقيقِ الشبكة . كانت البياناتُ الناتجةِ على شكلِ غازات العادم وسرعة المحرك المسجلة لكل حالة على حدة في ظل ظروف تحميل مختلفة . بعد انتهاء عملية التعلم، تكون الشبكة قادرة على التشخيص عند استخدام بيانات جديدة. البيانات المدخلة إلى الشبكة العصبية الصناعية تتضمّن مجموعة من أعطال المحرك التي نفذت بشكل تجريبي لهذا الغرض. ومن هذه الأعطال، خلل في نظام المكربنة، نظام ترشيح الهواء، شمعة القدح، الصمامات، وحلقات المكبس ... الخ. أظهرت النتائج أن خوارزمية تدريب الرجوع العكسي هي كافية في تشخيص أعطال المحرك تحت ظروف التحميل المختلفة . وجد أن قيم معامل الارتباط هي ٠.٩٩٩ و ١ لبيانات الاختبار و التدريب ، على التوالي . وأظهرت النتائج في هذا التحقيق أن نظام تشخيص الأخطاء المستند على الشبكات العصبية قادر على اكتشاف الخطأ والتشخيص مع موثوقية عالية.

1. Introduction

All real systems in nature – physical, biological and engineering – can malfunction and fail due to faults in their components. The chances for failures are increasing with the system's complexity. System fault can be defined as any undesirable or abnormal behavior. The complexity of engineering systems is permanently growing due to the growing of size of the systems and the degree of automation, which increases the chances for faults and aggravating their consequences for man and environment. Therefore, increased attention has to be paid to the reliability, safety and fault tolerance in the design and operation of engineering systems [1].

The cognitive approach of artificial intelligence is focusing on imitating the rational thinking of human. Artificial intelligence system such as fuzzy logic, neural networks and genetic programming had been integrated with the conventional fault diagnosis system to produce intelligent diagnostic systems [2].

2. Fault Diagnosis Methods

Fault diagnosis is the act of detecting and isolating faults present in a system. With the rising demand for safety and reliability of technical systems, driven by economical and environmental incentives, fault diagnosis has become increasingly important. One example is automotive systems that are by regulations required to have on-board diagnosis of all faults that may lead to increased emissions. In addition, fault diagnosis in automotive systems is essential to maintain high vehicle uptime, low fuel consumption, high safety, and efficient service and maintenance [3]. Fault diagnosis methods can be classified into two general categories, model-based and data driven methods. The hierarchy of fault diagnosis approaches is shown in Fig. 1.

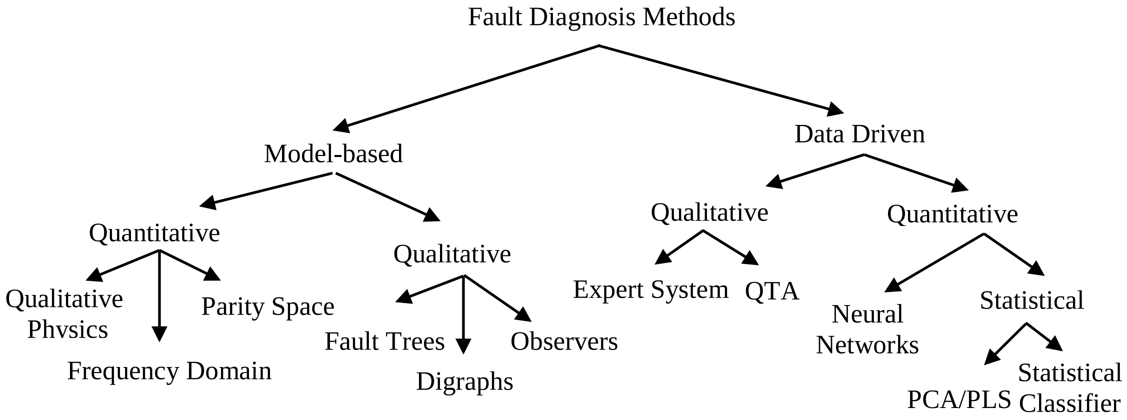


Fig. 1: Classification of Fault Diagnosis Methods [4]

3. Artificial Neural Networks

Artificial Neural Networks (ANNs) are a branch of the field known as "Artificial Intelligence" which may also consists of Fuzzy Logic (FL) and Genetic Algorithms (GA) [5]. Artificial neural networks are non-linear mapping structures based on the function of the human brain [6]. They are powerful tools for modeling, especially when the underlying data relationship is unknown. ANNs can identify and learn correlated patterns between input data sets and corresponding target values [7].

Neural networks are composed of simple elements operating in parallel. These elements are inspired by biological nervous systems. As in nature, the network function is determined largely by the connections between elements. It can train a neural network to perform a particular function by adjusting the values of the connections (weights) between elements. The weights adjusting process is usually called learning process. Commonly neural networks are adjusted, or trained, so that a particular input leads to a specific target output [8].

The block diagram of Fig. 2 shows the model of a neuron, which forms the basis for designing (artificial) neural network. Here we identify three basic elements of the neuronal model [9].

1. A set of synapses or connecting links, each of which is characterized by a weight or strength of its own. Specifically, a signal (p_i) at the input of synapse (i) connected to neuron (j) is multiplied by the synaptic weight w_{ij}.
2. An adder for summing the input signals, weighted by the respective synapses of the neuron.
3. An activation (or transfer) functions for limiting the amplitude of the output of a neuron.

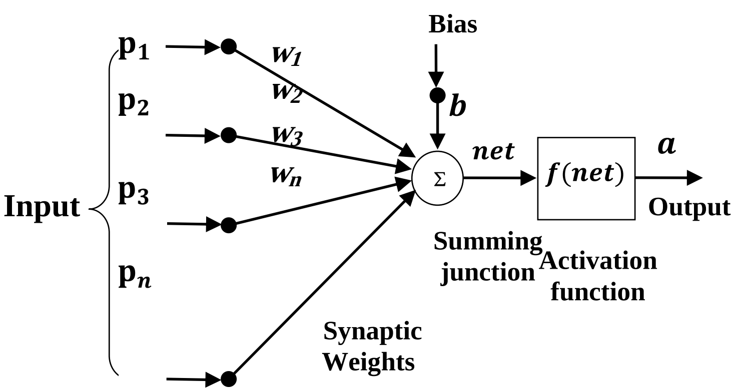


Fig. 2: Structure of a Single Artificial Neuron

3.1 Training of Multi-Layer Perceptron

The Multi-Layer Perceptron (MLP) is a nonparametric technique for performing a wide variety of estimation tasks. Error Back-Propagation is one of the most important and widely used algorithms for training multilayer perceptron's [10]. Training process of MLP networks continues until a certain number of iterations or a desired error rate is reached. The most common error approximation method used in MLP networks is mean square error (MSE) and it is defined by the following formula [11]:

$$\mathrm{MSE} = \left(\sum_{k=1}^{n}\sum_{j=1}^{m}\left({t_j}^k - {a_j}^k\right)^2\right) \Big/ s\,.\,l \qquad (1)$$

3.2. Error Back propagation (EBP)

MLP have been applied successfully to solve some difficult diverse problems by training them in a supervised manner with a popular algorithm known as the EBP algorithm. BP algorithm was first defined by Werbos (1974) and later improved by Rumelhart et al. (1986) [12] as a euphemism for generalized delta rule. The BP generalize delta rule is a decent method to minimize the total squared error of the output computed by the net. The Back propagation algorithm is used in layered feed-forward ANNs. This means that the artificial neurons are organized in layers, and send their signals "forward", and then the errors are propagated backwards [13]. The training procedure consists of two main steps: Feed-forward and back-propagation. During forward pass the synaptic weights of network are all fixed. During backward pass, on the other hand, the synaptic weights are all adjusted in accordance with the error-correction rule. Specifically, actual response of the network is subtracted from a desired response to produce an *error signal*. This error signal is then propagated backward through the network, against direction of synaptic connections - hence the name "error back-propagation" [14].

3.3. Back propagation Training Algorithm

The back propagation (BP) training algorithm involves three stages the feed forward of the input training pattern, the calculation and back propagation of the associated weight error and the weight adjustments. The training algorithm (for one hidden layer) is as following [15]:

Step 1: Initialize the weights.

Step 2: While squared error is greater than a tolerance, execute steps 3 to 11.

Step 3: For each training pair, do steps 4 to 11.

Step 4: Sums the weighted input and bases and apply the activation function to compute the output of hidden layer.

$$a_j = f\left(\Sigma\ \mathrm{p_i}\ w_{ij} + b_j\right) \tag{2}$$

Step 5: Sums weighted output of hidden layer and apply activation function to compute output.

$$a_k = f(\Sigma\ a_j\ w_{jk} + b_k) \tag{3}$$

Step 6: Compute back propagation error

$$\delta_{\ k} = (t_k - a_k) f'\left(\Sigma\ a_j\ w_{jk} + b_k\right) \tag{4}$$

Step 7: Calculate weight correction term

$$\Delta w_{jk}\ (\tau) = \eta \delta_k\ a_j\ +\ \mu \Delta w_{jk}(\tau\ - 1) \tag{5}$$

Step 8: Sums delta input for each hidden unit and calculate error term

$$\delta_j = \ \Sigma_k \delta_k\ a_k = \ \Sigma_k \delta_k \left(w_{jk}\ f\left(\Sigma_i\ \mathrm{p_i}\ w_{ij} + b_j\right) + b_k\right) \tag{6}$$

Step 9: Calculate weight correction term

$$\Delta w_{ij}\ (\tau) = \eta \delta_j\ \mathrm{p_i}\ +\ \mu \Delta w_{ij}(\tau\ - 1) \tag{7}$$

Step 10: Update weights

$$w_{jk}(new) = \ w_{jk}\ (old) + \ \Delta w_{jk} \tag{8}$$

$$w_{ij}\ (new) = w_{ij}\ (old) + \ \Delta w_{ij} \tag{9}$$

Step 11: Compute the mean squared error.

4. Experimental Work

The current work concerns the study of engine faults and suggesting the proper diagnosis for each fault. There are certain requirements necessary to achieve this task:

- Proper selection of test engine, in our case engine is selected due to its wide spread domestic use in Iraq.
- Selection of the suitable gas analyzer (with new factory calibration).

- Selecting set of engine faults, which will be discussed below.
- Framing the diagnosing program (ANN techniques)

The experimental procedure can be summarized in two steps. The first step is to prepare test equipments and connect the engine to the gas analyzer which is in turn connected to a personal computer. The analyzer sample gas pipe is to be connected to exhaust tail pipe. The RPM clamp counter is connected to the engine spark plug cable. All the above mentioned steps were followed according to the user manual of the gas analyzer. The second step consists of three processes:

1. Turn on the gas analyzer while the engine is running at speed of 3800 rpm and wait until the oil temperature reaches 80 $°C$.
2. Take measurements of the five exhaust concentrations plus (λ) at four loading cases: no load, one third load, two thirds load and full load (2.5 kW), this performed by loading the engine with electrical heater.
3. Repeat process 2 fifteen times: one with normal operation and fourteen with faulty operation.

Then the readings were collected and fed to the ANN as an input data. Fig. 3 shows the map of the experimental work as explained before.

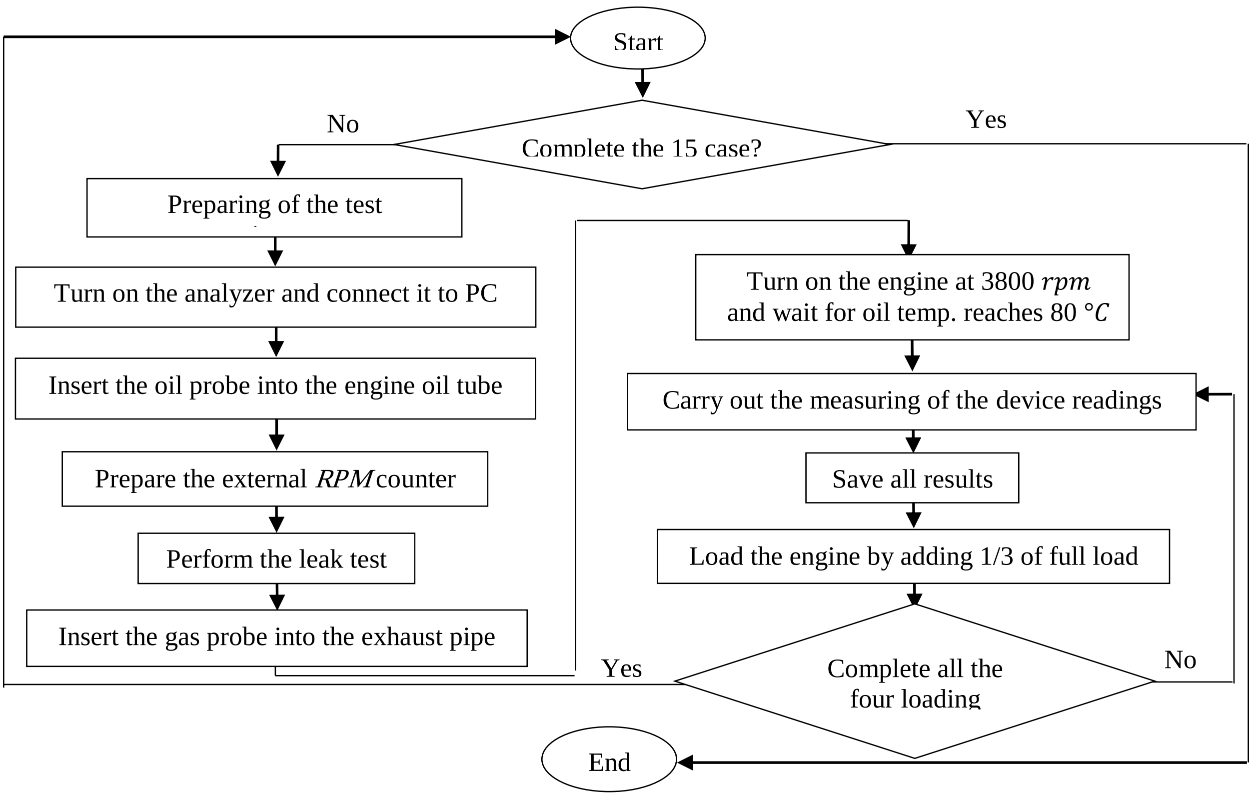


Fig. 3: Map of the Experimental Work

5. Selected Engine Faults

In order to perform engine faults diagnosis, it is required to run the engine under normal operation first and then to run it under faulty operation for each case. Fourteen engine faults were suggested. These faults were selected according to their frequent occurrence in practice. The faults can be listed hereunder with summarized description of the purpose of selecting each fault.

1. *Bad Lubricant:* Unchanging oil for time longer than specified can lead to increase engine temperature and reduce its efficiency, which leads to significant effect on engine operation.
2. *No Air Cleaner Used:* Air cleaner is very important element; it cleans the air entering the carburetor from the dust and dirt. Running of the generator without this element may be resulted in rapid engine wear [16]. This case is considered as engine fault since it affects carburetor operation, which in turn reduces the efficiency of engine operation.
3. *Dusty Air Cleaner*: Dusty air cleaner may result in lake of the air supplied to the carburetor. This in turn will give effect similar to chocking of the engine. This fault affects the air/fuel ratio directly, thus, it leads to increase the emissions.
4. *Removed Evaporated Oil Tube*: Oil tube is the tube in which evaporated oil can be transferred from the overhead valve to the carburetor inlet, this process causes to lubricate the piston crown beyond the piston rings. Removing this element affects the lubricating system, and then it affects engine efficiency.
5. *Small Spark Plug Air Gap:* Spark plug gap affects directly the spark intensity. Unsuitable air gap leads to improper spark intensity, which in turn leads to incomplete combustion due to ignition misfire. By reducing air gap, the ignition will start before normal ignition time. That leads to ignition misfire and causes engine fault.
6. *Large Spark Plug Air Gap:* The increase of the spark plug gap leads to lack in spark intensity, then incomplete combustion due to ignition misfire.

   Partially Blocked Carburetor Main Fuel Jet: The main jet is one of the carburetor elements, in which the fuel is transferred through it from the orifice. By making an obstruction through the orifice to make its diameter (1.4 *mm*), the fuel transferred for mixing with air will be reduced causing lean mixture introduced to the combustion chamber.

7. *Carburetor Main Jet Orifice Diameter is (0.7 mm):* This case was similar to the previous, but here the orifice diameter was (0.7 *mm*) which making the mixture leaner than the previous case, so we can consider this case as a carburetor fault as the former but greater degree.
8. *Leak in Inlet Valve:* The leak in the both seat or face valve escape the compression mixture, with reduce engine efficiency. Also, the valve spring gives the same fault.
9. *Leak in Exhaust Valve:* This fault is considered as important situation, since it causes to lose some of the combustion chamber charge through the tailpipe, which in turn reduces the pressure in the combustion chamber. This fault is introduced by replacing the exhaust valve by an old valve subjected to remarkable erosion level.
10. *Crack in the Tailpipe:* Tailpipe crack is undesirable; it causes an offhand manner because of generated noise. This fault is introduced in this study because it could be diagnosed easily.
11. *The Gasket is ruptured:* This case will result to burn some of lubrication oil with combustion chamber charge. Then, many problems will happen due to this fault such as temperature rise, oil lacking and emission increase.
12. *Wearing of the Rings:* When the piston rings become exhausted, as it used in this case, a complex problem will happen such as oil burning, pressure leaking and unstable operation.
13. *A Problem in the Cooling System:* Cooling air is coming from the fan will spread by untying the screw which ties the plate to the engine fins body. This case reduces cooling system efficiency; therefore, which affects directly on the engine operation and emissions.

6. Results and Discussion

ANN is developed to diagnose the engine faults using exhaust gas analysis. This section describes the data selection for training and testing patterns, the topology of the constructed network, the training process and the verification of the neural network results. The total experimental data are divided into two sets: a training set and a testing set. The training set is used for computing the gradient and updating the network weights and biases to diminish the training error, and find the relationship between the input and output parameters.

The testing set is used to evaluate the generalization ability of the learning process. In this study the testing set contains approximately (15) % of total data base; however, other ratios were executed in this investigation. The parameters used in this study are shown in Table 1. The experimental values used to train the neural network as training data are those measured using exhaust gas analyzer. The total number of (300) test cases were utilized, each five readings in this data represent one loading cases and each twenty columns in this data represent four loading cases of one fault case. The training set contains (254) cases and the testing set comprises of (46) cases.

Table 1: Input and Output Parameters

Item		***Range of Parameters***		***Unit***
		From	***To***	
Input	Carbon Dioxide (CO_2)	9.04	13.47	%
	Carbon Monoxide (CO)	0.253	5.024	%
	Oxygen (O_2)	0.34	2.81	%
	Unburned hydrocarbon (HC)	4	673	ppm
	Nitrogen Oxidize NO_X	47	873	ppm
	Relative air/fuel ratio (λ)	0.856	1.124	
Output	Case type code	1	15	

The output data is one parameter assumed arbitrarily. This vector represents the case type code. Each code in this vector refers to the selected cases (normal case and faulty cases) as shown in Table 2. Therefore, the nodes in the input layer and output layer are (6) and (1), respectively.

Table 2: Case Type and Code

	Case Type	*Code*
	No Problems	1
Faulty Cases	Bad oil	2
	No air cleaner used	3
	Dusty air cleaner	4
	Removed evaporated oil tube	5
	Gap of the spark plug is small	6
	Gap of the spark plug is large	7
	Carburetor main jet is 33 % blocked	8
	Carburetor main jet is 66 % blocked	9
	Leak in inlet valve	10
	Leak in exhaust valve	11
	Crack in tailpipe	12
	Gasket is crushed	13
	Exhaustion of the piston rings	14
	A problem in cooling system	15

In this study the network is tested with one and two hidden layer configurations with an increasing number of nodes in each hidden layer(s). Different activation function arrangements are investigated. The optimal topology is determined first by using one hidden layer with activation function as hyperbolic tangent (*tansig*) function for hidden and output layers.

Different numbers of nodes from (7 to 26) are investigated and the training performance and correlation coefficient for both training and testing dataset of these topologies are shown in Figs. 4 and 5. From these figures the network with (16) nodes in the hidden layer gives best performance and correlation coefficient than other.

Then two hidden layers are used with different activation functions for both first hidden layer, second hidden layer and output layer. Different node numbers in each hidden layer from (3 to 14) nodes are used. The performance and regression of these topologies of network for both training and testing are shown in Figs. 6 and 7. From these figures the network with (9 nodes in the first and 7 nodes in the second hidden layer) gives best performance and regression for both training and testing than other. The results show that a network with two hidden layers is significantly better than that with one hidden layer. Fig. 8 shows the comparison between the training and testing performance for the best one hidden layer network (16 nodes) and the best two hidden layer network (9,7 nodes). From this figure, it can be seen that the network (9-7) gives best performance and regression for both training and testing. The configuration of the ANN is shown in Fig. 9. The choice of the training function depends on the network application. Three types of training function are used in the neural networks modeling of the present study. Although, using *Traincgp* algorithm might be sufficient in solving many functional approximation problems, some other problems may be easier to be solved with *Traingdm* or *Trainrp* training function.

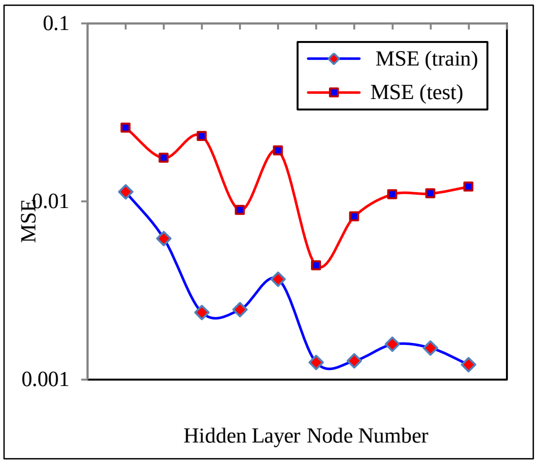
7. Proposed Network

From the above analysis, for different arrangements of neural networks it can be seen that (*Traincgp*) training function with (*tansig*) activation functions for the two hidden layers and output layer gives the best MSE and correlation coefficient for both training and testing than other. Therefore, this network can be selected as a proposed network for the present study. Table 3 and Fig. 10 show the comparison between the best different algorithms for network model. Fig. 11 shows training and testing performance of the proposed neural network.

The performance of a trained network can be measured to some extent by the errors on the training and testing sets, but it is often useful to investigate the network response in more detail. One option is to perform a regression analysis between the network response and the corresponding targets. Figs. 12 and 13 show the regression analysis between the output of neural network and the corresponding target for training and testing data respectively. In the figures, outputs are plotted versus the targets as open circles. The solid line indicates the best linear fit and the broken line indicates the perfect fit (output equals target). The values of the slope are (0.996) and (0.99) respectively, interceptions with y-axis are (0.0282) and (0.129) respectively, and correlation coefficients are (1) and (0.999) respectively. These values indicate that the mapping of neural network for the training and testing data is very good. Fig. 14 shows the graphical user interface operating mode for fault diagnosis model, it is coded using *Microsoft Visual Studio (2010).*

Table 3: MSE and Regression of Three Different Algorithms for Network Model

Training Function	*MSE (train)*	*MSE (test)*	*R (train)*	*R (test)*
Traingdm	0.00569	0.01438	0.992	0.985
Traincgp	0.00019	0.00088	1.000	0.999
Trainrp	0.00077	0.00361	0.999	0.996



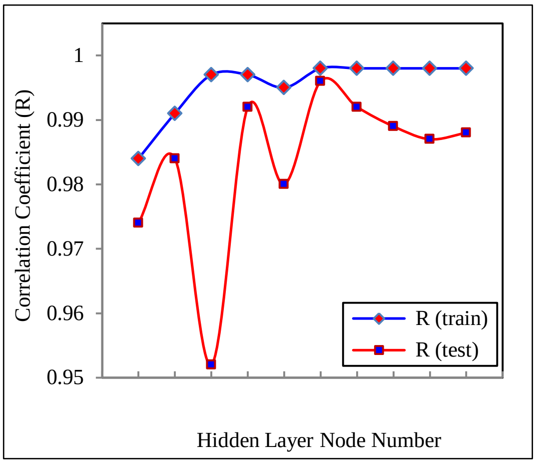
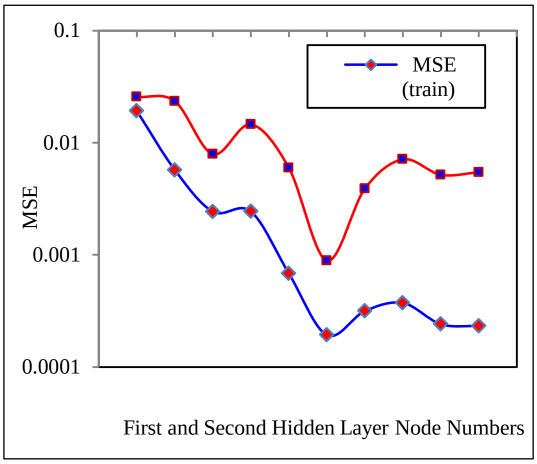


Fig. 5: Regression of Network with One Hidden Layer

Fig. 4: Performance of Network with One Hidden Layer



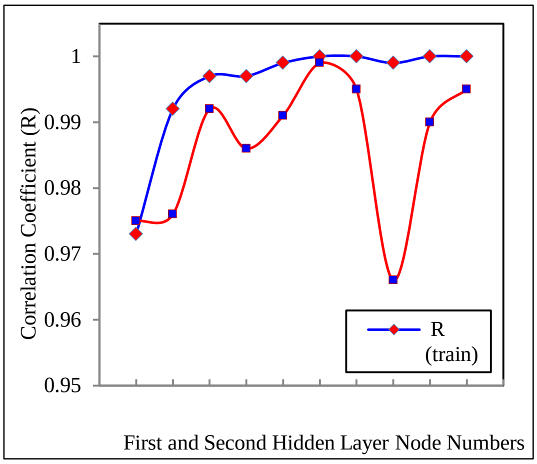


Fig. 6: Performance of Network with Two Hidden Layers

Fig. 7: Regression of Network with Two Hidden Layers.

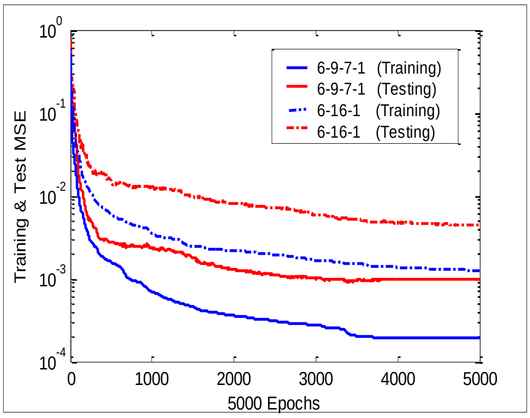


Fig. 8: Compare between One and Two Hidden Layers Performance

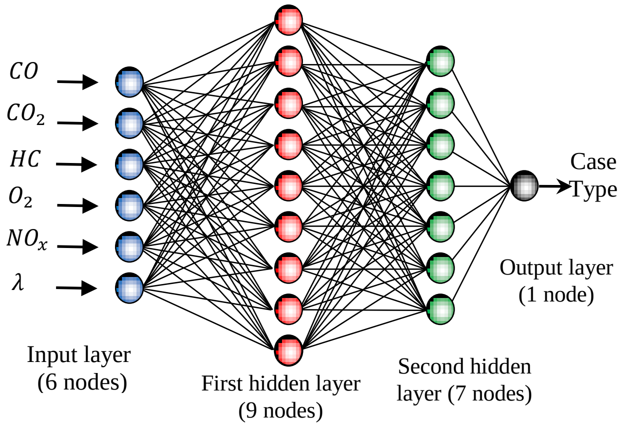


Fig. 9: Configuration of Neural Network (6-9-7-1)

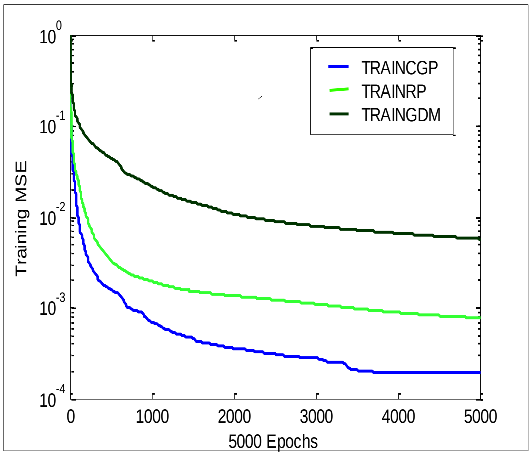


Fig. 10: Compare between the Best Three Types of Network Performance

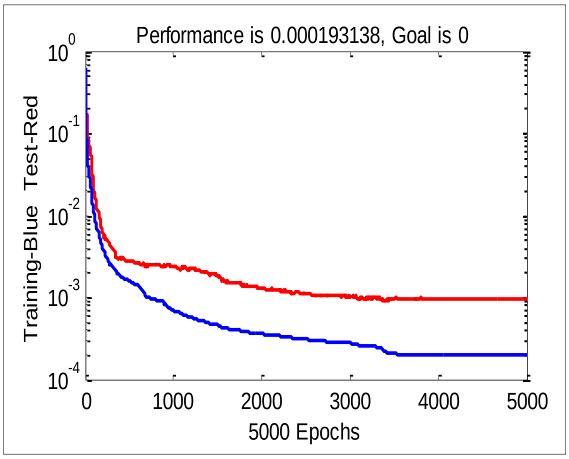


Fig. 11: Training and Testing MSE vs. Epochs of the Proposed Neural Network

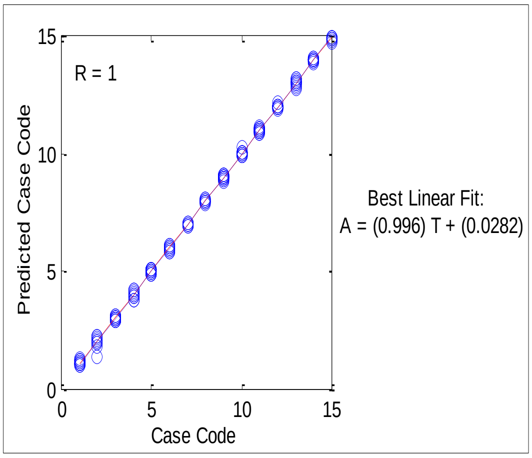


Fig. 12: Training Regression of the Proposed Network

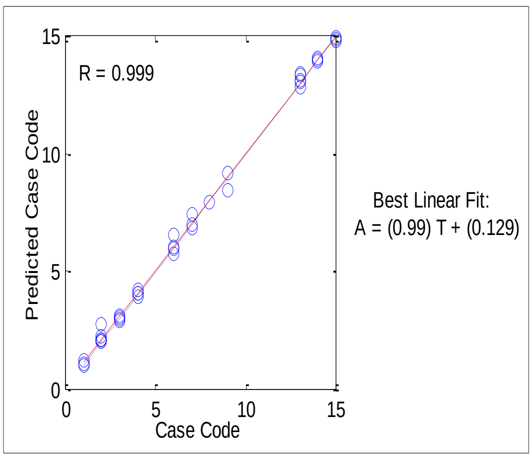


Fig. 13: Testing Regression of the Proposed Network

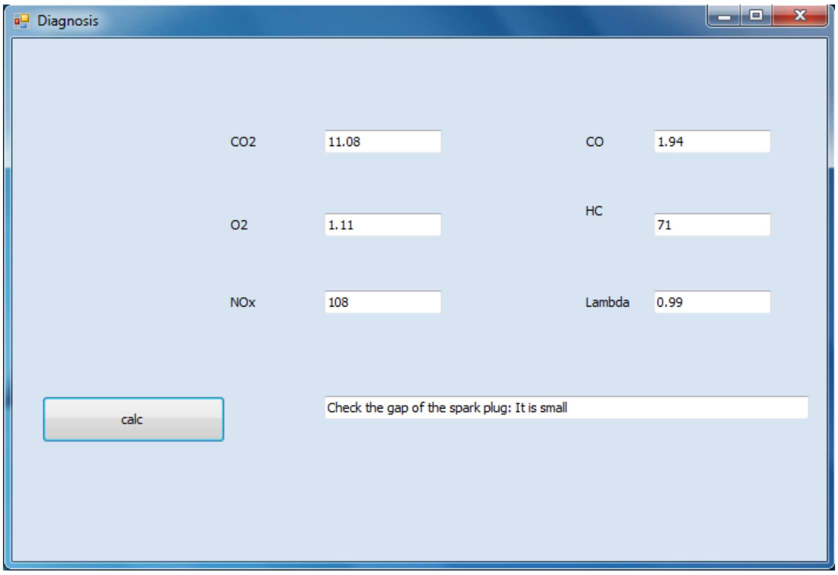


Fig. 14: Graphical User Interface Using Visual Basic Program

8. Conclusion

The most important conclusions that can be drawn from the present study are the followings:

1. Exhaust gas analysis of an Internal Combustion engine plays a significant role in fault diagnosis under various loading conditions.
2. The multilayer (two hidden layers) feed forward neural network model, rather than single hidden layer, significantly improves performance of network. The configuration (9-7) is proved to be very efficient for diagnosing different engine faults using its emissions.
3. The number of hidden layers, the number of nodes in the hidden layers, the arrangement of activation function and the type of training function had significant effects on the performance of the ANN models.
4. Type and arrangement of activation function affect the response of network. The [*tansig*, *tansig*, *tansig*] activation function is found to give a minimum mean square error and maximum regression for the most investigated networks.
5. The conjugate gradient back propagation with Polak-Ribiere updates (*Traincgp*) algorithm gives best performance and regression for both training and testing phases better than the other training algorithms.

References

[1] Frank, P. M. , 2002, "Handling Modeling Uncertainty in Fault Detection and Isolation systems", Journal of Control Eng. and Applied Informatics, Vol. 4, No. 4, pp. 29-46.

[2] Zhang, Y. M. and Li, X. R. 1998, "Detection and diagnosis of sensor and actuator failures using IMM estimator", Journal of AES, Vol. 34, No. 4, pp.1293-1312, Oct.

[3] Svard, C., 2009, "Residual Generation Methods for Fault Diagnosis with Automotive Applications", Thesis No. 1406, Electrical Engineering, Linkoping University, Sweden.

[4] Rengaswamy, R., et al, 2003, "A Review of Process Fault Detection and Diagnosis Part I: Quantitative Model-based Methods", J. of Comp. and Chem. Eng., Vol. 27, pp. 293-311.

[5] Gadoue, S. , 1998, "Artificial Neural Networks", EEE 8005 - SDL Industrial Automation Lecture Notes, Newcastle University, School of Electrical, Electronic & Computing Eng.

[6] Jha, G. K., 2007, "Artificial Neural Networks", Journal of Indian Agricultural Research Institute, PUSA, New Delhi, 110-12, pp. 1-10.

[7] Jha, G. K., 2007, "Artificial Neural Networks and Its Applications", Journal of Indian Agricultural Research Institute, New Delhi, 110-012, pp. 41-49.

[8] Demuth, H., Beale, M. and Hagan, M., 2006, "Neural Network Toolbox User's Guide Ver5", The Math works, Inc., Natick, MA, USA.

[9] Haykin, S., 1999, "Neural Networks - A Comprehensive Foundation", Second Edition, Prentice Hall, Indian Branch.

[10] Verma, B., "Fast Training of Multilayer Perceptrons", Proceedings of IEEE & IASTED, School of Information Technology,Griffith University,Gold Coast,Queensland, Australia.

[11] Jain, A. K., Mao, J. and Mohiuddin, K, 1996, "Artificial Neural Network: A Tutorial", Journal of IEEE Computer, Vol. 29, No. 2, pp. 31-44.

[12] Gullu, M., et al, 2011, "Application of Back Propagation ANN for Modeling Local GPS/Leveling Geoid Undulations", Geoid and GNSS Heighting, Turkey.

[13] Rumelhart, D. and McClelland, J., 1986, "Parallel Distributed Processing", MIT Press, Cambridge, Mass.

[14] Jain, A. K., Mao J. and Mohiuddin K. M., 1996, "Artificial Neural Networks: A Tutorial", Journal of IEEE, 0018-9162/96, pp. 31-44.

[15] Hema, C. R. et al., 2007, "Motor Imagery Signal Classification for a Four State Brain Machine Interface", International J. of Biological and Life Sci., Vol. 3, No. 1, pp. 76-81.

[16] Astra Korea Company, "Gasoline Generator", AST11770, Operation Manual.

List of Abbreviations and Symbols

Symbol	*Description*
a	Actual output
ANN	Artificial Neural Network
b	Bias value
BP	Back-Propagation
EBP	Error Back-Propagation
f	Activation function
FDI	Fault Detection and Isolation
FL	Fuzzy Logic
GA	Genetic Algorithms
MLP	Multi-Layer Perceptron
MSE	Mean Squared Error
p	Input signal
PCA	Principal Component Analysis
PLS	Partial Least Squares
pn	Normalized value of input signal
QTA	Qualitative Trend Analysis
R	Correlation coefficient
RPM	Revolution Per Minute
s	No. of training examples
t	Desired output (Target)
v	The difference between the target and the output
w	Weight value
δ	Back propagation error
λ	Stoichiometric or normalized Air to Fuel Ratio
μ	Momentum coefficient
τ	Time step
η	Learning rate parameter

List of Subscripts

Symbol	*Description*
i	No. of input parameters or species
j	No. of neurons of first layer
k	No. of neurons of second layer
l	No. of neurons of output layer