

Study of Genetic Algorithm to Fully-automate the Design and Training of Artificial Neural Network

Osman Ahmed[†], Mohd Nordin^{††}, Suziah Sulaiman^{†††} and Wan Fatimah^{†††}

[†] *Computer & Information Sciences Department, Universiti Teknologi PETRONAS, Perak, MALAYSIA*

Summary

Optimization of artificial neural network (ANN) parameters design for full-automation ability is an extremely important task, therefore it is challenging and daunting task to find out which is effective and accurate method for ANN prediction and optimization. This paper presents different procedures for the optimization of ANN with aim to: solve the time-consuming of learning process, enhancing generalizing ability, achieving robust and accurate model, and to reduce the computational complexity. A Genetic Algorithm (GA) has been used to optimize operational parameters (input variables), and we plan to optimize neural network architecture (i.e. number of hidden layer and neurons per layer), weight, types, training algorithms, activation functions, learning rate, momentum rate, number of iterations, and dataset partitioning ratio. A hybrid neural network and genetic algorithm model for the determination of optimal operational parameter settings based on the proposed approach was developed. The preliminary result of the model has indicated that the new model can optimize operational parameters precisely and quickly, subsequently, satisfactory performance.

Key words:

Artificial Neural network, Genetic algorithm, Optimization.

1. Introduction

Optimization of artificial neural network (ANN) parameters design for full-automation ability is an extremely important task, so as to solve the time-consuming of learning process, enhancing generalizing ability, achieving robust and accurate model, and to reduce the computational complexity. A genetic algorithm (GA) has been used to attain optimum ANN parameters design.

Generally, the successful application of ANN for purpose of prediction and modeling in science and engineering domains is tremendously affected by the consequent main factors: network type (i.e. feed-forward back-propagation, recurrent networks, radial basis functions, wavelet neural network,

hopfield network, etc.), network architecture (i.e. number of hidden layers, number of neurons in each hidden layer, and NN weights), training algorithms (i.e. gradient descent, quasi-Newton, conjugate gradient, Levenberg Marquardt, resilient backpropagation algorithm, etc), activation functions (i.e. Log-sigmoid, Softmax, Linear, etc.), input selection, neural network weight, momentum rate, number of iterations, and dataset partitioning ratio.

ANN [1,2, 3], in particular, was utilized for modeling and prediction purpose due to non-linearity, time variability, and difficulty in inferring input-output mapping, whereas GA is search algorithm for optimization, based on genetic and evolution principles [4,5,6]. Current literatures on ANN show that the selection of optimal design parameters is major obstacles for their daily usage, accurate, and effective. The trial-and-error approach for determining the process parameters for ANN is no longer good, since conducting along number of experiments is required so as an optimum ANN design and training parameters get done.

Input variables selection is an important phase in neural network configuration and performance. To choose a set of significant inputs from a lot of plant processing variables is crucially important to the success, and the selection method can be extremely time-consuming.

Number of hidden layer, and number of neurons in each layers, has a significant impact in network training performance, training time and its generalization abilities. Furthermore, the use of huge number of neurons in hidden layer may overfit the data, which may causes the loss of generalization capability of network, besides, small number of neurons in hidden layer may underfit the data, subsequently the network may not be able to learn.

Selection the best ANN weights usually affects the

accuracy of producing accurate predictions even though when the optimal/near weight values are undefined due to ill-conditioning. As well, more number of ANN weight means more complex computational and is therefore prone overfitting, whereas a network with less weight may not be sufficient to model the network.

GA is widely-used technique to get the optimal ANN design parameters. Recently many research works has been conducted to solve prior problems, however, a combination of all ANN design parameters had never been conducted. Such integration is highly expected to be useful for certain problems in petroleum and power energy domains.

The main objective of this work is to examine the suitability of GA to achieve accurate, and effectiveness ANN prediction model in order to get fully-automated ANN design parameters. The sub objectives includes: prevent and reduce the time-consuming problem, reduce the computational complexity, examine the effect of scaling the input training dataset to the range of 0-1 and -1-1 on ANN predicting performance efficiency, examine the effect of partitioning strategy on ANN prediction performance efficiency, enhance the generalizing ability, and examine the effect of input data size on the ANN predicting performance efficiency.

1.1 Artificial Neural Network

ANN is a technique that used to emulate the human decision and prediction ability. In recent year, ANN technique has been applied successfully in many fields, such as automotive [7], banking [8],

electronics [9], financial [10], Industrial [11], telecommunication [12], oil and gas [13], and robotic [14].

ANN consist of several layers, which are input, hidden, and output layers. Feedforward is most popular neural network type, while backpropagation algorithm is widely used as training algorithm.

ANN modeling has three commonly used stage operators, which are training, validation, and testing network. Training stage involves adjustment to the connection (weight) that exists between neurons. To achieve this stage a numbers of data set (usually huge) are used. Validation is usually used to tune parameters, i.e., architecture, for example to choose the number of hidden layers. Eventually testing stage is used to ensure the well network generalization.

1.1.2 Genetic Algorithm

GA is a combinatorial optimization technique and general-purpose optimization method based on Darwin Theory of evolution, that searches for an optimal/near value of a complex objective function by simulation of the natural evolutionary process. GA has been successfully used in a wide variety of problem domains (Goldberg, 1989).

In brief, GA consists of three basic operators: selection, crossover, and mutation. The algorithm starts with a set of solutions to the problem under examination, the solutions set (represented by chromosomes in GA) is called the population.

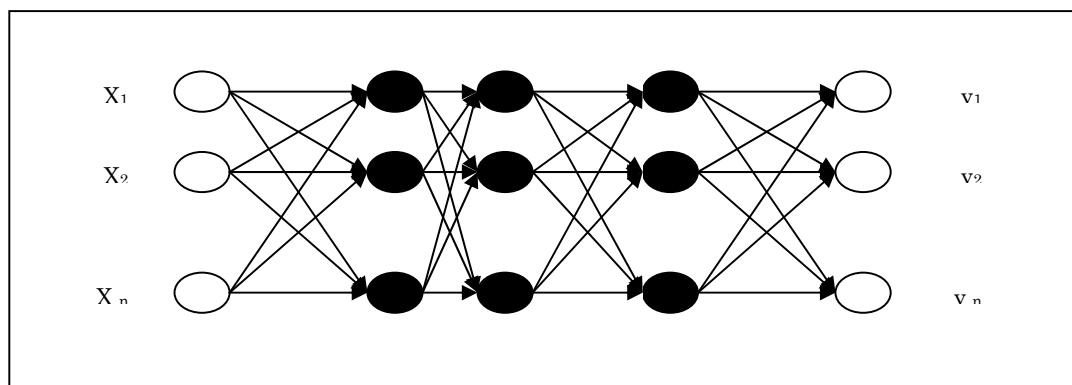


Figure1.structure of neural network of n inputs, n hidden, and n output layers

Crossover operation is used to obtain a new solution by combining different chromosomes to generate new better offspring, as well a new solution by altering existing member of the population, this operation called mutation.

2. Related Works

In recent years, many works in science and engineering domains were successfully applied, but there still no general method to design ANN. In this paper we will mainly concern with the optimization of ANN designing parameters in order to get fully-automated ANN design through GA, numerical, and statistical methods. In this aspect there is several works that have been done; these can be roughly summarized as follow:

To a void inadequate current trial-and-error practice that face of global competition a hybrid neural network and genetic algorithm approach is described to determine a set of initial process parameters for injection moulding [15]. To map the complex and nonlinear relationships among the parameters involved in the initial process parameter of injection moulding neural network (NN) was developed and implemented, and genetic algorithm (GA) was applied to search and identify a set of optimal/near optimal process parameters for injection moulding. The developed NN-GA approach consists of two parts: an NN prediction and a GA part. First, an initial population is randomly generated, which contains a number of sets of initial process parameters, Then the strings stored in it are individually fed into a NN-based prediction unit for the quality prediction of moulded parts. The developed hybrid NN-GA system significantly reduce the time that required to generate initial process parameters for injection moulding in comparison with the mould flow simulation for injection moulding it takes less than 2 minutes including the time for user input to obtain a set of initial process parameters corresponding to an input problem..

[16] Are used genetic algorithms to search for optimal hidden layer architectures, connectivity, and training parameters for ANN for predicting community acquired pneumonia among patients with respiratory complaints?

Feed-forward ANN that use back-propagation algorithm with 35 nodes in the input layer, one node in the output layer, and between 0 and 15 nodes in each of 0, 1, or 2 hidden layers,

determined by the developed genetic algorithm. Neural network structure and training parameters were represented by haploid chromosomes consisting of "genes" of binary numbers. Each chromosome had five genes. The first two genes were 4-bit binary numbers, representing the number of nodes in the first and second hidden layers of the network, which could each range from 0 to 15 nodes. The third and fourth genes were 2-bit binary numbers, representing the learning rate and momentum with which the network was trained, which each could assume discrete values of 0.01, 0.05, 0.1, or 0.5. The fifth gene was a 1-bit binary number, representing whether implicit within-layer connectivity using the competition algorithm was enabled.

Two feed-forward neural network for fault detection NN of a deep-trough hydroponic system model and a predictive modeling NN system of a similar hydroponic system are developed by , while genetic algorithm (GA) is used to encoding of NN topologies and training parameters in the field of biological engineering and more specifically in modern hydroponic plant production systems [17]. A back-propagation algorithm for the training of the NN is used as minimization algorithm. The NN mode is used to detect and diagnose possible faults in the operation by measured conditions of the system and in real time. The second NN model is used to predict one-step-ahead values of the pH and the electrical conductivity of a modern hydroponic system for plant production GA is used to optimize the selection of the minimization algorithm used by the backpropagation training algorithm, NN architecture, types of the activation functions of the hidden nodes and of the output nodes. Satisfactory performance is achieved by used both of models, and GA system successfully replaced the problematic trial-and-error method that used usually in this task.

3. Methodology

To overcome the ANN designing problems that occurred when using a cumbersome trial-and-error procedure, a methodology based on GA method have been adopted. This involves optimally designing the ANN parameters including, ANN architecture, weights, input selection, activation functions, ANN types, training algorithm, numbers of iterations, and dataset partitioning ratio, as depicted in Fig2.

3.1 Data Collection

In order to generalize the new model, three different dataset were collected. One from Universiti Teknologi PETRONAS GDC plant (TAURUS 60 gas turbine single-shaft generator set), collected on (Jan to Feb 2008) period, another dataset is collected from PETRONAS Penapisan (Melaka) Sdn Bhd from January to February 2007, the third dataset has already been published as an experimental dataset of flank wear for drilling process [18].

A 60 data set from Universiti Teknologi PETRONAS GDC plant, 500 data set from PETRONAS Penapisan (Melaka) Sdn Bhd, and 64 experimental dataset of flank wear for drilling process were used to develop the new proposed model.

3.2 pretreatment and analysis of the dataset

All data were crosschecked visually and statistically to ensure accuracy and validity of the input data both. New function has been used to remove non-number values and to apply a cut-off-error percentage (relative error) method, the latest used to remove invalid observations and to avoid the outlier error [19, 20].

Furthermore, in order to achieve best performance of ANN model the experimental dataset are further randomly partitioned into three different sets: training, validation, and testing using 3-1-1 ratio.

3.3 Preparations and scaling of the dataset

The training, validation and testing dataset are scaled to the range of (0–1) using the modified MATLAB functions 'premnmx' and 'trmnmx'. The following equation used for this purpose:

$$x_{ni} = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \quad (1)$$

Where x_i is the real-world input value, x_{ni} is the scaled input value of the real-world input value x_i and $x_{\min}$ and $x_{\max}$ are the minimum and maximum values of the unscaled dataset.

The network predicted values, which are in the range of (0–1), are transformed to real-world values using the modified MATLAB function

'postmnmx'. The equation below has been used for this purpose:

$$x_i = x_{ni}(x_{\max} - x_{\min}) + x_{\min} \quad (2)$$

3.4 ANN development model

In this step a feed-forward network ANN with back-propagation as minimization function is obtained to train three different networks (three dataset). First network is used to predict isolate pentene (iC₅) and normal pentene (nC₅) of debutanizer CRU, it consists of input layer one hidden layer, and output layer. The input layer contains three input neurons and represent the input variables, which are temperature, reflux flow, and flow rate. The hidden layer contains six nodes. The output layer contains two neurons (iC₅ and nC₅). Second network is used to predict net power and turbine outlet temperature (T4) of gas turbine single shaft; the inputs of the model are ambient conditions (T1 and P1) and turbine inlet temperature (T3), while the output are net power and T4. Furthermore the same number of hidden layers and neurons in each layer of previous network were used. Third network is used to predict flank wear of drilling process; seven neurons are used to represent input variables, i.e. diameter, speed, feed, thrust, torque, feed vibration, and radial vibration, the same number of hidden layers and neurons in each layer of first network were used, whereas flank wear is used as out variable.

In this work a Quasi Newton Algorithm based on Lavenberg Marquardt (LM) is used as training algorithm along with the sigmoid ('tansig') activation function for input and hidden layer, while the linear ('purelin') activation function is used for output layer.

To computes a single output from multiple real-valued inputs a MLP network of simple neurons called perceptrons is used. Mathematically this can be written as:

$$Y = \phi\left(\sum_{i=1}^n w_i x_i + b\right) \quad (3)$$

Where w denotes the vector of weights, x is the vector of inputs, b is the bias, and ϕ is the activation function.

3.5 ANN-GA combination

Combination of ANN and GA technique is powerful method for modeling and optimization of petroleum/energy process. In this paper, the first part of our work has already been achieved, which is to apply the GA to get the optimal operating condition for the input variables in order to achieve maximum process performance and better output values. The next part, which we are working on, is to apply GA to find other optimal ANN designing parameters.

Genetic Algorithm Toolbox included in MATLAB software was used for genetic algorithm optimization.

3.5.1 Generation of Initial Population

ANN-GA model is started with an initial population of elements which contains a predefined number of chromosomes (strings). To obtain ANN-GA model, the population size, mutation rate, and the crossover rate of the GA optimization parameters are set as 50, 0.001, and, 0.6 respectively.

3.5.2 Neural network prediction part

After the initial population is generated, the population is fed into a trained ANN for prediction purpose. The input of the network is a set of initial process parameters generated by the GA optimization part.

3.5.3 Fitness Evaluation

A fitness function has been recognized to evaluate the fitness of individual population. In this work the fitness function for the ANN evaluation is the sum squared error (SSE). If y_t is target output and y_c is the actual output, the fitness (J) can be defined as:

$$J = (w, \theta) = \sum e^2 \quad (4)$$

where $e = y_t - y_c$ and W, θ denote to the weights and biases connecting neuron unit to the preceding neuron layer. Evidently the goal is to minimize J subject to weights and biases.

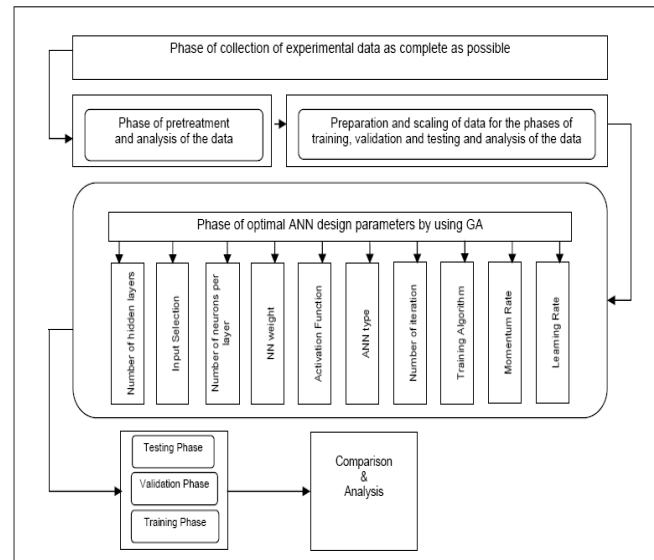


Figure2. ANN-GA Methodology

3.5.4 Termination Criteria

The evaluation process continue until some stopping applies, in this case the settling boundary and the maximum numbers of iteration are define as termination criteria in GA optimization.

3.5.5 Creation of a New Population

Generation of new population by applying reproduction operator (selection) and recombination operators (crossover and mutation), such operators are repeated until some condition satisfied.

Fig2 and 3 show ANN-GA methodology and flowchart respectively.

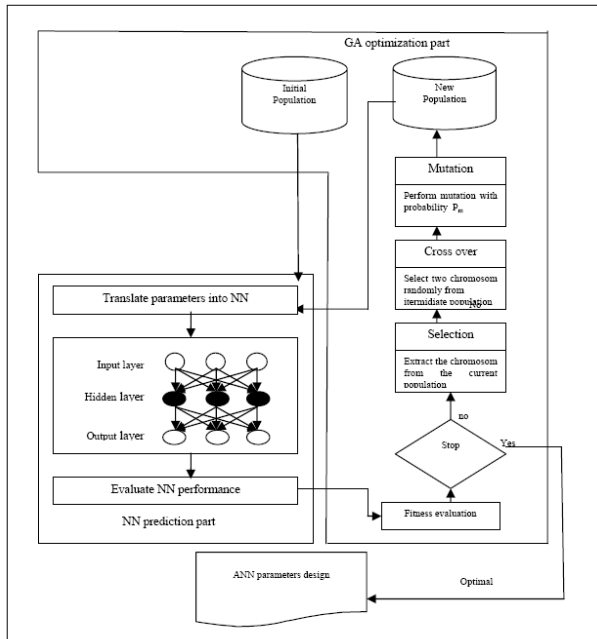


Figure3. ANN-GA Flowchart

4. Performance and accuracy measurement

Mean square error (MSE), root mean square error (RMSE), and determination coefficients (R^2) were used to evaluate ANN-GA performance. When the MSE as well as RMSE are at the minimum, and 'R' value

closer value to 1 represents high performance and perfect accuracy. The MSE, RMSE, and R^2 can be calculated by the below-mentioned equations Where T denotes the number of data patterns, Y_k is prediction output of k th pattern, T_k is target output of k th pattern, and $\bar{Y}_k$ is average of target output of k th pattern.

$$MSE = \frac{1}{T} \sum_{k=1}^T (Y_k - T_k)^2 \quad (5)$$

$$RMSE = \sqrt{\frac{1}{T} \sum_{k=1}^T (Y_k - T_k)^2} \quad (6)$$

$$R^2 = 1 - \frac{\sum_{k=1}^T (Y_k - T_k)^2}{\sum_{K=1}^T (Y_k - \bar{Y}_k)^2} \quad (7)$$

5. Results

To achieve fully-automated ANN design the methodology developed to find optimal ANN operational parameters (inputs). The new ANN-GA model was tested for prediction of three different dataset thereafter compared with ANN using trial-and-error procedure.

The maximum, minimum, and mean of the three dataset variables that used to train the network are given in Table1. The GA parameters that used to carry out above results are given in Table2. Table3 shows that ANN-GA method has achieved the highest R^2 , lowest MSE and RMSE compared with ANN using trial-and-error procedure. Computational time of ANN-GA model is illustrated in Table5.

The value of training determination coefficient (R^2) for iC_5 (0.9984), nC_5 (0.9983), net power (0.9958), T4 (0.9983), and flank wear yielded (0.9997), which are shown in Fig. 4, 5, 8, 9, 12, respectively.

GA is applied to find optimal operational parameters (input) leading to obtain the maximum outputs. Fig.6, 10, and 13 show the best fitness function value in each generation versus iteration number and convergence of the problem. Fig.7, 11, and 14 plot the best individual, which shows the vector entries of the individual among the best fitness function value in each generation.

Table1. Three Dataset Analysis

Dataset	unit	Minimum	maximum	Mean
Debutanizer variables				
Temperature	deg.c	183.57	195.67	190.99
Flow rate	m ³ /hr	59.48	93.83	87.66
Reflux flow	m ³ /hr	25.11	36.00	31.06
iC_5	%	2.70	3.28	2.66
nC_5	%	1.34	2.06	1.59
Gas turbine variables				
T1	deg.c	299.50	308.00	303.08
P1	KPag	647	922.20	765.25
T3	deg.c	572.50	581.00	576.08
T4	deg.c	500.74	909.43	658.10
Net power	KW	2110.00	3236.00	2754.58
Drilling process variables				
Diameter	mm	9.00	12.00	10.50
Speed	rpm	250.00	500.00	366.25
Feed	mm/rev	0.13	0.36	0.23
Thrust	force(N)	1088.10	3323.10	19993.61
Torque	Nm	10.67	33.11	20.14
Feed vibration	m/s ²	15.46	61.51	35.16
Radial vibration	m/s ²	16.21	62.36	36.60
Flank wear	mm	0.06	0.24	0.14

Table2. GA parameters

parameters	value
Crossover probability	0.8
Mutation probability	0.05
Population size	250
Next generation	20

Table4. ANN-GA computational time

Model	Time(s)
Debutanizer CRU	38.8152
Gas Turbine	21.9347
Flank wear	52.9617

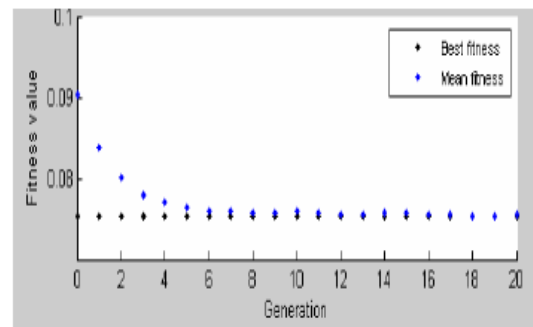


Figure.6. Best fitness convergence of the problem.

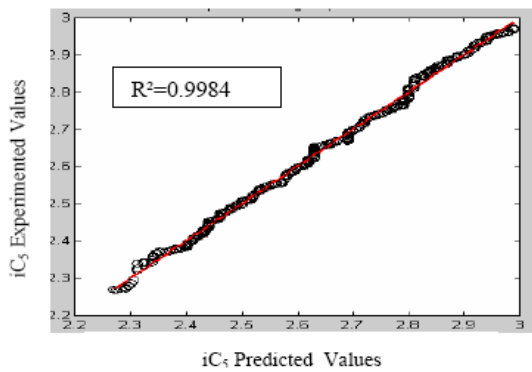
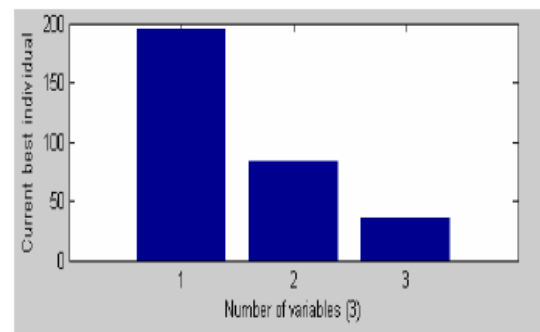
Figure.4. Predicted values of the iC_5 versus experimental values and correlation coefficient of

Figure.7. Best individual

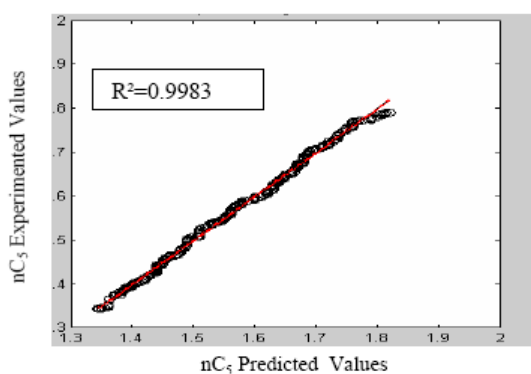
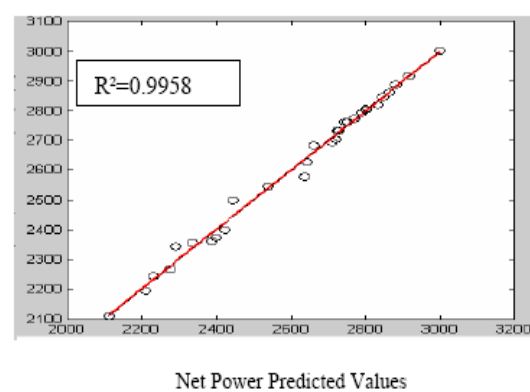
Figure.5. Predicted values of the nC_5 versus experimental values and correlation coefficient of

Figure.8. Predicted values of the net power versus experimental values and correlation coefficient of network

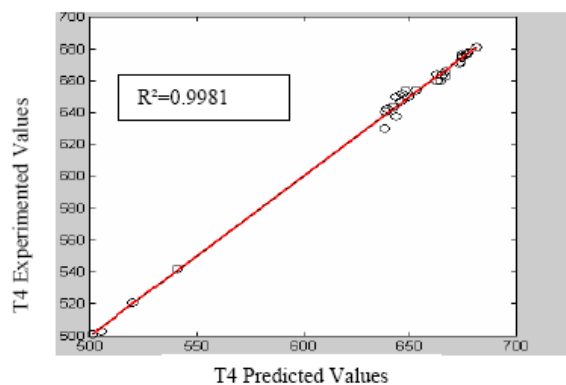


Figure.9. Predicted values of the T4 versus experimental values and correlation coefficient of network

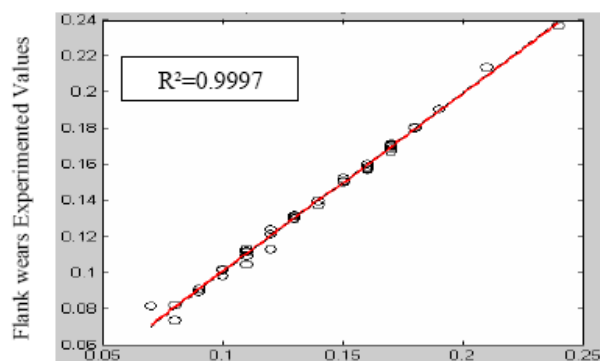


Figure.12. Predicted values of the flank wear versus experimental values and correlation coefficient of

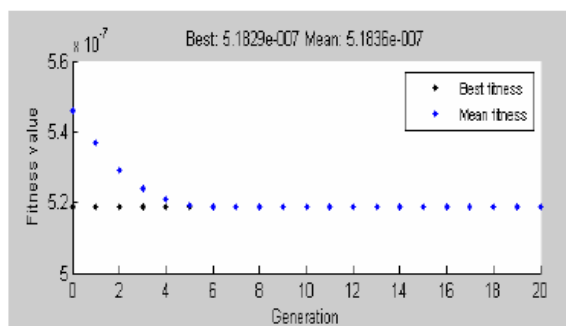


Figure10. Best fitness convergence of the problem

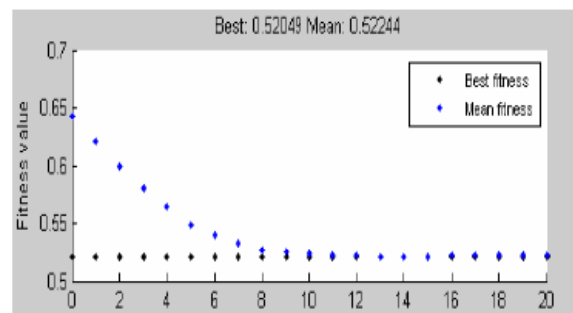


Figure13. Best fitness convergence of the problem

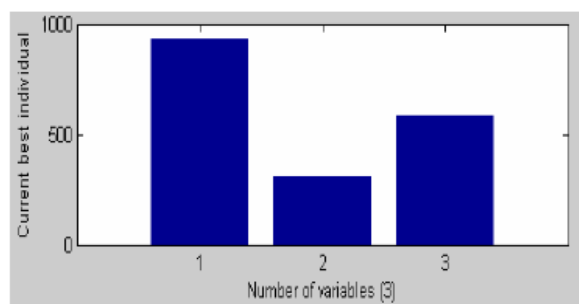


Figure11. Best individual

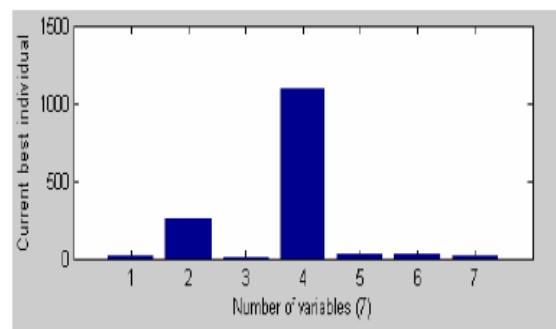


Figure14. Best individual

6. Conclusion

In this paper an ANN-GA model for fully automated ANN design and training parameters is presented. The target is to make the process of designing ANN is less human dependent and more sophisticated. Different datasets have been collected, treated, analyzed, and scaled. The GA has been applied to the ANN model so as to get optimal operating condition for input variables in order to maximize outputs. Neural network and Genetic algorithm matlab toolboxes have been used to obtain the results which proof that the new model can optimize operational parameters precisely and quickly, subsequently, satisfactory performance. As continue for this work, applying GA to find other optimal ANN designing parameters is the on progress step.

Acknowledgment

The authors would like to express sincere thanks to the staff of PETRONAS Penapisan (Melaka) Sdn Bhd and Teknologi PETRONAS GDC, for assisting in data collection and comments. In addition, the research could not be done without the support from Computer & Information Sciences Department and Postgraduate Department, Universiti Teknologi PETRONAS.

References

- [1] H. Patrick C.L., Application of artificial neural networks to the prediction of sewing performance of fabrics, *International Journal of clothing and Technology*, vol. 19 No.5. , 2007, pp. 291318.
- [2] F. Fred F., G. James D., and L. Juliet N., predicating temperature profiles in producing oil wells using artificial neural networks, *Engineering Computation*, Vol. 17 No. 6, 2000, pp. 0264-4401.
- [3] K. Okyay, Springer, 2003, "Artificial Neural Networks and Neural Information", ISBN 3540404082.
- [4] D. Satyanarayana, K. Kamarajan, and M. Rajappan, Genetic Algorithm Optimized Neural Networks Ensemble for Estimation of Mefenamic Acid and Paracetamol in Tablets, *Genetic Algorithm Optimized Neural Networks Ensemble*, Acta Chim. Slov. 2005, Volume 52, pp. 440-449.
- [5] M. Izadifar, M. Zolghadri Jahromi, Application of genetic algorithm for optimization of vegetable oil hydrogenation process, *Journal of Food Engineering*, Volume 78, Issue 1, 2007, pp. 1-8.
- [6] F. Konstantinos P., Biological engineering applications of feedforward neural networks designed and parameterized by genetic algorithms, *Neural Networks*, Volume 18, Issue 7, 2005, pp. 934-950.
- [7] M. Majors, J. Stori, C. Dong-il, Neural network control of automotive fuel-injection systems, *Control Systems Magazine*, IEEE, Volume: 14, Issue: 3, 2002, pp. 31-36.
- [8] C. Arzum E., K. Yalcin, Evaluating and forecasting banking crises through neural network models: An application for Turkish banking sector, *Expert Systems with Applications*, Volume 33, Issue 4, 2007, pp. 809-815.
- [9] L. Bor-Ren, and R. G. Hoft, Neural networks and fuzzy logic in power electronics, *Control Engineering Practice*, 2003, Volume 2, Issue 1, 2003, pp. 113-121.
- [10] Z. Xiaotian, X. Hong Wang, Li and L. Huaizu Li, Predicting stock index increments by neural networks: The role of trading volume under different horizons, *Expert Systems with Applications*, Volume 34, Issue 4, 2008, pp. 3043-3054.
- [11] G.R. Cheginia, J. Khazaeia, B. Ghobadianb and A.M. Goudarzi, 2008, Prediction of process and product parameters in an orange juice spray dryer using artificial neural networks, *Journal of Food Engineering*, Volume 84, Issue 4, pp 534-543.
- [12] N. Perambur S. and A. Preechayasomboon, development of a neuroinference engine for ADSL modem applications in telecommunications using an ANN with fast computational ability, *Neurocomputing*, Volume 48, Issues 1-4, 2002, pp. 423-441.
- [13] F. Fred F., G. James D., and L. Juliet N., predicating temperature profiles in producing oil wells using artificial neural networks, *Engineering Computation*, Vol. 17 No. 6, 2000, pp. 0264-4401.
- [14] .N. Huang, K.K. Tan and T.H. Lee, Adaptive neural network algorithm for control design of rigid-link electrically driven robots, *Neurocomputing*, Volume 71, Issues 4-6, 2008, pp 885-894.
- [15] S.L. Mok, C.K. Kwong¹ and W.S. Lau, A Hybrid Neural Network and Genetic Algorithm Approach to the Determination of Initial Process Parameters for Injection Moulding, *The International Journal of Advanced Manufacturing Technology*, Vol.18, 2001, pp. 404-409.
- [16] H. Paul S., G. Ben S., T. Thomas G., W. Robert S., Use of genetic algorithms for neural networks to predict community-acquired pneumonia, *Artificial Intelligence in Medicine*, Vol. 30, Issue 1, 2004, pp. 71-84.
- [17] F. Konstantinos P., Biological engineering applications of feedforward neural networks designed and parameterized by genetic algorithms, *Neural Networks*, Vol. 18, Issue 7, 2005, pp.934-950.
- [18] S.S. Panda a, D. Chakraborty , S.K. Pal, Flank wear prediction in drilling using back propagation neural network and radial basis function network, *Applied Soft Computing*, issue 8, (2008), pp. 858-871.
- [19] J. P. Marques de Sá, Joaquim P. Marques de Sa, Joaquim P. Marques de Sá, Springer, 2007 "Applied Statistics Using SPSS, STATISTICA, MATLAB and R", ISBN 3540719717, pp 78-83.
- [20] E.A. Osman, M.A. Ayoub, and M.A. Aggour, Artificial Neural Network Model for Predicting Bottomhole Flowing

Pressure in Vertical Multiphase Flow, SPE Middle East Oil and Gas Show and Conference, 2005.



Osman Ahmed received the BSc. and MSc. degrees, from Gezira Univ. in 2001 and 2007, respectively. After working as engineering (from 2003) in the Dept. of Informatics, Gezira Univ., His research interest includes neural networks, genetic algorithm, and their applications. Currently he working toward PhD degree at Universiti Teknologi PETRONAS.