IJCTA, 9(31), 2016, pp. 255-261 ISSN: 0974-5572
© International Science Press

Nodes Variation Analysis In Hidden Layer For Agarwood Oil Grade Prediction Using Multi-Layer Feed-Forward Neural Network

N. Ismail*, N. A. M. Ali, M. Jamil**, M. H. F. Rahiman*, S. N. Tajuddin***, and M. N. Taib***

ABSTRACT

The application of Multi-Layer Feed-forward Neural Network (MLFNN) for agarwood oil grade prediction has been presented in this paper. The work consists of extended analysis from the previous study where it is focusing on the nodes variation in hidden layer of MLFNN. The data collection started with the agarwood oil chemical compounds extraction using Gas Chromatography – Mass Spectrometry (GC-MS). Then, it was followed by identifying its significant compounds using Z-score technique. After that, the significant compounds were applied for its synthetic data in order to increase the performance of MLFNN training network. The number of nodes in hidden layer were varied from one to ten accompanied by Mean Square Error (MSE) for every node. The result shows the performance of MLFNN were different at each node. The MSE value varied between 0.0011 and 0.0088. It is found that two nodes in hidden layer yielded the smallest MSE among all. This finding is important as it contributes to the further work especially for agarwood oil grading classification.

Keywords: Agarwood Oil, Quality, Prediction, Multi-Layer Feed-Forward Neural Network.

1. INTRODUCTION

Agarwood oil is well known as precious and expensive oil [1]. It is extracted from the infected heartwood of Aquilaria species of the tree. The agarwood oil has been applied in wide areas in our life such as incense, in perfumery ingredients, in religious ceremony and traditional medical preparation where the oil extracted from agarwood trees. Mostly agarwood oil is use as perfumery ingredients as is have a strong smells.

Technically, the grading of agarwood oil is performed by professional and trained human graders depends on its physical properties such as color and odor [1, 2]. However, this method is limited due to human nose cannot accept many samples in one time and easily get fatigues especially when dealing with continuous production of the agarwood oil [3]. The human sensory panel also is limited in terms of subjectivity, poor reproducibility, time consumption and large labour expense [2]. These are constraining factors that refrain the growth of the agarwood oil trade and penetrate the market [4].

A review from literature domain suggested agarwood oil can be graded according to their chemical properties and so that there is a common standard recognized worldwide on grading the agarwood oil [4]. Analysis based on chemical profiles is required to ensure that agarwood oil can be classified based on their respective classes or grades where accurate results can be measured and automated [4-7]. The applications of MLFNN in essential oils are widely used [7-12]. One of the parameters to be adjusted during the network training in MLFNN is nodes in hidden layer size [14]. Previous study showed that the nodes variation in hidden layer gives significant impact for response range of an optical fibre pH sensor [13]. An algorithm was developed to predict antioxidant activity of known

* Faculty of Electrical Engineering, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia, Email: nrk_my@yahoo.com

** Herbal of Product Development Programme, Natural Product Division, Forest Research Institute Malaysia, 52109 Kepong, Selangor, Malaysia.

*** Faculty of Industrial Sciences & Technology, Universiti Malaysia Pahang, Lebuhraya Tun Razak, 26300 Gambang, Pahang, Malaysia.

chemical profiles in essential oil. The study brought great impact especially for food industry where its implementation for preservatives or nutraceuticals is potentially been used [12]. The research came out with satisfactory that MLFNN model is efficient in classifying different type of oils; Garda and not-Garda oil [11]. Not limited to that, there is one investigation on the behaviour of chemical profiles and the finding from MLFNN implementation contributed for the natural product [10]. The objective of this paper is to analyse the nodes in hidden layer using MLFNN for agarwood oil grade prediction. The Mean Square Error (MSE) for each node is calculated and plotted. The performance of training data on network built was tested and reported. The objective of this paper is to analyse the nodes in hidden layer using MLFNN for agarwood oil grade prediction. The Mean Square Error (MSE) for each node is calculated and plotted. The performance of training data on network built was tested and reported.

2. METHODS

This study extends the previous analysis that can be referred to Refs. [14, 15]. The methods used covered the chemical extraction using Gas Chromatography Mass Spectrometry (GC-MS), identification of significant chemical compounds by Z-score and nodes variation in hidden layer size using MLFNN.

2.1. The Agarwood Oil Sample Preparation, GC-MS Data Extraction and Z-Score

The agarwood oil samples used in this study were collected from the previous work [15, 16]. There were 16 agarwood oil samples are obtained from Forest Research Institute Malaysia (FRIM), Kepong, Selangor and Universiti Malaysia Pahang (UMP). The chemical compounds of samples are extracted by GC-MS using Standard Operation Procedure (SOP) practised at FRIM [2]. After that these compounds are fed to Z-score technique to identify their significant compounds. The concept of Z-score technique is based on the calculation of mean and standard deviation of any attribute [17]. It is also a measurement of difference between individual value and the mean population, and then divided by standard deviation of population. The computed Z-score, Z provides each feature with a zero mean and a unit variance. The foundation of Z-score is where the mathematical Gaussian curve or 'Bell Shaped' curve is applied to the data under study [3] and it is shown in Eq. (1) [4].

$$Z = \frac{x_i - \overline{X}}{SD} \tag{1}$$

Where x_i is an individual value, $\overline{X}$ is mean of population and SD is standard deviation on population. The following criteria are detailed as below:

1. The areas included between $Z = -1$ *and* $+1$ is equal to 68.27 percent. This means that 68.27 percent of the measurements will have a z-score between -1 and +1.
2. The areas included between $Z = -2$ *and* $+2$ is equal to 95.45 percent.
3. The areas included between $Z = -3$ *and* $+3$ is equal to 99.73 percent.

One of the popular training algorithms is the Leverberg Marquardt (LM), is well explained by Hagan [8]. More reading and understanding also can be found in [10-11]. The algorithm applies the function to minimize the mean square error of residual, *MSE*. The function can be expressed as in Eq. (2) [14].

$$MSE = \frac{1}{n}\sum_i (y_i - \hat{y}_i)^2 \tag{2}$$

where y_i is target, (y_i) is predicted target and n is number of sample.

2.2. Application of MLFNN

The input output obtained by GC-MS data is fed to MLFNN prediction to see their performance. The input is the abundances (%) of significant compounds and the output is agarwood quality either '0' for low or '1' for high. Due

to the network training, the data is splitted to the training and testing data with the ratio of 80:20 as recommend by the other researchers [17]. During network building, the nodes/neurons in hidden layer size is varied from 1 to 10 with the incremental of 1. Leverberg Marquardt algorithm is performed during the analysis via Matlab ver.R2010a. After that the network is tested and their MSE is recorded and reported.

3. RESULTS AND DICSUSSION

The GC-MS data consists of seven significant chemical compounds as identified by Z-score, from many compounds as extracted by GC-MS. Their abundance values (%) are fed to MLFNN as input and their quality either '0' (for low) or '1' (for high) as output. There are seven compounds identified by Z-score. They areβ-agarofuran,

Table 1
Number of Nodes in Hidden Layer with MSE.

Number of nodes in hidden layer	*MSE*
1	0.0031
2	0.0011*
3	0.0096
4	0.0073
5	0.0062
6	0.0088
7	0.0041
8	0.0059
9	0.0049
10	0.0074

Note: * is the smallest MSE

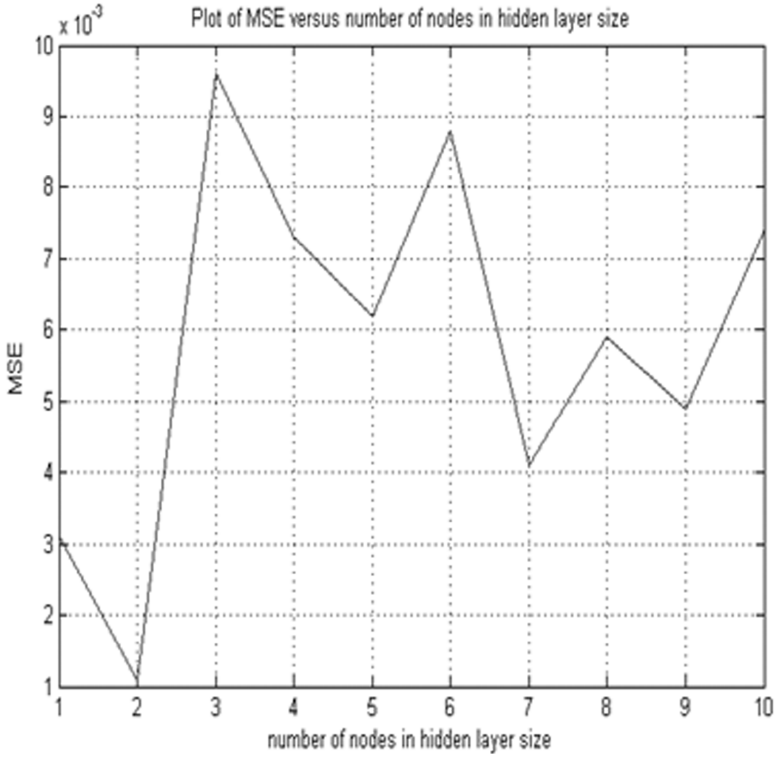

Figure 1: MSE versus Number of Nodes in Hidden Layer Size.

α-agarofuran, 10-epi-Ò-eudesmol, Ò-Eudesmol, longifolol, hexadecanol and eudesmol. The number of nodes in hidden layer with respect to MSE is tabulated in Table I. It shows that there are different MSE from one to ten nodes in hidden layer. The MSE values are ranging from 0.0011 to 0.0096. Among all, it can be seen that two nodes in hidden layers outperformed others. It is means that the MSE value for this node is the smallest, 0.0011. The graphical observation for Table I is displayed Fig. 1 for easier understanding. For the purpose documentation, only training and testing data using 2 nodes in hidden layer will be reported in this paper.

Figures 2 shows the plot of predicted target and measured target for training data using two nodes in hidden layer of MLFNN. The results show that for 77 measured target of training data, the predicted target is separated into two ranges. Range 1 is within 1.7 to 2 for Group 2 (high quality) and range 2 is within 1.0 to 1.3 for Group 1 (low quality). It also can be observed that there is big gap of predicted target for Group 1 and Group 2 and this observation contributed for further work i.e. especially for threshold value decision to classify the agarwood oil.

Figure 3 shows the prediction errors for MLFNN training using 2 nodes in hidden layer. It is observed that the errors are within 0.00 to 0.03, which is very small and acceptable for seventy eight numbers of training targets.

Figure 4 shows the prediction of the MLFNN testing to the optimized training network. It is clear that the predicted targets fall under two groups; 1 (low quality) and 2 (high quality). No errors are detected outside the region. Based the observation, the region for group 1 is between 0.8 and 1.4 and the region for group 2 is between 1.8 and 2.0. No errors are observed, i.e. no targets outside the region.

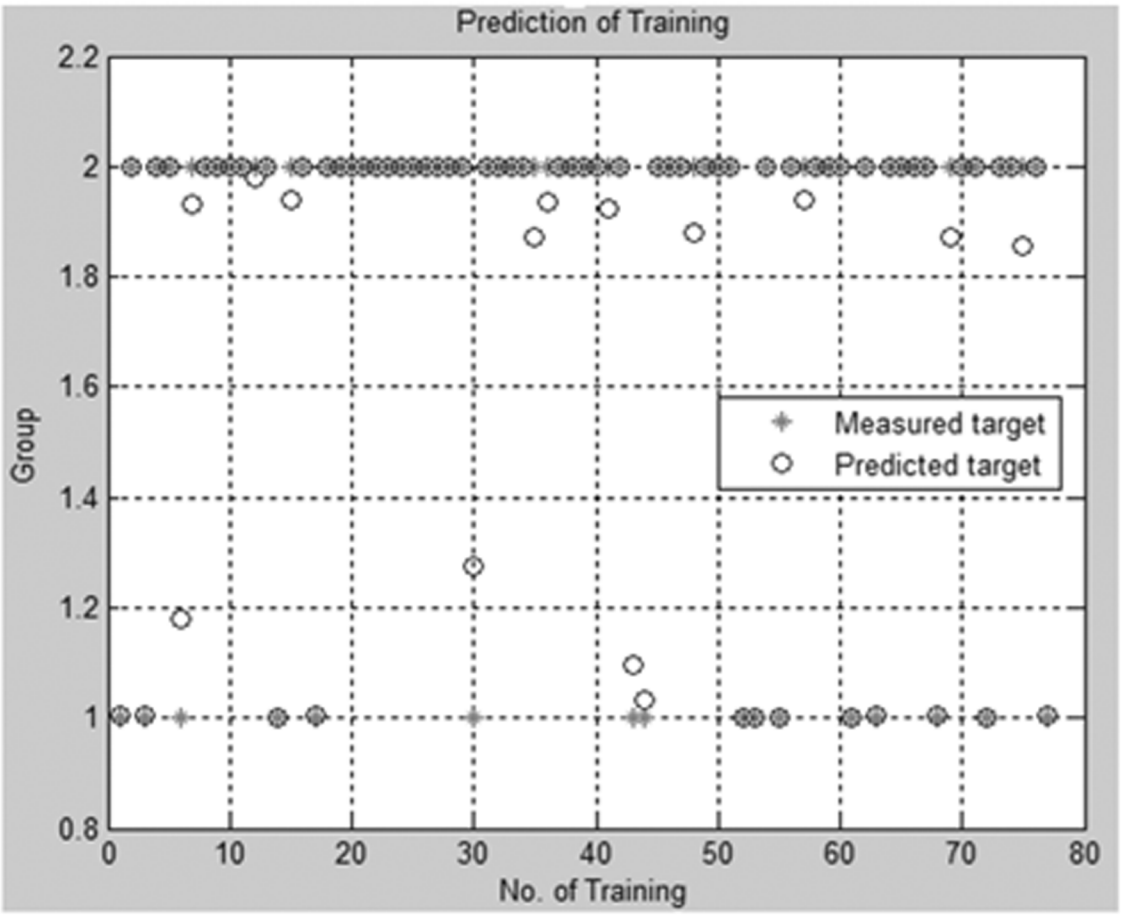

Figure 2: The Prediction Plot of Predicted Target and Measured Target for Training Data using Two Nodes in Hidden Layer of MLFNN

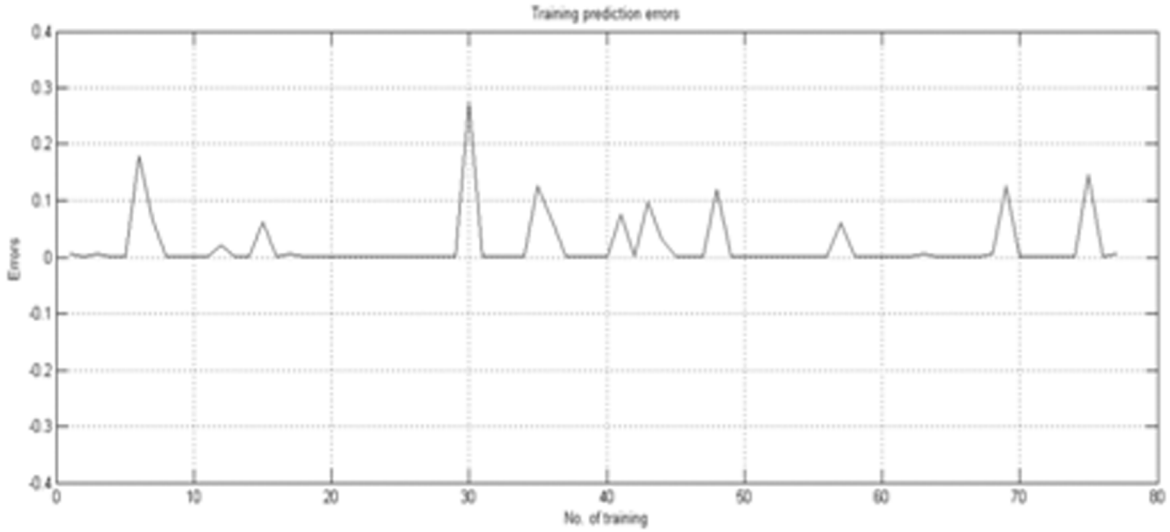

Figure 3: The Prediction Errors for MLFNN Network Training

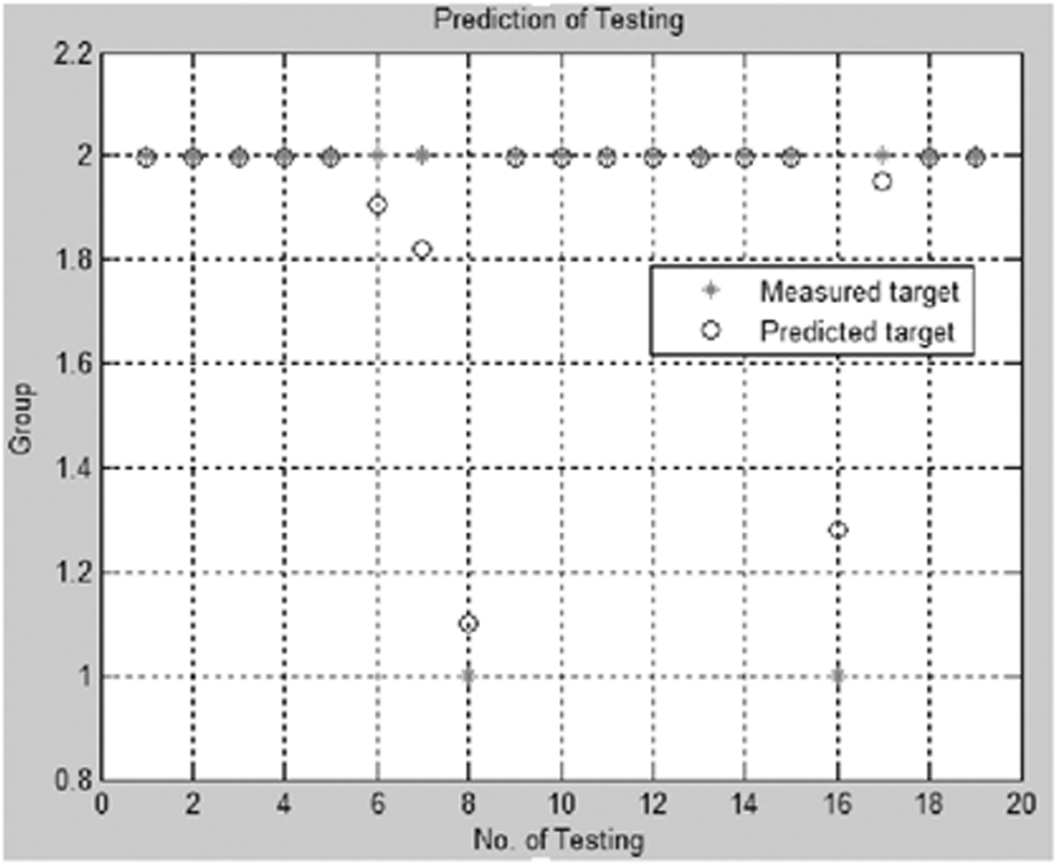

Figure 4: The Prediction Plot of Predicted Target and Measured Target for Testing Data using Two Nodes in Hidden Layer of MLFNN

Figure 5 shows the prediction errors for MLFNN testing. It is observed that the errors are within -0.03 to 0.02, which is very small and acceptable for nineteen numbers of testing targets of agarwood oil quality.

4. CONCLUSION AND FUTURE RECOMMENDATION

The application of MLFNN in agarwood oil grade prediction with nodes variation analysis in hidden layer has been presented in this study. This study showed that the nodes in hidden layer play important role as it affects the

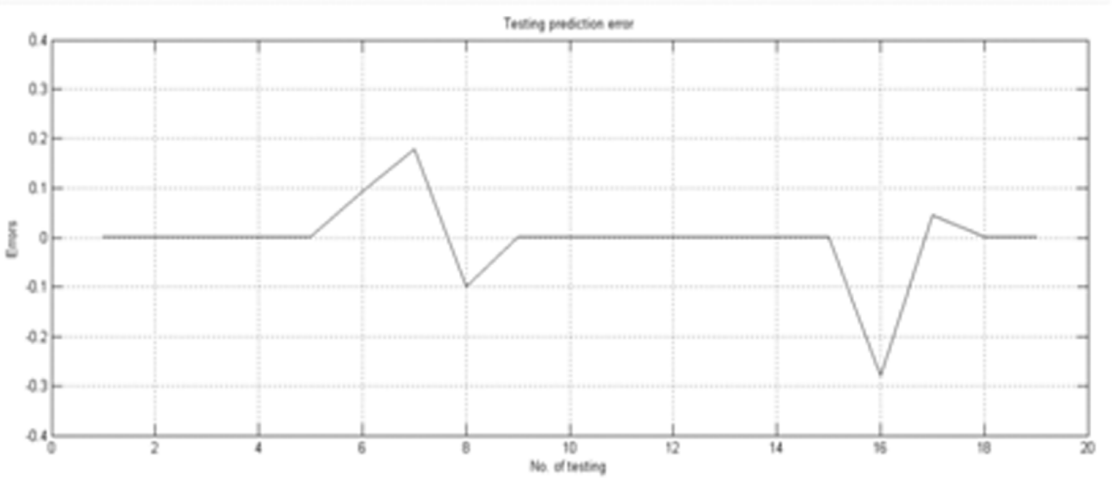

Figure 5: The Prediction Errors for MLFNN Testing

performance of the argawood oil grade prediction. The result showed that two nodes in hidden layer afforded the lowest MSE value among others. The finding is significant and useful for extended application for agarwood oil especially for oil quality classification.

REFERENCES

[1] R. Naef. (2011). The volatile and semi-volatile constituents of agarwood, the infected heartwood of Aquilaria species: a review. *Flavour and Fragrance Journal,* 26, 73-87.

[2] M. A. NorAzah, Y. S. Chang, J. Mailina, A. A. Said, J. A. Majid, S. S. Husni, H. N. Hasnida, and Y. N. Yasmin. (2008). Comparison of Chemical Profiles of Selected Gaharu Oils from Peninsular Malaysia. *The Malaysian Journal of Analytical Sciences.* 12, 338-340.

[3] M. S. Najib, M. N. Taib, N. A. M. Ali, M. N. M. Arip, and A. M. Jalil. (2011). Classification of Agarwood grades using ANN. *2011 International Conference on Electrical, Control and Computer Engineering (INECCE)*. 367-372.

[4] A. K. Pravina. (2008). *Application of solid phase microextraction in gaharu essential oil analysis*. B.Ce. Dissertation. Faculty of Chemical Engineering and Natural Resources, Univerisiti Malaysia Pahang (UMP), Pahang.

[5] M. Ishihara, T. Tsuneya, and K. Uneyama. (1993). Fragrant sesquiterpenes from agarwood. *Phytochemistry*. 33. 1147-1155.

[6] T. Nakanishi, E. Yamagata, K. Yoneda, and T. Nagashima. (1984). Three fragrant sesquiterpenes of agarwood. *Phytochemistry.* 23. 2066-2067.

[7] M. Ishihara, T. Tsuneya, and K. Uneyama. (1993). Components of the agarwood smoke on heating. *Journal of Essential Oil Research*. 5. 419-423.

[8] M. S. Najib, N. A. M. Ali, M. N. M. Arip, A. M. Jalil, and M. N. Taib. (2011). Classification of Agarwood region using ANN. *IEEE Control and System Graduate Research Colloquium (ICSGRC 2011).* 7-13.

[9] M. Najib, N. A. M. Ali, M. N. M. Arip, M. A. Jalil, and M. Taib. (2012). Classification of Agarwood using ANN *International Journal of Electrical and Electronic Systems Research*. 5.

[10] M. Jalali-Heravi and M. Fatemi. (2001). Artificial neural network modeling of Kovats retention indices for noncyclic and monocyclic terpenes. *Journal of Chromatography A*. 915. 177-183.

[11] M. S. Cosio, D. Ballabio, S. Benedetti, and C. Gigliotti. (2006). Geographical origin and authentication of extra virgin olive oils by an electronic nose in combination with artificial neural networks. *Analytica Chimica Acta*. 567. 202-210.

[12] A. C. Cabrera and J. M. Prieto. (2010). Application of artificial neural networks to the prediction of the antioxidant activity of essential oils in two experimental in vitro models. *Food Chemistry*. 118. 141-146.

[13] M. N. Taib, R. Andres, and R. Narayanaswamy. (1996). Extending the response range of an optical fibre pH sensor using an artificial neural network. *Analytica Chimica Acta*. 330. 31-40.

[14] N. Ismail, M. H. F. Rahiman, M. N. Taib, N. A. M. Ali, M. Jamil, and S. N. Tajuddin. (2014). Application of ANN in agarwood oil grade classification. *IEEE 10th International Colloquium on Signal Processing & its Applications (CSPA)*. 216-220.

[15] N. Ismail, N. A. M. Ali, M. Jamil, M. H. F. Rahiman, S. N. Tajuddin, and M. N. Taib. (2013). Analysis of High Quality Agarwood Oil Chemical Compounds by means of SPME/GC-MS and Z-Score Technique. *Malaysian Journal of Analytical Sciences.* 17. 403-413.

[16] R. W. Thatcher, C. J. Biver, and D. M. North. (2011). Z-Score EEG Biofeedback: Technical Foundations. Applied Neurosciences Inc. 2004.

[17] T. Jayalakshmi and Santhakumaran. (2011). Statistical Normalization and back propagation for classifiction. *International Journal of Computer Theory and Engineering*. 3. 89-93.