Noise-Induced Hearing Loss (NIHL) Prediction in Humans Using a Modified Back Propagation Neural Network

M. Z. Rehman[#], N. M. Nawi[*], M. I. Ghazali[#]

[#] *Fakulti Sains Komputer dan Teknologi Maklumat (FSKTM), Universiti Tun Hussein Onn Malaysia*

P. O. Box 101, 86400 Parit Raja, Batu Pahat, Johor, Malaysia
Tel.:+6074538093, E-mail: hi090004@siswa.uthm.edu.my, nazri@uthm.edu.my, imran@uthm.edu.my

Abstract—**Noise-Induced Hearing Loss (NIHL) has become a major source of health problem in industrial workers due to continuous exposure to high frequency sounds emitting from the machines. In the past, several studies have been carried-out to identify NIHL industrial workers. Unfortunately, these studies neglected some important factors that directly affect hearing ability in human. Artificial Neural Network (ANN) provides very effective way to predict hearing loss in humans. However, the training process for an ANN required the designers to arbitrarily select parameters such as network topology, initial weights and biases, learning rate value, the activation function, value for gain in activation function and momentum. An improper choice of any of these parameters can result in slow convergence or even network paralysis, where the training process comes to a standstill or get stuck at local minima. Therefore, this current study focuses on proposing a new framework on using Gradient Descent Back Propagation Neural Network model with an improvement on the momentum value to identify the important factors that directly affect the hearing ability of industrial workers. Results from the prediction will be used in determining the environmental health hazards which affect the workers health.**

Keywords— **Noise Induced Hearing Loss, adaptive momentum, back propagation neural network.**

I. INTRODUCTION

In the past four decades, World's Industry has progressed a lot and has not only benefited human kind in many ways but it also has caused adverse health effects on the human industrial workers. One of the major health problems that an Industrial worker faces today is Noise Induced Hearing Loss (NIHL). NIHL usually occurs due to continuous exposure to the noise levels of 90 decibels emitting from the heavy machines.

NIHL is a common problem identified among the workers working in the textile plants, basic metal industry, chemical industry, beverages and non-metallic mineral product industry. It was revealed in 1990's Audiometric (hearing loss test) survey by Department of Safety and Occupational Health, Malaysia (DOSH) that about 26.9 percent of industrial workers had a hearing threshold of 3000 Hz to 6000 Hz which was greater then normal and 21.9 percent of workers were already suffering from detectable hearing loss [1].

Human ear plays a vital role in the human body; it is not only a source of hearing in humans but it also helps human body in maintaining its balance. Any problem with the hearing ability damages the human's life by reducing the quality of communication [2]. Hearing loss is defined mathematically as in Equation (1):

$$hl = 10 \log \frac{I}{Io} \, dB \qquad (1)$$

where,

I : threshold sound intensity for the persons ear and,
Io : threshold sound intensity of the normal hearing

NIHL when detected at early stages can be stopped but in later stages hearing loss becomes permanent. Various studies have been carried out to detect NIHL in humans, but the recent improvements in the technology especially in Neural Networks has paved a way for researchers to predict various harmful effects of noise on humans such as human work

efficiency in noisy environment, noise induced sleep disturbance, speech interference in noisy environment, noise induced annoyance [3]-[8].

In a study carried out on NIHL [9], three variables such as age, work duration and noise exposure were selected and Levenberg-Marquadt (LM) model was used for hearing impairment prediction in industrial workers. In another study, on tympanic membrane perforation, three factors were identified that directly affect human workers (i.e. noise level, frequency and duration of exposure). It also negated the fact that age; an important factor in permanent hearing loss in older people can play the same effect on the young people [2]. Both studies on NIHL are in full-agreement that noise levels in excess of 90 decibels can cause permanent hearing loss but still some important factors that can be helpful in finding harmful effects of NIHL in human hearing are neglected.

Mostly the input parameters that have been used by the audiometric experts for detecting NIHL is unclear and not standardized as the data collected is often not precise and the environmental conditions are not suitable for the collection. For the sake of precision, this research proposes a new framework to improve the working performance of Back Propagation Gradient Descent Neural Network (BPGD-NN) model proposed by Nazri [10], [11] that will change adaptively the momentum coefficient during the training. The proposed framework will be implemented using the input parameters (e.g. noise level, frequency, duration of exposure, age, type of activity, individual's sensitivity to noise, health conditions and heat) to classify/predict the NIHL and its effects on workers.

The rest of the paper is organized as follows: the next sections describe the Artificial Neural Network (ANN), Back Propagation Neural Network (BPNN), the effect of using the momentum coefficient in BPNN. Section-3 introduces the proposed adaptive momentum algorithm for BPGD-NN Model proposed by Nazri [10], [11]. Finally the paper is concluded in the Section-4.

II. ARTIFICIAL NEURAL NETWORK

Artificial Neural Networks (ANN) are modelled on the human brain and consists of processing units known as artificial neurons that can be trained to perform complex calculations like human brain. Unlike traditional methods in which an output is based on the input it gets, a neuron can be trained to store, recognize and estimate patterns without having the information about the form of function. Due to ANN's high success rate in solving many complex real-world problems such as predicting future trends on the basis of huge historical data of an organization they have been successfully implemented in all engineering fields such as biological modelling, decision and control, ocean exploration and so on. [12]- [16]

A. Back-Propagation Neural Network

The Back-Propagation Neural Network (BPNN) is the most novel and oldest supervised learning ANN algorithm proposed in 1986 by Rumelhart, Hinton and Williams [17]. BPNN learns by calculating the errors of the output layer to find the errors in the hidden layers. Due to this ability of Back-Propagating, it is highly suitable for problems in which

no relationship is found between the output and inputs. The gradient descent method is utilized to calculate the weights and adjustments are made to the network to minimize the output error. The error function at the output neuron is defined as;

$$E = \frac{1}{2}\sum_{k=1}^{n}(t_k - o_k(\alpha_k))^2 \qquad (1)$$

Where,

n : number of output nodes in the output layer.

t_k : desired output of the kth output unit.

o_k : network output of the kth output unit.

α_k : momentum coefficient

1) BPNN with momentum coefficient (α)

Multilayer feed-forward Neural Network training using gradient descent BPNN requires parameters such as network topology, initial weights and biases, learning rate value, activation function, and value for the gain in the activation function should be selected carefully. An improper choice of these parameters can lead to slow network convergence, network error or failure. Seeing these problems, many variations in gradient descent BPNN algorithm have been proposed by previous researchers to increase the training efficiency. Some of the variations are the use of learning rate and momentum to speed-up the network convergence and avoid getting stuck at local minima. These two parameters are frequently used in the control of weight adjustments along the steepest direction and for controlling oscillations [18].

Momentum (α) is a modification based on the observation that convergence might be improved if the oscillation in the trajectory is smoothed out, by adding a fraction of the previous weight change [17], [19]. It has been revealed through various studies that Back-propagation with Fixed Momentum Coefficient (BPFM) shows acceleration results when the current downhill gradient of the error function and the last change in weights are in the similar directions, when the current gradient is in an opposing direction to the previous update, BPFM will cause the weight direction to be updated in the upward direction instead of down the slope as desired, so in that case it is necessary that the momentum coefficient should be varied adaptively instead of being kept fixed [20].

To overcome Static Momentum problem various methods for adaptive momentum have been developed by researchers. One such variation used a momentum step and dynamically selects the momentum rate. Using one-dimensional error minimization technique, the proposed BPNN algorithm was able to successfully converge on problems like 8-3-8 and 10-5-10 encoders [21]. Xiangui rejected the idea of using one-dimensional error minimization technique stating that error function is a very complex non-linear function with respect to the learning rate but it can be proved that optimal gradient vectors in two successive iteration steps are orthogonal. Based on this property one can use the Graham-Schmidt Orthogonalization method to ensure the orthogonality of the successive gradient vectors. This results in automatic

updating of momentum term in each successive iteration and oscillations are suppressed and error is greatly reduced at the end of final convergence [22]. In another study, relatively large momentum and learning rate was used on problems like XOR, the convergence rate was greatly accelerated but the use of larger momentum and learning rate was not found feasible as iterations were found highly unstable [23]. In 1994, Simple Adaptive Momentum (SAM) was proposed as a way of further improving the performance of BPNN. The momentum term is scaled according to the similarity between the changes in the weights at the current and previous iterations. If the change in the weights is in the similar 'direction' then the momentum term is increased to accelerate the convergence otherwise they are decreased. SAM has been found to have lower computational overheads then the Conjugate Gradient Descent and conventional BPNN algorithm and it converges in considerably less iteration on XOR and SINEWAVE period problems. Although found better then the conventional BPNN and Conjugate Gradient method, its success and failure rate is same like BPNN [24].

Concerned with the effect of learning rate and momentum on network training time, an efficient Back Propagation and Acceleration Learning Method (BPALM) was introduced to reduce the training time of conventional BPNN. The method was tested on Parity problem, Optical Character Recognition (OCR) and 2-Spirals problem, the results were found to be far superior then any other previous improvements on BPNN [25]. In 2009, R. J. Mitchell considered adjusting momentum differently in SAM [24] in which the scaling of the momentum term is found by considering all the weights in the Multi-layer Perceptrons (MLP). The momentum term was adjusted differently in each part of the MLP, by considering the weights only in that part of the MLP. This technique helps improve convergence speed to the global minimum [26]. Hongmei Shao and Gaofeng Zheng introduced a Back Propagation momentum Algorithm (PBPAM), where the momentum coefficient is adjusted dynamically by combining the information about the current gradient and the weight change in the previous step. When the angle between the current negative gradient and the last weight change is less than 90°, the momentum coefficient is defined as a positive value to accelerate learning. Otherwise, to guarantee the descent of the error function the momentum coefficient is termed as zero. The performance of the new algorithm was applied to the typical benchmark problem, i.e. XOR; the new algorithm not only outperforms the previous BPNN's by reducing training iterations as well as it smooth out oscillations in the network [20].

III. THE PROPOSED FRAMEWORK

Nazri [11], states that there are various reasons for the slow convergence in gradient descent. Sometimes the magnitude and direction components of the gradient vector are responsible for the slow convergence. When the error surface is fairly flat along a weight dimension, the derivative of the weight is small in magnitude. Therefore many steps are required and weights are adjusted by a small value to achieve a significant reduction in error. On the other hand, if the error surface is highly curved along a weight dimension, the derivative of the weight is large in magnitude.

Thus, large weight value adjustments may overshoot the minimum of the error surface along that weight dimension. Another reason for the slow rate convergence of the gradient descent method is that the direction of the negative gradient may not point directly toward the minimum error surface. Based on previous researches on the effect of momentum, to speed-up the convergence and to make weight adjustments efficiently on the gradient descent, a new framework is proposed to change the momentum adaptively.

A. Algorithm

The proposed algorithm uses batch mode of training in which momentum, weights and biases are updated for the complete training set which is presented to the network:

For each epoch,
    For each input vector,
        Step-1:
        Calculate the weights and biases using the previous momentum value
        Step-2:
        Use the weights and biases to calculate new momentum value.

Repeat the above steps until the network reaches the desired value.

B. The Derivation of the proposed framework

Adaptive Momentum is used to avoid oscillations in the network while searching the global minimum on the error surface. It smooths-out the descent path and helps your network in avoiding getting stuck in the local minima due to extreme changes in the gradient [27]. Adaptive Momentum, generates a value for the weight updates in a network. Here, the weight updates are limited to [0,1] as Log-sigmoid activation function is used to find the output on the jth node;

$$O_j = \frac{1}{1 + e^{-\alpha_j a_{net,j}}} \qquad (2)$$

where,

$$a_{net,j} = \left[\sum_{i=1}^{l} w_{ij} O_i \right] + \theta j \qquad (3)$$

where,

O_j : Output of the *jth* unit.

O_i : Output of the *ith* unit.

W_{ij} : weight of the link from unit *i* to unit *j*.

$a_{net,j}$: net input activation function for the *jth* unit.

θ_j : bias for the *jth* unit.

187

$\dfrac{\partial E}{\partial \alpha_k}$, needs to be calculated for the output units and $\dfrac{\partial E}{\partial \alpha_j}$ is also required to be calculated for hidden units, so that the respective momentum value can be updated in the Equation (6):

$$\Delta \alpha_k = \left(-\frac{\partial E}{\partial \alpha_k} \right) \qquad (4)$$

$$\Delta \alpha j = \left(-\frac{\partial E}{\partial \alpha_j} \right) \qquad (5)$$

$$\frac{\partial E}{\partial \alpha_k} = (t_{k\_}O_k)O_k(1-O_k)\left(\sum w_{jk}O_j + \theta_k\right) \qquad (6)$$

The momentum update expression from input to output nodes becomes;

$$\Delta\alpha_k(n+1) = (t_{k\_}O_k)O_k(1-O_k)\left(\sum w_{jk}O_j + \theta_k\right) \qquad (7)$$

$$\frac{\partial E}{\partial \alpha_j} = \left[-\sum_k \alpha_k w_{jk}(t_{k\_}O_k)O_k(1-O_k) \right]O_j(1-O_j)\left[\left[\sum_j w_{ij}O_i\right] + \theta_j\right] \qquad (8)$$

Therefore, the momentum update expression for the hidden units is:

$$\Delta\alpha_j(n+1) = \left[-\sum_k \alpha_k w_{jk}(t_{k\_}O_k)O_k(1-O_k) \right]O_j(1-O_j)\left[\left[\sum_j w_{ij}O_i\right] + \theta_j\right] \qquad (9)$$

Weights and biases are calculated in the same way, the weight update expression for the links connecting to the output nodes with a bias is;

$$\Delta w_{jk} = (t_{k\_}O_k)O_k(1-O_k)\alpha_k O_j \qquad (10)$$

Similarly, bias update expression for the output nodes will be;

$$\Delta\theta_k = (t_{k\_}O_k)O_k(1-O_k)\alpha_k \qquad (11)$$

The weight update expression for the input node links would be:

$$\Delta w_{ij} = \left[\sum_k \alpha_k w_{jk}(t_{k\_}O_k)O_k(1-O_k)\right]\alpha_j O_j(1-O_j)O_i \qquad (12)$$

And, finally the bias update expression for hidden nodes will be like this;

$$\Delta\theta_j = \left[\sum_k \alpha_k w_{jk}(t_{k\_}O_k)O_k(1-O_k)\right]\alpha_j O_j(1-O_j) \qquad (13)$$

IV. CONCLUSIONS

NIHL is detected as a major health problem in the workers of the present times. Many studies have been conducted by local as well as international regulatory and private bodies and they have come-up with standards for noise exposure time periods for a person. But still people are getting affected with NIHL, which means that there is a need of some standard that can predict NIHL precisely, so that health conditions can be improved in the industries. Back-Propagation Neural Network has been used widely in the practical fields and has a strong capability of classifying problems, but it has problems of slow convergence and network stagnancy, which still needs to be answered. So to predict NIHL in a better way, and to speed-up the BPGD-NN [10], [11] a new framework to improve current working BPGD-NN is introduced which modify adaptively the momentum coefficient during the training. In the next publication the performance criteria of the proposed adaptive momentum algorithm will be evaluate based on the speed of convergence, CPU time and the percentage level of the correct predictions for diagnosing NIHL in industrial workers. The simulations will be carried out on a Pentium Dual Core PC with 3GHz processor speed and 1GB RAM. The proposed algorithm's performance will be compared with the standard Gradient Descent Momentum (traingdm) from MATLAB Neural Network Toolbox version 4.01. Steps will be taken to make this algorithm efficient enough to predict NIHL in humans effectively and according to the criteria set by the World Health Organization (WHO) and DOSH. The results based on the NIHL prediction will be published in the next publication.

REFERENCES

[1] M. S. Leong, "Noise and Vibration Problems: How they effect us and the industry in the Malaysian Context," University Teknologi Malaysia, Skudai, Johar, Malaysia, 2003, pp.01-13.
[2] Zaheeruddin and V.K. Jain, "An intelligent system for noise-induced hearing loss," Proc. ICISIP 2004, 24 August, pp. 379-384.
[3] Zaheeruddin, G.V. Singh, V.K.Jain, "Fuzzy modelling of human work efficiency in noisy environment," Proc. Fuzzy Systems 2003, 25-28 May, pp.120-124.
[4] Zaheeruddin and Garima, "Application of Artificial Neural Networks for Prediction of Human Work Efficiency in Noisy Environment," Proc. CIMCA 2005, Vienna, Australia, 25-28 November, pp.842-846.
[5] Zaheeruddin, V.K.Jain, "A fuzzy expert system for noise-induced sleep disturbance," Expert Systems with Applications, vol. 30(4), pp.761-771, May. 2006.
[6] Zaheeruddin and V. K. Jain, "Fuzzy modeling of speech interference in noisy environment," Proc. ICISIP 2005, 4-7 January, pp.409-414.
[7] Zaheeruddin, V. K. Jain, G. V. Singh, "A Fuzzy Model for Noise-Induced Annoyance," IEEE Transaction on System, Man and Cybernetics, vol. 36(4), pp. 697 – 705, July 2006.
[8] Zaheeruddin, "Modelling of Noise-induced Annoyance: A Neuro-fuzzy Approach," Proc. ICIT 2006, Mumbai, India, 15-17 December, pp. 2686-2691.
[9] M.N. Yahya, M.I. Ghazali, "Hearing Impairment Prediction on Malaysia Industrial Workers by Using Neural Network," Proc. 8th International Conference on Quality in Research (QIR), Bali, Indonesia, 9-10 August 2005.

[10] N. M. Nawi, M. R. Ransing, and R. S. Ransing, "An improved Conjugate Gradient based learning algorithm for back propagation neural networks," International Journal of Computational Intelligence, vol. 4(1), pp. 46-55, 2007.

[11] N. M. Nawi, "Computational Issues in Process Optimization using historical data," Ph.D Eng. Thesis, Swansea University, United Kingdom, 2007.

[12] B. Kosko, "Neural Network and Fuzzy Systems," 1st Edition, Prentice Hall of India, 1994.

[13] V.M. Krasnopolsky and F. Chevallier, "Some Neural Network application in environmental sciences. Part II: advancing computational efficiency of environmental numerical models," Neural Networks, vol. 16(3-4), pp.335-348, April-May 2003.

[14] B. Coppin, "Artificial Intelligence Illuminated," Jones and Bartlet Illuminated Series, USA, Chapter 11, pp.291-324, 2004.

[15] I. A. Basheer, M. Hajmeer, "Artificial neural networks: fundamentals, computing, design, and application." Journal of Microbiological Methods, vol. 43(1), pp. 03-31 December 2000.

[16] He. Zheng, Wu. Meng, B. Gong, "Neural Network and its Application on Machine fault Diagnosis," Proc. ICSYSE 1992, 17-19 September, pp.576-579.

[17] D. E. Rumelhart, G.E. Hinton, R. J. Williams, "Learning Internal Representations by error Propagation," Parallel Distributed Processing: Explorations in the Microstructure of Cognition, vol. 1, 1986.

[18] Y.H. Zweiri, L D. Seneviratne, K. Althoefer, "Stability Analysis of a Three-term Back-propagation Algorithm," Neural Networks, vol. 18(10), pp.1341-1347, Dec. 2005.

[19] M. A. Fkirin, S. M. Badwai, S. A. Mohamed, "Change Detection Using Neural Network in Toshka Area," Proc. NSRC, 2009, Cairo, Egypt, 17-19 March, pp.1-10.

[20] H. Shao and G. Zheng, "A New BP Algorithm with Adaptive Momentum for FNNs Training," Proc. GCIS 2009, Xiamen, China, 19-21 May, pp. 16-20.

[21] G. Qiu, M.R. Venley and T.J.Terell, "Acceleration training of Backpropagation Networks by using adaptive momentum step," IET Electronic letters, vol. 28(4), pp.377-379, Feb. 1992.

[22] X. Yu, N.K. Loh and W.C. Miller, "A new acceleration technique for the back propagation algorithm," Proc. ICNN 1993, San Francisco, USA, 28March-01April, pp.1157-1161.

[23] X.-H.Yu, G.-A.Chen and S.-X.Cheng, "Acceleration of Backpropagation learning using optimised learning rate and momentum," Electronic Letters, vol. 29(14), pp:1288-1290, July.1993.

[24] D. J. Swanston, J.M. Bishop, and R. J. Mitchell, "Simple adaptive momentum: New algorithm for training multilayer Perceptrons," Electronic Letters, vol. 30(18), pp. 1498-1500, Sept 1994.

[25] C. Yu and B. Liu, "A Backpropagation algorithm with adaptive learning rate and momentum coefficient," Proc. IJCNN 2002, Honolulu, USA, 12-17 May, pp. 1218-1223.

[26] R. J. Mitchell, "On Simple Adaptive Momentum," Proc. CIS 2008, London, United Kingdom, 9-10 September, pp.01-06.

[27] H. Yan, Y. Jiang, J. Zheng, C. Peng and Q. Li, "A Multilayer Perceptron-Based Medical Decision Support System for Heart Disease Diagnosis, " *Expert Systems with Applications,* vol. 30(2), pp. 272–281, 2006.