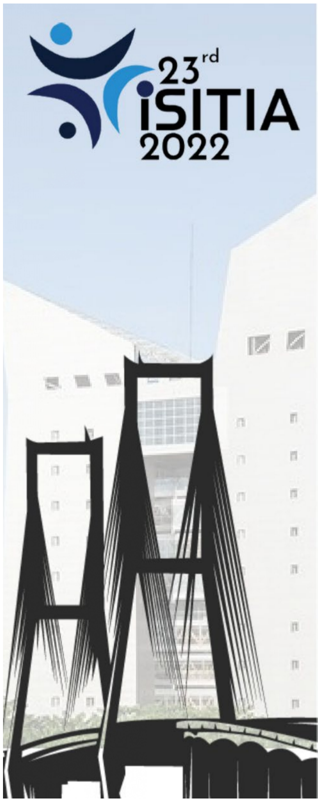

CERTIFICATE

No. : 266/ISITIA22.CERT/VII/2022

It is hereby certified that

Ahmed Mellad

has contributed as

Presenter

of the paper with the title

"Accelerated Adaptive Learning Rate"

at the 23rd International Seminar on Intelligent Technology and Its Applications (ISITIA) 2022 on Virtual Conference July 20th-21st 2022

Surabaya, July 21st 2022

PROF. DR. IR. SOEDIBYO, M.MT.

ISITIA 2022 General Chair

Accelerated Adaptive Learning Rate

Ahmed Nasraden Mellad
Department of Electrical Engineering
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia
ahmed.302216@mhs.its.ac.id

Gamantyo Hendrantoro
Department of Electrical Engineering
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia
gamantyo@ee.its.ac.id

Wirawan
Department of Electrical Engineering
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia
wirawan@ee.its.ac.id

***Abstract*— An artificial neural network (ANN) is a kind of human brain learning method software simulation. ANN is suitable for classification issues and has proven successful in many practical applications. Signal interference, power consumption, and signal noise are the main issues in signal processing research. This paper proposes AALR; this project attempts to develop a neural network demodulator and machine learning technology. As a result of the new concepts, the neural network demodulator processing speed has improved by 88%, increasing output accuracy by 7%.**

Keywords—learning rate, machine learning, neural network demodulator, signal processing.

I. Introduction

Wireless communications have become one of the fastest-growing areas in our modern lives and have created an enormous impact on nearly every feature of our daily lives. Signal modulation is mixing information or data with an electric signal to transform the data into form an electromagnetic wave to transmit it via the atmosphere. Digital modulation techniques are distinguished from analog modulation techniques by their capacity for communicating superior amounts of information. The process of extracting data from the modulated signal is called signal demodulation. Researchers have recently used an artificial neural network as a signal demodulator for its ability to enhance signal symbols classification. Still, researchers end up with the slowness issue of neural network processing, making it challenging to use it in existing communication systems.

To overcome this issue, Zeng et al. [2] presented an algorithm called automatic noise reduction, which identifies and removes noisy data using the ANNs multi-layer framework. The mechanism performed well for the noise level below 30% but degraded as the noise level exceeded 50 %. Zhao et al. [3] proposed convolutional neural networks (CNN). However, the error rate of quadrature phase shift keying (QPSK) demodulation is still higher than that of binary phase shift keying (BPSK).

II. Learning Rate

Learning rate is a hyper-parameter that manages the adjustment of the network weights concerning the loss gradient. If the *lr* value is too small, the network needs more learning loops to reach the lowest possible error rate. Otherwise, the neural network will exceed the minimum error rate. Therefore, the result is inaccurate output. In other words, the amount of change in the weight variable of each learning epoch is called a learning rate or a learning step. A primary factor determines the number of iterations to achieve the most accurate result with the fewest learning iterations possible.

In this paper, the accelerated adaptive learning rate concept is proposed. This concept improves the signal processing field by developing neural network demodulators. This research will focus mainly on ANNDs as an application of neural networks in the communication field.

III. Supervised Learning

The goal of supervised learning methods is to detect the relation between independent variables and a dependent variable, also known as input attributes and a target attribute, respectively. The relation is embodied structurally in a model. In general, models describe phenomena inherent in the dataset, making it useful for predicting the target attribute given the input attributes. Hence, the supervised learning methods offer potential applications in various fields, such as finance and manufacturing.

The supervised models can be classified into two: classification models and regression models. The former plot the input space into predefined classes. As an example, classifiers can distinguish between mortgage consumers based on their payback timeliness. On the other hand, the former is able to map the input space into the real valued domain. In this case, given the characteristics of a product, regression can predict the demand for that product.

Other methods to represent classifiers are available, such as support vector machines (SVM), decision tree, summaries of probabilistic, etc. Classification is one of the most researched models and has found many applications. The benefits of classification have been increasing owing to its contribution to other fields in data mining.

IV. Neural Network Demodulator

A neural network demodulator (NND) is an ANN explicitly customized for digital signal processing. NND is used extensively for the demodulation process of the digital signal. The demodulation process of NND is injecting the modulated digital signal into the input layer nodes of NND. The input values ($X_1, X_2, \ldots, X_i$) indicate the modulated signal. The input, neural network weights, and output values can be real, binary, or bipolar numbers.

The each node output in the hidden layer is expressed as follows:

$$X_j = \frac{1}{1+e^{-\sum_i W_i X_i}} \tag{2.2}$$

Where $\frac{1}{1+e^{-x}}$ is known as Sigmoid Activation Function.

Figure 2 illustrate the proposed multi-layer NND. The input layer comprises a y(t) sample of each symbol from the transmitted QAM signal.

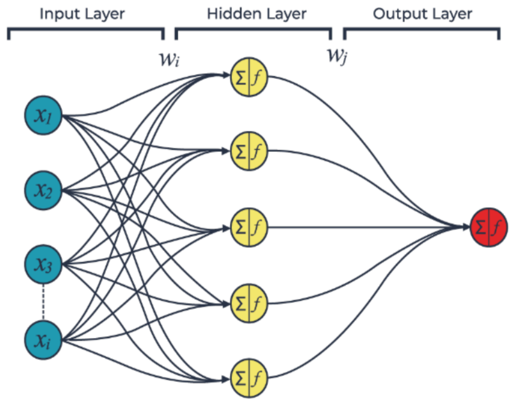

Fig. 2: NND with multi-layer for QAM signal.

V. LEARNING ALGORITHM

A. *Error Function and Gradient Descent*

In the learning process, an error function E is defined by quantifying the difference between result of computation by the perceptron and the actual value of the input over a set of input-output pairs. An error function is typically defined over the set: the typical way is to define an error function E over the pairs set $X=\{(\vec{x_1}, \vec{y_1}),........,(\vec{x_n}, \vec{y_n})\}$ such that $E(X, \theta)$ is small when $f_\theta(\vec{x_1}) \approx (\vec{y_1})$ for all i where θ is ANN's parameters (weight vectors and biases) and f_θ is an activation function.

Mean squared error (MSE) commonly used for E in regression problems or to use the cross-entropy in classification. Hence, given X, the training serves to minimize error $E(X, \theta)$ with respect to the parameter values. For instance, for the MSE function with two input-output pairs $X = \{(\vec{x_1}, \vec{y_1}), (\vec{x_2}, \vec{y_2})\}$ and an ANN with parameters θ that outputs $f_\theta(\vec{x_1})$ for input $\vec{x}$, the error function is expressed as follows:

$$E(X,\theta) = \frac{(y_1 - f_\theta(\vec{x}_1))^2}{2} + \frac{(y_2 - f_\theta(\vec{x}_2))^2}{2} \tag{2.3}$$

The minimum of the error function can be found by invoking the gradient descent method, which is chosen due to its capability in minimizing differentiable functions. The gradient descent evaluates the function f gradient at a point theta, finds the direction with the minimum gradient, and moves toward the direction, moving depends on step size η, finding a nearby value $x' = x - \eta * \nabla f(x)$ which $f(x')<f(x)$, by repeating this process until a minimum is found, or becomes smaller than some threshold.

B. *Backpropagation*

Backpropagation or backward-propagation is a training algorithm used for neural networks supervised learning method by using gradient descent. giving the neural network an error function to determines the error function gradient regarding the weights of the neural network.

Using gradient descent to train a neural network requires the error function gradient calculation $E(X, \theta)$ with respect to the weights w_{ij}^k and biases b_i^k. Then, each iteration of gradient descent updates the weights and biases (collectively denoted θ) according to the learning rate η, the neural network parameters update can expressed as follows:

$$\theta_{t+1} = \theta_t - \eta \frac{\partial E(X,\theta_t)}{\partial \theta} \tag{2.4}$$

where θ_t presents the neural network parameters at iteration t in gradient descent.
The error function in classic backpropagation is the mean squared error

$$E(X,\theta) = \frac{1}{2N}\sum_{i=1}^{N}(\acute{y}_i - y_i)^2 \tag{2.5}$$

Where y_i is the target value and $\acute{y}_i$ is the network output.

Backpropagation is a solution to a long-time problem of gradient determination caused by the complexity of the ANN structure. Backpropagation is applicable for both feed-forward and recurrent ANNs and makes it possible to train non-trivial ANNs.

C. *Accelerated Adaptive Learning Rate*

The updated value of learning rate for each epoch can be expressed as follows:

$$\eta_t = \begin{cases} lr \tanh\left(\dfrac{\eta_{t-1}}{\sqrt{1-\dfrac{\alpha^2}{lr^2}}}\right) & \text{if } \left|\dfrac{\partial E(X,\theta_t)}{\partial \theta}\right|_t \le \left(\dfrac{\partial E(X,\theta_t)}{\partial \theta}\right)_{t-1} \\ lr \tanh\left(\left(\dfrac{\eta_{t-1}}{\sqrt{1-\dfrac{\alpha^2}{lr^2}}}\right)^{-1}\right) & \text{if } \left|\dfrac{\partial E(X,\theta_t)}{\partial \theta}\right|_t > \left(\dfrac{\partial E(X,\theta_t)}{\partial \theta}\right)_{t-1} \end{cases}$$

$$\alpha_t = \begin{cases} lr \tanh(\alpha_{t-1} + \delta) & \text{if } \left|\dfrac{\partial E(X,\theta_t)}{\partial \theta}\right|_t \le \left(\dfrac{\partial E(X,\theta_t)}{\partial \theta}\right)_{t-1} \\ lr \tanh((\alpha_{t-1} + \delta)^{-1}) & \text{if } \left|\dfrac{\partial E(X,\theta_t)}{\partial \theta}\right|_t > \left(\dfrac{\partial E(X,\theta_t)}{\partial \theta}\right)_{t-1} \end{cases} \tag{2.6}$$

Where η represents the learning rate, α denotes the accelerated step of updating the learning rate value and $\delta = lr * 10^{-2}$ where lr is the range of learning rate for lr>0 and $\frac{\partial E(X,\theta_t)}{\partial \theta}$ is known as gradient of error function and t is the iteration.

VI. QUADRATURE AMPILTUDE DEMODULATION

Quadrature Amplitude Modulation (QAM) is a modulation scheme used in digital communication systems. It uses two orthogonal carriers, the in-phase and the quadrature, that are each amplitude-modulated.

QAM is used extensively as a digital telecommunication systems modulation scheme. Setting constellation size of QAM can achieve arbitrarily high spectral efficiency, limited by the linearity of the communication channel and the noise level. QAM is usually square, with the most common forms are 16-QAM, 64-QAM, and 256-QAM.

VII. Simulation and Discussions

A. *Simulation Experiment*

ANN demodulator in this experiment contains three-layer. The input layer to receive the modulated signal, the hidden layer, and the output layer to output the desired data. PYTHON programming language is used to build the ANN demodulator modules, in which the Sigmoid transfer function is applied as the activation function of all nodes for all network layers.

The learning algorithm of the neural network is the backpropagation algorithm. The simulation details are as follows:

Type: 16-QAM coherent demodulator.
Filters: Band-pass and Low-pass filter.
Channel model: AWNG.
Data size = 1000 bit.
fc = 4 kHz.
fs = 80 kHz.
snr = -15 dB
Phase offset at received signal = $\frac{\pi}{4}$.
NN demodulator parameters:
Iteration = 1000
Input data: Signal Samples, 100 sample per symbol
Activation function: Sigmoid
Learning rate: Accelerated Adaptive Learning Rate (AALR)

B. *Simulation Result*

Figure 3 shows the difference of the neural network performance using constant learning rate and Accelerated Adaptive Learning Rate, and figure 4 presents the BER versus SNR. This experiment resulted in a single hidden layer neural network for 16-QAM signal with SNR = 10.

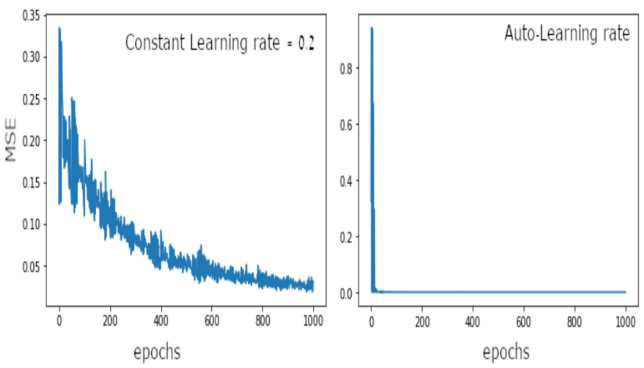

Fig. 3: The difference between a constant learning rate and the new AALR in the learning rate range = 10

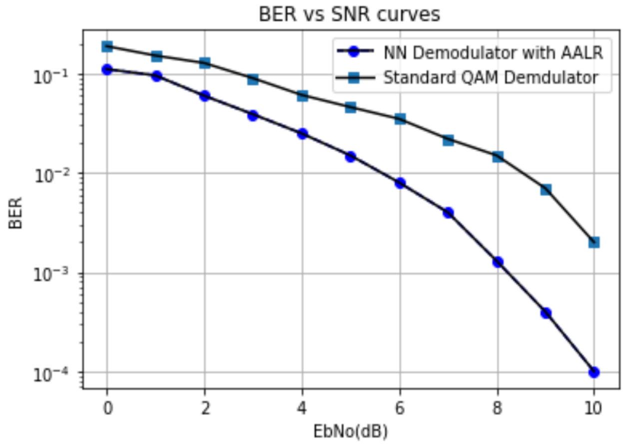

Fig. 4: The BER vs SNR comparison between standard 16-QAM demodulator and neural network 16-QAM demodulator AALR

VIII. Conclusion

A constant learning rate (*lr*) has shown to be less effective in obtaining fast and accurate output in signal processing. Artificial Neural Networks Demodulators (ANNDs) use the accelerated adaptive learning rate concept to enhance the accuracy of the output signal through the backpropagation machine learning algorithm. The learning process requires an error function *E* to quantify the difference between the computed output of the perceptron and the actual value for the input. Since finding the minimum analytically is impossible, gradient descent is used to find a function gradient at a particular value and then update that value. A single hidden layer neural network for 16-QAM signal with SNR = 10 results in BER 10^{-2} based on 55 errors compared to the standard demodulator which results in BER 10^{-1} based on 128 errors, the neural network demodulator processing speed has improved by 88%, increasing output accuracy by 7%.

References

[1] Lu, Xingyu & Lu, Chao & Yu, Weixiang & Qiao, Liang & Shangyu, Liang & Lau, Alan Pak Tao & Chi, Nan. (2019). Memory-controlled deep LSTM neural network post-equalizer used in high-speed PAM VLC system. Optics Express. 27. 7822. 10.1364/OE.27.007822.

[2] Xinchuan Zeng and T. Martinez, "A noise filtering method using neural networks," IEEE International Workshop on Soft Computing Techniques in Instrumentation, Measurement and Related Applications, 2003. SCIMA 2003., 2003, pp. 26-31, doi: 10.1109/SCIMA.2003.1215926.

[3] Ruobing Zhao *et al* 2021 *J. Phys.: Conf. Ser.* 2026 01200

[4] K. Elissa, "Title of paper if known," unpublished.

[5] R. Nicole, "Title of paper with only first word capitalized," J. Name Stand. Abbrev., in press.

[6] Behrouz A. Forouzan. (2007). Date communications and Networking. 4th Edition. McGraw-Hill.

[7] Yussy Perdana, Dina Herdiana, Hanny M, Andrian B, Achmad Munir. (2015). "Linear Frequency Modulated Continuous Wave Radar Using GNU Radio and USRP". 9784673846/15 IEEE.

[8] Yu Hen Hu, Jenq-Neng Hwang. (2002). "Handbook Of Neural Network Signal Processing". by CRC Press LLC.

[9] Ershad Junus Amin, Andriyan Bayu Suksmono, and Achmad Munir. (2014). "Accuracy Analysis of FM Chirp in GNU Radio-based FMCW Radar for Multiple Target Detection". 9781-4799-4575-7/14 IEEE.

[10] Mohammad Reza Amini. (2010). "A BFSK Neural Network Demodulator With Fast Training Hints". 978-0-7695-3961-4/10 IEEE.

[11] Song, wenmiao. (2009). "Configure Cognitive Radio using GNU Radio and USRP". 978-1-4244-4076-4/09 IEEE.

[12] Nitish Srivastava, Geo rey Hinton. (2014). "Dropout: A Simple Way to Prevent Neural Networks from Overfitting. Journal of Machine Learning Research 11-13.

[13] M. Amini and E. Balarastaghi. (2011). "Universal Neural Network Demodulator for Software Defined Radio". International Journal of Machine Learning and Computing, Vol. 1, No. 3, August.

[14] Sooyong Choi and Daesik Hong. (2000). "A Hybrid Structured Neural Network Receiver in Digital Communication Systems". 0-7695-0619-4/00 lEEE.

[15] Michael L. Seltzer, Dong Yu, Yongqiang Wang. (2013). "An Investigation of Deep Neural Networks for Noise Robust Speech Recognition". 978-1-4799-0356-6/13 IEEE.

[16] Lubna Badri. (2010). Development of Neural Networks for Noise Reduction. The International Arab Journal of Information Technology, Vol. 7, No. 3, July.

[17] Rosa M. Giraldo. (2008). Quadrature Amplitude Modulation (QAM) Microware Signal Transmission over a Radio-Over-Viber Link Using Semiconductor Optical Amplifiers. University of Limerick, Dept. Electronic and Computer Engineering.

[18] TRI T. HA. (2011). Theory and Design of Digital Communication Systems. Cambridge University.

[19] Dr. K.V. Prasad. (2003). Principles of Digital Communication Systems and Computer Networks. Charles River Media.

[20] R. E. Ziemer, W. H. Tranter. (2002). Principles of Communications Systems, Modulation and Noise. 5th Edition, John Wily & Sons, Inc.