IMPLEMENTATION OF ARTIFICIAL NEURAL NETWORK (ANN) USING BACKPROPAGATION ALGORITHM BY COMPARING FOUR ACTIVATION FUNCTIONS IN PREDICTING GOLD PRICES

Dian Kurniasari[1*], Ranti Vidia Mahyunis[2], Warsono[3], Aang Nuryaman[4]
[1,2,3,4]Department of Mathematics, Faculty of Mathematics and Natural Science,
University of Lampung
Jl. Prof. Dr. Ir. Sumantri Brojonegoro No.1, Gedong Meneng, Rajabasa, Bandar Lampung, , Lampung , 35141, Indonesia
*Email correspondece: [1]dian.kurniasari@fmipa.unila.ac.id

Abstract

The trend in global currency values is speedy and fluctuating due to the recession caused by the Covid-19 pandemic. That causes investors to flock to buy gold assets. Therefore, it is necessary to predict the price of gold from a business and academic perspective to obtain a reasonable gold price prediction model. This study applies the Backpropagation Algorithm by determining the best ANN model structure based on four activation functions: Sigmoid, Tanh, ReLU, and Linear, as well as learning rate values, namely 0.01 and 0.001. The results are the best ANN model structure with four nodes in the input layer, four nodes in the hidden layer and the output layer using the Linear activation function and a learning rate of 0.01. Based on the structure of the model, the MSE value is 0.00051, the MAPE value is 1.9798%, and the accuracy is 98%.

Keywords*: Artificial Neural Network, Backpropagation, Gold Price Prediction, Activation Function, Model Structure*

Abstrak

Trend nilai mata uang global sangat cepat dan fluktuatif akibat terjadinya resesi yang disebabkan oleh pandemi Covid-19. Hal ini menyebabkan, para investor berbondong-bondong untuk membeli aset emas. Oleh sebab itu, perlu dilakukan prediksi harga emas, baik dari perspektif bisnis maupun akademis agar memperoleh model prediksi harga emas yang baik. Penelitian ini menerapkan Algoritma Backpropagation dengan menentukan struktur model ANN terbaik berdasarkan empat fungsi aktivasi yaitu, Sigmoid, Tanh, ReLU, dan Linear serta nilai learning rate, yaitu 0,01 dan 0,001. Hasil yang diperoleh berupa struktur model ANN terbaik dengan empat node pada input layer, empat node pada hidden layer dan output layer dengan menggunakan fungsi aktivasi Linear dan learning rate sebesar 0,01. Berdasarkan struktur model tersebut, diperoleh nilai MSE sebesar 0.00051, nilai MAPE sebesar 1,9798% dan akurasi sebesar 98%.

Kata Kunci: *Artificial Neural Network, Backpropagation, Prediksi Harga Emas, Fungsi Aktivasi, Struktur Model*

1. INTRODUCTION

Gold is one of the primary commodities traded in global financial markets. Gold is considered a safe asset compared to fiat currencies due to the physical value it has, and from ancient times investors flocked to gold assets when the market was not performing well.

Recently due to the global economic recession caused by the Covid-19 pandemic, investors have flocked to buy gold-related assets as global currencies lose value rapidly and are highly volatile.

In this scenario, predicting the gold price is very important from both a business and academic perspective because it will enable financial practitioners to predict the future financial situation [1].

Predicting what will happen based on the available information is called prediction. Prediction aims to recognize systematic patterns and find patterns of trend relationships based on historical data. The first stage in predicting is collecting historical data [2]. One approach in predicting that has been widely used is the Artificial Neural Network (ANN). Artificial Neural Network is a programming algorithm applied to computer computing processes by adapting the structure and characteristics of human biological neural networks [3]. There are various algorithms in ANN, one of which is the backpropagation algorithm.

The Backpropagation Algorithm is a systematic ANN method that uses a supervised learning algorithm which has an outstanding advantage in conducting training, especially on large-scale and complex data [4]. The process in the backpropagation algorithm is divided into 3 main stages: feedforward, backpropagation, and changes in weight values. The backpropagation algorithm reduces the error rate by adjusting the weights based on the desired output [5]. The activation function is one of the crucial components in building an ANN model using the backpropagation algorithm. The activation function has a vital role in an artificial neural network which will determine the magnitude of the weight and act as a signal in determining the output to several other neurons [6].

Research using the ANN model with the backpropagation algorithm model has been carried out before, including predicting the number of tourists in NTB using the backpropagation method by Lestari et al., which obtained a meagre error difference value [7]. Untoro research in predicting stock prices using ANN. This study obtained maximum results in making predictions with a relatively low percentage of error values [8]. Then, research by Nafi'ah to predict gold prices using the Linear Regression, Backpropagation and Fuzzy Mamdani methods, with an accuracy of 95% [9]. That indicates that the backpropagation method is superior to the other two methods because it has the highest accuracy. Therefore, this study will apply the ANN model with the backpropagation algorithm in predicting gold prices using a different structure by comparing the 4 activation functions and adding components such as batch size, optimizer, and learning rate in building the model. Based on this structure, the level of effectiveness, accuracy, slightest error difference, and gold price prediction results will be assessed.

2. RESEARCH METHODS

The data used in this study is secondary data in the form of historical data on the gold price index for 6 years which was downloaded directly from the Yahoo Finance website: https://finance.yahoo.com/. This study uses the ANN model with a backpropagation algorithm to predict the closing price of gold prices with the help of the Python programming language. The steps of this research, namely:

a. Input the gold price data obtained from the yahoo finance website into the Python programming language using Google Colab.
b. Conduct descriptive analysis by visualizing gold price data into time series plots.
c. Performing data preprocessing, namely:
   1) Check if any data is missing.
   2) Perform data transformation using Min-Max Normalization.
   3) Dividing 80% of training data, 13, 59% of data validation 6.41% of testing data.
d. Build the ANN model with the backpropagation algorithm:
   1) Determine the selection of input variables and output variables.
   2) Determine the components and parameters of the network model, namely the hidden layer and its nodes, batch size, learning rate, epoch, and based on four activation functions.
   3) Build a model using predefined components based on 4 activation functions.
e. Check for overfitting and underfitting using the learning curve of the loss and validation values divided in step 3c.
f. Make predictions using testing data.
g. Validate the model based on the MSE value, MAPE value, and the accuracy obtained.
h. Perform data visualization on the results of gold price predictions with models best based on four activation functions.

Some of the theories used are as follows.

2.1. Data Mining

Data mining is an information search technique and exciting patterns from data collections in databases, data warehouses, or other information repositories [10]. Several techniques in data mining that are often used in various studies include Clustering, Classification, Association Rules, and Neural Networks [11]. Data mining is often called knowledge discovery in databases (KDD) [12]. Knowledge discovery in the database consists of data cleaning, data integration, data selection, data transformation, mining process, pattern evaluation and knowledge presentation. (knowledge presentation) [13].

2.2. Machine Learning

Machine Learning is an approach in Artificial Intelligence (AI) widely used to replace or imitate human behaviour in solving problems or automation [14]. Machine learning requires learning data, which is known as training data. This method has two kinds of training: supervised learning and unsupervised learning. The difference between the two types of training depends on how the training algorithm uses the type of pattern [3].

2.3. Artificial Neural Network

An artificial Neural Network is an information processing system of an artificial intelligence system that has the same characteristics as the workings of the human biological nervous system. Artificial Neural Network is formed based on the following assumptions [3]:

a. The simple elements used in processing information are called neurons.
b. There is a link between neurons that is used as a signal carrier.
c. Each connecting joint has an appropriate weight which will be processed with the connected signal.
d. Each neuron has an activation function by adding the input weights to determine the output signal.

2.3.1. Artificial Neural Network Architecture

In general, there are three layers in the ANN architecture: the input layer, the hidden layer, and the output layer. Each of these layers can have a different number of nodes and neurons.

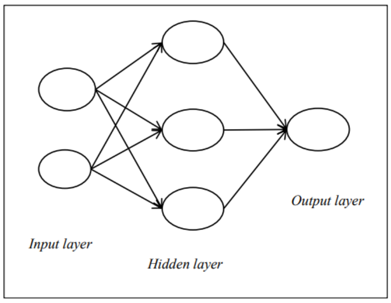


Figure 1. Artificial Neural Network Architecture

Based on Figure 1, the ANN architecture layer consists of [3]:

a. Single Layer Network
   This network consists of one input layer and one output layer. This network only accepts input and then directly processes it into output without going through the hidden layer.
b. Multilayer Network
   This network has 3 types of layers, namely: an input layer, an output layer, and a hidden layer. This multilayer network can solve complex problems compared to single-layer networks.

2.3.2. Components and Parameters of Artificial Neural Network Model

Building an ANN model with the backpropagation algorithm uses the following components and parameters.

2.3.2.1. Hidden Layer

Several rules can be used to determine the number of nodes in the hidden layer [15].

a. The number of nodes is the number of input layers and the number of output layers.
b. The number of nodes can be determined using the formula: (number of inputs ×2/3) + number of outputs.

The number of nodes must be less than double the number of input layers.

2.3.2.2. Activation Function

The activation function is a function that describes the relationship between the levels of internal activation (summation function), which may be linear or non-linear [6]. Several activation functions are often used in the backpropagation algorithm, namely [16]:

a. Sigmoid Activation Function

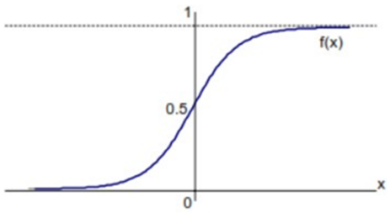


Figure 2. Sigmoid Function

Sigmoid Activation Function; is a non-linear function with a range of 0 to 1. The Sigmoid function is often used for input values in the form of real numbers and outputs that lie in the interval 0 to 1. The following is the formula for the Sigmoid function:

$$F(x) = \frac{1}{1+e^{-x}}$$

b. ReLU Activation Function

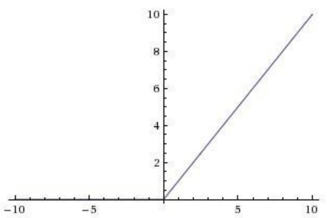


Figure 3. ReLU Function

The ReLU Activation Function is an activation layer in the model that applies the function (x) = max $(0, x)$, meaning that if x ≤ 0 then x = 0 and if x > 0 then x = x.

c. Linear Activation Function

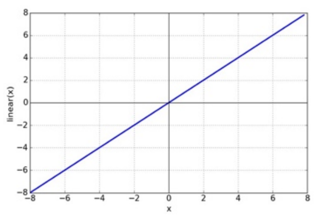


Figure 4. Linear Function

The Linear Activation Function is a function that has the same output value as the input value, and is formulated as follows:

Linear Function: f(x) = x

d. Tanh Activation Function

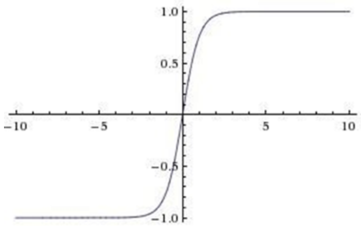


Figure 5. Tanh Function

The Tanh Activation Function is a non-linear function that will change its input value to a value that has a range between -1 to 1. The following is the formulation of the Tanh function:

$$F(x) = \tanh(x) = \frac{2}{1+e^{-x}} - 1$$

2.3.2.4. Learning rate

One of the training parameters for calculating weight correction values during the training process is the learning rate. This learning rate value is in the range of zero from 0 to 1. The greater the learning rate value, the faster the training process will run. The greater the learning rate, the less accurate the network is, and vice versa [17].

2.3.2.5. Batch Size and Optimizer

Batch size is a hyperparameter that defines the number of samples to work with before updating the internal model parameters. The training dataset can be divided into one or several batches. Generally, the batch size used are 32, 64, and 128. Optimizer or optimization is an auxiliary function in determining a more precise weight to drastically reduce the number of trials and errors when doing backpropagation. Adam is one of the optimizers, an adaptive learning rate method where Adam calculates individual learning rates for different parameters [18].

2.3.2.6. Epoch and Iteration

The process of training a dataset on a neural network for one round is called an epoch. Because one epoch is too large to be fed into the computer, it is necessary to divide it into small units (batches). While iteration is the number of batches needed to complete one epoch. One iteration has been done if the forward and backward pass process has been done once.

2.4. Algoritma Backpropagation

Backpropagation is an ANN method that uses a supervised learning algorithm. The backpropagation method was initially designed for feedforward neural networks, but in its development, it was adapted for learning in other neural network models [19]. The feature of this method is to minimize errors in the output generated by the network. In conducting training on a network using the backpropagation algorithm, there are 3 stages carried out, namely as follows:

The algorithm I: the feedforward phase is where each input is calculated forward from the hidden layer to the output layer using the specified activation function.
Algorithm II: In the backpropagation phase, the error (difference between output and the desired target) is counted backwards from the line directly related to the units in the output layer.

Algorithm III: Changes in weight and bias is the phase of weight modification to reduce errors.

2.5. Model Validation

Model validation is essential for prediction, namely how to measure the suitability between existing and predicted data. Calculating the error value is very important to see the training results on the ANN. That is because, at the training stage, the error value is expected to be the smallest value. Several ways can be used to evaluate the model (function) that has been generated [20]:

a. Mean Squared Error (MSE.) Value

Mean Squared Error is generally used to evaluate the average error in predictive value search. The MSE calculation formula is as follows:

$$MSE = \frac{\sum_{i=1}^{n}(y_i - t_i)^2}{n}$$

With $y_i =$ actual value, $t_i =$ predictive value, $n =$ and lots of data.

Since MSE is measured in square units, root means square error (RMSE) can be interpreted as the average distance between the predicted and observed values, measured in units of the target variable with the following formula:

$$RMSE = \sqrt{MSE}$$

b. Mean Square Percentage Error (MAPE.) Value

The MAPE calculation aims to measure statistics about prediction accuracy using the absolute error in each period divided by the actual value and then averaging the percentage error. The MAPE calculation formula is as follows:

$$MAPE = \frac{\sum_{t=1}^{n}\frac{|y_i - t_i|}{y_i}}{n} \times 100\%$$

With $y_i =$ actual value, $t_i =$ predictive value, $n =$ and lots of data.

3. Results and Discussion

The first step is a descriptive analysis of gold price data from 2020 to 2022 with a total of 1554 data. Descriptive analysis was carried out to provide an overview of the data in verbal and numerical form. A descriptive analysis of gold price data is shown in Figure 6.

Figure 6. Gold Price Trend Chart 2016-2022

Based on Figure 6, it is clear that the price of gold continues to increase. However, the price of gold has also decreased, although not significantly. The peak of the highest gold price occurred on August 6, 2020, at US$ 2051.5 per troy ounce. In contrast, the lowest gold price fell on January 25, 2016, worth US $ 1106.2 per troy ounce, which means that gold experienced a significant price increase over the past 4 years. Meanwhile, in 2021-2022, gold prices will fluctuate, namely the increase and decrease in gold prices, but with insignificant fluctuations. The average gold price from 2016 to 2022 is US$ 1476.1 per troy ounce.

After doing descriptive analysis, the next step is to do data preprocessing. The first step is to check for missing data. The results of the missing data check are shown in Table 1.

Table 1. Checking Missing Data on Gold Price Data

Variable	Amount of Missing Data
Open	2
High	2
Low	2
Close	2
Volume	31

Table 1 shows that the amount of missing data is not too large. Therefore preprocessing of lost data is done by deleting all rows or records that contain missing data in the dataset, namely as many as 31 rows. The next stage is to transform the data. The transformation used is the Min-Max transformation

After performing data transformation, the next preprocessing stage is data sharing. Data distribution is divided into training, testing, and validation. The composition used is 80% training with 1213 training data, while 13.59% validation with 212 validation data and 6.41% testing with 100. The model is trained with training data, while the evaluation is done with data validation and testing with data testing.

The next step is to prepare to build an ANN model with the backpropagation algorithm. Building a backpropagation algorithm ANN model requires the division of input and output variables on the gold price data as well as the components and parameters of the network model. The division of input and output variables is presented in Table 2.

Table 2. Division of Input and Output Variables

Input Variable	Output Variable
Open High Low Volume	Close

The next stage is to determine the structure or network components. The built ANN structure will be assessed based on its accuracy by determining the best number of neurons or nodes in the hidden layer. Based on Heaton's rules, the number of nodes in the hidden layer that meet the requirements is 2, 3, and 4. The next component is the learning rate. The learning rate directly affects the speed of the training process. The learning rate values used in this study were 0.01 and 0.001.

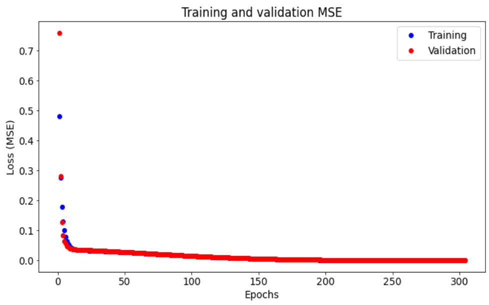
Then determine the maximum epoch of 1000 epoch by applying the early stopping technique to stop the training process when the performance difference between training and validation has exceeded a predetermined limit. Then determine the component batch size, iteration, optimizer, and activation function. The batch size used is 64 batch sizes with a total of 1312 training data, so it takes 19 iterations to complete one epoch. The optimizer used is the Adam optimizer because Adam is relatively easy to configure, computationally efficient, and suitable for parameters that have big problems. Four activation functions used are Sigmoid, Tanh, ReLU, and Linear. The activation functions will be compared to obtain the best error and accuracy values.

The following process is to build an ANN model to obtain prediction results, accuracy and error difference. The ANN model obtained is one hidden layer model with a learning rate of 0.01, a two hidden layer model with a learning rate of 0.01, a one hidden layer model with a learning rate of 0.001, and a two hidden layer model with a learning rate of 0.001. The four models produce the best structure, which is determined based on the loss value obtained in Table 3.

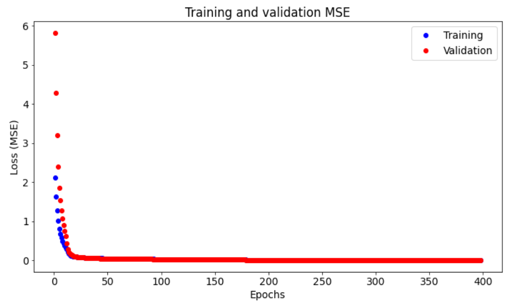
Table 3. Loss Value with the Best Structure in 4 Models

Model	Activation Function	Batch Size	Epoch	Iteration	Loss Value
Model 1 Hidden Layer (4 nodes) with LR = 0,01	Linear	64	304	5776	0,00051
Model 2 Hidden Layer (4 and 3 nodes) with LR = 0,01	ReLU	64	398	7562	0,00125
Model 1 Hidden Layer (4 nodes) with LR = 0,001	Linear	64	251	4769	0,00061
Model 2 Hidden Layer (3 and 2 nodes) with LR = 0,001	ReLU	64	274	5206	0,00061

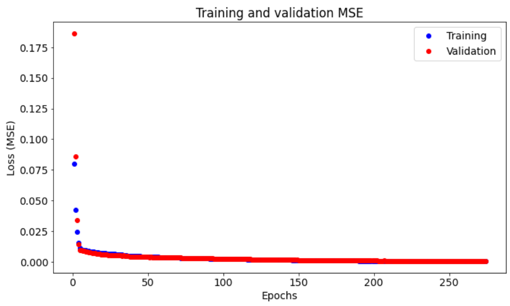
Based on the models in Table 3, the structure and the best model are based on the lowest loss value of 0.00051, namely the 1 hidden layer model with 4 nodes and a learning rate of 0.01. Determination of the best model can also be done by doing overfitting and underfitting checks to determine the best model based on the loss and validation loss graphs in Figure 7.



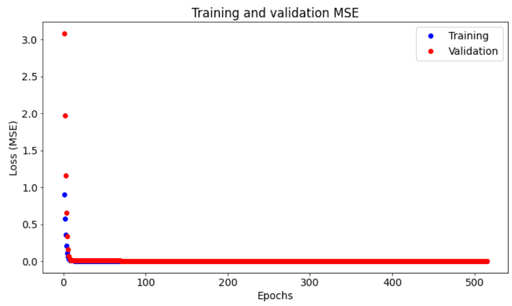
Model 1 Hidden Layer with LR = 0,01



Model 2 Hidden Layer with LR = 0,01



Model 1 Hidden Layer with LR = 0,001



Model 2 Hidden Layer with LR = 0,001

Figure 7. Loss Graph and Validation Loss for Each Model

Based on the graphs of the four models, the best good fit model graph is obtained, namely the model with a hidden layer structure with four nodes and a learning rate of 0.01 and a Linear activation function. That is because the graph of a hidden layer model with a learning rate = 0.01 tends to be stable and decreases to a faster stability point, which coincides with epoch 15 onwards.

After getting the best good fit model, the next step is to predict the gold price with data testing using the best model structure that has been built. The results of gold price predictions using a hidden layer model structure and learning rate = 0.01 with a Linear activation function are visualized using a time series graph shown in Figure 8.

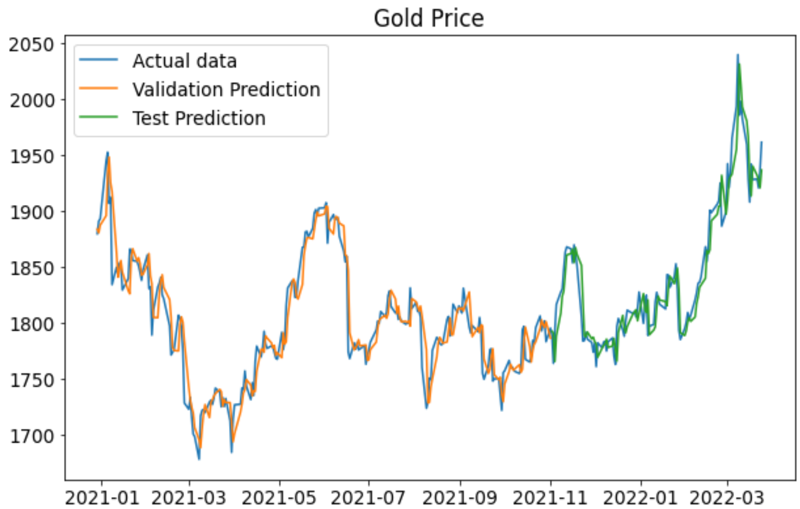


Figure 8. Visualization of Gold Price Prediction Results Using the Best Structure

Figure 8 shows that the blue line represents the actual data on the gold price, while the orange and green lines show the plot of the predicted gold price. The graph shows that the prediction line with the solid line coincides, so it can be said that the model is quite good.

After visualizing the results of gold price predictions, the next step is to conduct model testing to determine the performance of the best model built by looking at the accuracy and error in the model. Model testing indicators by looking for the mean absolute error (MAPE), mean squared error (MSE), and accuracy. Testing the ANN model is presented in Table 4.

Table 4. Model Testing with M.A.P.E., M.S.E., and Accuracy

Model	MSE	MAPE (%)	Akurasi (%)
Model 1 Hidden Layer with LR = 0.01	0,00051	1,97989	98,02011
Model 2 Hidden Layer with LR = 0.01	0,00125	1,99351	98,00649
Model 1 Hidden Layer with LR = 0.001	0,00061	2.06090	97,98147
Model 2 Hidden Layer with LR = 0.001	0,00061	2.01639	97,95553

Based on Table 4, the model test resulted in the best structure: one hidden layer model with four nodes and a learning rate of 0.01 using the Linear activation function. One hidden layer model with a learning rate = 0.01 produces an MSE value of 0.00051 and a MAPE value of 1.97989%, and the highest accuracy is 98.02011%. According to Lewis, according to the criteria for the MAPE value, the predictive model ability is outstanding if the value is < 10%. So, it can be said that the model that has been built is excellent.

4. CONCLUSION

The best structure of the ANN model with a backpropagation algorithm based on four activation functions in predicting the price of gold is to use four nodes in the input layer, one hidden layer with four nodes, an output layer, a learning rate of 0.01, 64 batch size, 274 epochs, 5206 iterations and activation function. Linear.

The prediction of gold prices based on the ANN model using the best network structure obtains quite good results because the prediction results with actual data show a slight difference in price differences.

Testing the ANN model using the best network structure resulted in an MSE value of 0.00051 and a MAPE value of 1.97989% with an accuracy of 98.02011%. The model built with the best structure also does not experience overfitting or underfitting and can be said to be a good fit model. Therefore, the performance of the ANN model is excellent and optimal.

DAFTAR PUSTAKA

[1] S. Patalay and M. R. Bandlamudi, "Gold Price Prediction Using Machine Learning Model Trees," in *International Conference on Changing Business Paradigm(ICCBP)*, Murshidabad, 2021.

[2] S. Makridakis, S. C. Wheelwright and V. E. McGee, Metode dan Aplikasi Peramalan, Edisi ke-2, Jakarta: Erlangga, 1999.

[3] L. V. Fausett, Fundamentals of Neural Networks: Architectures, Algorithms, and Applications, New Jersey: Prentice-Hall, 1994.

[4] S. P. Siregar and A. Wanto, "Analysis of Artificial Neural Network Accuracy Using Backpropagation Algorithm In Predicting Process (Forecasting)," *International Journal Of Information System & Technology,* vol. 1, no. 1, pp. 34-42, 2017.

[5] Y. A. Lesnussa, C. G. Mustamu, F. K. Lembang and M. W. Talakua, "Application Of Backpropagation Neural Networks In Predicting Rainfall Data In Ambon City," *International Journal Of Artificial Intelegence Research,* vol. 2, no. 2, 2018.

[6] D. Puspitanigrum, Pengantar Jaringan Syaraf Tiruan, Yogyakarta: Andi, 2006.

[7] K. T. N. Lestari, M. A. Albar and R. Afwani, "Application of the Backpropagation Method in Predicting the Number of Tourist Visits to Nusa Tenggara Barat Province," *Journal of Computer Science and Informatics Engineering (J-Cosine),* vol. 3, no. 1, pp. 39-48, 2019.

[8] A. B. Bayu, "Prediksi Harga Saham Dengan Menggunakan Jaringan Syaraf Tiruan," *Jurnal Teknologi Informatika dan Komputer MH Thamrin,* vol. 6, no. 2, pp. 103-111, 2020.

[9] N. Nafi'iyah, "Perbandingan Regresi Linear, Backpropagation, dan Fuzzy Mamdani dalam Prediksi Harga Emas," in *Seminar Nasional Inovasi dan Aplikasi Teknologi Industri (SENIATI)*, Malang, 2016.

[10] J. Han, M. Kamber and J. Pei, Data Mining: Concept and Techniques, 3rd Edition, Waltham: Morgan Kaufmann Publishers, 2012.

[11] E. Turban, J. E. Aronson and T.-P. Liang, Decision Support System and Intelligent Systems, 7th Edition, Prentice Hall, 2004.

[12] B. Santosa, Data Mining: Teknik Pemanfaatan Data untuk Keperluan Bisnis, Graha Ilmu, 2007.

[13] M. Ayub, "Proses Data Mining dalam Sistem Pembelajaran Berbantuan

Komputer," *Jurnal Sistem Informasi,* vol. 2, no. 1, pp. 21-30, 2007.

[14] T. Masayuki and M. Okutomi, "A Novel Inference of a Restricted Boltzmann Machine," in *2014 22nd International Conference on Pattern Recognition (ICPR)*, Tokyo, 2014.

[15] J. Heaton, Introduction to Neural Networks with Java, 2nd Edition, Heaton Research, Inc., 2008.

[16] Y. Lu, "Deep Neural Networks and Fraud Detection," U.U.D.M. Project Report 2017, Sweden, 2017.

[17] D. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," in *Proceeding Conference Paper at the 3rd International Conference for Learning Representations*, San Diego, 2015.

[18] S. Sena, "Pengenalan Deep Learning Part 1: Neural Network," 19 Maret 2018.

[19] A. D. Erna, Pengantar Jaringan Syaraf Tiruan, Wonosobo: Star Publishing, 2009.

[20] F. Pakaja, A. Naba and P. Purwanto, "Peramalan Penjualan Mobil Menggunakan Jaringan Syaraf Tiruan dan Certainty Factor," *Jurnal EECCIS,* vol. 6, no. 1, pp. 23-28, 2012.