Technical Data Analysis for Movement Prediction of Euro to USD Using Genetic Algorithm-Neural Network

Ary Sespajayadi[1,a], Indrabayu[2,b] and Ingrid Nurtanio[3,c]
Artificial Intelligent and Multimedia Processing Research Group
[1]Graduate of Informatics Engineering Study Programme, Hasanuddin University
[2]Informatics Engineering Study Program, Electrical Engineering Department, Hasanuddin University
[3]Electrical Engineering Study Program, Electrical Engineering Department, Hasanuddin University
Makassar, Indonesia 90245
[a]ary.sespajayadi.id@ieee.org, [b]indrabayu@unhas.ac.id, [c]ingrid@unhas.ac.id

***Abstract*— In the foreign currency exchange (FOREX), a technical data analysis system for predicting currency movements is needed to help traders in decision making. Thus, this study proposes a system of technical data analysis to movement prediction of Euro to USD using Genetic Algorithm-Neural Network (GANN). To generate a predicted value, Genetic Algorithm searching for the best value of Feed Forward Neural Network (FFNN) trained with the Neural Network method that produced a net to predict. The Validation of predicted results with GANN method based on the degree of accuracy as follows. RMSE values of open is 0.00043; The RMSE values of high is 0.00068; The RMSE value of low is 0.00075; and RMSE values of close is 0.00070.**

Keywords—genetic algorithm; neural network; gann; prediction; forex.

I. INTRODUCTION

In Artificial Intelligence technological developments there are many things can be done to predict certain issues. One of the most interesting subject to predict is foreign exchange market. In foreign currency trading, currency trade one type country currency towards another country e.g Euro to USD.

Forex trading has a very high risk, if a trader mistaken in doing the trade steps; the trader will fall into a loss. Thus a trader should have a basic methods through the forex trading process. The term *open*, *high*, *low* and *close* in currency movements, where *open* value is the first transaction price conducted on the day. *Open* value is all market information occur between closing price of previous day and the last moment investors allowed to put orders into the stock engine, while the *high* price and *low* price is the daily price range movement of stock where investors have the courage or rationality to do buy position or sell position. If there is a happening information that the stock price could go up with buying so that the price closed in the up position, as well as vice-versa. And *close* value is the most important price in performing technical analysis. Close price is the whole existing information on all market participants when currency trading ended. More than 90% technical indicators used by analysis technician, using the *close* value as its main input. This causes close price position, could trigger a signal to buy or sell signal.

Performed research about Using Recurrent Neural Networks To Forecasting of Forex [1], Towards the Automatic Evolutionary Prediction of the Forex Market Behaviour [2], Forex Trading Prediction Using Linear Regression Line, Artificial Neural Network and Dynamic Time Warping Algorithms [3].

In this study, technical data analysis to foreign currency movement prediction by GANN whereas Genetic algorithm is an algorithm applied to initial parameters for optimization [4] as well as trying to find desired solution through previously obtained solution improvements, then neural network is an algorithm that can be used in modeling complex relationships between inputs and outputs to find patterns in data. By blending Genetic Algorithm and Neural Network it is expected to generate a prediction of foreign currency movement to reduce risk in forex trading.

II. GENETIC ALGORITHM-NEURAL NETWORK

A. Genetic Algorithm

Genetic Algorithms are search algorithms based on natural selection mechanism and natural genetics, along with algorithms that seek to apply an understanding of natural evolution in problem solving tasks. Attempted approach by this algorithm is randomly combine wide selection of solutions in a population then evaluate it in getting the best solution [5].

By doing this process is repeated, this algorithm simulates the process of evolution as the desired number of generations. This generation will represent improvements on previous population. In the end, we will get the solutions most appropriate to the problems faced. To use genetic algorithms, problem solutions are represented as a set of genes that make up chromosomes. This chromosome is formed by random based coding techniques used. The entire set of chromosomes observed represent a population.

Chromosomes will evolve in several stages iterations called generations. The new generation is obtained by crossover and mutation. Crossover includes cutting two pieces

of chromosomes based on the desired number of points then combine half of each chromosome with other couples. While mutations include the replacement value of a gene on a chromosome with the value of other genes from other chromosomes that become his partner. These chromosomes further evolved with a suitability criteria (fitness) are defined and the best results will be selected while others are ignored. Furthermore, the process is done repeatedly until you have a chromosome that has the best fit (best fitness) to be taken as the best solution of the problem.

B. *Neural network*

Artificial neural networks (ANN) is a computer science field with a different method that tries to solve problems in the real world by offering a powerful solution. ANN has the ability to learn and produce their own knowledge from their environment [6]. ANN can be used to model complex relationships between inputs and outputs in finding data patterns. ANN which also known as Backpropagation network training is to get a balance between network's ability to recognize patterns used during training as well as the network's ability to provide the correct response towards similar input pattern but different with the pattern used for training.

Backpropagation network training includes three stages that is feedforward input pattern, calculation and backpropagation of errors, and weights adjustment. On feed-forward phase, input signal (x_i) are propagated to hidden layer. Weighted input signals that entered each unit in the hidden layer (z_j, $_j$ = 1,2, .., p), is determined by the equation:

$$z\_in_j = \delta_{1j} + \sum_{i=1}^{n} \quad x_i v_{ij} \tag{1}$$

Output signal from hidden layer (z_j) is determined by activation function used (sigmoid function) with the equation:

$$z_j = 1/(1 + e - (z\_inj)) \tag{2}$$

Furthermore, z_j is propagated forward again to output layer by equation:

$$y\_in = \delta_2 + \sum_{j=1}^{p} \quad z_j w_j \tag{3}$$

Output signal is determined by the activation function used in the output layer that is identity function with the equation:

$$y = f(y\_in) = y\_in = \delta_2 + \sum_{j=1}^{p} \quad z_j w_{j1} \tag{4}$$

Next, network output is compared with target that must be achieved. Network output target gap with the network output is occurred error, backpropagation error calculated by the equation:

$$MSE(\theta) = 1/N \sum_{j=1}^{n} \quad (\theta_j - \theta)^2 \tag{5}$$

Where n is data amount, Θ_j is the j-target, while Θ is the network *output*. If this error is smaller than specified tolerance limits, then epoch stopped. However, if error is greater than tolerance limit, then weight of each line in network will be modified to reduce occurred errors.

Occurred errors propagated backwards, starting from line directly related to units in the output layer. Weights and biases modification to reduce occurred errors is done by changes in weight and bias on line that leading to output unit:

$$w_j\,(new) = w_j\,(old) + \Delta w_j \tag{6}$$

$$\delta_2\,(new) = \delta_2\,(old) + \Delta \delta_2 \tag{7}$$

Changes in weight and bias on the line towards the hidden units:

$$v_{ij}\,(new) = v_{ij}\,(old) + \Delta v_{ij} \tag{8}$$

$$\delta_{1j}\,(new) = \delta_{1j}\,(old) + \Delta \delta_{1j} \tag{9}$$

These three stage is continuously repeated until termination condition is met. Generally, termination condition often used are number of epochs or errors. Epoch will be terminated if the number of epochs done already exceeded defined maximum number of epochs, or if the occurred error is smaller than desired tolerance limits.

III. Proposed Scheme

Steps taken in making predictions of Euro currency to USD using GANN, starting from data acquisition, train process and prediction test, until steps taken in completing this method. Performed prediction system is M30 history data prediction (data price changes every 30 minutes), which is prediction process for the next 30 minutes, by taking price of 30 minutes earlier.

A. *Input*

Data acquisition, whereas data obtained from data history of *Metatrader* application. Therefore, obtained data has five variables, *open*, *high*, *low*, *close* and *volume*. But not all these variables are used. Variables used as inputs are *open*, *high*, *low* and *close*, while output data also *open*, *high*, *low* and *close*.

Following are prediction scheme process of foreign currency exchange rates GANN Method.

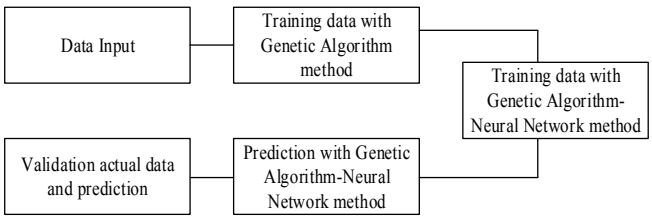

Fig. 1. Prediction scheme process of foreign currency exchange rates GANN Method.

B. System Design

This stage is divided into two processe, that is Training and Prediction.

- Training
  In this case, the training data used are data from 2013 through 2014, month 12 date 19 time 09:00. In the training phase, training data input are in input variables type such as *open*, *high*, *low* and *close.*
- Prediction
  The next stage is finding the prediction value after training process. Prediction with GANN method are using Matlab program. The output of this process will be validated with actual data.

C. Validation

To validate prediction results, comparing with actual data of *MetaTrader*. Validation is an important part to evaluate the prediction performance. In order to get the results of the predictive test each value of *open, high, low* and *close*, the results validated by the respective actual data on the same date and time, so that the accuracy obtained for each outcome prediction test the *open, high, low* and *close* 30 minutes ahead. Thus, to determine the accuracy of the prediction performance, the importance of knowing the *Root Mean Square Error* (RMSE). The smaller the RMSE it will get better prediction accuracy, the following formula for the validation of the accuracy of prediction:

$$RMSE = \frac{\sqrt{\frac{1}{N}\sum_{t=h}^{N}(y_t - \hat{y}_t)^2}}{y\ \max - y\ \min} \quad (10)$$

D. Output

The output of this system is prediction result and average error. Steps in training process GANN are as follows.

- Load data from each input variable from year 2013 to 2014, where x training data (input data) is data from year 2013 until December 2014 and y data (target data) is Forex 2013 to December 2014 data.
- Finding weights and biases value of Feed Forward Neural Network (FFNN) with number of neurons in the hidden layer is used together with the number of neurons in the input layer as many as four and one neuron in the output layer, so required variable amount (weights and biases) is total (number of inputs + 1) where in this study the number of inputs are four parameters, namely: *open*, *high*, *low* and *close* to the number of variables in a chromosome is as much as 25 pieces of genes, that is: sixteen pieces of weights and four biases from input layer to hidden layer and four weights and a bias from hidden layer to the output layer.
- Of formation FFNN network with Genetic Algorithm, combined with Neural Network-Backpropagation, to be retrained so that the combined algorithm between GANN can generate predictions train foreign currency with a small error and produce a net, then the results of a trained net can be used to predict the next 30 minutes.

The data already trained previously to generate predictive value forex on December 29, 2014. The steps in process of Forex prediction methods GANN are as follows.

a. Data training parameter containing the *open*, *high*, *low,* and *close*.
b. Training process of GANN used for prediction.

Calculate output of prediction date on December 29, 2014 to acquire accuracy value of the system.

IV. Results and Discussion

Prediction test results for each variable that is, the value of the price of the *open*, *high*, *low* and *close* for 30 minutes ahead can be seen in the following graph:

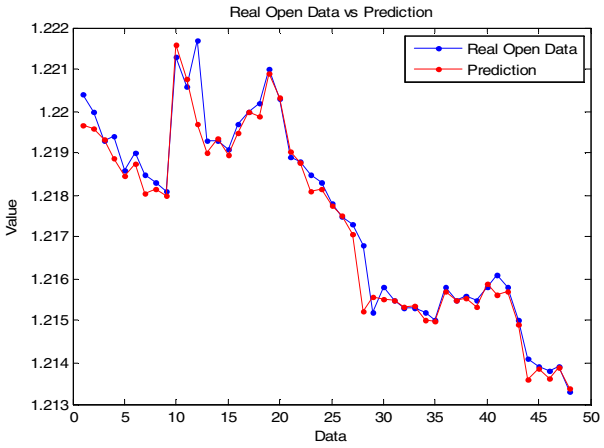

Fig. 2. Comparison of *real open* data with test results prediction *open* foreign exchange currency GANN Method.

With the *open*-variable predictive test, where the data amount as 48 with RMSE 0.00043.

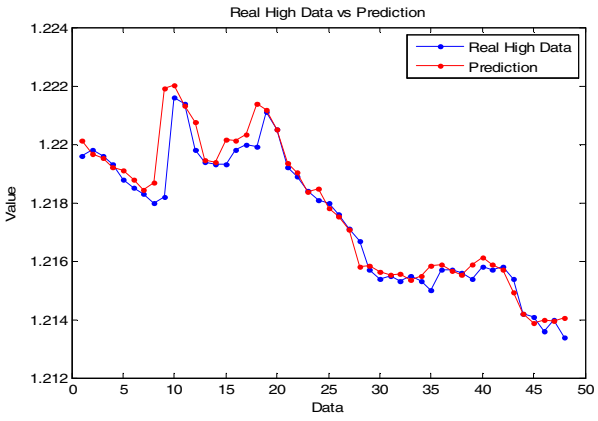

Fig. 3. Comparison of *real high* data with *high* value test result prediction of foreign currency exchange rates with GANN Method.

Test predictions with *high* variable, where the data amount as 48 with with RMSE 0.00068.

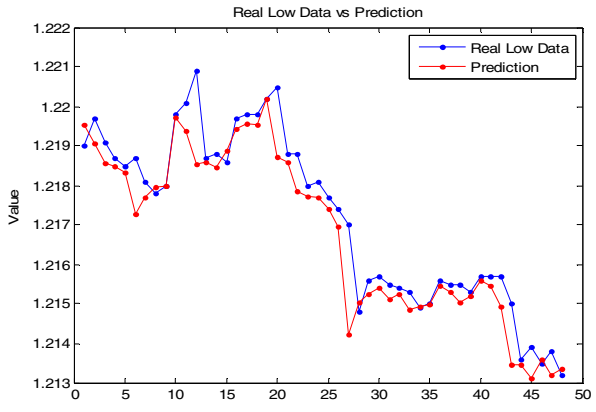

Fig. 4. Comparison of *real low* data with *low* value test result prediction of foreign currency exchange rates with GANN Method.

By test predictions with *low* variable, where the data amount as 48 with RMSE 0.00075

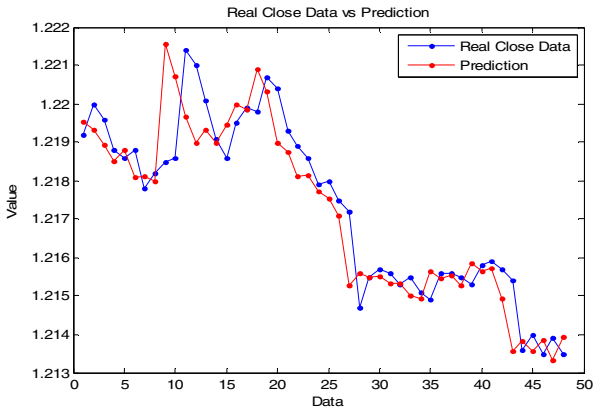

Fig. 5. Comparison of *real close* data with *close* value test result prediction of foreign currency exchange rates with GANN Method.

With *close* variable prediction test, where the data amount as 48 with RMSE 0.00070.

V. CONCLUSION

From performed research in prediction of foreign exchange currency rates *open*, *high*, *low* and *close* using GANN method (Euro to USD). can be concluded that, test prediction results validation of Euro to USD on the value of *open*, *high*, *low* and *close* using GANN method against the actual data M30 from four variables each have RMSE value and accuracy level as follows, *open* with RMSE value of 0.00044, *high* RMSE value of 0.00068, *low* has a RMSE value of 0.00075 and *close* RMSE value of 0.00070.

From the analysis of predicted results and the results of RMSE, the prediction accuracy value of the *open* always high due to predictive value of the *open* based on the value of the previous *close*, or in other words the value of the *open* in the next 30 minutes always see the value of *close* on 30 minutes earlier. With the technical data predicted results, the next study will be conducted research with fundamental data.

REFERENCES

[1] V.V.Kondratenko and Yu. A Kuperin. "Using Recurrent Neural Networks To Forecasting of Forex". Rusia. 2003.

[2] Karel Slaný. "Towards the Automatic Evolutionary Prediction of the Forex Market Behaviour". Czech Republic. 2009

[3] Leslie C.O. Tiong, David C.L. Ngo, and Yunli Lee. "Forex Trading Prediction Using Linear Regression Line, Artificial Neural Network and Dynamic Time Warping Algorithms". Proceedings of the 4th International Conference on Computing and Informatics, ICOCI 2013 28-30 August, 2013 Sarawak, Malaysia. Universiti Utara Malaysia (http://www.uum.edu.my) Paper No. 092

[4] Lanz CWJ Chan and Wing-Keung Wong, "Automated Trading with Genetic-Algorithm Neural-Network Risk Cybernetics: An Application on FX Markets". 2012.

[5] Gopesh Joshi, "Review of Genetic Algorithm: An Optimization Technique", International Journal of Advanced Research in Computer Science and Software Engineering Research Paper Available online at: www.ijarcsse.com ISSN: 2277 128X, Volume 4, Issue 4, April 2014.

[6] Alvarez, E. G. (2006). "Artificial neural networks", 1–48.