International Journal
of Applied Information Technology
http://journals.telkomuniversity.ac.id/ijait/

Artificial Neural Network Model with PSO as a Learning Method to Predict Movement of the Rupiah Exchange Rate against the US Dollar

Eko Verianto [a, *], Budi Sutedjo Dharma Oetomo [b]

[a,b] *Faculty of Information Technology, Duta Wacana Christian University, Yogyakarta*
eko.verianto@mail.ugm.ac.id, budi@staff.ukdw.ac.id

ARTICLE INFO

Received October 31st, 2020
Revised February 19th, 2021
Accepted February 23rd, 2021
Available online April 19th, 2021

Keywords
Artificial Neural Network, Particle Swarm Optimization, Prediction

ABSTRACT

The movement of currency exchange rate can be predicted in the next few days, this is used by economic actors to get profit. Artificial Neural Network with the backpropagation learning method is good enough to use for forecasting time series data, it's just that in its application this method was considered to have shortcomings such as a long training time to achieve convergence. The purpose of this research is to form a Multilayer Perceptron Artificial Neural Network model with the Particle Swarm Optimization (PSO) algorithm as a learning method in the case of currency exchange rate prediction. This research produced a model that can predict the movement of the Rupiah exchange rate against the US Dollar, while the model formed was the MLP-PSO model with an error rate of 5.6168×10^{-8}, slightly better than the MLP-BP model with an error rate of 6.4683×10^{-8}. These results indicated that the PSO algorithm can be used as a learning algorithm in the Multilayer Perceptron Artificial Neural Network.

* Corresponding author at:
Faculty of Information Technology, Duta Wacana Christian University
Jl. Dr. Wahidin Sudirohusodo No. 5-25, Kotabaru, Gondokusuman, Kota Yogyakarta, D. I. Yogyakarta, 55224
Indonesia
E-mail address: eko.verianto@mail.ugm.ac.id

ORCID ID:
First Author: 0000-0002-1769-5983

https://doi.org/10.25124/ijait.v4i02.3381
Paper_reg_number IJAIT000040209 2020

1. Introduction

At present, the demand for globalization makes a country able to improve itself in maintaining macroeconomic stability. The currency exchange rate is one of the indicators in macroeconomic that need attention because the currency exchange rate is an agreement on current and future payments between two countries.

Many factors can affect a country's currency exchange rate, including changes in taste, prices for imported and exported goods, inflation rates, changes in interest rates, and rates of return on investment, and the country's economic growth [1].

The increasing volume of international trade will increase the use of the foreign exchange, so it is not surprising if the country's currency exchange rate is determined in trading on the foreign exchange market [2]. This futures trading is an investment in the future that invites huge profit potential for business people. On the other hand, trading like this happens continuously for twenty-four hours. In the foreign exchange market, the exchange rate of each country often fluctuates.

Two types of analysis can be used to analyze the foreign exchange market, namely Fundamental Analysis and Technical Analysis [3]. Fundamental analysis is used to observe currency exchange rate movement by considering various factors that influence changes in the currency exchange rate, while Technical Analysis is used to observe exchange rate movement based on the historical value of the previous exchange rate data.

Among the two techniques, the one that is quite widely used by economic actors to observe the historical movement of the currency exchange rate is Technical Analysis. Through this analysis, the movement of the currency exchange rate can be predicted in the next few days.

Meanwhile, several methods that are used to predict currency exchange rates include ARIMA, Exponential Smoothing, and Nonlinear Model.

For the closing data, the currency exchange rate has two main attributes, namely the time and the exchange rate itself, so that the data, in this case, is classified into time-series data. The BP algorithm is one of the supervised learning methods that is usually used for modeling time series data. The algorithm performs backpropagation to change the network weights that are connected to each iteration. The algorithm is good enough to be used for modeling nonlinear data, but this method takes too long to achieve convergence [4].

This research will use Particle Swarm Optimization as a learning algorithm for ANN. Particle Swarm Optimization is a stochastic and parallel optimization algorithm. Moreover, this algorithm is simple and has fast convergence [5].

Particle Swarm Optimization is widely used as a method for ANN weights optimization, such as research with the topic of backpropagation neural network optimization with particle swarm optimization to predict tide level. This research uses Particle Swarm Optimization to optimize the minimum error value on ANN [6]. Another similar research with the topic of NN-PSO modeling to predict thermal properties, this study also uses Particle Swarm Optimization to optimize tissue weight using trial and error techniques [7]. Likewise, other research on a similar topic, Particle Swarm Optimization in this research is used to optimize the first test only; for further testing, it will be optimized with backpropagation [8].

Based on the sentences above, the focus is on ANN modeling with PSO as a learning method for predicting the movement of the Rupiah exchange rate against the US Dollar.

2. Literature Review

2.1. Prediction of Currency Exchange Rate

Prediction is a process of systematically estimating or forecasting something that is most likely to happen in the future based on the past and present information that is owned [9]. There are two methods to predict the currency exchange rate, namely qualitative and quantitative prediction methods. Qualitative prediction is carried out based on knowledge, while quantitative prediction is carried out based on past quantitative data.

Meanwhile, the currency exchange rate is the value or price of a country's currency that is compared to the price of another country's currency [10]. The exchange rate always changes every time (fluctuation) depending on the economic and political conditions of a country. Therefore, it is necessary to predict the exchange rate so that company managers can make a decision that is closer to the actual exchange rate.

2.2. Artificial Neural Network

An Artificial Neural Network (ANN) (Figure 1) is an architecture system, and its operation is inspired by knowledge of biological nerve cells in the brain which can be described as a mathematical and computational model for non-Linear Approximation function, cluster data classification, and non-Parametric Regression or simulation from a collection of Biological Neural Network model [11]. In short, ANN is a representation of the human brain [12]. ANN has an architecture consisting of several layers, neurons, and also a connector, the connector on each connector has its own weight.

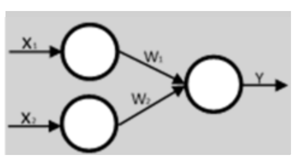


Figure 1 Artificial Neural Network Architecture

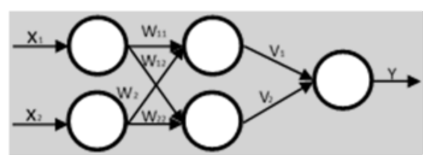


Figure 2 MLP ANN Architecture

The neurons that are contained in each layer have an activation function for each incoming value, then this activation function will determine the output value. Several kinds of activation functions is available on ANN: Identity Function, Binary Function with Threshold θ, Binary Sigmoid Function, Bipolar Sigmoid Function, and ReLU Function.

2.3. Multilayer Perceptron

Multi Layer Perceptron (MLP) (Figure 2) is a Feedforward ANN that has one or more hidden layers and a number of neurons that are connected to each other [13]. Neurons in the multilayer perceptron artificial neural network are arranged in the input layer, hidden layer, and output layer. The working of neurons in the input layer will receive the signal from outside, then they are forwarded to one or more hidden layers until they finally reach the output layer and the final result is found based on the learning that has been done.

2.4. Backpropagation Algorithm

Learning is the most important stage in building a model, through this learning process, a model can do its job. The BP algorithm is a supervised learning algorithm that is used to change the weights of ANN. The working of this algorithm is by changing the weights that are on the ANN, based on the error that is produced at the forward propagation stage.

The BP algorithm is used alongside the Gradient Descent Algorithm which aims to minimize the value of a function.

There are two Gradient Descent Algorithms that are often used in building models using Backpropagation ANN, namely Stochastic Gradient Descent (SGD) and Adam. SGD works by changing value after evaluating one or several pairs of training data, while Adam is a combination of the advantages of AdaGrad and RMSProp, so Adam adapts the learning rate parameter based on the average of the first moment and also takes advantage of the second moment of the gradient.

2.5. Particle Swarm Optimization

Particle Swarm Optimization is a Stochastic optimization technique based on Swarm. The PSO algorithm simulates the social behavior of animals, including insects, birds, and fish. This herd forms a cooperative way to find food, and each member will continue to change the search pattern according to the learning experiences of their own members and other members [5].

The PSO is assumed to be like a herd of a certain size that is placed in a random multidimensional space. Each particle will then move in the space to find the right position, that position will be informed to other particles and will adjust the position and speed of each particle to reach its destination.

In the PSO Algorithm, changes in velocity are influenced by global best and local best solutions. The global best solution is related to the lowest cost ever obtained by particles, while the local best is related to the lowest cost in the initial population.

2.6. Sliding Window

The sliding window is a technique that is used for forecasting time series data. The working of this technique is by paying attention to the previous data, for example, if the size of the sliding window is 3, then what is used as data for the independent variable is the previous 3 data, while the 4th data will act as data for the dependent variable.

2.7. Data Transformation

Data transformation is the act of changing the measurement scale of the data into another form, usually lower than the original data. This is done for the purposes of analysis. In addition, in this research, currency exchange rate data has a different range, so it is necessary to transform data.

Many methods can be used to perform data transformation, one of which is vector normalization. Here is the vector normalization equation.

$$r_{ij} = \frac{x_{ij}}{\sum_{i=1}^{m} x_{ij}^2} \qquad \text{Equation 1 [14]}$$

Where r_{ij} is the result of data transformation, x_{ij} is the data in row i and column j matrices, while $\sum_{i=1}^{m} x_{ij}^2$ is the sum of each data in row i and column j matrices that are squared.

2.8. Data Splitting

Data splitting is one of the most important parts of building a model. The right data splitting will produce a good model too. Two data splitting methods are often used, namely Holdout and K-Fold Cross-Validation. The Holdout method divides the data by separating the data into two parts, namely the training dataset and test dataset, while the K-Fold Cross Validation divides the data into k groups, where each group has the same size, then the first k group will be used as data for training. Meanwhile, the other k groups will be used as data for model evaluation.

2.9. Evaluation Method

The predicted value may not always have the same value as the original data. The difference between the predicted value and the value in the original data can be used as a reference to determine which model has the smallest error rate which will then be used to predict the test data. There are many methods that can be used to calculate the difference between the predicted data and the original data.

This research used Mean Absolute Error (MAE) to calculate the difference between the predicted value and the original data. The following is an equation of the mean absolute error model evaluation method.

$$MAE = \frac{1}{n}\sum_{j=1}^{n} |y_j - y_j'| \qquad \text{Equation 2 [15]}$$

MAE is an absolute value from the prediction result of the model and the original data which is then divided by the amount of data available.

3. Methodology

There are several steps taken in this research to produce a model with a low error rate. The methodology that was used in this research will be discussed in this section.

3.1. Research Stages

The flowchart in Figure 3 shows the data sample to be used. Data samples will be divided into three parts, namely training data, validation data, and testing. In the learning process, the data will be normalized. Next is the data training process carried out using multilayer perceptron and particle swarm optimization. The training process uses data training, and validation using data validation. This process will continue to get the best model, and then this model is to make predictions on the data testing. Research stages are as follows.

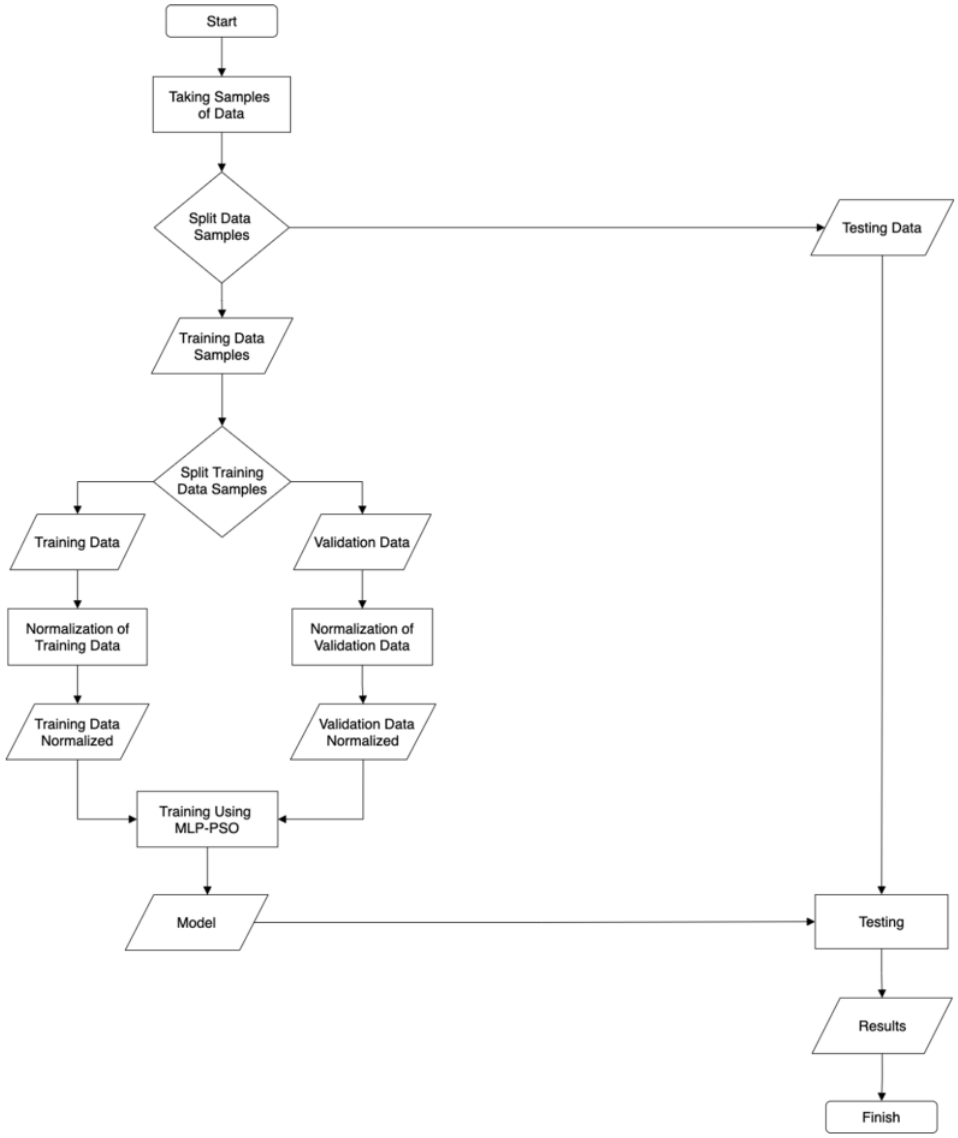


Figure 3 Research Process Flowchart

3.2. Data Collection

The data that was used in this research was historical data on currency exchange rate closing that was obtained through the website http://fx.sauder.ubc.ca/data.html which contained historical data on currency exchange rate closing from several countries, including countries in the Asia Pacific. In this research, data on the closing Rupiah exchange rate against the US Dollar was used.

3.3. Data Splitting

Data splitting uses the Holdout method and data was divided into three parts, namely training dataset, validation dataset, and test dataset. This data splitting was chosen to avoid repeated test dataset and to minimize model overfitting [16].

This research used a training dataset for the learning process to obtain the best model. Meanwhile, a validation dataset was used to test models that have previously been formed through the training process, while a test dataset is only used when the best model has been found and cannot be used repeatedly.

3.4. Data Transformation

Data transformation was carried out on data samples, training data, validation data, and test data. Data transformation was carried out before learning was carried out so that the data used for the learning process until the testing was data that has been transformed. Data transformation uses Vector Normalization because Vector Normalization combined with ANN (Artificial Neural Network) has good results in time series data predicting [14].

3.5. Multilayer Perceptron Design

The architecture of Multi Layer Perceptron (MLP) ANN in this research consists of an input layer, one hidden layer, and an output layer. Based on the trial and error process, the number of neurons in the input layer, hidden layer, and output layer was 5, 20, and 1 respectively. The number of neurons in the input layer shows the number of neurons in the output layer, indicating that the output result of the model is the prediction result of the model.

The weight that was used in the MLP ANN in this research was a vector solution when generating the solution. The dimension of the vector for each solution that produced was the same as the weights contained in the Multi Layer Perceptron ANN architecture.

3.6. Learning Design

The learning stage is a stage for training data that has previously been shared. The purpose of learning is to get a good model, a good model in the case of regression is a model with a low error rate. In this research, the PSO algorithm was used as stated in the flowchart in Figure 4.

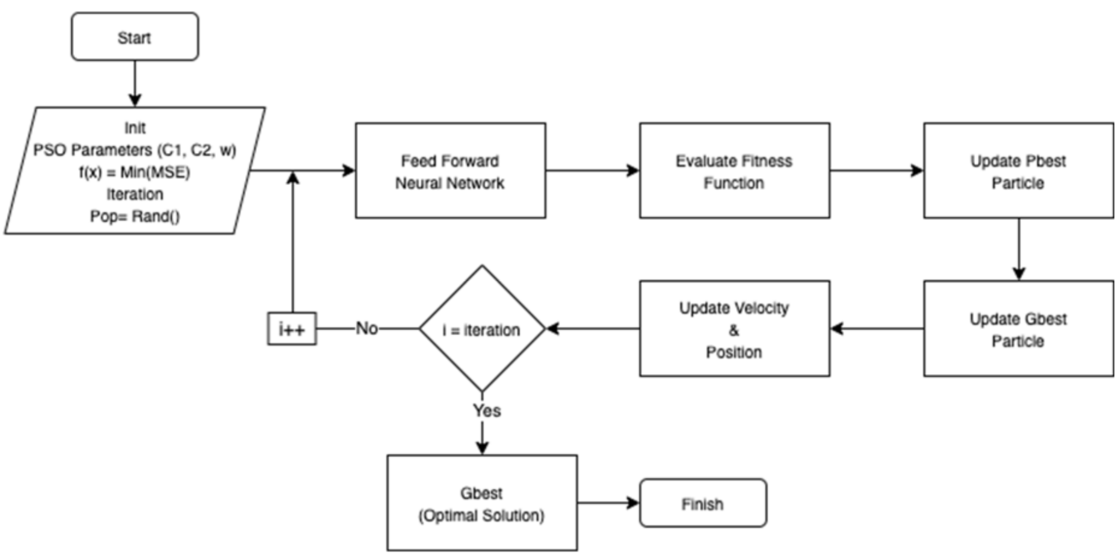


Figure 4 Learning Design Flowchart

3.7. Testing Design

The testing design is the stage in a series of testing activities that were carried out in this research. The first series of the stage was to determine the set of iteration, then to test the MLP model with PSO. The last part of the testing phase was convergence analysis, overfitting analysis, training time analysis, and best model analysis.

4. Results and Discussion

4.1. Iteration Testing

The first test that was carried out in this research was iteration testing. Iteration testing was carried out by changing the iteration value during training. This test was carried out to see the relationship between the number of iteration and the error rate that was produced by the model. Table 1 below shows the MLP hyperparameter used in this test.

Table 1 Hyperparameter Multilayer Perceptron

Input Layer	Hidden Layer	Output Layer	Bias
5 Neurons	20 Neurons	1 Neuron	2 Neurons

Table 2 Hyperparameter Particle Swarm Optimization

C1	C2	W	Dimension
0.5	0.3	0.9	141

Table 1 shows the MLP Hyperparameter that was used in all experiments in this research. Based on this table, the number of dimensions of each population that will be raised can be calculated later. In the PSO Algorithm, there were several Hyperparameters that will be used for the calculation process.

The PSO Hyperparameter in Table 2 is the hyperparameter that will be used in all experiments in this research so that the hyperparameter will not be changed later. In this research, BP also used as a comparison method.

As a comparison method in this research, the hyperparameter in the BP algorithm is also arranged, where the hyperparameter was also used in all experiments. The hyperparameter that changed in each test was iteration. The number of iteration in this test was varied.

Model testing was done using training data with a varying number of iterations. The result of this test was the error rate calculated using the Mean Squared Error (MSE). The test results are as shown in Table 3.

Table 3 shows the error rate in the MLP model with PSO optimization and the MLP model with BP as the learning algorithm. The difference in error rate resulting from the two models can be seen from the iteration variations given. The MLP-PSO model in the first experiment with iteration 50 produced an error rate of $1.0576 \text{ x } 10^{-6}$, then the iteration was increased to 100 in the second experiment with the same model produced in an error rate of $1.3357 \text{ x } 10^{-7}$, in this second experiment there was a decrease in the error rate. Likewise, with the third and fourth experiments on the same model with the addition of each iteration to 500 iterations and 1000 iterations, the error rate produced by the model has decreased gradually.

Table 3 MSE of Each Iteration

Method	Iteration	Error Rate
MLP-PSO	50	$1.0576 \text{ x } 10^{-6}$
	100	$1.3357 \text{ x } 10^{-7}$
	500	$1.0130 \text{ x } 10^{-7}$
	1000	$7.9096 \text{ x } 10^{-8}$
MLP-BP	50	$3.4694 \text{ x } 10^{-6}$
	100	$1.2795 \text{ x } 10^{-6}$
	500	$9.5095 \text{ x } 10^{-8}$
	1000	$1.3917 \text{ x } 10^{-7}$

Table 4 Comparison of Error Rate between Training Data and Validation Data

Model	Iteration	Training Data	Validation Data
MLP-PSO	50	1.0576×10^{-6}	1.0892×10^{-6}
	100	1.3357×10^{-7}	1.0780×10^{-7}
	500	1.0130×10^{-7}	1.0442×10^{-7}
	1000	7.9096×10^{-8}	6.7627×10^{-8}
MLP-BP	50	3.4694×10^{-6}	3.4694×10^{-6}
	100	1.2795×10^{-6}	6.7716×10^{-7}
	500	9.5095×10^{-8}	6.7215×10^{-8}
	1000	1.3917×10^{-7}	1.5346×10^{-7}

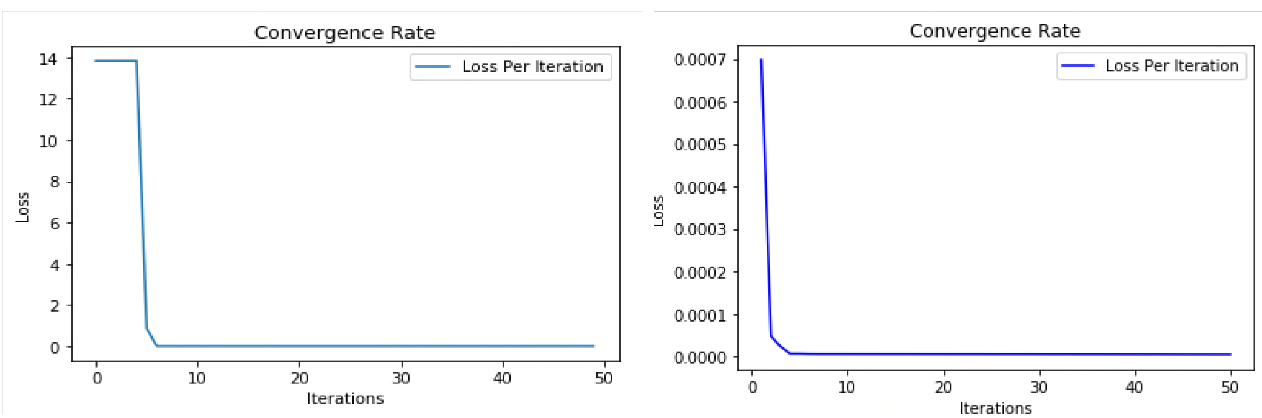
Unlike the previous model, the addition of the number of iterations in the MLP-BP model does not always reduce the error rate, such as giving 500 iterations in the model which results in an error rate of 9.5095×10^{-8}, but the error rate increases when there is an addition in a number of 1000 iterations.

4.2. Convergence Analysis

Convergence analysis is intended to see how fast the model reaches the minimum error, the faster the model reaches the minimum error, the better the model's performance.

The model that was used in the convergence analysis was the MLP-PSO model and the MLP-BP model with 50 iterations. This is intended to make it easier to see the convergence rate of the model.

Figure 5 is a graph of the convergence rate in the MLP-FPA model and the MLP-BP model, this graph shows that a drastic decrease in the error rate occurred between iteration 1 to iteration 10, the MLP-PSO model has a drastic decrease in the error rate in the 7th iteration with produced an error rate of 2.1199×10^{-5}, while the MLP-BP model also has a drastic decrease in the error rate in the 6th iteration with an error rate of 6.1583×10^{-6}.



(a) (b)

Figure 5 Convergence Rate Graph: MLP-PSO (a) and MLP-BP (b)

4.3. Overfitting Analysis

A model with a minimum error rate in training data does not necessarily perform well when given validation data and test data. This kind of case is known as overfitting. A good model should be able to provide good predictive results too so that the objectives of the model can be achieved. In this analysis, each model was used to see the model's performance in predicting validation data.

Based on Table 4, there were error rates for two different models and two different data. Each of the models was also differentiated based on the number of iterations, for more details in analyzing overfitting.

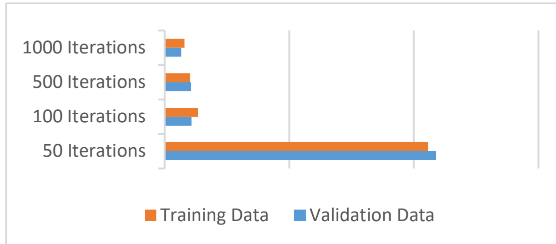


Figure 6 Comparison Graph of Error Rate for Training Data and Validation Data in the MLP-PSO Model

Figure 6 shows that there was an overfitting in the trained model with 50 iterations and 500 iterations, where the minimum error rate occurred in the training data, while in the validation data there was an increase in errors. This shows that the patterns produced in the generalization process can predict training data well, while the validation data cannot be predicted well.

4.4. Analysis of Training Time

Based on previous tests, in addition to the variation in the number of iterations, the selection of a learning algorithm will also greatly affected the training time as shown in Table 5 below.

Table 5 shows that a model that used the BP Algorithm has a faster training time than using the PSO Algorithm, This, of course, can occur because of the initial population generation in the PSO Algorithm.

Table 6 shows that population size greatly affected training time, where the larger the population size, the longer the training process will take.

Table 5 Comparison of Training Time in Each Model for

Model	Iteration	Training Time (s)
MLP-PSO	50	10.1411
	100	12.3130
	500	100. 9110
	1000	180.1336
MLP-BP	50	3.6251
	100	5.6565
	500	23.4074
	1000	55.4871

Table 6 Relationship of Training Time and Population Size in PSO Each Iteration Algorithm

Population Variation	Training Time (s)
20	4.0684
40	6.4080
60	8.6413
80	10.8200
100	12.9393

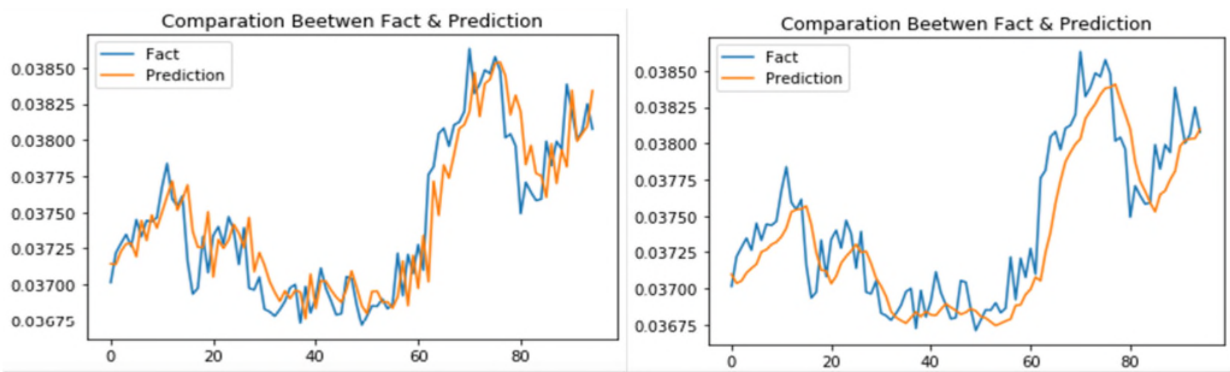
4.5. Best Model Analysis

The best model was carried out based on the error rate that was produced in the validation data, this was to avoid overfitting the model. Based on Table 7, the minimum error rate in the MLP-PSO model was in 1000 iterations, while the MLP-BP model was in 500 iterations. Furthermore, testing will be carried out on each of these best models using test data.

Table 7 Comparison of Error Rate MLP-PSO Model and MLP-BP Model

Model	Error Rate
MLP-PSO	5.6168×10^{-8}
MLP-BP	6.4683×10^{-8}

Table 7 shows that the MLP-PSO model produced a smaller error rate than the MLP-BP model. The comparison of the two models can also be seen through graphic visualization as shown in Figure 7 below.



(a) (b)

Figure 7 Performance Comparison Graph: MLP-PSO Model (a) and MLP-BP Model (b)

Figure 6 shows the performance of the model in predicting the test data, based on this graph it can be concluded that the prediction results with the MLP-PSO model showed a more dominant direction to the original data, while the prediction results with the MLP-BP model did not show the right direction in detail.

5. Conclusions

Based on the research that has been done, the following conclusions can be drawn:

1. Convergence testing shows that the BP Algorithm has a drastic decrease in the error rate in the 6th iteration with an error rate of 6.1583×10^{-6}, much better than the particle swarm optimization algorithm which has a drastic decrease in the error rate in the 7th iteration with an error rate of 2.1199×10^{-5}.
2. Based on the best model analysis performed in this research, the PSO Algorithm has a slightly better error rate, of 5.6168×10^{-8}, compared to the BP Algorithm which has an error rate of 6.4683×10^{-8}.
3. The PSO algorithm can be used as a learning algorithm in the multilayer perceptron ANN network.

Bibliography

[1] D. A. S. Fauzi, “Faktor – Faktor Yang Mempengaruhi Nilai Tukar Rupiah Periode 2013 – Triwulan I 2015,” *Jurnal Nusantara Aplikasi Manajemen Bisnis*, *1*(2), 64, 2016.

[2] S. Ichsan, S. Suhadak and S. Sulasmiyati, “Pengaruh Pergerakan Nilai Tukar Yuan Terhadap Ekspor dan Impor Indonesia (Studi Pada Bank Indonesia Periode 2012-2014),” *Jurnal Administrasi Bisnis S1 Universitas Brawijaya*, *35*(2), 94–103, 2016.

[3] J. Kuepper, “*Basic Of Technical Analysis*,” [Online]. Available: www.investopedia/university/technical/. [Accessed 10 September 2017]

[4] C. D. Suhendra and R. Wardoyo, “Penentuan Arsitektur Jaringan Syaraf Tiruan Backpropagation (Bobot Awal dan Bias Awal) Menggunakan Algoritma Genetika,” *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, *9*(1), 77, 2015.

[5] D. Wang, D. Tan and L. Liu, “Particle swarm optimization algorithm: an overview,” *Soft Computing*, *22*(2), 387–408, 2018.

[6] N. Nikentari, H. Kurniawan, N. Ritha, and D. Kurniawan, “Optimasi Jaringan Syaraf Tiruan Backpropagation Dengan Particle Swarm Optimization Untuk Prediksi Pasang Surut Air Laut,” *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 5, no. 5, pp. 605–612, 2018.

[7] J. A. Lazzús, “Neural network-particle swarm modeling to predict thermal properties,” *Mathematical and Computer Modelling*, vol. 57, no. 9–10, pp. 2408–2418, 2013.

[8] M. R. Lubis, “Metode Hybrid Particle Swarm Optimization - Neural Network Backpropagation Untuk Prediksi Hasil Pertandingan Sepak Bola,” *J-SAKTI (Jurnal Sains Komputer dan Informasi.*, vol. 1, no. 1, p. 71, 2017.

[9] F. Rohmawati, M. G. Rohman and S. Mujilahwati, “Sistem Prediksi Jumlah Pengunjung Wisata Wego Kec.Sugio Kab. Lamongan Menggunakan Metode Fuzzy Time Series,” *Jouticla*, *2*(2), 2017.

[10] T. Abbas and D. Irayani, “Pengaruh Nilai Tukar Rupiah dan Produk Domestik Bruto Terhadap Nilai Ekspor Tembakau di Indonesia Tahun 1986-2016,” *Jurnal Ekonomi Pertanian Unimal*, *1*(1), 8–16, 2018.

[11] N. Lestari and L. L. Van FC, “Implementasi jaringan syaraf tiruan untuk menilai kelayakan tugas akhir mahasiswa (studi kasus di amik bukittinggi),” *Digital Zone: Jurnal Teknologi Informasi Dan Komunikasi*, *8*(1), 10–24, 2017.

[12] A. S. Ritonga and S. Atmojo, "Pengembangan Model Jaringan Syaraf Tiruan untuk Memprediksi Jumlah Mahasiswa Baru di PTS Surabaya (Studi Kasus Universitas Wijaya Putra),” *Jurnal Ilmiah Teknologi Informasi Asia*, *12*(1), 15, 2018.

[13] D. A. Navastara, J. Safitri and D. Purwitasari, “Kombinasi Metode Multilayer Perceptron dan Teori Fuzzy untuk Klasifikasi Data Medis,” *IKRAITH Informatika*, *2*, 2018.

[14] S. Panigrahi, Y. Karali and H. S. Behera, “Normalize Time Series and Forecast using Evolutionary Neural Network,” *International Journal of Engineering Research & Technology (IJERT)*, vol. 2, 2013.

[15] Doreswamy and C. M. Vastrad, “Performance Analysis of Neural Network Models for Oxazolines and Oxazoles Derivatives Descriptor Dataset,” International Journal of Information Sciences and Techniques (IJIST), vol. 3, no. 6, 2013.

[16] S. Raschka, “*Python Machine Learning*,” Packt Publishing Ltd, 2015.