An Empirical Study in Forecasting Bitcoin Price Using Bayesian Regularization Neural Network

Rina Sriwiji[1*], Arum Handini Primandari[2]

[1]*Department of Statistics, Universitas Islam Indonesia, Yogyakarta, 55584, Indonesia*
[2]*Department of Statistics, Universitas Islam Indonesia, Yogyakarta, 55584, Indonesia*

[*)]Corresponding author: primandari.arum@uii.ac.id

Abstract. In recent years, Bitcoin has attracted a lot of attention because of its nature that supports encryption technology and monetary units. For traders, Bitcoin becomes a promising investment since its fluctuating prices potentially draw high profit (the higher the risk the higher the return). Unlike conventional stock, Bitcoin trades for 24 hours a day without a closing period, so that it escalates the risk. Predicting the value of Bitcoin is expected to minimize the risk by considering some information such as blockchain information, macroeconomic factors, and global currency ratios. However, the multicollinearity among these independent variables causes regression method cannot be used. This research employs Bayesian Regularization Neural Network (BRNN) which is a free assumption. This method is Single Hidden Layer Feed Forward Neural Network (SLNN) that utilize Bayesian concept to optimize weights, biases, and connection strengths. The data is time series data from January 23, 2017, to January 23, 2019. Regression with subset selection is employed to reduce independent variables, from a total of 25 variables to 14 variables. As a result, the predicted value is not much different from the actual data, with an accuracy of 91.1% based on the MAPE value.

Keywords: Bayesian Regularization Neural Network**,** Bitcoin, Regression.

1 Introduction

Market price predictions have attracted and challenged investors and researchers, because a lot of uncertainty is involved and manny variables affect the market, such as economic conditions and polotic events[7]. In recent years, the market is not only about stocks and foreign exchange (foreign exchange) but also cryptocurrency. Cryptocurrency is a digital currency where transactions can be made online (online) and fund transfer verification, operating independently from the central bank.

Currently there are 904 types of crypto currencies and the most famous is Bitcoin, which was created in 2009 by Satoshi Nakamoto. Bitcoin is a peer-to-peer electronic cash system that allows online payments to be sent directly from one party to another without going through a financial institution[10]. All Bitcoin transactions are posted in ledgers called Blockchain. Blockchain is a distributed database that stores central transaction data that is not regulated by banks or third parties[1]. Based on data from Coinmarketcap.com, as of September 2017, Bitcoin exceeds US $ 60 billion in market capitalization and trading volume is in the range of $ 150-400 per day [2]. With a large amount of trading volume per day, it

makes sense to think of it as the right financial instrument as part of a quantitative trading strategy.

Bitcoin and other cryptocurrency markets can be traded at any time because it does not have a closing period, this is what distinguishes it from other markets. Bitcoin is easier to change and risky for traders. The existing uncertainty factors need to be reduced by traders to minimize risk. One way to do this is to accurately predict the price of Bitcoin.

When making predictions, the right method is needed. Neural network or also called artificial neural network is a computational method that can be used to predict[9]. One type of artificial neural network is Bayesian Regularization Neural Network (BRNN). Bayesian Regularization (BR) is an artificial neural network training algorithm that improves weight values and bias based on Levenberg-Marquardt optimization. This algorithm minimizes the combination of error squares and weights, then determines the correct combination to produce a good network [11]. Good network results can then be used for predictions.

A number of studies have been conducted recently on modeling the timing of the price of Bitcoin as a new market variable with certain technical rules. An empirical study in modelling and prediction of bitcoin price apply Bayesian neural network [7]. Greaves and Au (2015) analyzed data on the Blockchain Bitcoin to predict the price of Bitcoin using SVM and ANN (Multilayer Perceptron) and obtained an accuracy score of 55%. Jang and Lee (2017) conducted an empirical study comparing Bayesian neural networks (BNN) with other linear and non-linear benchmark models in modeling and predicting Bitcoin processes, concluded that the BNN worked well in predicting Bitcoin prices and explained the high volatility of Bitcoin prices . The BRNN method is also well utilized in the field of prediction, as was done by Qiu et al. (2011) in predicting the speed of short-term traffic, shows that the BRNN method is able to improve the ability to generalize neural networks, and can achieve better predictive results than some traditional prediction models[12].

Based on some previous studies, researchers are interested in conducting practical analysis on Bitcoin price modeling and prediction using BRNN, where BRNN can handle an increase in the number of relevant features in the evaluation. BRNN incorporated the term regularization into an objective function to prevent overfitting problems. When machines consider many input variables, trained machines can be complex and suffer from overfitting problems. According to Jang and Lee (2017) in addition to macroeconomic variables, direct use of Blockchain information, such as hash rates, difficulties, and mining revenues, has not been investigated to describe the process of Bitcoin prices[7]. Based on this, a research on Bitcoin price predictions is conducted using Blockchain information, macroeconomic factors, and global currency ratios.

2 Material

2.1 Bitcoin and Blockchain

Bitcoinabbreviated as "BTC" is a digital currency, which is not issued by institutions, organizations or governments in its regulations. Bitcoin utilizes a peer-to-peer network as a distribution media using sophisticated cryptographic protocols. First proposed by Satoshi Nakamoto in 2008, bitcoin software was created and started to run in 2009[10].

All the transaction of bitcoin which is posted in block of legder known as blockchain is verified by miners using cryptography. The verification happens in the system by members (it can be thousand or million computers distributed in the net). The verified block is added to a chain, which is stored across the net, creating not just a unique record. Falsifying a single record would mean falsifying the entire chain. That is virtually impossible.

Blockchain is distributed database stores an immutable record of data. Generally the blockchain acts as a trusted and reliable third party to maintain a shared state, mediate exchanges, and provide a secure computing engine [1].All trading history is recorded in the Blockchain and shared by the network, and all previous transaction history is verified by network participants.
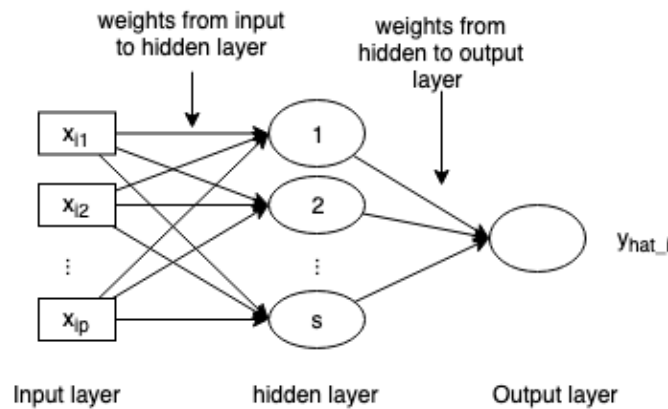
2.2 Neural Network

The basic neural network is the Feed Forward Neural Network which has only one Single Hidden Layer also known as perceptron. In the **fig. 1.,** the prediction is obtained by two step)[5][6]:
1. The input is transformed nonlinearly to hidden layer
2. The output of hidden layer is linearly combined to obtain prediction.

The model written as follow:

$$y_i = \mu + \sum_{k=1}^{s} w_k f_k \left(b_k + \sum_{j=1}^{p} x_{ij} \beta_j^{[k]} \right) + \varepsilon_i \tag{1}$$

Where $(w_1, w_2, \ldots, w_s)$ is the network weights (from hidden layer to output), $(b_1, b_2, \ldots, b_s)$ is bias; $(\beta_1^{[1]}, \ldots, \beta_p^{[1]}; \ldots; \beta_1^{[s]}, \ldots, \beta_p^{[s]})'$ is the strength connections (weights from input to hidden layer) where $\beta_j^{[k]}$ is the parameter for the j-th input in neuron, k=1, 2, …, s.

Fig. 1. Single Hidden Layer Feed Forward Neural Network

The activation function $f_k(\cdot)$ define as follows:

$$f_k(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{2}$$

This function which is known as tanh (tangent hyperbolic) maps the real value to bounded interval (-1,1).

2.2 Bayesian Regularization Neural Network

The number of parameter to be estimated increase, when both the number of inputs (p) and neurons (s) also increase. When that happens, the estimation of neural network become complex. In order to avoid overfitting, the optimization is applied. Supposed that $\theta = (w_1, w_2, \ldots, w_s; b_1, b_2, \ldots, b_s; \beta_1^{[1]}, \ldots, \beta_p^{[1]}; \ldots; \beta_1^{[s]}, \ldots, \beta_p^{[s]})$ is the parameters. The algorithm for fitting model (1) can be summarized as follows:

1. Initialize tuning parameter that are β, α, δ and the weights using the Nguyen and Widrow algorithm.
2. Take 1 step of the Levenberg-Marquardt algorithm to minimize the objective function $Q(\varphi)$:

$$Q(\varphi) = \beta \sum_{i=1}^{n} e_i^2 + \alpha \theta_a' \theta_a - \delta \theta_d' \theta_d \tag{3}$$

3. Update β, α, δ by maximizing Eq. (19) using the Nelder and Mead algorithm.
   $$G(\beta, \alpha, \delta, \psi = \psi^{map})$$
   $$= \log p(y|\psi, \sigma_e^2) \approx c + \frac{n}{2}\log\beta + \frac{m}{2}\log\alpha + \frac{q}{2}\log\delta - \frac{1}{2}\log|\Sigma|_{\varphi=\varphi^{map}} - Q(\varphi)|_{\varphi=\varphi^{map}}$$
4. Iterate Steps 2 and 3 until convergence.

3 Methods

The population in this study is all data on Bitcoin prices, blockchain information, macroeconomic indicators, and global currency ratios. While the research sample is BTC/USD data, Average Block Size, Transaction per Block, Median Confirmation Time, Hash Rate, Difficulty, Cost% per Transaction, Miners Revenue, Confirmed Transaction, Unique Addresses, S & P 500, Euro Stoxx 50 , Dow30, Nasdaq, Crude Oil, Gold Futures, VIX, Nikkei 225, FTSE 100, JKSE, JKLQ45, GBP/USD, JPY/USD, CHF/USD, CNY/USD, EUR/USD each of these variables from period 23 January 2017 to January 23, 2019.

Table 1. The Variables

Type of Variable	Data category	Data
Dependent (y_i)	-	BTC/USD
independent (x_{ip})	blockchain information	Average Block Size, Transaction per Block, Median Confirmation Time, Hash Rate, Difficulty, Cost% per Transaction, Miners Revenue, Confirmed Transaction, Unique Addresses,

Type of Variable	Data category	Data
	macroeconomic indicators	S & P 500, Euro Stoxx 50 , Dow30, Nasdaq, Crude Oil, Gold Futures, VIX, Nikkei 225, FTSE 100, JKSE, JKLQ45,
	global currency ratios	GBP / USD, JPY / USD, CHF / USD, CNY / USD, EUR / USD

The data is obtained from several websites, namely from www.blockchain.com to obtain the bitcoin blockchain information data, then from the website www.finance.yahoo.com and www.investing.com to obtain bitcoin data, macroeconomic variables, and global currency ratios. This study has a total of 520 data, for each variable, so that the total data from this study is 13,520.The following is a description of the steps of data analysis in this study:

1. Preprocessing data

There is a difference in the time interval between bitcoin information and others financial variable. Bitcoin is traded for seven days a week, while others are only in business days. In order to equalize the time interval, Saturday and Sunday data (in bitcoin information) is removed. The cleaning data which imputes missing values comes afterwards.

There are holidays cause missing values. The process to fill out missing value use Moving Average with the length of period 3 (MA(3)). Missing values are imputed by taking the average of k period data before and k period data after missing value. Table 2. shows the example of imputation missing data in February, 20th 2017.

Table 2. Missing Value Imputation Example

Date	Before	After	Condition
14/02/2017	2326.12	2326.12	
15/02/2017	2335.58	2335.58	
16/02/2017	2349.64	2349.64	Before
17/02/2017	2343.01	2343.01	
20/02/2017	NA	2351.958	Average(before:after)
21/02/2017	2354.91	2354.91	
22/02/2017	2361.11	2361.11	After
23/02/2017	2367.5	2367.5	
24/02/2017	2355.73	2355.73	

2. Descriptive analysis of historical bitcoin data.

The line graphics is applied to visualize the data. The visualization helps to understand the component of the data such as trend and seasonal.

3. Perform variable selection using linear regression.

The use of linear regression with stepwise selectionis to obtain the significant variables. This method use the t-value for testing the hypothesis that this variable should be added to, or deleted from, the model. The null hypothesis for t-test assumes the parameter is equal to zero ($\beta_i=0$). The statistics test for t-test is

$$t_0 = \frac{\hat{\beta}_i}{\mathrm{se}(\hat{\beta}_i)}$$

<div align="right">(3)</div>

Where $\mathrm{se}(\beta_i)$ is standar error. The library employed in determine the significant variable is MASS. The function for stepwise selection is stepAIC(model, direction = "both",trace = FALSE), then the sintax summary(stepmodel) is used to show the result.

4. Carry out BRNN analysis, with the following stages:
   a. Divide data into training data and testing data.
   b. Determine the number of neurons.
   c. Determine the model
      The estimation of parameters in the BRNN model is calculated using the help of package BRNN in R[6]. The function used is BRNN(model, data = the data frame contains dependent and independent variable, n = the number of neurons, verbose = FALSE).
   d. Data prediction.
5. Comparing the MAPE value between BRNN and all variables and BRNN with significant variables.

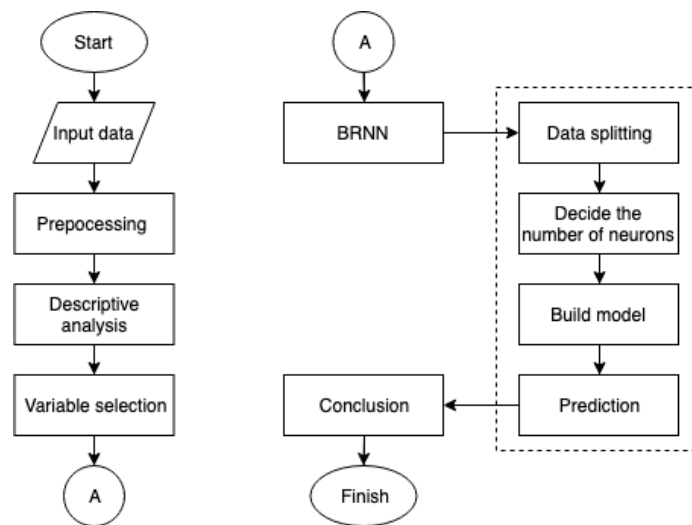The following is a flowchart from the analysis step above:

Fig. 2. Data Analysis Flowchart

4 Result and Discussion
4.1 Descriptive Analysis of Bitcoin
Before analyzing the data using the bayesian regularization neural network method, researchers first conducted a descriptive analysis of historical Bitcoin data to find out an overview of Bitcoin over the past two years, from January 23, 2017 to January 23, 2019.

Fig. 3. Bitcoin Daily Price Chart from January 23, 2017 until January 23, 2019

Based on **fig.3.**it can be seen that the movement of the opening price of bitcoin for 2 years tends to increase or have an upward trend. Where on January 23, 2017 the opening price of Bitcoin was US $ 918.84. The most significant increase occurred at the end of 2017, namely on December 17, 2017 with the price of Bitcoin reaching the highest value of US $ 19,870.62. As quoted on diskartes.com, stated that the price of Bitcoin soared to its peak in 2017. This is because more and more people are interested in becoming miners and coupled with the limitation of the maximum number of Bitcoin circulating is 21 million pieces.

The destruction of the world currency also contributes to rising pricesBitcoin. At that time the pound's decline was accompanied by an increase in the price of Bitcoin. Besides that, it turns out that countries in the world react differently to Bitcoin. Japan choose pro Bitcoin, and successfully raised the price of Bitcoin by 2% in 24 hours. But at the beginning of 2018 the price of Bitcoin began to experience a significant decline. One reason is that China prefers not to want Bitcoin to date. The decision caused the price of Bitcoin to fall by more than 20% in a short time(Diskartes, 2018). In addition, the cause of the decline in the price of Bitcoin is due to theft cases experienced by some of the world's leading exchangers such as Mt.Gox, Bitfinnex, and others. People become disbelievers to save money on exchangers in the form of Bitcoin. In a short time, people who sell Bitcoin are more than those who buy. Until January 23, 2019 the opening price of Bitcoin reached US $ 3,602.04.

4.2 Variable Selection Using Regression

Regression analysis in this study is only used to find out significant variables. From the results of the analysis it can be concluded that there are eleven insignificant variables, namely the median confirmation time, hash rate, difficulty, percentage of cost per transaction, total unique address number, VIX, JKSE, JKLQ45, GBP / USD, JPY / USD, CHF / USD. There are fourteen significant variables in the model, namely variable average block size, transaction per block, miners revenue, confirmed transactions, S & P 500, Euro Stoxx 50, Dow 30, Nasdaq, Crude Oil, Gold Futures, Nikkei 225, FTSE 100, CNY / USD, and EUR / USD.

4.3 Bayesian Regularization Neural Network (BRNN) Analysis

A total of 25 explanatory variables included in the three categories were used as inputs for BRNN learning. The researcher also discussed another set of inputs consisting of 14 input variables by removing non-significant variables as mentioned in the previous subsection.

a. DetermineThe Training Data and TestingData

The research data is divided into training data and testing data by ratio 85% for training data and 15% for testing data which means training data amounts is 442 and testing data is 78. The division for training data is greater than testing data due to learning machines or learning algorithms need to be more trained with data patterns from training data. The training process

is carried out using the BRNN method so that a model is formed which will be tested for its performance on testing data.

b.   Determination of the Number of Neurons

Networks with only a few neurons in hidden layers may not be able to capture the complex relationship between dependent variables and input. However, if too many neurons are assigned to the hidden layer of the network, it will make poor predictive abilities from data that are not visible due to overparameterization. Therefore, a different number of neurons in hidden layers must be tried.

In this study, the network was formed with 25 input variables and 1 output layer, also 14 input variables and 1 output layer. The number of neurons in the hidden layer that will be used for trial-error is in range from 1 to 10 neurons. Therefore, to determine the number of neurons used in the model for forecasting, the comparison takes the smaller MSE value. The following is a table of test results with a number of these neurons.

Table 3. Determination of the Number of Neurons

BRNN	Number of Neurons in the Hidden Layer	Amount of Epochs produced	MSE
25 Input Variables	1	15	815242,8
	2	29	1055285
	3	73	856119,8
	4	57	12672667
	5	58	2948530
	6	74	8248582
	7	119	15301754
	8	209	12185988
	9	141	87091024
	10	107	13892124
14 Input Variables	1	17	384797,3
	2	51	339089
	3	39	2139563
	4	45	2102295
	5	47	395981,3
	6	81	4001355
	7	105	960671,2
	8	148	2485130
	9	136	685773,6
	10	158	1872920

Based on **Table 3**, it can be seen that the model with 2 and 5 neurons produces the smaller MSE compared to the others. The number of epochs (iterations) represents the length of the learning process carried out on the network being observed. Too few epochs cause the network to be formed to be too general, meaning that the networks' ability to recognize patterns is too little or none at all. While too many epochs will cause the network to experience an overfit condition (the network is too specific to the training data).

c. Model Determination

After determining the number of neurons in the hidden layer and getting the number of epochs, then determine the architecture of model. The following is the program code when establishing the BRNN model.

```
brnn(da_train$BTC_USD ~.,data =
da_train,neurons=2,verbose=TRUE)
```

```
brnn(da_train$BTC_USD      ~.,data      =      da_train,
neurons=5,verbose=TRUE)
```

Vebose is set to true means that you will print the iteration history. Whereas in Figure 4 the neurons used are 5.

4.4 Data Prediction

After getting the model in the training process, the model will be tested using testing data. The performance of each model that is trained is measured by the Mean Absolute Percentage Error (MAPE). The following is the MAPE value generated from the prediction using the models.

Table 4.MAPE values from the Model for Prediction

BRNN Model	BRNN	MAPE
First Model	25 Input Variables	19,99905
	14 Input Variables	10,70526
Second Model	25 Input Variables	107,9427
	14 Input Variables	8,917875

Based on **Table 4.**, it is better to use the second model with 14 input variables which has 5 neurons. In the second model with 25 input variables yields the MAPE value more than 100%. This is because the error is far greater than the actual value (e.g. the actual value is 1, while the prediction result is 3, will yield 200% of MAPE).

To see whether the results of the model formed can make good predictions, the following is a plot of the actual data, fitting data, and prediction data.

Fig. 4. Actual Data Plots, Data Fitting, and Prediction Data

Based on **fig.4.**, it can be concluded that the model can follow the pattern of actual data. Fitting data is obtained from the prediction of training data. While predictive data is obtained from the prediction of testing data. Moreover, the following are the results of the Bitcoin opening price prediction for the last three days.

Table 5. Data Testing Prediction Results

Date	Prediction for the Last Three Days	Actual Data
21/01/2019	3303,975	3567,72998
22/01/2019	3467,067	3571,919922
23/01/2019	3319,767	3602,040039

The prediction of testing data has only 8.9% of MAPE which means it has 91.1% accuracy. When this research is completed, there is new data. Then predictions are made on new data and results are obtained as shown in **Table 6.**

Table 6. Price Comparison of Predictions with Actual Prices

Date	Prediction Results	Actual Data	e^2
24/01/2019	3459,306	3572,050049	12711,22058
25/01/2019	3440,368	3598,52002	25012,06143
28/01/2019	3295,472	3565,080078	72688,51572
29/01/2019	3366,896	3453,419922	7486,389078
30/01/2019	3377,238	3418,25	1681,984144

Based on Table 5, it can be seen that the prediction results for new data using 14 input variables and 5 neurons produce predictive values that are not much different from the actual data. The MSE value obtained is 23916.03419 and MAPE is 3.763785.

5 Conclusion

There are total 25 independent variables contains blockchain information, macroeconomic factors, and global exchange ratio in this research. The variable selection employs linier regression with stepwise method which result only fourteen significant variables. These variables are blockchain information(average block size, transaction per block, miners revenue, confirmed transactions), macroeconomics (S & P 500, Euro Stoxx 50, Dow 30, Nasdaq, Crude Oil, Gold Futures, Nikkei 225, FTSE 100), dan global exchange ratio(CNY / USD, and EUR / USD).

The trial and error method to determine the number of neurons use 1 till 10 neurons. Comparing the error of training data, the use of 2 and 5 neurons produces the smaller MSE than the others. Based on MAPE of testing data, we obtain the best model of BRNN which has 5 neurons. In the process of finishing this research, there are a set of new data. The forecasting of the new data has 3.7% MAPE. Thus, the BRNN is good enough to make prediction.

References

[1] Cachin, C., & Vukolić, M: Blockchain Consensus Protocols in the Wild. https://doi.org/10.4230/LIPIcs.DISC.2017.1. (2017).

[2] Coinmarketcap.com: Bitcoin (BTC) Price, Charts, Market Cap, and Other Metrics. Retrieved March 15, 2019, from https://coinmarketcap.com/currencies/bitcoin/. (2018).

[3] Diskartes: Apa Penyebab Harga Bitcoin Naik Selangit Kemudian Terjun Bebas? Retrieved June 18, 2019, from http://diskartes.com/2018/11/harga-bitcoin/. (2018).

[4] Foresee, F. D., & Hagan, M. T.: Gauss-Newton Approximation To Bayesian Learning. International Joint Conference on Neural Networks. pp 1930–1935. (1997).

[5] Gianola, D., Okut, H., Weigel, K. A., & Rosa, G. J. M.: Predicting complex quantitative traits with Bayesian neural networks: A case study with Jersey cows and wheat. BMC Genetics, 12(February 2019). https://doi.org/10.1186/1471-2156-12-87. (2011).

[6] Gianola, D., Rosa, G. J. M., Pérez-Rodríguez, P., Gianola, D., Weigel, K. A., Rosa, G. J. M., & Crossa#, J.: Technical Note: An R package for fitting Bayesian regularized neural networks with applications in animal breeding Investigating functional relationships among fertility and health traits in dairy cows View project Seeds of Discovery -MasAgro Biodiversida. Article in Journal of Animal Science J. Anim. Sci, 91(May).pp. 3522–3531. https://doi.org/10.2527/jas2012-6162. (2013).

[7] Huisu, Jang & Jaewook, Lee: An Empirical Study on Modelling and Predicting of Bitcoin Prices With Bayesion Neural Network Based on Blockchain Information. IEEE-Access. Vol. 6. Pp. 5427-5437. (2017).

[8] Khaidem, L., Saha, S., & Dey, S. R.: Predicting the Direction of Stock Market Prices using Random Forest, (May 2016). Retrieved from http://arxiv.org/abs/1605.00003. (2016).

[9] Kusumodestoni, R. H., & Zyen, A. K.: Prediksi Kecepatan Angin Menggunakan Neural Network Untuk Mengetahui Besar Daya Listrik Yang Dihasilkan. Vol. 6(1).(2015).

[10] Nakamoto, S.: Bitcoin: A Peer-to-Peer Electronic Cash System. Retrieved from www.bitcoin.org. 1–9. (2008).

[11] Pan, X., Lee, B., & Zhang, C.: A Comparison of Neural Network Backpropagation Algorithm for Electricity Load Forecasting. In IEEE International Workshop on Inteligent Energy Systems (IWIES). pp. 22–27. Vienna, Austria: IEEE. (2013).

[12] Qiu, Chenye & Wang, Chunlu & Zuo, Xingquan & Fang, Binxing: A Bayesian regularized neural network approach to short-term traffic speed prediction. Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics. 2215-2220. 10.1109/ICSMC.2011.6084006. (2011).

[13] Wellmann, R., & Bennewitz, J.: Bayesian models with dominance effects for genomic evaluation of quantitative traits. Genetics Research, Vol. 94(1), pp. 21–37. (2012)