Adaptive neuro-fuzzy inference system for forecasting rubber milk production

R F Rahmat[1], Nurmawan[1], S Sembiring[1], M F Syahputra[1] and Fadli[2]

[1] Department of Information Technology, Faculty of Computer Science and Information Technology, Universitas Sumatera Utara, Medan, Indonesia
[2] Faculty of Management and Business, Universitas Sumatera Utara, Medan, Indonesia

Email: romi.fadillah@usu.ac.id

Abstract. Natural Rubber is classified as the top export commodity in Indonesia. Its high production leads to a significant contribution to Indonesia's foreign exchange. Before natural rubber ready to be exported to another country, the production of rubber milk becomes the primary concern. In this research, we use adaptive neuro-fuzzy inference system (ANFIS) to do rubber milk production forecasting. The data presented here is taken from PT. Anglo Eastern Plantation (AEP), which has high data variance and range for rubber milk production. Our data will span from January 2009 until December 2015. The best forecasting result is 1,182% in term of Mean Absolute Percentage Error (MAPE).

1. Introduction

Rubber is one of the leading plantation commodities that contribute significantly to non-oil exports. The increase in rubber exports is very significant, of export volume of 1.496 million tons valued at USD 1038 million in 2002 increased to 2.35 million tons valued at USD 7330 million in 2010 or an increase of 57%. The export contribution of plantation sector above is expected to be the driver of national economic development in the future [1]. The increased demand for rubber spurred an increase in the rubber milk production in rubber producing countries in the world, including Indonesia. Rubber is a vital need for human life, and it is associated with the mobility of human and goods that require components made of rubber such as tires, conveyor belt, transmission belt, etc. Rubber industries have to prepare raw materials in accordance with market demand to meet those needs. Rubber milk production is paramount to rubber industries. Rubber companies expect the accurateness of rubber milk production in term of its amount and time. Therefore, state-owned or private plantation companies require a system that can predict the production precisely.

Forecasting is an attempt to predict future circumstances through examination of the past circumstances [2]. One of the quantitative forecasting techniques is time series data forecasting, a set of observations built in sequence within a specified time in which data can be data daily, weekly, monthly, etc. [3]. Time series data forecasting has been developed in soft computing such as fuzzy method, genetic algorithm, neural networks and hybrid method.

ANFIS is a method that uses a neural network to implement fuzzy inference system. In other words, ANFIS is the incorporation of fuzzy inference system mechanisms depicted in artificial neural network architecture. In statistical modeling, ANFIS is applied to classification, clustering, regression, and forecasting in time series data [10].

Several kinds of research have been conducted. In 2013, Agus Harianto did a research on the forecast of rubber milk production using Autoregressive Integrated Moving Average (ARIMA) method [4]. The best ARIMA model for rubber milk production forecasting is ARIMA model (2,1,1) with an estimated R2-squared parameter of 93.20, AIC value (Akaike Information Criterion) of 25.64, SIC value (Schwarz Information Criterion) of 25.84 and significant probability close to zero. With the measurement result of forecasting error with MAPE (Mean Absolute Percentage Error) parameter of 1.37, MAD (*Mean Absolute Deviation*) value of 380,5, MSE (*Mean Squared Error*) value of 5.194.584.081 and MPE (*Mean percentage Error*) value of 0.01. The analysis of production forecasting shows that production of 2012 to 2013 is predicted to be increasing year by year with a small error rate.

In 2015, Hani'ah did adaptive neuro-fuzzy inference system (ANFIS) implementation for water use forecasting at Tirta Moedal Semarang drinking water company [5]. Ria in 2014, did weather forecasting using case-based reasoning and adaptive neuro-fuzzy inference system methods [6]. Gernowo and Yenny in 2013 used ANFIS to predict the variability of Indonesian Rupiah exchange value against US dollar. The result shows that the proposed method is capable in predicting the rise and fall of exchange rate fluctuations [7]. Saputra in 2012 applied ANFIS to analyze time series data [8]. Widyapratiwi in 2012 did short-term power load forecasting in Bali using adaptive approach neuro-fuzzy inference system (ANFIS) [9]. Several other methods to do forecasting also has been analyzed such as exponential smoothing method for palm oil production [11], evolving multilayer perceptron [12][16], distributed and adaptive neural network [13] [14] [17] and fast learning algorithm [15].

A company needs a precise forecasting of rubber milk production to be used as a reference to compose a business plan. The forecasting of rubber milk production done manually usually experiences a problem in the form of production prediction results that are not in accordance with the actual production results, so it required a proper method that generates a précised forecasting of the production to be used as a reference to compose a business plan of plantation company.

2. Methodology

Forecasting system of rubber milk production using adaptive neuro-fuzzy inference system (ANFIS) is a system that delivers results in rubber milk production forecasting based on time series data in the past. The system will receive input in the form of monthly production data. Afterwards, the data will be processed using adaptive neuro-fuzzy inference system method to generate the forecast result. The general flowchart of the proposed method can be seen in Figure 1. The principal work of the forecasting system of rubber milk production to be built is designed as follows:

1. Input CSV file data containing rubber milk production data through input form.
2. Set the input parameters in the form of with maximum epoch and maximum error.
3. Form network structure of ANFIS.
4. Classify rubber milk production data inputted into two clusters using fuzzy c-means algorithm and determine the type of *membership* function to be used. The system will use generalized bell membership function, as follows:

$$\mu(Z) = \frac{1}{1 + \left|\frac{Z - c}{a}\right|^{2b}}$$

Where: $\mu(Z)$ = membership degree
 Z = the amount of production
 c = premise parameter, the initial value is determined using mean cluster
 a = premise parameter, the initial value is determined using the standard deviation of the cluster
 b = premises parameter value 1

> The output value of generalized bell membership function is layer 1 output of ANFIS.
5. Compute layer 2 output by multiplying weight values in the corresponding layer 1.
6. Compute layer 3 output by dividing the weight of each w1 or w2 by the amount of weights w1 + w2 in the second layer.
7. Compute output layer 4 by multiplying the output weights of layer 3 3 w_1f_1 and w_2f_2 also set the consequent parameter values using LSE Recursive.
8. Compute layer 5 output, in the form of output prediction which error value can be calculated by subtracting the actual output value with the value of output prediction. If the error is smaller than the maximum error specified in the input form, the system will stop and generate an output prediction.
9. The system will store the production data into a database along with all premise and consequent parameters.
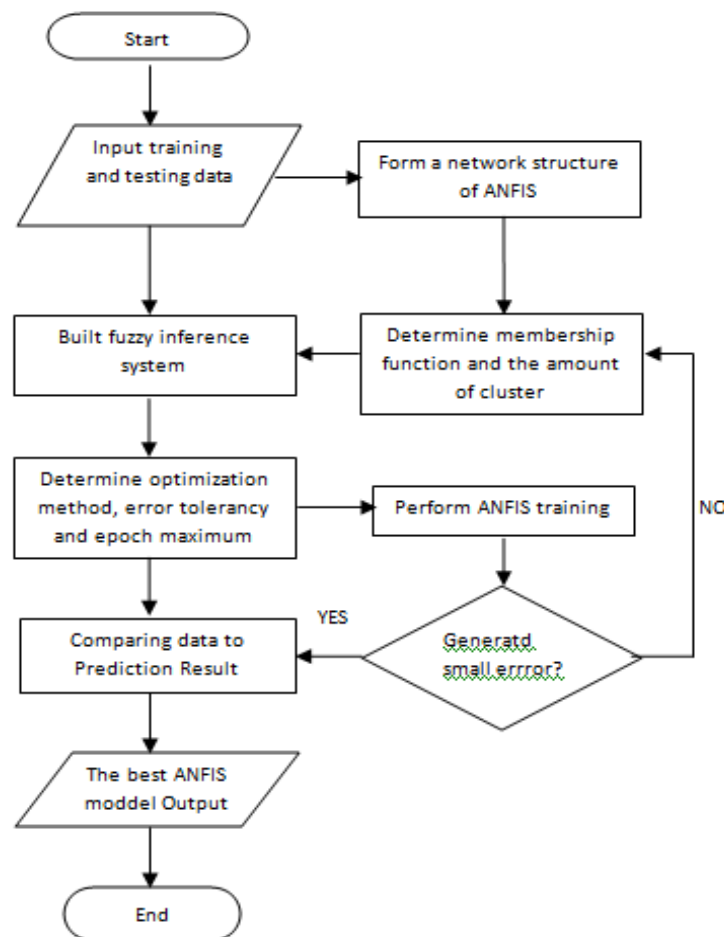

Figure 1. Flowchart of the System

2.1. ANFIS Simulation

The implementation of ANFIS calculation in the prediction of rubber milk production can be simulated as follows. Supposed the given time series data of 8 rubber milk production data are shown in Table 1

Table 1. Rubber milk Production Data's

t	1	2	3	4	5	6	7	8
Z_t	3875	3672	2697	2226	2364	2498	3111	4849

Given ANFIS input in the form of Z_{t-1} and Z_{t-2} while the output is in the form of Z_t as shown in Table 2.

Table 2. Input and Output Data of ANFIS

Z_{t-2}	Z_{t-1}	Z_t
3875	3672	2697
3672	2697	2226
2697	2226	2364
2226	2364	2498
2364	2498	3111
2498	3111	4849

Since the number of clusters is 2, then Sugeno model of ANFIS rule is as follows:

If Z_{t-2} is A_1 and Z_{t-1} is B_1 then $Z_t = p_1Z_{t-2} + q_1Z_{t-1} + r_1$

If Z_{t-2} is A_2 and Z_{t-1} is B_2 then $Z_t = p_2Z_{t-2} + q_2Z_{t-1} + r_2$

Afterwards, clustering process will be performed using fuzzy c-means clustering. Clustering result can be seen in Table 3.

Tabel 3. FCM Cluster Data

Membership Degree		Cluster Tendency	
Z_{t-1}	Z_{t-2}	*Cluster 1*	*Cluster 2*
0.0567545497136	0.9432454502863		*
0.1624564377419	0.8375435622581		*
0.9366214242975	0.0633785757024	*	
0.9778185681115	0.0221814318884	*	
0.9972202565087	0.0027797434912	*	
0.8155083611785	0.1844916388214	*	

Then input the value of clustering to ANFIS network, as follows:

Layer 1:

Since the membership function used is generalized bell:

$$\mu(Z) = \frac{1}{1 + \left|\frac{Z - c}{a}\right|^{2b}}$$

Determined the value of *b*=1 also as the initial value of a and c, mean value and standard deviation were used can be seen in Table 4.

Table 4. Mean and Standard Deviation

Data to-	Cluster 1		Cluster 2	
	Z_{t-1}	Z_{t-2}	Z_{t-1}	Z_{t-2}

1			3875	3672
2			3672	2697
3	2697	2226		
4	2226	2364		
5	2364	2498		
6	2498	3111		
m	2446.25	2549.75	3773.5	3184.5
sd	200.68943	390.29764	143.54267	689.42911

Then neurons of layer 1 can be calculated 1 as follows:

$$\mu_{A1}(x) = \frac{1}{1 + \left|\dfrac{Z_{t-2} - 2446.25}{200.689436692}\right|^2}$$

$$\mu_{A2}(x) = \frac{1}{1 + \left|\dfrac{Z_{t-2} - 2549.75}{390.297642831}\right|^2}$$

$$\mu_{B1}(x) = \frac{1}{1 + \left|\dfrac{Z_{t-1} - 3773.5}{143.542676580}\right|^2}$$

$$\mu_{B2}(x) = \frac{1}{1 + \left|\dfrac{Z_{t-1} - 3184.5}{689.429111656}\right|^2}$$

The result of layer 1 output in the form of membership degree of each data can be seen in Table 5.

Table 5. Membership Degree

Cluster 1		*Cluster* 2	
μ_{A1}	μ_{A2}	μ_{B1}	μ_{B2}
0.006685503	0.02289485	0.00144492	0.04477173
0.006685503	0.02289485	0.00144492	0.04477173
0.390456057	0.87539817	0.00853060	0.34096172
0.453631295	0.59239523	0.01026481	0.41384322
0.856188908	0.81532900	0.01250648	0.50212881
0.937653233	0.98272332	0.044840082	0.98876204

Layer 2:
Layer 2 output can be computed as follows:
$$w_1 = \mu_{A1}(x).\mu_{B1}(X); \quad w_2 = \mu_{A2}(x).\mu_{B2}(X);$$

The calculation result of layer 2 output can be seen in Table 6.

Table 6. Layer 2 Output

Layer 2 output	
w_1	w_2

9.6600532144628E-6	0.0010250424471626
9.6600532144628E-6	0.0010250424471626
0.0033308282542147	0.29847727756169
0.004656439238402	0.24515875557234
0.010707917147754	0.4094001885337
0.042044448460943	0.97167952783949

Layer 3:
Layer 3 output can be computed as follows:

$$\overline{w}_1 = \frac{w_1}{w_1 + w_2}; \qquad \overline{w}_2 = \frac{w_2}{w_1 + w_2}$$

The calculation result of layer 2 output can be seen in Table 7.

Table 7. Layer 3 Output

Layer 3 output	
$\overline{w}_1$	$\overline{w}_2$
0.0093360683007364	0.99066393169926
0.0093360683007364	0.99066393169926
0.011036245183707	0.98896375481629
0.018639535685288	0.98136046431471
0.025488480233879	0.97451151976612
0.041475243206127	0.95852475679387

Layer 4:
Layer 4 output can be computed as follows:
$$\overline{w}_1 f_1 = (\overline{w}_1 Z_{t-2})p_1 + (\overline{w}_1 Z_{t-1})q_1 + (\overline{w}_1)r_1$$
$$\overline{w}_1 f_2 = (\overline{w}_2 Z_{t-2})p_2 + (\overline{w}_2 Z_{t-1})q_2 + (\overline{w}_2)r_2$$

The calculation result of layer 4 output can be seen in Table 8.

Tabel 8. Layer 4 Output

$\overline{w}_1 Z_{t-2}$	$\overline{w}_1 Z_{t-1}$	$\overline{w}_1$	$\overline{w}_2 Z_{t-2}$	$\overline{w}_2 Z_{t-1}$	$\overline{w}_2$
36.17	34.28	0.0093	3838.82	3637.71	0.99
34.28	25.17	0.0093	3637.71	2671.82	0.99
29.76	24.56	0.0110	2667.23	2201.43	0.98
60.25	44.06	0.0186	2184.50	2319.93	0.98
60.2547	63.6702	0.0254	2303.74	2434.32	0.97
103.6051	129.0294	0.0414	2394.39	2981.97	0.95

The estimation of consequent parameters (p_1, p_2, r_1, p_2, q_2, and r_2) can be computed using LSE recursive method as follows:

$$\theta = (A^T A)^{-1} A^T y$$

Where A is the matrix of layer 4 output, y is production data output and θ is the consequent parameter. Therefore, the parameters value obtained are:

p_1 = 205.2208783291, p_2 = -8.7312130543796,
r_1 = -415469.33821946, p_2 = -2.1247969600763,
q_2 = 0.60963523398073dan r_2 = 5440.6762735378.

Layer 5:
layer 5 output can be computed by multiplying the obtained consequent parameters with layer 4 output (A matrix), layer 5 output can be seen in Table 9.

Table 9. Layer 5 Output

Z_t target	$Z_t output$	Error
2697	2696.9999999889	1.1070369E-8
2226	2225.9999999885	1.1506019E-8
2364	2363.9999999906	9.4355527E-9
2498	2497.9999999881	1.1867541E-8
3111	3110.9999999842	1.5775185E-8
4849	4848.9999999772	2.28219505-8

The estimation of premise parameters (a,c) is using error gradient descent propagation model.

3. Experimental Results
Data used for the testing in this research is monthly rubber milk production data of private company period January 2009 - December 2015. Production data consists of 696 data. The composition of training data amounted to 588 data and data testing amounted to 108 data.

The training result of ANFIS method can be tested using mean absolute percentage error (MAPE) as follows:

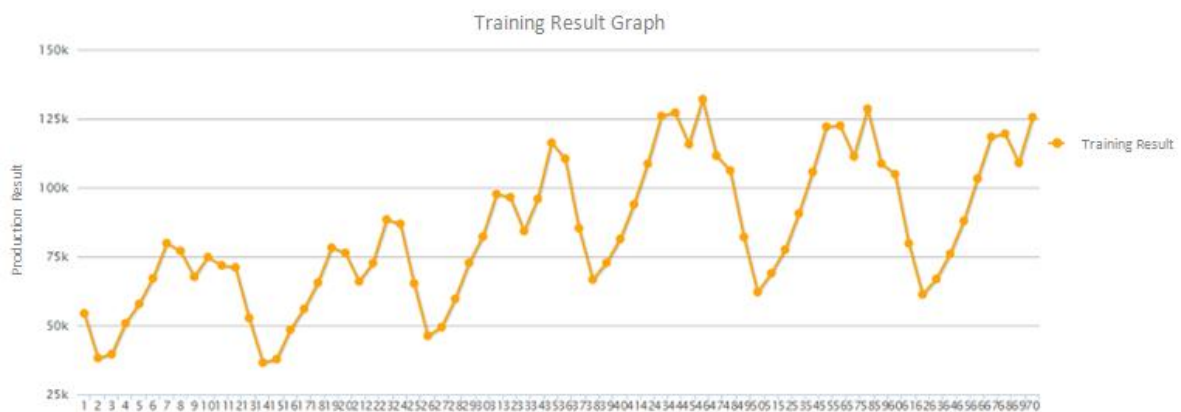$$Error = \frac{\Sigma_1^n(\frac{|y - p|}{y})}{n} \times 100\%$$

Where y is the actual data, p is the prediction data, and n is the amount of data. From the result, obtained MAPE value of 1.182%.

Figure 2. Training and Testing Results from the System

Data training aims to find the premise and consequent parameter values to be to test the data that do not train (See Figure 2 and 3). Training data used is rubber milk production data period January 2009 – December 2014 with total data of 588. Graphic of rubber milk production training result can be seen in Figure 3.

Figure 3. Graphic of rubber milk production training result

After training result was obtained, continued by performing a test on 108 data testing of rubber milk production data period January 2015 – December 2015. Based on training data, produced the smallest error value during training process which is 1.198%. The testing aims to test the accuracy of ANFIS architecture in identifying time series data patterns. From the testing results, obtained the smallest MAPE error value of 1.182%. The testing result of rubber milk production prediction that generates MAPE error value of 1.182% can be seen in Table 10, Figure 2 and Figure 4.

Table 10. Result of data testing

No.	Targeted Output (Kg)	Predicted Output (Kg)	MAPE Error (%)
1	105474	106615.81749049	
2	100846	101905.17110266	
3	77513	78371.711026616	
4	58620	59414.448669202	
5	64565	65486.3878327	
6	72846	73732.243346008	1,182 %
7	85016	85953.231939163	
8	99000	100129.65779468	
9	114608	115797.87072243	
10	116237	117578.3269962	

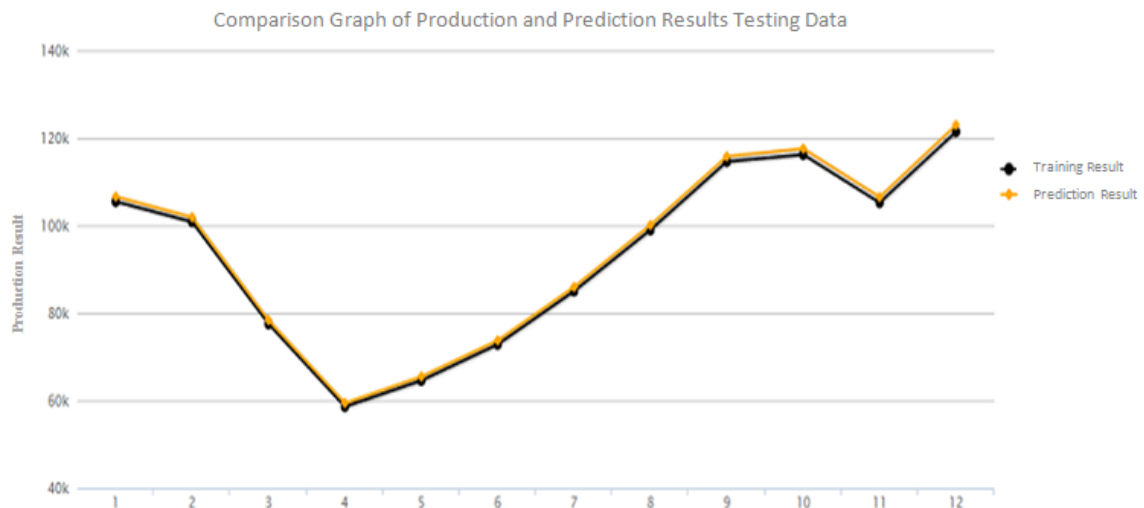| 11 | 105235 | 106521.90114068 | |
| 12 | 121414 | 122975.05703422 | |

Figure 4. Graphic of testing result

The prediction testing is conducted to find out the result of rubber milk production for the next 12 periods. The testing result will show the predictive production data and prediction testing graphic. In Figure 4.5. Can be seen the results of production predictions and the graphs for production period January 2016 - December 2016.
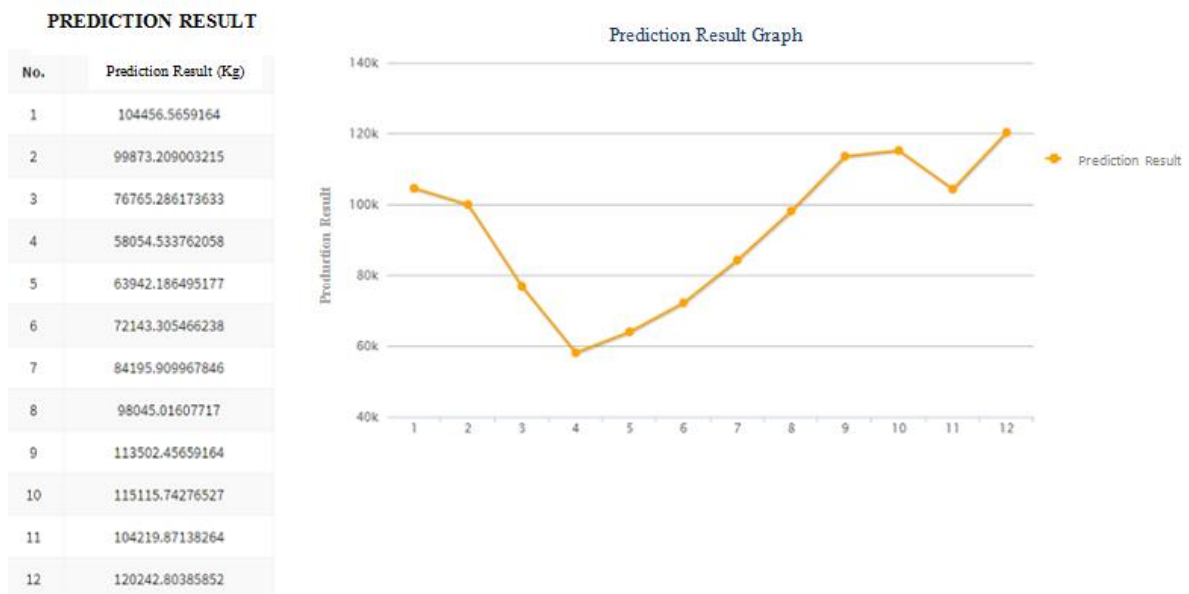

Figure 4.5. 2016 production prediction result

Based on the results of data training and data testing that have been conducted, it can be seen that the ANFIS prediction of rubber milk production generated MAPE error of 1.182%. When compared to Harianto's research in 2013 which researched rubber milk production using ARIMA method that produces MAPE error of 1.370%, the prediction of rubber milk production using ANFIS generates relatively smaller error compared to ARIMA method.

4. Conclusion

Based on the result, ANFIS is successfully implemented in a system to predict the production of rubber milk. The error rate of prediction result using ANFIS is relatively smaller, which is MAPE error value of 1,182%, if compared to previous research done by Harianto using ARIMA with MAPE error value of 1,370%.

For further research, we can apply other fuzzy determination methods to determine the value of fuzzy in the early stage of adaptive neuro-fuzzy inference system method. It is expected that adaptive neuro-fuzzy inference system method can be analyzed and compared to other methods in similar research data, to get more accurate prediction results by comparing the mean error between those methods.

5. References

[1] General Directorate of International Trade Cooperation. 2011. *ASEAN Community in a Global Community of Nations*.

[2] Rambe M F 2002 Market Demand and Sales Prediction Analysis. Scientific Journal of Management and Business, Management Department, Faculty of Economy **1** pp 1-12

[3] Halim S 2006 Time series analysis. Diktat. Universitas Petra Christian.

[4] Harianto A 2013 Forecasting Analysis of Rubber Production in PT Perkebunan Nusantara IX(PERSERO) Batujamus Plantation of Karanganyar Regency. *Agribusiness Review* **1** pp 45-64.

[5] Hani'ah U 2015 *Adaptive Neuro-Fuzzy Inference System* (ANFIS) Implementation for water use forecasting at Tirta Moedal Semarang drinking water company. Thesis. Universitas Negeri Semarang.

[6] Ria 2014 Weather Prediction Using Case-Based Reasoning and Adaptive Neuro-Fuzzy Inference System Methods. *Journal of Informatics* **12**(2): 90-95.

[7] Gernowo R dan Ernitawati Y 2013 Fuzzy Inference System Of The Variability of Indonesian Rupiah Exchange Value against US Dollar Prediction on Adaptive Neural Network as Economical Analysis. *Journal of Science and Mathematics* **21**(2): 39-41.

[8] Saputra A H 2012 Time series Data Analysis Using Adaptive Neuro-Fuzzy Inference System Method. Thesis. Universitas Diponegoro.

[9] Widyapratiwi L K 2012. Short-term Power Load Forecasting In Bali Using Adaptive Neuro-Fuzzy Inference System (ANFIS). Thesis. Universitas Udayana.

[10] Jang, JSR., CT Sun, dan E Mizutani. 1997. *Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence*. London: Prentice-Hall, Inc.

[11] Siregar B, Butar-butar I A, Rahmat R F, Andayani A and Fahmi F 2017 Comparison of Exponential Smoothing Methods in Forecasting Palm Oil Real Production *Journal of Physics: Conference Series* **801** p. 12004.

[12] Rahmat R F, Rizki A, Alharthi A F and Budiarto R 2016 Big Data Forecasting Using Evolving Multi-layer Perceptron *4th Saudi International Conference on Information Technology (Big Data Analysis) (KACSTIT)* pp. 1–5.

[13] Pasha M F, Rahmat R F, Budiarto R, and Syukur M 2010 A distributed autonomous neuro-gen learning engine and its application to the lattice analysis of cubic structure identification problem *International Journal of Innovative Computing, Information and Control* **6** pp. 1005–1022

[14] Rahmat R, Pasha M F, Syukur M, Budiarto R 2015 A gene-regulated nested neural network *International Arab Journal of Information Technology* **12** pp. 532–539

[15] Alwajeeh T, Alharthi A F, Rahmat R F and Budiarto R 2017 Fast Learning for Big Data Using Dynamic Function *IOP Conference Series: Materials Science and Engineering* **190** p. 12015

[16] Nababan E B, Iqbal M and Rahmat R F 2016 Breast Cancer Identification on Digital Mammogram Using Evolving Connectionist Systems *International Conference on Informatics and Computing (ICIC)* pp. 132–136

[17] Aaron, Sitompul O S and Rahmat R F 2014 Distributed autonomous Neuro-Gen Learning Engine for content-based document file type identification *International Conference on Cyber and IT Service Management* pp. 63–68