

Failure prediction of e-banking application system using Adaptive Neuro Fuzzy Inference System (ANFIS)

Yuwono Abdillah, Suharjito

Computer Science Department, Binus Graduate Program-Master in Computer Science,
Bina Nusantara University, Indonesia

Article Info

Article history:

Received Mar 2, 2018

Revised Jul 30, 2018

Accepted Aug 16, 2018

Keyword:

ANFIS

e-banking failure

FCM clustering

ABSTRACT

Problems often faced by IT operation unit is the difficulty in determining the cause of the failure of an incident such as slowing access to the internet banking url, non-functioning of some features of m-banking or even the cessation of the entire e-banking service. The proposed method to modify ANFIS with Fuzzy C-Means Clustering (FCM) approach is applied to detect four typical kinds of faults that may happen in the e-banking system, which are application response times, transaction per second, server utilization and network performance. Input data is obtained from the e-banking monitoring results throughout 2017 that become data training and data testing. The study shows that an ANFIS modeling with FCM optimized input has a RMSE 0.006 and increased accuracy by 1.27% compared to ANFIS without FCM optimization

Copyright © 2019 Institute of Advanced Engineering and Science.
All rights reserved.

Corresponding Author:

Suharjito,

Department of Computer Science,

Bina Nusantara University,

Angrek Campus, Jl. Kebon Jeruk Raya No. 27, Jakarta Barat, Indonesia.

Email: suharjito@binus.edu

1. INTRODUCTION

Monitoring e-banking services is one of the activities involves several input variables at once i.e applications, databases and infrastructure. With the variety of monitoring methods used to monitor incidents and problems by the Operation Team, escalation processes become increasingly difficult in determining the right course of action. It takes a prediction model that can take into account the factors of uncertainty to reduce errors in taking an action. Inputs can be application performance, transaction per second (tps), server utilization and network performance. Problems often encountered by the IT operations unit is the difficulty of determining the cause of the failure of an incident problems that could disrupt banking operations services. This causes the service recovery process to be longer because it needs to identify the potential failure of service one by one.

ANFIS can build input output relationships based on human knowledge in the form of fuzzy rules [1]. ANFIS method proved able to solve problems faced by financial institutions such as fraud on credit card services [2] and customerloan assessment issues [3]. In ANFIS learning process by combining method of Last Square Estimator (LSE) and Error Backpropagation (EBP). The EBP method is applied to layer-1 while the LSE method is applied to layer-4. The aim of this paper is to model the prediction of e-banking system failure using ANFIS-FCM.

2. RELATED WORKS

ANFIS approach with Fuzzy C-means is used in predicting the number of airport passengers [4]. In the case study, ANFIS modeling using FCM method uses time series input over a period of time.

The combination of ANFIS using FCM in predicting the number of passengers proved better because it has a small MAPE value which means the accuracy of the modeling is very good.

ANFIS approach as a technique in texture classification of images [5], where ANFIS is used as a classifier. The methodology used is to collect the image to be processed by DCT and GLCM method to generate input and output. The process data-set is divided into training data and data testing for use in the ANFIS process. The resulting modeling is able to obtain excellent accuracy for both combinations of input data in the form of DCT and GLCM

To obtain pest and classification of rice disease based on existing symptoms, [6] applies Fuzzy logic optimization with bee colony. Bee colony is used as the initial centres for fuzzy logic. This research concludes that fuzzy logic (fuzzy C-means and Fuzzy K-means) optimization using bee colony can improve model accuracy compared to 13.87% for FCM and 18.39% without optimization.

Studies [7] have shown to estimate the thunderstorm activity based on Multiple Linear Regression (MLR), Dvorak technique, and Adaptive Neuro-Fuzzy Inference System (ANFIS). Input used is a combination of Pressure (P), Temperature (T), Relative Humidity (H), Cloud (C), Precipitable Water Vapor (PWV) and Precipitation (Pr). The results show that the activity of thunderstorms during intermonsoon is higher than in other seasons this corresponds to the data obtained manually with an error rate below 50%.

Adaptive Neuro Fuzzy Inference (ANFIS) and Multiple Linear Regression (MLP) method is used to use to predict of solar radiation condition because the condition of solar irradiation will indirectly affect the electrical grid system [8]. The result of study proves that ANFIS method gives better RMSE and MAE value compared with MLP method that is 12.51% and 11.71%.

To identify the regions around the ice infested areas, four methods are used to classify the SAR ice image are performed i.e. Back Propagation Neural Network (BPN), Fuzzy Classifier, Adaptive Neuro Fuzzy Inference System (ANFIS), and ANFIS with Particle Swarm Optimization (PSO) [9]. Input is taken are the pixel values of an image ANFIS with Particle Swarm Optimization (PSO) classifier is best suitable for classifying the SAR ice image than other algorithms with accuracy 93.33 %.

3. RESEARCH METHOD

This study consists of several sections, background, objectives and literature review, which is used as the basis for ANFIS-FCM development. The input data used in this study is derived from the output value generated by BMC Control-M period of 2017. The researcher also implements shell scripts to optimize the output results obtained from various other monitoring tools that cannot be applied to primary monitoring tools. In this paper, the data will be grouped into training and testing data. The main variables tested in decision making in e-banking applications, i.e.: application performance (application response time in processing a request), transactions per second (the number of transactions capable of processing an application), CPU usage, and Network I/O (the number of network connections that occur in a given time).

The steps of the research process start from study literature, the literature study was conducted to obtain a comprehensive overview of the ANFIS method by optimizing fuzzy c-means input (FCM). Next steps is collect data, the data obtained comes from the results of e-banking monitoring during 2017. The third step is ANFIS, in layer-1 FCM process will group the data into multiple clusters that have been determined to optimize parameter rules. The fourth step is analyse output, this stage it will be known whether ANFIS modelling has been made capable of recognizing different inputs with inputs in the previous training process. The dataset used in the testing is the second data period (20% of total data). The last step is testing result, there are several methods such as MSE, MAPE and RMSE to measure the performance of the model that has been made whether the decision is accurate.

4. PROPOSED METHOD

The method proposed in this study is to model the prediction of e-banking system failure using ANFIS-FCM. The process of generating FIS is a process based on the clustering results using FCM. In ANFIS, the FCM process is a process that occurs on the first layer.

4.1. Fuzzy C-Means Clustering

Fuzzy C-Means (FCM) is a clustering algorithm that applies to various issues related to feature analysis, classification and classification design [10]. Fuzzy c-means (FCM) is a method for classifying that allows a data to have more than one cluster with different membership values [11]. Fuzzy c-means algorithm divides the available data from each finite data element and then puts it into the part of the cluster collection that is affected by some given criteria. At first, the cluster centers $c_i = 1, 2, \dots, C$ = randomly from the n points. In the initial conditions, the center of the cluster is still not accurate. Each data has a membership degree for

each cluster. The cluster center and membership value will be recurrently repaired until the cluster center is compliant. After that, the membership matrix U using the following equation is computed. The complete algorithm used in FCM is as follows [12] [13].

1. Input data is matrix $n \times m$ (n =sample data, m =atribute each data)
2. Determine :
 - Cluster amount = c
 - Rank value = w
 - Maximum iteration = $\max I_{\text{ter}}$
 - Smallest error = ξ
 - Objective function = P_0
 - Initial iteration = $t=1$
3. Generate initial matrix

$$U_0 = \begin{bmatrix} \mu_{11} x_1 & \mu_{12} x_2 & \mu_{1c} x_c \\ \mu_{21} x_1 & \mu_{22} x_2 & \mu_{2n} x_n \\ \vdots & \vdots & \vdots \\ \mu_{c1} x_1 & \mu_{c2} x_2 & \mu_{cn} x_n \end{bmatrix} \quad (1)$$

Matrix must match the following conditions:

$$\mu_{ij} = [0,1], 1 \leq i \leq n; 1 \leq k \leq c \quad (2)$$

$$\sum_{i=1}^n \mu_{ik} = 1; 1 \leq k \leq c \quad (3)$$

$$0 \leq \sum_{i=1}^n \mu_{ik} < c, 1 \leq i \leq n \quad (4)$$

4. Calculating the center of the k -cluster

$$v_{ij} = \frac{\sum_{k=1}^n (\mu_{ik}) x_{kj}}{\sum_{k=1}^n (\mu_{ik})^w} \pi r^2 \quad (5)$$

5. Calculates the degree of membership of each data in each cluster

$$v_{ij} = d(x_k - v_i) \left[\sum_{j=1}^c (x_{kj} - v_{ij}) \right]^{1/2} \quad (6)$$

6. Calculating the matrix change in the iteration

$$J(U, V; X) = \sum_{k=1}^n \sum_{j=1}^c (\mu_{ik})^w (d_{ik})^2 \quad (7)$$

7. Check stop condition

If: $(|P_t - P_{t-1}| < \xi)$ or $(t > \max I_{\text{ter}})$;

Then stop

If Not: $t=t+1$, repeat step d

$$X = \frac{x_1 + x_2 + x_1 + \dots + x_n}{n} = \frac{\sum_{i=1}^n x_i}{n} \quad (8)$$

X = mean, n = total data, x_i = data value i

$$\left(\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \right)^{1/2} \quad (9)$$

where; n = total data, x_i = data value i

4.2. Adaptive Neuro Fuzzy Inference System (ANFIS)

ANFIS was introduced by Jang [14], ANFIS is a method that uses artificial neural networks to implement a fuzzy inference system. FCM is used in initial initialization of ANFIS process especially in identifying fuzzy rules. A typical architecture of an ANFIS with FCM clustering is shown in Figure 1.

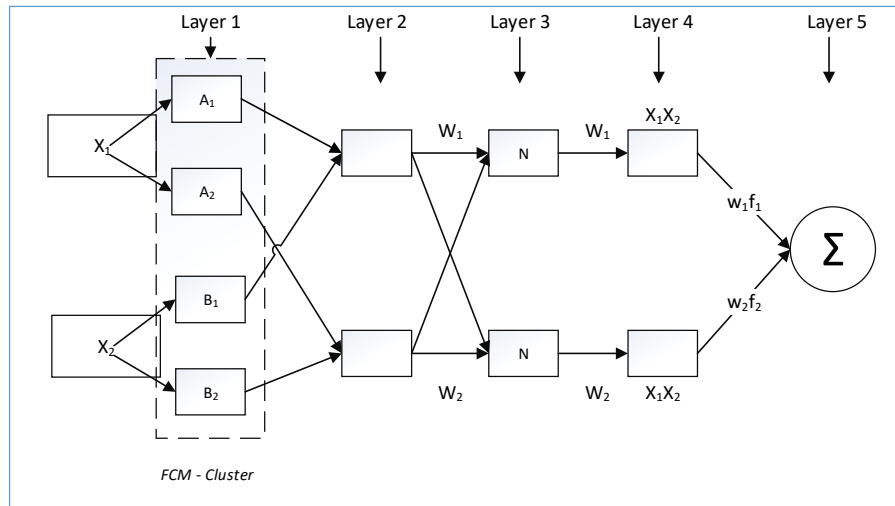


Figure 1. ANFIS with FCM architecture

- a. Layer 1: The first layer is a fuzzification layer where each neuron adaptive to the parameters of an activation. The output generated on this layer is the degree of membership assigned by the input membership function.

$$O_i^1 = \mu_{A_i}(x) \quad (10)$$

$$\mu_{A_i} = \frac{1}{1 + \left[\frac{(x - c_i)^2}{a_i} \right]} \quad (11)$$

- b. Layer 2: Every node in this layer labelled by Π is a fixed node. The output of this layer is the multiplication of all incoming signals. Each node output states the degree of activation of the fuzzy rule

$$O_i^2 = w_i = \mu_{A_i}(x) \times \mu_{B_i}(y), \quad i=1,2 \quad (12)$$

- c. Layer 3: Each node in this layer is a non-adaptive node that displays a function of activated enabling degree labelled by N.

$$O_i^3 = \varpi_i = \frac{w_i}{w_1 + w_2}, \quad i=1,2 \quad (13)$$

- d. Layer 4: In this layer the LSE calculation is done to obtain the coefficient parameter value. Every node in this layer marked by a square.

$$O_i^4 = \varpi f_i = \varpi(p_i x + q_i y + r_i), \quad i=1,2 \quad (14)$$

- e. Layer 5: This layer of a single neuron (given the symbol Σ) is the sum of all the outputs of the fourth layer

$$O_i^5 = \text{overall output} = \sum_i \varpi f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \quad (15)$$

5. RESULTS AND ANALYSSIS

In this case 10 epoch are sufficient to train and will take a level of an error to convergence.

5.1. Dataset

The data-set obtained comes from the results of e-banking monitoring during 2017 . The input data used as training are 80% of data and 20% data used as testing which is a combination of data causing the

failure conditions and normal conditions, predictions are represented in '0' for normal condition and '1' for abnormal condition as shown in Table 1.

Table 1. Data Input

Variable					Observation
Application performance	Transaction per second	Cpu utilization	Network i/o		
100	100	50	20	1	
100	150	50	15	1	
160	100	25	31	1	
170	150	30	60	1	
150	100	31	76	1	
100	150	21	0	0	

5.2. Fuzzy inference system

After the process of training data, the learning process is conducted by forming a ANFIS architecture and FIS rules as shown in Figure 2. Figure 1 demonstrates how the structure ANFIS model created FIS looks. Fuzzy C-Mean clustering process (FCM) with Sugeno structure minimizes the fuzzy if-then rules as follows:

- a. Rule 1: If (application_performance is low) and (transaction_per_second is low) and (cpu_utilization is low) and (network is low) then (target is normal)
- b. Rule 2: If (application_performance is high) and (transaction_per_second is high) and (cpu_utilization is high) and (network is high) then (target is high)

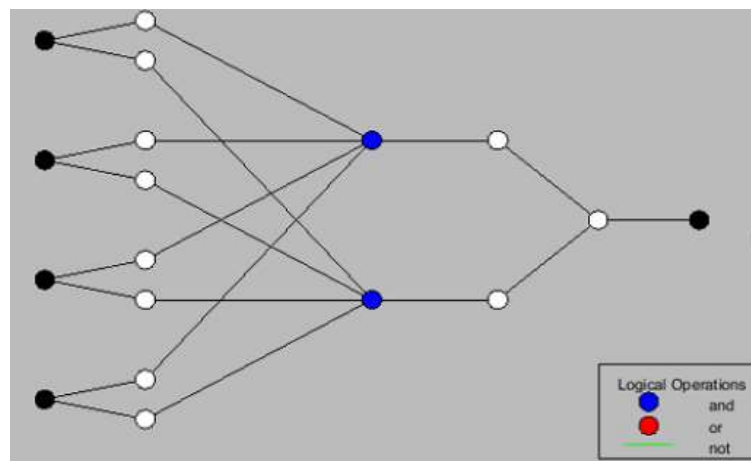


Figure 2. Structure of the proposed ANFIS

The ANFIS structure generated in Figure 1 shows the optimization performed by FCM is able to simplify the process which means that processing time becomes shorter than plain ANFIS. In the figure shows that ANFIS consists of four inputs, with each consisting of two membership Function. There is two rule with the operator 'and'. ANFIS consists of one output with MF output formed as many as two. The connecting lines between nodes represent a corresponding rule.

The membership function is generated in Figure 2 shows used is Gaussian which consists of four main variables (application performance, transactions per second, CPU usage and Network I/O) with two linguistic variables "low' and 'high'. Membership function obtained from the process of generating FIS done before. Algorithm used in Gauss is as follow:

$$G(x;k, \gamma)=e^{-k(y-x)^2} \tag{16}$$

Membership Function (MF) and plot rules of each input variable is presented in Figure 3-7:

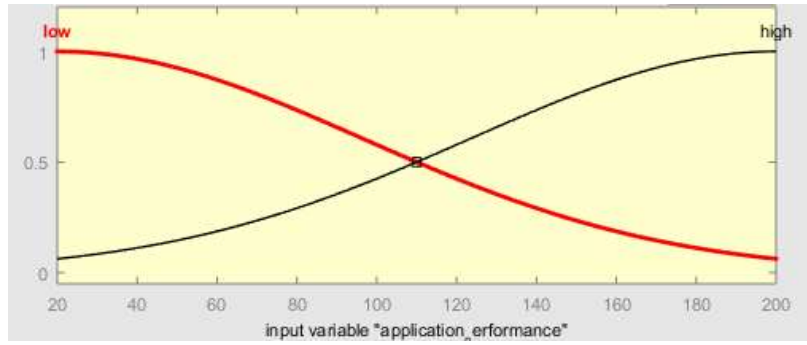


Figure 3. MF application performance

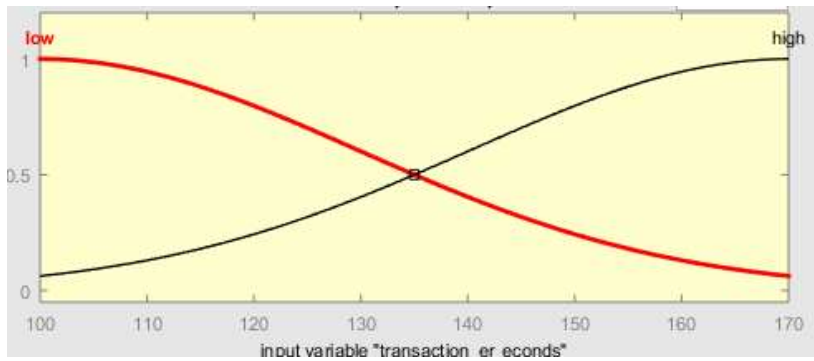


Figure 4. MF transaction per seconds

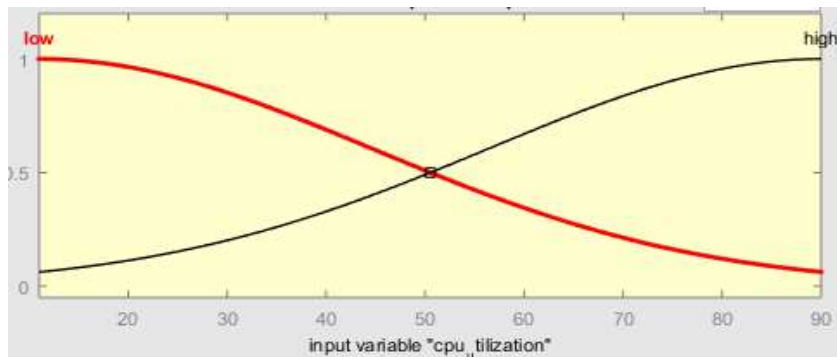


Figure 5. MF CPU utilization

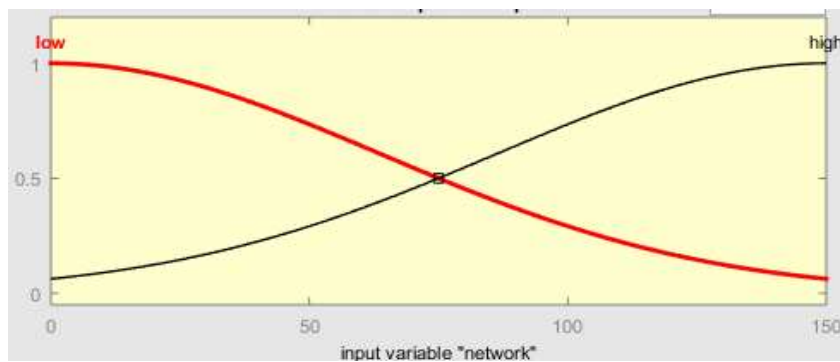


Figure 6. MF network

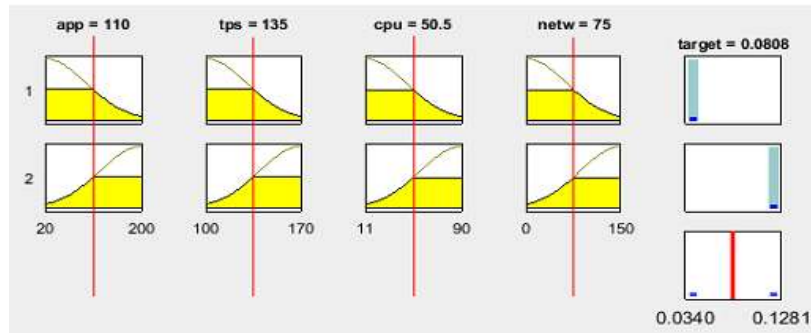


Figure 7. Plot generated rules

5.3. Comparison of the results

As can be seen in Figure 8, an average error testing FIS is 0.24462 it means having excellent accuracy over the actual data compared to the predicted data. The comparison ANFIS and ANFIS+FCM is shown in Figure 9, blue color represents the actual value and the orange color represents the predicted value, as the predicted data approaches the actual data, it means that the model accuracy is better.

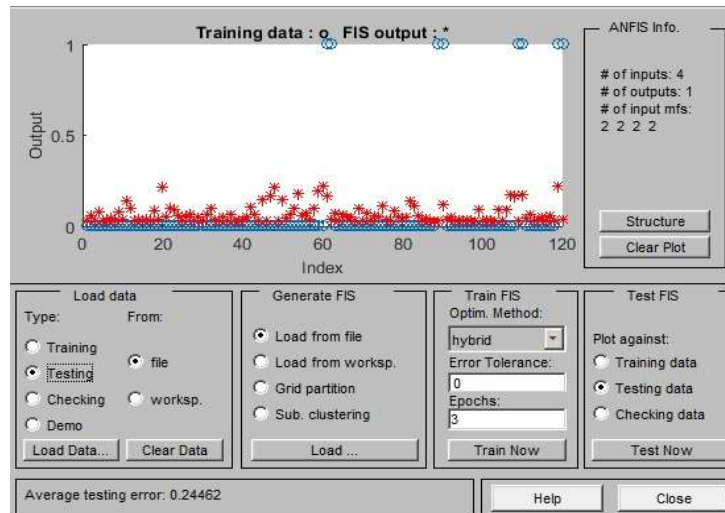


Figure 8. Error plotting of the proposed model testing result

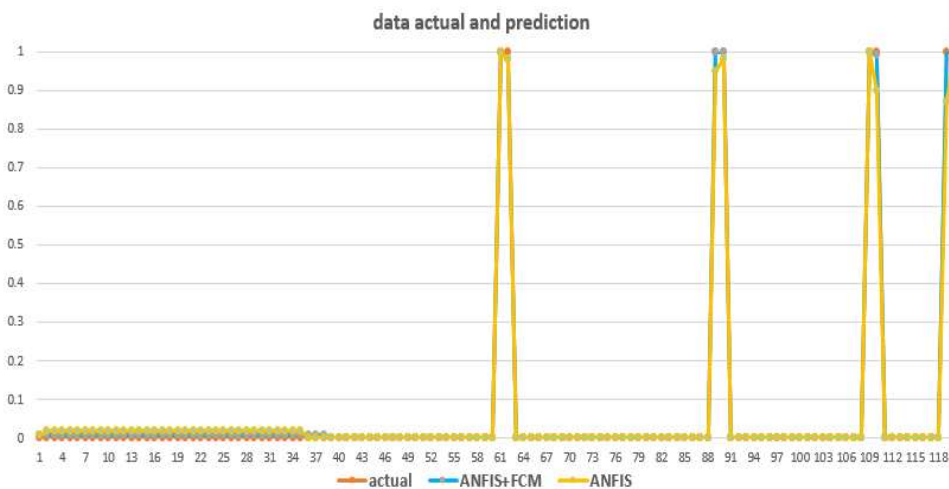


Figure 9. Comparison of prediction result

Most common metrics used to measure accuracy is MSE, RMSE and MAPE. Test method is used in determining how consistent the resulting model. In Table 2 and Figure 10 shows comparison of test results accuracy of several techniques presented using FCM optimized data compared to unoptimized data.

Table 2. Performance evaluation comparison

Total Data	MSE		RMSE		MAPE	
	ANFIS	ANFIS+FCM	ANFIS	ANFIS+FCM	ANFIS	ANFIS+FCM
120	0.00391	0.00004	0.063	0.006	0.022	0.017

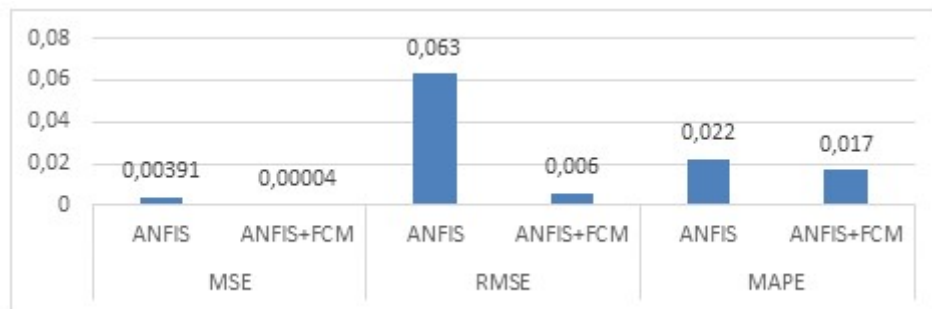


Figure 10. Performance comparison of ANFIS+FCM with ANFIS

Based on performance evaluation using MSE, obtained MSE value 000391 for ANFIS and 0.00004 for ANFIS with FCM. The result of RMSE method shows the value of 0.063 for ANFIS and 0.006 for ANFIS with FCM. MAPE is the last method used, obtained value 0.022 for ANFIS and 0.017 for ANFIS with FCM. The whole method used it can be concluded that FCM is able to improve ANFIS performance with excellent accuracy compared to ANFIS without using FCM optimization.

6. CONCLUSION

Fuzzy C-means with gaussian membership function improves ANFIS performance by optimizing input variables i.e. application response times, transaction per second, server utilization and network performance so that modeling becomes more effective. Testing with several test methods showed relatively good performance evidenced by the value of MSE 0.00004, RMSE 006 and MAPE 0.017. Added new input variables to better ANFIS process, such as: application core banking, database performance and 3rd party application performance. To increase the confidence level of the modeling can be added other testing methods such as Mean Absolute Error (MAE), Moving Range (MR) and others.

REFERENCES

- [1] J. Shing e . R. Jang, "ANFIS: Adap tive-Network-Based Fuzzy Inference System," *IEEE Transactions On Systems, Man, And Cybernetics*, Vol. 23, NO. 3, , pp. 665-685, 1993.
- [2] M. Ghodsi e M. S. Abadeh, "Fraud Detection of Credit Cards Using Neuro-fuzzy Approach Based on TLBO and PSO Algorithms," *Journal of Computer & Robotics* 10 (2),, pp. 57-68, 2017.
- [3] O. Margheni e M. Benrejeb, "Credit Scoring of Personal Loans based on Adaptive Neuro-Fuzzy Inference System and Artificial Neural Networks," *2nd International Conference on Automation, Control, Engineering and Computer Science*, pp. 54-60, 2015.
- [4] S. Sitohang, S. e A. S. Girsang, "Prediction of the Number of Airport Passengers Using Fuzzy C-Means and Adaptive Neuro Fuzzy Inference System," *International Review of Automatic Control (I.R.E.A.CO.)*, Vol. 10, N. 3, pp. 280-287, 2017.
- [5] A. Kuncoro e Suharjito, "Adaptive Neuro-Fuzzy Inference System for Texture Image Classification," *International Conference on Automation, Cognitive Science, Optics, Micro Electro-Mechanical System, and Information Technology (ICACOMIT)*, pp. 196-200, 2015.
- [6] Y. Tunardi e Suharjito, "Fuzzy Expert System for Classifying Pests and Diseases of Paddy Using Bee Colony Algorithms," *International Review on Computers and Software (I.R.E.CO.S.)*, Vol. 11, N. 5, pp. 381-387, 2016.

- [7] W. Suparta e W. S. Putro, "Comparison of Tropical Thunderstorm Estimation between Multiple Linear Regression, Dvorak and ANFIS," *Bulletin of Electrical Engineering and Informatics*, Vols. %1 de %2Vol. 6, No. 2, pp. 149-158, 2017.
- [8] H. Suyono, R. N. Hasanah, S. P. . Mudjirahardjo, A. Wijoyo e I. Musirin, "Comparison of Solar Radiation Intensity Forecasting Using ANFIS and Multiple Linear Regression Methods," *Bulletin of Electrical Engineering and Informatics*, Vols. %1 de %2Vol. 7, No. 2, pp. 191-198, 2018.
- [9] V. G e S. P, "Pixel Classification of SAR Ice Images Using ANFISPSO Classifier," *Indonesian Journal of Electrical Engineering and Informatics (IJEI)*, vol. 4, pp. 264-280, 2016.
- [10] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function*, New York: Plenum Press, 1981.
- [11] J. C Bezdek, . W. Full e R. Ehrlich, "FCM: The Fuzzy c-Means Clustering Algorithm," *Computers & Geosciences* Vol. 10, No. 2-3, pp. 191-203, 1984.
- [12] J. Yen e L. Reza, *Fuzzy Logic: Intelligence, Control, and Information*, NJ: Prentice-Hall, Inc, 1999.
- [13] Zimmermann e H. Jürgen, "Fuzzy Set Theory and Its Applications," New York: Springer Science Business Media, 2001.
- [14] Jang, "ANFIS: Adaptive-network-based fuzzy inference system.," *IEEE Transactions on Systems, Man, and Cybernetic* 23, 1993.

BIOGRAPHIES OF AUTHORS



Yuwono Abdillah is an employee of one of the national banking institutions in Indonesia. He received bachelor degree in mathematics from the Sepuluh Nopember Institut of Technology , Surabaya, Indonesia in 2003. His main areas of interest are artificial intelligence and IT infrastructure.



Suharjito is the Head of Information Technology Department in Binus Graduate Program of Bina Nusantara University. He received under graduated degree in mathematics from The Faculty of Mathematics and Natural Science in Gadjah Mada University, Yogyakarta, Indonesia in 1994. He received master degree in information technology engineering from Sepuluh November Institute of Technology, Surabaya, Indonesia in 2000. He received the PhD degree in system engineering from the Bogor Agricultural University (IPB), Bogor, Indonesia in 2011. His research interests are intelligent system, Fuzzy system, image processing and software engineering