



Contents lists available at DJM

DESIMAL: JURNAL MATEMATIKA

p-ISSN: 2613-9073 (print), e-ISSN: 2613-9081 (online), DOI 10.24042/djm
<http://ejournal.radenintan.ac.id/index.php/desimal/index>



Forecasting Indonesian inflation using a hybrid ARIMA-ANFIS

Nina Fitriyati*, Mahmudi, Madona Yunita Wijaya, Maysun

UIN Syarif Hidayatullah Jakarta, Indonesia

ARTICLE INFO

Article History

Received : 19-10-2022
Revised : 31-10-2022
Accepted : 14-11-2022
Published : 20-12-2022

Keywords:

Clusters; Gaussian Membership Function; Fuzzy C Means; Linear Components; Non-Linear Components.

*Correspondence: E-mail:
nina.fitriyati@uinjkt.ac.id

Doi:
[10.24042/djm.v5i3.14093](https://doi.org/10.24042/djm.v5i3.14093)

ABSTRACT

This paper discusses the prediction of the inflation rate in Indonesia. The data used in this research is assumed to have both linear and non-linear components. The ARIMA model is selected to accommodate the linear component, while the ANFIS method accounts for the non-linear component in the inflation data. Thus, the model is known as the hybrid ARIMA-ANFIS model. The clustering method is performed in the ANFIS model using Fuzzy C-Mean (FMS) with a Gaussian membership function. Consider 2 to 6 clusters. The optimal number of clusters is assessed according to the minimum value of the error prediction. To evaluate the performance of the fitted hybrid ARIMA-ANFIS model, it can be compared to the classical ARIMA model and with the ordinary ANFIS model. The result reveals that the best ARIMA model for inflation prediction in Indonesia is ARIMA(2,1,0). In the hybrid ARIMA(2,1,0)-ANFIS model, two clusters are optimal. Meanwhile, the optimum number of clusters in the ordinary ANFIS model is six. The comparison of prediction accuracy confirms that the hybrid model is superior to the individual model alone of either ARIMA or ANFIS model.

<http://ejournal.radenintan.ac.id/index.php/desimal/index>

INTRODUCTION

The Central Bureau of Statistics (BPS) reported that the inflation rate in Indonesia in 2020 was about 1.68%, which was recorded as the lowest annual rate since BPS released inflation rate data (Fauzia, 2021). In addition, BPS also recorded that the COVID-19 pandemic has disrupted Indonesia's inflation rate pattern. The spread of COVID-19 also led to a decrease in demand and circulation of money stemming from a decline in economic activity resulting from a surge in layoffs and changes in the working scheme

to "Work from Home" (WFH). This low inflation rate can be interpreted as economic sluggishness because of economic actors' difficulty in raising prices to maintain demand. The decline in the inflation rate at the end of 2021 can be seen in Figure 1 (Bank Indonesia, 2022). The inflation rate has been unstable for the past ten years. The unstable inflation rate harms the social, economic, and mental conditions of the community. If not evaluated and appropriately handled, this harmful condition can influence the quality of human resources (Fitriyati &

Wijaya, 2022). Therefore, an excellent method to predict the inflation rate in the future is necessary so that the government

can take strategic policies to control the inflation rate.

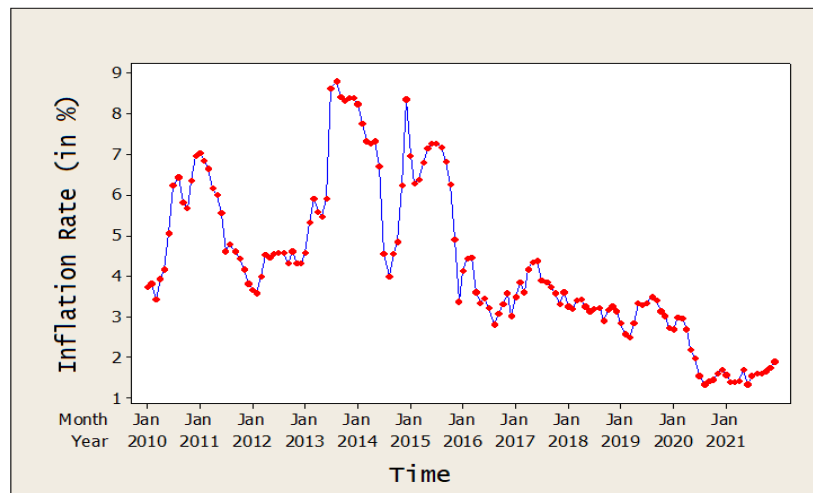


Figure 1. Inflation Rate in Indonesia from January 2010 to January 2021 (Bank Indonesia, 2022)

Several methods have been developed to predict the inflation rate in several countries. Wen et al. (2014) predicted the inflation rate by predicting the consumer price index (CPI) using the Back Propagation (BP) neural network model. The result revealed that the selected model could predict the CPI at least six months ahead and was an excellent model for forecasting inflation. Charemza et al. (2019) made an inflation prediction using the Copula approach, and Enke & Mehdiyev (2014) made an inflation prediction using the hybrid neuro-fuzzy approach.

Zhang (2019) built an inflation prediction model in the United States using real-time macroeconomic variables and combination forecasts of time-varying weights and equal weights. The combination forecasts compare three sets of commonly used time-varying coefficient autoregressive models: Gaussian distributed errors, errors with stochastic volatility, and errors with moving average stochastic volatility. John et al. (2020) empirically examined the performance of the combined approach to inflation forecasting on the individual

model, the benchmark random walk model, and median/mean inflation forecasts from the Survey of Professional Forecasters (SPF) conducted by the Reserve Bank of India (2017). It was found that the combined approach resulted in better accuracy as compared to individual models, such as random walk (RW), autoregressive (AR), moving average with stochastic volatility (MA-SV), vector autoregression (VAR), Bayesian VAR, VAR and BVAR with exogenous variables (VAR-X and BVAR-X, respectively), and Phillips curves (PC). Işığçok et al. (2020) forecasted the inflation rate in Turkey using the combination of Box-Jenkins (ARIMA) Models and the Artificial Neural Network.

In our previous research, the inflation rate was modeled through an Integrated Moving Average (ARIMA) model. The ARIMA model assumed linearity. However, real-world data rarely has a linear structure (Barak & Sadegh, 2016). Therefore, to account for the non-linear structure in the inflation data, the *Adaptive Neuro-Fuzzy Inference System* (ANFIS) and the ARIMA model are combined. This model is known as a

hybrid ARIMA-ANFIS model. For the research, monthly inflation rate data in Indonesia was collected from January 2010 to January 2021, as depicted in Figure 1. The data is divided into two parts, i.e., 92% as training data and 8% as testing data. The clustering method used in the hybrid ARIMA-ANFIS model is Fuzzy C-Means (FCM) with 2 - 6 clusters. Different clusters are explored to find the optimal cluster and minimize the prediction error. Furthermore, the resulting inflation rate prediction from the hybrid model is compared to the classical ARIMA model and the ordinary ANFIS model.

METHOD

The data used in this research is monthly inflation rate data from January 2010 to December 2021 obtained from www.bi.go.id (Bank Indonesia, 2022). The data consists of 144 samples divided into training and testing sets with a composition of 92% and 8%, respectively. This composition was taken based on inflation predictions, which obtained the best accuracy in several experiments. The training set is used to build the hybrid ARIMA-ANFIS model, and the testing set is used to assess the accuracy of the inflation prediction.

The inflation rate depicted in Figure 1 indicates non-stationary behavior in both mean and variance. Thus, Box-Cox transformations and differencing are applied to the data to overcome these non-stationary issues. An Augmented Dickey-Fuller (ADF) test is further conducted as a formal test to check whether the time series data is stationary or not. The data can be modeled using the ARIMA method if the stationary assumption is met.

ARIMA Model Building

The Autoregressive-Moving Average (ARMA) model is a stationary model, which can be expressed as follows:

$$Z_t = \phi_1 Z_{t-1} + \dots + \phi_p Z_{t-p} + e_t - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q} \quad (1)$$

or equivalently can be written as

$$\phi_p(B)Z_t = \theta_q(B)e_t, \quad (2)$$

where $\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$, $\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$, B is a backshift operator, and e_t is the model residuals. However, if the data is not stationary, then differencing is required. The time series data after differencing d times can be expressed as an Integrated Autoregressive Moving Average or ARIMA(p, d, q) model:

$$\phi_p(B)(1 - B)^d Z_t = \theta_q(B)e_t. \quad (3)$$

where $(1 - B)^d$ is a differencing function of order d . The parameters in the ARIMA model can be estimated using the Maximum Likelihood Estimation (MLE) method. The best ARIMA model is selected with the lowest AIC (Akaike Criteria Information) value. Afterward, diagnostic testing for the selected model is carried out to evaluate whether the model assumptions are satisfied or not. Diagnostic testing includes testing the normality and autocorrelation of the model residuals.

The normality of the residuals is evaluated using the Kolmogorov-Smirnov test (Massey, 1951) with the following hypothesis:

$H_0: F(x) = F_0(x)$ (the residual is normally distributed),

$H_1: F(x) \neq F_0(x)$ (the residual is not normally distributed).

The test statistic is $D = \max_x |F_0(x) - S_N(x)|$. The null hypothesis is rejected if $D > D_{tabel}$ or $p\text{-value} < \alpha$, where α is the significance level.

To test whether autocorrelation is still present in the residuals, the Ljung-Box

test is carried out with the following hypothesis (Hillmer & Wei, 1991):

$$H_0: \rho_1 = \rho_2 = \dots = \rho_K = 0$$

$$H_1: \text{at least one } \rho_i \neq 0, i = 1, 2, \dots, K.$$

The test statistic:

$$Q^* = n(n + 2) \sum_{k=1}^K \frac{\hat{r}_k^2}{n-k}, \quad (4)$$

where K is the maximum number of lag, r_k is the estimated autocorrelation at lag k . The test statistic Q^* follows a Chi-square distribution. Thus, H_0 is rejected if $Q^* > \chi_{table}^2$ or $p\text{-value} < \alpha$. The best ARIMA model will be selected by the smallest Akaike's Information Criterion (AIC) values (Cryer & Chan, 2008).

Hybrid ARIMA-ANFIS Model Building

After obtaining the best-fit ARIMA model, the combined hybrid ARIMA-ANFIS model can be built. The first step is to determine the number of clusters using the FCM method. FCM is a data clustering approach where every data point in a cluster is determined by the degrees of membership (Benmouiza & Cheknane, 2019). The fundamental concept of FCM is to determine the cluster's center in order to identify each cluster's average location. In the initial condition, the cluster center is not yet accurate. By improving the cluster center and the degree of membership of each data point repeatedly, the cluster center will move to the correct location (Tarno et al., 2013). The clustering process using the FCM method is used to obtain the mean and standard deviation values. Next, these two values become the premise parameters in the ANFIS modeling.

The ANFIS model consists of one input layer, three hidden layers, and one output layer, where a square symbol indicates an adaptive layer and a circle symbol indicates a fixed layer. Figure 2 illustrates the ANFIS model's architecture.

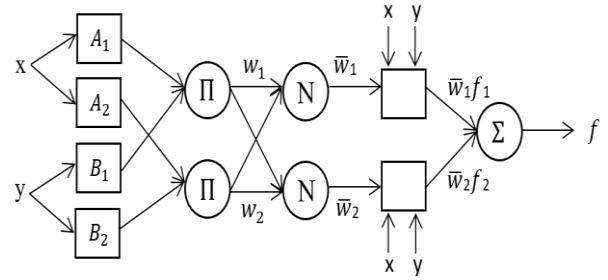


Figure 2. The Architecture of the ANFIS Model

Each layer in Figure 2 is described as follows:

Layer 1 (Fuzzification)

Fuzzification is the inversion of a set of crisp values into a set of fuzzy values. Every i neuron in the first layer is adaptive and has a function:

$$O_i^1 = \mu_{A_i}(x) \quad (5)$$

where x is the i neuron input and A_i is the linguistic label that corresponds to the function of this neuron, and O_i^1 is the membership function of A_i . For example, a membership function that can be used is a Gaussian function, which is written as follows:

$$\mu_{A_i}(x) = e^{-\frac{1}{2} \left(\frac{x-\mu}{\sigma} \right)^2}, \quad (6)$$

where μ and σ are the premise parameters.

Layer 2 (Fuzzy Logic Operation)

Each neuron in the second layer is fixed, denoted by a label Π , where the output is the product of all inputs, which can be expressed as follows:

$$O_i^2 = w_i = \mu_{A_i}(x) \times \mu_{B_i}(y), i = 1, 2. \quad (7)$$

In general, the fuzzy operation "AND" is used. The output for each neuron describes the weight for every rule.

Layer 3 (Normalized Firing Strength)

Each neuron in the third layer is fixed and denoted by N. The process includes normalizing the weight function

obtained from the second layer, which can be expressed as follows:

$$O_i^3 = \bar{w}_i = \frac{w_i}{w_1 + w_2}, i = 1, 2. \quad (8)$$

Layer 4 (Defuzzification)

Each neuron in the fourth layer is adaptive with the following function:

$$O_i^4 = \bar{w}_i f_i = \bar{w}_i (p_i x + q_i y + r_i), \quad (9)$$

where w_i is the normalization output from the third layer; and p_i, q_i , and r_i are the consequent parameters.

Layer 5 (Output Operation)

A single neuron in the fifth layer is fixed and denoted by Σ . It is the output from all layers, which can be expressed as follows:

$$O_i^5 = \Sigma_i \bar{w}_i f_i = \frac{\Sigma_i w_i f_i}{\Sigma_i w_i}. \quad (10)$$

The workflow of ANFIS utilizes a hybrid learning algorithm. The hybrid algorithm will use the backpropagation gradient descent method to govern the consequent parameters with forward pass calculation using the Least Square Estimator (LSE) and the premise parameters with backward pass calculation (Zhang, 2019). Each process of a single forward pass and a single backward pass is counted as one epoch.

The input for the hybrid ARIMA-ANFIS model is the significant AR lag from the best-fit ARIMA model residuals. A hybrid model is a combination of one or more models in a system of functions. Zhang (2019) stated that a general time series model consists of a combination of linear and non-linear components, which can be written as follows:

$$Z_t = L_t + N_t, \quad (11)$$

where Z_t is the actual data, L_t is the linear component, and N_t is the non-linear component. The linear component can be resolved through an ARIMA model. The residual of the ARIMA model contains a

non-linear component (Zhang, 2019), which can be expressed as $e_t = Z_t - \hat{L}_t$ where \hat{L}_t is the prediction value. Next, the residual of the ARIMA model is modeled using the ANFIS method. The prediction results from the ANFIS model are accumulated with the prediction results from the ARIMA model. Thus the formula for total prediction from hybrid ARIMA-ANFIS can be written as follows:

$$\hat{Z}_t = \hat{L}_t + \hat{N}_t. \quad (12)$$

where \hat{N}_t is the prediction from the non-linear component. The prediction accuracy is computed using the Mean Absolute Percentage Error (MAPE) (Hillmer & Wei, 1991).

The hybrid ARIMA-ANFIS model is a method that combines a time series approach with a neural network model. This combination has been applied in many areas, including rainfall prediction (Faulina & Suhartono, 2013), energy consumption prediction (Barak & Sadegh, 2016), and modeling red onion production (Diarsih et al., 2019). The research flowchart is shown in Figure 3.

RESULTS AND DISCUSSION

The Box-Cox transformation indicates that no transformation is required for the inflation data, but the ADF test reveals that the series is not stationary in the mean. After the first difference is applied to the data, the series becomes stationary in the mean, as displayed in Figure 4. The second step in ARIMA modeling is model identification through the ACF and PACF plots (see Figure 5).

According to Figure 5, the ACF plot cuts off after lag 1, and the PACF plot cuts off after lags 1 and 2. Therefore, a number of potential models are selected based on ACF and PACF plots, such as ARIMA(1,1,0), ARIMA(0,1,1), ARIMA(1,1,1), ARIMA(2,1,0), and ARIMA(2,1,1). These ARIMA models follow equation (1). The

estimated parameters from the selected models are summarized in Table 1. The table shows that at a 5% significance level, there are three models that have significant parameters, i.e., ARIMA(1,1,0), ARIMA(0,1,1), and ARIMA(2,1,0), since the p-values are lower than 5%. Diagnostic

test results for the model residuals are presented in Table 2. At a 5% significance level, all three models satisfy the assumption of normality since their p-values are greater than 5%. The residuals also do not exhibit autocorrelation since p-values are greater than 5%.

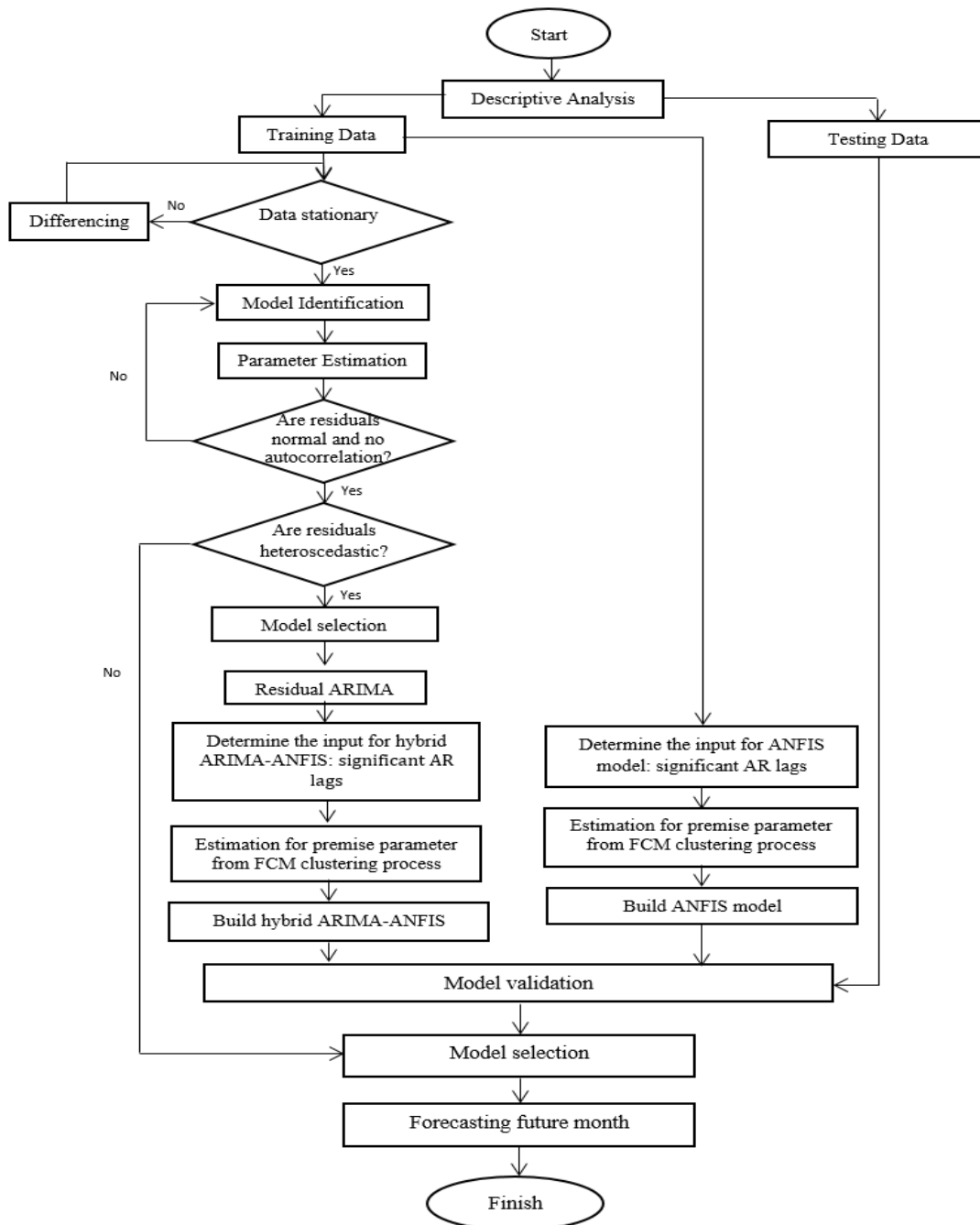


Figure 3. The Research Flowchart

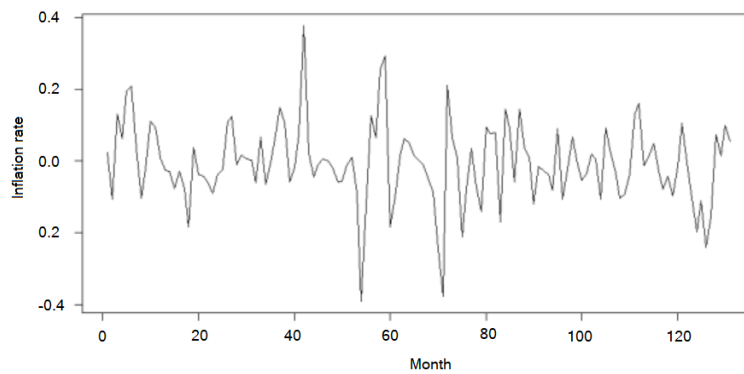


Figure 4. A Plot of Stationary Inflation Data after Applying the First Difference

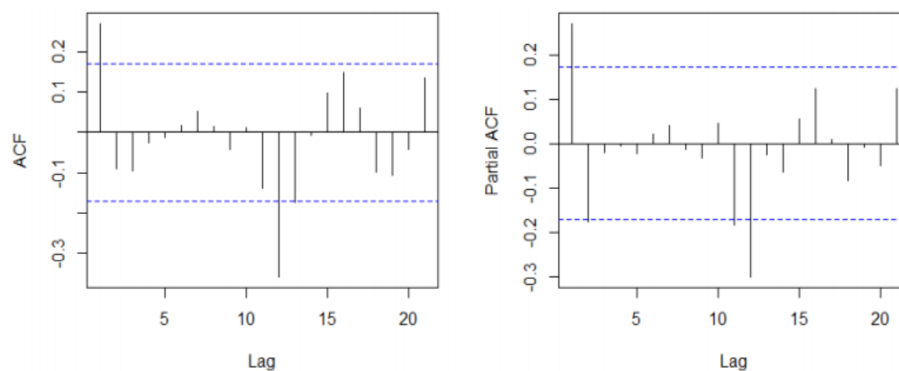


Figure 5. A Plot of ACF and PACF of Stationary Inflation Data

Table 1. The Estimated Parameters of ARIMA Models

No	ARIMA (p, d, q)	Estimate	Standard Error	p-value	Conclusions
1.	(1,1,0)	ϕ_1 0.2718	0.08374	0.001	Significant
2.	(0,1,1)	θ_1 -0.3371	0.07820	0.000	Significant
3.	(1,1,1)	ϕ_1 -0.1210	0.22740	0.595	Not Significant
		θ_1 -0.4686	0.20199	0.022	Significant
4.	(2,1,0)	ϕ_1 0.3351	0.0823	0.000	Significant
		ϕ_2 -0.2283	0.0823	0.006	Significant
5.	(2,1,1)	ϕ_1 0.3415	0.3615	0.347	Not Significant
		ϕ_2 -0.2301	0.1263	0.071	Not Significant
		θ_1 0.0067	0.3715	0.796	Not Significant

Table 2. Diagnostic Test Results

Model Assumption	p-value		
	ARIMA (1,1,0)	ARIMA (0,1,1)	ARIMA (2,1,0)
Normality	0.1702	0.07379	0.07133
No Autocorrelation	0.5821	0.88350	0.96650

Table 3 presents AIC values for the selected candidate models. The best-fitted model is ARIMA(2,1,0) since it has the lowest AIC value. The estimated model can be expressed as follows:

$$\ln Z_t = 1,31921 \ln Z_{t-1} - 0,49231 \ln Z_{t-2} + 0,1731 \ln Z_t + e_t,$$

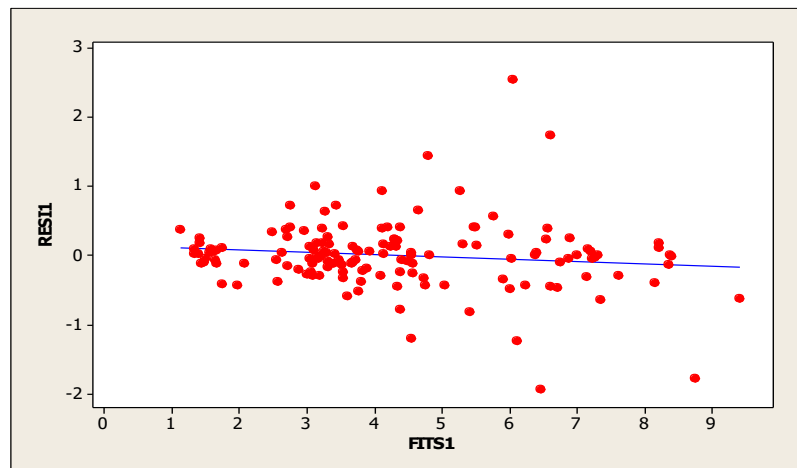
where Z_t is the inflation rate at time t .

Table 3. AIC Value for ARIMA Models

No.	Model	AIC
1.	ARIMA(1,1,0)	-210.9
2.	ARIMA (0,1,1)	-212.12
3.	ARIMA (2,1,0)	-212.89

Model ARIMA(2,1,0) is a linear model. However, the inflation data contains a non-linear component that cannot be accounted for by the ARIMA model. The non-linear component can be

observed through a scatterplot between fitted values and residuals of ARIMA(2,1,0) (Nau, 2014), as shown in Figure 6. The plot shows that the residuals increase as the fitted values increase or as the spread of points varies around the regression line. This indicates non-constant variance, and therefore the non-linear model is necessary. A method to handle non-linear structures is the ANFIS model.

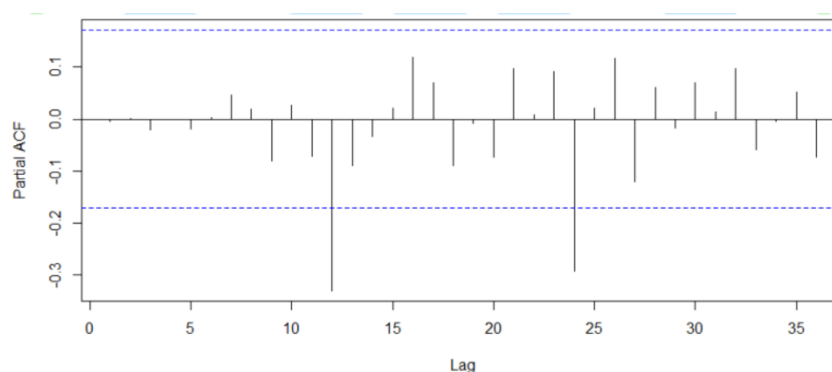
**Figure 6.** Scatterplot between Fitted Values and Residuals of ARIMA(2,1,0) Model

Prediction with Hybrid ARIMA(2,1,0)-ANFIS Model

The first step for hybrid ARIMA(2,1,0)-ANFIS model building is to determine the input, i.e. significant AR lags. Significant lag can be identified through the PACF plot from the residuals of the ARIMA(2,1,0) model, as depicted in Figure 7. It shows that significant lags

appear at lags 12 and 24, and these two lags are the inputs in the hybrid ARIMA-ANFIS model.

Next, clustering is performed using FCM with 2, 3, 4, 5, and 6 clusters. A Gaussian function is used as the membership function. In the FCM method, the number of rules is equal to the specified number of clusters, so there is no combination in the formation of rules.

**Figure 7.** PACF Plot for Residual of ARIMA(2,1,0) Model

The computation of the ANFIS model on the residuals is carried out by calculating each layer using equations (5)–(10). This research uses zero error tolerance and ten epochs in the calculation. Afterward, the prediction result is accumulated with the prediction result from ARIMA(2,1,0) using equation

(12). The architecture of the ANFIS model on the residual series with two inputs and two membership functions is displayed in Figure 8, which consists of five layers, i.e., fuzzification, fuzzy logic operation, normalized firing strength, defuzzification, and output calculation.

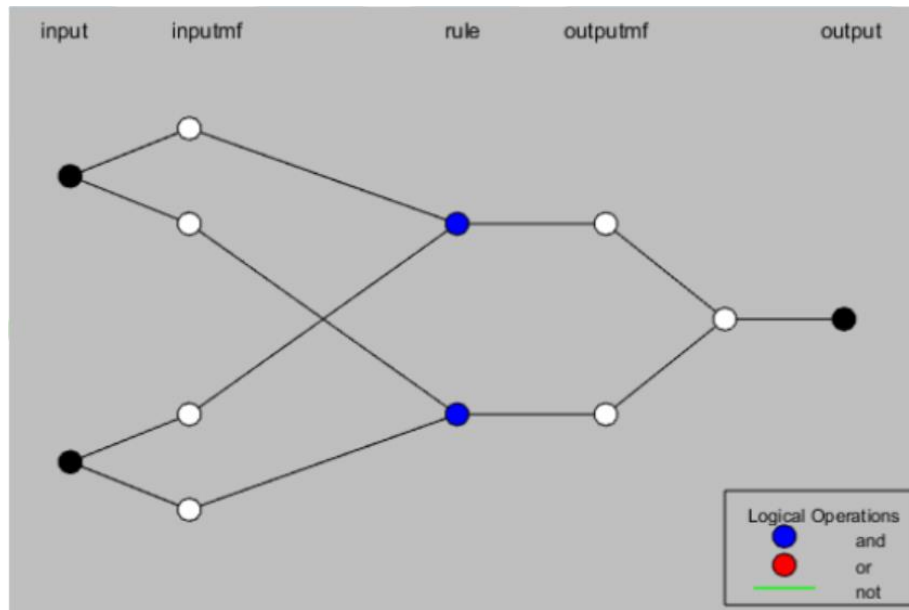


Figure 8. The Architecture of the ANFIS Model with the Two-Cluster Solution

In the first layer (fuzzification), the estimated mean (μ) and standard deviation (σ) obtained from the FCM clustering process are summarized in Table 4. These values are used as premise parameters, which are then used to make the membership functions. Mathematically, the membership function for each input can be written as follows:

$$\mu_{A_1} = \exp \left\{ -\frac{1}{2} \left(\frac{e_{t-24} + 0.03226}{0.0589} \right)^2 \right\},$$

$$\mu_{A_2} = \exp \left\{ -\frac{1}{2} \left(\frac{e_{t-24} - 0.04135}{0.1031} \right)^2 \right\},$$

$$\mu_{B_1} = \exp \left\{ -\frac{1}{2} \left(\frac{e_{t-12} - 0.02992}{0.08958} \right)^2 \right\},$$

$$\mu_{B_2} = \exp \left\{ -\frac{1}{2} \left(\frac{e_{t-12} + 0.04427}{0.05019} \right)^2 \right\}.$$

Table 4. Premise Parameters of ANFIS Model with the Two-Cluster Solutions for Residuals of ARIMA(2,1,0) Model

Cluster	e_{t-24}	e_{t-12}
1	$\mu_{11} = -0.03226$	$\mu_{21} = 0.02992$
	$\sigma_{11} = 0.0589$	$\sigma_{21} = 0.08958$
2	$\mu_{12} = 0.04135$	$\mu_{22} = -0.04427$
	$\sigma_{12} = 0.1031$	$\sigma_{22} = 0.05019$

In the second layer, the multiplication operation is applied to all inputs. In the FCM method, the number of rules equals the number of clusters. Since a two-cluster solution is used, two rules are applied and are written as follows:
 Rule 1 : If (e_{t-24} is A_1) and (e_{t-12} is B_1) then
 $(f_1 = p_1 e_{t-24} + q_1 e_{t-12} + r_1)$.
 Rule 2 : If (e_{t-24} is A_2) and (e_{t-12} is B_2) then
 $(f_2 = p_2 e_{t-24} + q_2 e_{t-12} + r_2)$.

The output for layer two is denoted by w_i ($i = 1, 2$). Next, the process of normalizing firing strength proceeded in layer 3. The output in layer 3 is the ratio from the i^{th} rule to the overall w_i denoted by \bar{w}_i , where the number of outputs equals the number of outputs in layer 2.

The fourth layer (defuzzification process) yields consequent parameters from the LSE learning, as shown in Table 5. The consequent parameter can be expressed in a linear equation (9) for each rule as follows:

$$f_1 = -1.069 e_{t-24} - 0.8086 e_{t-12} - 0.01161,$$

$$f_2 = 0.0278 e_{t-24} + 0.04587 e_{t-12} - 0.001012.$$

Table 5. Consequent Parameter of ANFIS Model with the Two-Cluster Solution for Residual of ARIMA(2,1,0) Model

Rule (i)	p_i	q_i	r_i
1	-1.069	-0.8086	-0.01161
2	0.0278	0.04587	-0.001012

The last step in layer five is obtaining the output values of the ANFIS network. The ANFIS model for residuals follows equation (10):

$$\hat{f} = \frac{w_1}{w_1+w_2} (-1.069 e_{t-24} - 0.8086 e_{t-12} - 0.01161) + \frac{w_2}{w_1+w_2} (0.0278 e_{t-24} + 0.04587 e_{t-12} - 0.001012),$$

where

$$w_1 = \exp \left\{ -\frac{1}{2} \left(\frac{e_{t-24} + 0.03226}{0.0589} \right)^2 \right\} \times \exp \left\{ -\frac{1}{2} \left(\frac{e_{t-12} - 0.02992}{0.08958} \right)^2 \right\},$$

$$w_2 = \exp \left\{ -\frac{1}{2} \left(\frac{e_{t-24} - 0.04135}{0.1031} \right)^2 \right\} \times \exp \left\{ -\frac{1}{2} \left(\frac{e_{t-12} + 0.04427}{0.05019} \right)^2 \right\}.$$

Next, the fitted hybrid ARIMA(2,1,0)-ANFIS model based on equation (12) is:

$$\hat{Z}_t = \ln \hat{L}_t + \hat{N}_t,$$

where

$$\ln \hat{L}_t = 1.31921 \ln Z_{t-1} - 0.49231 \ln Z_{t-2} + 0.1731 \ln Z_t + e_t,$$

$$\hat{N}_t = \frac{w_1}{w_1+w_2} (-1.069 e_{t-24} - 0.8086 e_{t-12} - 0.01161) + \frac{w_2}{w_1+w_2} (0.0278 e_{t-24} + 0.04587 e_{t-12} - 0.001012).$$

Similar hybrid ARIMA(2,1,0)-ANFIS model can be obtained for a different number of clusters. These models are then used to predict the inflation rate according to equation (12). To assess model accuracy, the predicted inflation rate is compared to the actual series in the testing set using MAPE. The MAPE values for the hybrid ARIMA(2,1,0)-ANFIS model with a different number of clusters are presented in Table 6. It can be seen that the lowest MAPE value is achieved when considering two clusters. Thus, the best-fit model is a hybrid ARIMA(2,1,0)-ANFIS with a two-cluster solution.

Table 6. Accuracy of Inflation Prediction using Hybrid ARIMA(2,1,0)-ANFIS Model with 2, 3, 4, 5, and 6 Clusters

No.	Number of Clusters	MAPE (%)
1.	2	10.27
2.	3	10.65
3.	4	10.55
4.	5	11.14
5.	6	11.45

Prediction with Individual ANFIS Model

As opposed to the hybrid ARIMA(2,1,0)-ANFIS model where the input is a significant lag in the model residuals, the input in the individual ANFIS model is obtained from significant lag in the actual series, i.e., Z_{t-1} and Z_{t-2} . A similar number of clusters are used in this model. In the FCM method, the number of rules is equal to the specified number of clusters, so there is no combination in the formation of rules. In addition, the same

Gaussian function is applied as the membership function.

The computation of the ANFIS model is carried out by calculating each layer using equations (5)–(10). In this case, 0 error tolerance and 200 epochs are considered. Without loss of generality, six clusters are discussed in this section. The architecture of the ANFIS model with two inputs and six membership functions is displayed in Figure 9, which consists of five layers, i.e., fuzzification, fuzzy logic operation, normalized firing strength, defuzzification, and output calculation. The membership function for each input can be written as follows:

$$\mu_{A_1} = \exp \left\{ -\frac{1}{2} \left(\frac{Z_{t-2} - 4.365}{0.9754} \right)^2 \right\}$$

$$\mu_{A_2} = \exp \left\{ -\frac{1}{2} \left(\frac{Z_{t-2} - 3.164}{1.165} \right)^2 \right\},$$

$$\mu_{A_3} = \exp \left\{ -\frac{1}{2} \left(\frac{Z_{t-2} - 6.529}{0.3283} \right)^2 \right\},$$

$$\mu_{A_4} = \exp \left\{ -\frac{1}{2} \left(\frac{Z_{t-2} - 8.242}{1.238} \right)^2 \right\},$$

$$\mu_{A_5} = \exp \left\{ -\frac{1}{2} \left(\frac{Z_{t-2} - 1,985}{1.286} \right)^2 \right\},$$

$$\mu_{A_6} = \exp \left\{ -\frac{1}{2} \left(\frac{Z_{t-2} - 5.49}{1.745} \right)^2 \right\},$$

$$\mu_{B_1} = \exp \left\{ -\frac{1}{2} \left(\frac{Z_{t-1} - 3.806}{1.128} \right)^2 \right\},$$

$$\mu_{B_2} = \exp \left\{ -\frac{1}{2} \left(\frac{Z_{t-1} - 3.238}{0.8479} \right)^2 \right\},$$

$$\mu_{B_3} = \exp \left\{ -\frac{1}{2} \left(\frac{Z_{t-1} - 6.977}{0.6233} \right)^2 \right\},$$

$$\mu_{B_4} = \exp \left\{ -\frac{1}{2} \left(\frac{Z_{t-1} - 8.131}{1.385} \right)^2 \right\},$$

$$\mu_{B_5} = \exp \left\{ -\frac{1}{2} \left(\frac{Z_{t-1} - 1,599}{0,860} \right)^2 \right\},$$

$$\mu_{B_6} = \exp \left\{ -\frac{1}{2} \left(\frac{Z_{t-1} - 5.137}{0.9701} \right)^2 \right\}.$$

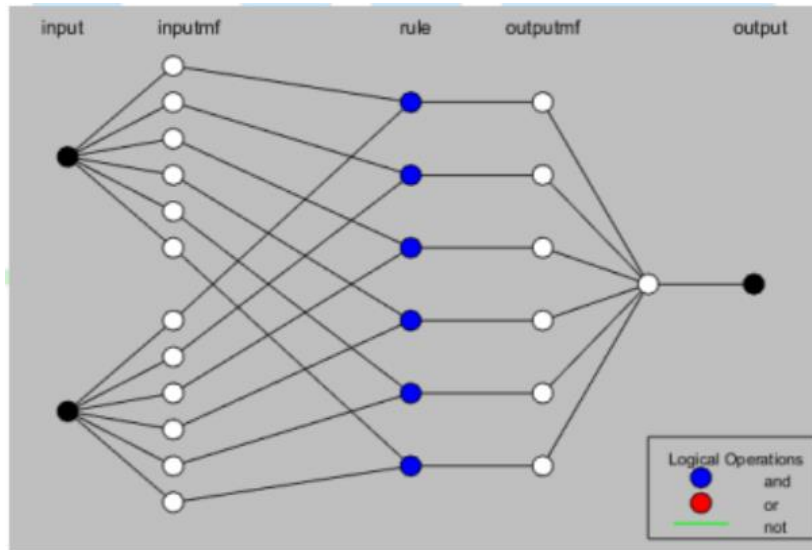


Figure 9. The Architecture of the ANFIS Model with Six Clusters

In the second layer, six rules are applied since six clusters are considered. They are:

Rule 1 : If (Z_{t-2} is A_1) and (Z_{t-1} is B_1) then
 $(f_1 = p_1 Z_{t-2} + q_1 Z_{t-1} + r_1)$

Rule 2 : If (Z_{t-2} is A_2) and (Z_{t-1} is B_2) then
 $(f_2 = p_2 Z_{t-2} + q_2 Z_{t-1} + r_2)$

Rule 3 : If (Z_{t-2} is A_3) and (Z_{t-1} is B_3) then
 $(f_3 = p_3 Z_{t-2} + q_3 Z_{t-1} + r_3)$

Rule 4 : If (Z_{t-2} is A_4) and (Z_{t-1} is B_4) then
 $(f_4 = p_4 Z_{t-2} + q_4 Z_{t-1} + r_4)$

Rule 5 : If (Z_{t-2} is A_5) and (Z_{t-1} is B_5) then
 $(f_5 = p_5 Z_{t-2} + q_5 Z_{t-1} + r_5)$

Rule 6 : If (Z_{t-2} is A_6) and (Z_{t-1} is B_6) then
 $(f_6 = p_6 Z_{t-2} + q_6 Z_{t-1} + r_6)$

The output for layer 2 is w_i ($i = 1, 2$). Next, the process of normalizing firing strength proceeded in layer 3. Layer 3 output is the ratio of the i^{th} rule to the overall w_i , denoted by \bar{w}_i , where the

number of outputs equals the number of outputs in layer 2.

The fourth layer (defuzzification process) yields a consequent parameter from the LSE learning. The consequent parameter can be expressed in a linear equation for each rule as follows:

$$\begin{aligned} f_1 &= -3.079Z_{t-2} - 19.98Z_{t-1} + 98.74 \\ f_2 &= -12.55Z_{t-2} + 1.835Z_{t-1} + 26.84 \\ f_3 &= 3.664Z_{t-2} - 5.022Z_{t-1} + 16.87 \\ f_4 &= -0.3716Z_{t-2} + 0.6656Z_{t-1} + 5.791 \\ f_5 &= 1.498Z_{t-2} - 2.672Z_{t-1} + 3.001 \\ f_6 &= -7.531Z_{t-2} - 1.774Z_{t-1} + 63.88 \end{aligned}$$

The last step in layer five is obtaining the output values of the ANFIS network. The ANFIS model for residuals can be generally used using equation (10), as follows:

$$\begin{aligned} \hat{f} &= \frac{w_1}{w_1 + w_2 + w_3 + w_4 + w_5 + w_6} (-3.079Z_{t-2} - 19.98Z_{t-1} + 98.74) + \\ &\frac{w_2}{w_1 + w_2 + w_3 + w_4 + w_5 + w_6} (-12.55Z_{t-2} + 1.835Z_{t-1} + 26.84) + \\ &\frac{w_3}{w_1 + w_2 + w_3 + w_4 + w_5 + w_6} (3.664Z_{t-2} - 5.022Z_{t-1} + 16.87) + \\ &\frac{w_4}{w_1 + w_2 + w_3 + w_4 + w_5 + w_6} (-0.3716Z_{t-2} + 0.6656Z_{t-1} + 5.791) + \\ &\frac{w_5}{w_1 + w_2 + w_3 + w_4 + w_5 + w_6} (1.498Z_{t-2} - 2.672Z_{t-1} + 3.001) + \\ &\frac{w_6}{w_1 + w_2 + w_3 + w_4 + w_5 + w_6} (-7.531Z_{t-2} - 1.774Z_{t-1} + 63.88), \end{aligned}$$

where,

$$\begin{aligned} w_1 &= \exp \left\{ -\frac{1}{2} \left(\frac{Z_{t-2} - 4.365}{0.9754} \right)^2 \right\} \times \\ &\exp \left\{ -\frac{1}{2} \left(\frac{Z_{t-1} - 3.806}{1.128} \right)^2 \right\}, \\ w_2 &= \exp \left\{ -\frac{1}{2} \left(\frac{Z_{t-2} - 3.164}{1.165} \right)^2 \right\} \times \\ &\exp \left\{ -\frac{1}{2} \left(\frac{Z_{t-12} - 3.238}{0.8479} \right)^2 \right\} \end{aligned}$$

$$\begin{aligned} w_3 &= \exp \left\{ -\frac{1}{2} \left(\frac{Z_{t-2} - 6.529}{0.3283} \right)^2 \right\} \times \\ &\exp \left\{ -\frac{1}{2} \left(\frac{Z_{t-1} - 6.977}{0.6233} \right)^2 \right\} \\ w_4 &= \exp \left\{ -\frac{1}{2} \left(\frac{Z_{t-2} - 8.242}{1.238} \right)^2 \right\} \times \\ &\exp \left\{ -\frac{1}{2} \left(\frac{Z_{t-1} - 8.131}{1.385} \right)^2 \right\} \\ w_5 &= \exp \left\{ -\frac{1}{2} \left(\frac{Z_{t-2} - 1.985}{1.286} \right)^2 \right\} \times \\ &\exp \left\{ -\frac{1}{2} \left(\frac{Z_{t-12} - 1,599}{0,86} \right)^2 \right\} \\ w_6 &= \exp \left\{ -\frac{1}{2} \left(\frac{Z_{t-2} - 5.49}{1.745} \right)^2 \right\} \times \\ &\exp \left\{ -\frac{1}{2} \left(\frac{Z_{t-12} - 5.137}{0,9701} \right)^2 \right\}. \end{aligned}$$

Similarly, the ANFIS modeling is carried out with 2, 3, 4, and 5 clusters. These models are then used to predict the inflation rate. To assess model accuracy, the predicted inflation rate is compared to the actual series in the testing set using MAPE. The MAPE values for the ANFIS model with a different number of clusters are presented in Table 7. The lowest MAPE value is achieved when considering six clusters.

Table 7. Accuracy of Inflation Prediction using ANFIS Model with 2, 3, 4, 5, and 6 Clusters

No.	Number of Clusters	MAPE (%)
1	2	32.78
2	3	22.24
3	4	30.06
4	5	31.62
5	6	18.49

Model Comparison

The two best models were chosen in the previous section: the hybrid ARIMA(2,1,0)-ANFIS model with a two-cluster solution and the individual ANFIS model with a six-cluster solution. The predicted inflation rate from the two models is presented in Table 8. According to the table, the hybrid ARIMA(2,1,0)-ANFIS model outperforms the individual

ANFIS model because it has a lower MAPE value of 10.30%. Therefore, it can be concluded that the hybrid model can provide better prediction models with higher accuracy. The accuracy performance of these models is also

presented in Figure 10. When compared to the individual ANFIS model, the predicted inflation rate obtained by the hybrid model is very close to the actual value.

Table 8. Accuracy Comparison between Hybrid ARIMA(2,1,0)-ANFIS Model with Two Clusters and Individual ANFIS Model with 6 Clusters to Predict the Inflation Rate

Month	Actual	Hybrid ARIMA(2,1,0)-ANFIS	ANFIS with 6 clusters	ARIMA (2,1,0)
Jan-21	1.55	1.75	1.73	1.88
Feb-21	1.39	1.67	1.99	1.84
Mar-21	1.37	1.65	2.03	1.82
Apr-21	1.42	1.57	1.82	1.81
May-21	1.68	1.61	1.70	1.80
Jun-21	1.33	1.68	1.47	1.79
Jul-21	1.52	1.61	2.33	1.77
Aug-21	1.59	1.63	1.49	1.76
Sep-21	1.6	1.72	1.70	1.75
Oct-21	1.66	1.66	1.80	1.74
Nov-21	1.75	1.68	1.76	1.72
Dec-21	1.87	1.70	1.79	1.71
MAPE (%)		10.30	18.49	17.28

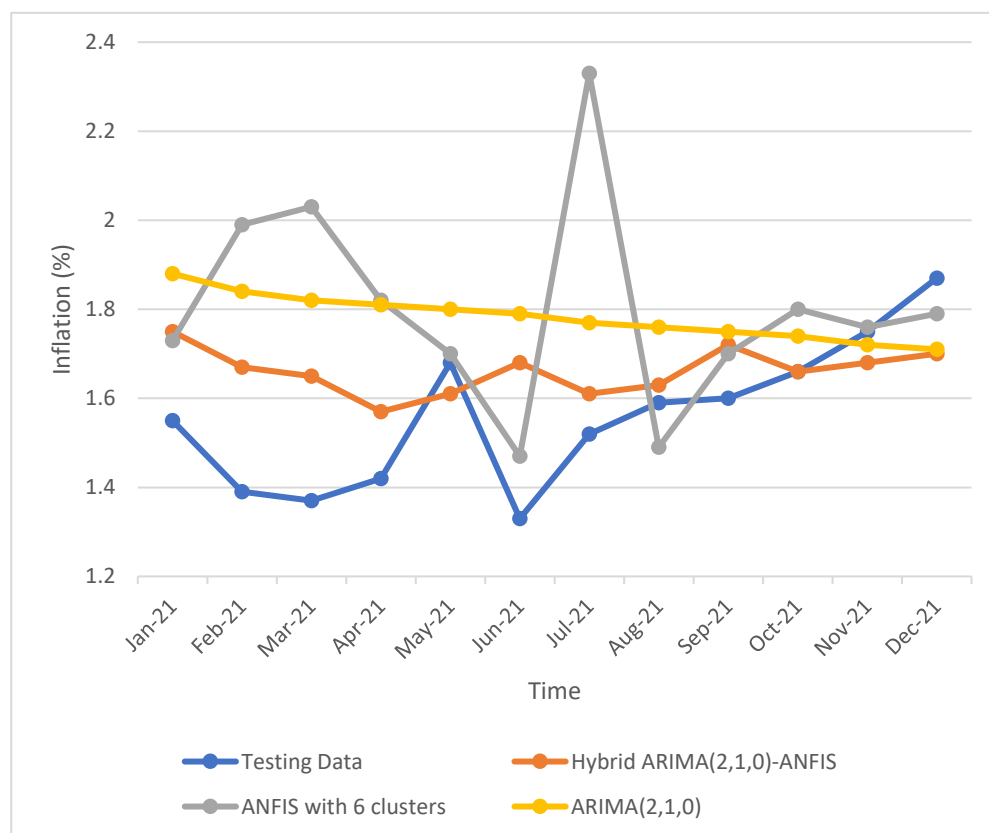


Figure 10. Comparison of Inflation Rate Prediction using Hybrid ARIMA(2,1,0)-ANFIS Model with Two Clusters, Individual ANFIS Model with Six Clusters, and Classical ARIMA(2,1,0) Model

The prediction of the inflation rate for January – December 2022 is presented in Table 9. On average, the predicted inflation in Indonesia is about 1.70%, with

a standard deviation of 0.096%. Figure 11 shows the testing data and forecast generated by the hybrid ARIMA(2,1,0)-ANFIS model with two clusters.

Table 9. The Results of Inflation Prediction in January - December 2022

Month	Prediction of Inflation Rate
Jan-22	1.67
Feb-22	1.67
Mar-22	1.82
Apr-22	1.79
May-22	1.74
Jun-22	1.89
Jul-22	1.67
Aug-22	1.72
Sep-22	1.65
Oct-22	1.68
Nov-22	1.61
Dec-22	1.53

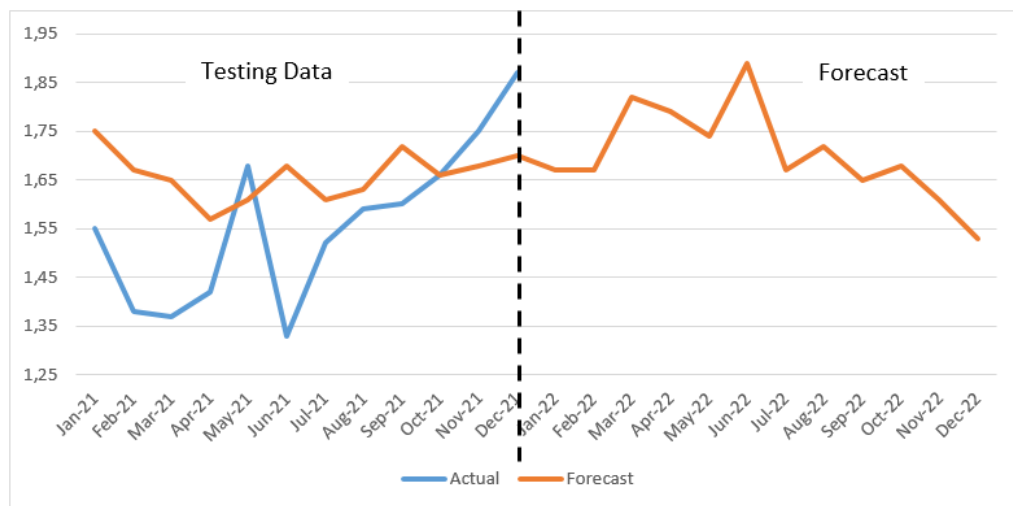


Figure 11. Inflation Rate (Testing Data and Prediction) using Hybrid ARIMA(2,1,0)-ANFIS Model with Two Clusters

CONCLUSIONS AND SUGGESTIONS

This paper successfully compares the inflation rate prediction results using hybrid ARIMA-ANFIS, individual ANFIS, and classical ARIMA models. The ARIMA model is able to account for linear structure in the inflation rate data, and the ANFIS model is considered to accommodate for non-linear structure in the data. The clustering method used is FCM, which involves varying the number of clusters. Based on the MAPE values, the hybrid ARIMA(2,1,0)-ANFIS model with a

two-cluster solution outperforms the other two competing models. This indicates that the proposed hybrid model is able to improve prediction accuracy more than the individual model alone. The MAPE values obtained from individual models are about 17-18%, decreasing to about 10% with the combined approach. Thus, it is evident that the hybrid model is capable of decreasing the error prediction to almost half.

REFERENCES

- Bank Indonesia. (2022). *Data inflasi*. <https://www.bi.go.id/id/statistik/indikator/data-inflasi.aspx>
- Barak, S., & Sadegh, S. S. (2016). Forecasting energy consumption using ensemble arima-anfis hybrid algorithm. *International Journal of Electrical Power and Energy Systems*, 82. <https://doi.org/10.1016/j.ijepes.2016.03.012>
- Benmouiza, K., & Cheknane, A. (2019). Clustered anfis network using fuzzy c-means, subtractive clustering, and grid partitioning for hourly solar radiation forecasting. *Theoretical and Applied Climatology*, 137(1-2). <https://doi.org/10.1007/s00704-018-2576-4>
- Charemza, W., Díaz, C., & Makarova, S. (2019). Conditional term structure of inflation forecast uncertainty: The copula approach. *Romanian Journal of Economic Forecasting*, 22(1).
- Cryer, J. D., & Chan, K.-S. (2008). Time series analysis: With applications in R. In *Design* (2nd ed., Issue January). Springer Science.
- Diarsih, I. H., Tarno, & Rusgiyono, A. (2019). Modeling of red onion production in central java using hybrid arima-anfis. *Journal of Physics: Conference Series*, 1217(1). <https://doi.org/10.1088/1742-6596/1217/1/012080>
- Enke, D., & Mehdiyev, N. (2014). A hybrid neuro-fuzzy model to forecast inflation. *Procedia Computer Science*, 36(C). <https://doi.org/10.1016/j.procs.2014.09.088>
- Faulina, R., & Suhartono. (2013). Hybrid arima-anfis for rainfall prediction in indonesia. *International Journal of Science and Research*, 2(2).
- Fauzia, M. (2021). *Inflasi 2020 1.68 persen terendah sepanjang sejarah*. [kompas.com.:](https://money.kompas.com/read/2021/01/04/125828426/inflasi-2020-168-persen-terendah-sepanjang-sejarah?page=all)
- Fitriyati, N., & Wijaya, M. Y. (2022). A monte carlo simulation study to assess estimation methods in cfa on ordinal data. *CAUCHY: Jurnal Matematika Murni Dan Aplikasi*, 7(3), 332-344.
- Hillmer, S. C., & Wei, W. W. S. (1991). Time series analysis: Univariate and multivariate methods. *Journal of the American Statistical Association*, 86(413). <https://doi.org/10.2307/2289741>
- Işığışok, E., Öz, R., & Tarkun, S. (2020). Forecasting and technical comparison of inflation in turkey with box-jenkins (arima) models and the artificial neural network. *International Journal of Energy Optimization and Engineering*, 9(4). <https://doi.org/10.4018/ijeoe.2020100106>
- John, J., Singh, S. R., & Kapur, M. (2020). *Inflation forecast combinations – the indian experience*.
- Massey, F. J. (1951). The kolmogorov-smirnov test for goodness of fit. *Journal of the American Statistical Association*, 46(253). <https://doi.org/10.1080/01621459.1951.10500769>
- Nau, R. (2014). *Notes on linear regression analysis*. https://people.duke.edu/~rnau/Notes_on_linear_regression_analysis--Robert_Nau.pdf
- Reserve Bank of India. (2017). *Monetary policy report*.
- Tarno, Subanar, Rosadi, D., & Suhartono. (2013). Analysis of financial time series data using adaptive neuro fuzzy inference system (anfis). *IJCSI: International Journal of Computer Science*, 10(2).
- Wen, Y. C., Tan, L. N., & Wuc, H. L. (2014). Inflation forecast based on bp neural

network model. *Advanced Materials Research*, 989–994.
<https://doi.org/10.4028/www.scientific.net/AMR.989-994.5536>

Zhang, B. (2019). Real-time inflation

forecast combination for time-varying coefficient models. *Journal of Forecasting*, 38(3).
<https://doi.org/10.1002/for.2563>