Construction of Fuzzy Inference Model to Predict Percentage of Poor Population in Indonesia

R Rustanuarsi¹ and A M Abadi²

¹Graduate Program of Mathematics Education, Yogyakarta State University, Colombo Street 01, Sleman, Yogyakarta, Indonesia
²Department of Mathematics, Faculty of Mathematics and Natural Sciences, Yogyakarta State University, Colombo Street 01, Sleman, Yogyakarta, Indonesia

Abstract. Fuzzy inference system is a soft computing model which has been widely applied to build a prediction model. This model applied the fuzzy sets theory and believed providing high accuracy for prediction. This study aims to construct the fuzzy inference model to predict the percentage of poor population in Indonesia based on unemployment rate and Gini index. The data of unemployment rates and Gini index are used as input while the data of poor population percentage is used as output. This fuzzy inference model consists of 5 rules. It used Mamdani Max-Min method for inference and centroid defuzzifier for defuzzification. The result demonstrated that the performance of this fuzzy inference model can predict the percentage of poor population with an accuracy level of 94.34%. Therefore, it can be concluded that fuzzy inference system can be used as an appropriate alternative model for predicting the percentage of poor population because it provides high accuracy.

1. Introduction

Poverty is a global issue which is still a big problem in developing countries. One of the national development success main indicators is the declining number of poor population. According to Central Bureau of Statistics [1] in September 2017, the number of poor population in Indonesia has reached 26.58 million (10.12%). It means there is a decline of 1.19 million poor population compared to March 2017 which is 27.77 million. From this data, it can be seen that there is a declining of poor population from period to period.

Based on that phenomenon, it allows the government to propose prevention programs of the poor population accretion in the future. In order to make these programs not being redundant, it is necessary to predict the percentage of poor population. Prediction which known as forecasting plays a notable role in making crucial decisions about the future [2]. Therefore, many decisions are made based on predictions [3]. The result of prediction can be used by the policymakers as a consideration to create the poverty alleviation programs.

The important thing to be considered in predicting the percentage of poor population is how to construct an accurate prediction model. It is important to focus on some factors that affect the number of poor people. The scientific literature on poverty-related causes identifies one of the primary factors
causing poverty is unemployment [4]. There is a positive relationship between high unemployment and widespread poverty [5]. In Indonesia, the term of unemployment rate known as open unemployment rate. In addition, the number of poor people is also affected by income inequality [6]. There are important relationships between the level of income inequality and the extent of poverty. The gap size of income distribution is expressed by the Gini index.

The main step in predicting is constructing a modeling using the available data [7]. In the last decade, there has been a rapidly growing model for predicting based on soft computing, such as Neural Networks (NN) and fuzzy model. Both models are flexible, they do not require assumptions on the data and they have high accuracy forecasts [8]. There are many types of NN which have been performed by other researchers. For example, Elman Recurrent Neural Network (ERNN) model that has been applied for predicting consumer price index of education, recreation and sports [9] and Radial Basis Function Neural Network (RBFNN) model that has been applied for predicting the foreign tourist flows to Yogyakarta [10]. However, the weakness of NN model is the process that is in a black box so it is not transparent [11]. The researchers also have developed a neural-fuzzy model which combines fuzzy and NN model. It has been applied to forecast the number of train passenger [11].

In this study, we consider to use fuzzy model to predict the percentage of poor population. Besides it does not require any assumption, it is able to approximate any function. It also able to model the data based on a combination of empirical data and expert knowledge represented as fuzzy logic [11]. Although in the using of fuzzy model, the generated estimation model can be relatively higher or lower than the real data [7], it attracts researchers to try improving the quality of forecasting in many ways. The researchers reducing error using model translation for time series data [12].

Another type of fuzzy model besides time series model is fuzzy inference system. A Fuzzy Inference System (FIS) is one of fuzzy model type which uses the fuzzy sets theory [13]. Fuzzy set theory is another useful tool to increase prediction efficiency and effectiveness [3]. In FIS, there are three methods, which are Mamdani, Takagi-Sugeno, and Tsukamoto. Furthermore, Mamdani method is applied in this study since Mamdani method is widely known and commonly used in developing fuzzy model [14].

Recently, the application of fuzzy inference of Mamdani method has been widely performed in various fields of life. In geology, fuzzy inference with Mamdani method applied to determine hydrocarbon prospective zone [15]. In medical, it applied to classify the toddlers’ nutritional status [16] and to detect the type of thalassemia disease in patients [17]. In climatology, it has been applied to predict the rainfall [18] and to predict the air quality index [19]. In economic sector, it has been applied to predict the number of pottery souvenir production [20] and the number of fertilizer ordering [21] by considering the demand and supply as input variables. In this study, we will construct the fuzzy inference model to predict the percentage of the poor population in Indonesia. The factors used as inputs are open unemployment rate and Gini index.

2. Fuzzy Inference System
Fuzzy inference system contains the expert’ knowledge and experience, in the design of a system that controls a process which input-output relations defined by fuzzy rules. [22]. FIS contains the process of employing fuzzy logic to formulate the mapping from the input to an output. In classical logic, the truth value of a proposition pertaining to either 0 or 1. While in fuzzy logic, the truth value of a proposition ranging from 0 to 1. Fuzzy logic is widely used due to its ability to express the vagueness and imprecise information [14]. It utilizes fuzzy variables such as long, medium, and short to explain the membership [23].

Two types of fuzzy inference system model which already well-known and commonly used are Mamdani and Takagi Sugeno. In Mamdani, both inputs and outputs can be represented with fuzzy sets. Whereas in Takagi-Sugeno model, output is a numeric or linear [13]. In this study, we used Mamdani model.
2.1. Fuzzy Inference Model to Predict the Percentage of Poor Population  
This study focuses on the construction of fuzzy inference model to predict the percentage of the poor population in Indonesia. The proposed fuzzy inference model consists of five procedures i.e (1) define the universal set of each input and output variable, (2) define the fuzzy set of each input and output, (3) determine fuzzy rule bases, (4) fuzzy inference, and (5) defuzzification. The scheme of the fuzzy inference model which used in this study adapted from Wang [24] which is illustrated in Figure 1.

![Diagram of fuzzy inference model]

**Figure 1.** Model of fuzzy inference for predicting the percentage of poor population in Indonesia.

Based on the Figure 1 the procedures in using fuzzy inference model of Mamdani’s method is given by the following steps.

Step 1: Define the universal set of each input and output. In this study, we assigned the unemployment rates and Gini index as input variables, while the percentage of poor population as the output variable. The universal set contains all the possible element or value of each input and output variables [24].

Step 2: Define the fuzzy set of each input and output. A fuzzy set in a universal set $U$ is characterized by a membership function $\mu_A(x)$ that takes values in the interval $[0, 1]$ [24]. In this study, some of the linguistic terms (e.g., low, medium, high) referred to as fuzzy sets, are assigned to each input and output variables.

Step 3: Determine fuzzy rule bases. A fuzzy rule base consists of a set of fuzzy IF-THEN rules [24]. In order to determine the fuzzy rules, we have to fuzzify every pair of input and output data. Fuzzification is a process to compare the input variables with the membership function on the premise part to obtain the membership value of each linguistic fuzzy set [18]. Therefore, we should convert the crisp values of input and output data to a fuzzy value using the membership function as defined in step 2. For each input and output variables, we should determine the fuzzy set which has the highest membership degree. If there are conflicting rules, then chosen rule is the rule which has the highest degree [25].

Step 4: Inference using Mamdani method. Mamdani method also known as the max-min method [26]. In the max-min method, the minimum value from each rule is selected using fuzzy min operator. Next, it selects the maximum value from that combined consequents of any such rules.

Step 5: Defuzzify using centroid defuzzifier [24]. The purpose of defuzzify is to compute the crisp value on the output. Defuzzify was known as a reverse process of fuzzification. In this study, we used centroid defuzzifier.

### 3. Empirical Result
This study aims to construct the fuzzy inference model to predict the percentage of poor population in Indonesia based on unemployment rate and Gini index. In this section, we would implement the fuzzy inference model. This study used 14 empirical data per semester during the 2011-2017 period, which consist of the unemployment rate, Gini index, and percentage of poor population data in Indonesia. The data were collected from the Central Bureau of Statistics (BPS) Indonesia [1]. Those data can be seen in Table 1.
Table 1. Unemployment rate, Gini index, and percentage of poor population.

| Semester | Unemployment Rate (%) | Gini Index | Poor Population (%) |
|----------|-----------------------|------------|---------------------|
| 2011     | 6.96                  | 0.410      | 12.49               |
|          | 7.48                  | 0.388      | 12.36               |
| 2012     | 6.37                  | 0.410      | 11.96               |
|          | 6.13                  | 0.413      | 11.66               |
| 2013     | 5.88                  | 0.413      | 11.37               |
|          | 6.17                  | 0.406      | 11.47               |
| 2014     | 5.7                   | 0.406      | 11.25               |
|          | 5.94                  | 0.414      | 10.96               |
| 2015     | 5.81                  | 0.408      | 11.22               |
|          | 6.18                  | 0.402      | 11.13               |
| 2016     | 5.5                   | 0.397      | 10.86               |
|          | 5.61                  | 0.394      | 10.70               |
| 2017     | 5.33                  | 0.393      | 10.64               |
|          | 5.5                   | 0.391      | 10.12               |

From Table 1, it can be seen there are 14 pairs of input and output data. Then, let the unemployment rate variable is denoted $A$, so the universal set for this variable is $A = [4, 9]$. The Gini index variable is denoted $B$ and the universal set for it defined as $B = [0.3, 0.5]$. The percentage of poor population is denoted $C$ with the universal set $C = [9, 14]$. We employed the linguistic terms in this study such as low ($L$), medium ($M$) and high ($H$). Therefore, each of the universal set $A$, $B$, and $C$ are defined by three fuzzy sets of low ($A_L$, $B_L$, and $C_L$), medium ($A_M$, $B_M$, and $C_M$), and high ($A_H$, $B_H$, and $C_H$). Those fuzzy sets characterized by triangular and trapezoid membership function can be seen in the following equations.

\[
\mu_{A_L}(x) = \begin{cases} 
1 & \text{if } x \leq 5.5 \\
6.5 - x & \text{if } 5.5 \leq x \leq 6.5 \\
0 & \text{if } x \geq 6.5 
\end{cases} \quad (1)
\]

\[
\mu_{A_M}(x) = \begin{cases} 
x - 5.5 & \text{if } 5.5 \leq x \leq 6.5 \\
7.5 - x & \text{if } 6.5 \leq x \leq 7.5 \\
0 & \text{if } x \leq 5.5 \text{ or } x \geq 7.5 
\end{cases} \quad (2)
\]

\[
\mu_{A_H}(x) = \begin{cases} 
1 & \text{if } x \geq 7.5 \\
x - 6.5 & \text{if } 6.5 \leq x \leq 7.5 \\
0 & \text{if } x \leq 6.5 
\end{cases} \quad (3)
\]

\[
\mu_{B_L}(x) = \begin{cases} 
0.1 - x & \text{if } x \leq 0.38 \\
0.02 & \text{if } 0.38 \leq x \leq 0.4 \\
0 & \text{if } x \geq 0.4 
\end{cases} \quad (4)
\]

\[
\mu_{B_M}(x) = \begin{cases} 
x - 0.38 & \text{if } 0.38 \leq x \leq 0.4 \\
0.02 & \text{if } x \leq 0.38 \text{ or } x \geq 0.42 
\end{cases} \quad (5)
\]
\[
\mu_{B_H}(x) = \begin{cases} 
1 & \text{if } x \geq 0.42 \\
0.42 - x & \text{if } 0.4 \leq x < 0.42 \\
0 & \text{if } x < 0.4 
\end{cases}
\]
(6)

\[
\mu_{C_L}(x) = \begin{cases} 
1 & \text{if } x \leq 10.5 \\
11.5 - x & \text{if } 10.5 \leq x < 11.5 \\
0 & \text{if } x \geq 11.5 
\end{cases}
\]
(7)

\[
\mu_{C_M}(x) = \begin{cases} 
x - 10.5 & \text{if } 10.5 \leq x < 11.5 \\
12.5 - x & \text{if } 11.5 \leq x < 12.5 \\
0 & \text{if } x \geq 12.5 \text{ or } x \leq 10.5
\end{cases}
\]
(8)

\[
\mu_{C_H}(x) = \begin{cases} 
x - 11.5 & \text{if } 11.5 \leq x < 12.5 \\
0 & \text{if } x \geq 12.5 \text{ or } x \leq 11.5
\end{cases}
\]
(9)

The membership function of fuzzy set \(A, B,\) and \(C\) are presented in Figure 2.

\[\text{Figure 2. (a) Graph of membership function of unemployment rate, (b) Graph of membership function of Gini Index, (c) Graph of membership function of poor population}\]

Then, we have to fuzzify every pair of input and output data value using the membership function as defined in equation (1) - (9). This process generate 14 rules as many as input-output data pairs. The rules are can be seen in Table 2.

\[\begin{array}{|c|c|c|c|c|c|c|}
\hline
\text{Rules} & \text{IF} & \text{THEN} & \text{Membership Degree} \\
\hline
1 & A_M & B_H & C_H & 0.54 & 0.5 & 0.99 & 0.267 \\
2 & A_H & B_L & C_H & 0.98 & 0.6 & 0.86 & 0.506 \\
3 & A_M & B_M & C_M & 0.87 & 0.5 & 0.54 & 0.235 \\
4 & A_M & B_H & C_M & 0.63 & 0.65 & 0.84 & 0.344 \\
\hline
\end{array}\]

Table 2. Unemployment rate, Gini index, and percentage of poor population.
As shown in Table 2, we found two cases. First, there are some rules which have the similar antecedent, but different consequent. Second, there are some rules which have the similar antecedent and similar consequent. In this cases, we just should select one rule with the highest multiplication membership degree of each case. This process is reduced the number of rules, then we obtain 5 rules as shown in Table 3.

Table 3. Rules after reduction.

| Rules | IF | THEN |
|-------|----|------|
| 1     | $A_L$, $B_M$ | $C_l$, $C_M$, $C_l$ |
| 2     | $A_L$, $B_H$ | $C_M$, $C_l$, $C_H$ |
| 3     | $A_M$, $B_M$ | $C_M$, $C_M$, $C_M$ |
| 4     | $A_M$, $B_H$ | $C_M$, $C_M$, $C_H$ |
| 5     | $A_H$, $B_L$ | $C_M$, $C_M$, $C_H$ |

Table 3 shows that there are 5 rules after the reduction process. The linguistic expression of the rules above are as follows:

- If the unemployment rate is low and the Gini index is medium then the percentage of poor population is low (Rule 1).
- If the unemployment rate is low and the Gini index is high then the percentage of poor population is medium (Rule 2).
- If the unemployment rate is medium and the Gini index is medium then the percentage of poor population is medium (Rule 3).
- If the unemployment rate is medium and the Gini index is high then the percentage of poor population is medium (Rule 4).
- If the unemployment rate is high and the Gini index is low then the percentage of poor population is high (Rule 5).

The predicted result obtained using fuzzy inference model with Mamdani method and centroid defuzzifier is shown in Figure 3.
Figure 3 presents summary result predicted value of percentage of poor population. Then, it is important to measures the error rate in this study using MAPE (Mean Absolute Percentage Error). MAPE is the mean of the overall percentage error across the actual value and the predicted value. Based on calculation, the error rate in predicting the percentage of the poor population in Indonesia using fuzzy inference model of Mamdani method is 0.0566 or 5.66%. In other words, this fuzzy inference model is giving the accuracy about 94.34%. It should be noted, the proposed of this fuzzy inference model can provide high accuracy for predict the percentage of poor population.

4. Conclusion
Based on this study, to predict the percentage of poor population using fuzzy inference, we can use unemployment rates and Gini index as the input variables. We employed fuzzy inference of Mamdani method to predict the percentage of poor population. There are some steps in doing this study, those are define the universal set of each input and output variable, define the fuzzy set of each input and output, determine fuzzy rule bases, inference using Mamdani method, and defuzzify using centroid defuzzification approach. This fuzzy inference model consists of 5 rules. This fuzzy inference model can successfully predict the the percentage of poor population in Indonesia with an accurate rate of 94.34%. The recommendation for further study is to increase the accuracy of the prediction model by increasing the number of data, using other variables as input that affects the poor population, or using other fuzzy system methods.

References
[1] Badan Pusat Statistik [Central Bureau of Statistics] Indonesia Persentase Penduduk Miskin, Tingkat Pengangguran Terbuka, dan Gini Rasio 2011-2017 [Percentage of Poor Population, Open Unemployment rate, and Gini Index 2011-2017] (Jakarta: Badan Pusat Statistik [Central Bureau of Statistics] Indonesia)
[2] Sah M and Degtiarev K Y 2007 Int. J. Comput. Inf. Eng. 1 221-224
[3] Kahraman C, Yavuz M and Kaya I 2010 Fuzzy and Grey Forecasting Techniques and Their Applications in Production Systems Production Engineering and Management under Fuzziness (Study In Fuzzines and Soft Computing vol 252) ed C Kahraman and M Yavuz (Berlin: Springer) chapter 1 pp 1-24
[4] Šileika A and Bekerytė J 2013 J. Security and Sustainability Issues 2 59-70
[5] Quy N H 2016 Int. J. Econ. Finance 8 113-119
[6] Ncube M, Anyanwu J C and Hausken K 2014 African Development Review 26 435–453
[7] Nurhayadi, Subanar, Abdurakhman and Abadi A M 2014 Appl. Math. Sci. 8 2113-24
[8] Wutsqa D U and Abadi A M 2012 The 2nd Int. Conf. on Computation for Science and Technology (Nidge, Turkey) pp 1-9
[9] Wutsqa D U, Kusumawati R and Subekti R 2014 The 10th Proc. Int. Conf. on Natural Computation (Xiamen, China: IEEE) pp 192-196
[10] Wutsqa D U and Yasi S M 2015 Proc. The 1st Int. Conf. on Statistical Methods in Engineering, Science, Economy and Education (Yogyakarta: Department of Statistics Universitas Islam Indonesia) pp 1-6
[11] Abadi A M and Wutsqa D U 2014 Int. Conf. Fuzzy Systems and Knowledge Discovery (Xiamen, China: IEEE) pp 178-182
[12] Nurhayadi, Subanar, Abdurakhman and Abadi A M 2014 J. Math. Stat. 10 267-274
[13] Vafaei L E and Sah M 2017 Procedia Computer Science 120 221-228
[14] Moahmmed S A and Sadkhan S B 2013 Int. J. Sci. Eng. Res. 4 366-371
[15] Zen H N H, Trimartanti L W, Abidin Z and Abadi A M 2017 J. Syst. Sci. Syst. Eng. 26 463-474
[16] Permatasari D, Azizah I N, Hadiat H L and Abadi A M 2017 AIP Conf. Proc. 1868 040007-1
[17] Thakur S, Raw S N, Sharma R and Mishra P 2016 Int. J. Appl. Pharm. Sci. Res. 1 88-95
[18] Agboola A H, Gabriel A J, Aliyu E O and Alese B K 2013 Int. J. Eng. Tech. 3 427-435
[19] Ganesh S S, Reddy N B and Arulmozhiwarman P 2017 Int. Conf. on Trends in Electronics and Informatics (Tirunelveli: IEEE) pp 338-341
[20] Indiani V, Yanianti A, Widiaswari S D and Abadi A M 2017 Proc. 4th Int. Conf. on Research, Implementation and Education of Mathematics and Sciences (Yogyakarta: Faculty of Mathematics and Natural Sciences Yogyakarta State University) pp 117-124
[21] Fritiani, Kurniasih N R, Mandini G W and Abadi A M 2017 Proc. 4th Int. Conf. on Research, Implementation and Education of Mathematics and Sciences (Yogyakarta: Faculty of Mathematics and Natural Sciences Yogyakarta State University) pp 27-34
[22] Cavallaro F 2015 Sustainability 7 12359-71
[23] Lei M, Shiyan L, Chuanwen J, Hongling L and Yan Z 2009 Renewable and Sustainable Energy Reviews 13 915-920
[24] Wang L X 1997 A course in fuzzy systems and control (Upper Saddle River: Prentice-Hall International)
[25] Abadi A M, Nurhayadi and Musthofa 2017 J. Eng. Appl. Sci. 12 5672-78
[26] Harliana P and Rahim R 2017 J. Phys.: Conf. Ser. 930 012029