ANN : Prediction of Per Capita Income Rural Community on Poverty Line Based on Province

Solikhun\textsuperscript{1}, Mochamad Wahyudi\textsuperscript{1}, M. Safii\textsuperscript{1}, Syahril Efendi \textsuperscript{2}, Suci Ramadhani\textsuperscript{3}, Yogi Yunefri\textsuperscript{1}, Ahmad Zamsuri\textsuperscript{4}

\textsuperscript{1}Doctoral Program, Faculty of Computer Science and Information Technology, Universitas Sumatera Utara, Sumatera Utara, Indonesia
\textsuperscript{2}Faculty of Computer Science and Information Technology, Universitas Sumatera Utara
\textsuperscript{3}Universitas Pembangunan Panca Budi, Medan, North Sumatera, Indonesia
\textsuperscript{4}Universitas Lancang Kuning, Pekanbaru, Indonesia

*solikhun@amiktunasbangsa.ac.id

Abstract. The problem of poverty is a fundamental problem that is of concern to every country. The Indonesian state has established a poverty reduction program as the main program. Poverty occurs in urban and rural communities. This research raises the problem of poverty in rural communities. The contribution of research to the government is to predict the per capita income of rural communities according to the poverty line based on the future provinces. The data used is data from the National Statistics Agency. These data are 2015 semester 1 data up to 2018 semester 1. The algorithm for its completion uses the artificial neural network backpropagation method. Input data is 2015 data for the 1 to 2017 semester 2. The training and testing architecture model is 4, namely 6-2-1, 6-3-1, 6-2-3-1, and 6-3-2-1. Target data is 2018 semester 1 data. The best architectural model is 6-2-1 with 79 epoch, MSE 0.004801 and 100% accuracy rate. From this model, a prediction of rural income per capita in the poverty line is based on the provinces of each province in Indonesia.

1. Introduction

Poverty is a fundamental problem that is of concern to the Indonesian government. Every new government always brings the idea of change in the economic field. The government needs to predict rural income per capita in the poverty line based on the province as a material consideration to take policy on poverty alleviation programs in rural areas. The method used in this study is an artificial neural network (ANN)\textsuperscript{[1]} with a backpropagation algorithm\textsuperscript{[2], [3]}. Backpropagation algorithms are able to generalize and extract from a certain pattern and create new patterns through the process of learning ability (self organizing)\textsuperscript{[4]–[6]}. In addition, ANN can also process input data without having to have a target\textsuperscript{[3][7]}. Based on these advantages, this technique is able to predict rural income per capita in the poverty line by province. The results of this study provide input for the government to take policies in poverty alleviation programs in rural areas by making activities that support increased rural income so that income in rural areas continues to increase.
2. Methodology

Data on rural per capita income in the poverty line based on the province will be processed by the Artificial Neural Network with the backpropagation method. Data is represented in numerical form between 0 and 1 because the network uses the binary sigmoid activation function (logsig) whose range is from 0 to 1. The values used are obtained based on the categories of each variable as well as to make it easier to remember in defining it. Input data is per capita income data of rural communities in the 2015 poverty line semester 1 to 2017 semester 2. Input data is obtained from the National Statistics Agency’s website on the per capita income of rural communities on the poverty line by province. This data will be transformed into a data between 0 to 1 before training and testing using artificial neural network backpropagation method with the formula:

\[ x' = \frac{0.8(x-a)}{b-a} + 0.1 \]

(1)

The target data is data on rural per capita income in the poverty line based on the province in 2018 Semester 1.

2.1. Artificial Intelligence

Artificial intelligence [8] is a branch of computer science in representing knowledge through symbols and processing information based on a number of rules[9]–[11] [12] [13]. The process of knowledge gained through education and experience organized with each other that can be applied to decision making and problem solving[14] [15][16].

2.2. Backpropagation Neural Network

The Backpropagation method [17][18] is one of the supervised learning methods that minimizes errors in the output generated by the network originating from the network against the network weights that can be modified[19]–[21]. This error is named Error gradient which is used to find the weight value that will minimize error[22][23].

3. Results and Discussion

3.1. Data processing

Data processing is carried out with the help of Matlab 6.1 software applications. The sample data is the per capita income of rural people on the poverty line by province. This data will be used in training data and test data. Samples of data that have been processed and transformed are as follows.

| No | Name | X1  | X2  | X3  | X4  | X5  | X6  | Target |
|----|------|-----|-----|-----|-----|-----|-----|--------|
| 1  | Data 1 | 0.3852 | 0.4099 | 0.4292 | 0.4530 | 0.4729 | 0.5073 | 0.5312 |
| 2  | Data 2 | 0.2843 | 0.3260 | 0.3764 | 0.3985 | 0.4132 | 0.4355 | 0.4645 |
| 3  | Data 3 | 0.3604 | 0.4034 | 0.4489 | 0.4724 | 0.5000 | 0.5044 | 0.5422 |
| 4  | Data 4 | 0.4124 | 0.4549 | 0.4730 | 0.4894 | 0.5228 | 0.5364 | 0.5652 |
| 5  | Data 5 | 0.2540 | 0.2803 | 0.3049 | 0.3202 | 0.3418 | 0.3529 | 0.4181 |
| 6  | Data 6 | 0.2314 | 0.2604 | 0.2837 | 0.3003 | 0.3157 | 0.3328 | 0.3700 |
3.4. Design of Artificial Neural Network architecture

The following stages will be performed in the user of the backpropagation algorithm with the sigmoid activation function. The steps that must be done are as follows:

1. Initialization, is the stage where the value variables will be set or defined in advance, for example: input data value, weight, expected output value, learning rate and other data values.

| No | Keterangan | Error Minimum |
|----|------------|---------------|
| 1  | True       | 0.05 - 0.001  |
| 2  | False      | > 0.05        |
2. Activation, is the process of calculating the actual value of the output in the hidden layer and calculating the actual output value in the output layer.

3. Weight Training, is the process of calculating the gradient error value in the output layer and calculating the value of the gradient error in the hidden layer.

4. Iteration, is the final stage in testing, where if there is still a minimum error that is expected not to be found then return to the activation stage.

3.5. Selection of the best architectural neural network

The results of the Matlab 6.1 application software used for the 6-2-1 architectural model, 6-3-1 architecture, 6-2-3-1 architecture and 6-3-2-1 architecture are obtaining the best architectural patterns. From this pattern, it will be used to predict rural income per capita in the poverty line by province. The best assessment of architectural models is seen from several aspects such as epoch, minimum error and accuracy of truth. For more details can be seen in the following:

Table 3. Recapitulation of Architectural Models

| Model      | 6-2-1 | 6-3-1 | 6-2-3-1 | 6-3-2-1 |
|------------|-------|-------|---------|---------|
| Epochs     | 79    | 35    | 124     | 50      |
| MSE        | 0.004801 | 0.009977 | 0.006745 | 0.025842 |
| Accuracy   | 100%  | 100%  | 100%    | 75%     |

From table 3 it can be seen that the best architectural model that will be used to make predictions from a series of model trials is 6-2-1 with an 79 epoch, MSE 0.004801 and 100% accuracy rate.

4. Conclusion

Based on the results and discussion above, the writer can draw the following conclusions:

a) The Artificial Neural Network model used is 6-2-1 architecture, 6-3-1 model, 6-2-3-1 model and 6-3-2-1 model, can get the best results by looking at the smallest MSE Testing is 6-2-1.

b) With a 6-2-1 architectural model, it can predict urban per capita income at the poverty line by province by showing 100% performance.

References

[1] B. Febriadi and A. Zamsuri, “RDBMS Applications as Online Based Data Archive: A Case of Harbour Medical Center in Pekanbaru,” IOP Conf. Ser. Earth Environ. Sci., vol. 97, no. 1, pp. 1–5, 2017.

[2] A. P. Windarto, L. S. Dewi, and D. Hartama, “Implementation of Artificial Intelligence in Predicting the Value of Indonesian Oil and Gas Exports With BP Algorithm,” Int. J. Recent Trends Eng. Res., vol. 3, no. 10, pp. 1–12, 2017.

[3] Sumijan, A. P. Windarto, A. Muhammad, and Budiharjo, “Implementation of Neural Networks in Predicting the Understanding Level of Students Subject,” Int. J. Softw. Eng. Its Appl., vol. 10, no. 10, pp. 189–204, 2016.

[4] J. Tarigan, R. Diedan, and Y. Suryana, “Plate Recognition Using Backpropagation Neural Network and Genetic Algorithm,” Procedia Comput. Sci. Comput. Intell. ICCSCI, vol. 116, no. 2, pp. 365–372, 2017.

[5] O. F. Ertugrul, “A Novel Type of Activation Function in Artificial Neural Networks: Trained Activation Function,” Neural Networks, vol. 99, pp. 1–22, 2018.

[6] N. M. Nawi, a. Khan, M. Z. Rehan, M. a. Aziz, J. H. Abawajy, and T. Herawan, “Neural network training by hybrid accelerated cuckoo particle swarm optimization algorithm,” Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 8835, pp. 237–244, 2014.

[7] B. Supriyadi, A. P. Windarto, T. Soemartono, and Mungad, “Classification of natural disaster prone areas in Indonesia using K-means,” Int. J. Grid Distrib. Comput., vol. 11, no. 8, pp. 87–98, 2018.

[8] A. Wanto, M. Zarlis, Sawaluddin, and D. Hartama, “Analysis of Artificial Neural Network Backpropagation Using Conjugate Gradient Fletcher Reeves in the Predicting Process,” J. Phys. Conf. Ser., vol. 930, no. 1, pp. 1–7, 2017.

[9] H. Chiroma, S. Abdulkareem, A. I. Abubakar, and T. Herawan, “Kernel functions for the support vector machine: Comparing performances on crude oil price data,” Adv. Intell. Syst. Comput., vol. 287, pp.
[10] R. B. F. Hakim, E. N. Sari, and T. Herawan, “Soft Solution of Soft Set Theory for Recommendation in Decision Making,” *Adv. Intell. Syst. Comput.*, vol. 287, pp. 313–324, 2014.

[11] A. Lasisi, R. Ghazali, and T. Herawan, “Comparative Performance Analysis of Negative Selection Algorithm with Immune and Classification Algorithms,” *Adv. Intell. Syst. Comput.*, vol. 287, pp. 441–452, 2014.

[12] P. Hamet and J. Tremblay, “Artificial Intelligence in Medicine,” *Metabolism.*, pp. 1–14, 2017.

[13] J. Adnan *et al.*, “Multilayer Perceptron Based Activation Function On Heart,” *J. Fundam. Appl. Sci.*, vol. 9, no. 3, pp. 417–432, 2017.

[14] A. Wanto *et al.*, “Levenberg-Marquardt Algorithm Combined With Bipolar Sigmoid Function To Measure Open Unemployment Rate In Indonesia,” *Int. Conf. Comput. Environ. Agric. Soc. Sci. Heal. Sci. Eng. Technol. (3rd ICEST)*, vol. 3, no. 1, 2018.

[15] A. Ehret, D. Hochstuhl, D. Gianola, and G. Thaller, “Application of neural networks with back-propagation to genome-enabled prediction of complex traits in Holstein-Friesian and German Fleckvieh cattle,” *Genet. Sel. Evol.*, vol. 47, no. 1, p. 22, 2015.

[16] M. Fauzan *et al.*, “Epoch Analysis And Accuracy 3 ANN Algorithm Using Consumer Price Index Data In Indonesia,” *Int. Conf. Comput. Environ. Agric. Soc. Sci. Heal. Sci. Eng. Technol. (3rd ICEST)*, vol. 3, no. 1, 2018.

[17] P. Antwi *et al.*, “Estimation of Biogas and Methane Yields in an UASB Treating Potato Starch Processing Wastewater With Backpropagation Artificial Neural Network,” *Bioresour. Technol.*, vol. 228, pp. 106–115, 2017.

[18] S. P. Siregar and A. Wanto, “Analysis of Artificial Neural Network Accuracy Using Backpropagation Algorithm In Predicting Process (Forecasting),” *Int. J. Inf. Syst. Technol.*, vol. 1, no. 1, pp. 34–42, 2017.

[19] D. Huang and Z. Wu, “Forecasting outpatient visits using empirical mode decomposition coupled with backpropagation artificial neural networks optimized by particle swarm optimization,” *PLoS One*, vol. 12, no. 2, pp. 1–17, 2017.

[20] A. Wanto, A. P. Windarto, D. Hartama, and I. Parlina, “Use of Binary Sigmoid Function And Linear Identity In Artificial Neural Networks For Forecasting Population Density,” *Int. J. Inf. Syst. Technol.*, vol. 1, no. 1, pp. 43–54, 2017.

[21] B. Febriadi, Z. Zamzami, Y. Yunefri, and A. Wanto, “Bipolar function in backpropagation algorithm in predicting Indonesia’s coal exports by major destination countries,” *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 420, no. 12089, pp. 1–9, 2018.

[22] A. Wanto *et al.*, “Analysis Of Standard Gradient Descent With GD Momentum And Adaptive LR For SPR Prediction,” *Int. Conf. Comput. Environ. Agric. Soc. Sci. Heal. Sci. Eng. Technol. (3rd ICEST)*, vol. 3, no. 1, 2018.

[23] N. Nasution, A. Zamsuri, L. Lisnawita, and A. Wanto, “Polak-Ribiere updates analysis with binary and linear function in determining coffee exports in Indonesia,” *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 420, no. 12089, pp. 1–9, 2018.