IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⋇ Impact Factor 7.918 ⋇ Vol. 11, Issue 11, November 2022

DOI: 10.17148/IJARCCE.2022.111101

Prediction of Poverty Rates with Nonlinear Autoregressive Neural Networks with External Input

Prima Dina Atika[1], Herlawati[2], Rahmadya Trias Handayanto[3]

Faculty of Computer Science, Universitas Bhayangkara Jakarta Raya, Jakarta, Indonesia[1,2]

Computer Engineering Department, Universitas Islam 45, Bekasi, Indonesia[3]

Abstract: The indicator of the success of a region's development is determined by the reduction in the poverty rate of its citizens. Based on these problems, a method is needed to determine the level of poverty in Indonesia, both in urban, rural and national areas, one of which is by using the prediction method. In this study, a Levenberg-Marquardt Neural Network model to predict the level of poverty in Bekasi Regency in the future is proposed, namely nonlinear autoregressive neural networks with external input (NARX). The poverty dataset used is sourced from the Central Statistics Agency (BPS) of Bekasi Regency with test data from 2012 to 2020. The test results show an error of 6.9% compared to the real poverty rate.

Keywords: Poverty Rates, Neural Networks, Bekasi, MATLAB, Levenberg-Marquardt.

I. INTRODUCTION

Indonesia is one of the developing countries that has a main problem, namely poverty. In the life of the state, social problems such as poverty can affect the value of a nation in the welfare of its people. Poverty is a complex and multidimensional problem because the indicators of the success of a country's development depend on the level of community welfare [1].

Based on data from the Central Bureau of Statistics of Bekasi Regency that from 2016 to 2020 the percentage of the poverty rate in Bekasi Regency has decreased, only in 2015 there was an increase in the poverty rate [2].

Therefore, it is necessary to make predictions to get the prediction results of the number of poor people in Bekasi Regency for the following years, this is done so that the Bekasi Regency government has references and considerations in determining policies and in making appropriate steps to overcome this poverty. Making predictions is not easy, it takes the right data, methods, and steps so that the prediction results can be accounted for later.

The purpose of this research is in accordance with the third mission of the Universitas Bhayangkara Jakarta Raya, which is to conduct research in the context of developing science and technology, both in basic and applicable scientific fields and in accordance with IKU 3, namely the percentage of lecturers who carry out three service activities (TRIDHARMA).

Previous research applied the Bayesian regulation method, this is because this method can predict data based on previous data, so that prediction results are obtained after learning and training based on data that has already happened. the result is that the best architectural model is 10-25-25-2. The accuracy rate of this architectural model is 94.1% and 61.8% with MSE values of 0.00013571 and 0.00005189, respectively. The results of this study are estimates of the poor for the next 5 years [3].

Other researchers use a model from Box Jenkins, namely the Auto Regressive Moving Average (ARIMA) to predict the level of poverty in Indonesia in the future. The poverty dataset used is sourced from the Central Statistics Agency (BPS) with test data from 2011 to 2020. Researchers will use 3 error parameters to evaluate the results of poverty rates in urban, rural, and national levels, namely RMSE, MAE and MAPE. Based on the tests conducted, the urban dataset produces the ARIMA (2,2,5) model as the best ARIMA model with RMSE=1.246582, MAE=0.923255 and MAPE=12%, for the rural dataset, the ARIMA (1,2,1) model as the best with RMSE=0.392650, MAE=0.311529 and MAPE=2%. Meanwhile, for the national dataset, the ARIMA (0,2,5) model is the best with RMSE=2.533166, MAE=2.090505 and MAPE=20%. From the 3 tests, it was concluded that the ARIMA model succeeded in producing forecasting values for the poverty level in Indonesia, both in urban, rural, and national areas with good results [4].

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁜ Impact Factor 7.918 ⁜ Vol. 11, Issue 11, November 2022
DOI: 10.17148/IJARCCE.2022.111101

This study aims to predict data on the number of poor people in Indonesia using the autoregressive neural networks with external input (NARX) method. This type of research is descriptive quantitative. The data used for forecasting from 2012-2019, with MSE and MAPE accuracy parameters. Based on the results of the data simulation, the prediction results for 2020 the number of poor people in 33 provinces in Indonesia is 332,005 people. The accuracy parameter of the Back Propagation architecture is to obtain an MSE of 0.119 and a MAPE of 2.298 [5].

Predicting the amount of poverty in the future can use several types of forecasting methods. By implementing the data on the number of poor people in the previous period as training data, which was tested using a quantitative forecasting method, namely the Single Moving Average which was then refined again with the Double Moving Average method where the error rate value resulting from the combination of the two methods was 3.47 % difference from the actual real amount [6].

Weight and bias training on Levenberg-Marquardt is optimized using genetic algorithms to get more optimal results. In this study, the Levenberg-Marquardt artificial neural network method whose training weight was optimized using genetic algorithms resulted in an AFER value of 8.744579% [7].

II. DATA AND METHODS

This study used time-series data on poverty in the Bekasi district from 2012 to 2020 from the Indonesian Bureau of Statistics. With the Machine Learning model, it is hoped that it can predict the poverty level of the area for years to come.

A. Data

Fig 1 shows the Bekasi district as the study area. This area is a member of the Jakarta Metropolitan Area (JABOTABEK). It has latitude and longitude of -6.241586 and 106.992416, respectively [8].

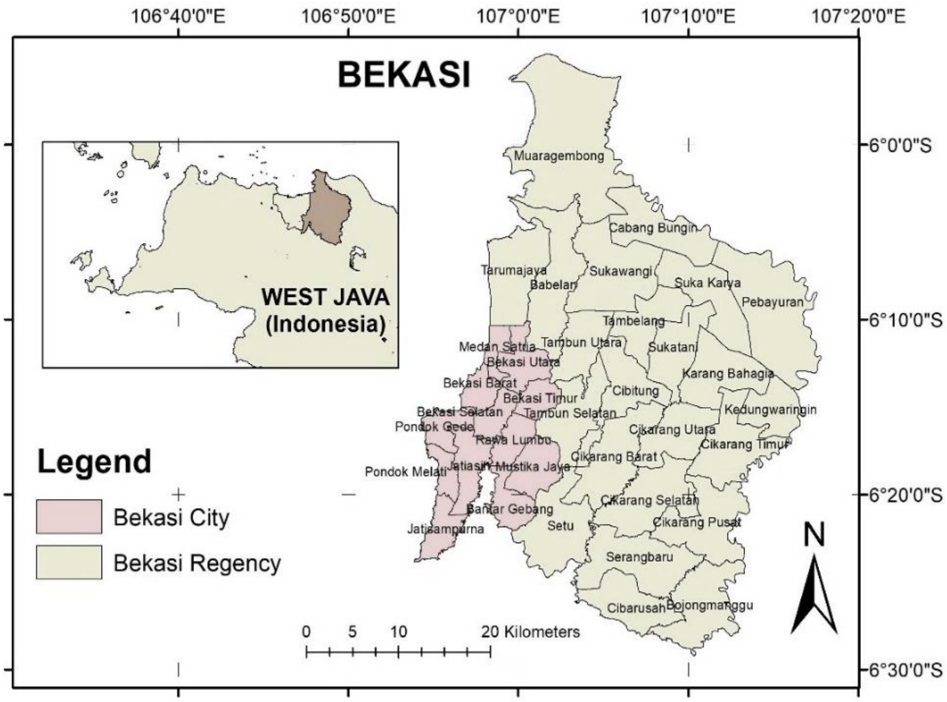

Fig. 1 Bekasi, West J
ava, Indonesia

On the border with Jakarta, some areas become cities, i.e., Bekasi City. This suburbanized area is experiencing rapid growth, both in terms of economy and population. However, the existing conditions show a real growth of the poor population. Table 1 shows the poverty level of Bekasi District. There appears to be considerable growth, especially during the COVID-19 pandemic [9].

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁙ Impact Factor 7.918 ⁙ Vol. 11, Issue 11, November 2022
DOI: 10.17148/IJARCCE.2022.111101

TABLE 1. POOR POPULATION DATA OF BEKASI DISTRICT

Year	Poverty Line (rupiah/capita/month)	Number of Poor People (thousand)	Percentage of Poor People
2012	331032	151.60	5.25
2013	361510	157.70	5.20
2014	374255	156.60	4.97
2015	394513	169.20	5.27
2016	416058	164.40	4.92
2017	434112	160.00	4.73
2018	463507	157.20	4.37
2019	481732	149.40	4.01
2020	490112	186.30	

B. Methods

This study uses a machine learning base to predict time series data. Figure 2 shows the flow of this study. First the timeseries data should be convert to a MATLAB file, namely, MAT-file. Using MATLAB Toolbox, a Nonlinear Autoregressive Neural Networks with External Input (NARX) was created. Several important parameters need to be set, including the number of neurons involved. Furthermore, the training process is run until a model is obtained that is ready to be used for prediction.

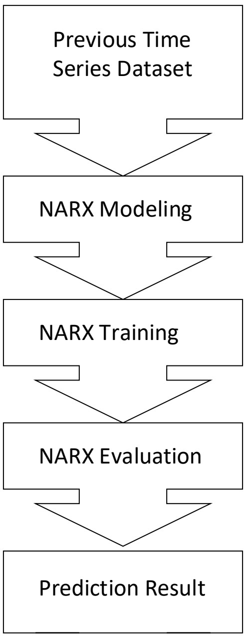

Fig. 2 Change Analysis Framework

Figure 3 shows the MATLAB toolbox for NARX. One of the parameters that distinguishes NARX from regular Neural Networks is the delay [10], [11]. Delay functions to set input and output where some outputs are used as input by NARX. Unlike ARIMA which only uses delay for prediction, NARX uses another variable known as External Input or Exogenous Variable. In this study, the poverty line is used as an external input. Figure 3 also show the network of NARX with 10 hidden neurons with sigmoid function and pure linear function for output layer.
The equation to be formed is:

$$Y(t)=f(x(t-1),\ldots,x(t-d),y(t-1),\ldots,y(t-d)) \quad (1)$$

where y represents the output, x is input, and d represents a delay.

 ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 7.918 ⁜ Vol. 11, Issue 11, November 2022

DOI: 10.17148/IJARCCE.2022.111101

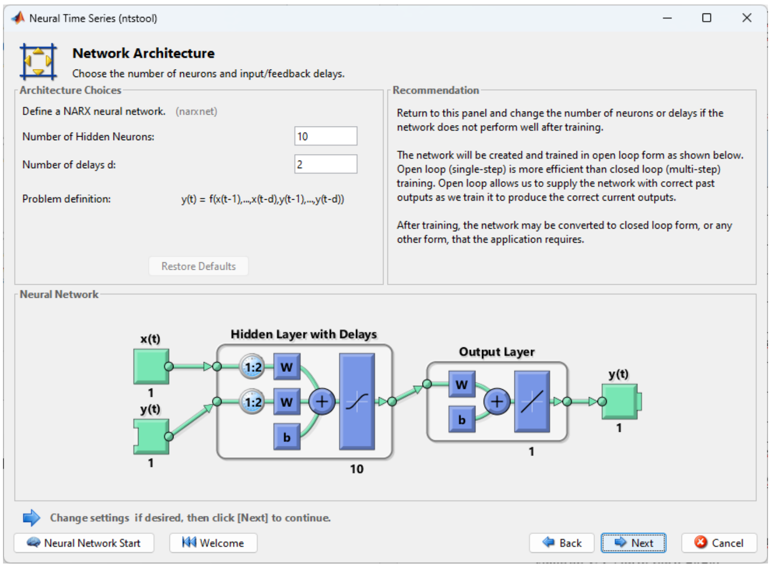

Fig. 3. Change Analysis Framework

There is little setting of input and output matrices whether column or row. Just select the settings, and make sure the next process can be executed by pressing the next button. Then the training process is carried out. With the Levenberg-Marquardt method, training results are obtained with a certain level of accuracy. After the training is complete, the trained model can be saved in the form of a MAT-file. Furthermore, the evaluation stage is carried out with the SIM function which produces predictions with a level of accuracy which will be discussed in the following section. Equation 2 shows the prediction:

$$\text{Pred} = \text{sim}(\text{net1},\text{xtest}) \quad (2)$$

where *Pred* represents the prediction results, *sim* shows the MALAB simulation function of the *net1* model, and the test data is *xtest* [12], [13].

III. RESULT AND DISCUSSION

Small data requires several runs to ensure good accuracy before being used for prediction. Figure 4 shows the performance of the training process.

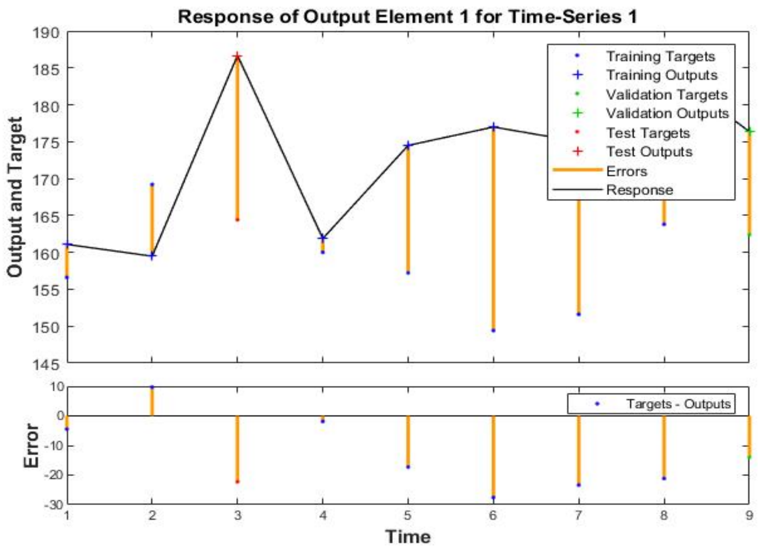

Fig. 4. Training Performance

 ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⋇ Impact Factor 7.918 ⋇ Vol. 11, Issue 11, November 2022

DOI: 10.17148/IJARCCE.2022.111101

Parameters measured include training targets, training outputs, validation targets, validation outputs, test targets, test outputs, errors, and response. another graph is also available in the form of an error histogram with 20 bins. Errors represent the difference between the target and the output.

Fig. 5. Error Histogram

To test the NARX model, from the previous data it can be predicted the poverty rate in the next years, e.g., 2021 (Table 2).

TABLE 2. POOR POPULATION IN BEKASI

Year	Poverty Line (rupiah/capita/month)	Number of Poor People (thousand)	Percentage of Poor People
2012	331032	151.60	5.25
2013	361510	157.70	5.20
2014	374255	156.60	4.97
2015	394513	169.20	5.27
2016	416058	164.40	4.92
2017	434112	160.00	4.73
2018	463507	157.20	4.37
2019	481732	149.40	4.01
2020	490112	186.30	
2021	Prediction	165.65	
	Target	154.08	
	% Error	6.9%	

The results of the study can be used further, i.e., Poverty Line, Head Count Index, Poverty Gap Index, and Poverty Severity Index. The poverty line or the poverty line is the minimum level of income that is deemed necessary to meet to obtain an adequate standard of living in a country. The poverty line is useful as an economic tool that can be used to measure the poor and consider socio-economic reforms, such as welfare improvement programs and unemployment insurance to reduce poverty.

Head Count Index is the percentage of the population below the Poverty Line. The Poverty Gap Index (Poverty Gap Index-P1) is a measure of the average expenditure gap of each poor population towards the poverty line. The higher the index value, the farther the population's average expenditure is from the poverty line.

Poverty Severity Index provides an overview of the distribution of expenditure among the poor. The higher the index value, the higher the disparity of expenditure among the poor.

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⋇ Impact Factor 7.918 ⋇ Vol. 11, Issue 11, November 2022
DOI: 10.17148/IJARCCE.2022.111101

IV. Conclusion

Time series data is one of the sequential data where the order affects the meaning. Most of the researches on forecasting use conventional methods, one of the famous ones is ARIMA. However, the development of artificial intelligence, one of which is a neural network, has created a model that applies this method for projection, one of which is the Nonlinear Autoregressive Neural Network with External Input (NARX). This method is the development of Nonlinear Autoregressive Neural Networks (NAR). With the addition of one external input, or also known as an exogenous variable, an accuracy of 6.9% is obtained for the prediction of the poor in the Bekasi district, West Java, Indonesia. The next research will be compared with current research using Long-Short Term Memory (LSTM) and other sequential deep learning methods.

ACKNOWLEDGMENT

The authors thank to Universitas Bhayangkara Jakarta Raya in supporting this research as internal research grant. Also, for the reviewers who have given the insightful comments.

References

[1] O. LA, S. A, and B. JS, "Children, Families and Poverty," Physiology & behavior, vol. 176, no. 5, pp. 139–148, 2017.
[2] BPS Provinsi Jawa Barat, "Provinsi Jawa Barat dalam Angka 2017," p. 718, 2017.
[3] A. Wanto and J. T. Hardinata, "Estimasi Penduduk Miskin di Indonesia sebagai Upaya Pengentasan Kemiskinan dalam Menghadapi Revolusi Industri 4.0," CESS (Journal of Computer Engineering System and Science), vol. 4, no. 2, pp. 198–207, 2019.
[4] R. I. Prasetyono and D. Anggraini, "Analisis Peramalan Tingkat Kemiskinan Di Indonesia Dengan Model Arima," Jurnal Ilmiah Informatika Komputer, vol. 26, no. 2, pp. 95–110, 2021.
[5] Syaharuddin, E. Pujiana, I. P. Sari, V. M. Mardika, and M. Putri, "Analisis Algoritma Back Propagation Dalam Prediksi Angka Kemiskinan Di Indonesia," J. Pendidik. Berkarakter, vol. 3, no. 1, pp. 11–17, 2020.
[6] F. Kusuma, M. Ahsan, and S. Syahminan, "Prediksi Jumlah Penduduk Miskin Indonesia menggunakan Metode Single Moving Average dan Double Moving Average," Jurnal Informatika dan Rekayasa Perangkat Lunak, vol. 3, no. 2, p. 105, 2021.
[7] A. J. R. Ashshiddiqi, Indriati, and Sutrisno, "Implementasi Jaringan Saraf Tiruan Backpropagation untuk Memprediksi Jumlah Penduduk Miskin di Indonesia dengan Optimasi Algoritme Genetika," Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer (J-PTIIK) Universitas Brawijaya, vol. 2, no. 11, pp. 4638–4646, 2018.
[8] Bureau of Statistics, "Population of Bekasi City," 2015. [Online]. Available: http://bekasikota.bps.go.id/. [Accessed: 24-Nov-2015].
[9] R. T. Handayanto and H. Herlawati, "Efektifitas Pembatasan Sosial Berskala Besar (PSBB) di Kota Bekasi Dalam Mengatasi COVID-19 dengan Model Susceptible-Infected-Recovered (SIR)," Jurnal Kajian Ilmiah, vol. 20, no. 2, pp. 119–124, 2020.
[10] Herlawati, R. T. Handayanto, and Solikin, "Neural network regression with support vector regression for land-use growth prediction," in Proceedings of the 3rd International Conference on Informatics and Computing, ICIC 2018, 2018.
[11] J. Buitrago and S. Asfour, "Short-term forecasting of electric loads using nonlinear autoregressive artificial neural networks with exogenous vector inputs," Energies, vol. 10, no. 1, pp. 1–24, 2017.
[12] R. T. Handayanto and H. Herlawati, Data Mining Dan Machine Learning Menggunakan Matlab Dan Python. Bandung: Informatika Bandung, 2020.
[13] P. Kim, MATLAB Deep Learning. New York: Apress, 2017.

BIOGRAPHY

Prima Dina Atika has published some papers about datamining and database. She is now a lecturer of Informatics department in Universitas Bhayangkara Jakarta Raya.

DOI: 10.17148/IJARCCE.2022.111101

Herlawati earned Magister Management and Magister of Computer Science. She has been published some books and papers in Data Mining, Machine Learning, and Geographic Information Systems. She is also a doctoral student of Computer Science in Binus University. She is still active as lecturer at Bhayangkara University.

Rahmadya Trias Handayanto has published some papers about Geo-AI. After finishing his PhD in Information Management, he held academic position as head of computer engineering department in Universitas Islam 45 Bekasi. He is still active in reviewing some international journal and writing books.

