

Prediction of Life Expectancy in Maluku Province Using Artificial Neural Networks Backpropagation

YOPI ANDRY LESNUSSA, FRANCIS YUNITO RUMLOWANG,
ENDRO RISAMASU, CHARLITA FHILYA

Jurusan Jurusan Matematika, FMIPA Universitas Pattimura, Ambon
e-mail: yopi_a_lesnussa@yahoo.com, rumlawang@gmail.com,
risamasue@gmail.com, chfhilya@gmail.com

Abstract

Life Expectancy at Birth (LE) is defined as the average estimated number of years a person can live to since their birth. The purpose of LE is to represent the health rate of a community. Backpropagation is an algorithm in artificial neural networks (ANN) used to predict or forecast data. This study aims to predict Life Expectancy in Maluku Province. The results of this research obtained the forecasting success of 99.65% with the smallest error $MAPE = 0,0035$. Forecasting for the next 5 years shows that the Life Expectancy value tends to increase over the next 5 years from 2019-2023 at 65.7828 in 2019 and increasing to 66.6632 in 2023.

Keywords: Artificial neural network; Backpropagation; Life Expectancy.

1. INTRODUCTION

Artificial Neural Network is one of the methods resulting from the development of the world of information technology that is widely used to predict or forecast [2]. This method uses time series data analysis which is a time sequence observation of quantitative characteristics of events that occur [1, 4]. One of architecture in Artificial Neural Networks is Backpropagation. Backpropagation Artificial Neural Networks consist of a group of small processing units that are modeled in basis of human neural networks [11]. Generally, Backpropagation Architecture uses multi-layer network that aims to minimize errors that will be generated at the network output [10, 5]. The more complex the network is built, the better results of the resulting editions [3].

This research will use the Backpropagation Artificial Neural Network method for Prediction of Life Expectancy (LE) in Maluku Province in the future. Life expectancy data is one measure of life quality of a population [7]. The Indonesian government makes LE as an indicator of the assessment of the degree of public health [12]. This is used as an evaluation material for government work programs in the health sector. The higher of Life Expectancy a region has, the better its public health [6].

2. METHOD

The type of data used in this research is quantitative data. The data used for this research is the Maluku Province Life Expectancy (LE) data from 2010-2018. LE data used in the study were obtained from the Badan Pusat Statistika (BPS) of Maluku Province. Collecting data through direct observation methods in the Badan Pusat Statistika (BPS) of Maluku

Province and literature studies obtained from literature studies on e-books, scientific journals, and websites. The method used is Backpropagation Neural Network method. The following is a flow chart of the training stages in Figure 1 and testing process stages in Figure 2 of the Backpropagation method.

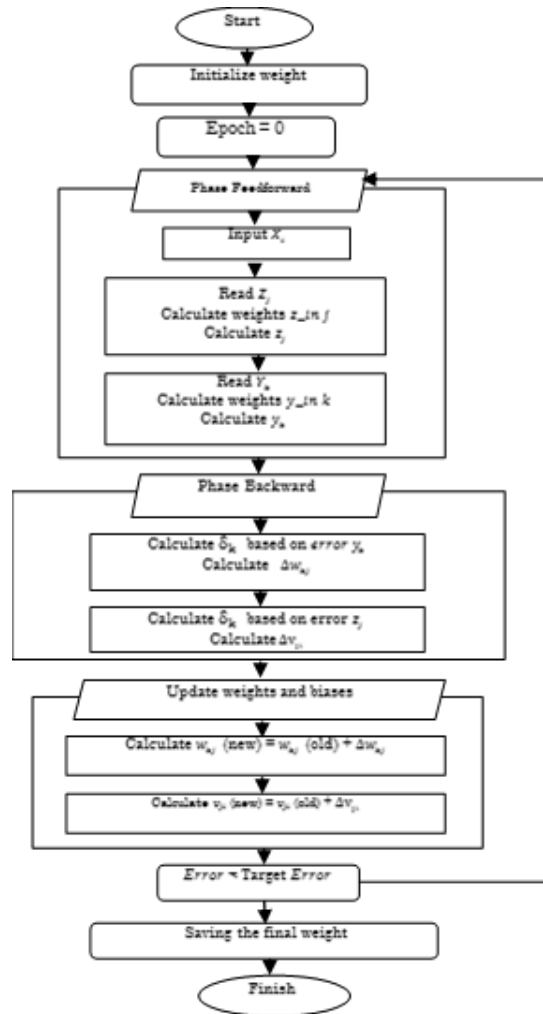


FIGURE 1. Training process flow chart on backpropagation networks[8]

3. RESULT AND DISCUSSION

3.1. Normalization Data. In this section, the data of Life Expectancy (LE) in Maluku Province that obtained from Badan Pusat Statistika (BPS) of Maluku Province shown in Table 1:

LE data in Table 1, contains LE Maluku Province data, which then will be normalized in the form of weighting so it can be used in the process of testing and training data using the Backpropagation method. This is the data preparation stage. Data normalized into interval [0,1; 0.9] using the formula [9], as follows:

$$x_{new} = \frac{0.8(x - x_{min})}{x_{max} - x_{min}} + 0.1 \tag{1}$$

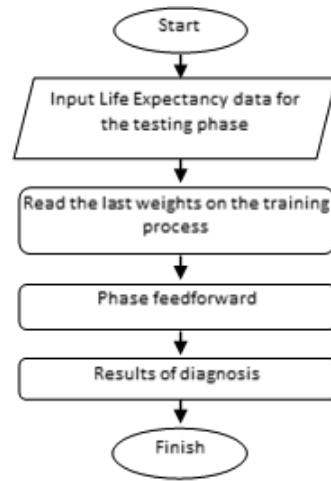


FIGURE 2. Testing process flow chart on backpropagation networks ([8])

TABLE 1. Life Expectancy (LE) Indicator Data

Year	Life Expectancy
2010	64.46
2011	64.61
2012	64.77
2013	64.93
2014	65.01
2015	65.31
2016	65.35
2017	65.4
2018	65.59

TABLE 2. Normalized life expectancy indicator data

Tahun	Life Expectancy
2010	0.1
2011	0.2062
2012	0.3195
2013	0.4327
2014	0.4894
2015	0.7018
2016	0.7301
2017	0.7655
2018	0.9

The data that has been transformed by normalization process above (Table 2) will be used as input parameters and the data has gone through the preprocessing stage. At this stage, the data is formed in time series data with a vulnerable time of 2 years as training data and the third year as target data. This data is then made into a pattern or scheme of data sharing in general as in Table 3, below:

Based on Table 3, we obtain 7 data from LE data which will be divided according to the annual data sharing scheme. This data will then be divided into training data and testing data.

TABLE 3. Data division scheme each year

X1	X2	Target
2010 Data	2011 Data	2012 Data
2011 Data	2012 Data	2013 Data
...
2015 Data	2016 Data	2017 Data
2016 Data	2017 Data	2018 Data

TABLE 4. Distributions of normalized annual life expectancy indicator data

X1	X2	Target
0.1	0.2062	0.3195
0.2062	0.3195	0.4327
0.3195	0.4327	0.4894
0.4327	0.4894	0.7018
0.4894	0.7018	0.7301
0.7018	0.7301	0.7655
0.7301	0.7655	0.9

The data in Table 4, above is then divided into training data and test data with a percentage of 80% (5 data) and 20% (2 data) respectively. Therefore, from the total data of 7 data, it is divided into 5 training data and 2 data of testing data for each LE data and the indicators are given in Table 5-6. This division was taken after going through several stages of testing the distribution of training data and testing data, the best distribution composition was obtained from the accuracy of the prediction results and the smallest error is 80% and 20%.

TABLE 5. Normalized annual life expectancy training data scheme

X1	X2	Target
0.1	0.2062	0.3195
0.2062	0.3195	0.4327
0.3195	0.4327	0.4894
0.4327	0.4894	0.7018
0.4894	0.7018	0.7301

TABLE 6. Normalized annual life expectancy testing data scheme

X1	X2	Target
0.7018	0.7301	0.7655
0.7301	0.7655	0.9

3.2. Research Data Processing Stage - Training. The training stage of Life Expectancy data with ANN Backpropagation is given as follows:

- (1) Network Architecture consists of input, two hidden layers, and output
- (2) Each hidden layer has 3 neurons
- (3) The initial weight of input to the first hidden layer
 $net.IW\{1,1\} =$

0.4 0.4
 0.4 0.4
 0.4 0.4

The initial weights of the first hidden layer to the second hidden layer
 $net.LW\{2,1\} =$

0.4 0.4 0.4
 0.4 0.4 0.4
 0.4 0.4 0.4

The initial weight of the second hidden layer to the output
 $net.LW\{3,2\} =$

0.4 0.4 0.4

The initial bias of the first hidden layer
 $net.b\{1,1\} =$ The initial bias of the first hidden layer

0.4
 0.4
 0.4

$net.b\{2,1\} =$ Initial bias output

0.4
 0.4
 0.4

$net.b\{3,1\} = 0.4$

(4) Establish parameters for the training process

Learning rate=0.1

Target *Error*=0.0000001

Maximum *epoch*=500

The training process is carried out and stops at the 5th epoch with the following results:

3.3. Research Data Processing Stage Testing. After the analysis of training process, then in this part the data is given the analysis of testing stage by Artificial Neural Networks (ANN) Backpropagation: In Table 7, the following is the value of the testing data processing

Original Test Target	ANN Output Target
65.4	65.709
65.59	65.7515

results with MATLAB software. The measurement of forecasting/prediction accuracy is using the following mean absolute percentage error (MAPE) formula:

$$MAPE = \frac{\sum |e_i|}{x_{asti}} (100\%) \tag{2}$$

From the calculation is given the value of MAPE = 0.0035. It is means that the success of predictions/forecasting is 99.65%.

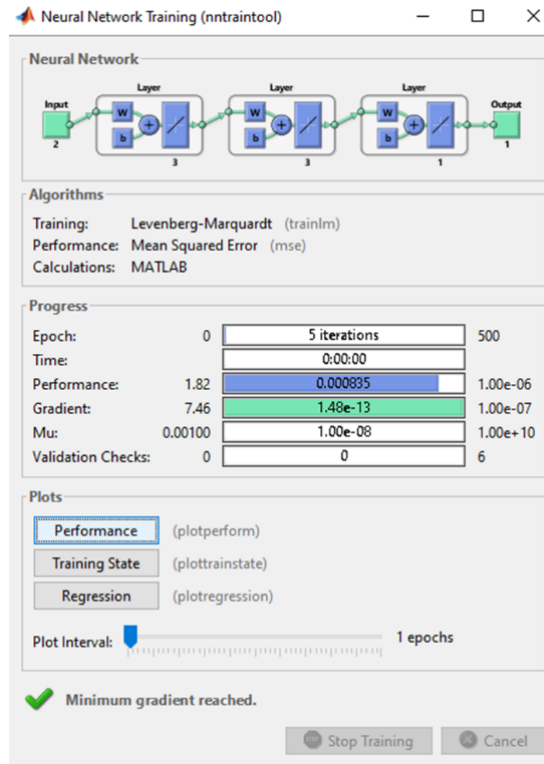


FIGURE 3. Training results graph (training) life expectancy data

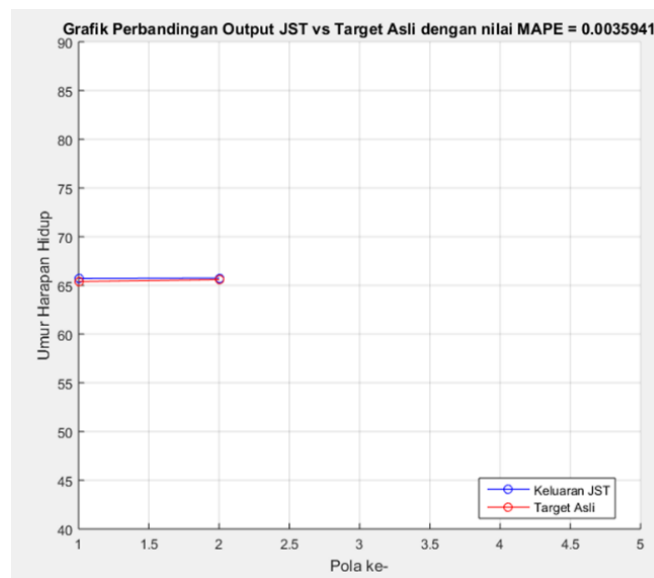


FIGURE 4. The testing result graph for life expectancy data (comparison of prediction result with original targets)

3.4. LE Data Prediction Stages. From the results of the testing stages in Table 7 and by using a network that has been trained, we can predict the Life Expectancy in Maluku Province for the next 5 years from 2019-2023. Prediction results can be obtained from the processing using Matlab Software, as follows:

TABLE 7. Predicted results of life expectancy data for years 2019-2023

Year	Prediction Result of Life Expectancy
2019	65.7828
2020	65.9924
2021	66.2027
2022	66.4344
2023	66.6632

Based on the prediction results in Table 8 can be concluded that the value of Life Expectancy tends to increase over the next 5 years started at 65.7828 in 2019 and will increasing to 66.6632 in 2023.

4. CONCLUSION

Based on the research it can be concluded that the average prediction/forecasting success is 99.65%, with MAPE = 0.0035, and the learning rate value is 0.1. The architecture of this research network uses input layer 2 neurons, hidden layer 3-3-1, and output layer 1 neurons, with epoch/iteration: 500. In addition, the Forecasting results Life expectancy (LE) tends to increase over 5 years from 2019-2023 at 65.7828 in 2019 and will increasing to 66.6632 in 2023. It means that the government of Maluku Province has improved the health status of the community, including improving access to and quality of health services.

REFERENCES

- [1] Igor Aizenberg, Leonid Sheremetov, Luis Villa-Vargas, and Jorge Martinez-Muñoz. Multi-layer neural network with multi-valued neurons in time series forecasting of oil production. *Neurocomputing*, 175:980–989, 2016.
- [2] Imad A Basheer and Maha Hajmeer. Artificial neural networks: fundamentals, computing, design, and application. *Journal of microbiological methods*, 43(1):3–31, 2000.
- [3] Abdollahi Ghaffari, H Abdollahi, MR Khoshayand, I Soltani Bozchalooi, A Dadgar, and M Rafiee-Tehrani. Performance comparison of neural network training algorithms in modeling of bimodal drug delivery. *International journal of pharmaceuticals*, 327(1-2):126–138, 2006.
- [4] KI Hoi, KV Yuen, and KM Mok. Prediction of daily averaged pm10 concentrations by statistical time-varying model. *Atmospheric Environment*, 43(16):2579–2581, 2009.
- [5] Joarder Kamruzzaman and Ruhul A Sarker. Ann-based forecasting of foreign currency exchange rates. *Neural Information Processing-Letters and Reviews*, 3(2):49–58, 2004.
- [6] Khomarun Khomarun, Maharso Adhi Nugroho, and Endang Sri Wahyuni. Pengaruh aktivitas fisik jalan pagi terhadap penurunan tekanan darah pada lansia dengan hipertensi stadium i di posyandu lansia desa makamhaji. *Interest: Jurnal Ilmu Kesehatan*, 3(2), 2014.
- [7] J Koto. Pengaruh angka melek huruf dan angka harapan hidup terhadap jumlah penduduk miskin di propinsi sumatera barat. *economica*, 2(2):126–133.
- [8] Yopi Andry Lesnussa, CG Mustamu, F Kondo Lembang, and MW Talakua. Application of backpropagation neural networks in predicting rainfall data in ambon city. *International Journal of Artificial Intelligence Research*, 2(2):41–50, 2018.
- [9] Darnisa Azzahra Nasution, Hidayah Husnul Khotimah, and Nurul Chamidah. Perbandingan normalisasi data untuk klasifikasi wine menggunakan algoritma k-nn. *Computer Engineering, Science and System Journal*, 4(1):78–82, 2019.
- [10] H Hasan Örkü and Hasan Bal. Comparing performances of backpropagation and genetic algorithms in the data classification. *Expert systems with applications*, 38(4):3703–3709, 2011.

- [11] Bhavna Sharma and K Venugopalan. Comparison of neural network training functions for hematoma classification in brain ct images. *IOSR Journal of Computer Engineering*, 16(1):31–35, 2014.
- [12] Ayuk Putri Sugiantari and I Nyoman Budiantara. Analisis faktor-faktor yang mempengaruhi angka harapan hidup di jawa timur menggunakan regresi semiparametrik spline. *Jurnal Sains dan Seni ITS*, 2(1):D37–D41, 2013.