Vol.7 No.1 Tahun 2026
Jurnal Kecerdasan Buatan, Komputasi dan Teknologi Informasi
https://ejournal.unuja.ac.id/index.php/core
E-ISSN : 2774-7875 and P-ISSN : 2775-0124

Forecasting the Population of Balikpapan City Using the Artificial Neural Network Method

Vira Oktavia[1], Andri Azmul Fauzi[2], Yuki Novia Nasution[3]
[1,2,3] Universitas Mulawarman, Indonesia

Article Info

Article history:

Received : 03 09, 2026
Revised : 04 29, 2026
Accepted : 05 25, 2026

Keywords:

Artificial Neural Network;
Backpropagation;
Population Prediction;
Population Growth;
Balikpapan City;

Abstract

The continuous growth of the population requires prediction methods capable of generating accurate population estimates as a basis for development planning. This study aims to predict the population of Balikpapan City for the period 2025–2029 using an Artificial Neural Network (ANN) with the Backpropagation algorithm. The dataset consists of annual population data of Balikpapan City from 2010 to 2024. The data were processed using the sliding window technique, followed by normalization, model training, and testing. The proposed ANN model employed a 3–4–1 architecture with a learning rate of 0,5, a sigmoid activation function in the hidden layer, a linear activation function in the output layer, and 150 training epochs. The model performance was evaluated using the Mean Absolute Percentage Error (MAPE). The experimental results show that the proposed model produces predictions that closely match the actual population data, achieving a MAPE value of 0,57%, which indicates a high level of prediction accuracy. The trained model was subsequently used to forecast the population of Balikpapan City for the period 2025–2029. The prediction results are expected to provide useful references for local governments in formulating policies and development planning that are aligned with future population growth.

Corresponding Author:

Andri Azmul Fauzi,
andriazmul161022@fmipa.unmul.ac.id

INTRODUCTION

Population growth is one of the key factors influencing the success of regional development. An increasing population affects various sectors, including the economy, education, healthcare, and infrastructure (Rachma et al., 2025). Rapid population growth may also lead to various demographic challenges that influence the social and economic conditions of a region (Sari et al., 2023). Therefore, governments are required to formulate comprehensive development strategies to ensure that public needs can be sustainably fulfilled. According to Statistics Indonesia (BPS), Indonesia's population is projected to reach approximately 284 million in 2025, accompanied by demographic changes that require appropriate public policy adjustments. Demographic trend analysis is essential not only for supporting the achievement of the Sustainable Development Goals (SDGs) but also for facilitating long-term development planning (Danil & Rochayati, 2025).

Balikpapan City is one of the major cities in East Kalimantan Province and plays a strategic role, particularly following the development of Indonesia's New Capital City (Ibu Kota Nusantara/IKN) in the surrounding region. This development is expected to stimulate population growth through migration and other non-natural demographic factors (Muliati et al., 2025). The relocation of the national capital has also accelerated economic activities in Balikpapan, especially in the service and hospitality sectors, while encouraging infrastructure development. Furthermore, increasing interactions among people from diverse backgrounds are expected to bring social and cultural changes to the local community (Danil & Rochayati, 2025).

Population projection is not merely an estimation of future population size but also an essential instrument for development planning. It is based on three fundamental demographic components, namely fertility, mortality, and migration, whose future patterns can be estimated to predict population growth in subsequent periods (Karyana & Rusliana, 2021). Consequently, data-driven planning has become increasingly important in supporting sustainable regional development. One of its essential components is population projection, which enables local governments to anticipate future demographic dynamics more effectively (Sianturi et al., 2025).

Artificial Neural Networks (ANNs) are machine learning algorithms designed to simulate the learning mechanism of the human brain in solving complex problems. Learning is achieved through iterative adjustments of synaptic weights based on historical data, enabling ANNs to recognize patterns, classify information, and generate predictions for previously unseen data (Yuberta, 2022). ANNs possess adaptive learning capabilities, allowing them to continuously learn from available data and adjust to changing patterns. In general, ANN architectures are classified into two main categories: single-layer networks and multilayer networks (Hardoyo & Parmadi, 2022). In a single-layer network, input neurons are directly connected to output neurons without hidden layers, whereas multilayer networks consist of input, hidden, and output layers, making them more suitable for solving complex prediction problems (Suahati et al., 2022).

The Backpropagation algorithm is one of the most widely used supervised learning algorithms for training multilayer neural networks. It computes the gradient of the error function with respect to network weights using the chain rule of differentiation (J. E. Napitupulu & Solikhun, 2024). Through this learning mechanism, the network iteratively updates its weights to minimize prediction errors, thereby improving model performance (P. N. Napitupulu & Safii, 2024). The training process consists of two main stages: forward propagation, which produces the network output, and backward propagation, which calculates prediction errors and updates the network parameters. The performance of the model is strongly influenced by several hyperparameters, including the learning rate, activation functions, number of hidden neurons, and number of training epochs (Milawati et al., 2025).

The sliding window technique is a widely used preprocessing approach for transforming time-series data into supervised learning datasets suitable for machine learning models (Radho et al., 2021). This technique constructs input-output pairs by continuously shifting a fixed-size window across the chronological data sequence, thereby preserving temporal dependencies while enabling the model to learn historical patterns effectively (Dwi Kartini et al., 2021). In Artificial Neural Networks, each input vector consists of observations from several previous time periods, while the corresponding target represents the subsequent observation. Consequently, the sliding window technique plays an important role in constructing meaningful input patterns that improve the effectiveness of the ANN training process (Setiyorini & Rianto, 2025).

Population growth has become an increasingly important consideration in regional development planning, highlighting the need for accurate prediction methods. Although Artificial Neural Networks have been widely applied to forecasting problems with promising accuracy, studies focusing on population prediction for Balikpapan City remain limited. Therefore, this study applies an Artificial Neural Network trained using the Backpropagation algorithm to predict the population of Balikpapan City based on historical data from 2010–2024, with the objective of forecasting population growth for the 2025–2029 period.

METHOD

Dataset

This study used annual population data of Balikpapan City obtained from Statistics Indonesia (BPS) of East Kalimantan Province. The dataset covers the period from 2010 to 2024 and consists of the total population recorded each year. The historical data served as the primary dataset for training and evaluating the Artificial Neural Network model.

Table 1. Annual Population Data of Balikpapan City

No.	Year	Population
1.	2010	557.600
2.	2011	579.100
3.	2012	596.000
4.	2013	594.300
5.	2014	605.100
6.	2015	615.600
7.	2016	626.000
8.	2017	636.000
9.	2018	645.700
10.	2019	655.200
11.	2020	688.300
12.	2021	695.300
13.	2022	703.600
14.	2023	710.000
15.	2024	717.200

Research Procedure

This research employed a quantitative approach using an Artificial Neural Network (ANN) trained with the Backpropagation algorithm. Prior to the training process, the historical population data were preprocessed using the sliding window technique to transform the time-series data into supervised learning samples.

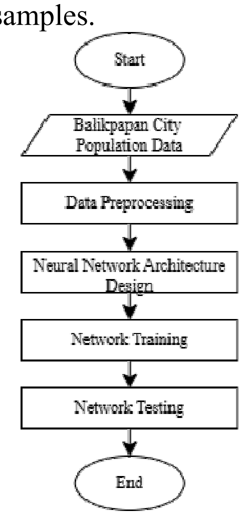

Figure 1. Research Flowchart

The study was conducted from November 2025 to April 2026 at the Computational Mathematics Laboratory and the Mathematical Modeling and Machine Learning Laboratory, Department of Mathematics, Universitas Mulawarman. Initially, the dataset was divided into 80% training data and 20% testing data. The data were then normalized using the min–max normalization method before being transformed into input–target pairs through the sliding window technique.

The proposed ANN model adopted a 3–4–1 architecture, consisting of three input

DOI: https://doi.org/10.33650/coreai.v6i2

neurons, four hidden neurons, and one output neuron. The three input neurons represent the population data from the previous three years, which are used to predict the population of the following year. The hidden layer contains four neurons, while the output layer consists of a single neuron that produces one prediction value for each forecasting process. To generate population forecasts for the 2025–2029 period, the prediction process was performed iteratively. Specifically, the prediction generated for one year was subsequently used as one of the input values to predict the population of the following year. During the training stage, the Backpropagation algorithm was implemented using the logsig (logistic sigmoid) activation function in the hidden layer and the purelin (linear) activation function in the output layer. The optimal network parameters obtained during the training process were subsequently used for model testing and future population forecasting.

RESULTS AND ANALYSIS

Normalisasi Data

Before the training process, the population data were normalized to ensure that all input values were within the same numerical range. In this study, min–max normalization was applied to transform the original data into the interval 0,1–0,9, which is suitable for the ANN using the logistic sigmoid activation function. The normalization process can be expressed as follows:

$$x' = \left(\frac{(0{,}8)(x - x_{min})}{x_{max} - x_{min}}\right) + 0{,}1 \qquad (1)$$

Where :

x' : Normalized population value

x : Original population value

x_{min} : Minimum value of the population dataset

x_{max} : Maximum value of the population dataset

Weight Initialization

$$weight\ v = (n_{input} + 1) \times n_{hidden} \qquad (2)$$

$$weight\ w - (n_{hidden} + 1) \times n_{output} \qquad (3)$$

Network Architecture

Based on the selected numbers of neurons in the input, hidden, and output layers, the architecture of the proposed Backpropagation Artificial Neural Network is presented in Figure 2.

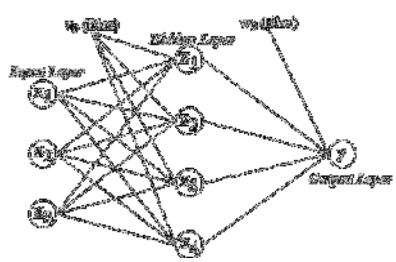

Figure 2. Backpropagation Artificial Neural Network Architecture

Feedforward Phase

Computation of Hidden Layer Neurons

The output of each neuron in the hidden layer is computed by first calculating the weighted sum of the input signals. The net input to the hidden neuron is determined using the following equation:

$$z_{in_j} = v_{0j} + \sum_{i=1}^{n} x_i v_{ij} \quad (4)$$

After obtaining the net input value, the output of each hidden neuron is calculated using the logistic sigmoid (logsig) activation function as follows:

$$z_j = f(z_{in_j}) = \frac{1}{1 + e^{-z_{in_j}}} \quad (5)$$

Where :

v_{0j} : Bias of the j-th hidden neuron
x_i : Input of the i-th neuron
v_{ij} : Weight connecting the i-th input neuron to the j-th hidden neuron
z_{in_j} : Net input received by the j-th hidden neuron
z_j : Output of the j-th hidden neuron
$f(z_{in_j})$: Activation function

Computation of Output Layer Neurons

The output layer receives signals from the hidden layer. The net input of the output neuron is calculated using the following equation:

$$y_{in_k} = w_{0k} + \sum_{j=1}^{n} z_j w_{jk} \quad (6)$$

The final network output is then obtained by applying the purelin (linear) activation function:

$$y_k = y_{in_k} \quad (7)$$

Where :

w_{0k} : Bias of the output neuron
w_{jk} : Weight connecting the j-th hidden neuron to the output neuron
y_{in_k} : Net input of the output neuron
y_k : Network output

Backpropagation Phase

Error Propagation from the Output Layer to the Hidden Layer

This stage updates the weights connecting the hidden layer to the output layer. The correction factor for the output neuron is computed using the following equation:

$$\delta_k = t_k - y_k \quad (8)$$

The weight correction is determined as follows:

$$\Delta w_{jk} = \alpha \delta_k z_j \quad (9)$$

The bias correction is calculated using the following equation:

$$\Delta w_{0k} = \alpha \delta_k \quad (10)$$

Where :

δ_k : Error correction factor of the output neuron

t_k : Target output

y_k : Predicted output

Δw_{jk} : Correction applied to the weight connecting the hidden neuron to the output neuron

Δw_{0k} : Bias correction of the output neuron

α : Learning rate

Error Propagation from the Hidden Layer to the Input Layer

The next step is to update the weights connecting the input layer and the hidden layer. The error at each hidden neuron is calculated using the following equation:

$$\delta_{in_j} - \sum_{k=1}^{n} \delta_k w_{jk} \quad (11)$$

The weight correction is obtained as follows:

$$\delta_j = \delta_{in_j} f'\left(z_{in_j}\right) = \delta_{in_j} z_j \left(1 - z_j\right) \quad (12)$$

The weights are updated using the following equation:

$$\Delta v_{ij} = \alpha \delta_j x_i \quad (13)$$

The hidden-layer bias correction is computed using:

$$\Delta v_{0j} = \alpha \delta_j \quad (14)$$

Where :

δ_{in_j} : Propagated error signal from the output layer

δ_j : Error correction factor of the hidden neuron

Δv_{ij} : Correction applied to the weight connecting the input neuron to the hidden neuron

Δv_{0j} : Bias correction of the hidden neuron

<u>Weight and Bias Update</u>

After the correction values have been obtained, the network weights and biases are updated. For the output layer, the updated weights and biases are computed using the following equations:

$$w_{jk}(new) = w_{jk}(old) + \Delta w_{jk} \quad (15)$$

$$w_{0k}(new) = w_{0k}(old) + \Delta w_{0k} \quad (16)$$

Where:

$w_{jk}(new)$: Updated weight connecting the hidden neuron to the output neuron

$w_{jk}(old)$: Previous weight connecting the hidden neuron to the output neuron

$w_{0k}(new)$: Updated output-layer bias

$w_{0k}(old)$: Previous output-layer bias

Δw_{jk} : Weight correction

Δw_{0k} : Bias correction

Similarly, the weights and biases between the input layer and the hidden layer are updated using the following equations:

$$v_{ij}(new) = v_{ij}(old) + \Delta v_{ij} \quad (17)$$

$$v_{0j}(new) = v_{0j}(old) + \Delta v_{0j} \quad (18)$$

Where:

$v_{ij}(new)$: Updated weight connecting the input neuron to the hidden neuron
$v_{ij}(old)$: Previous weight connecting the input neuron to the hidden neuron
$v_{0j}(new)$: Updated hidden-layer bias
$v_{0j}(old)$: Previous hidden-layer bias
Δv_{ij} : Weight correction
Δv_{0j} : Bias correction

Stopping Criterion

The training process is terminated once the predefined maximum number of epochs has been reached. The final optimized weights and biases obtained at the end of the training process are subsequently used during the testing and forecasting stages.

Denormalization

$$\hat{x} = \left(\frac{(x - 0{,}1)(x_{max} - x_{min})}{0{,}8}\right) + x_{min} \qquad (19)$$

Where :
$\hat{x}$: Denormalized population value
x : Normalized prediction value
x_{max} : Maximum population value
x_{min} : Minimum population value

Mean Absolute Percentage Error (MAPE)

$$MAPE = \frac{1}{n}\sum_{i-1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \times 100\%$$

Keterangan:
n : Number of testing samples
y_i : Actual value of the i-th observation
$\hat{y}_i$: Predicted value of the i-th observation

The training process began by normalizing the historical population data using the min–max normalization method. In this study, 80% of the population data from 2010 to 2024 were used as the training dataset. The normalization transformed the original data into the range of 0,1–0,9, making them suitable for the Artificial Neural Network model. The normalized training data are presented in Table 2.

Table 2. Normalized Training Data

No.	Year	Normalized Population
1.	2010	0,1
2.	2011	0,21
3.	2012	0,29
4.	2013	0,28
5.	2014	0,34
6.	2015	0,39
7.	2016	0,44
8.	2017	0,49
9.	2018	0,54
10.	2019	0,59
11.	2020	0,76
12.	2021	0,79

DOI: https://doi.org/10.33650/coreai.v6i2

The normalized training data were subsequently transformed into input–target pairs using the sliding window technique. This preprocessing step enables the neural network to learn temporal relationships from historical observations. The resulting training dataset is presented in Table 3.

Table 3. Training Dataset

X_1	X_2	X_3	Target
0,1	0,21	0,29	0,28
0,21	0,29	0,28	0,34
0,29	0,28	0,34	0,39
0,28	0,34	0,39	0,44
0,34	0,39	0,44	0,49
0,39	0,44	0,49	0,54
0,44	0,49	0,54	0,59
0,49	0,54	0,59	0,76
0,54	0,58	0,76	0,79

The input–target pairs generated through the sliding window technique were then used to train the Artificial Neural Network using the Backpropagation algorithm. During the training process, the network iteratively updated its weights and biases until the predefined stopping criterion was reached. The optimized weights and biases obtained during this stage were subsequently used to evaluate the model on unseen testing data.

The testing phase employed the remaining 20% of the dataset, which had not been used during training. Similar to the training data, the testing dataset was first normalized into the range of 0,1–0,9 before being transformed into input–target pairs using the sliding window technique. The normalized testing data and the corresponding testing dataset are presented in Tables 4 and 5, respectively.

Table 4. Normalized Testing Data

No.	Tahun	Normalized Population
1.	2022	0,83
2.	2023	0,86
3.	2024	0,9

Table 5. Testing Dataset

X_1	X_2	X_3	Target
0,59	0,76	0,79	0,83
0,76	0,79	0,83	0,86
0,79	0,83	0,86	0,9

Using the optimized weights and biases obtained during training, the trained model generated predictions for the testing dataset. The prediction results are presented in Table 6.

Table 6. Testing Results

Year	Testing Results
2022	702.178
2023	716.442
2024	721.489

The prediction performance of the proposed model was evaluated using the Mean Absolute Percentage Error (MAPE) by comparing the predicted values with the corresponding actual population data. The evaluation results are summarized in Table 7.

Table 7. MAPE Evaluation Results

Year	Actual Value	Predicted Value	Prediction Error	MAPE (%)
2022	703.600	702.417	702.178	0,2
2023	710.000	711.335	716.442	0,91

2024	717.200	716.582	721.489	0,6
		Rata-rata MAPE		0,57

The evaluation results indicate that the proposed Artificial Neural Network model achieved a MAPE value of 0,57%. According to commonly accepted forecasting accuracy criteria, a MAPE value below 10% indicates highly accurate forecasting performance. Therefore, the obtained MAPE demonstrates that the proposed model is capable of accurately capturing the historical population trend of Balikpapan City and producing reliable predictions.

After achieving satisfactory performance during the testing stage, the trained Artificial Neural Network model was employed to forecast the population of Balikpapan City for the period 2025–2029. The prediction results are presented in Table 8.

Table 8. Population Forecasts for 2025–2029

Year	Population
2025	726.816
2026	732.143
2027	737.589
2028	742.317
2029	745.669

CONCLUSIONS

This study demonstrates that an Artificial Neural Network trained using the Backpropagation algorithm can effectively be applied to predict the population of Balikpapan City. Based on the testing results, the proposed model achieved a Mean Absolute Percentage Error (MAPE) of 0,57%, indicating a high level of forecasting accuracy. Furthermore, the forecasting results suggest that the population of Balikpapan City is expected to continue increasing throughout the 2025–2029 period. These findings indicate that the proposed model is capable of capturing historical population patterns and generating reliable future population estimates. However, this study is limited to the use of historical population data without considering other demographic factors that may influence population growth. Future research is therefore recommended to incorporate a longer historical dataset as well as additional demographic variables, such as birth rates, mortality rates, and migration, in order to improve forecasting accuracy.

REFERENCES

Danil, D., & Rochayati, N. (2025). Analisis Dinamika Penduduk dan Perubahan Komposisi Demografi di Indonesia: Implikasi bagi Perencanaan Pembangunan. Education, Social Sciences, and Linguistics: Conference Series, 1(2), 106–111.

Dwi Kartini, Friska Abadi, & Triando Hamonangan Saragih. (2021). Prediksi Tinggi Permukaan Air Waduk Menggunakan Artificial Neural Network Berbasis Sliding Window. Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi), 5(1), 39–44. https://doi.org/10.29207/resti.v5i1.2602

Hardoyo, T., & Parmadi, E. H. (2022). Klasifikasi Usaha Mikro Kecil Menengah Menggunakan Jaringan Syaraf Tiruan Backpropagation. KONSTELASI: Konvergensi Teknologi Dan Sistem Informasi, 2(1). https://doi.org/10.24002/konstelasi.v2i1.5625

Karyana, Y., & Rusliana, N. (2021). Proyeksi Penduduk Jawa Barat Tahun 2025 â€“ 2035 Menggunakan Metode Campuran dengan Data Dasar Sensus Penduduk 2020. WELFARE Jurnal Ilmu Ekonomi, 2(1), 26–35. https://doi.org/10.37058/wlfr.v2i1.2824

Milawati, Sofinah, L., Nabila, P. S., & Maghfira, Z. (2025). Optimasi Arsitektur Jaringan Syaraf Tiruan untuk Prediksi Curah Hujan Berdasarkan Data Meteorologi Indonesia. Saturnus: Jurnal Teknologi Dan Sistem Informasi, 3(1), 167–180. https://doi.org/10.61132/saturnus.v3i1.1274

Muliati, Gaffar, E. U. A., Aprianti, Y., Auliansyah, Cahyana, I., & Sulindrina, A. (2025). Penyuluhan Pengembangan Kesejahteraan Ditinjau dari Komposisi Kependudukan di Kota Balikpapan. Jurnal Pengabdian Kepada Masyarakat, 4(1), 32–38.

Napitupulu, J. E., & Solikhun, S. (2024). Implementasi Algoritma Backpropagation Dalam Prediksi Laju Pertumbuhan Penduduk Di Kabupaten Sinjai. Jurnal Manajemen, Pendidikan Dan Ilmu Komputer, 1(1), 15–21. https://doi.org/10.65309/evgesw76

Napitupulu, P. N., & Safii, M. (2024). PENERAPAN ALGORITMA BACKPROPOGATION DALAM PREDIKSI JUMLAH PENDUDUK DI PROVINSI SUMATERA UTARA. Jurnal Manajamen Informatika Jayakarta, 4(1), 15–27. https://doi.org/10.52362/jmijayakarta.v4i1.1300

Rachma, H., Nafisah, M., & Hidayati, N. (2025). ESTIMASI LAJU PERTUMBUHAN PENDUDUK DI KABUPATEN JEPARA DENGAN PENDEKATAN REGRESI LINIER BERGANDA. Jurnal Sistem Informasi Dan Informatika (Simika), 8(1), 57–66. https://doi.org/10.47080/simika.v8i1.3790

Radho, A. E., Sugiartawan, P., & Santiago, G. A. (2021). Prediksi Jumlah Kasus COVID-19 Menggunakan Teknik Sliding Wondows dengan Metode BPNN. Jurnal Sistem Informasi Dan Komputer Terapan Indonesia (JSIKTI), 4(1), 12–23. https://doi.org/10.33173/jsikti.123

Sari, A. P., Rahmadini, G., Charlina, H., Pradani, Z. E., & Ramadan, M. I. (2023). ANALISIS MASALAH KEPENDUDUKAN DI INDONESIA. Journal of Economic Education, 2(1), 29–37. https://doi.org/10.22437/jeec.v2i1.23180

Setiyorini, T., & Rianto, H. (2025). Perbandingan Neural Network dan K-Fold Cross Validation dengan Neural Network dan Sliding Window Validation untuk Estimasi Kuat Tekan Beton. Jurnal Nasional Komputasi Dan Teknologi Informasi (JNKTI), 8(1), 1–10.

Sianturi, M. D., Lubis, M. C., Sinaga, G. T., & Manihuruk, J. N. (2025). Prediksi Pertumbuhan Jumlah Penduduk Indonesia Menggunakan Interpolasi Polinomial Lagrange. J-CEKI : Jurnal Cendekia Ilmiah, 4(2), 1315–1321. https://doi.org/10.56799/jceki.v4i2.6766

Suahati, A. F., Nurrahman, A. A., & Rukmana, O. (2022). Penggunaan Jaringan Syaraf Tiruan – Backpropagation dalam Memprediksi Jumlah Mahasiswa Baru. Jurnal Media Teknik Dan Sistem Industri, 6(1), 21–29. https://doi.org/10.35194/jmtsi.v6i1.1589

Yuberta, A. (2022). Jaringan Syaraf Tiruan dengan Algoritma Backpropagation dalam Memprediksi Hasil Asesmen Nasional Berbasis Komputer (ANBK) SMP Se Kota Sawahlunto. Jurnal Informasi dan Teknologi, 200–205. https://doi.org/10.37034/jidt.v4i4.234