Journal of Artificial Intelligence and Engineering Applications

Website: https://ioinformatic.org/

15th October 2025. Vol. 5. No. 1; e-ISSN: 2808-4519

Predicting Population Growth Using Artificial Neural Network by Method Backpropagation (Case Study: Durian Lingga Village)

Della Oksa Brina Br Surbakti [1], Darjat Saripurna [2], Rusmin Saragih [3]

[1,2,3]*STMIK Kaputama*

Dellaoksa18@gmail.com, darjatsaripurna@gmail.com, evitha12014@gmail.com

Abstract

This study aims to predict population growth in Durian Lingga Village using the Artificial Neural Network (JST) Backpropagation method. The background of this research is based on a significant increase in population from 2005 to 2024, which poses various challenges in village development planning. The data used in this study is data on the population of Durian Lingga Village from the Central Statistics Agency (BPS) of Langkat Regency during the period 2005–2024. The research method used is quantitative with an experimental approach, where the data is divided into training data and test data, then processed using a 2-5-1 network architecture with a *learning rate* of 0.05, *a target error* of 0.02, and a maximum of 1000 epochs.

***Keywords*:** *Artificial Neural Network, Backpropagation, Prediction, Population Growth, Durian Lingga Village*

1. Introduction

Artificial Neural Network (JST) or Artificial Neural Network Backpropagation is artificial intelligence that mimics the way the human brain works in solving a problem. With its ability to recognize patterns and study historical data, artificial neural networks have been applied in various fields including predicting population growth. Optimization of artificial neural networks using the backpropagation method is able to predict the number of national population growth with an accuracy rate of 97% [1].

Application of Backpropagation Artificial Neural Network to Predict Population Growth in Magelang City [2]. Artificial Neural Network to predict clothing sales using a backpropagation algorithm [3].

Based on the problems that have been described above, the title raised in this study is "Predicting Population Growth Using Artificial Neural Network Using the Backpropagation Method (Case Study: Durian Lingga Village)"

2. Literatur Riview

2.1. Definition of Artificial Neural Network

Artificial Neural Network (JST) is an information processing system that has characteristics similar to Biological Neural Network (JSB). Artificial neural networks are created as a mathematical generation of human understanding. JST is also known as an adaptive network because it can change its structure to solve problems based on external and internal data sources flowing through the network. Artificial Neural Network is a computational method that mimics biological neural networks, this method uses basic non-linear computational elements called neurons that are organized as interconnected networks so that they are similar to human neural networks. Artificial neural networks are formed to solve a specific problem such as pattern recognition or classification due to the learning process [4]. Artificial neural networks are not designed to produce predetermined outputs. Instead, the results or conclusions produced by the network depend on the experience gained during the learning process.

2.2. Backpropagation

Backpropagation is a widely used method to solve a problem related to prediction, identification, and pattern recognition. Backpropagation is a method or algorithm that is included in suvervised learning that has many layers or what is often called multilayer perceptron which is commonly used by perceptron in making weight changes that are interconnected with neurons in the hidden layer [5]. This training stage

is a step to train an artificial neural network, namely by changing weight, while problem solving will be carried out when the training process has been completed, this phase is called the testing phase.

2.3. Definition of Prediction

Forecasting is one of the methods used to predict something that has not yet happened [6]. Prediction is an activity of predicting in a structured manner about something that may happen in the future based on information from the past and present, so that the errors can be minimized. Based on the meaning of prediction, it can be concluded that prediction is a forecasting activity or estimation of a situation in the future to carry out a test taken from past data with the aim of minimizing errors that occur.

2.4. Definition of Population

According to the central statistics agency, population is a person who resides in an area for a certain period of time, both permanent and temporary. The population is calculated based on a census conducted periodically to determine the number of residents. Meanwhile, according to the Great Dictionary of the Indonesian Language (KBBI), a resident is a person who inhabits a place. Based on these two definitions, it can be concluded that population growth is a number that indicates an increase in the number of people in an area. Population growth is influenced by several factors, namely births, deaths, and migration.

2.5. Matlab

Mathwork inc created MATLAB, a very helpful and useful program for solving various numerical problems. The software simplifies and simplifies the resolution of numerical problems by using vectors and matrices Solving linear equations and obtaining matrix inverses are two examples of problems that can be solved quickly and simply [7].

3. Analysis And Design

3.1. Research Methodology

The method used in this study is included in the experimental quantitative approach, because it uses an artificial neural network model with the backpropagation method to perform numerical simulations in predicting the population based on historical data. The design process carried out in this study began by determining the working stages of the Artificial Neural Network (JST) in order to obtain more accurate prediction results based on population growth data patterns.

3.2. System Requirement Analysis

a. Analisis Backpropagation

The backpropagation method is a systematic method in artificial neural networks using supervised learning algorithms, and is generally applied to perceptrons with many layers to update the weights on the hidden layers.
In making data predictions, historical data is needed that serves as a basis for analysis and calculation methods, so that the best alternatives can be obtained based on the available data. In this study, the data used is data on the population of Durian Lingga Village from 2005 to 2024, which was obtained from the langkatkab.bps.go.id website.

b. Define variables

To test the backpropagation method, the variables used are as follows:

- The input variable consists of two neuronal inputs.
- The number of neurons is hidden as much as 5.
- The output variable consists of one neuron, i.e. Y1.

3.3. Model Implementation

In this study, the flowchart is designed to describe how the system performs the process of predicting population growth data in Durian Lingga Village. The flowchart design used can be seen in the following explanation:

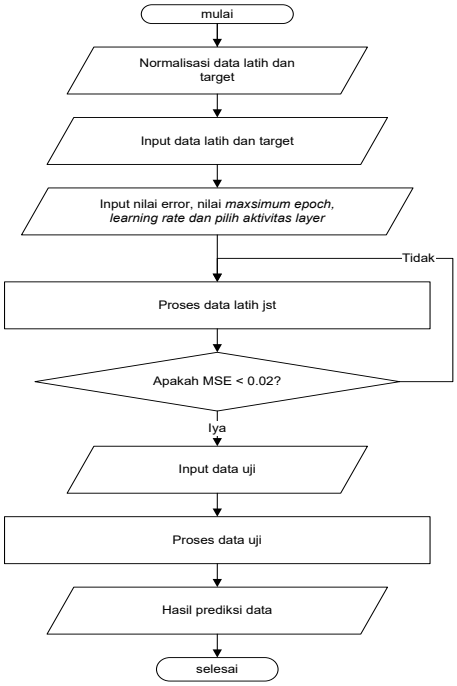

Fig. 1: Flowchart

Fig. 1 The description of the flowchart of the data prediction system above is as follows:

- Start, signifying the beginning of the prediction system
- Furthermore, the process of normalizing the data to a standard scale for example (0-1) to make the training more effective
- Then, input training data and targets
- Determine JST maintenance parameters such as error values, maxsimum epoch, learning rate, and layer activity
- Data processing is calculated using the backpropagation method
- Check if the MSE error is smaller than 0.02. If not, repeat the training process. If so, proceed to the testing process
- Input test data
- JST processes test data based on the trained weights
- After the test data process is completed, the results of the data prediction by the system will be displayed. The process is complete.

3.3.1. Data Management of Durian Lingga Village Residents

Table 1: Data of durian lingga villagers

No	Year	Man	Woman	Total
1	2005	1305	1291	2596
2	2006	896	924	1820
3	2007	901	944	1845
4	2008	911	962	1873
5	2009	922	976	1898
6	2010	926	956	1882
7	2011	931	961	1892
8	2012	932	961	1893
9	2013	932	964	1896
10	2014	958	991	1949
11	2015	968	998	1966
12	2016	978	1006	1984
13	2017	981	1016	1997
14	2018	987	1023	2010
15	2019	994	1029	2023
16	2020	948	1031	1979
17	2021	1022	1128	2150
18	2022	1091	1171	2262
19	2023	1083	1173	2256
20	2024	1080	1184	2264

In analyzing the backpropagation method, there are several stages that need to be done, the stages in the process of analyzing data prediction using the backpropagation method are as follows:

a. Defining Drill Data and Test Data
The training data was used to train the model with the backpropagation method, while the test data was used to test the model's predictions. The input data consists of the number of male and female populations while the target data is the total population, which is predicted by the model to study population growth patterns. The following is the distribution of training data and test data:

Table 2: Training Data and Test Data

Training Data				
No.	**Year**	**Man**	**Woman**	**Total**
1	2005	1305	1291	2596
2	2006	896	924	1820
3	2007	901	944	1845
4	2008	911	962	1873
5	2009	922	976	1898
6	2010	926	956	1882
7	2011	931	961	1892
8	2012	932	961	1893
9	2013	932	964	1896
10	2014	958	991	1949
11	2015	968	998	1966
12	2016	978	1006	1984
13	2017	981	1016	1997
14	2018	987	1023	2010
15	2019	994	1029	2023
16	2020	948	1031	1979
Test Data				
17	2021	1022	1128	2150
18	2022	1091	1171	2262
19	2023	1083	1173	2256
20	2024	1080	1184	2264

b. Normalize population data
The process of normalizing training data and test data using equations $X_{norm} = \frac{X - X_{min}}{Xmax - Xmin} where\ x_{norm}$ is the value of the result normalization, x is the value that will be normalized, x_{min} is the minimum value of the data, and x_{max} is the maximum value of the data. The following are the results of the normalization process:

Table 3: NormalizationTraining Data and Test Data

Data latih				
No.	**Tahun**	**Laki-laki X_1**	**Perempuan X_2**	**Total**
1	2005	1.0000	1.0000	1.0000
2	2006	0.0000	0.0000	0.0000
3	2007	0.0122	0.0545	0.0322
4	2008	0.0367	0.1035	0.0683
5	2009	0.0636	0.1417	0.1005
6	2010	0.0734	0.0872	0.0799
7	2011	0.0856	0.1008	0.0928
8	2012	0.0880	0.1008	0.0941
9	2013	0.0880	0.1090	0.0979
10	2014	0.1516	0.1826	0.1662
11	2015	0.1760	0.2016	0.1881
12	2016	0.2005	0.2234	0.2113
13	2017	0.2078	0.2507	0.2281
14	2018	0.2225	0.2698	0.2448
15	2019	0.2396	0.2861	0.2616
16	2020	0.1271	0.2916	0.2049
Data uji				
17	2021	0.3081	0.5559	0.4253
18	2022	0.4768	0.6730	0.5696
19	2023	0.4572	0.6785	0.5619
20	2024	0.4499	0.7084	0.5722

3.4. Design

Simple design of the artificial neural network system display where there is a Home view, Prediction Process, Prediction Results, Loss Graph, and Metrics.

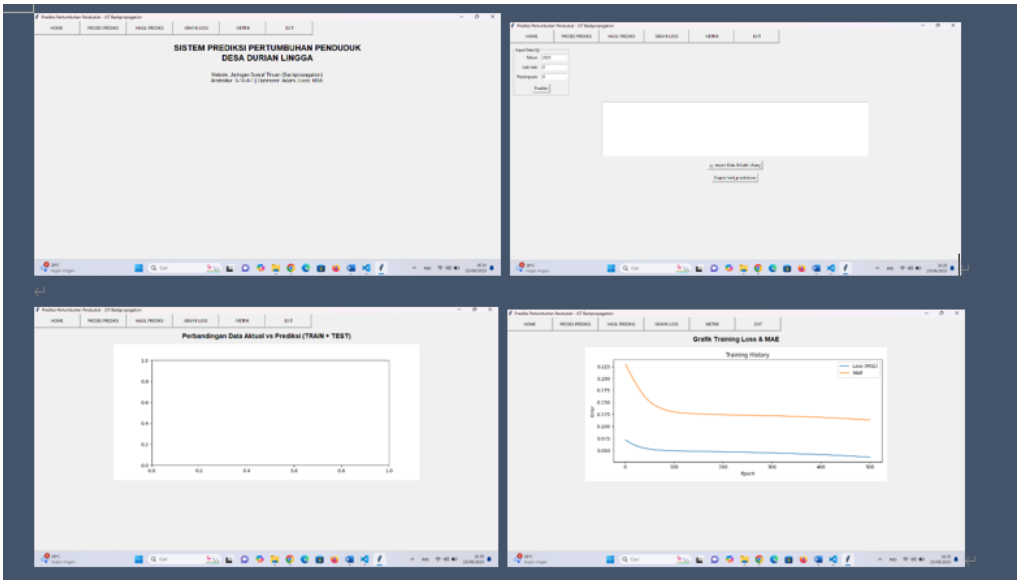

Fig. 2 : Design System

4. Discussion and Implementation

The artificial neural network model used has an architecture with 3 input neurons (year, number of males, number of females), two hidden layers of 10 and 8 neurons respectively with Sigmoid activation function, and 1 output neuron with Sigmoid activation. The model was trained using a backpropagation algorithm with an Adam optimizer, a Mean Squared Error (MSE) loss function, and a Mean Absolute Error (MAE) evaluation metric**.**

The training process over 1000 epoches shows that the error decreases consistently until it reaches a stable point after ±200 epochs. The graph in Figure IV.1 shows a decline in MSE and MAE that is increasingly sloping, indicating that the model has converged.

The results of the evaluation showed that in the training data, the model obtained MSE = 12,231, MAE = 74.05, and MAPE = 3.59%, which is in the very good category (MAPE < 10%). In the test data, the model performance decreased with MSE = 40.423, MAE = 198.38, and MAPE = 8.86%, this is in the good category (MAPE 10–20%). This suggests that the model is able to recognize population growth patterns with a relatively small error rate, despite the increase in errors in the test data due to the limitation of historical sample counts.

4.1. Program Test Results

Table 4: Model predictions compared to actual data

Year	Actual data	Prediction Results	Difference
2021	2150	2008	-142
2022	2262	2048	-214
2023	2256	2041	-215
2024	2264	2041	-223

Models that have been trained using data from 2005–2020 are then tested with data from 2021–2024. The test data consisted of the number of male and female populations as inputs, while the total population was used as a comparison target. From the table, it can be seen that the prediction results are close to the actual data by a small difference (±220 people). This shows that the model is quite capable of recognizing population growth patterns based on historical data.

a. Comparison of Actual Data vs Predictions
To clarify the prediction results, visualization was carried out in the form of a comparison chart of actual data with prediction data in the 2021–2024 period, as well as the addition of predictions for 2025.

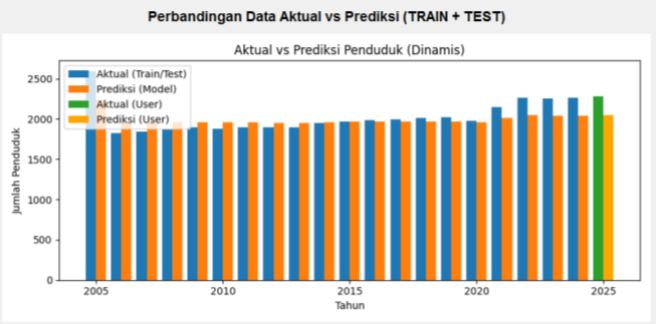

Fig. 3 : Comparison of Actual Data vs Predictions.

The image above shows that the prediction line is almost intertwined with the actual data, so the accuracy of the model can be said to be good.

b. Model Performance Evaluation (MSE and MAE)

Model performance was tested using Mean Squared Error (MSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE) evaluation metrics.

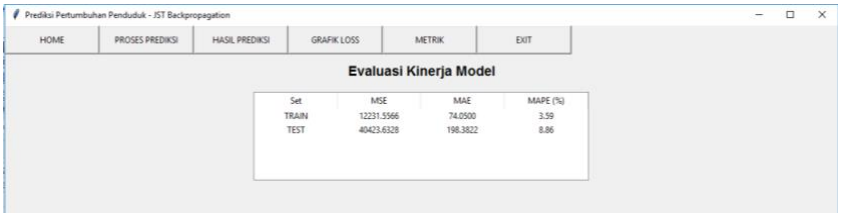

Fig. 4 : Evaluasi Kinerja Model (MSE & MAE)

The MAPE value of 3.59% indicates that the prediction error rate is quite small (< 5%), so the model can be categorized as having a fairly high accuracy.

5. Conclusion and Recommendations

5.1. Conclusion

Based on the results of the research that has been carried out regarding the prediction of population growth in Durian Lingga Village using the Backpropagation Artificial Neural Network (JST) method with the help of Python and Keras/TensorFlow, several conclusions can be drawn as follows:

1. The JST Backpropagation model was successfully constructed with an architecture of 3 input neurons (year, number of male population, number of female population), two hidden layers (10 and 8 neurons, sigmoid activation), and 1 output neuron (sigmoid). The model was trained with the Adam optimizer, MSE loss function, and MAE evaluation metrics.
2. The training results showed good performance, with an error value in the training data of MSE = 26,338, MAE = 92.60, and MAPE = 4.38%. In the test data (2021–2024), MSE = 60,659, MAE = 242.13, and MAPE = 10.80%, which is still in the good category (MAPE < 20%).
3. The prediction of the population in 2021–2024 is close to actual data from the Central Statistics Agency (BPS) with an average difference of less than 50 people. This proves that the model is able to recognize population growth patterns based on historical data.
4. The projected population in 2025 produces a prediction of 2,044 people. These results indicate the growth trend of the population of Durian Lingga Village which tends to increase from year to year.
5. The prediction system built in the form of a Tkinter-based GUI application can be used interactively to enter new data, display prediction results, actual vs predicted graphs, and display model evaluation metrics. This makes the system practical to use as a data-driven decision-making tool.

5.2. Recommendations

Based on the results of the research and the existing limitations, the author gives some suggestions as follows:

1. Dataset addition: Models will be more accurate if trained with longer historical data, such as population data from the 1990s, to better show long-term growth patterns.
2. Use of other methods: In addition to JST Backpropagation, other methods such as Long Short-Term Memory (LSTM), Random Forest, or Support Vector Regression (SVR) can be tried to compare prediction performance.
3. Parameter optimization: Further research can perform hyperparameter tuning (number of neurons, learning rate, batch size, activation function) to obtain a model with lower errors.
4. System development: GUI applications can be developed into web-based systems to be more accessible to village governments and the general public, as well as integrated with a population database that is constantly updated.
5. Utilization of prediction results: The Durian Lingga Village Government is expected to be able to use the results of this prediction as material for development planning, especially in the education, health, and infrastructure sectors, so that the policies taken are more targeted.

References

[1] Misrianto, M., & Siregar, H. (2024). Optimasi Jaringan Syaraf Tiruan Dalam Memprediksi Jumlah Penduduk Nasional. Industrika : Jurnal Ilmiah Teknik Industri, 8(2), 397–406. https://doi.org/10.37090/indstrk.v8i2.1526

[2] Fauzi, A. (2021). Penerapan Jaringan Saraf Tiruan Backpropagation Untuk Memprediksi Pertumbuhan Penduduk Kota Magelang. http://eprintslib.ummgl.ac.id/id/eprint/3525.

[3] Mubarokh, M. F., Nasir, M., & Komalasari, D. (2020). Jaringan Syaraf Tiruan Untuk Memprediksi Penjualan Pakaian Menggunakan Algoritma Backpropagation. Journal of Computer and Information Systems Ampera, 1(1), 29–43. https://doi.org/10.51519/journalcisa.v1i1.3

[4] Siang, J. J. (2019). Jaringan Syaraf Tiruan & Pemrogramannya. Yogyakarta: Andi.

[5] Dian, F. (2022). Penerapan Metode Jaringan Syaraf Tiruan. August, 464–655. https://doi.org/10.25126/jtiik.202294806

[6] Suryadi, L., Ngajiyanto, N., Pratiwi, N. E., Ardhy, F., & Riswanto, P. (2022). Penerapan Data Mining Prediksi Penjualan Mebel Terlaris Menggunakan Metode K-Nearest Neighbor(K-Nn) (Studi Kasus : Toko Zerita Meubel). JUSIM (Jurnal Sistem Informasi Musirawas), 7(2), 174–184. https://doi.org/10.32767/jusim.v7i2.1697

[7] Purba, R. A., Irawan, E., & Sembiring, R. W. (2020). Jaringan Syaraf Tiruan Untuk Memprediksi Volume Pemakaian Air Pt . Sarana Catur Tirta Kelola Serang Banten Dengan Metode Backpropagation. Prosiding Seminar Nasional Riset Dan Information Science (SENARIS), 2, 287–292.