Prediction of arrival domestic and foreign tourists based on regions using neural network algorithm based on genetic algorithm

Mohamad Ilyas Abas* and Alter Lasarudin

Faculty Science and Technology, Universitas Muhammadiyah Gorontalo, Gorontalo, Indonesia

*ilyasabas@umgo.ac.id

Abstract. Tourists are an integral part of the world of tourism. Generally, tourists visit to see the diversity of an area. In Gorontalo, several tourist attractions have been visited by domestic and foreign tourists. This is certainly a large amount so that it can help improve economic growth in Gorontalo from the tourism sector. Therefore the need for knowledge of the number of tourists for the coming year. So that, it can provide an analysis of the consideration of the decision of the government to be able to prepare steps in building the economy of the tourism sector. The number of tourists can be made a prediction using the method in data mining namely the Neural Network. Neural Network is a good method for predicting non-linear datasets such as the number of tourists, with the Neural Network method, it can be done. Not only that, Genetic Algorithm will be used to optimize the parameters of the Neural Network so that it can increase the accuracy value that can be measured with the Root Mean Square Error (RMSE) value. The results of this study indicate that the value of RMSE for domestic tourist data as follows: Gorontalo City: 0.116, Gorontalo Regency: 0.220, Boalemo: 0.073, Pohuwato: 0.142, Bone Bolango: 0.078, North Gorontalo: 0.093. For foreign tourists, Gorontalo City: 0.117, Gorontalo Regency: 0.178, Boalemo: 0.075, Pohuwato: 0.099, Bone Bolango: 0.124, North Gorontalo: 0.155.

1. Introduction
Tourists are one of the many sectors that can be proud of in every region. The arrival of tourists from various regions and countries is a matter of pride in the area. of course, besides that, it can increase the economy of an area. As is the case in several other areas, tourist attractions have become the wheels of the economy that move every month and can become new income for local governments and the surrounding population in particular. The number of tourists that increases over time can attract investors to build and renovate the tourist attractions. Facilities in the form of hotels, lodging and transportation will experience an increase in every tourist visit to the area. In Gorontalo, the number of tourists is relatively large with a number reaching 1,267,581 consisting of domestic and foreign tourists. These tourists are spread throughout the regions in Gorontalo province. The visit of tourists who come for each month in each district/city in Gorontalo always changes from 2013 to 2016. Data that often changes in the period of the month makes the Gorontalo government difficult to predict the arrival rate of tourists to Gorontalo. With the prediction of government arrivals, in this case, the Gorontalo Tourism Office can avoid the decline in the number of domestic tourists and foreign tourists to come to visit to enjoy the diversity of tourism in Gorontalo.
According to data from the Gorontalo Tourism Office, several tourist destinations that can be visited are divided into regions, including olele beach, lombongo tourism, whale shark tours located in Bone Bolango, otanaha fortress, Soekarno landing museum, bongo religious tourism (City), water baths Pentadio Resort (Regency), Nantu Animal Sanctuary, Bolihutuo Beach (Boalemo), Pulau Saronde (North Gorontalo), Bajo Torosiaje Tribe and also love located in Pohuwato area. From the data obtained, it was noted that the area of Gorontalo City was the destination area that was visited by domestic tourists and foreign tourists as many as 6,967 tourists. This certainly can be a consideration for the government to make accurate predictions. Services, infrastructure development and tourism facilities can be improved not only in terms of tourist income, this also provides investors with an investment in each tourist location which is predicted to increase tourist arrivals. Prediction can be done using one of the data mining algorithms, namely the neural network. Therefore several studies propose to optimize the parameters of neural networks using genetic algorithms such as research by Noersasongko [1] showing that genetic algorithms are able to optimize neural network parameters so as to provide good RMSE results.

Hartono[2] conduct research on Rentet Prediction Model for Health Drink Sales Based on Neural Network. Research focuses on the best architecture using a neural network, where the neural network parameters are set by using the number of hidden layers no more than the number of inputs and outputs. Four stages are done, namely adjustment parameters, models, objects and measurements. Parameters used include learning rate, momentum and neuron size. The model used in this experiment is the neural network with backpropagation. The object under study is the prediction of the rented time of the business of selling health drinks. While the measurement uses RMSE (Root Mean Square Error). Based on the results of the experiments that have been done, the best models for predictions with architectural training cycles are: 300, learning rate: 0.3, momentum: 0.7, number of hidden layers: 1, and size of hidden layer: 23, low RMSE level is 0.152. And the error resulting from the test results is 0.0189%. This shows that neural networks with backpropagation learning algorithms can produce time-rent models for accurate predictions.

Ruliah[3] conduct research on Electricity Forecasting in the operation of electric power systems ranging from generation planning, power flow analysis, unit commitment, thermal hydro, and power system economical operations. Therefore forecasting the use of electricity is important. In this study using time series data from electricity usage in South and Central Kalimantan in the period 2008 - 2012 because it has a time attribute, the method used is Backpropagation Neural Network. The results of this study were able to produce Electric Usage Forecasting with a structure of 12-3-1 Root Mean Square Error generated reaching 0.024, and on a 12-25-1 structure capable of producing Root Mean Square Error 0.011, and finally in a 12-100-1 structure capable of generating Root Mean Square Error 0.0098.

Varahrami[4] conduct research on forecasting short-term water needs in planning and managing water and waste facilities such as pump scheduling, reservoir volume control, pressure management and water conservation programs. This helps deep network managers to reduce system and consumer vulnerabilities to increase network reliability. This paper describes the application of genetic algorithms to forecasting short-term water demand neural networks. Two types of artificial neural networks: MLFF with backpropagation and GMDH learning algorithms with genetic learning algorithms were investigated to model water demand estimates. This comparison shows that GMDH artificial neural network with GA produces results that are close to the actual data.

Makvandi[5] conduct research to determine effective variables regarding forecasting future dividends from companies that are members of the stock exchange in Tehran. The results of research conducted that the NNGA model can be used in various fields such as financial forecasting, market variable prediction, decision making and so forth.

Islam[6] conduct research to develop hybrid prediction models based on neural networks and genetic algorithms that integrate the benefits of these two techniques to increase the estimated electric load. Genetic algorithms can optimize neural network architecture with initial neuron weights, selection of training algorithms and critical analysis and selection of the most appropriate input parameters are some of the considerations in the analysis in predictions.
2. Neural Network
Artificial neural networks embody subsystematic paradigms to represent and process information. The field of science that deals with methods and systems for processing information using artificial neural networks is called neurocomputation [7]. Neural Network (Neural Network) is a method whose pattern follows the work of the human nervous system. Where the main processing of the human nervous system lies in the brain. The smallest part of the human brain is the nerve cell which is the basic unit for information processing. This unit is also called a neuron. There are about 10 billion neurons in every human brain and about 60 trillion connections between neurons in the human brain [8]. By using these neurons simultaneously, the human brain can process information in parallel and quickly, even faster than the fastest computer at this time. Departing from the analogy of the nervous system of the brain, the neural network adopts it into an algorithm. Neural networks consist of a processing unit called a neuron that contains an adder and an activation function, a number of weights and a number of input vectors. The input vector consists of a number of values given as input values in the neural network, the input vector consists of 3 values (x1, x2, x3) as features in the vector that will be processed in the neural network, each input value pass through a weighted relationship, then all values are combined. The combined value is then processed through the neuron by the activation function to produce the y signal as the output. The activation function uses a threshold value to limit the output value so that it is always within the specified value limit.

3. Genetic Algorithm
Genetic Algorithm is the computational process for optimization that was first introduced by John Holland of the University of Michigan in 1975. Genetic Algorithms is a process of finding optimization in random searches [8]. In the process this algorithm is carried out for the selection process naturally or known as the process of evolution and operation of the genetics of chromosomes. So that in its use can be used to optimize the value of Root Mean Squared Error (RMSE) to the Neural Network Algorithm. This algorithm utilizes the natural selection process which is known as the evolution process and genetic operation of chromosomes.

Neural networks are algorithms that are often used for time series datasets. Selection of input variables for effective training data is one of the most important dilemmas in the field of forecasting and decision making [5]. In the development of the parameters of the neural network can be optimized using genetic algorithms to automate neural network parameters consisting of training cycles, learning rate and momentum. Research results by [9] shows that the design of neural networks and genetic algorithms can process and predict time series data so that it can get good results.

4. Results and Discussion
Data collection was conducted at the Gorontalo Provincial Tourism Office through data on both domestic and foreign tourists. Data is obtained by entering a research permit to request data collection for data processing in the study. From the data obtained, it was noted that the tourism agency carried out data records from 2013 to 2016. The data was divided into each district / city. So that in the process of processing and analyzing data later, the data taken meets the criteria in processing with neural network based genetic algorithm. The collection of tourist arrivals taken is divided into monthly periods. Here are the charts of domestic and international tourists in Figures 2 and 3 below:
Figure 1. Domestic tourist data 2013-2016 (Gorontalo Tourism Office)

Figure 2. Data on foreign tourists 2013-2016 (Gorontalo Tourism Office)

The data pattern used to convert univariate data into multivariate can be seen in the following data pattern table.

**Table 1.** Time series data patterns (univariate to multivariate) [10]

| Pattern | Input lag                  | Output/Target |
|---------|---------------------------|---------------|
| 1       | x₁, x₂, x₃, x₄, ..., xₚ    | xₚ₊1          |
| 2       | x₂, x₃, x₄, x₅, ..., xₚ₊1  | xₚ₊2          |
| 3       | x₃, x₄, x₅, x₆, ..., xₚ₊2  | xₚ₊3          |
| ...     | ...                       | ...           |
| m-p     | xₘ₋p, xₘ₋p₊1, xₘ₋p₊2, ..., xₘ-1 | xₘ               |

Table 1 above is a pattern used to convert univariate data to multivariate where:

\( m-p \) : value of pattern  
\( x_{m-p} \) : input lag  
\( x_m \) : Output/Target  

\( model \) : \( x_m = x_{m-p}, x_{m-p+1}, x_{m-p+2}, ..., x_{m-1} \) ................................. (1)

Normalization is carried out aimed at changing the time series dataset into binary numbers ranging from 0 to 1[11]. This is done so that data processing can be better because it avoids the missing value or dominant value of a dataset. The following normalization formulas are used:

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

................................. (2)

Where,

\( x' \) = data resulting from normalization  
\( x \) = original data / initial data
The pattern is done as shown in Table 1 by determining the period \( p = 3 \) and \( x_p \) that are the target of the training process. This process is carried out similarly to other data processing. From the results of the prediction shows the trend of tourists in each region, the highest is in Gorontalo Regency which will experience an increase in terms of tourists. This is in line with the existence of several new tourist attractions that are developing at this time.

**Figure 3.** Graph of domestic tourists (2017)

**Figure 4.** Graph of foreign tourists (2017)
While for foreigners also in the Regency area which experienced an increase compared to other regions. This was carried out by several international seminars on halal tourism and the charm of limboto lake located in the Gorontalo Regency region.

In 2018 the prediction results also experienced an increase in the Gorontalo Regency area compared to other regions. The results of this prediction can be a reference for the government to continue to improve processing and development in the tourism sector. To get an economic increase from the results of tourist visits. The results of this prediction also prove that Gorontalo Province can be a good tourist area and can compete with other large regions such as Bali, Bandung, Jakarta and other big cities.
If seen in the picture above the trend of tourist visitors has increased and decreased in each year. This can occur due to tourism management and also the promotion of tourist attractions to the public. For this reason, it is necessary to increase the promotion and development of new tourism which is more attractive for tourists to visit Gorontalo. For domestic tourists, there has been a positive increase in 2017 and 2018. On the contrary, foreign tourists have decreased for foreign tourists in 2018.

5. Conclusion
Based on research that has been conducted on the prediction of the number of domestic and foreign tourists based on Regency / city. In the analysis process, it was proven that the optimization performed using Genetic Algorithm could reduce RMSE. So that from the combination of Neural Network pattern and Genetic Algorithm parameter determination it produces the best RMSE which will then be used to make predictions. From the analysis carried out without using Genetic, the Neural Network method algorithm shows that the RMSE value is greater than the optimization. The results of this study can be a reference for developing tourist attractions and further promotion to attract the attention of domestic tourists and foreign tourists.

The total number of domestic tourists reaching 1,740,111, this confirms that tourists have experienced a significant increase from 2013 to 2018. While foreign tourists reach the journal 23,242 there needs to be an increase in promotions for foreign tourists. All data from 2013 to 2018 will be the basis for achieving predictions in the coming year. Increasing the number of domestic tourists and tourists will be greatly influenced by tourism management in Gorontalo. Improved services to attract...
tourists to visit Gorontalo. More and more tourists visiting will improve the economy of the Gorontalo community starting from transportation, buying and selling typical Gorontalo goods and also the community near tourist attractions. This is certainly valuable both in terms of the economic community.

6. References

[1] Noersasongko E Julfia F T Syukur A , P Pramunendar R A and Supriyanto C, 2016 A Tourism Arrival Forecasting using Genetic Algorithm based Neural Network Indian J. Sci. Technol. 9, 4 p. 3–7.

[2] Hartono, Daniel dan Satrio Wahono R, 2013 Model Prediksi Rentet Waktu Penjualan Minuman Kesehatan Berbasis Neural Network J. Teknol. Inf. 9 p. 12–21.

[3] Ruliah S and Rolyadely R, 2014 Prediksi Pemakaian Listrik Dengan Pendekatan Back Propagation J. Tek. Inform. dan Sist. Inf. 3, 1 p. 465–476.

[4] Varahrami V, 2010 Application of Genetic Algorithm To Neural Network Forecasting of Short – Term Water Demand E ζ p. 783–788.

[5] Makvandi P Jassbi J and Khanmohammadi S, 2005 Application of Genetic Algorithm and Neural Network in Forecasting with Good Data 2005 p. 56–61.

[6] B . Islam , Z . Baharudin Q . R and P . N, 2014 A Hybrid Neural Network and Genetic Algorithm Based Model for Short Term Load Forecast Res. J. Appl. Sci. Eng. Technol. 7, 13 p. 2667–2673.

[7] Kasabov N K, 1996 Foundations of neural networks, fuzzy systems, and knowledge engineering .

[8] Dewi S K, 2005 Penyelesaian Masalah Optimasi Menggunakan Teknik-teknik Heuristik Yogyakarta: Graha Ilmu.

[9] Peralta J Gutierrez G and Sanchis A, 2007 Design of Artificial Neural Networks based on Genetic Algorithms to Forecast Time Series.

[10] Purwanto Eswaran C and Logeswaran R, 2011 Improved Adaptive Neuro-Fuzzy Inference System for HIV/AIDS Time Series Prediction Informatics Eng. Inf. Sci. Pt Iii 253 p. 1–13.

[11] Prasetyo E, 2014 Mengolah Data Menjadi Informasi Menggunakan Matlab Yogyakarta.