Jurnal Elektronik Ilmu Komputer Udayana
Volume 8, No 3. February 2020

p-ISSN: 2301-5373
e-ISSN: 2654-5101

Prediction Of The Number Of Tourists To Visit Bali Province Using Backpropagation Artificial Neural Network (Case Study: Data 1990-2016)

I Gusti Agung Ngurah Panji Palguna[a1], I Made Widiartha[a2]

[a]Informatics Department, Mathematics and Science Faculty, Udayana University
Jalan Raya Kampus Unud, Bukit, Jimbaran, Bali, Indonesia
panji_palguna@student.unud.ac.id[1], imadewidiartha@cs.unud.ac.id[2]

Abstract

Bali Island is the most popular tourist destination in Indonesia. The total number of foreign tourists visiting Indonesia through the entrance of Ngurah Rai Airport reached 40% as of October 2016, with the value of Bali's foreign exchange receipts for Indonesia from the tourism sector amounting to 70 Trillion Rupiah. Minister of Tourism (Menpar) Arief Yahya always uses the password "Bali" in promoting destinations throughout the world. Because in tourism, Bali is a gate that is passed by 40 percent of foreign tourists (tourists) to Indonesia. In support of more accurate decision making, the author makes a system of forecasting numbers of foreign tourists visiting Bali Province by taking a sample of Japan. Factors that are used as input to make predictions include the number of tourists visiting before, the population of the country of origin of foreign tourists, Gross Domestic Product, and the Relative Consumer Price Index of the countries of origin of foreign tourists. In this research, optimization of the activation function, hidden neuron, and learning rate parameters is performed. Forecasting results using the backpropagation method produce a pretty good accuracy with an accuracy of Mean Square Error = 0.0050558, and test data accuracy of MSE = 0.031695. ANN architecture in the training process is then used to calculate predictions of visits by foreign tourists in the testing process

Keywords: *Artificial Neural Networks, Mean Square Error, Prediction, Backpropagation*

1. Introduction

Based on data from the United Nations World Tourism Organization (UNWTO), tourism has experienced continued growth and deepened its diversification to become one of the fastest growing economic sectors in the world. At present, the volume of tourism businesses is equal to or even exceeds the volume of exports of oil, food products, or automotive. Tourism has become a major player in international trade, and is also a major source of income for many developing countries.

All forms of efforts to improve quality and quantity in the tourism sector are inseparable from the budget costs that must be prepared by both the government and the private sector. Directly budget planning without strategy can provide potential losses to budget providers. One of the basics in strategizing and making decisions is forecasting. Forecasting is an activity of estimating what will happen by considering past events and the influence of current conditions. Forecasting an accurate number of foreign tourist visits will certainly provide many benefits to managers and investors in making decisions regarding operational, planning, and marketing as well as assisting the government in designing the budget appropriately.

In this study a prediction is made by taking into account 3 factors as training data, namely the population in the country, the relative consumer price index in the country, and the country's gross domestic product. In this study the focus is on the data of foreign tourist arrivals in Japan.

2. Research Methods

This study will predict the number of tourist arrivals. The research data used are secondary data obtained from the Central Statistics Agency of Bali and the Official Website of the World Bank

Predictions of tourist visits with backpropagation neural networks use the following steps:

a. Separate training data and test data. Where training data used are visits from 1990-2007 and 2008-2016 as test data
b. Design of ANN
   ANN design is done to predict the visit of foreign tourists starting by determining the number of hidden screens used, and the number of outputs desired. The data used as input data are 3 features, namely population, relative consumer price index, and gross domestic product and the target is the number of tourist visits in the Province of Bali. Predicting the visit of foreign tourists using artificial neural network modeling in the process can use functions in MATLAB. Compare the foreign tourist visits obtained by calculating the relative errors resulting from the results obtained with the Artificial Neural Network.

3. Results and Discussion

3.1 Data Collection Stage

In this study, a system was made to predict the visit of foreign tourists to the Province of Bali, to implement the data needed related to foreign tourist visits for several years. These data are used to conduct training programs as training data and test data for the system. As the output is a prediction of rice prices in the following month. The data used are data from 1990-2016. With training data is the year 1990-2007 and test data from 2008-2016. 18 training data patterns and 9 test data. Or if it is presented as much as 0.67% of training data and 0.33% of test data

The following data and sources are presented in the following table.

Table 1 Data and Sources

No	Data	Sourcess
1	Foreign tourist arrivals to Bali (annual)	Central Bureau of Statistics
2	Population of countries of origin of foreign tourists (annual)	World Bank Official Website
3	GDP at Current Prices (ADHB)	World Bank Official Website
4	GDP Deflator (GDP DEF)	World Bank Official Website
5	Consumer Price Index by Country (Annual)	World Bank Official Website
6	Average Currency Exchange Rates	World Bank Official Website

These data will be grouped into 3 features that will determine how the pattern of foreign tourists visiting Bali based on the following factors or features:

1. Population of the country of origin of foreign tourists
2. Gross Domestic Product; this is obtained by constant GDP = GDP ADHB / GDP DEF.
3. Relative Consumer Price Index obtained by means of CPI = CPI INA / (Japanese CPI * Exchange Rate (Indonesia-> Japan))

3.2 Data Analysis

In this study the type of ANN activation function used for data normalization is binary sigmoid. Where this function works in the range [0,1]. However, this function never reaches a value of 0 or 1 so the data needs to be normalized by transforming the data into the range 0.1-0.9 by using the following equation:

p-ISSN: 2301-5373
e-ISSN: 2654-5101

$$x' = 0.8\left(\frac{x - \min(x)}{\max(x) - \min(x)}\right) + 0.1$$

where x 'is data after normalization and x is data before normalization

With the following features:

X1 = Total Population

X2 = Constant Domestic Product (GDP) Constant

X3 = Consumer Price Index

The target is tourist visits in the year (1990-2016). The results of data normalization are as follows.

Table 2 *Training Data (1990-2007)*

Training Data				
Pattern	X1	X2	X3	Target
1	0.1	0.6513	0.1	0.1
2	0.16777	0.6512	0.107667	0.144724
3	0.222127	0.6577	0.114372	0.35679
4	0.276307	0.7044	0.123652	0.381597
5	0.351313	0.9000	0.132603	0.483849
6	0.435672	0.5813	0.143358	0.191956
7	0.491794	0.5842	0.153301	0.354914
8	0.544739	0.6958	0.161694	0.564406
9	0.605272	0.1000	0.245256	0.494457
10	0.64604	0.6219	0.2917	0.726635
11	0.683455	0.6216	0.301861	0.9
12	0.737459	0.6194	0.334444	0.718519
13	0.789698	0.6259	0.371965	0.732738
14	0.837878	0.6263	0.395239	0.530765
15	0.845467	0.6165	0.418755	0.799835
16	0.847584	0.6153	0.460582	0.756634
17	0.86188	0.6120	0.518526	0.729845
18	0.887823	0.6056	0.55056	0.87186

Table 3. Testing Data (2008-2016)

Testing Data				
Pattern	X1	X2	X3	Target
1	0.8987646	0.6119	0.60257	0.893273
2	0.8959409	0.5912	0.630681	0.82199
3	0.9	0.6248	0.662098	0.578066
4	0.8581734	0.6207	0.696575	0.405279
5	0.8221707	0.5938	0.72559	0.422677
6	0.7896978	0.5497	0.770934	0.476042
7	0.759872	0.6596	0.819046	0.500914
8	0.7360468	0.6559	0.869981	0.531238
9	0.7101939	0.7474	0.9	0.542146

In processing data in this study using back propagation ANN with 3 inputs and 1 target (output). Input in the form of features that affect tourist visits occur such as Gross Domestic Product, Consumer Price Index, and Population in the Country of Origin of Foreign Tourists. In ANN, the input values of these 3 features are propagated forward to neurons in the hidden layer. In the hidden layer, the values are processed using the activation function then propagated forward through output weights so as to get an output value that is prediction of foreign tourist visits. The next output value is compared with the target value so an error value is obtained. If the error value is smaller than the error goal, the iteration will stop.

However, if the error value is still greater than the error goal, back propagation is performed by updating the output weight and input weight. In ANN training, there are several parameters that are set. These parameters are shown in the following table.

Table 4. ANN Parameters with Reverse Propagation

No	Parameter	Score
1	Hidden Layer	10
2	Activation Function	Sigmoid
3	Error goal	0.001
4	Number of Epoch	1000
5	Momentum	0.95
6	Learning Rate	0.1

3.3. Training Stage

At this stage 18 consecutive sample data patterns (1990-2007) were given to be trained as a learning process. To get the output value, the first thing to do is to determine the target matrix. Next determine the formula of the artificial neural network by regulating the activation function between the hidden layer input by using logical (binary sigmoid), the activation function from the hidden layer to the output using the purelin (linear) activation function and the network training function using the training function. At the training stage, the first thing to do is to call the data you have in Excel format into matlab, here is the source code.

```
filename = 'dataset_wismanjpn.xlsx';
sheet = 2;
xlRange = 'D6:G23';

Data = xlsread(filename, sheet, xlRange);
data_latih = Data(:,1:3)';
target_latih = Data(:,4)';
[m,n] = size(data_latih);
```

Figure 1. Call Training Data

Figure 1 is a piece of source code that aims to call training data in excel form consisting of 3 parameters, namely filename, sheet, and xlRange. Then determined training data and target data on Excel that has been called.

```
net = newff(minmax(data_latih),[10 1],{'logsig','purelin'},'traingdx');
```

Figure 2. Making Artificial Neural Network

The script snippet in Figure 2 aims to create a network using newff tools, where there are several parameters, namely the training data, hidden layer, output, and activation functions.

p-ISSN: 2301-5373
e-ISSN: 2654-5101

```
net = newff(minmax(data_latih),[10 1],{'logsig','purelin'},'traingdx');
net.performFcn = 'mse';
net.trainParam.goal = 0.001;
net.trainParam.show = 20;
net.trainParam.epochs = 1000;
net.trainParam.mc = 0.95;
net.trainParam.lr = 0.1;
```

Figure 3. Determine the value to influence the training process

The script snippet in Figure 3 aims to influence the training process with the backpropagation method, namely the number of epochs, error goals, learning rates and momentum. The results of training conducted with artificial neural networks are as follows:

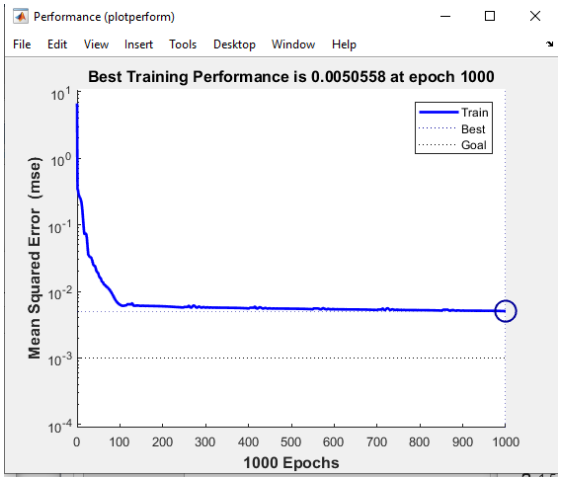

Figure 4. Training Performance

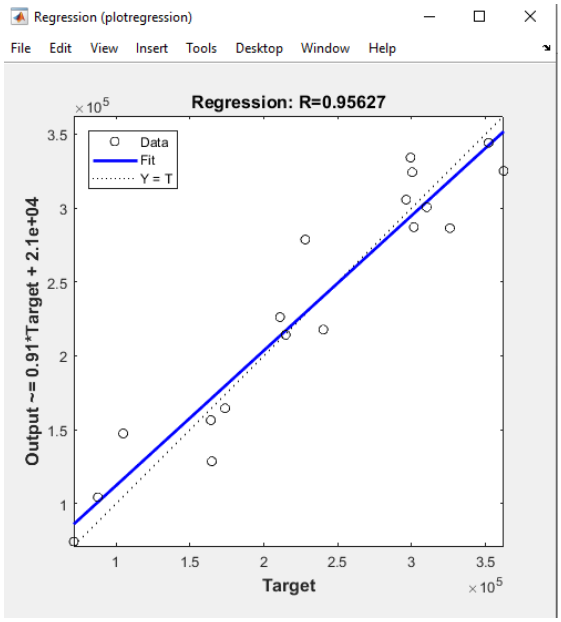

Figure 5. Regression of Training Results

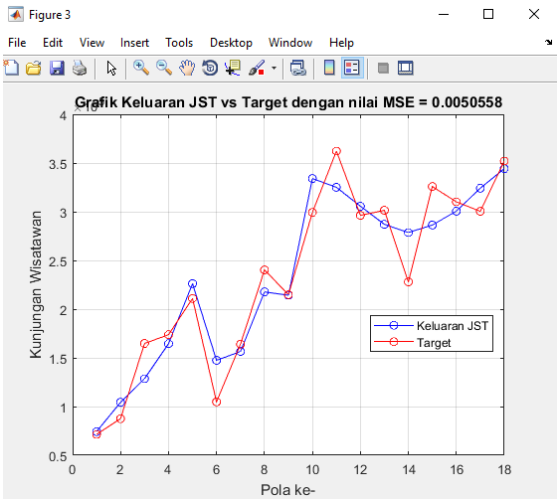

Figure 6. Graphics Output of ANN Result and Target

Based on figures 5 and 6, it appears that the level of accuracy obtained is the value of the correlation coefficient 0.95267 and MSE = 0.0050558

3.4 Testing Stage

At this stage the test data testing process is carried out. Here are some source code for the testing process.

```
filename = 'dataset_wismanjpn.xlsx';
sheet = 2;
xlRange = 'D28:G36';

Data = xlsread(filename, sheet, xlRange);
data_uji = Data(:,1:3)';
target_uji = Data(:,4)';
[m,n] = size(data_uji);
```

Figure 7. *Source Code Testing Stage*

Figure 7 is a piece of source code that aims to call test data in excel form, where data in excel form consists of 3 parameters, namely filename, sheet, and xlRange. Then determined the test data and target training data on excel that has been called, which aims to simulate the prediction results of foreign tourist visits. The results of tests conducted with artificial neural networks are as follows:

p-ISSN: 2301-5373
e-ISSN: 2654-5101

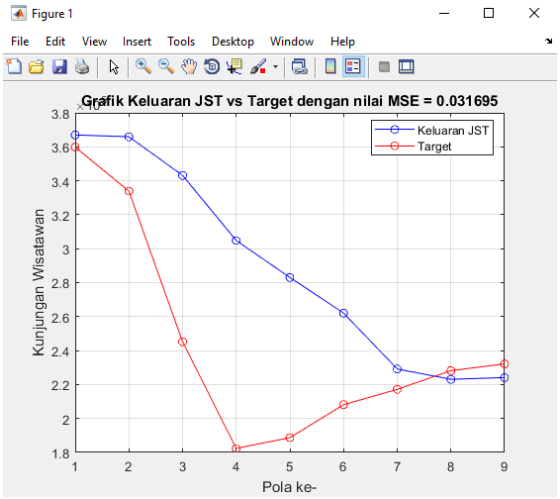

Figure 8. Output of ANN Result and Target

Based on Figure 8 it appears that the level of accuracy obtained is the value of MSE = 0.031695.

4. Conclusion

Based on the research conducted, the following conclusions can be drawn:

1. The results obtained from the calculation of predictions of foreign tourist visits by using the back propagation ANN algorithm. The input value is in the form of 3 factors or features that determine the visit of foreign tourists, among others: Population of the country of origin of foreign tourists, the consumer price index (CPI), and the Gross Domestic Product of the Constant Country of Origin of Foreign Tourists (Constant GDP). ANN architecture used includes 3 input layers, 10 hidden layers and 1 output. The maximum number of epochs used is 1000 with a learning rate of 0.1, an error goal of 0.001 and momentum of 0.95.
2. The training process uses 18 training data patterns of training data (1990-2007) and 9 test data (2008-2016) which yields a MSE value = 0.0050558. ANN architecture in the training process is then used to calculate predictions of visits by foreign tourists in the testing process. The resulting MSE value is 0.031695

Suggestion

The correlation coefficient and MSE values can be improved by increasing training data and changing parameters that affect network performance such as error goals, number of epochs, network architecture, types of activation functions, etc.

References

[1] Siang J..J 2009. Jaringan Syaraf Tiruan dan Pemrogramannya Menggunakan MATLAB. Yogyakarta:Andi

[2] Andrian, Y.& Ningsih, E.2014. Prediksi Curah Hujan di Kota Medan. Seminar Nasional Informatika

[3] Made Agung Raharja & I Made Teja Gani Astra.2018.Prediksi Ketinggian Gelombang Laut Menggunakan Metode Backpropagation pada Pantai Lebih, Gianyar. Seminar Nasional Teknologi dan Aplikasinya Ilmu Komputer Unud

[4] Nikki Asvikarani.2018.Peramalan Angka Kunjungan Wisatawan Mancanegara ke Provinsi Bali Menggunakan Metode Cascade Forward Backpropagation.Tidak Diterbitkan(Skripsi)