Foreign Tourist Arrivals Forecasting Using Recurrent Neural Network Backpropagation through Time

Wayan Oger Vihikan*, I Ketut Gede Darma Putra, I Putu Arya Dharmadhi
Information Technology Department, Faculty of Engineering, Udayana University,
Bukit Jimbaran, Bali, Indonesia-803611, Telp/fax: (0361) 701954, 704845 / (0361) 701907
*Corresponding author, e-mail: ogervihikan@outlook.com

Abstract
Bali as an icon of tourism in Indonesia has been visited by many foreign tourists. Thus, Bali is one of the provinces that contribute huge foreign exchange for Indonesia. However, this potential could be threatened by the effectuation of the ASEAN Economic Community as it causes stricter competition among ASEAN countries including in tourism field. To resolve this issue, Balinese government need to forecast the arrival of foreign tourist to Bali in order to help them strategizing tourism plan. However, they do not have an appropriate method to do this. To overcome this problem, this study contributed a forecasting method using Recurrent Neural Network Backpropagation Through Time. We also compare this method with Single Moving Average method. The results showed that proposed method outperformed Single Moving Average in 10 countries tested with 80%, 70%, and 70% better MSE results for 1, 3 and 6 months ahead forecast respectively.

Keywords: Backpropagation Through Time, forecasting, tourism, Recurrent Neural Network

1. Introduction
The tourism industry is one of the huge industries that is growing very rapidly throughout the world. Many countries put the tourism sector as a strategic industry to increase revenue, including Indonesia, which has a great potential. Indonesia's tourism industry has become an important part of national development, as it can hold a lot of manpower, improve the local economy, and state revenues.

Bali as an icon of tourism in Indonesia and also as a world tourist destination that has a variety of art, culture and hospitality of the community has been visited by many foreign tourists. Data from Bali Tourism Office shows that since 2011 until 2015 the amount of foreign tourist arrivals to Bali is always increasing. For example, in 2011 the number of foreign tourist arrivals is 2,756,579, in 2012 it increased to 2,892,019, in 2013 it went up to 3,278,598, in 2014 it increased about 15% to 3,766,638 and in 2015 it climbed up to 4,001,835 [1]. In addition to that, Bali is the only province in Indonesia which has the highest amount of foreign tourist staying in star hotels since 2003 until 2015 [2]. Moreover, Bali won as the best island for holidaymakers and it is Bali’s 12th times [3]. These make Bali as one of the provinces that contribute huge foreign exchange for Indonesia.

In the end of 2015, ASEAN Economy Community (AEC) was applied. This will cause strict competition among ASEAN countries in a lot of sectors including tourism. Indonesia as part of ASEAN had already prepared a presidential decree which ask local government to improve competitiveness in order to face AEC. It means Balinese government have a good opportunity to maintain and improve the amount of foreign tourist arrivals to Bali by promoting Bali’s tourism intensively and improving Bali’s tourism facilities. These can be achieved by strategizing Bali’s tourism based on past data of foreign tourist arrivals which lead to forecasting to get overview of foreign tourist arrivals in the future.

The problem is Balinese government do not have an appropriate method to forecast foreign tourist arrival to Bali. Their forecast for the current year is based on the last year foreign tourist arrival growth with analysis of several factors such as political situation and natural disasters. To help overcome that problem, this paper proposed Recurrent Neural Network as forecasting method with Backpropagation Through Time as the training algorithm. Previous
researches using Recurrent Neural Network Backpropagation through Time show good results in forecasting time series data [4-7].

Therefore, the aims of this research is to contribute an accurate forecasting method of foreign tourist arrivals for Balinese government in order to anticipate the increase or decrease in foreign tourist arrivals to Bali. This will help Balinese government in setting up tourism facilities such as accommodation facilities (hotels, villas), infrastructure (roads, water, electricity), tourist attraction, restaurants, transportation, travel agencies, money changers and others.

2. Research Method

This research uses data of foreign tourist arrivals to Bali and their factors. The factors of foreign tourist arrivals are population of origin country, Gross Domestic Product (GDP) real of origin country and Consumer Price Index (CPI) in Indonesia relative to CPI of origin country [8]. Those data are taken from 2005 to 2015 and has time series characteristic. Time series data is a set of data generated sequentially in time [9].

Before beginning the process, the data have to be normalized and scaled to range between -1 and 1 using equation below. Then, the data is divided into 2 parts, those are training and test data.

\[
\hat{x} = \frac{(x - \text{minVal}) \times (\text{maxRange} - \text{minRange})}{(\text{maxVal} - \text{minVal})} + \text{minRange} \quad (1)
\]

With:
- \(x\) : Original data.
- \(\hat{x}\) : Normalized data.
- \(\text{maxVal}\) : The maximum value of original data.
- \(\text{minVal}\) : The minimum value of original data.
- \(\text{maxRange}\) : The maximum value of normalized data (1).
- \(\text{minRange}\) : The minimum value of normalized data (-1).

Recurrent Neural Network is adapted from standard feed forward neural network that can model sequential data [10]. It allows the network to map from all previous inputs to each output [11]. It is easier to understand by looking to unfolded RNN in Figure 1, where previous time step is needed to calculate current time step which is shown as \(t-1\) and \(t\) respectively. RNN BPTT is started by doing Forward Propagation process. This process will calculate the network output from the input data given using equation below [10].

\[
h_t = \tanh(Ux_t + Wh_{t-1} + b_h) \quad (2)
\]
\[
o_t = \tanh(Vh_t + b_o) \quad (3)
\]

With:
- \(h_t\) : Hidden state at time step \(t\).
- \(U\) : Input to hidden state weight matrix.
- \(x_t\) : Input vector at time step \(t\).
- \(W\) : Hidden to previous hidden state (recurrent) weight matrix.
- \(h_{t-1}\) : Hidden state at previous time step.
- \(b_h\) : Bias at hidden state.
- \(V\) : Hidden state to output weight matrix.
- \(b_o\) : Bias at output state.

After that the loss between network's output and targeted output is calculated. In this case Mean Squared Error is used as the loss function. If the value close to 0, it means the network and targeted output or factual data value are close.

TELKOMNIKA Vol. 15, No. 3, September 2017 : 1257 – 1264
Each of the weight matrices have to learn from the error (loss). This can be achieved by updating each weight matrices for each training data using Stochastic Gradient Descent. To do that, first Backpropagation Through Time should be applied. This algorithm calculates the gradient of loss function with respect to each weight matrix as formula below.

\[ L = \frac{1}{2N} \sum_{i=1}^{N} (\hat{o}_t - o_t)^2 \]  

With:
- \( L \) : Loss value.
- \( N \) : The amount of training data.
- \( \hat{o} \) : Network’s output.
- \( o \) : Targeted output.

Finally, each weight matrix is updated with BPTT calculation results and learning rate. In order to avoid over-fitting and improve generalization, weight decay is used by multiplying each weight with regularization parameter before updating them [12]. The new weight matrices value will be used in the next time step. Formula for updating weight matrix is shown below.

\[
\frac{\partial L}{\partial v} = \frac{\partial L}{\partial \hat{o}} \frac{\partial \hat{o}}{\partial z} \frac{\partial z}{\partial \hat{o}} \frac{\partial \hat{o}}{\partial v}
\]

(5)

\[
\frac{\partial L}{\partial b_0} = \frac{\partial L}{\partial \hat{o}} \frac{\partial \hat{o}}{\partial z} \frac{\partial z}{\partial \hat{o}} \frac{\partial \hat{o}}{\partial b_0}
\]

(6)

\[
\frac{\partial L}{\partial w} = \sum_{t=0}^{T} \frac{\partial L}{\partial \hat{o}} \frac{\partial \hat{o}}{\partial z} \frac{\partial z}{\partial \hat{o}} \frac{\partial \hat{o}}{\partial w} \frac{\partial w}{\partial z} \frac{\partial z}{\partial h_T} \frac{\partial h_T}{\partial z}
\]

(7)

\[
\frac{\partial L}{\partial b_h} = \frac{\partial L}{\partial \hat{o}} \frac{\partial \hat{o}}{\partial z} \frac{\partial z}{\partial \hat{o}} \frac{\partial \hat{o}}{\partial b_h} \frac{\partial z}{\partial h_T} \frac{\partial h_T}{\partial h_T}
\]

(8)

\[
\frac{\partial L}{\partial u} = \sum_{t=0}^{T} \frac{\partial L}{\partial \hat{o}} \frac{\partial \hat{o}}{\partial z} \frac{\partial z}{\partial \hat{o}} \frac{\partial \hat{o}}{\partial u} \frac{\partial u}{\partial z} \frac{\partial z}{\partial h_T} \frac{\partial h_T}{\partial h_T} \frac{\partial h_T}{\partial u}
\]

(9)

With:
- \( z \) : Network’s output before activation function (tanh) is applied.
- \( T \) : Current time step.

Finally, each weight matrix is updated with BPTT calculation results and learning rate. In order to avoid over-fitting and improve generalization, weight decay is used by multiplying each weight with regularization parameter before updating them [12]. The new weight matrices value will be used in the next time step. Formula for updating weight matrix is shown below.

\[
U = U - \eta \ast \left( \frac{\partial L}{\partial u} + \lambda U \right)
\]

(10)

\[
W = W - \eta \ast \left( \frac{\partial L}{\partial w} + \lambda W \right)
\]

(11)

\[
b_h = b_h - \eta \ast \frac{\partial L}{\partial b_h}
\]

(12)
\[ V = V - \eta \cdot \left( \frac{\partial L}{\partial V} + \lambda V \right) \]  
(13)

\[ b_o = b_o - \eta \cdot \frac{\partial L}{\partial b_o} \]  
(14)

With:
\[ \eta \] : Learning rate.
\[ \lambda \] : Regularization parameter.

All these processes are repeated for all training data and epoch. The test will be done by using test data and final weight matrices of training process into Forward Propagation process. Mean Squared Error will be used to see how good the network output when tested using test data.

Single Moving Average is a statistic method that can be used to forecasting time series data. It works by calculating the mean of a constant number of observations in a sliding time span and it will smooth the irregular fluctuation of data. The number of observations used is called the order of the series [13]. Previous researches showed that Single Moving Average can be used to forecast time series data [14-15]. In this research Single Moving Average with order 2 or SMA(2) is used and calculated with formula below.

\[ MA(2) = \sum_{t=t_1}^{t_2} x_t \]  
(15)

With:
\[ x_t \] : Number at index t.

3. Results and Analysis

There are 132 monthly data of foreign tourist arrival and 11 annual data for each foreign tourist arrival factor for each country which is taken from 2005 until 2015. Foreign tourist arrival factor data are divided by 12 to get the mean value for each month. All these data are divided into training and test data after normalizing process. There are 3 types of forecasting time period tested in this research, those are 1, 3, and 6 months ahead. Each time period has different amount of training data. Forecasting 1, 3, and 6 months ahead are done using 119, 117 and 114 data respectively and then tested using 12 data outside of the training data.

Web based application written in PHP and HTML is made to implement RNN BPTT and SMA(2) method. This application will train and test the network and showed the comparison results between those 2 methods and target data in chart form as shown in Figure 2 to 6.

The network consists of 4 input states where each of them represents input for foreign tourist arrivals, population of origin country, GDP of origin country and CPI of Indonesia relative to CPI of the origin country. There is 1 hidden layer with specified number of hidden state. The number of output state depends on the time period of forecasting, for example forecasting 3 months ahead will use 3 output states.

There are 36 network test configuration for each time period. The configuration variables are the amount of hidden state, number of epoch and learning rate. The number of hidden states are 5, 10 and 15. The number of epoch are 300, 400, and 500 and learning rate is 0.01, 0.05, 0.1, and 0.5. In Backpropagation Through Time process, truncated backpropagation is applied. It reduced the calculation cost because it only calculates \( k_2 \) timesteps [16]. In this research \( k_2 \) is 2, so the backpropagation process will calculate up to 2 time step only. Each of network configuration results are compared with Single Moving Average results in MSE form.

There are 10 countries tested, they are Australia, China, Malaysia, Japan, Singapore, South Korea, United Kingdom, United State of America, France and Germany. These are top 10 countries with highest foreign tourist arrival to Bali since 2011 to 2015 [1].

Table 1, 2, and 3 show the best network configuration with the smallest RNN BPTT MSE compared with SMA(2) MSE when testing those methods to forecast foreign tourist arrival to Bali. Table 1 shows the comparison between test data MSE of RNN BPTT and SMA with order 2 in 1 month ahead forecast. The comparison reveals that Australia, China, Malaysia, Japan, Singapore, South Korea, France and Germany RNN BPTT MSE are smaller than
SMA(2) MSE. This means 8 out of 10 countries or 80% of them gives better results when tested using RNN BPTT method.

<table>
<thead>
<tr>
<th>Country</th>
<th>Hidden State</th>
<th>Epoch</th>
<th>Learning Rate</th>
<th>MSE of RNN BPTT Test Data</th>
<th>MSE of SMA(2) Test Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>5</td>
<td>500</td>
<td>0.1</td>
<td>0.04835</td>
<td>0.08901</td>
</tr>
<tr>
<td>China</td>
<td>15</td>
<td>300</td>
<td>0.1</td>
<td>0.07835</td>
<td>0.09859</td>
</tr>
<tr>
<td>Malaysia</td>
<td>10</td>
<td>300</td>
<td>0.01</td>
<td>0.04479</td>
<td>0.09926</td>
</tr>
<tr>
<td>Japan</td>
<td>15</td>
<td>400</td>
<td>0.1</td>
<td>0.02258</td>
<td>0.03689</td>
</tr>
<tr>
<td>Singapore</td>
<td>10</td>
<td>300</td>
<td>0.01</td>
<td>0.06471</td>
<td>0.15260</td>
</tr>
<tr>
<td>South Korea</td>
<td>5</td>
<td>300</td>
<td>0.5</td>
<td>0.05690</td>
<td>0.10034</td>
</tr>
<tr>
<td>UK</td>
<td>10</td>
<td>400</td>
<td>0.01</td>
<td>0.09694</td>
<td>0.08209</td>
</tr>
<tr>
<td>USA</td>
<td>5</td>
<td>300</td>
<td>0.5</td>
<td>0.05883</td>
<td>0.05565</td>
</tr>
<tr>
<td>France</td>
<td>10</td>
<td>500</td>
<td>0.01</td>
<td>0.06447</td>
<td>0.08900</td>
</tr>
<tr>
<td>Germany</td>
<td>15</td>
<td>300</td>
<td>0.05</td>
<td>0.06988</td>
<td>0.09709</td>
</tr>
</tbody>
</table>

Table 2 shows the comparison between test data MSE of RNN BPTT and SMA with order 2 in 3 months ahead forecast. The comparison reveals that Australia, China, Malaysia, Japan, Singapore, South Korea and France RNN BPTT MSE are smaller than SMA(2) MSE. This means 7 out of 10 countries or 70% of them gives better results when tested using RNN BPTT method.

<table>
<thead>
<tr>
<th>Country</th>
<th>Hidden State</th>
<th>Epoch</th>
<th>Learning Rate</th>
<th>MSE of RNN BPTT Test Data</th>
<th>MSE of SMA(2) Test Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>5</td>
<td>300</td>
<td>0.1</td>
<td>0.04083</td>
<td>0.07674</td>
</tr>
<tr>
<td>China</td>
<td>10</td>
<td>300</td>
<td>0.1</td>
<td>0.07025</td>
<td>0.09814</td>
</tr>
<tr>
<td>Malaysia</td>
<td>15</td>
<td>300</td>
<td>0.01</td>
<td>0.06534</td>
<td>0.09826</td>
</tr>
<tr>
<td>Japan</td>
<td>15</td>
<td>300</td>
<td>0.1</td>
<td>0.02654</td>
<td>0.03303</td>
</tr>
<tr>
<td>Singapore</td>
<td>15</td>
<td>400</td>
<td>0.01</td>
<td>0.09591</td>
<td>0.15136</td>
</tr>
<tr>
<td>South Korea</td>
<td>5</td>
<td>300</td>
<td>0.5</td>
<td>0.05893</td>
<td>0.10125</td>
</tr>
<tr>
<td>UK</td>
<td>15</td>
<td>400</td>
<td>0.01</td>
<td>0.11138</td>
<td>0.06971</td>
</tr>
<tr>
<td>USA</td>
<td>5</td>
<td>300</td>
<td>0.5</td>
<td>0.06250</td>
<td>0.05087</td>
</tr>
<tr>
<td>France</td>
<td>10</td>
<td>400</td>
<td>0.01</td>
<td>0.07302</td>
<td>0.08615</td>
</tr>
<tr>
<td>Germany</td>
<td>5</td>
<td>400</td>
<td>0.1</td>
<td>0.10328</td>
<td>0.08465</td>
</tr>
</tbody>
</table>

Table 3 shows the comparison between test data MSE of RNN BPTT and SMA with order 2 in 6 months ahead forecast. The comparison reveals that Australia, China, Malaysia, Japan, Singapore, South Korea and France RNN BPTT MSE are smaller than SMA(2) MSE. This means 7 out of 10 countries or 70% of them gives better results when tested using RNN BPTT method.

<table>
<thead>
<tr>
<th>Country</th>
<th>Hidden State</th>
<th>Epoch</th>
<th>Learning Rate</th>
<th>MSE of RNN BPTT Test Data</th>
<th>MSE of SMA(2) Test Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>5</td>
<td>400</td>
<td>0.5</td>
<td>0.03370</td>
<td>0.05587</td>
</tr>
<tr>
<td>China</td>
<td>15</td>
<td>300</td>
<td>0.1</td>
<td>0.06417</td>
<td>0.09589</td>
</tr>
<tr>
<td>Malaysia</td>
<td>15</td>
<td>300</td>
<td>0.01</td>
<td>0.06969</td>
<td>0.09785</td>
</tr>
<tr>
<td>Japan</td>
<td>5</td>
<td>500</td>
<td>0.01</td>
<td>0.02304</td>
<td>0.02689</td>
</tr>
<tr>
<td>Singapore</td>
<td>10</td>
<td>400</td>
<td>0.1</td>
<td>0.07845</td>
<td>0.15263</td>
</tr>
<tr>
<td>South Korea</td>
<td>15</td>
<td>300</td>
<td>0.5</td>
<td>0.05356</td>
<td>0.08929</td>
</tr>
<tr>
<td>UK</td>
<td>5</td>
<td>400</td>
<td>0.01</td>
<td>0.09568</td>
<td>0.05388</td>
</tr>
<tr>
<td>USA</td>
<td>10</td>
<td>400</td>
<td>0.01</td>
<td>0.04830</td>
<td>0.04301</td>
</tr>
<tr>
<td>France</td>
<td>10</td>
<td>500</td>
<td>0.01</td>
<td>0.05771</td>
<td>0.09155</td>
</tr>
<tr>
<td>Germany</td>
<td>5</td>
<td>500</td>
<td>0.1</td>
<td>0.08870</td>
<td>0.06597</td>
</tr>
</tbody>
</table>
Figure 2, 3, 4, 5 and 6 show one of test data result of first 5 countries tested in chart form that represent forecast test comparisons of foreign tourist arrival to Bali by RNN BPTT, SMA(2) and real data. The green line represents real data, pink line represents RNN BPTT forecast and blue line represents SMA(2) forecast. The values of tourist arrivals are still in range between -1 to 1.

Figure 2. Forecast comparison chart between RNN BPTT, SMA(2), and real data of Australia

Figure 3. Forecast comparison chart between RNN BPTT, SMA(2), and real data of China

Figure 4. Forecast comparison chart between RNN BPTT, SMA(2), and real data of Malaysia
4. Conclusion

Based on the research results, the following points can be concluded:

1. Recurrent Neural Network Backpropagation Through Time can be used as a forecasting method for foreign tourist arrivals to Bali by the Balinese government due to the forecast test results for most countries tested in 1, 3, and 6 months ahead forecasts are closer to factual data than the Single Moving Average method, as shown by their MSE values.

2. From 10 countries tested, 80% of MSE for RNN BPTT is smaller than MSE of SMA(2) for 1 month ahead forecast. For 3 and 6 months ahead forecasts, 70% of MSE for RNN BPTT is smaller than SMA(2). These mean that the forecast test results are closer to factual data. These mean that the RNN BPTT method outperformed the SMA order 2 in forecasting foreign tourist arrival to Bali.

References


