Comparison Between Conventional Methods, Neural Network, and Support Vector Regression in Forecasting Foreign Tourists in Indonesia

Purwati\textsuperscript{1}, Donny Montreano\textsuperscript{2}, Alina Cynthia Dewi\textsuperscript{3}

\textsuperscript{1,2,3}Industrial Engineering, Universitas Pembangunan Nasional Veteran Jakarta, Jakarta, Indonesia
\textsuperscript{1}ppurwati997@gmail.com, \textsuperscript{2}montreano@gmail.com, \textsuperscript{3}alinacynthia@gmail.com

Abstract. Tourism is one of the important sectors in the growth of Indonesia’s economy. To be able to achieve the target and increase foreign tourist visits to Indonesia it is necessary to plan appropriate promotion and sustainable development that must be in line with the development of foreign tourists to be in target, effective and efficient. In this study forecasting to the level of foreign tourists visiting Indonesia in order to obtain accurate data. Forecasting is done by comparing 3 forecast methods, viz. traditional method, Support Vector Regression (SVR), and Backpropagation Neural Network (BPNN) method. In this research, there are 36 monthly visit data from 2017 to 2020 that are used to forecast. The result of this research indicates that the best forecasting is done by the Support Vector Regression (SVR) method with a MAPE value is 2.5614\% whereas, Backpropagation Neural Network (BPNN) has a MAPE value 31.3777\%.

1. Introduction
Tourism has an important role in economic growth in Indonesia. BPS has done the calculation and stated that the foreign tourist visit rate of 2019 is 16.1 million people while the desired target is 18 million people. At the beginning of the year, 2020 visit has decreased by 90\% because of the COVID pandemic – 19 currently being experienced, so it is necessary to do the planning and also the development of Indonesian tourism, where the planning should be in line with the level of foreign tourists visit to fit and precisely target. So it is necessary to view and information on the level of foreign visitors (tourists) in the future, it can be obtained by forecasting. Because each product has a life time limit consisting of introduction, growth, maturity, and decline [1]. The forecasting itself is an activity of making predictions on an incident that may be happening in the future and is a significant problem in various areas such as business and industry, government, economics, environmental sciences, medicine, social sciences, politics, and finance [2]. Some of the methods that will be used are Support Vector Regression (SVR) that has been performed in 2019 and obtained an error value of 2.513\% [3], and other methods used are the Backpropagation Neural Network which was also performed in 2019 and resulted in a 9.5\% error value [4]. The purpose of this research is to find out which method is suitable for forecasting on the level of foreign tourists visit Indonesia.

2. Research Methodology
This research methodology is a process or steps are taken to get the data to be used in this research.

2.1 Data needs
The data used in this research is the historical data of foreign tourists visit Indonesia from all entrances and origin of the country from May 2017 to April 2020 with the country of origin of...
the tourists themselves. Data obtained from the official page of the BPS (Central statistic) the Republic of Indonesia.

2.2 Data Processing techniques

Data processing is done with three methods namely traditional methods, support vector regression, and backpropagation. Here are the steps of each method:

1. Traditional Methods

   In conducting traditional methods forecasting, the methods used are as follows:
   a) Moving Average Method, that is evenly-flatten the demand based on the data of the past [5].

   $F_t = \frac{A_t + A_{t-1} + \cdots + A_{t-(N-1)}}{N}$  \hspace{1cm} (1)

   With,
   $F_t = \text{The result of the forecasting}$
   $A_t = \text{Actual data}$
   $N = \text{Period}$

   b) Weight Moving Average Method, that is even–flatten using weights [5].

   $F_t = \frac{W_1A_{t-1} + W_2A_{t-2} + \cdots + W_nA_{t-n}}{N}$  \hspace{1cm} (2)

   $W$ is a weight with the total weight used is 1

   c) Single Exponential Smoothing Method, i.e. forecasting using level weights ($\alpha$) [6].

   $F_t = F_{t-1} + \alpha(A_{t-1} - F_{t-1})$  \hspace{1cm} (3)

   With, $\alpha$ is a smoothing weight or constants with a value of $0 \leq \alpha \leq 1$.

   d) Double Exponential Smoothing Method, namely the development of the previous method with the level weights ($\alpha$) and trend ($\beta$) with a value of $0 \leq \alpha$ or $\beta \leq 1$ [6].

   $F_t = \alpha(A_{t-1}) + (1 - \alpha)(F_{t-1} + T_{t-1})$  \hspace{1cm} (4)

   $T_t = \beta(F_{t-1} + T_{t-1}) + (1 - \beta)T_{t-1}$  \hspace{1cm} (5)

   $FIT_t = F_t + T_t$  \hspace{1cm} (6)

   With, $FIT_t = \text{The result of forecasting}$

2. Backpropagation Method

   BackPropagation is one of the neural network models with multiple layer architecture and is composed of finding the optimal weight [7]. In the backpropagation, there is an input layer, a hidden layer, and an output layer, and the number of hidden layers is no limit but more or equal to one [8].

   The steps taken on forecasting using this method are as follows:
   a) Normalizing Data

   $x_m = \frac{x - x_{\text{min}}}{x_{\text{max}} - x_{\text{min}}}$  \hspace{1cm} (7)

   b) Practice data using the Matlab R2015a software

   c) Experiment with sigmoid binary activation function and bipolar sigmoid to select the network architecture.

   $f(\cdot) = \frac{1}{1 + e^{-\cdot}}$  \hspace{1cm} (8)

   $f(\cdot) = \frac{1 - e^{-\cdot}}{1 + e^{-\cdot}}$  \hspace{1cm} (9)

   With, $\cdot$ Is the initialization weight

   d) Validate with the selected architecture

   e) Perform denormalization of data

   $x = (x_m(x_{\text{max}} - x_{\text{min}})) + x_{\text{min}}$  \hspace{1cm} (10)

   f) Calculate MAD, MSE, and MAPE error values
3. Support Vector Regression Method

Support Vector Regression is a development of the Support Vector Machine that applies to the issue of both linear and non-linear regression [9]. The concept of SVR is to obtain values of good forecasting, this is because SVR can adjust the problem of overfitting that is the data behavior in the training phase that will obtain the value of high forecasting accuracy [10].

The steps taken on this method are as follows:

a) Normalizing data
b) Calculate the values of the kernel

\[ K(x_i, x_j) = \exp \left( -\frac{1}{2\sigma^2} \right) (x_i - x_j) \]  

With, \( x_i \) = Train data, \( x_j \) = Test data, and \( \sigma \) = Standard Deviation
c) Calculating the Hessian matrix

\[ R_{ij} = \left( \frac{\partial^2}{\partial x_i \partial x_j} + \lambda^2 \right) \]  

With, \( R_{ij} \) = Hessian Matrix, \( K \) = Value of kernel, and \( \lambda \) = Scalar variable
d) Calculating the value of E

\[ E = y_i - \sum_{j=1}^{n} (\alpha_j^* - \alpha_j)R_{ij} \]  

With the initial \( \alpha \) initialization is 0, repeat until obtaining the optimal \( \alpha \) value.

\[ \delta \alpha_i = \min \{ \max \{ y_i (E_i - \varepsilon) - \alpha_i \}, C - \alpha_i \} \]  

\[ \delta \alpha_i^* = \min \{ \max \{ y_i (E_i - \varepsilon) - \alpha_i^* \}, C - \alpha_i^* \} \]  

\[ \alpha_i^* = \alpha_i^* + \delta \alpha_i^* \]  

\[ \alpha_i = \alpha_i + \delta \alpha_i \]  

With, \( E \) = Error value, \( Y_i \) = Train data, \( \alpha = Lagrange\ multiplier, \gamma = Value\ of\ learning\ rate, \varepsilon = Epsilon\ parameter, \text{ and } C = Complexity\ parameter \]

cLR =  \[ \frac{1}{\gamma} \max(\text{Matriks Hessian}) \]  

cLR is the value of the learning rate coefficient.
e) Calculate prediction value

\[ f(x) = \sum_{j=1}^{n} (\alpha_j^* - \alpha_j) \left( K(x_i, x_j) + \lambda^2 \right) \]  

f) Perform denormalization of data
g) Calculate MAD, MSE, and MAPE error values

3. Result and Discussion

Data on foreign tourists visit in the period of 1 (May 2017 and April 2018) have an average of 1,219,864 people with the nation of Malaysia as the country with the largest visit is 16.17%, in period 2 (May 2018 – April 2019) has an average of 1,323,424 people with the state of Malaysia as the country with the largest level of visit is 16.47%, and in the period 3 (May 2019 – April 2020) has an average of 1,153,511 people with the country Malaysia as the country with the largest rate of visit is 18.31%.

3.1 Moving Average Method

Forecasting is done using an average 3-month average period obtained with the following error values:

| Tabel 1. Error value of Moving Average |
|----------------------------------------|
| MSE         | MAD    | MAPE   |
| 51,132,338,195 | 153,477 | 0.26893 |
Based on the error values obtained forecasting using this method has a moderate level of accuracy, with the results of forecasting in May 2020 obtained is 498.324 inhabitants.

3.2 Weight Moving Average Method

In this method the treatment is carried out with an average period of 3 monthly and the weight value used is W1 0.1, W2 0.1, and W3 0.8. The error values obtained are as follows:

Table 2. Error value of Weight Moving Average

| MAD     | MSE          | MAPE  |
|---------|--------------|-------|
| 129,722 | 30,291,428,599 | 0.19885 |

Based on the error values obtained forecasting using this method has a good level of accuracy, with forecasting in May 2020 is 261.527 inhabitants.

3.3 Single Exponential Smoothing Method

In this method, the α value used is 0.05, and the resulting value error is as follows:

Table 3. Error value of Single Exponential Smoothing

| MAD     | MSE          | MAPE  |
|---------|--------------|-------|
| 166,704 | 75,139,977,280 | 0.32689 |

Based on the error value, this method has a moderate level of accuracy, with forecasting results in May 2020 is 1,175,002 inhabitants.

3.4 Double Exponential Smoothing Method

This method is a continuation of the previous method in which the method uses α 0.05 and β values of 0.6, and the following error values are obtained:

Table 4. Error value of Double Exponential Smoothing

| MAD     | MSE          | MAPE  |
|---------|--------------|-------|
| 57,553  | 7,715,384,997 | 0.10157 |

Based on the error values obtained by forecasting using this method has good accuracy, and the forecasting results obtained in May 2020 is 233.782 inhabitants.

3.5 Backpropagation Method

Forecasting calculations using this method are done in several steps, namely:

1. Network training

The training was conducted using Matlab R2015a software with the provisions of learning rate 0.1, momentum 0.9, and epoch 100. This stage is done to compare the matching activation function and the number of hidden neurons.

a. Binary sigmoid

Table 5. The network

| Architecture | Hidden Neuron | MSE       |
|--------------|--------------|-----------|
| 1 – 2 – 1    | 2            | 0.0030013 |
| 1 – 3 – 1    | 3            | 0.065175  |
| 1 – 4 – 1    | 4            | 0.010959  |
| 1 – 10 – 1   | 10           | 0.0082761 |

b. Bipolar sigmoid

Table 6. The network

| Architecture | Hidden Neuron | MSE       |
|--------------|--------------|-----------|
| 1 – 2 – 1    | 2            | 0.080783  |
| 1 – 3 – 1    | 3            | 0.008364  |
| 1 – 4 – 1    | 4            | 0.018923  |
| 1 – 10 – 1   | 10           | 0.067188  |

The training results of each activation function with the smallest MSE value are as follows:
2. Validation
This stage is done to keep the output value with the desired target and select the network with the smallest error value. The results obtained are as follows:

| Table 8. Network selection |
|-----------------------------|
|                            | Training | Validation |
| Activation | Network   | MSE | MSE   |
| Sig. Binary | 1 – 2 – 1 | 0.0030013 | 0.03559 |
| Sig. Bipolar | 1 – 3 – 1 | 0.008364 | 0.03806 |

According to the table, the selected network is 1 – 2 – 1 with the binary sigmoid activation function.

3. The result
Based on the network architecture and activation function used is obtained the error value as follows:

| Table 9. Error value of Backpropagation |
|-----------------------------------------|
| MAD | MSE | MAPE |
| 161,284 | 68,486,289,879 | 0.313777 |

Based on the error values obtained forecasting using this method has a moderate level of accuracy, and the forecasting result in May 2020 is 1,176,567 inhabitants.

3.6 Support Vector Regression Method
In this method of calculation is done by experimenting on each parameter used, namely:

1. Sigma parameters (\( \Sigma \))
   In the test Sigma value there are 10 values tested, and specified in advance the value of other parameters, namely:
   a. Lamda = 0.01
   b. cLR = 0.4
   c. Epsilon = 0.001
   d. C = 10

| Table 10. Sigma value testing |
|------------------------------|
| No | Sigma | MAD    |
|----|-------|--------|
| 1  | 5     | 0.734749 |
| 2  | 2.4   | 0.223539 |
| 3  | 2.3   | 0.184188 |
| 4  | 1     | 2.998340 |

Based on the results, the value of sigma that has the smallest error value is 2.3 and the smaller the value of the Sigma error is greater.

2. Lamda Parameter (\( \lambda \))
   At the test Lamda value is also performed as much as 10 lamda values, with a predefined sigma value, here is another parameter value used:
   a. Sigma = 2.3
   b. Clr = 0.4
   c. Epsilon = 0.001
   d. C = 10

| Table 11. Lamda value testing |
|-----------------------------|
| No | Lamda | MAD |
|----|-------|-----|
| 1  | 1     | 0.576721 |
| 2  | 0.04  | 0.182094 |
| 3  | 0.01  | 0.184180 |
Based on a test the LAMDA value that has the smallest error value is 0.04. And it can be known that the large value of the LAMDA is the greater the error value gained.

3. Coefficient of Learning Rate (Clr) parameters
On testing the Clr parameter is done with 10 values, and specified the previous determinations, here is another parameter value:

| No | Clr | MAD    |
|----|-----|--------|
| 1  | 0.1 | 0.025148 |
| 2  | 0.09| 0.024971 |
| 3  | 0.07| 0.052353 |
| 4  | 0.01| 0.499810 |

Based on the test done, the Clr that generates the smallest error value is 0.09, and it can be seen that the smaller the Clr value the larger the value of the indicator.

4. Epsilon (ε) parameter
In testing this parameter, the value of the other parameter used is the value that has been tested, namely:

| No | Epsilon | MAD    |
|----|---------|--------|
| 1  | 1       | 0.73255 |
| 2  | 0.1     | 0.115776 |
| 3  | 0.01    | 0.033227 |
| 4  | 0.001   | 0.024971 |

From the tests that have been done, the CLR value with a small error value is 0.001 and it is also known that the greater the CLR value eats the larger the resulting error value.

5. Complexity (C) parameter
Testing the last parameter to be performed is C, and the other parameter value used is the previously tested parameter:

| No | C    | MAD    |
|----|------|--------|
| 1  | 10   | 0.0243646 |
| 2  | 1    | 0.0243646 |
| 3  | 0.01 | 0.548520 |
| 4  | 0.001| 0.710609 |

From the test results done the correct C value to produce a small error value is 1, and seen from the test result that the smaller the C value then the error value will be greater.

Based on the testing of the parameters, the parameters obtained with the smallest error value are as follows:

| a. Sigma | 2.3 | c. Clr  | 0.09 |
| b. Lamda | 0.04| d. Epsilon | 0.0001 | e. C | 1 |
Based on the parameters used, the error value is obtained as follows:

**Table 15. Error value of Support Vector Regression**

| Method               | MAD      | MSE                  | MAPE    |
|----------------------|----------|----------------------|---------|
| Support Vector       | 33,798   | 2,528,148,044        | 0.025614|
| Regression           |          |                      |         |

From the error value obtained, this method has a high level of forecasting accuracy, and the result of forecasting in May 2020 is 164,723 inhabitants.

3.7 Evaluate Error Value

The results obtained from the forecasting calculations of each method are as follows:

**Table 16. Error Value**

| Methods                  | MAD      | MSE                  | MAPE    | Note    |
|--------------------------|----------|----------------------|---------|---------|
| Moving Average           | 153,477  | 51,132,338,195       | 0.26893 | Moderate|
| Weight Moving Average    | 129,722  | 30,291,428,599       | 0.19885 | Good    |
| Single Exponential       | 166,704  | 75,139,977,280       | 0.32689 |         |
| Smoothing                |          |                      |         |         |
| Double Exponential       | 57,553   | 7,715,384,997        | 0.10157 | Good    |
| Smoothing                |          |                      |         |         |
| Backpropagation Neural   | 161,284  | 68,486,289,879       | 0.313777| Moderate|
| Network                  |          |                      |         |         |
| Support Vector Regression| 33,798   | 2,528,148,044        | 0.025614| High    |

Based on the results of the error value obtained from each forecasting method, it can be seen that the Support Vector Regression method has the smallest error value of MAD 33.798, MSE 2,528,148,044, and MAPE 0.025614, so this method of forecasting has a high level of accuracy.

4. Conclusion

From the calculations that have been done can be said that:

1. Forecasting foreign tourists visit Indonesia is done using several methods, namely:
   a. Moving Average Method
      The best Moving average models are obtained with a period of 3 monthly and produce an error value of MAD 153.477, MSE 51,132,338,195, as well as the value of MAPE 0.26893, so forecasting using this method can be said to have a moderate level of accuracy.
   b. Weight Moving Average Method
      The best Moving Average Weight Model with a period of 3 monthly, and the weight of W1 0.1, W2 0.1, and W3 0.8 resulted in MAD 129.722, MSE 30,291,428,599, and MAPE 0.19885 error values, so forecasting has a good level of accuracy.
   c. Single Exponential Smoothing Method
      The best Single Exponential Smoothing Model with α 0.05 results in MAD 166.704, MSE 75,139,977,280, and MAPE 0.32689 error values, so forecasting has a moderate level of accuracy.
   d. Double Exponential Smoothing Method
      The best Double Exponential Smoothing Model with α 0.05 and β 0.6 weighs an error value of MAD 57.553, MSE 7,715,384,997, and MAPE value 0.10157, so forecasting has a good level of accuracy.
   e. Backpropagation Method
      Best Backpropagation Model with 1 – 2 – 1 network and the binary sigmoid activation function results in MAD 161.284, MSE 68,486,289,879, and MAPE 0.313777 error values, thus having a moderate level of accuracy.
   f. Support Vector Regression Method
The best Vector Regression Support models with Parameters Sigma 2.5, Lamda 0.04, CLR 0.09, Epsilon 0.0001, and C 10 generate the error values of MAD 33.798, MSE 2,528,148,044, and MAPE 0.025614 so this forecasting has a high level of accuracy.

2. Based on research known that the best method for forecasting the level of foreign tourists visit Indonesia is the method of Support Vector Regression with the smallest error rate of 33.798, MSE 2,528,148,044, and MAPE value of 2.5614%. So this forecasting has a high degree of accuracy.

3. The forecasting obtained in May 2020 with the Support Vector Regression method is 164,723 inhabitants.

References

[1] Kotler, P., & Keller, K. L., 2016. "Marketing Management", Global Pearson, USA.
[2] Montgomery, D, Jennings, C & Kulahci, M, 2015. “Introduction to Time Series Analysis and Forecasting”, 2nd edn, John Wiley & Sons, Inc., New Jersey.
[3] Surtiningsih, L., Furoq, M & Adinugroho, S., 2018. ‘Prediksi Jumlah Kunjungan Wisatawan Mancanegara ke Bali Menggunakan Support Vector Regression dengan Algoritma Genetika’. Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer, vol. 2, no. 8, pp. 2578-2586.
[4] Misriati, T, 2016. ‘Peramalan Kunjungan Wisatawan Mancanegara ke Lombok Menggunakan Jaringan Syaraf Tiruan’, SNIPTEK.
[5] Makridakis, S. & Wheelwright, S., 1999. “Forecasting Methods for Management”, John Wiley & Sons Inc, New Jersey.
[6] Jacobs, Chase & Aquilano, 2009. “Operation & Supply Management”, Mc Graw Hill, New York.
[7] Sawitri, M, Sumarjaya, I, W & Tastrawati, N, 2018, ‘Peramalan Menggunakan Metode Backpropagation Neural Network’, Jurnal Matematika, vol. 7, no. 3, pp. 264-270.
[8] Anderson, D & McNeill, G, 1992. “Artificial Neural Network Technology”, Kaman Sciences Corporation, New York.
[9] Hu, X, 2017. “Support Vector Machine and Its Application to Regression and Classification”, MSU Graduate Theses, Missouri State University, US.
[10] Maulana, N, D., Setiawan, B, D & Dewi, C, 2019. ‘Implementasi Metode Support Vector Regression (SVR) dalam Perusahaan Penjualan Roti (Studi Kasus: Harum Bakery)’, Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer, vol. 3, no. 3, pp. 2986-2995.