



## Travel Time Estimation Using Support Vector Regression on Model with 8 Features

Rifki Kosasih<sup>1\*</sup>, Iffatul Mardhiyah<sup>2</sup>

<sup>1,2</sup>Department of Informatics, Faculty of Computer Science and Information Technology, Gunadarma University, Indonesia

### Abstract.

**Purpose:** In travelling, we need to predict travel time so that itinerary is as expected. This paper proposes Support Vector Regression (SVR) to build a prediction model. In this case, we will estimate travel time in the Bali area. We propose to use a regression model with 8 features, i.e., time, weather, route, wind speed, day, precipitation, temperature and humidity information.

**Methods:** In this study, we collect real-time data from Global Positioning System (GPS) and weather applications. We divide our data into two types: training dataset consisting of 177 data and testing dataset comprising 51 data. The Support Vector Regression (SVR) method is used in the training stage to build a model representing data. To validate the model, error measurements were carried out by calculating the values of  $R^2$ , Accuracy, MAE (Mean Absolute Error), RMSE (Root Mean Square Error) and Accuracy.

**Result:** From the research results, the model obtained is the SVR model with parameters  $\gamma = 0.125$ ,  $\epsilon = 0.1$  and  $C = 1000$ , which has a value of  $R^2 = 0.9860528612283006$ . Later, we predict travel times on testing data using the SVR model that has been obtained. Based on the result of the research, our model has a 0.8008 MAE (Mean Absolute Error), 1.2817 RMSE (Root Mean Square Error) and 95.3369% Accuracy.

**Novelty:** In this study, we use 8 features to estimate travel time in the Bali area. Furthermore, we will compare the KNN regression method (previous studies) with Support Vector Regression (SVR) (proposed method) on a model with 8 features to estimate travel time.

**Keywords:** Travel Time, Support Vector Regression, Features, GPS

**Received** June 2022 / **Revised** June 2022 / **Accepted** October 2022

This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).



### INTRODUCTION

In the tourism industry, especially in travel agents, it is very important to make travel plans, i.e., determining the number of attractions to be visited in one day, estimating travel time to the destination place and others. Predicting travel time is very important for travel agents to determine travel rates [1]. To estimate the travel time to attractions, travel agents use patterns of the trip they've ever been through earlier in the day but have not paid attention to the factors that affect the travel time. Many factors can affect the duration of the trip. In this study, we propose 8 features that affect travel time, i.e., time, weather, route, wind speed, day, precipitation, temperature and humidity information.

Time is a factor that can affect travel time because there is a pattern of rush hours, i.e., in the morning, many people will move to work, which can cause the volume of vehicles on the road to be more crowded and affect travel time [2]. Furthermore, weather can also affect travel time, i.e., when it rains, a lot of vehicles will slow down their speed vehicle so that it can cause the travel time to be longer. Routes information can also affect travel times, such as during peak hours on certain routes there will be congestion, so alternative routes are needed, so that travel times are faster. Later wind speed can affect the travel time, i.e., if the wind speed is fast, it will cause the vehicle's speed to slow so that the travel time becomes longer [2].

---

\*Corresponding author.

Email addresses: [rifki\\_kosasih@staff.gunadarma.ac.id](mailto:rifki_kosasih@staff.gunadarma.ac.id) (Kosasih), [iffatul@staff.gunadarma.ac.id](mailto:iffatul@staff.gunadarma.ac.id) (Mardhiyah)

DOI: [10.15294/sji.v9i2.37215](https://doi.org/10.15294/sji.v9i2.37215)

The next factor, information on the day also affects because there is usually a travel pattern on certain days such as on weekends, the road to the tourist attraction may be crowded so that the travel time will be longer. Furthermore, precipitation information can also affect travel time, i.e., if precipitation in your area is high, it can make people not want to leave the house. This can result in the volume of vehicles coming out to be a little so that travel time becomes faster. Then the temperature information, i.e., if the temperature in your area is extreme, can make people stay at home, which causes the volume of the vehicle to be less crowded and speed up travel time. The last factor is humidity, if the humidity in your area is high then it can cause the environment to feel hot so it can make people not want to leave the house which will cause the vehicle volume to be less crowded so that travel time becomes faster [2]. After the 8 features are determined, then we will predict the travel time with cases in the Bali area. In this research, we propose using the Support Vector Regression method to predict the travel time.

Several studies have been carried out to predict travel time, i.e., [3] uses a dynamic model to predict the bus waiting time in the last halte with the K-Nearest Neighbor (KNN) algorithm. The data used are Global Positioning System (GPS) data on the bus, Transit Capacity and Quality of Service Manual, passenger volume up and down, and fare payment methods. After that, an evaluation is performed by calculating Mean Absolute Error (MAE). From the research, the MAE was calculated at each bus stop with MAE values varying from 0.11 to 8.5 [3].

Research [4] discusses travel time of various kinds on city streets and variance estimation methods by analyzing the data set derived from the travel time detected from automatic license plate readers installed in the Beijing region. Several factors affect travel time, i.e., traffic incidents, rush hours, work zones, bad weather, special events, and fluctuations in traffic demand. High variability indicates unpredictable travel times and reduced reliability of traffic services [4].

Furthermore, [5] use the random forest to predict travel time. There are 14 attributes that affect travel time obtained from VISSIM simulation. VISSIM is a software used for traffic simulation developed by PVP in Germany. Later, feature selection was performed using Random Forest and obtained 7 variables that were important in predicting travel time, i.e., average of travel time, traffic conditions, vehicle density, median of travel time, vehicle speed, car density and speed. His research obtained an error estimate with OOB (Out-of-Bag) at 5.586 [5]. However, this research focused on travel time estimation from 1 city to another using 1 route.

The next study, [6] predicts travel time using the K-Nearest Neighbor algorithm. The data used comes from GPS placed on buses in the Chennai region, India. In his research, an evaluation is performed by calculating the MAPE. MAPE calculations were carried out on several trip variations and obtained errors varying from 11.68 to 29.60% [6].

The next research, [7] builds a model to predict travel time using vehicle data and the detector system on the automatic toll road. By using these two data, the model can increase prediction Accuracy. The data used are the number plate detection system, moving vehicle detection system, and GPS located on the vehicle. Furthermore, the evaluation of the model is carried out by calculating the MAE and MAPE. The MAE and MAPE values obtained were 4.74 and 7.28 [7].

Later, [8] uses a fuzzy nervous system to estimate travel times. The data used is traffic flow data, i.e., the number of vehicles and vehicle speed obtained from the loop detector. The test is done by calculating MAE, RMSE, and Mean Absolute Relative Error (MARE). In his research, the MAE value was 3.23, RMSE was 3.96 and MARE was 0.97 [8].

Research [9] predicts vehicle travel time using a regression model. The data used is vehicle data which consists of 4 features, i.e., personal, traffic, temporal and spatial information [9]. Furthermore, [2] uses a regression model consisting of 8 features/variables, i.e., zone, time, day, weather, temperature, wind speed, humidity and rainfall information using the KNN regression method to predict travel time. From the results of his research, it was found that the prediction Accuracy rate of 88.19% [2].

A few researchers focused on travel time estimation using a small number of features (less than 8 features), however, there are still many features that should be used to predict travel times. Therefore, this research is focused on using many features that can affect travel time. In this study, we propose using 8 features. A

previous researcher [2] focused on travel time estimation with 8 features and KNN regression. Therefore, as a comparison, this research intends to estimate the travel time with 8 features using support vector regression. Previous research [4]–[6] focused on travel time estimation from 1 city to another using 1 route. While in this study intends to predict travel time from 1 city to another city by using 3 routes so that from these predictions, transportation users such as travel agents can choose the fastest route to their destination. This research aims to compare travel time using the KNN regression method (previous method) and Support Vector Regression (proposed method).

Regression models usually use the Ordinary Least Square (OLS) method to create models that fit the data. However, if the data used is nonseparable data, the SVR method can produce better predictions than the regression model [10], [11].

## METHODS

In this study, we propose the Support Vector Regression (SVR) method to predict travel time. The stages of this study as shown in Figure 1.

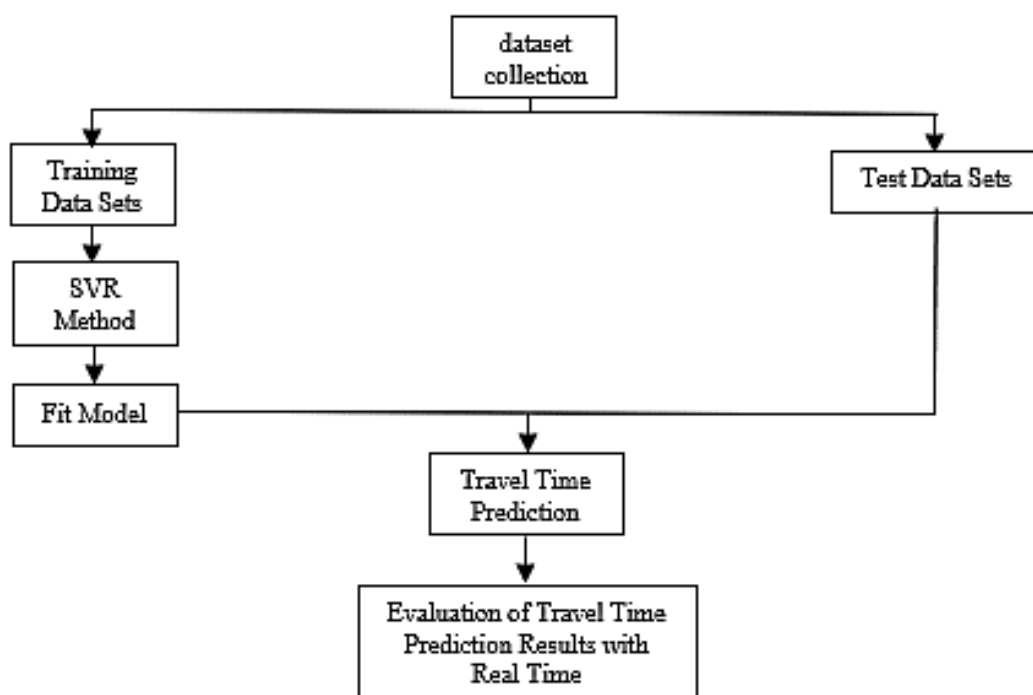


Figure 1. The stages of this research

Based on Figure 1, the first step is collecting data. We collect travel data via GPS and weather information based on weather applications in the Bali region, as shown in Figure 2. In this study, we use 228 data and divide the data into two types with a composition of 177 data as training data collected in the range 25 July 2019 to 1 August 2019 in the Bali region, i.e., from Ngurah Rai Airport to Kuta Beach while 51 data as testing data that collected in the range 9 August 2019 to 12 August 2019.

In this study, we collect 8 features based on GPS, i.e., time, route, day as in Figure 2(a) and weather applications, i.e., weather, wind speed, precipitation, temperature and humidity information as in Figure 2(b). Time information is categorized into three types, i.e., morning, afternoon and evening. This time information is very important to know because there is a time pattern that often occurs in traffic jams, such as when people are active in the morning which can affect congestion.

To get to the Kuta Beach location, the route used has 3 options so that users can choose the best route to arrive on time. For route information, three routes are used from the Ngurah Rai Airport to Kuta Beach, while the days are categorized into 7 days from Monday to Sunday. In addition, if when travelling the wind speed is very high, precipitation, temperature and humidity are also high, it will affect the travel time.

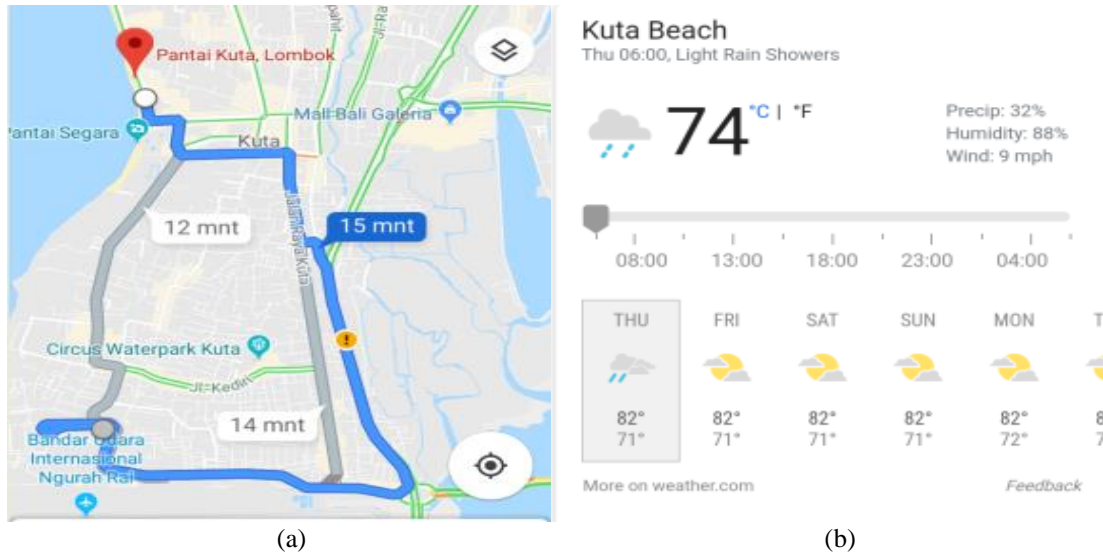


Figure 2. Feature retrieval using (a) GPS and (b) weather application

Weather information is categorized into 30 types as shown in Table 1. Changes in the weather in the area can also affect travel times.

Table 1. Code of weather

Code	Weather	Code	Weather	Code	Weather
1	sunny (morning)	11	clear with periodic (day)	21	mostly cloudy (night)
2	sunny (day)	12	clear with periodic (night)	22	change of shower (morning)
3	sunny (night)	13	cloudy (morning)	23	change of shower (day)
4	mostly sunny (morning)	14	cloudy (day)	24	change of shower (night)
5	mostly sunny (day)	15	cloudy (night)	25	shower (morning)
6	mostly sunny (night)	16	partly cloudy (morning)	26	shower (day)
7	clear (morning)	17	partly cloudy (day)	27	shower (night)
8	clear (day)	18	partly cloudy (night)	28	light rain shower (morning)
9	clear (night)	19	mostly cloudy (morning)	29	light rain shower (day)
10	clear with periodic (morning)	20	mostly cloudy (day)	30	light rain shower (night)

In the next stage, we apply the Support Vector Regression (SVR) method to training data to build a model that can represent data.

### Support Vector Regression (SVR)

Support Vector Machine (SVM) is one of the methods of machine learning used in classification problems that maps  $N$  samples that are free from each other to the higher dimensional space [12]–[15]. However, in its development the SVM method can also be used in regression problems. We apply the SVM method to the regression problem is called Support Vector Regression (SVR). This method is very good because it can be used in data overfitting problems.

There is a training set  $S = \{(x_1, y_1), (x_2, y_2), \dots (x_m, y_m)\}$  with  $x = \{x_1, x_2, \dots, x_n\} \in R^n$  is input data and  $y = \{y_1, y_2, \dots\} \in R$  is output data.  $A(x)$  is a nonlinear mapping that maps the  $x$  vector input to the vector space, so linear regression in high dimensional vector space can be performed [16]. The regression function can be seen in Equation (1) [17].

$$y = f(x) = w \cdot A(x) + b \quad (1)$$

The regression model of the linear function in Equation (1) is Equation (2)

$$L(w, b) = \sum_{i=1}^n (y_i - wx_i - b)^2 \quad (2)$$

with:

$$s. t. y_i - wx_i - b \leq \varepsilon + \xi, (w, b) \in R^{n+1}, \xi_i \in R^m$$

The regression function  $f(x)$  is used to get a model that matches the training data. The Ordinary Least Squares approach is used to select parameters  $(w, b)$  that minimize the sum of the squares deviations from data [10], [18]. These problems can be formulated as an unconstrained minimization problems as in Equation (3)

$$\min_{w \in R^n} \frac{1}{2} \|w\|^2 + C \left( \frac{1}{l} \sum_{i=1}^l (\xi_i + \xi_i^*) \right) \quad (3)$$

with:

$$\begin{aligned} s. t. y_i - wx_i - b &\leq \varepsilon + \xi_i, i = 1, 2, \dots, l \\ y_i - wx_i - b &\leq \varepsilon + \xi_i^*, i = 1, 2, \dots, l \\ \varepsilon, \xi_i, \xi_i^* &\geq 0, i = 1, 2, \dots, l \end{aligned}$$

where:

$$w = \sum_{j=1}^n (\alpha_j - \hat{\alpha}_j) A(x_j) \quad (4)$$

$w$  is the weight vector with dimension  $n+1$  and constant  $C > 0$  to determine the trade-off between differences in the function of the decision, where the tolerance satisfies the upper limit of the deviation of more than  $\varepsilon$  [19], [20]. If deviations are more than  $\varepsilon$  then penalty of  $C$  will be given. We substitute Equation (4) into Equation (2) so that we get Equation (5).

$$y = \sum_{j=1}^n (\alpha_j - \hat{\alpha}_j) A(x_j) \cdot A(x) + b \quad (5)$$

$A(x_j) \cdot A(x)$  is defined as  $k(x_j, x)$  which is a kernel function so that Equation (5) becomes Equation (6)

$$y = \sum_{j=1}^n (\alpha_j - \hat{\alpha}_j) k(x_j, x) + b \quad (6)$$

Examples of commonly used kernel functions are polynomial kernel function, radial basis function, Sigmoid function, etc [21]–[23]. In this study, we use the radial basis function kernel that can be seen in Equation (7) because it provides the best performance to predict the load compared to other kernels [24]–[26].

$$K(x, x') = e^{-\gamma \|x - x'\|^2} \quad (7)$$

In SVR, support vectors is the value of training data that is on or outside the boundary decision function. Therefore, the number of support vectors will decrease as the error value  $\varepsilon$  increases [27], [28]. In a dual formulation, the SVR optimization problem can be seen in Equation (8)

$$\max \left[ -\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (\alpha_i - \hat{\alpha}_i) (\alpha_j - \hat{\alpha}_j) k(x_i, x_j) + \sum_{i=1}^n (\alpha_i - \hat{\alpha}_i) y_i - \varepsilon \sum_{i=1}^n (\alpha_i - \hat{\alpha}_i) \right] \quad (8)$$

with: s.t.  $\sum_{i=1}^n (\alpha_i - \hat{\alpha}_i) = 0, 0 \leq \alpha_i \leq C, 0 \leq \hat{\alpha}_i \leq C,$   
where:  $\alpha_i$  and  $\hat{\alpha}_i$  are Lagrange multiplier.

The following is the algorithm of SVR:

Algorithm of Support Vector Regression:

1. Training data  $S = \{(x_1, y_1), \dots, (x_m, y_m)\}$ , with  $x = \{x_1, x_2, \dots, x_n\} \in R^n$  and  $y = \{y_1, y_2, \dots\} \in R$
2. Select the values  $C, \gamma$  and  $\varepsilon$
3. Use the RBF Kernel function
4. Build and solve the optimization problem in (3) to get the optimal solution:  $\alpha = (\alpha_1, \hat{\alpha}_1, \dots, \alpha_l, \hat{\alpha}_l)^T$
5. Build the regression function as in Equation (4), namely:  $y = \sum_{j=1}^n (\alpha_j - \hat{\alpha}_j) k(x_j, x) + b$

### Fit Model

In regression problems, making a model in accordance with the training data is needed to get a good predictive value. To find out that the model used is in accordance with the training data, it is necessary to test, one of which is to calculate the coefficient of determination  $R^2$  [29]. The calculation of  $R^2$  can be seen in Equation (9)

$$R^2 = \frac{(\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}))^2}{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2} \quad (9)$$

The value of  $R^2$  represents what percentage of the model can represent the data [29].

### Evaluation Model

In the final stage of this research, we perform evaluate on predicted travel time by using RMSE [30] in Equation (10), Mean Absolute Error (MAE) [6] in Equation (11) and Accuracy [2], [6] in Equation (12).

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{N}} \quad (10)$$

$$MAE = \frac{\sum_{i=1}^N |y_i - \hat{y}_i|}{N} \quad (11)$$

$$Accuracy = \left( 1 - \frac{\sum_{i=1}^N \frac{|y_i - \hat{y}_i|}{y_i}}{N} \right) \times 100\% \quad (12)$$

Where  $N$  is the number of testing dataset,  $y_i$  is real travel time and  $\hat{y}_i$  is travel time prediction.

## RESULT AND DISCUSSION

In previous studies, travel time was predicted by using the KNN regression method [2]. As a comparison, in this study, we develop another model using the support vector regression (SVR) method to predict travel times. We collect the data which consists of 8 features in the Bali area and divide the data into 177 training data as in Table 2 and 51 testing data as in Table 3.

Table 2. Training dataset (21 of 177 training dataset)

No.	Date	Route	Day	Coded Day	Coded Weather	Temperature	Wind	Humidity	Precipitation	Time Travel
1.	25/07/2019 6:09	1	Thursday	5	28	74	9	0.88	0.32	12
2.	25/07/2019 6:09	2	Thursday	5	28	74	9	0.88	0.32	14
3.	25/07/2019 6:10	3	Thursday	5	28	74	9	0.88	0.32	15
4.	26/07/2019 7:27	1	Friday	6	16	75	5	0.88	0.30	13
5.	26/07/2019 7:27	2	Friday	6	16	75	5	0.88	0.30	14
6.	26/07/2019 7:27	3	Friday	6	16	75	5	0.88	0.30	14
7.	27/07/2019 7:40	1	Saturday	7	16	77	11	0.84	0.10	12
8.	27/07/2019 7:40	2	Saturday	7	16	77	11	0.84	0.10	13
9.	27/07/2019 7:40	3	Saturday	7	16	77	11	0.84	0.10	14
10.	28/07/2019 8:18	1	Sunday	1	19	77	11	0.84	0.00	14
11.	28/07/2019 8:18	2	Sunday	1	19	77	11	0.84	0.00	14
12.	28/07/2019 8:18	3	Sunday	1	19	77	11	0.84	0.00	15
13.	29/07/2019 13:04	1	Monday	2	20	79	9	0.83	0.24	17
14.	29/07/2019 13:04	2	Monday	2	20	79	9	0.83	0.24	19
15.	29/07/2019 13:04	3	Monday	2	20	79	9	0.83	0.24	17
16.	30/07/2019 7:27	1	Tuesday	3	16	77	9	0.88	0.16	12
17.	30/07/2019 7:27	2	Tuesday	3	16	77	9	0.88	0.16	14
18.	30/07/2019 7:27	3	Tuesday	3	16	77	9	0.88	0.16	14
19.	31/07/2019 6:24	1	Wednesday	4	16	76	8	0.90	0.24	12
20.	31/07/2019 6:24	2	Wednesday	4	16	76	8	0.90	0.24	13
21.	31/07/2019 6:24	3	Wednesday	4	16	76	8	0.90	0.24	13

Based on Table 2, training data were obtained from 25 July 2019 to 1 August 2019 in the Bali region. In Table 2, there are 8 features which used, i.e., time, route, day, weather, temperature, wind speed, humidity and precipitation information. The time feature has data of the numeric type and is taken from 6 am to 23

o'clock. The route feature has data of category type consisting of 3 different routes obtained from GPS. The day feature has data of category type which consists of 7 categories of days from Monday to Sunday. The weather feature has data of category type which consists of 30 categories as in Table 1.

The temperature has data of the numeric type and the temperature unit used is Fahrenheit. The wind speed feature has data of the numeric type and the wind speed unit used is miles per hour (mph). The humidity and precipitation have data of numeric type. These features are obtained based on observations from the weather application. Each training data contains travel time which is used as historical data. The last column in Table 2 is the travel time from Ngurah Rai Airport to the Kuta Beach which has units in minutes.

Table 3. Testing dataset (10 of 51 testing dataset)

No.	Date	Route	Day	Coded Day	Coded Weather	Temperature	Wind	Humidity	Precipitation
1.	09/08/2019 8:59	1	Friday	6	19	77	16	0.87	0.33
2.	10/08/2019 5:50	3	Saturday	7	10	74	5	0.90	0.10
3.	10/08/2019 9:07	1	Saturday	7	16	78	11	0.84	0.07
4.	10/08/2019 10:00	2	Saturday	7	4	81	11	0.73	0.06
5.	10/08/2019 12:10	1	Saturday	7	5	83	14	0.69	0.01
6.	11/08/2019 8:25	1	Sunday	1	19	76	3	0.87	0.10
7.	11/08/2019 11:27	2	Sunday	1	16	82	11	0.74	0.09
8.	11/08/2019 18:32	3	Sunday	1	12	79	11	0.84	0.21
9.	12/08/2019 8:34	1	Monday	2	4	79	7	0.87	0.00
10.	12/08/2019 13:36	1	Monday	2	5	84	17	0.75	0.10

Based on Table 3, data was obtained from 9 August 2019 to 12 August 2019. Table 3 is a sample of testing data which consists of 8 features used to predict travel time based on training data. In the next stage, we build a regression model with Equation (13).

$$y = f(x) = w \cdot A(x) + b \quad (13)$$

where:

$A(x)$  is a matrix that represents features.

The support vector regression (SVR) method is used to build this prediction model. In this study, the parameter can be calculated by Equation (14).

$$\gamma = \frac{1}{\text{number of feature}} \quad (14)$$

$$\gamma = \frac{1}{\text{number of feature}} = \frac{1}{8}$$

was used, parameter  $\varepsilon = 0.1$  and parameter C which varied from 1, 2, 3, ..., 1000. To find out the performance of the proposed model, the  $R^2$  value for the various C values is calculated as in Figure 3.

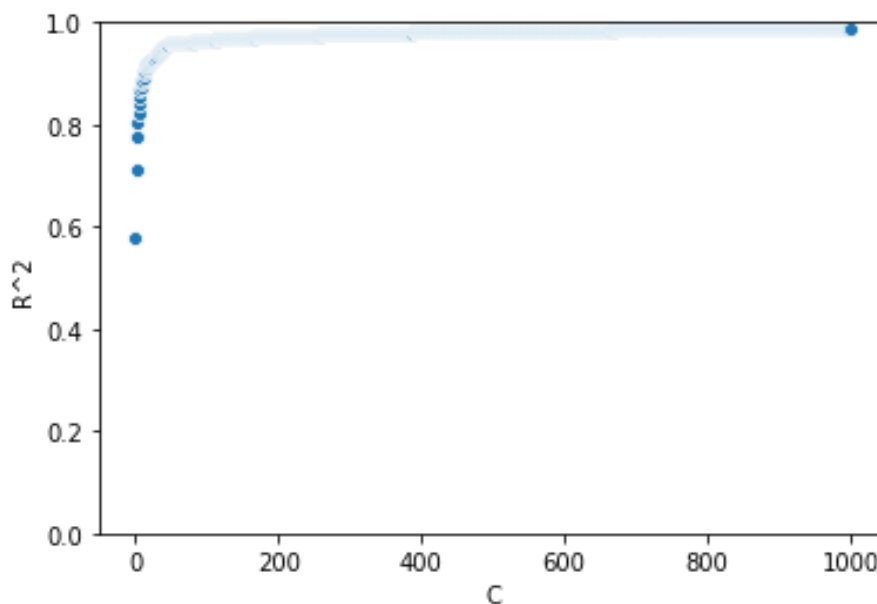


Figure 3. The scatterplot of  $R^2$  values for  $C = 1$  to  $C = 1000$ .

Based on Figure 3, the greater the value of  $C$ , the greater the value of  $R^2$  and closer to 1. When the value of  $C = 1$ , the value of  $R^2$  is 0.5801677881641607, when the value of  $C = 2$ , the value of  $R^2$  is 0.7115087665074087 and when the value of  $C = 1000$  then the value of  $R^2$  is 0.9860528612283006.

Furthermore, we evaluate by calculating the RMSE, MAE and Accuracy values. We compare the SVR method (proposed method) with KNN Regression (previous method) as in Table 4.

Table 4. Evaluation result of model SVR

Indicator	Model KNN Regression (Previous Method [2])	Model SVR (Proposed Method)
RMSE	2.9770	1.2817
MAE	2.1961	0.8008
Accuracy	88.1819 %	95.3369 %

In Table 4, the RMSE value of the SVR method is 1.2817 smaller than the RMSE value of the KNN regression method, indicating that the SVR method is better than the KNN method. For the MAE value, the model with the SVR method is 0.8008 also smaller than the KNN regression method, it also indicates that the SVR method is better than the KNN method. For accuracy, the model with the SVR method has a higher Accuracy rate of 95.3369% which means the SVR method is better. Based on these three measurements, it can be concluded that the SVR method is better than the KNN method in travel time estimation.

## CONCLUSION

The tourism sector is a very important sector to be developed because it can produce a promising business. One of the businesses in the tourism sector is to become a travel agent. To become a travel agent requires good planning in making a travel schedule. In making travel plans, travel agents usually estimate travel time based on previous travel experience without considering other factors. Therefore, in this study, we use 8 factors, i.e., time, weather, route, wind speed, day, precipitation, temperature and humidity information. This study collected data on travel from Ngurah Rai Airport to Kuta Beach, Bali. The data obtained were 228 which were divided into 177 training data and 51 test data. The Support Vector Regression (SVR) method is used to estimate travel time. To find out the SVR method can represent the data, the value of  $R^2$  is calculated. Based on the results of the study, it was found that the value of  $R^2 = 0.9860528612283006$  which interpreted that about 98.60528612283006 % of the SVR method could represent the data. After that, a test is carried out using test data and validation by measuring the error using RMSE, MAE and Accuracy.



Based on this research, we found that the RMSE value is 1.2817, the MAE value is 0.8008 and the Accuracy value is 95.3369%, indicating that the SVR method is a very good travel time prediction model.

## REFERENCES

- [1] I. C. Athira, C. P. Muneera, K. Krishnamurthy, and M. V. L. R. Anjaneyulu, "Estimation of Value of Travel Time for Work Trips," *Transp. Res. Procedia*, vol. 17, 2014, pp. 116–123, 2016.
- [2] Murni, R. Kosasih, A. Fahrurrozi, T. Handhika, I. Sari, and D. P. Lestari, "Travel Time Estimation for Destination in Bali Using KNN-Regression Method with Tensorflow," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 854, pp. 1–7, 2020.
- [3] J. Xin and S. Chen, "Bus Dwell Time Prediction Based on KNN," *Procedia Eng.*, vol. 137, pp. 283–288, 2016.
- [4] R. Li, H. Chai, and J. Tang, "Empirical Study of Travel Time Estimation and Reliability," *Math. Probl. Eng.*, vol. 2013, pp. 1–9, 2013.
- [5] J. Cheng, G. Li, and X. Chen, "Developing a Travel Time Estimation Method of Freeway based on Floating Car Using Random Forests," *J. Adv. Transp.*, vol. 2019, pp. 1–13, 2019.
- [6] B. A. Kumar, R. Jairam, S. S. Arkatkar, and L. Vanajakshi, "Real Time Bus Travel Time Prediction Using K-NN Classifier," *Transp. Lett.*, vol. 11, no. 7, pp. 362–372, 2019.
- [7] J. Myung, D. Kim, S. Kho, and C. Park, "Travel Time Prediction Using K Nearest Neighbor Method with Combined Data from Vehicle Detector System and Automatic Toll Collection System," *Transp. Res. Rec.*, vol. 2256, no. 1, pp. 51–59, 2012.
- [8] J. Tang, Y. Zou, J. Ash, S. Zhang, F. Liu, and Y. Wang, "Travel Time Estimation Using Freeway Point Detector Data Based on Evolving Fuzzy Neural Inference System," *PLoS One*, vol. 11, no. 2, pp. 1–15, 2016.
- [9] Z. Wang, K. Fu, J. Ye, D. A. I. Labs, and D. Chuxing, "Learning to Estimate the Travel Time," in *In KDD '18: 24th ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, 2018, pp. 858–866.
- [10] Z. Rustam and P. Kintandani, "Application of Support Vector Regression in Indonesian Stock Price Prediction with Feature Selection Using Particle Swarm Optimisation," *Model. Simul. Eng.*, vol. 2019, 2019.
- [11] Z. Rustam and J. Diani, "Support Vector Regression Implementation for Indonesian Private External Debt Analysis," *Eur. J. Electr. Eng. Comput. Sci.*, vol. 3, no. 3, pp. 1–4, 2019.
- [12] D. Setiyowati, Alamsyah, and M. A. Muslim, "Water Consumption Prediction of Semarang Water Utilities Using Support Vector Regression Radial Basic Function Kernel Method," *J. Adv. Inf. Syst. Technol.*, vol. 1, no. 1, pp. 55–63, 2019.
- [13] N. P. Ririanti and A. Purwinarko, "Implementation of Support Vector Machine Algorithm with Correlation-Based Feature Selection and Term Frequency Inverse Document Frequency for Sentiment Analysis Review Hotel," *Sci. J. Informatics*, vol. 8, no. 2, pp. 297–303, 2021.
- [14] S. Fransiska, Rianto, and A. I. Gufroni, "Sentiment Analysis Provider by.U on Google Play Store Reviews with TF-IDF and Support Vector Machine (SVM) Method," *Sci. J. Informatics*, vol. 7, no. 2, pp. 203–212, 2020.
- [15] Muhathir, E. P. Wanti, A. Pariyandani, S. Z. S. Idrus, and A. H. Lubis, "Utilization of SVM Method and GLCM Feature Extraction in Classifying Fish Images with Formalin," *Sci. J. Informatics*, vol. 8, no. 1, pp. 168–175, 2021.
- [16] A. Prahutama and H. Yasin, "Prediction of Weekly Rainfall in Semarang City Use Support Vector Regression (SVR) with Quadratic Loss Function," *Int. J. Sci. Eng.*, vol. 9, no. 1, pp. 13–16, 2015.
- [17] T. M. N. Utami, D. C. R. Novitasari, F. Setiawan, N. Ulinnuha, Y. Farida, and G. I. P. Sari, "Tide Prediction in Prigi Beach Using Support Vector Regression ( SVR ) Method," *Sci. J. Informatics*, vol. 8, no. 2, pp. 194–201, 2021.
- [18] E. Fradinata, S. Suthummanon, and W. Suntiarnortut, "Initial Optimal Parameters of Artificial Neural Network and Support Vector Regression," *Int. J. Electr. Comput. Eng.*, vol. 8, no. 5, pp. 3341–3348, 2018.
- [19] D. Novitasari, I. Cholissodin, and W. F. Mahmudy, "Hybridizing PSO with SA for Optimizing SVR Applied to Software Effort Estimation," *TELKOMNIKA*, vol. 14, no. 1, pp. 245–253, 2016.
- [20] Z. Rustam and R. Hidayat, "Indonesia Composite Index Prediction Using Fuzzy Support Vector Regression with Fisher Score Feature Selection," *Int. J. Adv. Sci. Eng. Inf. Technol.*, vol. 9, no. 1, pp. 121–128, 2019.
- [21] B. M. Henrique, V. A. Sobreiro, and H. Kimura, "Stock Price Prediction Using Support Vector Regression on Daily and Up to the Minute Prices," *J. Financ. Data Sci.*, vol. 4, no. 3, pp. 183–201,

- 2018.
- [22] Q. Guo, Y. Feng, X. Sun, and L. Zhang, "Power Demand Forecasting and Application based on SVR," *Procedia Comput. Sci.*, vol. 122, pp. 269–275, 2017.
  - [23] Y. Wang, Y. Yang, J. Jiao, Z. Wu, and M. Yang, "Support Vector Regression Approach to Predict the Design Space for the Extraction Process of Pueraria Lobata," *Molecules*, vol. 23, no. 10, 2018.
  - [24] A. Zakrani, M. Hain, and A. Idri, "Improving Software Development Effort Estimation Using Support Vector Regression and Feature Selection," *Int. J. Artif. Intell.*, vol. 8, no. 4, pp. 399–410, 2019.
  - [25] U. K. Das *et al.*, "SVR-Based Model to Forecast PV Power Generation Under Different Weather Conditions," *Energies*, vol. 10, no. 7, pp. 1–17, 2017.
  - [26] E. M. Priliani, A. T. Putra, and M. A. Muslim, "Forecasting Inflation Rate Using Support Vector Regression ( SVR ) Based Weight Attribute Particle Swarm Optimization ( WAPSO )," *Sci. J. Informatics*, vol. 5, no. 2, pp. 118–127, 2018.
  - [27] H. Yasin, R. E. Caraka, Tarno, and A. Hoyyi, "Prediction of Crude Oil Prices Using Support Vector Regression (SVR) with Grid Search - Cross Validation Algorithm," *Glob. J. Pure Appl. Math.*, vol. 12, no. 4, pp. 3009–3020, 2016.
  - [28] J. Tong, J. Du, P. Chen, J. Yuan, and Z. Huan, "The Price Forecasting of Military Aircraft based on SVR," *J. Comput. Commun.*, vol. 03, no. 05, pp. 234–237, 2015.
  - [29] E. Kasuya, "On the Use of R and R Squared in Correlation and Regression," *Ecol. Res.*, vol. 34, no. 1, pp. 235–236, 2019.
  - [30] S. B. Imandoust and M. Bolandraftar, "Application of K-Nearest Neighbor ( KNN ) Approach for Predicting Economic Events : Theoretical Background," *Int. J. Eng. Res. Appl.*, vol. 3, no. 5, pp. 605–610, 2013.