

SHORT-TERM PREDICTION OF TRAFFIC FLOW WITH SUPPORT VECTOR MACHINE TO REDUCE CONGESTION ON JAKARTA INNER-CITY TOLL ROADS

Ginanjari Bakti Rakhmanto
Universitas Indonesia, Indonesia
Email : bektiginanjar@gmail.com

DATE	ABSTRACT
<i>Accepted:</i> <i>Revised:</i> <i>Published:</i>	Traffic congestion on Jakarta's inner-city toll roads requires serious attention. The proposal is to offer an accurate short-term prediction solution to support the traffic engineering management system. This research uses the Support Vector Machine (SVM) model to predict traffic based on several variables, namely average speed, vehicle volume, weather, and time, by comparing it with the existing ARIMA model. The data is then examined for accuracy using Mean Absolute Error (MAE), Mean Percentage Error (MPE), and Root Mean Squared Error (RMSE). The results showed that SVM model 1, which only uses the input variable of vehicle volume, has a better prediction accuracy with MAE = 79.13, MPE = 5.51%, and RMSE = 99.90 compared to SVM model 2 (which uses average speed, vehicle volume, weather, and time variables) and the existing ARIMA model. The short-term prediction results can also be implemented in a traffic management system to reduce congestion on the Jakarta inner-city toll road. Solving congestion problems can have a positive impact on PT Jasa Marga, including increased satisfaction of toll road users and increased toll revenues through increased vehicle volumes that continue to flow without congestion.
	KEYWORDS Congestion, Short term traffic prediction, traffic management Support Vector Machine, ARIMA
	 <i>This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International</i>

INTRODUCTION

Traffic congestion is one of the crucial problems in big cities, including Jakarta, which continues to experience rapid growth in the number of vehicles and urbanisation. In this context, Jakarta's inner-city toll roads play an important role as an alternative for motorists who want to avoid congestion on arterial roads. However, these toll roads are not free from similar problems, especially during peak hours or when traffic accidents

occur. According to Anggoro & Anas (2024), the main cause of traffic congestion on Jakarta inner-city toll road is the imbalance between the increase in the number of vehicles compared to the capacity of the existing road network infrastructure. In this section, the increase in the number of vehicles is increasing rapidly but this is not in line with the increase in the number and length of toll roads, so that there is often a density of vehicles on a road, especially on toll roads during peak hour periods.

Table 1. Toll and non-toll road length growth table 2022-2025

Type	Year				% growth 2022- 2025
	2022	2023	2024	2025 (until April)	
Arteri Primer	57,70	58,82	59,36	59,36	2,80%
Kolektor Primer	2,16	2,16	2,16	2,16	0%
Arteri Sekunder	692,40	701,88	702,57	702,57	1,45%
Kolektor Sekunder	790,68	792,45	792,45	792,45	0,22%
Lokal	4.949,39	4.949,39	4.949,39	4.949,39	0%
Jakarta inner-city toll road	23,55	23,55	23,55	23,55	0%
TOTAL	6.515,88	6.528,25	6.529,48	6.529,48	0,19%

Table 2. Table of growth in the number of vehicles in 2022-2025

Type	Year				% growth 2022-2025
	2022	2023	2024	2025 (untill April)	
Passenger	4.210.390	4.354.843	4.490.548	4.520.927	6,24%
Buses	344.357	344.736	345.151	345.781	0,23%
Goods	849.132	865.461	886.338	891.778	4,20%
TOTAL	5.403.879	5.565.040	5.722.037	5.758.486	6,16%

From 2022 to 2025 from tables 1 and 2, the growth in the number of vehicles is 6.16% and the growth in road length is only 0.19%. The total length of roads from 2024 to April 2025 has not increased and has the potential not to increase considering that road construction requires an increasingly limited amount of land. Especially for toll roads within Jakarta, from 2022 to 2025 there is no increase in road length while the number of passenger vehicles, buses, and goods is increasing.

Congestion on Jakarta's inner-city toll roads has several impacts. Some of the reasons why congestion must be minimised are because it has a negative impact on road users and the Toll Road Business Entity (BUJT) or the toll road manager itself. According to Sari et al., (2018), congestion on toll roads has a negative impact on road users, including causing stress, causing radiator damage, increasing travel time, and causing high fuel consumption. As for toll road managers, according to Anggoro and Anas (2024), the level of customer satisfaction is reduced and the reduced toll revenue from BUJT when compared to smooth traffic flow conditions.

The congestion that occurs must be addressed and anticipated or minimised with various approaches. Intelligent Traffic Management Systems (ITMS) can now be applied in traffic engineering management by integrating the latest technology into the transport

infrastructure. According to Sakr and El-Afifi (2023) today's intelligent traffic management systems utilise new technologies and integrate various connectivity technologies, hardware, and software to manage the complexity of modern urban traffic in an effort to offer solutions to transport and logistics problems. According to Ait Ouallane et al., (2022), current ITMS already use communication and information technology, artificial intelligence, Internet of Things, Big Data, blockchain, and cloud computing.

Various efforts have been made and implemented in intelligent traffic management. One of the approaches to be applied in intelligent traffic management system is Support Vector Machine (SVM), a machine learning technique that is able to produce robust prediction models under varying data conditions. Currently, the prediction used in traffic prediction at PT Jasa Marga as the toll road manager in the inner city of Jakarta is using the ARIMA model which needs to be tested.

Studies conducted by Toan and Truong (2021) show that SVM has a good ability to capture complex and non-linear traffic patterns especially in dynamic urban scenarios and is able to predict short-term traffic.

Research related to traffic prediction using SVMs has shown significant potential, especially in the context of urban traffic. Some studies, such as the one by Kalaivanan and Brindha (2024) highlighted that SVMs can detect and manage traffic congestion in intelligent transport systems. The application of SVM and other machine learning methods in traffic flow prediction, provides high accuracy results in intelligent transport systems and is more efficient than traditional methods such as linear regression (Li and Xu, 2021). In Jakarta, the implementation of this model can help reduce the impact of congestion on toll roads, especially when vehicle volumes increase drastically. Studies on the use of SVMs have also shown superiority in handling variations in traffic data on toll roads (Wang et al., 2015). In the context of this research, traffic prediction with SVM not only aims to improve the accuracy of vehicle flow prediction, but also as a proactive step in more efficient traffic management.

External factors such as weather conditions play an important role in influencing traffic patterns. Studies conducted by F. Tseng and Hsueh (2018) confirmed that traffic prediction models that consider weather variables perform better. This is also in line with the study of Li and Xu (2021) which states that the developed model can work well in various weather conditions.

In making a traffic prediction, prediction models with multiple variables can make up for the shortcomings of using single variable data and improve prediction accuracy (Long et al., 2019). This suggests that a comprehensive traffic prediction model should incorporate various variables that affect traffic flow. Thus, the incorporation of these factors in an SVM-based prediction model is expected to provide a better solution for traffic management on Jakarta's inner-city toll roads.

Jakarta, as one of the most congested cities in the world, requires an innovative approach to traffic management. Solving congestion problems can certainly have a positive impact on PT Jasa Marga, including increased satisfaction of toll road users and increased toll revenues through increased vehicle volumes that continue to flow without congestion.

This research aims to develop a short-term traffic flow prediction model using SVM method, by considering various variables such as average speed, vehicle volume, weather, and time. The results of this research are expected to be useful and more accurate prediction results can provide data support for monitoring, early warning, and decision

analysis for freeway operation status (Long et al., 2019). With more accurate predictions, it is expected to assist toll road managers in optimising traffic flow and reducing chronic congestion.

METHOD

This research is a study with an SVM-based short-term prediction model to predict the volume of vehicles per hour on the toll road in the city of Jakarta, precisely at KM 4 + 500 A. Short-term prediction serves to provide the latest information about traffic conditions for one or more future time intervals based on pre-existing data sets in real-time (Williams et al., 2014). In the first stage, observations were made of the initial conditions at the location of KM 4+500 A of the Jakarta inner-city toll road. In the second stage, all data were collected together. Data was collected from Jasa Marga for traffic data and BMKG for weather data. The output variable in this study is vehicle volume, while the input variables include vehicle volume, average speed, weather, and time. In this research, 2 (two) scenarios were carried out, namely SVM scenario 1 using the input variable of vehicle volume. While SVM scenario 2 uses input variables of average speed, vehicle volume, weather, and time. In the third stage, the data obtained is then analysed using RStudio software for SVM 1, SVM 2 models, then the existing ARIMA model is used as a comparison. Model evaluation is carried out using error metrics such as MAE, MPE, and RMSE to assess the prediction accuracy of each model.

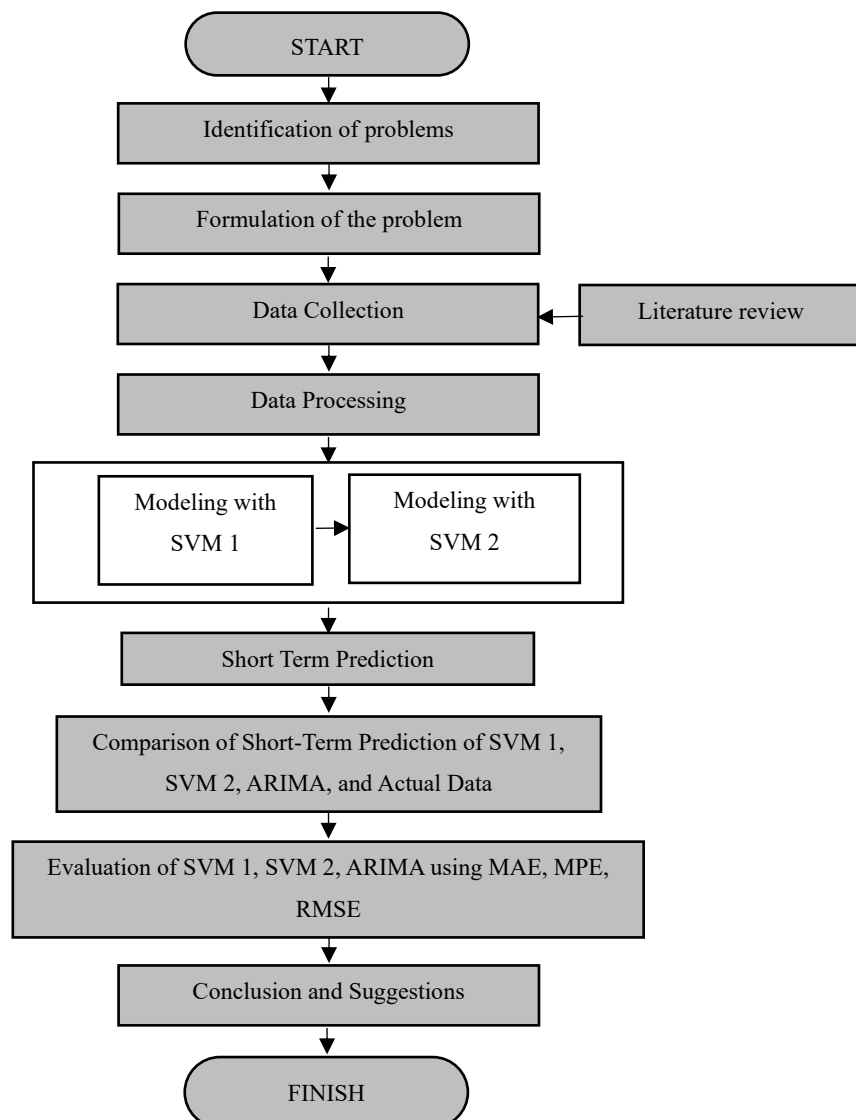


Figure 2 Framework

Data was collected from Jasa Marga for traffic data and BMKG for weather data, then preprocessed, including cleaning, and merging to ensure the data is ready for analysis. The type of data used in this research is secondary data collected from official institutions:

1. Jasa Marga provides data on average speed, V/C Ratio, and vehicle volume on toll roads from traffic counting radar. This tool uses microwave-based technology. The location of the traffic counting radar that will be used in the study is at KM 4+500 A.
2. Meteorology, Climatology and Geophysics Agency (BMKG): provides daily weather data to identify weather conditions at any given time.

The variables used in this study are as follows:

1. Output Variables

Vehicle volume is the output variable in this study because the main objective is to predict the number of vehicles passing a point on the toll road per hour. According to Tang et al., (2019); Toan and Truong, (2021); Xu et al., (2017) used vehicle volume as the predicted variable in the study. Vehicle volume prediction is important because it can provide an early indication of potential congestion. By predicting vehicle volume, this research aims to provide relevant information for traffic managers to take actions that can reduce or manage traffic congestion.

2. Input Variables

These variables are used as inputs in the prediction model to influence vehicle volumes. The following are the input variables and their explanations:

- a. Vehicle Volume of the previous period. In this study, 24-hour daily vehicle volume data for 1-28 February 2025 is used. According to Toan & Truong (2021) the previous vehicle volume data is used to predict vehicle volumes in the future according to the rolling horizon concept.
- b. Average vehicle speed is the average of the speeds of all vehicles travelling on a given point or road section at a given time. This variable can affect vehicle volumes because when the average speed is high, the volume of vehicles travelling through a particular road section per hour is usually higher, whereas low speeds tend to be associated with higher density and potential congestion. According to Kalaivanan and Brindha (2024) vehicle speed data is used to determine the prediction of vehicles experiencing congestion or not.
- c. Daily Weather: Weather conditions such as sunny, rainy, fog, or bad weather have a significant influence on traffic. A study from Tseng & Hsueh (2018) and Zafar & Haq (2020) explains that weather is a variable that affects traffic congestion conditions. Meanwhile Li and Xu (2021) divided weather conditions into sunny, cloudy, foggy, and rainy. Bad weather, such as rain or fog, tends to lower vehicle speeds and reduce traffic volumes as vehicles move more slowly or even avoid travelling. By including weather conditions as inputs, the model can consider the effect of weather conditions on vehicle volumes.
- d. Time: is a dummy variable indicating whether it is congested or not, the morning congested time is 06.00 - 09.00, while the afternoon congested time is 16.00 -

19.00. Toan & Truong (2021b) used time interval prediction in making SVM models. Zafar & Haq (2020) took into account the time period of congested and uncongested hours.

Support Vector Regression (SVR)

SVR is used to predict vehicle volume in short-time. SVR is a regression method adapted from Support Vector Machine (SVM) and is part of machine learning that utilises margin optimisation principles in data classification (V. N. Vapnik, 2000). SVR was introduced to handle continuous value prediction problems, different from SVM which was originally focused on classifying data. The main goal of SVR is to find the optimal hyperplane that can accurately estimate the value of the dependent variable, with a permissible margin of error. In the context of SVR, this margin of error is known as the epsilon-tube, which defines the maximum distance between the predicted and actual values that is still considered a valid estimate (Smola & Schölkopf, 2004). SVR was chosen for its ability to solve non-linearly divisible problems, which is suitable for handling traffic flow. This model uses RBF (Radial Basis Function) kernel which is proven to be better than linear kernel for traffic prediction (Zheng et al., 2020). According to Li et al., (2019), the SVR algorithm also has the ability to solve the problems of nonlinearity, small sample, and high dimension, which conforms to the characteristics of short-term traffic flow prediction research. Therefore, the SVR algorithm is adopted to build a short-term traffic flow prediction model.

Mathematically, SVR works by minimising the following function:

$$\frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \quad (3.1)$$

with constraints:

$$\begin{aligned} y_i - (w, x_i) - b &\leq \epsilon + \xi_i \\ (w, x_i) + b - y_i &\leq \epsilon + \xi_i^* \\ \xi_i, \xi_i^* &\geq 0 \end{aligned}$$

Description:

w is the weight vector,

b is the bias,

ϵ is the margin of error tolerated in the ϵ -tube,

ξ_i and ξ_i^* is a slack variable that allows some data to fall outside the margin ϵ ,

C is a regulatory parameter that controls the trade-off between model complexity and error tolerance.

The goal of SVR is to minimise the value $\|w\|^2$ to keep the hyperplane simple while adjusting the margins so that the model does not unduly influence data outside the margins (Cristianini & Shawe-Taylor, 2000). SVR works by setting margins that minimise the influence of small errors around the model, allowing the model to be tolerant

of small amounts of noise in the data. This approach reduces sensitivity to outliers, making SVR superior to traditional linear regression methods under conditions of noisy data or outliers (Drucker et al., 1997).

In addition, SVR also supports kernel trick, which is a technique to transform the data into a higher dimensional space, where the data becomes easier to separate linearly. With kernel trick, SVR can handle non-linear relationships in the data. The kernel function $K(x_i, x_j)$ replaces the dot product (x_i, x_j) , allowing operation in a higher feature space without calculating the coordinates of the data in that space. Commonly used kernel functions include:

- Linear Kernel: $K(x_i, x_j) = (x_i, x_j)$ (3.2)

- Polynomial Kernel: $K(x_i, x_j) = (\langle x_i, x_j \rangle + 1)^d$ (3.3)

- *Radial Basis Function (RBF) Kernel*: $K(x_i, x_j) = \exp(-\gamma \|x_i, x_j\|^2)$ (3.4)

- Sigmoid Kernel: $K(x_i, x_j) = \tanh(\kappa \langle x_i, x_j \rangle + \theta)$ (3.5)

With the kernel trick, SVR can model more complex relationships, making it superior in handling non-linear data and producing more accurate models under varying data conditions (Schölkopf & Smola, 2002).

In SVR, the error is calculated only if the deviation between the prediction and the actual value exceeds a margin of ϵ . This ϵ -insensitive loss function is defined as:

$$L_\epsilon(y, f(x)) = \max(0, |y - f(x)| - \epsilon)$$
(3.6)

Description:

y is the actual value,

$f(x)$ is the predicted value,

ϵ is the tolerable error margin.

This loss function allows the model to ignore small errors that fall within the ϵ -tube margin, which increases the model's robustness to outliers (V. Vapnik, 1998). It also allows SVR to have the flexibility to deal with large data or high-dimensional data without losing computational efficiency. As a robust regression method, SVR has been applied in various fields, including economics, finance, and engineering, where it is able to provide competitive performance even under complex data conditions (Awad & Khanna, 2015). To solve optimisation problems, SVR uses Lagrange duality to maximise the margin while minimising the error. By converting the primal equation into the dual form, the model can be written as:

$$f(x) = \sum_{i=1}^n (\alpha_i - \alpha_i^*) K(x_i, x) + b \quad (3.7)$$

α_i and α_i^* are Lagrange coefficients,

$K(x_i, x)$ is the kernel function.

This equation models the SVR prediction as a linear combination of kernels computed between the training data and the data to be predicted, where α_i and α_i^* are only non-zero for support vectors on the edge of the ϵ -tube.

Comparative Methods

As a comparative method, this research uses Autoregressive Integrated Moving Average (ARIMA) time series analysis which has been used previously by PT Jasa Marga in their application.

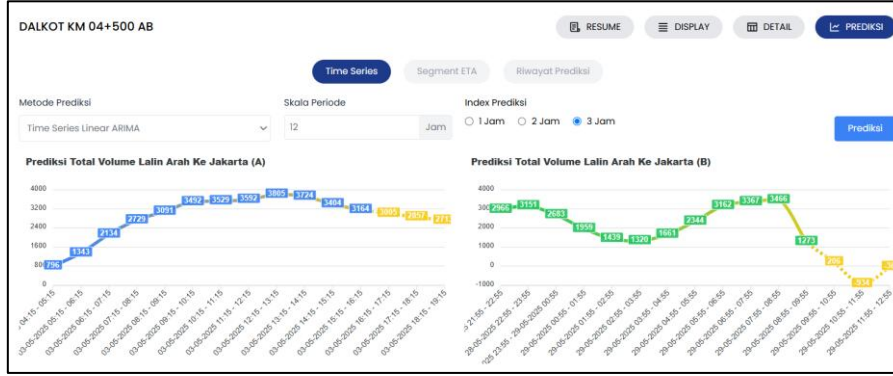


Figure 3 PT Jasa Marga's traffic prediction application using ARIMA

ARIMA models are a subset of statistical models that are often used to predict time series data. The ARIMA model acts as a comparison in assessing the performance of SVM in predicting traffic flow. ARIMA combines Auto Regressive (AR) and Moving Average (MA) components, which is suitable for data that is stationary or has no seasonal pattern. The ARIMA (p,d,q) formula is:

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \dots + \theta_q \epsilon_{t-q} \quad (3.8)$$

Description:

- Y_t is the time series value at time t after differencing d times
- ϕ_i is the i-th order autoregressive (AR) parameter
- θ_j is the j-th order moving average (MA) parameter
- ϵ_t is the white noise error at time t
- p is the AR order, and q is the MA order

To measure the prediction accuracy of the SVM model and the comparison method, several error metrics were used, namely:

1. Mean Absolute Error (MAE) is to measure the absolute average of the difference between the prediction and the actual value. The MAE formula is:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (3.9)$$

Description:

- y_i = Actual Value
- $\hat{y}_i$ = Predicted Value
- n = the number of data samples

2. Mean Percentage Error (MPE), which measures the average percentage error between actual and predicted values in a model. MPE is useful for understanding the direction of bias in predictions: if the APE value is positive, the model tends to underestimate; if the MPE value is negative, the model tends to overestimate. The Mean Percentage Error (MPE) formula is:

$$MPE = \frac{1}{n} \sum_{i=1}^n \left(\frac{y_i - \hat{y}_i}{y_i} \right) \times 100\% \quad (3.10)$$

Description:

- y_i = Actual Value
- $\hat{y}_i$ = Predicted Value
- n = the number of data samples

3. Root Mean Squared Error (RMSE) is the square root of MSE, which gives higher weight to large errors. The RMSE formula is:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (3.11)$$

Description:

- y_i = Actual Value
- $\hat{y}_i$ = Predicted Value
- n = the number of data samples

RESULTS AND DISCUSSION

Overview of Research Objects

The data source of this research is the toll road in the city of Jakarta, precisely at KM 04 + 500 A by using radar traffic counting. This tool uses microwave-based technology to count the number of vehicles passing by. Currently there are 26 radar traffic counting locations installed on the Inner City Toll Road. Initial data that can be obtained are vehicle volume and average speed. Data calculations will be processed using Rstudio software with package SVM 1071.



Figure 4 Radar Traffic Counting

The Jakarta inner-city toll road at KM 04+500 A is part of the Jakarta inner-city toll road network which is one of the main routes connecting various strategic areas in the capital city. This section has an important role in flowing the flow of vehicles from Cawang, Tomang to Pluit.

Data obtained from radar traffic counting KM 04+500 A will be processed using SVM to predict the value of traffic volume. The location that will be studied is the Tegal Parang off ramp KM 5+000 A. As one of the points with high density, this location often experiences congestion due to congestion on the arterial road (after exiting the Tegal Parang off ramp KM 5+000 A) whose queue has an impact on the toll road.



Figure 5 Tegal Parang Off Ramp Situation (KM 5+000 A)

Congestion at Tegal Parang Off Ramp KM 5+000 A to Kuningan Junction is caused by:

- The meeting or joining of vehicles from the direction of the toll road (Tegal Parang off ramp KM 5+000A) and from Jalan Gatot Subroto
- The movement of Transjakarta vehicles that had to switch lanes to get to the Tegal Parang bus stop, causing obstacles for toll road users who had just exited the Tegal Parang KM 5+000A off ramp.



Figure 6 Transjakarta's movement (switch lane) to Tegal Parang bus stop

- c. The movement of Transjakarta vehicles that must change lanes from the Tegal Parang bus stop to the Simpang Kuningan bus stop, causing obstacles for road users who will lead to the Kuningan Elevated Road.



Figure 7 Transjakarta's movement (switch lane) to Kuningan Junction bus stop

- d. Queue at Kuningan Junction due to a traffic light at Kuningan Junction



Figure 8 Kuningan Junction

These problems contribute to slowing down the flow of vehicles, especially during the morning and evening rush hours. Therefore, the implementation of more effective and technology-based transportation policies, such as intelligent transportation systems, is necessary to optimise the flow of vehicles on this toll road.

Descriptive Statistics

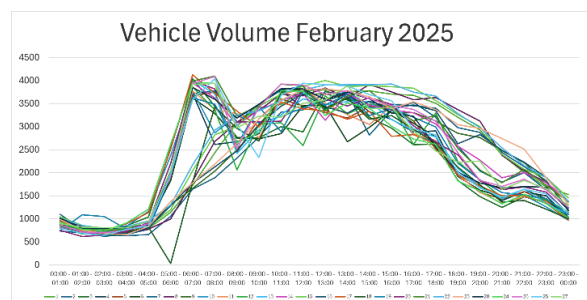


Figure 9 Graphic of Daily hourly vehicle volume February 2025

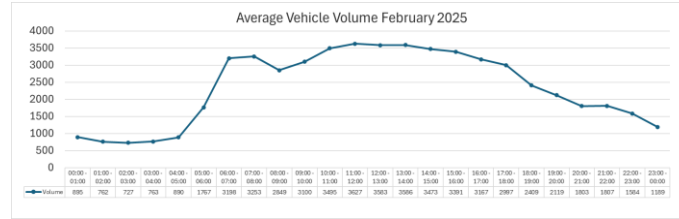


Figure 10 Graphic of Average Vehicle Volume February 2025

Based on Figures 9 and 10 above, it can be described that the lowest vehicle volume occurs in the early morning hours (00:00-05:00) with an average of 700-900 vehicles per hour. There is a significant spike in the morning from 05:00-06:00, increasing from around 900 to 3,100 vehicles per hour. Peak vehicle volumes occur in the afternoon between 11:00-14:00, reaching around 3,500-3,600 vehicles per hour. The volume starts to decline gradually after 14:00, with a faster decline after 17:00 (working hours). Meanwhile, at night (20:00-23:00) volumes are again low with an average of 1,500-2,000 vehicles per hour. This pattern shows the typical characteristics of urban toll roads with peak volumes during working hours and low volumes at night/early morning.

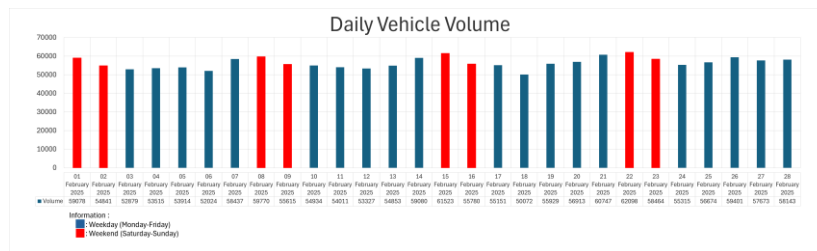


Figure 11 Graphic of daily vehicle volume February 2025

Based on the graphic of daily vehicle volumes in February 2025, noting that the red bars indicate Saturdays and Sundays, the following patterns can be seen: (i) Vehicle volumes on weekdays (Monday-Friday) tend to be stable in the range of 50,000-55,000 vehicles per day; (ii) On weekends (Saturday-Sunday), vehicle volumes are generally higher than weekdays, reaching around 55,000-60,000 vehicles per day; (iii) The peak volume occurred on 22 February 2025, which was 62,098 vehicles. This pattern shows the characteristics of toll road usage with a tendency for higher volumes on weekends (especially Saturdays).

Table 3. Variable Descriptive Statistics

Variabel	N (amount of data)	Minimum	Maximum	Mean	Std. Dev
Average speed	671	14	79	53,44	15,32
Vehicle volume	671	36	4.126	2.352,85	1.120,87
V/C Ratio	671	0,01	0,58	0,33	0,15

Source: Author's Processed Results

Modelling with SVM

The Support Vector Machine (SVM) model is used in this study to predict the volume of vehicles on the Jakarta inner-city toll road. The model was built using a radial

basis function (RBF) kernel and configured parameters, namely cost = 100, gamma = 0.1, and epsilon = 0.1. The model utilises input variables such as vehicle volume, average speed, time of day, and weather conditions to capture traffic movement patterns and provide short-term estimates of vehicle volume. The comparison graph of realised and predicted vehicle volumes in this model can be seen in Figures 12 and 13

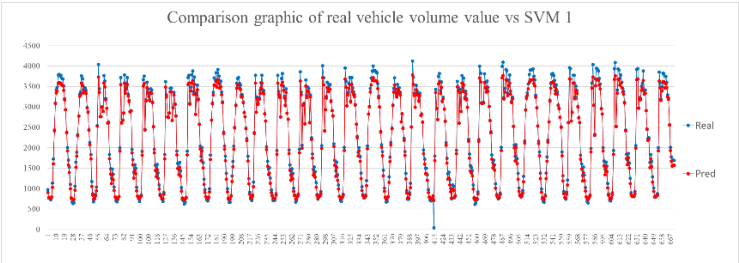


Figure 12 Comparison of realised vs predicted vehicle volume (SVM model 1)
Source: Author's Processed Results

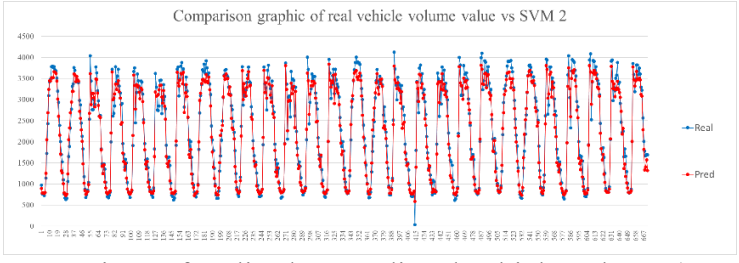


Figure 13 Comparison of realised vs predicted vehicle volume (SVM model 2)
Source: Author's Processed Results

Comparison with Existing Model (ARIMA)

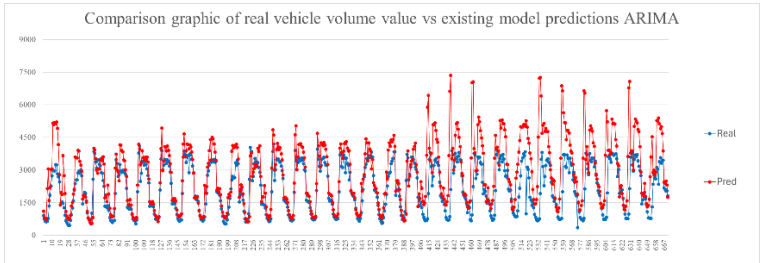


Figure 14 Comparison of realised vs existing model ARIMA
Source: Author's Processed Results

Figure 14 shows the comparison between the realised vehicle volume and the prediction results using the Autoregressive Integrated Moving Average (ARIMA) model. In this graph, it can be seen that the red line represents the ARIMA predicted value, while the blue line shows the realised vehicle volume. From the visualisation, it can be observed that while the ARIMA model is able to capture the general pattern of traffic movement, there are quite noticeable differences at some points, especially when there are sharp spikes or drops in vehicle volumes.

Tabel 0 Comparison of Model Error Measures

Model	MAE	MPE	RMSE
-------	-----	-----	------

SVM 1	79,13	5,51	99,90
SVM 2	141,04	8,50	184,23
ARIMA	804,68	58,40	1315,68

Source: Author's Processed Results

Table 4 presents a comparison of the performance of the prediction models, namely SVM 1, SVM 2 and ARIMA, based on three main prediction error measures: Mean Absolute Error (MAE), Mean Percentage Error (MPE), and Root Mean Squared Error (RMSE). These three matrices are used to evaluate the extent to which the model predictions are close to the actual values in predicting vehicle volumes.

Short Term Prediction

After making SVM 1 and SVM 2 models, short-term predictions from both methods are compared with existing predictions (ARIMA) and their realised values. In application, the mechanism of short-term prediction with the SVM method can be described by the following pattern.

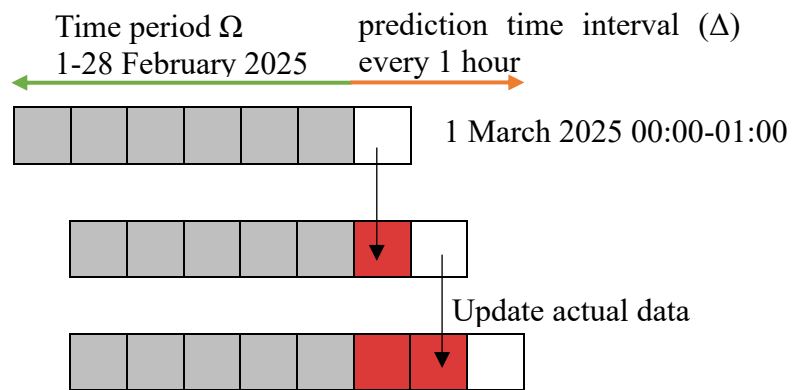


Figure 15 Rolling Horizon Approach for short term prediction

Source: Author's Processed Results

In short-term data prediction according to the rolling horizon approach where the value of the time period Ω is data from 1-28 February 2025 with a rolling step δ of 1 hour) and the prediction time interval (Δ) every 1 hour ahead. The initial data predicted is traffic data on 1 March 2025 at 00:00 - 01:00. Then to predict the traffic data on 1 March 2025 at 01:00 - 02:00 can be done by filling in the realisation data from the data on 1 March 2025 at 00:00 - 01:00.

In short-term prediction, each subsequent prediction step, the data window used for prediction shifts to replace the old data with the latest data, while the prediction data afterwards will be updated with new prediction results based on more recent information. To illustrate a wider range of data, short-term predictions were made at 06:00-09:00 and 17:00-21:00 during March 2025.

The results of traffic prediction using SVM models can be used to support more efficient traffic management on the Jakarta Inner City toll road, especially to overcome congestion at the Tegal Parang Off Ramp KM 5+000 A. With more accurate predictions, transportation authorities or policy makers can plan traffic engineering management

systems, such as open and close implementation, odd-even enforcement, and traffic light timings that are more adaptive to road density.

By comparing the conditions during heavy traffic compared to smooth traffic, the following losses are obtained:

1. Assumed impact of potential toll revenue loss during heavy traffic compared to smooth traffic based on Toll Gate origin
2. Impact of pollution in congested or heavy traffic conditions
3. Impact on customer satisfaction

By applying the right prediction, congestion is expected to be resolved quickly so that losses due to congestion can be minimised.

RESUME

The conclusions from the results of this study. (i) The Support Vector Machine (SVM) model is able to predict short-term traffic flow patterns on the Jakarta inner-city toll road with a rolling horizon pattern. In this approach, every next prediction step, the data window used for prediction shifts to replace the old data with the latest data, while the prediction data afterwards will be updated with new prediction results based on more recent information. (ii) SVM 1 has the best performance in terms of prediction accuracy compared to SVM 2 and the existing model used by PT Jasa Marga, namely ARIMA. SVM 1 model in this study can be considered to replace the existing prediction model currently used, namely ARIMA. (iii) SVM-based short-term traffic prediction results can be used to support better traffic management, especially in reducing congestion. More accurate prediction values can be used by toll road managers or external policy makers to optimise traffic engineering management policies at KM 4+500 A to Tegal Parang Off Ramp area. By applying the right prediction, congestion is expected to be resolved quickly so that losses due to congestion can be minimised.

REFERENCES

- Ait Ouallane, A., Bakali, A., Bahnasse, A., Broumi, S., & Talea, M. (2022). Fusion of engineering insights and emerging trends: Intelligent urban traffic management system. *Information Fusion*, 88(February), 218–248. <https://doi.org/10.1016/j.inffus.2022.07.020>
- Anggoro, Y. T., & Anas, E. P. (2024). Analisa Kemungkinan Penerapan Tarif Tol Dinamis: Studi Kasus Pada Jalan Tol dalam Kota Jakarta. *Syntax Literate ; Jurnal Ilmiah Indonesia*, 9(4), 2370–2388. <https://doi.org/10.36418/syntax-literate.v9i4.15481>
- Awad, M., & Khanna, R. (2015). *Efficient Learning Machines*. apress.
- Chen, L., Zheng, L., Yang, J., Xia, D., & Liu, W. (2020). Neurocomputing Short-term traffic flow prediction: From the perspective of traffic flow decomposition. *Neurocomputing*, 413, 444–456. <https://doi.org/10.1016/j.neucom.2020.07.009>
- Chi, Z., & Shi, L. (2018). Short-Term Traffic Flow Forecasting Using ARIMA-SVM Algorithm and R. *Proceedings - 2018 5th International Conference on Information Science and Control Engineering, ICISCE 2018*, 517–522. <https://doi.org/10.1109/ICISCE.2018.00114>
- Cristianini, N., & Shawe-Taylor, J. (2000). *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge University Press.
- Drucker, H., Surges, C. J. C., Kaufman, L., Smola, A., & Vapnik, V. (1997). Support

- vector regression machines. *Advances in Neural Information Processing Systems*, 1, 155–161.
- Kalaivanan, E., & Brindha, S. (2024). *Support Vector Machine Based Traffic Congestion in Intelligent Transportation System*. 28(1), 2026–2035.
- Li, C., & Xu, P. (2021). Application on traffic flow prediction of machine learning in intelligent transportation. *Neural Computing and Applications*, 33(2), 613–624. <https://doi.org/10.1007/s00521-020-05002-6>
- Li, C., Zhang, H., Zhang, H., & Liu, Y. (2019). Short-term traffic flow prediction algorithm by support vector regression based on artificial bee colony optimization. *ICIC Express Letters*, 13(6), 475–482. <https://doi.org/10.24507/icicel.13.06.475>
- Lint, H. van, & Hinsbergen, C. van. (2012). Short-Term Traffic and Travel Time Prediction Models. *Transportation Research Circular Number E-C168*, November, 22–41.
- Long, K., Yao, W., Gu, J., Wu, W., & Han, L. D. (2019). Predicting freeway travel time using multiple- source heterogeneous data integration. *Applied Sciences (Switzerland)*, 9(1). <https://doi.org/10.3390/app9010104>
- Mustikarani, W., & Suherdiyanto. (2016). Analisis Faktor-Faktor Penyebab Kemacetan Lalu Lintas Di Sepanjang Jalan H Rais a Rahman (Sui Jawi) Kota Pontianak. *Jurnal Edukasi*, 14, 143–155.
- Sakr, H. A., & El-Afifi, M. I. (2023). Intelligent Traffic Management Systems: A review. *Nile Journal of Communication & Computer Science Journal Webpage*, 5(June). <https://njccs.journals.ekb.eg>
- Samal, S. R., Gireesh Kumar, P., Cyril Santhosh, J., & Santhakumar, M. (2020). Analysis of Traffic Congestion Impacts of Urban Road Network under Indian Condition. *IOP Conference Series: Materials Science and Engineering*, 1006(1). <https://doi.org/10.1088/1757-899X/1006/1/012002>
- Sari, M., Saidah, D., & Wahyuni, E. (2018). Dampak Kemacetan di Jalan Tol Brebes Timur The Effect of Traffic Jam in Brebes East. *Jurnal Manajemen Transportasi & Logistik*, 05(01), 1–6.
- Schölkopf, B., & Smola, A. J. (2002). Support Vector Machines and Kernel Algorithms. *The Handbook of Brain Theory and Neural Networks*, April 2002, 1119–1125.
- Smola, A. J., & Schölkopf, B. (2004). A tutorial on support vector regression. *Statistics and Computing*, 14(1), 199–222. <https://doi.org/10.1186/s12984-021-00806-6>
- Tamara, S., & Sasana, H. (2017). Analisis Dampak Ekonomi Dan Sosial Akibat Kemacetan Lalu Lintas Di Jalan Raya Bogor-Jakarta. *Jurnal REP (Riset Ekonomi Pembangunan)*, 2(2), 185–196. <https://doi.org/10.31002/rep.v2i3.529>
- Tang, J., Chen, X., Hu, Z., Zong, F., Han, C., & Li, L. (2019). Traffic flow prediction based on combination of support vector machine and data denoising schemes. *Physica A: Statistical Mechanics and Its Applications*, 534, 120642. <https://doi.org/10.1016/j.physa.2019.03.007>
- Thakre, Y. K., & Pawade, P. Y. (2024). Traffic Congestion at Urban Road - Review. *IOP Conference Series: Earth and Environmental Science*, 1326(1). <https://doi.org/10.1088/1755-1315/1326/1/012094>
- Toan, T. D., & Truong, V. (2021a). *Support Vector Machine for Short-Term Traffic Flow Prediction and Improvement of Its Model Training using Nearest Neighbor Approach*. December 2020. <https://doi.org/10.1177/0361198120980432>
- Toan, T. D., & Truong, V. H. (2021b). Support vector machine for short-term traffic flow prediction and improvement of its model training using nearest neighbor approach.

- Transportation Research Record*, 2675(4), 362–373.
<https://doi.org/10.1177/0361198120980432>
- Tseng, F. H., Hsueh, J. H., Tseng, C. W., Yang, Y. T., Chao, H. C., & Chou, L. Der. (2018). Congestion prediction with big data for real-time highway traffic. *IEEE Access*, 6, 57311–57323. <https://doi.org/10.1109/ACCESS.2018.2873569>
- Tseng, F., & Hsueh, J. (2018). Congestion Prediction With Big Data for Real-Time Highway Traffic. *IEEE Access*, 6, 57311–57323. <https://doi.org/10.1109/ACCESS.2018.2873569>
- Vapnik, V. (1998). *The Support Vector Method of Function Estimation*. 55–85.
- Vapnik, V. N. (2000). *Statistics for Engineering and Information Science* (Vladimir N. Vapnik - The Nature of Statistical Learning Theory-Springer (2000).pdf.
- Wang, X., An, K., Tang, L., & Chen, X. (2015). Short Term Prediction of Freeway Exiting Volume Based on SVM and KNN. *International Journal of Transportation Science and Technology*, 4(3), 337–352. <https://doi.org/10.1260/2046-0430.4.3.337>
- Williams, B. M., Asce, M., Hoel, L. A., & Asce, F. (2014). *Modeling and Forecasting Vehicular Traffic Flow as a Seasonal ARIMA Process : Modeling and Forecasting Vehicular Traffic Flow as a Seasonal ARIMA Process : Theoretical Basis and Empirical Results*. 6(November 2003). [https://doi.org/10.1061/\(ASCE\)0733-947X\(2003\)129](https://doi.org/10.1061/(ASCE)0733-947X(2003)129)
- Xu, Y., Hu, D. W., & Su, B. (2017). Short-term traffic flow prediction based on optimised support vector regression. *International Journal of Applied Decision Sciences*, 10(4), 305–314. <https://doi.org/10.1504/IJADS.2017.087176>
- Zafar, N., & Haq, I. U. (2020). Traffic congestion prediction based on Estimated Time of Arrival. *PLoS ONE*, 15(12 December), 1–19. <https://doi.org/10.1371/journal.pone.0238200>
- Zheng, L., Yang, J., Chen, L., Sun, D., & Liu, W. (2020). Neurocomputing Dynamic spatial-temporal feature optimization with ERI big data for Short-term traffic flow prediction. *Neurocomputing*, 412, 339–350. <https://doi.org/10.1016/j.neucom.2020.05.038>