Inserach
Information System Research Journal

E-ISSN: 2775-4669
Volume 1, Nomor 2, Agustus 2021

# Stock Price Prediction using Support Vector Machine in the Second Wave of Covid-19 Pandemic

**Ahmad Fauzi**
Faculty of Science and Technology, UIN Imam Bonjol Padang
Kampus III Universitas Islam Negeri Imam Bonjol Padang,
Sungai Bangek Koto Tangah, Padang, West Sumatera, Indonesia
ahmadfauzi@uinib.ac.id

Diterima: 30 Aug 2021 | Direvisi: 04 Sep 2021
Disetujui: 06 Sep 2021 | Dipublikasi: 20 Sep 2021

**Abstract**

*The aim of the research is to predict The Jakarta Composite Index (JKSE) when the second wave of the Covid-19 occurred in Indonesia. The method used in this research is Support Vector Machine (SVM). The data used in this study is JKSE for period April 2021 to August 2021. The data is divided into two parts, the training data and the testing data. The training data starts from April 1st to July 30th 2021. The data used as ground truth is the data from August 2nd to August 31st 2021. To get the best hyperparameter, the research uses Grid Search with RBF kernel. The results show the best hyperparameters using RBF kernel are C=100, ε =0.01, and γ = 3. The prediction result obtained tends to be stable and has a slight decrease. This result if we compared with the cumulative cases and the total recovery cases of the Covid-19, the investor has a positive sentiment towards the Indonesian government in solving the Covid-19 pandemic.*

***Keywords: Stock Price, Support Vector Machine, Covid-19***

## I. INTRODUCTION

In March 2020, the Indonesian government announced that the Sars-Cov-2 virus that causes Covid-19 has entered Indonesia. As a result of the Covid-19, various sectors were collapse, including the economic sector. One of the indicators of the decline in the economic sector is the stock prices is collapse at the beginning of Covid-19 pandemic.

The stock price is basically influenced by sentiment or external factors. The decline of the stock price that occurred during the Covid-19 pandemic indicates the level of investor confidence is decline too. The policies taken by the government in dealing with the Covid-19 will cause positive or negative sentiment among the investors.

The Jakarta Composite Index (JKSE) as one of the largest indexes in Indonesia shows the index has a significant decline on March 24th where the close index is 3937.63 (Figure 1). However, during the second wave of the Covid-19 in Indonesia, the JKSE tends to be in a safe position despite in June 2021, the stock price is decreased. The decline of JKSE that occurred in the second wave of the Covid-19 was not as extreme as the decline that occurred from February to April 2020.

The Covid-19 pandemic basically won't go away in a short time. The research conducted by (Qureshi, Baskett, Huang, Lobanova, Naqvi, & Shyu, 2021) (To, et al., 2021) (Wang, Kaperak, Sato, & Sakuraba, 2021) shows that someone who has already been exposed by the Covid-19 can be suspected again. The research conducted by (Bjørnstad, Shea, Krzywinski, & Altman, 2020) using the SEIRS model and a report issued by the World Health Organization (WHO) about the variants of the Covid-19 adds to the possibility that the Covid-19 pandemic will last a long time.

One of the strategies taken to anticipate the risks caused by the Covid-19 to the economic sector is by analyzing the stock prices. This study aims to conduct modeling and analysis of JKSE

when the second wave of the Covid-19 occurred in Indonesia. The research uses the Support Vector Machine (SVM) to predict the stock prices.

## II. LITERATURE

Support Vector Machine (SVM) is basically used to find the maximum margin of hyperplane between classes. The general equation of SVM is shown by Equation (1)

$$f(\boldsymbol{x}) = \boldsymbol{w}\varphi(\boldsymbol{x}) + b \quad (1)$$

with $\varphi(\boldsymbol{x})$ is the function for mapping the input values into the feature space, b is the coefficient of bias, and W is the weighting factor. The minimum error achieved when

$$\min \frac{1}{2}\|w\|^2 \quad (2)$$

with the constraint is

$$y_i - \langle w.\varphi(x_i)\rangle - b \leq \varepsilon \quad (3)$$

$$\langle w.\varphi(x_i)\rangle + b - y_i \leq \varepsilon \quad (4)$$

Adding of the slack $\xi$ and $\xi*$ is used to error optimizing. Equation (2) will be

$$\min \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i{}^*) \quad (5)$$

with the constraint is

$$y_i - \langle w.\varphi(x_i)\rangle - b \leq \varepsilon + \xi_i \quad (6)$$

$$\langle w.\varphi(x_i)\rangle + b - y_i \leq \varepsilon + \xi_i{}^* \quad (7)$$

$$\xi_i, \xi_i{}^* \geq 0 \quad (8)$$

There are two main groups of SVMs, Support Vector Classifier (SVC) and Support Vector Regression (SVR). SVC is used to solve problems related to discrete data, while SVR is used to solve regression cases or continues data. SVR was proposed by Vapnik, Steven Golowich, and Alex Smola in 1997 (Vapnik, Golowich, & Smola, 1996). SVR in its implementation uses high dimensional feature space or called Kernel. The purpose of the Kernel is to transform data into a higher feature space. The SVR equation for the Non-linear case is shown by Equation (9)

$$f(x) = \sum_{i}^{n}(\alpha_i + \alpha_i{}^*)K(x_i, x) + b \quad (9)$$

with $\alpha$ and $\alpha^*$ are Lagrange Multiplier, $K(x_i, x)$ is Kernel function. One of the Kernels on the SVR is the Radial Basis Function (RBF). The RBF Kernel is shown by Equation (10)

$$K(\boldsymbol{x}_i, \boldsymbol{x}_j) = \exp\left(-\frac{1}{2\sigma^2}\|\boldsymbol{x}_i - \boldsymbol{x}_j\|^2\right) \quad (10)$$

## III. METHOD

The data used in this study is JKSE from April 2021 to August 2021. The data is divided into two parts, the training data and the testing data. The training data starts from April 1st to July 30th 2021. The data used as ground truth is the data from August 2nd to August 31st 2021.

The evaluation of the model is done by finding the Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE). The RMSE and MAPE are calculated based on Equation (11) and Equation (12)

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{A_i - F_i}{A_i}\right| \times 100\% \quad (11)$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(A_i - F_i)^2} \quad (12)$$

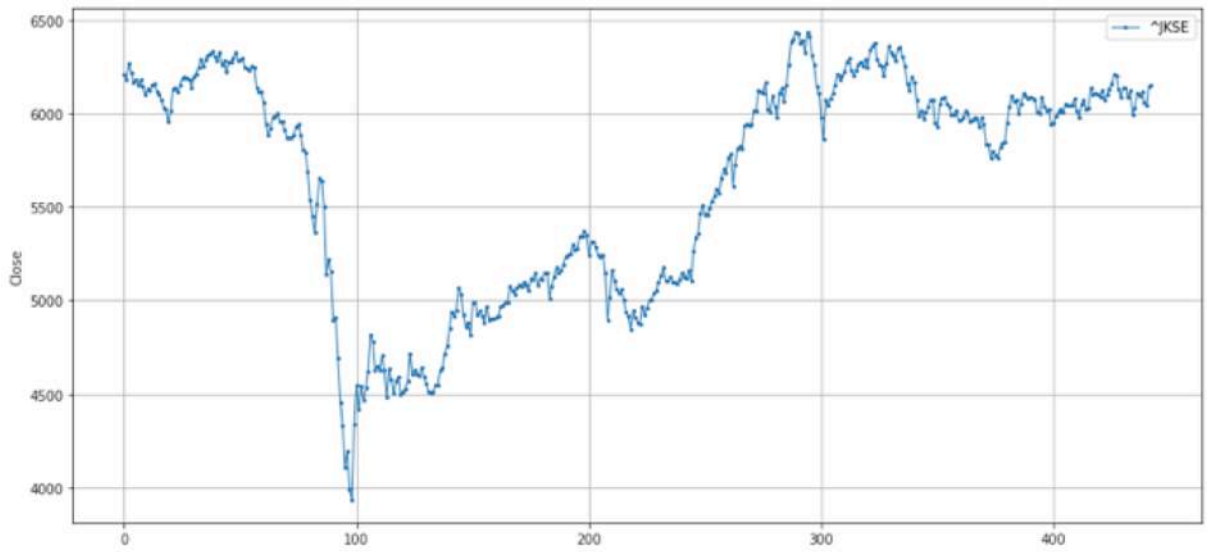


**Figure 1. Closing prices of the JKSE**

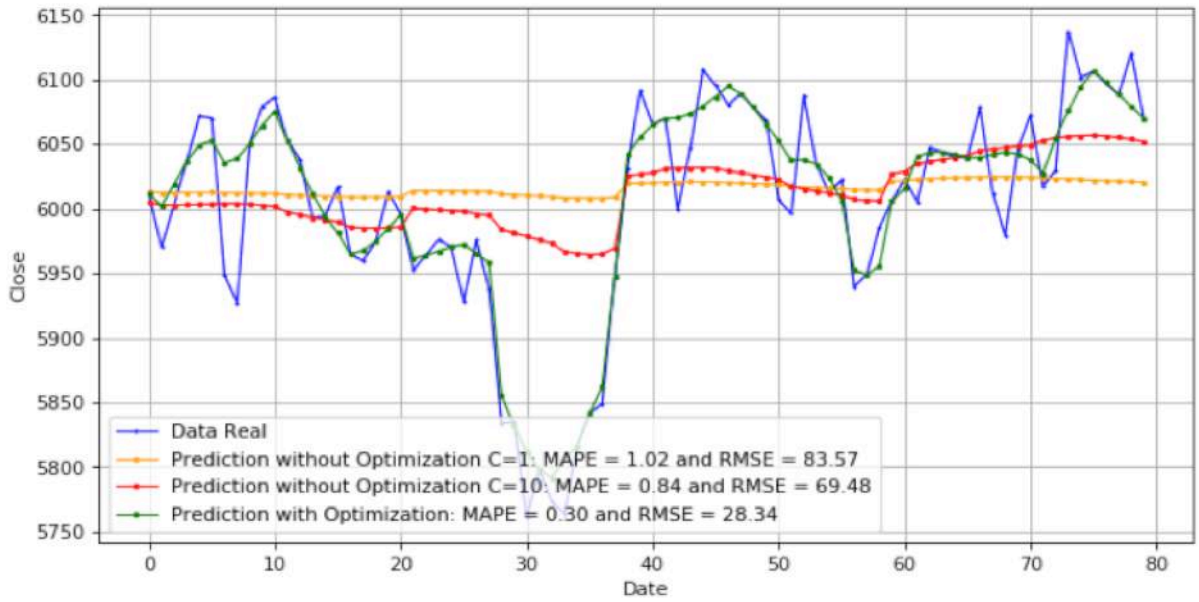


**Figure 2. Comparison between the results using optimization and without optimization**

To get the best hyperparameter, the research uses Grid Search with RBF kernel. The research (Yudhawan & Purwaningsih, 2020) (Shah, Shaikh, & Patel, 2017) shows that modeling using RBF for stock prices prediction has better accuracy than using polynomial or linear kernel. The optimal hyperparameters resulted by Grid Search are then used for data modeling.

## IV. RESULT AND DISCUSSION

The iteration process in Grid Search for searching the best hyperparameter is based on a combination of C, ε, and γ. The C values used in the cross-validation process are 1, 100, and 1000. The ε values are 0.01, 0.05, 0.1, 0.5, 1.0, and 5.0. The γ values are 1, 3, and 5. The results show the best hyperparameters using RBF kernel are C=100, ε =0.01, and γ = 3.

The hyperparameters obtained using Grid Search showed better performance than did not use Grid Search (Figure 2). At C=100, ε =0.01, dan γ = 3, the MAPE is 0.30 and the RMSE is 28.34. The use of Grid Search is basically a process of combining parameters so that the combination of

E-ISSN: 2775-4669

parameters will provide optimal results. Without using Grid Search, the optimal hyperparameters are obtained manually from a combination of parameters. In Figure 2, the increase of C affects the performance of the model. For C = 1, MAPE is 1.02 and RMSE is 83.57. For C=10, MAPE is 0.84 and RMSE is 69.48.

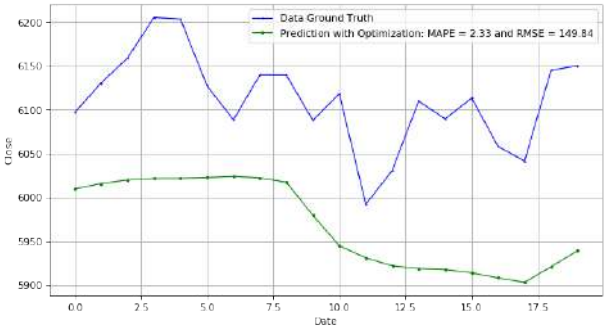


**Figure 4. Comparation between the ground truth and the prediction. The prediction model uses C=100, =0.01, dan γ = 3**

Figure 3 shows the comparison between the prediction results produced by using SVM and the ground truth for the next days (20 days). The MAPE and RMSE values obtained from the modeling towards the ground truth are 2.33 and 149.84. Although the testing results have relatively large MAPE and RMSE values compared to the modeling process, both the prediction results and the ground truth show the stable pattern and only decreases when certain point but not as severe as the decline in early days of the Covid-19 pandemic happened in Indonesia.

Stock prices are influenced by market sentiment. The positive or negative sentiment of the market is influenced by the level of investor confidence, one of which is trust in the policies taken by the government to solve the Covid-19. Figure 3 shows that although the daily cases on the second wave of the Covid-19 had reached 56,757, the number of recoveries of Covid-19 patients also increased (Figure 4).

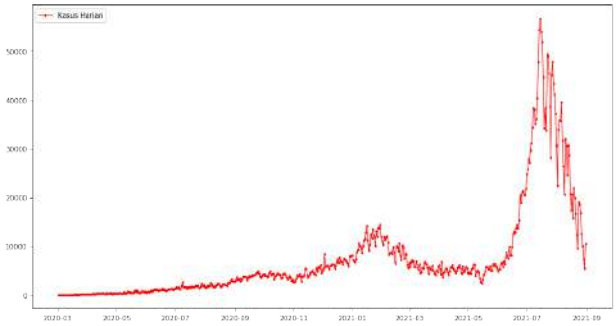


**Figure 3. Daily cases of the Covid-19 in Indonesia**

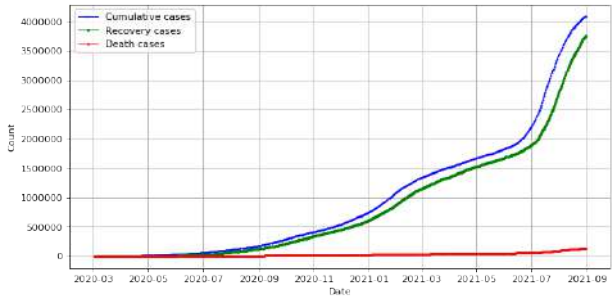


**Figure 5. Cumulative cases, recovery, and death cases of the Covid-19 in Indonesia**

## V. CONCLUSION

The best parameters to predict the JKSE stock price are C=100, ε =0.01, and γ = 3. MAPE and RMSE resulted by using RBF kernel are 0.30 and 28.34. The prediction result obtained tend to be stable and has a slight decrease. This shows that the investor has a positive sentiment towards the Indonesian government in solving the Covid-19 pandemic.

## REFERENCES

[1] A. I. Qureshi, W. I. Baskett, W. Huang, . I. Lobanova, S. H. Naqvi and C.-R. Shyu, "Reinfection with Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) in patients undergoing serial laboratory testing," *Clinical Infectious Diseases,* 2021.

[2] K. K.-W. To, I. F.-N. Hung, J. D. Ip, A. W.-H. Chu, W.-M. Chan, A. R. Tam, C. H.-Y. Fong, S. Yuan, H.-W. Tsoi, A. C.-K. Ng, L. L.-Y. Lee, P. Wan, E. Y.-K. Tso, W.-K. To, D. N.-C. Tsang, K.-H. Chan, J.-D. Huang, K.-H. Kok, V. C.-C. Cheng and K.-Y. Yuen, "Coronavirus Disease 2019 (COVID-19) Re-infection by a phylogenetically distinct severe acute respiratory syndrome coronavirus 2 strain confirmed by whole genome

sequencing," *Clinical Infectious Diseases,* 2021.

[3] J. Wang, C. Kaperak, T. Sato and A. Sakuraba, "COVID-19 Reinfection: A rapid systematic review of case reports and case series," *Journal of Investigative Medicine,* vol. 69, pp. 1253-1255, 2021.

[4] O. N. Bjørnstad, K. Shea, M. Krzywinski and N. Altman, "The SEIRS model for infectious disease dynamics," *Nature Methods,* vol. 17, pp. 557-558, 2020.

[5] V. Vapnik, S. E. Golowich and A. Smola, "Support vector method for function approximation, regression estimation and signal processing," in *In Proceedings of the 9th International Conference on Neural Information Processing Systems (NIPS'96)*, Cambridge, 1996.

[6] D. H. Yudhawan and T. Purwaningsih, "Developing support vector regression model to forcast stock prices of mining companies in Indonesia," *Jurnal Informatika,* vol. 14, no. 2, pp. 43-48, 2020.

[7] T. Shah, I. Shaikh and A. Patel, "Comparison of different kernels of support vector machine for predicting stock prices," *International Journal of Engineering and Technology (IJET),* vol. 9, no. 6, pp. 4288-4291, 2017.