

## Research Article

# Application of Support Vector Regression in Indonesian Stock Price Prediction with Feature Selection Using Particle Swarm Optimisation

Zuherman Rustam  and Puteri Kintandani

*Department of Mathematics, FMIPA Universitas Indonesia Kampus UI Depok, Depok 16424, Indonesia*

Correspondence should be addressed to Zuherman Rustam; [rustam@ui.ac.id](mailto:rustam@ui.ac.id)

Received 19 October 2018; Accepted 19 February 2019; Published 21 April 2019

Academic Editor: Michele Cali

Copyright © 2019 Zuherman Rustam and Puteri Kintandani. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Stock investing is one of the most popular types of investments since it provides the highest return among all investment types; however, it is also associated with considerable risk. Fluctuating stock prices provide an opportunity for investors to make a high profit. We can see the movement of groups of stock prices from the stock index, which is called Jakarta Composite Index (JKSE) in Indonesia. Several studies have focused on the prediction of stock prices using machine learning, while one uses support vector regression (SVR). Therefore, this study examines the application of SVR and particle swarm optimisation (PSO) in predicting stock prices using stock historical data and several technical indicators, which are selected using PSO. Subsequently, a support vector machine (SVM) was applied to predict stock prices with the technical indicator selected by PSO as the predictor. The study found that stock price prediction using SVR and PSO shows good performances for all data, and many features and training data used by the study have relatively low error probabilities. Thereby, an accurate model was obtained to predict stock prices in Indonesia.

## 1. Introduction

A stock is a sign of ownership of a company and indicates that a shareholder holds a share of the company's assets and earnings [1]. Stock is one of the most popular types of investment instruments; they are doubly beneficial since they provide both dividend and capital gain. Dividends are profits shared among shareholders based on the number of shares held by them. Further, capital gain is a benefit of the difference between the purchase and selling prices; however, it can become capital loss if the selling price is lower than the purchase price. Stock prices are determined by the marketplace, where the seller's supply meets the buyer's demand. Unfortunately, since there is no specific equation that can exactly determine how a stock price will behave, stock prices are always fluctuating. These fluctuations provide an opportunity for investors to make big profits; however, they present big risks, as well. This is because numerous factors, such as company news and performance, industry

performance, investor sentiment, and economic factors, influence stock price fluctuations.

A stock index denotes stock price movement. One of the stock indices prevalent in Indonesia is the Jakarta Composite Index (JKSE). It denotes the movement, whether increase or decrease, of the stock prices of all securities listed on the Indonesia Stock Exchange (IDX). This is an important concern for investors since JKSE affects the attitudes of investors regarding whether to buy, hold, or sell their shares. To ensure that this prediction model applies to other stock data as well, this study uses three real estate stock datasets listed on IDX.

Stock price prediction mechanisms are fundamental to the formation of investment strategies and development of risk management models [2]. Computational advances have led to the formulation of several machine learning algorithms that can be used to anticipate market movements consistently and, thereby, estimate future asset values, such as company stock prices [3]. Machine learning is a branch of

science that allows computers the ability to learn based on existing data. One type of machine learning is support vector regression (SVR). SVR is the development of the support vector machine (SVM) method for regression cases. One of its advantages when compared to other regression models, such as ordinary least square (OLS), is that SVR can handle nonseparable data, whereas the OLS method gives a poor prediction result for such data [4].

Feature selection will be done using particle swarm optimisation (PSO) to maximise the performance of SVR. Using PSO for feature selection will reduce computational time. It is an evolutionary computation technique, which is computationally less expensive than other evolutionary computation algorithm [5, 6].

This study focuses on one such machine learning algorithm, that is, PSO, to select any indicator that has an effect on stock prices. Further, it predicts stock prices by inserting the selected indicator into the SVR.

## 2. Literature Review

Over the past few decades, numerous researchers have conducted studies on the prediction of stock prices using machine learning and deep learning. Henrique, Sobreiro, and Kimura used SVR for stock price prediction on daily and up-to-the-minute prices [7]. Hiransha et al. [8] used deep-learning models for NSE stock market prediction [8].

**2.1. Technical Analysis.** Technical analysis is defined as the art of predicting stock prices based on current commodity offerings, stocks, indices, futures, or tradeable instruments. It inserts stock price and volume information into a chart and applies various patterns and indicators to assess future stock price movements [9]. Usually, the literature focusing on SVM and SVR uses technical analysis indicators. Table 1 depicts the formulas of technical analysis [10]:

The closing price on day  $t$  is denoted by  $C_t$ , and the number of trading days used is represented by  $n$ .  $EMA_{t-1}$  is the exponential moving average one day before  $t$ ,  $\alpha$  is the weight coefficient,  $Dw_{t-1}$  denotes the changes in price decreases, and  $Up_{t-1}$  changes in price increases.  $L_{t-x}$  is the low price  $x$  days before  $t$ , and  $H_{t-x}$  is the high price  $x$  days before  $t$ . Low is the lowest price, High is the highest price, and Close is the last price for day  $t$ . The expression  $\%K_i$  represents the stochastic  $\%K$  on day  $i$ . A typical price of day  $i$  is denoted by  $TP_i$ .  $SMA_1$  and  $SMA_2$  are simple moving average values for a certain period, and the differences between the typical price and simple moving average of day  $i$  are denoted by  $D_i$ .

**2.2. Particle Swarm Optimisation.** In 1995, Dr. Eberhart and Dr. Kennedy proposed the PSO algorithm [5]. This algorithm mimics the behaviour of a flock of birds. Similar to a bird in a flock or a particle in a group of particles, each individual in a group has his or her own intelligence and can influence group behaviour. Each individual or particle behaves in a mutually connected manner by using one's own intelligence and is influenced by the behaviour of the group. Therefore, if a particle or a bird finds the right path or

TABLE 1: Formulas of technical analysis.

Technical indicator	Formula
SMA	$(1/N)\sum_{k=1}^N C_t$
EMA	$(C_t \times \alpha) + (EMA_{t-1} \times (1 - \alpha)), c > n$
MACD	$EMA_{12} - EMA_{26}$
MOM	$C_t - C_{t-x}$
RSI	$100 - (100 / (1 + (\sum_{i=1}^{n-1} Up_{t-1}/n) / (\sum_{i=1}^{n-1} Dw_{t-1}/n)))$
ROC	$((C_t - C_{t-x}) / C_{t-x}) \times 100$
Stochastic %K	$((C_t - L_t) / (C_t - L_t)) \times 100$
Stochastic %D	$(\sum_{i=1}^n \%K_i/n) \times 100$
	$TP = ((High + Low + Close)/3)$
	$SMA_1 = (\sum_{i=1}^n TP_i/n)$
CCI	$D_i = TP_i - SMA_1$
	$SMA_2 = (\sum_{i=1}^n D_i/n)$
	$M_i = SMA_2 \times 0.015$
	$CCI = (M_i/D_i)$

SMA, simple moving average; EMA, exponential moving average; MACD, moving average convergence/divergence; MOM, momentum; RSI, relative strength index; ROC, rate of change; CCI, commodity channel index.

shortest path to a food source, then their group will be able to immediately follow the path, even though the location of the particle or bird is far away from the group. The particles in a group have a certain size, and each particle has two characteristics, namely, location and velocity.

PSO is a well-known tool for finding the optimal characteristics of a particle by performing local and global iterative searches in the feature search space [11]. In PSO, there is a group of random particles that moves around the solution space until convergence is reached. There are some features that are irrelevant and noisy and lead to high misclassification rates [11]. Therefore, PSO is used for feature selection to reduce noisy features and remove irrelevant features.

This model is simulated in space with a certain dimension; an increase in the number of iterations indicates that the particle is getting closer to the intended target. This is done until the maximum iteration is achieved, or another criterion can be used.

First, we generate the initial positions,  $\mathbf{X}_i$ , and initial velocities,  $\mathbf{V}_i$ , of  $n$  random particles; then, for each particle, evaluate the fitness function,  $f(x_i)$ , from each particle based on its position. We determine the particle position with the best fitness and set it as  $gbest$ . For each particle, the  $pbest$  is the particle position with the best fitness that has been obtained by each particle so far. After determining  $pbest$  and  $gbest$ , the positions and velocities will be updated as specified by Seal and colleagues [11]:

$$\begin{aligned} \mathbf{V}_i(t+1) &= w \times \mathbf{V}_i(t) + c_1 \times ud \times [pbest_i(t) - \mathbf{X}_i(t)] \\ &\quad + c_2 \times ud \times [gbest(t) - \mathbf{X}_i(t)], \quad (1) \\ \mathbf{X}_i(t+1) &= \mathbf{X}_i(t) + \mathbf{V}_i(t+1), \end{aligned}$$

where  $w$  is a factor used to control the balance of search between exploitation and exploration,  $c_1$  and  $c_2$  are cognitive and social parameters, respectively, which have a value between 0 and 1,  $ud$  is another random value bounded between 0 and 1,  $i = 1, 2, 3, \dots, n$ , and  $n$  is the size of population.

**2.3. Support Vector Machines for Regression.** SVM is a machine learning method invented by Vladimir Vapnik and team in 1992. Initially, an SVM was used to solve classification problems alone; however, now, it has been developed to solve regression problems, as well. It is noted that SVR involves the application of SVM in regression cases; in this case, the output of this method comprises real numbers. Further, since SVR can solve the problem of overfitting, it has a good performance [12].

SVM is a method of a classification that maps  $N$  independent samples to a higher space dimension, and it is used to classify observations between groups. This study employed SVR by using training data  $\{(\mathbf{x}_k, y_k)\}_{k=1}^N$  to build a linear model using nonlinear classification.

Suppose there are  $n$  training data  $\{(\mathbf{x}_k, y_k)\}_{k=1}^N$ , the input is  $\mathbf{x} = \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \dots\} \in R^N$  and the output is  $y = \{y_1, y_2, y_3, \dots\} \in R$ . The purpose of this method is to find a function of  $y(\mathbf{x})$  that gives the smallest error,  $\varepsilon$ , for all the learning data  $x_i$ .

The primary SVR problem can be defined as follows [12]:

$$\begin{aligned} & \text{minimize} && \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^l (\xi_i + \widehat{\xi}_i), \\ & \text{subject to} && \mathbf{w} \cdot \mathbf{x}_i + b - y_i \leq \varepsilon + \xi_i, \quad i = 1, \dots, l, \\ & && y_i - \mathbf{w} \cdot \mathbf{x}_i - b < \varepsilon + \widehat{\xi}_i, \quad i = 1, \dots, l, \\ & && \xi_i, \widehat{\xi}_i \geq 0, \quad i = 1, \dots, l, \end{aligned} \quad (2)$$

where  $\mathbf{w}$  is a  $d$  dimensional weight vector. The constant  $C > 0$  determines the trade-off between the differences in decision function, where the upper limit of deviation that is more than  $\varepsilon$  can still be tolerated [12]. A deviation greater than  $\varepsilon$  will be subject to a penalty of  $C$ . Further, high slack variable values cause empirical errors to affect regulatory factors significantly. In SVR, the support vector is a training data value that is located on or outside the boundary of the decision function; therefore, the number of support vectors decreases with an increase in error values,  $\varepsilon$ .

In dual formulations, the optimisation problem of SVR is represented as follows [12]:

$$\begin{aligned} & \max && \left[ -\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (\alpha_i - \widehat{\alpha}_i)(\alpha_j - \widehat{\alpha}_j) k(\mathbf{x}_i, \mathbf{x}_j) \right. \\ & && \left. + \sum_{i=1}^n (\alpha_i - \widehat{\alpha}_i) y_i - \varepsilon \sum_{i=1}^n (\alpha_i - \widehat{\alpha}_i) \right], \quad (3) \\ & \text{subject to} && \left\{ \sum_{i=1}^n (\alpha_i - \widehat{\alpha}_i) = 0, \quad 0 \leq \alpha_i \leq C, 0 \leq \widehat{\alpha}_i \leq C, \right. \end{aligned}$$

where  $k(\mathbf{x}_i, \mathbf{x}_j)$  denotes the kernel function, which is defined as  $k(\mathbf{x}_i, \mathbf{x}_j) = \varphi(\mathbf{x}_i) \cdot \varphi(\mathbf{x}_j)$ , in which  $\varphi$  is a mapping from data space to the feature space  $F$ .  $\alpha_i$  and  $\widehat{\alpha}_i$  are the Lagrange multipliers. Using the Lagrange multiplier and optimality conditions, the regression function can be explicitly formulated as follows:

$$y(x) = \sum_{i=1}^n (\alpha_i - \widehat{\alpha}_i) k(\mathbf{x}_i, \mathbf{x}) + b. \quad (4)$$

### 3. Materials and Methods

This study proposes the application of SVR with PSO for feature selection. The closing price is initially used as a raw input; subsequently, analysis techniques are used to transform the raw input to a technical analysis indicator. In case the indicator has a difference scale, the data in each indicator should be normalised. A normalised technical analysis indicator is used as a feature to predict the stock price and is then selected using PSO. It is noted that PSO is the best tool to determine the optimal characteristics of a feature by finding the local best and global best in the feature search space with the help of a local, as well as global, search in an iterative manner to make model prediction more effective.

**3.1. Data Set.** This study used the Stock Composite Index and several stocks from the real estate sector in Indonesia as data sets. The JKSE is one of the main indicators reflecting the performance of the capital market in Indonesia; it records the price movements of shares of all the securities listed on the IDX, as well. The data used comprise the adjusted closing price of the daily data history from Yahoo Finance, and the amount of data used is 650 datasets from 4 January 2016 to 10 September 2018. Using technical analysis, the data are processed into several indicators, such as simple moving average, exponential moving average, momentum, rate of change, moving average convergence divergence, commodity channel index, relative strength index, and stochastic %K and stochastic %D.

**3.2. Data Preprocessing.** In this study, the data were pre-processed using technical analysis, normalisation, and feature selection through PSO. The technical analysis indicator is obtained by applying a formula on the daily data history of stock prices [7]. The daily data history of stock price components used by the study includes close price, low price, and high price. Further, 14 indicators are included in this study using the formula in [10].

As discussed previously, we normalised the data to equalise the value scale. The formula used to normalise the data to range  $[-1, 1]$  is as follows [8]:

$$x'(t) = \left( \frac{x(t) - \min A}{\max A - \min A} (\max' A - \min' A) \right) + \min' A, \quad (5)$$

where  $x'(t)$  is the value of the technical indicator,  $A$ , on the normalised day,  $t$ ;  $x(t)$  is the value of the technical indicator,  $A$ , on day,  $t$ ;  $\min A$  is the smallest value of technical indicator,  $A$ ; and  $\max A$  is the largest value of technical indicator,  $A$ . Further,  $\min' A$  and  $\max' A$  denote the new smallest and new largest values of technical indicator,  $A$ , respectively. Then, the normalised value of the technical indicator is selected using PSO. The selected features are used as input data for the prediction process.

**3.3. Feature Selection.** Feature selection, or variable selection, can be briefly described as a tool to select the variable that can represent the original data set [13]. It is the process of selecting the best  $i$  features in a set of data by using an algorithm to evaluate the features [14]. It has many advantages, one of which is that it increases the accuracy of the resulting model [15]. PSO is one of the tools for feature selection. In a previous study, Seal et al. researched the use of PSO for feature selection in thermal face recognition [11]. Cases that used PSO provided markedly better performance results than those that did not use PSO; this inspired the author to try it in predicting the stock price model.

First, we input the variable into the PSO program and then obtain the output as a cost score from PSO. Subsequently, the score is sorted by the smallest score. After being sorted, we create groups of input data with one variable into data with 14 variables and input the data into the SVR program. This study performed 10%, 20%, and 90% data training to find the best model.

**3.4. Performance Criteria.** In this study, SVR was used to predict close prices. The results are compared with the returns obtained by the given variance [7]. The adequacy of the model's price prediction can be evaluated using methods such as the root-mean-squared error, mean absolute percentage error, and normalised mean squared error (NMSE). In this study, the NMSE was calculated according to the following equation [16]:

$$\text{NMSE} = \frac{1}{n} \left( \frac{1}{s^2} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \right), \quad (6)$$

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2,$$

where the observed value is represented as  $y_i$ , the predicted value is shown as  $\hat{y}_i$ , and the mean of the observed value is depicted as  $\bar{y}$ .

## 4. Experimental Results

The experimental results of stock price prediction using SVR and PSO showed good performances for all the data that were used, and many features and training data used by the study had relatively small NMSE values, which averaged below 0.1. The data used are JKSE and real estate stock data, comprising Alam Sutera Realty Tbk (ASRI), Agung Podomoro Land Tbk (APLN), and Serpong Damai Tbk (BSDE).

Table 2 reveals that the experiment results of using the SVR algorithm and the PSO for feature selection for JKSE data, with 70% training data, have the lowest NMSE value. By using 70% training data, the use of 12, 13, and 14 features give the lowest value of NMSE.

TABLE 2: PSO-SVR NMSE values of JKSE.

		Features			
		11	12	13	14
Data training	40%	0.0010	0.0010	0.0009	0.0008
	50%	0.0008	0.0007	0.0007	0.0006
	60%	0.0006	0.0005	0.0005	0.0005
	70%	0.0007	0.0004	0.0004	0.0004
	80%	0.0010	0.0007	0.0006	0.0005
	90%	0.0061	0.0046	0.0046	0.0045

TABLE 3: PSO-SVR NMSE values of Alam Sutera Realty Tbk (ASRI).

		Features			
		11	12	13	14
Data training	40%	0.0018	0.0013	0.0012	0.001
	50%	0.0011	0.0008	0.0007	0.0006
	60%	0.0008	0.0006	0.0005	0.0004
	70%	0.0006	0.0004	0.0004	0.0003
	80%	0.0004	0.0003	0.0003	0.0002
	90%	0.0003	0.0003	0.0002	0.0002

TABLE 4: PSO-SVR NMSE values of Agung Podomoro Land Tbk (APLN).

		Features			
		11	12	13	14
Data training	40%	0.0001	0.0001	0.0001	0.0001
	50%	0	0	0	0
	60%	0	0	0	0
	70%	0	0	0	0
	80%	0	0	0	0
	90%	0	0	0	0

TABLE 5: PSO-SVR NMSE values of Bumi Serpong Damai Tbk (BSDE).

		Features			
		11	12	13	14
Data training	40%	0.0002	0.0001	0.0001	0.0001
	50%	0.0001	0.0001	0.0001	0.0001
	60%	0.0001	0.0001	0.0001	0.0001
	70%	0.0001	0	0	0
	80%	0.0001	0	0	0
	90%	0.0002	0.0001	0	0

As shown in Table 3, the smallest NMSE value for ASRI uses 90% training data by using 13 features and 80% and 90% training data by using 14 features.

Table 4 clarifies that, by using the same number of features using 50% to 90% training data, the smallest NMSE value of APLN is obtained. On the other hand, by using the same number of training data using 11, 12, 13, and 14 features, the smallest value of NMSE is obtained.

Table 5 depicts that the smallest NMSE value of BSDE uses 12 features by using 70% and 80% training data; 13

features by using 70%, 80%, and 90% training data; and 14 features by using 70%, 80%, and 90% training data.

Although the NMSE value is the same for data with and without feature selection, the use of a smaller feature is more beneficial, since it will streamline the running time of the program. For example, when considering the use of 13 or 14 features to obtain the same value of NMSE, it is better to use 13 features. The same aspect applies to training data as well; for example, if the use of 80% and 90% training data gives the same NMSE value, then it is better to use the former, since it will streamline the running time of the program.

## 5. Conclusions

Based on the experimental results of this study, we conclude that the prediction model for stock price prediction using SVR with feature selection using PSO exhibits good performance, since it has relatively small NMSE values, which average below 0.1. Although some data provide the same smallest NMSE value with and without feature selection, it would be better to use a small feature, since it will streamline the running time of the program. Moreover, depending on the type of data, the order of features in each data varies. Thus, an accurate stock price prediction model is obtained so investors can predict future stock prices and gain profits. In subsequent studies, more technical indicators will be added, and another feature selection and another data input will be used as a comparison.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

## Acknowledgments

This research was financially supported by the Indonesia Ministry of Research and Higher Education, with a PDUPT 2018 research grant scheme (ID number 389/UN2.R3.1/HKP05.00/2018).

## References

- [1] H. K. Penawar and Z. Rustam, "A fuzzy logic model to forecast stock market momentum in Indonesia's property and real estate sector," in *Proceedings of 2nd International Symposium on Current Progress in Mathematics and Sciences*, Depok, Indonesia, November 2016.
- [2] R. Dash and P. K. Dash, "A hybrid stock trading framework integrating technical analysis with machine learning techniques," *Journal of Finance and Data Science*, vol. 2, no. 1, pp. 42–57, 2015.
- [3] E. A. Gerlin, M. McGinnity, A. Belatreche, and S. Coleman, "Evaluating machine learning classification for financial trading: an empirical approach," *Expert Systems with Applications*, vol. 54, pp. 193–207, 2016.
- [4] V. N. Vapnik, *The Nature of Statistical Learning Theory*, Springer, New York, NY, USA, 2000.
- [5] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of the International Conference on Neural Networks*, vol. 4, pp. 1942–1948, Washington, DC, USA, 1995.
- [6] Y. Shi and R. Eberhart, "A modified particle swarm optimizer," in *Proceedings of International Conference on Evolutionary Computation*, pp. 69–73, Anchorage, AL, USA, May 1998.
- [7] B. M. Henrique, V. A. Sobreiro, and H. Kimura, "Stock price prediction using support vector regression on daily and up to the minute price," *Journal of Finance and Data Science*, vol. 4, no. 3, pp. 183–201, 2018.
- [8] M. Hiransha, E. A. Gopalakrishnan, V. K. Menon, and K. P. Shoman, "NSE stock market prediction using deep-learning models," *Procedia Computer Science*, vol. 132, pp. 1351–1362, 2018.
- [9] C. Boobalan, "Technical analysis in select stocks of Indian companies," *International Journal of Business and Administration Research Review*, vol. 2, no. 4, pp. 26–36, 2014.
- [10] Z. Rustam, D. F. Vibranti, and D. Widya, "Predicting the direction of Indonesian stock price movement using support vector machines and fuzzy kernel C-means," in *Proceedings of 3rd International Symposium on Current Progress in Mathematics and Sciences*, Bali, Indonesia, July 2017.
- [11] S. G. Seal, D. Bhattacharjee, M. Nasipuri, and C. G. Martin, "Feature selection using particle swarm optimization for thermal face recognition," in *Proceedings of Applied Computation and Security System*, pp. 25–35, Kolkata, India, 2015.
- [12] J. Smola and B. Scholkopf, "A tutorial on support vector regression," *Journal of Statistics and Computing*, vol. 1, pp. 199–222, 2002.
- [13] H. Budak and S. E. Tasabat, "A modified t-score for feature selection," *Anadolu University Journal of Science and Technology A-Applied Sciences and Engineering*, vol. 17, no. 5, pp. 845–852, 2016.
- [14] G. Forman, "An extensive empirical study of feature selection metrics for text classification," *Journal of Machine Learning Research*, vol. 3, pp. 1289–1305, 2003.
- [15] L. Ladha and T. Deepa, "Feature selection methods and algorithms," *International Journal on Computer Science and Engineering*, vol. 3, pp. 1787–1797, 2011.
- [16] Y. Bao, Z. Liu, L. Guo, and W. Wang, "Forecasting stock composite index by fuzzy support vector machines regression," in *Proceedings of the Fourth International Conference on Machine Learning and Cybernetics*, Guangzhou, China, August 2005.



**Hindawi**

Submit your manuscripts at  
[www.hindawi.com](http://www.hindawi.com)

