Indonesian Journal of Electrical Engineering and Computer Science
Vol. 38, No. 1, April 2025, pp. 508~516
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v38.i1.pp508-516

Comparing machine learning models for Indonesia stock market prediction

Selly Anastassia Amellia Kharis[1], Arman Haqqi Anna Zili[2], Maulana Malik[2], Wahyu Nuryaningrum[2], Agustiani Putri[3]
[1]Department of Mathematics, Faculty of Science and Technology, Universitas Terbuka, Tangerang, Indonesia
[2]Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Indonesia, Depok, Indonesia
[3]Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Negeri Malang, Malang, Indonesia

Article Info

Article history:

Received Feb 19, 2024
Revised Oct 17, 2024
Accepted Oct 30, 2024

Keywords:

Investment
K-nearest neighbor
Random forest
Stock price prediction
Support vector regression

ABSTRACT

The financial market hold a significant role in the economy and the ability to accurately predict stock prices poses a major challenge, particularly in volatile markets like Indonesia. This study investigates the application of three supervised machine learning algorithms: random forest (RF), support vector regression (SVR), K-nearest neighbor (KNN) to predict the closing prices of stocks. The data used in this research consists of BBCA, PWON, and TOWR stocks. This study adopted daily historical stock prices from March 2017 to February 2020, which were normalized and segmented into training and testing datasets. The models were trained using machine learning techniques, and their predictive accuracy was evaluated using root mean square error (RMSE) and mean absolute error (MAE). The historical stock data includes Open, High, Low, and Close prices. The result indicated that SVR consistently outperforms RF and KNN in terms of RMSE and MAE across different stocks. The SVR method produced RMSE values of 4.79% for BBCA stock, 10.61% for PWON stock, and 15.14% for TOWR stock, and produces MAE values of 3.52% for BBCA stock, 8.49% for PWON stock, and 13.78% for TOWR stock.

Corresponding Author:

Selly Anastassia Amellia Kharis
Department of Mathematics, Faculty of Science and Technology, Universitas Terbuka
15418 South Tangerang, Banten, Indonesia
Email: selly@ecampus.ut.ac.id

1. INTRODUCTION

The welfare of any growing nation, economy, or community in the 21st century primarily rests over their stock price and market economy, by the financial market serving as the key pillar [1], [2]. Financial markets remain unquestionably foremost among the most exhilarating inventions of recent years. Getting precise forecasts for financial future time series recently remains a challenging undertaking for numerous scholars [3]-[5], especially owing to the presence of nonlinear, irregular, and unpredictable nature [6]. By the advent of quantitative financial management, accurate forecasts of stock price shifts are essential for investment approaches, which has captured the considerable enthusiasm from companies and scholars. Despite machine learning models are frequently used in emerging markets, a significant gap exists in understanding their effectiveness, particularly in volatile regions like Indonesia. This study addresses to bridge that gap by comparing the performance of three machine learning models: random forest (RF), support vector regression (SVR), and K-nearest neighbor (KNN) within the Indonesian stock market context.

Forecasting stock prices in emerging markets such as Indonesia be a difficult issue due to inherent volatility, nonlinear behavior, and the influence of various economic, political, and psychological factors.

Traditional statistical models often struggle to account for these complexities, leading to inaccurate predictions and potential financial losses [7], [8]. In the context of the Indonesian stock market, prevalent literature indicates that stock prices are influenced by various factors including political issues, economic conditions, commodity price indexes, investor expectations, shifts in different stock markets, and investor psychology [9]. These multifaceted influences contribute to the complexity of accurately forecasting stock price movements. Consequently, there is a need for machine learning methods that are capable of modelling and analyzing these intricate patterns. The significance of stock classification is often reflected in high market capitalization, and various technical metrics are available to derive insights from stock price data [10].

Commonly, stock index deals are derived from stocks prices with heavy market investments and these indices frequently will provide a forecast of the status of the economy of each nation. For instance, several literature evidence that the economic development in many nations has been significantly affected by the capitalization of their stock markets [11]. However, shifts in stock prices exhibit vague properties, which puts investments at risk for investors. Additionally, it is challenging to identify the market's status with respect to governments. In fact, stock prices are inherently volatile, nonlinear, and unpredictable. Hence, this frequently leads to underperformance in statistical forecasting models and failure to forecast values and shifts accurately [12], [13]. This challenge underscores the need to explore stronger and more adaptive forecasting methods.

The Indonesian stock market is renowned for its volatility, with prices changes frequently by the previous day's closing price. These circumstances make it difficult for traditional time series forecasting approaches, which depend on stable trends to perform. Over the short timeframe, the market performs likewise to the voting tool, but over the more prolonged timeframe, it is behaviorally similar to the weighing tool and therefore a scope exists for forecasting the market shifts for an extended period [14]. Machine learning has become the foremost influential instrument encompassing distinct algorithms for efficiently evolving their version on a specific issue. Machine learning is widely recognized as having noteworthy powers in recognizing accurate information and capturing trends from data sets [15]. In this study, three supervised machine learning algorithms: RF, SVR, and K-NN are employed to forecast the closing prices of stocks in the Indonesian market. These models utilize a newly created set of variables derived from financial datasets, including open, close, low, and high prices for specific enterprises. These indicators are selected to enhance the models' precision in forecasting the following day's closing prices.

Among all non-parametric models, KNN is a popular approach and broadly implemented in numerous prediction [16]-[18]. In this strategy, the number of nearest neighbors (K) determines the model's ability to capture relationships within the data, with the root mean square error (RMSE) serving as a key metric for performance evaluation. Thus, the nearest neighbors model presents the points of data with low RMSE and large resemblance. This approach provides excellent predictive power for both multidimensional and imperfect data. K-NNs approach is ideal for forecasting against the stock market [19], [20]. Furthermore, SVR is also one of the machine learning algorithms that considerably applied for forecasting both the values of stock market indices and stock price [21]. A number of studies have offered to utilize RF for the purpose of prediction. RF is a commonly used ensemble technique that performs the task of classification and regression. This instrument works by building several judgment trees in training time that produces the average regression from the single judgement trees [22].

Previous studies using various machine learning models to investigate stock price prediction. Henrique et al. predict stock prices for large and small capitalizations in three different markets using SVR [23]. Zheng *et al.* [24] employed RF for analyze and forecast the US Stock Market using optimal parameters. Furthermore, Sarala and Bhushan predict the stock price using KNN approach with a probabilistic method [25]. Unlike previous research that have predominantly focused on developed markets or single-model approaches, this study uses various machine learning methods on daily historical data to compare three stocks data in Indonesia, which is an emerging country and has high volatility. This study provides new insights for investors, particularly in emerging economies like Indonesia, regarding the use of machine learning to predict stock prices and effectively adapt to such conditions.

The effectiveness of these models is evaluated using two key metrics: RMSE and mean absolute error (MAE). RMSE is defined as the average value of the sum of squared errors that can be utilized to measure the magnitude of the error value in a model. MAE is the average value of the absolute difference between predicted and actual data to measure the magnitude of a model error without considering its direction. A model is accurate if it has lower RMSE and MAE values. This research utilizes daily historical data from IDX30 stocks for analysis. This study contributes to advancement of financial forecasting using machine learning. The following sections of this paper are organized as follows: the method section describes data historical stock price data collection, data preprocessing, model construction and justification, and performance evaluation. The results and discussion section describe the findings, comparing SVR, RF, and KNN performance of each model in the Indonesian market. Finally, the conclusion highlights the findings and study's contributions to the field and suggests directions for future research.

2. METHOD

Figure 1 shows the research step. This study utilized daily historical stock prices from March 2017 to February 2020. The research collecting historical stock price data, which is subsequently normalized and divided into training and testing datasets. Machine learning is employed on the training data to train the model. Subsequently, a predictive model is constructed using RF, SVR, and KNN. The testing data is then evaluated using each method to obtain prediction results. The predictive outcomes are compared with the actual results, and accuracy is calculated using RMSE and MAE. After obtaining the accuracy results for each method, a comparison is made to determine the best predictive model.

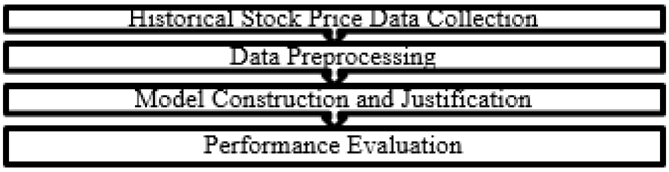


Figure 1. Research step

2.1. Historical stock price data collection

This study utilized a dataset comprising daily historical stock data totaling 725 samples, including three IDX30 stocks: PT. Bank Central Asia Tbk (BBCA), PT. Pakuwon Jati Tbk (PWON), and PT. Sarana Menara Nusantara Tbk (TOWR). The dataset includes key variables such as Open, High, Low, and Close prices with each variable representing specific aspects of daily stock transactions. ‘Open’ refers to the price at the beginning of the transaction activity, ‘High’ and ‘Low’ indicate the highest and lowest prices during the transaction activity on a given day, and ‘Close’ is defined as the last appearing stock price before the market closes as shown in Figure 2. The data was sourced from investing.com website, ensuring a reliable and comprehensive dataset for analysis. This study incorporated a dataset of historical daily stocks, in which the Open, High, and Low price were labelled as independent variables, whereas the Close price was the dependent variable. The data analysis, illustrated in Figure 2(a), demonstrates that BBCA exhibits a relatively stable stock fluctuation, demonstrating a bullish trend over time, as opposed to PWON in Figure 2(b), and TOWR in Figure 2(c).

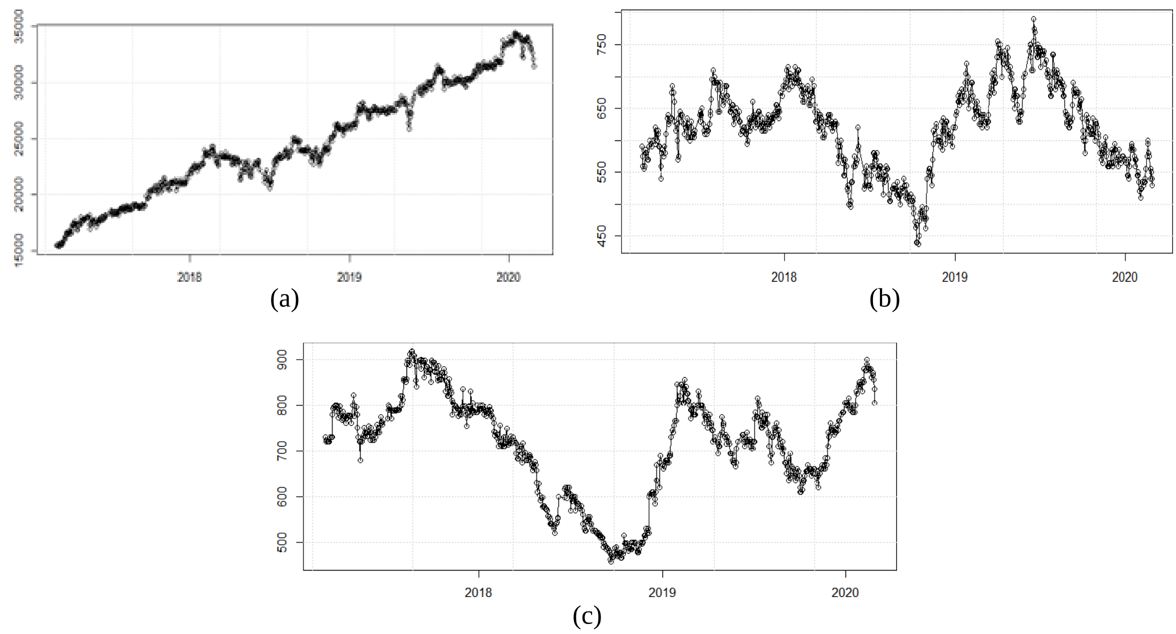


Figure 2. Stock price trends: (a) BBCA, (b) PWON, and (c) TOWR

2.2. Data preprocessing

Historical stock data frequently manifests non-linear characteristics. In this context, to minimize the error rate of the forecast output, data normalization is necessary. Normalization is a crucial preprocessing step to ensure the comparability of variables. This study employed z-score normalization to standardize the data, ensuring all variables are on a comparable scale. Z-score normalization was selected because it is less sensitive to outliers and better suited for algorithms like SVR, which assume that the data is normally distributed. The z-score normalization as shown in (1):

$$p_i^* = \frac{p_i - \mu_p}{\sigma_p} \tag{1}$$

where p_i represents the observed value, μ_p denotes the mean of variable, and σ_p represents the standard deviation. This process is applied uniformly to the open, high, low, and close prices across all selected stocks. The normalized data was then split into training and testing sets based on the chronological order of the transaction period. The training data included 705 samples from March 1, 2017 to January 31, 2020. The testing data comprised 20 samples from February 2020, intended to evaluate the predictive model by comparing the predicted outputs with the actual values to calculate the error rate.

2.3. Model construction and justification

This research employed three machine learning algorithms to generate forecasting models: RF, SVR, KNN regression. These models were selected because they offer a balance between complexity and interpretability, making them ideal for analyzing the volatile and multifaceted nature of the Indonesian stock market.

2.3.1. Random forest

RF displays numerous growing trees forming a forest, where each tree is built using a random subset of the data and features, providing robust predictions through the aggregation of multiple models. By averaging each tree's output, this ensemble technique improves prediction accuracy while lowering the variance often seen in single decision trees [18]. Furthermore, the prediction outcomes from each regression tree are averaged, producing the RF prediction output. The RF algorithm uses bootstrapping or random resampling of data to generate n bootstrap samples, each used to construct a regression tree. The predicted values generated by the RF approach can be expressed through (2) where y_i is the output of the prediction from the i^{th} regression tree and B is the number of regression trees.

$$y = \frac{1}{B}\sum_{i=1}^{B} y_i \tag{2}$$

2.3.2. Support vector regression

SVR is an approach stemming from support vector machines (SVM) used to regression problems. This approach generates real or continuous-valued outputs with the aim of finding a function as a separating line (hyperplane) in the form of a regression function, implementing the concept of ε-insensitive area. The error tolerance between predicted values and actual data can be specified by a value of ε. SVR has proven to deliver excellent performance as it addresses the issue of overfitting in data [26]. This method mechanism involves seeking the maximum distance between two classes to obtain the optimal hyperplane that separates the two classes [27]. In this context, the linear model to be used as the regression function in SVM to determine the hyperplane has a general form, as outlined in (3) where x^* as the vector of independent variables in the training data, w^T as the weight parameters in the model, $\emptyset(x^*)$ as the feature transformation function on the training data, and b as a bias term (in the form of a constant).

$$f(x^*) = w^T\emptyset(x^*) + b \tag{3}$$

2.3.3. K-nearest neighbor regression

KNN is one of the approaches used for predictive analysis on a dataset based on training data extracted from its k nearest neighbors within the dataset. These nearest neighbors are used to predict the response variable values for test examples. The hyperparameter k determines how many neighbors are included when predicting values for the testing data. The best value for k can be determined by parameter optimization. One common method for measuring the proximity or distance between neighbors is the Euclidean distance, calculated as shown in (4) where $d(x, x^*)$ as Euclidean distance, where x_i and x_i^* as the independent variable of the testing and training data points, respectively.

$$d(x, x^*) = \sqrt{\left(\sum_{i=1}^{n}(x_i - x_i^*)^2\right)} \tag{4}$$

2.4. Performance evaluation
2.4.1. Root mean squared error (RMSE)

The RMSE is employed to assess the error estimation of the performance generated by the RF, SVR, and KNN algorithms. In addition, these estimations are compared, the magnitude of errors produced in comparison to the actual data is calculated. A lower RMSE value shows higher accuracy in the model's estimations. RMSE measures the extent to which the predicted values from a model approximate the actual values, and a lower value indicates that the model predictions are more accurate. The RMSE can be expresses using (5) where y_i^* represents the actual value in the i^{th} data point, y_i represents the predicted value in the i^{th} data point, i is the variable of the i^{th} data point, and n is the total number of data samples.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i^* - y_i)^2} \quad (5)$$

2.4.2. Mean absolute error

MAE is an approach used to measure the accuracy level of a predictive model. It represents the average absolute error between the predicted values generated by the RF, SVR, and KNN and the actual values. A lower MAE indicates higher accuracy in the predictive model. MAE can be computed utilizing (6) y_i^* represents the actual value in the i^{th} data point, y_i represents the predicted value in the i^{th} data point, i is the variable of the i^{th} data point, and n is the total number of data samples.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i^* - y_i| \quad (6)$$

3. RESULTS AND DISCUSSION

RF, SVR, and KNN were employed to forecast the stock price trends of three IDX30 stocks: BBCA, PWON, and TOWR. Subsequently, the forecasted prices obtained from the three methods were compared with the actual data. Based on the results obtained from the simulation using the RStudio application, the first output comprises the closing stock price predictions generated on the testing data in Tables 1-3. Table 1 shows the predicted outcomes for BBCA stock, Table 2 for PWON stock, and Table 3 for TOWR stock.

Table 1. Predicted outcomes of RF, SVR, and KNN on BBCA stock

No	Date	Actual price	Forecasting price		
			RF	SVR	KNN
1	03/02/2020	32.200,00	32.679,65	32.325,43	31.940,00
2	04/02/2020	33.000,00	33.062,90	32.774,71	33.295,00
3	05/02/2020	33.650,00	33.358,59	33.443,98	33.690,00
⋮	⋮	⋮	⋮	⋮	⋮
20	28/02/2020	31.450,00	30.654,58	30.744,45	30.725,00

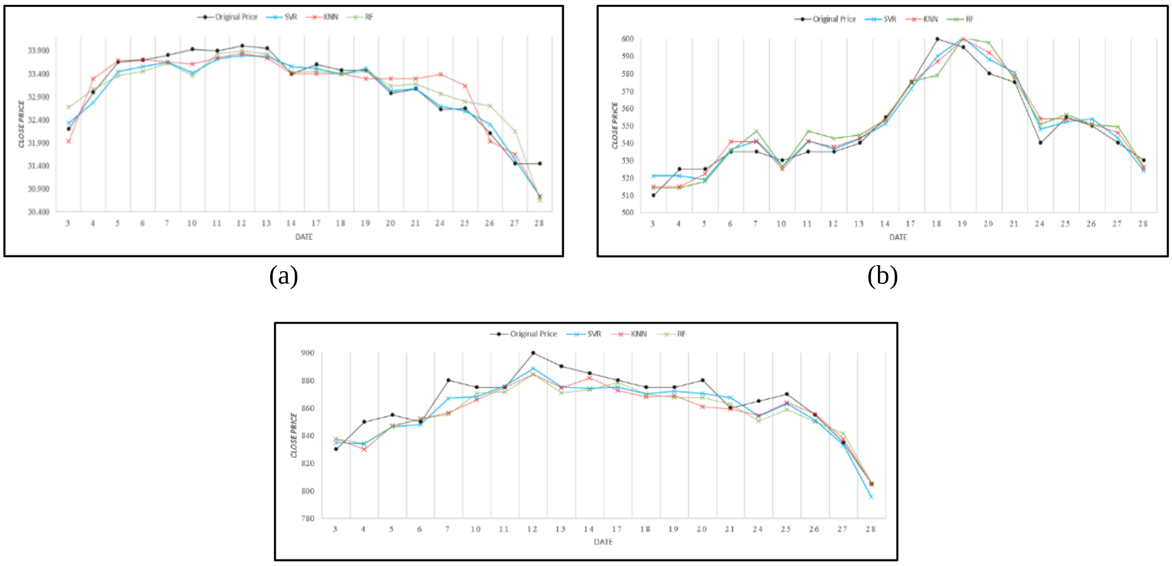
Table 2. Predicted outcomes of RF, SVR, and KNN on PWON stock

No	Date	Actual price	Forecasting price		
			RF	SVR	KNN
1	03/02/2020	510,0000	514,1742	521,1901	515,0000
2	04/02/2020	525,0000	514,1742	521,1901	515,0000
3	05/02/2020	525,0000	518,0602	519,0761	522,5000
⋮	⋮	⋮	⋮	⋮	⋮
20	28/02/2020	530,0000	526,5120	524,1113	525,7143

Table 3. Predicted outcomes of RF, SVR, and KNN on TOWR stock

No	Date	Actual price	Forecasting price		
			RF	SVR	KNN
1	03/02/2020	830,0000	837,6766	834,9473	837,6000
2	04/02/2020	850,0000	834,4189	833,9166	830,1429
3	05/02/2020	855,0000	846,8633	846,5214	847,4000
⋮	⋮	⋮	⋮	⋮	⋮
20	28/02/2020	805,0000	805,4791	795,6247	804,6667

The second result was a graphical representation of the stock price according to the testing data generated from the prediction results of the closing stock price. In Figure 3 the black line represented the trend formed based on actual data, the green line illustrated the trend based on the RF approach, the blue line indicated the trend according to SVR, and the red line denoted the trend based on the KNN. The stock price comparison of BBCA, PWON, and TOWR were sequentially depicted in Figures 3(a)-3(c). In evaluating the performance of the prediction models, RMSE and MAE were calculated. A model was more accurate if it had lower RMSE and MAE values compared to other models. Table 4 represented that the lowest RMSE values for BBCA, PWON, and TWOR were obtained when utilizing the SVR method. Similarly, the calculation of MAE values in Table 5, the lowest results were obtained when using the SVR approach.



(a) (b) (c)

Figure 3. Comparison of trend predictions: (a) BBCA, (b) PWON, and (c) TWOR stocks

Table 4. Comparison of RMSE values

Algorithm	Forecasting price		
	BBCA	PWON	TOWR
SVR	0.047935	0.106091	0.151406
RF	0.068012	0.120757	0.209977
KNN	0.069811	0.125458	0.238498

Table 5. Comparison of MAE values

Algorithm	Forecasting price		
	BBCA	PWON	TOWR
SVR	0.035155	0.084895	0.137808
RF	0.053465	0.101737	0.176353
KNN	0.052128	0.103811	0.175157

The comparison of RMSE values between the RF, SVR, and KNN methods was given in Figure 4. The SVR method generated RMSE values of 4.79% for BBCA stocks, 10.61% for PWON stocks, and 15.14% for TOWR stocks. Meanwhile, the RF approach produced an RMSE value of 6.80% for BBCA shares, 12.08% for PWON shares, and 21% for TOWR shares. Furthermore, the KNN approach generated RMSE values of 6.98% for BBCA stocks, 12.55% for PWON stocks, and 23.85% for TOWR stocks. These RMSE values provided a quantitative measurement of forecast accuracy for each method on each stock. RMSE represented the extent to which the predicted values deviated from the actual values. Lower RMSE values typically represented a closer match between predicted and actual values. In Figure 4, it can be noticed that the SVR algorithm provided the best forecast outcomes compared to RF and KNN from the RMSE value.

The comparison of MAE values among the RF, SVR, and KNN method was illustrated in Figure 5. The SVR method resulted MAE values of 3.52% for BBCA stock, 8.49% for PWON stock, and 13.78% for

TOWR stock. In contrast, the RF approach generated MAE values of 5.35% for BBCA stock, 10.17% for PWON stock, and 17.64% for TOWR stock. Additionally, the KNN method resulted in MAE values of 5.21% for BBCA stock, 10.38% for PWON stock, and 17.52% for TOWR stock. MAE was a metric that measures the average magnitude of errors between predicted and actual values. The lower MAE values demonstrated a better fit and higher accuracy in the predictions made by the respective methods. In Figure 5, it can be observed that the SVR algorithm yielded the best prediction outputs compared to RF and KNN from MAE values.

Figure 4. Comparison graph of RMSE values

Figure 5. Comparison graph of MAE values

The results indicate that the SVR model performed better in predicting stock prices than RF and KNN for all three of the evaluated equities, as seen by the smaller RMSE and MAE values. The better performance of the SVR model is consistent with previous research that indicates SVR can effectively capture intricate, non-linear relationships in financial data, especially in highly volatile markets [28]. The lower error rates associated with SVR can be attributed to its ability to model the data with an appropriate kernel function, optimizing the hyperparameters to suit the specific characteristics of the stock market being analyzed. This is particularly relevant in the context of the Indonesian stock market, where volatility and rapid changes in stock prices are common. The performance of RF, while strong, was slightly lower than SVR, likely due to its sensitivity to the depth and number of trees used in the ensemble, which may not have been optimized for the particular data patterns in this study. However, KNN's lower performance may be due to its reliance on local data points, which might not capture broader market trends as effectively as the other models. This study's finding was consistent with previous research that highlight the efficacy of SVR in stock price forecasting [29]. However, it's crucial to recognize that Indonesia has unique market circumstances, such greater volatility and less market maturity, which make these results particularly valuable. Studies conducted in more stable markets may find that model such as RF outperform SVR. Nevertheless, in the Indonesian context amplifies the benefits of non-linear, kernel-based approaches like SVR are more advantageous. This implies that even while SVR's superiority has been well studied, its use in emerging markets provides additional insights into its robustness under different economic conditions. This study limited the findings' generalizability by concentrating just on the Indonesian stock market, even if it offered insightful information. Future research should investigate the use of these models in other developing economies and throughout various time periods to validate the generalizability of these findings.

4. CONCLUSION

This research investigated the comparison of three approaches for forecasting stock price trends including RF, SVR, and KNN within the volatile Indonesian market. The analysis utilized daily historical stock data from three IDX30-listed companies: PT. Bank Central Asia Tbk (BBCA), PT. Pakuwon Jati Tbk (PWON), and PT. Sarana Menara Nusantara Tbk (TOWR). The evaluation metrics employed were RMSE and MAE. The findings revealed that the SVR method achieves the best performance, showed lower RMSE and MAE values than those of RF and KNN. These results suggest that SVR effective to capture complex and non-linear relationships. The study addresses a gap in the literature by focusing on an emerging market, offering new insights into the applicability of machine learning models beyond developed markets. The implications extend to investors and financial analysts, who may benefit from integrating non-linear, kernel-based approaches like SVR into their predictive strategies. Future studies should investigate similar models in more emerging markets, improve model parameters, and consider additional financial indicators to enhance predictive accuracy. This study contributes to the scholarly discourse on machine learning in financial forecasting and offers pragmatic instruments to enhance investment strategies in emerging markets.

REFERENCES

[1] M. Göçken, M. Özçalici, A. Boru, and A. T. Dosdoǧru, "Integrating metaheuristics and Artificial Neural Networks for improved stock price prediction," *Expert Systems with Applications*, vol. 44, pp. 320–331, Feb. 2016, doi: 10.1016/j.eswa.2015.09.029.

[2] A. K. Nassirtoussi, S. Aghabozorgi, T. Ying Wah, and D. C. L. Ngo, "Text mining for market prediction: A systematic review," *Expert Systems with Applications*, vol. 41, no. 16, pp. 7653–7670, Nov. 2014, doi: 10.1016/j.eswa.2014.06.009.

[3] A. A. Rodrigues and S. Lleo, "Combining standard and behavioral portfolio theories: a practical and intuitive approach," *Quantitative Finance*, vol. 18, no. 5, pp. 707–717, Dec. 2018, doi: 10.1080/14697688.2017.1401225.

[4] Z. Q. Jiang *et al.*, "Short term prediction of extreme returns based on the recurrence interval analysis," *Quantitative Finance*, vol. 18, no. 3, pp. 353–370, Oct. 2018, doi: 10.1080/14697688.2017.1373843.

[5] T. Fischer and C. Krauss, "Deep learning with long short-term memory networks for financial market predictions," *European Journal of Operational Research*, vol. 270, no. 2, pp. 654–669, Oct. 2018, doi: 10.1016/j.ejor.2017.11.054.

[6] Y. S. Abu-Mostafa and A. F. Atiya, "Introduction to financial forecasting," *Applied Intelligence*, vol. 6, no. 3, pp. 205–213, Jul. 1996, doi: 10.1007/BF00126626.

[7] S. Asadi, E. Hadavandi, F. Mehmanpazir, and M. M. Nakhostin, "Hybridization of evolutionary Levenberg-Marquardt neural networks and data pre-processing for stock market prediction," *Knowledge-Based Systems*, vol. 35, pp. 245–258, Nov. 2012, doi: 10.1016/j.knosys.2012.05.003.

[8] S. Akhter and M. A. Misir, "Capital markets efficiency: Evidence from the emerging capital market with particular reference to Dhaka stock exchange," *South Asian Journal of Management*, vol. 12, no. 3, pp. 35–51, 2005.

[9] K. Miao, F. Chen, and Z. Zhao, "Stock price forecast based on bacterial colony RBF neural network," *Journal of Qingdao University*, vol. 2, pp. 210–230, 2007.

[10] J. Lehoczky and M. Schervish, "Overview and history of statistics for equity markets," *Annual Review of Statistics and Its Application*, vol. 5, no. 1, pp. 265–288, Mar. 2018, doi: 10.1146/annurev-statistics-031017-100518.

[11] A. Aali-Bujari, F. Venegas-Martínez, and G. Pérez-Lechuga, "Impacto de la capitalización bursátil y del spread bancario sobre el crecimiento y el desarrollo en América Latina: una estimación de datos de panel con MGM-sistema," *Contaduria y Administracion*, vol. 62, no. 5, pp. 1427–1441, Dec. 2017, doi: 10.1016/j.cya.2017.09.005.

[12] M. P. Naeini, H. Taremian, and H. B. Hashemi, "Stock market value prediction using neural networks," in *2010 International Conference on Computer Information Systems and Industrial Management Applications, CISIM 2010*, Oct. 2010, pp. 132–136, doi: 10.1109/CISIM.2010.5643675.

[13] B. Qian and K. Rasheed, "Stock market prediction with multiple classifiers," *Applied Intelligence*, vol. 26, no. 1, pp. 25–33, Nov. 2007, doi: 10.1007/s10489-006-0001-7.

[14] D. Shah, H. Isah, and F. Zulkernine, "Stock market analysis: A review and taxonomy of prediction techniques," *International Journal of Financial Studies*, vol. 7, no. 2, p. 26, May 2019, doi: 10.3390/ijfs7020026.

[15] E. S. Olivas, "Handbook of research on machine learning applications and trends: Algorithms, methods, and techniques," Hershey, PA, USA: IGI Global, 2009.

[16] R. E. McRoberts, "Estimating forest attribute parameters for small areas using nearest neighbors techniques," *Forest Ecology and Management*, vol. 272, pp. 3–12, May 2012, doi: 10.1016/j.foreco.2011.06.039.

[17] S. A. A. Kharis, G. F. Hertono, S. R. Irawan, E. Wahyuningrum, and Yumiati, "Students' success prediction based on the fuzzy K-nearest neighbor method in Universitas Terbuka," in *Education Technology in the New Normal: Now and Beyond*, Routledge, 2023, pp. 212–218.

[18] S. A. A. Kharis and A. H. A. Zili, "Predicting life expectancy of lung cancer patients after thoracic surgery using SMOTE and machine learning approaches," *Jurnal Natural*, vol. 23, no. 3, pp. 152–161, Oct. 2023, doi: 10.24815/jn.v23i3.29144.

[19] R. E. McRoberts, "Diagnostic tools for nearest neighbors techniques when used with satellite imagery," *Remote Sensing of Environment*, vol. 113, no. 3, pp. 489–499, Mar. 2009, doi: 10.1016/j.rse.2008.06.015.

[20] S. Magnussen, E. Tomppo, and R. E. McRoberts, "A model-assisted k-nearest neighbour approach to remove extrapolation bias," *Scandinavian Journal of Forest Research*, vol. 25, no. 2, pp. 174–184, Mar. 2010, doi: 10.1080/02827581003667348.

[21] J. Patel, S. Shah, P. Thakkar, and K. Kotecha, "Predicting stock market index using fusion of machine learning techniques," *Expert Systems with Applications*, vol. 42, no. 4, pp. 2162–2172, Mar. 2015, doi: 10.1016/j.eswa.2014.10.031.

[22] M. Kumar and M. Thenmozhi, "Forecasting stock index movement: a comparison of support vector machines and random forest," *SSRN Electronic Journal*, 2011, doi: 10.2139/ssrn.876544.

[23] B. M. Henrique, V. A. Sobreiro, and H. Kimura, "Stock price prediction using support vector regression on daily and up to the minute prices," *Journal of Finance and Data Science*, vol. 4, no. 3, pp. 183–201, Sep. 2018, doi: 10.1016/j.jfds.2018.04.003.

[24] J. Zheng, D. Xin, Q. Cheng, M. Tian, and L. Yang, "The random forest model for analyzing and forecasting the US stock market under the background of smart finance," in *Proceedings of the 3rd International Academic Conference on Blockchain, Information Technology and Smart Finance (ICBIS 2024)*, Atlantis Press International BV, 2024, pp. 82–90.

[25] V. Sarala and G. N. V. P. Bhushan, "Stock market trend prediction using K-nearest neighbor (KNN) algorithm," *Journal of Engineering Sciences*, vol. 13, no. 8, 2022, [Online]. Available: www.jespublication.com.

[26] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," *Statistics and Computing*, vol. 14, no. 3, pp. 199–222, Aug. 2004, doi: 10.1023/B:STCO.0000035301.49549.88.

[27] Z. Rustam and S. A. A. Kharis, "Multiclass classification on brain cancer with multiple support vector machine and feature selection based on kernel function," in *AIP Conference Proceedings*, 2018, vol. 2023, p. 20233, doi: 10.1063/1.5064230.

[28] X. Yan, "Research on financial field integrating artificial intelligence: application basis, case analysis, and SVR model-based overnight," *Applied Artificial Intelligence*, vol. 37, no. 1, Jun. 2023, doi: 10.1080/08839514.2023.2222258.

[29] A. Ampountolas, "Enhancing forecasting accuracy in commodity and financial markets: insights from GARCH and SVR models," *International Journal of Financial Studies*, vol. 12, no. 3, p. 59, Jun. 2024, doi: 10.3390/ijfs12030059.

BIOGRAPHIES OF AUTHORS

Selly Anastassia Amellia Kharis is Assistant Professor at Universitas Terbuka, Indonesia. Selly's academic career is not only reflected in her academic achievements, but also in her contribution to research and development in the fields of Mathematics and Mathematics Education. As a dedicated academic, she has published numerous scientific articles and is a respected contributor in the scientific community. Behind her academic achievements, Selly is also known as a good mentor for students and colleagues. She often provides encouragement and guidance to those who are trying to achieve success in academics. With her passion for knowledge and dedication to science, Selly continues to move forward in her academic career. Her life is proof that with persistence and hard work, any academic dream and purposes can be realized. She can be contacted at email: selly@ecampus.ut.ac.id.

Arman Haqqi Anna Zili received his bachelor's and master's degrees in Mathematics from the University of Indonesia. Currently he serves as an Assistant Professor at the Department of Mathematics, University of Indonesia (UI). He has experience testing and guiding the final assignment writing of more than 30 students. His academic career is not only reflected in his academic achievements, but also in his contribution to research and development in the fields of Applied Mathematics and Actuarial Science. As a dedicated academic, he has published numerous scientific articles and is a respected contributor in the scientific community. He has authored or co-authored more than 15 publications. His research interests include applied mathematics, actuarial science, machine learning, and artificial intelligence. He can be contacted at email: armanhaz@sci.ui.ac.id.

Maulana Malik received the B.Sc. and M.Sc. degrees in Mathematics from Universitas Indonesia (UI), Indonesia, and the Ph.D. degree in Mathematics from Universiti Sultan Zainal Abidin (UniSZA), Malaysia. He is currently an Assistant Professor with the Department of Mathematics Universitas Indonesia. He has authored and coauthored a number of resarch articles in high impact journals. His current research interest iterative algorithms for solving nonlinear optimization, such as conjugate gradient method with application in portfolio selection, robotic motion control, image restoration, signal recovery, and carbon emission. He can be contacted at email: m.malik@sci.ui.ac.id.

Wahyu Nuryaningrum she is a student at Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Indonesia, Depok, Indonesia. She can be contacted at email: wahyu.nuryaningrum179@gmail.com.

Agustiani Putri is a Ph.D. candidate at Universitas Negeri Malang, Indonesia. She always concentrated on mathematics education during her higher education. She earned B.Ed. from Universitas Negeri Jakarta, Indonesia, and M.Ed. from Universitas Negeri Malang. She is currently on a research fellowship at Deakin University, Melbourne, Australia. She also has been involved in mathematics and education research both quantitatively and qualitatively. She has published many works both in reputable international journals indexed by Scopus and national journals. She can be contacted at email: agustianiputri15@gmail.com.