PILAR Nusa Mandiri: Journal of Computing and Information System
Vol. 21, No. 1 March 2025 | DOI: 10.33480/pilar.v21i1.6072

PERFORMANCE COMPARISON OF RANDOM FOREST REGRESSION, SVR MODELS IN STOCK PRICE PREDICTION

Maysas Yafi Urrochman[1*]; Hasyim Asy'ari[2]; Fadhel Akhmad Hizham[3]

Department of Engineering and Informatics[1, 2, 3]
Institut Teknologi Dan Bisnis Widya Gama Lumajang, Indonesia[1, 2, 3]
https://itbwigalumajang.ac.id/[1, 2, 3]
maysasyafi@gmail.com[1*], hasyimasyarimail@gmail.com[2], hizhamfadhel@gmail.com[3]
(*) Corresponding Author

Abstract— *The stock market is characterized by high volatility and complexity, making it an intriguing and challenging subject for researchers and practitioners. This study aims to predict stock prices by comparing the performance of two machine learning models: Random Forest Regression and Support Vector Regression (SVR). These models were selected for their ability to handle complex data and high volatility. The dataset used in this study consists of BNI stock data over the last five years (2019–2024), comprising a total of 1,211 data points. Testing was conducted using a cross-validation approach, and model performance was evaluated based on several metrics, including MSE, R^2, RMSE, MAPE, MAE, and Score. The results indicate that Random Forest Regression outperforms SVR. The model achieved an MAE of 17.766, an RMSE of 22.376, and an R^2 of 0.997. These findings suggest that Random Forest Regression is more effective in predicting stock prices, particularly in unstable market conditions. This study recommends Random Forest Regression as a reliable model for stock price prediction, with potential applications in other stock markets with similar characteristics.*

Keywords: *machine learning, random forest regression, stock price prediction, support vector regression.*

Abstrak— *Pasar saham memiliki volatilitas tinggi dan kompleksitas yang sulit diprediksi, menjadikannya subjek penelitian yang menarik dan menantang bagi para peneliti serta praktisi. Penelitian ini bertujuan untuk meramalkan harga saham dengan membandingkan kinerja dua model machine learning, yaitu Random Forest Regression dan Support Vector Regression (SVR). Kedua model ini dipilih karena kemampuannya dalam menangani data yang kompleks dan volatilitas tinggi. Data yang digunakan dalam penelitian ini adalah data saham BNI selama lima tahun terakhir (2019–2024), dengan total 1.211 data. Pengujian dilakukan menggunakan pendekatan cross-validation, dan kinerja model diukur berdasarkan beberapa metrik, yaitu MSE, R^2, RMSE, MAPE, MAE, dan Score. Hasil penelitian menunjukkan bahwa Random Forest Regression memiliki performa yang lebih unggul dibandingkan SVR. Model ini mencatat nilai MAE sebesar 17.766, RMSE sebesar 22.376, dan R^2 sebesar 0.997. Temuan ini mengindikasikan bahwa Random Forest Regression lebih efektif dalam memprediksi harga saham, terutama pada kondisi pasar yang tidak stabil. Penelitian ini merekomendasikan penggunaan Random Forest Regression sebagai model yang andal untuk prediksi harga saham, dengan potensi aplikasi lebih lanjut pada pasar saham lainnya yang memiliki karakteristik serupa.*

Kata Kunci*: machine learning, random forest regression*, prediksi harga saham, *support vector regression*,.

INTRODUCTION

In today's digital age, the stock market has become one of the most attractive and complex investment arenas (Patriya, 2020). Rapid changes in stock prices often make stock price prediction a challenge. The ability to accurately predict stock prices can give investors and portfolio managers a competitive advantage, as well as aid in better decision-making (Saputra & Koesrindartoto, 2024). Along with advances in technology and artificial intelligence, machine learning models are increasingly being used to predict stock price movements (D. Kumar, Sarangi, & Verma, 2022). Machine learning models offer a more sophisticated approach than traditional methods such as technical or fundamental analysis (Bansal, Goyal, & Choudhary, 2022).

Among the various machine learning models available, Random Forest Regression, Support

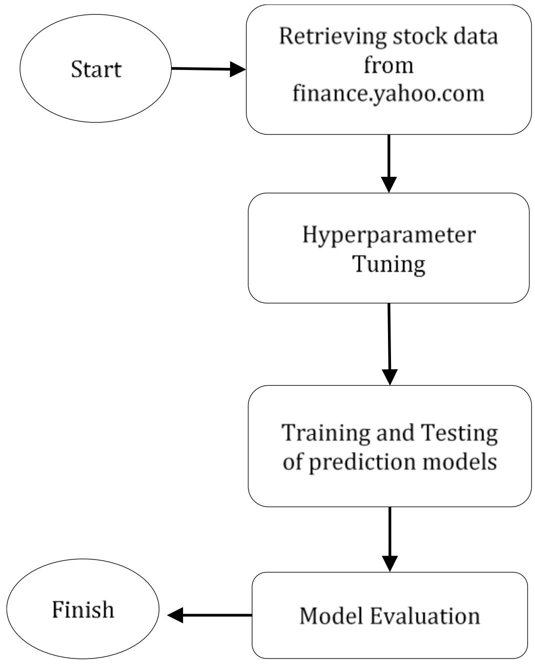
Vector Regression (SVR) are two frequently used methods in stock price prediction (Pashankar, Shendage, & Pawar, 2024). Random Forest Regression and SVR are often used in the prediction of stock price movements due to their respective advantages in handling complex data and non-linear patterns that are common in stock data. Random Forest Regression is an ensemble method that uses multiple decision trees to produce more stable and accurate predictions (Syafira & Saepudin, 2023). SVR, on the other hand, is a variant of Support Vector Machines (SVM) specifically designed for regression problems with the ability to handle non-linear data (Kurani, Doshi, Vakharia, & Shah, 2023).

In BNI shares, several interesting issues arise, namely in 2023 and 2024. In 2023, BNI shares received a positive recommendation from the analysis, because the projected net profit growth reached 19.2% driven by a selective increase in corporate credit in several sectors. But in the first quarter of 2024, BNI shares experienced a significant decline of up to 8%. This made investors sell BNI shares significantly. Therefore, researchers use BNI shares as the object of this research. Based on previous research, the results obtained are quite satisfying, but no one has specifically discussed stock predictions on BNI. In research conducted by (A. Kumar & Chaudhry, 2021) regarding the review and analysis of stock price predictions using data mining, shows the results that the use of linear regression gets an accuracy rate above 80%. Meanwhile, the research conducted by (Panggabean & Widyasari, 2023) on the prediction of stock price movements showed that random forest managed to get a high accuracy of 94%. In research conducted by (Raya, Srinivasan, Vishnu, Adedoyin, & Sathiyanarayanan, 2022) stated that random forest is the best algorithm based on the MAE (mean absolute error) measurement compared to the SVR, ARIMA and XGBoost algorithms.

In this research, we focus on the comparison of machine learning Random Forest Regression and SVR to predict the price movement of the same BNI. The data uses historical data obtained from yahoo finance. By comparing the two methods, researchers hope to provide a more accurate and effective solution in predicting the price movement of the same BNI.

MATERIALS AND METHODS

The research process begins with retrieving the data to be used, in this study it is historical data on BNI stock prices. Furthermore, the data is pre-processed, where it is split into two sets: training data and test data. The training data is used to train the model, while the test data is used to assess the model's performance. The flow of this research process can be seen in Figure 1.



Source: (Research Result, 2024)
Figure 1. Research Process Flow

Data Collection

The stock dataset used comes from https://finance.yahoo.com/. Covers BNI stock data for 5 years. The collected stock dataset consists of several columns, namely Date, Open, High, Low, Close, Adj Close, and Volume. The Date column contains date data, while Open reflects the stock price when the first transaction is made on that day. High and Low reflect the daily price fluctuations that enable rational action in buying or selling shares. Close indicates the share price when all trading activity on the stock exchange has been completed. Adj Close reflects the closing price of a stock that takes into account dividend and stock split factors. Volume indicates the total number of shares traded in a given period. In this study, attention is focused on the daily closing price (Close) as input data to train and test the model.

Table 1. 5-Year BNI Stock Data Sample

Date	Open	High	Low	Close	Adj. Close	Volume
2019-09-05	3750	3787	3725	3762	3255	3149
2019-09-06	3800	3825	3775	3800	3288	3769
2019-09-09	3800	3812	3750	3775	3266	2534
...	...	...	...	...	...	...
2024-09-02	5350	5400	5325	5350	5350	2298

Source: (Research Result, 2024)

PILAR Nusa Mandiri: Journal of Computing and Information System
Vol. 21, No. 1 March 2025 | DOI: 10.33480/pilar.v21i1.6072

Table 1 shows that price movements of a financial instrument, such as shares, can be analyzed through historical data which includes information on opening, closing, highest, lowest prices and trading volume. The following is a brief review based on available data.

In the data period starting from September 5 2019 to September 2 2024, the price of this instrument shows an interesting fluctuation pattern. The opening price on September 5 2019 was recorded at 3,750, with the highest price reaching 3,787 and the lowest price 3,725, before finally closing at level 3,762. Trading volume that day reached 3,149 units.

On the following day, namely September 6 2019, the opening price increased to 3,800, with the highest price being 3,825 and closing at the same level, namely 3,800. Trading volume increased to 3,769 units, indicating more intense market activity compared to the previous day.

This upward trend continued until it reached a more recent period. The last data recorded on September 2, 2024 shows that the opening price reached 5,350, while the highest price on that day was 5,400, and the closing was at the same level as the opening price, namely 5,350. Even though trading volume was recorded to be lower compared to previous days, namely 2,298 units, the price was stable at a high level indicating the potential for interesting movements for further analysis.

Data Pre-Processing

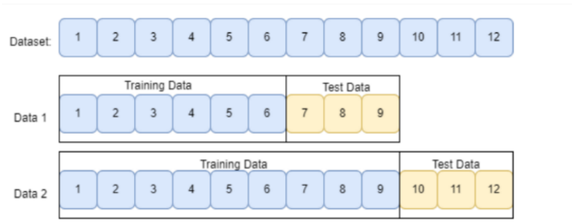
The data pre-processing stage is the stage used to prepare the dataset for use in the next stage. In stock data, there are corporate actions that are often carried out, namely stock splits. Therefore, it is necessary to adjust the stock data according to the ratio that applies to the stock split. The pre-processing process is needed to calculate the stock split ratio and adjust the stock price based on the predetermined ratio.

The normalization technique used in this research is the Min-Max method. Normalization Min-Max is one approach that aims to convert the values in a dataset to fall within a specific range, such as [0, 1]. This technique works by converting each value in the dataset into a predefined range, making it easier to analyze and apply the model. Here's the equation for the Min-Max normalization technique (Wahanani, Swari, & Akbar, 2020).

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad \text{..}(1)$$

Xnorm is the normalized value of the original value (X), where Xmin is the smallest value and Xmax is the largest value in the dataset. By using Min-Max normalization, all values in the dataset will be converted to be in the range [0, 1]. The smallest value will be converted to 0, while the largest value will be converted to 1. Other values that fall between Xmin and Xmax will be adjusted proportionally so that they fall within the range of 0 to 1, according to their position relative to Xmin and Xmax.

Cross-validation on time series data is a method used to evaluate the performance of predictive models on time series data (Bastian, Rahayudi, & Ratnawati, 2021). In contrast to cross-validation on non-time series data, cross-validation on time series data must consider the time sequence in data splitting and evaluation (Prasetya, Priyatno, & Nuraeni, 2023). This is due to the interrelationship between neighboring observations in the time series, which often have certain patterns or trends. Therefore, the process of cross-validation on time series data has slightly different steps than non-time series data. In this study, the time series cross-validation method is applied to both training and test data. Figure 2 shows an illustration of the application of cross-validation on time series data.



Source: (Research Result, 2024)
Figure 2. Concept of Cross-validation of Time Series Data

Prediction Model Random Forest Regression

Prediction of stock price movements in this study is to compare machine learning methods Random Forest Regression and SVR (Priyatno et al., 2023). Random forest Regression is one of the machine learning algorithms used for regression modelling (Yin, Li, Li, & Zhang, 2023). The Random Forest Regression algorithm is an ensemble learning technique that integrates multiple decision tree regression models to enhance accuracy and minimize overfitting. This algorithm is based on the concept of Random Forest which is an ensemble of many decision trees (Zheng, Xin, Cheng, Tian, & Yang, 2024). Algorithm Stages Random Forest Regression (Fitri & Riana, 2022).

1. Bootstrap Sampling: From the original training data, randomly generate several subsets of data (bootstrap samples) with replacement. Each subset will be used to train one decision tree.
2. Development Decision Trees: Create a decision tree for each data subset. At every node of the tree, randomly choose a subset of features

(rather than using all features) to find the optimal split.
3. Prediction of Each Tree: Once all the trees are built, each tree will make independent predictions on the test data or new data.
4. Combining the results of (Ensemble): averages the predictions from all trees to produce the final prediction of the Random Forest Regression model.
5. Model Evaluation: Use Utilize metrics like Mean Squared Error (MSE) or R-squared (R^2) to assess the model's performance on the test data.
6. Hyper parameter Tuning (Optional): Adjust parameters such as the number of trees (n_estimators), the maximum features considered (max_features), and the maximum tree depth (max_depth) to enhance the model's performance.
7. Model Usage: Use the trained model to make predictions on new data or in production applications.

Prediction Model Support Vector Regression (SVR)

The prediction model uses the SVR algorithm in its learning method to recognize patterns in the training data (Putra & Kurniawati, 2021). The optimal hyperparameters in SVR obtained from the hyperparameter tuning stage are used to assist the training of the prediction model resulting in precise model performance (Thumu, Nellore, & others, 2024). The researcher created a prediction model that aims to predict the closing price of stocks (Close) against the Composite Stock Price Index (JCI) denoted as (Yt),using 3 independent variables namely the opening price (Open), the highest (High) and the lowest (Low). In training the SVR model, researchers use training data and are assisted by optimal hyperparameters to produce a model that has the ability to predict the closing price of shares (Close). After that, testing the SVR model by entering the testing data as the period tested, so that the model can produce the expected predictions. In SVR, the regression function is expressed in the form of the following equation:

$$(x, w) = w^T(x) + b \text{ ....(2)}$$

Where (x) is a point in the feature space F resulting from mapping x in the input space. SVR also looks for the flattest possible regression equation by finding the minimum possible w value. One way is with Euclidian $||w||2$.

Evaluation Method

Evaluation methods are techniques or approaches used to assess or measure the effectiveness, efficiency, quality or impact of a particular programme, project, process or activity (FADLI & Saputra, 2023). In the context of stock price prediction, model valuation is the result of predicting the stock price index compared to the actual data. The model performance evaluation used is MSE, RMSE, MAPE, MAE. Normalization is done in order to see whether the evaluation results are high or low because each indicator has a different range of values.

1. Mean Squared Error (MSE)

MSE is to calculate the average squared error in prediction, the smaller the value of MSE, the better the quality of the model. Formula MSE:

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(\text{yi} - \hat{y}i)^2 \text{ .... (3)}$$

2. Root Mean Squared Error (RMSE)

RMSE is a derivative of MSE, which calculates the average of the squared difference between predicted and actual values and then takes the square root. The smaller the RMSE value, the better the quality of the model. Formula RMSE:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\text{yi} - \hat{y}i)^2} \text{ ....(4)}$$

3. Mean Absolute Percentage Error (MAPE)

MAPE is to calculate the average error in prediction as a percentage of the actual value. The smaller the MAPE value, the better the quality of the model. Formula *MAPE*:

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\text{yi}-\hat{y}i}{\text{yi}}\right| \times 100\% \text{ .... (5)}$$

4. Mean Absolute Error (MAE)

MAE is to calculate the average absolute error in prediction, the smaller the MAE value, the better the quality of the model. Formula *MAE*:

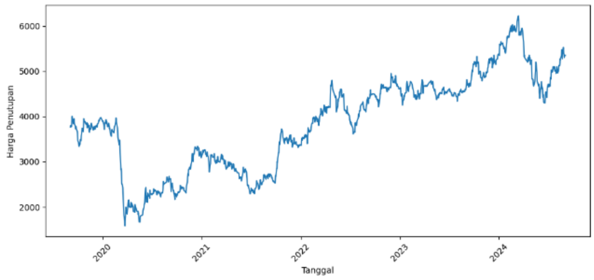
$$MAE = \frac{1}{n}\sum_{i=1}^{n}|\text{yi} - \hat{y}i| \text{ ....(6)}$$

In Formula 3, 4, 5 and 6, n represents the number of samples in the data, y_i denotes the actual value and ŷ_i refers to the predicted value.

RESULTS AND DISCUSSION

The initial stage before implementing the prediction model is by collecting data on BNI Shares obtained from https://finance.yahoo.com/. The results obtained were 1211 data. After the data is obtained, it is uploaded to the google colab platform. The graphical display of BNI stock movements over the past 5 years can be seen in Figure 3.

PILAR Nusa Mandiri: Journal of Computing and Information System
Vol. 21, No. 1 March 2025 | DOI: 10.33480/pilar.v21i1.6072



Source: (Research Result, 2024)

Figure 3. BNI Stock Movement Chart for the Last 5 Years

This graph depicts high volatility, especially in 2020 due to the pandemic, followed by a gradual recovery trend until 2024. Price corrections occurred at some points, but the overall trend shows a strong recovery and increase in BNI's share price over the past five years.

Furthermore, the dataset is then split into two for 80% training data while for 20% testing data. Next, a parameter grid is created containing hyperparameter values for prediction models using Random Forest Regression and SVR. The following parameter table is used in both prediction models. Table 2 shows the Prediction Model Parameters Random Forest Regression and SVR.

Table 2. Prediction Model Parameters

No	Model Prediction	Parameter
1	Random Forest Regression	n_estimators': [100, 200, 300] max_depth': [10, 20, 30, None] min_samples_split': [2, 5, 10] min_samples_leaf': [1, 2, 4] max_features': ['auto', 'sqrt', 'log2'] bootstrap': [True, False] grid_search=GridSearchCV(estimator=rf, param_grid=param_grid, cv=5, n_jobs=-1, verbose=2)
2	SVR	param_grid = {'C': [1, 10, 100, 1000] 'gamma': [0.001, 0.01, 0.1, 1] 'epsilon': [0.1, 0.2, 0.5] grid_search = GridSearchCV(SVR(kernel='rbf') param_grid, cv=5 scoring='neg_mean_squared_error' grid_search.fit(x_train, y_train)

Source: (Research Result, 2024)

Random Forest Regression

The Random Forest Regression model uses an ensemble-based approach that combines several decision trees to produce reliable predictions. For optimization, various parameters are tested with the following values:

1. n_estimators: Refers to the number of trees in the model, evaluated using values [100, 200, 300].
2. max_depth: Specifies the maximum allowable depth of the tree, tested with values [10, 20, 30, None].
3. min_samples_split: Defines the minimum number of samples required to split an internal node, tested with values [2, 5,10].
4. min_samples_leaf: Indicates the minimum number of samples allowed in a leaf node, evaluated with values [1, 2,4].
5. max_features: Determines the maximum number of features considered for splitting, with options ['auto,' 'sqrt,' 'log2'].
6. bootstrap: Specifies the sampling strategy, tested with the values [True, False].

Support Vector Regression (SVR)

The Support Vector Regression model uses an RBF (Radial Basis Function) kernel to capture non-linear relationships between variables. The parameters tested in this optimization include:

1. C: Regulatory parameters, tested at values [1, 10, 100, 1000].
2. Gamma: Parameter for RBF kernel, tested with values [0.001, 0.01, 0.1, 1].
3. Epsilon: Limit deviation for determining the margin of error, with values [0.1, 0.2, and 0.5].

The optimization process was also carried out with Grid Search CV using cross-validation of 5 folds and the negative mean squared error (neg mean squared error) evaluation metric. Training data (x_train, y_train) is used to train the model in each iteration.

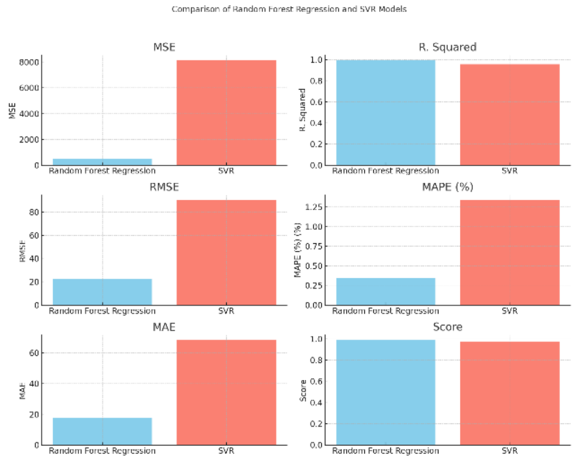
After the prediction model is made using the parameters that have been determined, the next step is the model evaluation stage. In this study using model evaluation metrics MSE, RMSE, MAPE, MAE and Score.

Table 3. Comparison of Prediction Model Evaluation Results

No	Model	MSE	R. Squared	RMSE	MAPE	MAE	Score
1	Random Forest Regression	500.700	0.997	22.376	0.34%	17.766	0.99
2	SVR	8166.247	0.958	90.367	1.34%	68.426	0.97

Source: (Research Result, 2024)

Table 3 presents the performance evaluation results, demonstrating that Random Forest Regression is more effective than the SVR prediction model. The table provides a comparison of the evaluation metrics for both Random Forest Regression and SVR. By analyzing these results, we can compare the performance of the two models across several evaluation metrics.

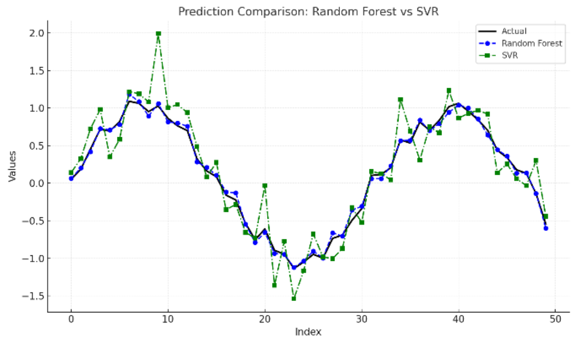


Source: (Research Result, 2024)
Figure 4. Graph of the Results of the Comparison Between Random Forest and SVR

Figure 4 explains each metric and a comparison of the two models:

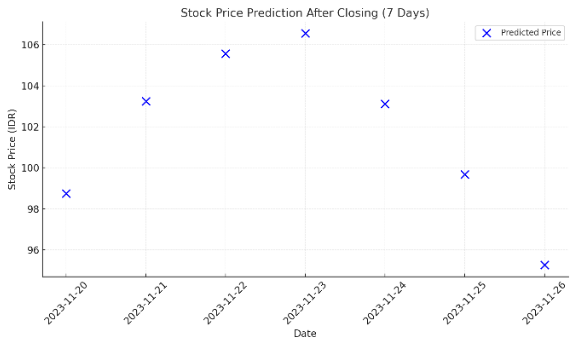
1. Mean Squared Error (MSE): Random Forest Regression has an MSE of 500,700, while SVR has a much larger MSE of 8166,247. This shows that Random Forest Regression has a much smaller error compared to SVR.
2. R-squared (R^2): Random Forest Regression has a R^2 value of 0.997, which indicates that almost 99.7% of the variation in the data can be explained by this model. SVR has a R^2 value of 0.958, which is also quite good, but lower compared to Random Forest Regression.
3. Root Mean Squared Error (RMSE): Random Forest Regression has a RMSE of 22.376, while SVR has a RMSE of 90.367. Random Forest shows much better performance as its prediction error is smaller.
4. Mean Absolute Percentage Error (MAPE): Random Forest Regression has a MAPE of 0.34%, which indicates that this model only has about 0.34% error in its prediction. Meanwhile, SVR has a MAPE of 1.34%, which is still quite low but greater than Random Forest Regression.
5. Mean Absolute Error (MAE): Random Forest Regression has a MAE of 17.766, while SVR has a MAE of 68.426. This means that Random Forest has a smaller average error than SVR.
6. Score: Random Forest Regression has a score of 0.99, which is almost perfect, indicating that the model is very reliable. SVR has a score of 0.97, which is also good but lower than Random Forest Regression.

The following graph shows the comparison between actual data and predicted data.



Source: (Research Result, 2024)
Figure 5. Comparison Chart of Predicted Stock Price Movements With Actual Data

Figure 5 shows a comparison between the actual values and the predicted results of the Random Forest (RF) and Support Vector Regression (SVR) models. The Random Forest model has higher accuracy than SVR, as seen from the prediction line which is closer to the actual value. This shows that RF is more effective in predicting the data in this case. The following graph shows the prediction of the Share Price after the 7-day close.



Source: (Research Result, 2024)
Figure 6. Stock Price Prediction Chart After 7-Day Close

Figure 6 visualises illustrates the predicted stock prices for the 7 days following market close. Each dot represents the projected price on a specific date, allowing for an easy visual comparison of predictions over time. The prices are displayed in Indonesian Rupiah (IDR).

CONCLUSION

From this evaluation, it can be concluded that Random Forest Regression outperforms SVR in all evaluation metrics used. Random Forest has significantly smaller error values (MSE, RMSE, MAE) and higher measures of variation in the data (R^2 and Score). This suggests that Random Forest

PILAR Nusa Mandiri: Journal of Computing and Information System
Vol. 21, No. 1 March 2025 | DOI: 10.33480/pilar.v21i1.6072

Regression is a more accurate and more reliable model to use on this data. On the other hand, Random Forest Regression requires high computation time, especially when applied to large and dynamic datasets such as stock prices. While both models rely on historical data for prediction, stock price movements are often influenced by external factors that cannot be predicted by past data such as government policy changes, market sentiment or global events. Taking into account these limitations, future research is expected to use other models such as the Deep Learning method, as well as consider external data such as social media sentiment or financial news that can improve the accuracy of predicting stock price movements.

REFERENCE

Bansal, M., Goyal, A., & Choudhary, A. (2022). Stock market prediction with high accuracy using machine learning techniques. *Procedia Computer Science*, *215*, 247–265. https://doi.org/10.1016/j.procs.2022.12.028

Bastian, M. E., Rahayudi, B., & Ratnawati, D. E. (2021). Prediksi trend harga saham jangka pendek berdasarkan fitur technical analysis dengan menggunakan algoritma random forest. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, *5*(10), 4536–4542.

FADLI, M., & Saputra, R. A. (2023). Klasifikasi Dan Evaluasi Performa Model Random Forest Untuk Prediksi Stroke. *Jurnal Teknik*, *12*(2). http://dx.doi.org/10.31000/jt.v12i2.9099

Fitri, E., & Riana, D. (2022). Analisa Perbandingan Model Prediction Dalam Prediksi Harga Saham Menggunakan Metode Linear Regression, Random Forest Regression Dan Multilayer Perceptron. *METHOMIKA: Jurnal Manajemen Informatika & Komputerisasi Akuntansi*, *6*(1), 69–78. https://doi.org/10.46880/jmika.Vol6No1.pp69-78

Kumar, A., & Chaudhry, M. (2021). Review and Analysis of Stock Market Data Prediction Using Data mining Techniques. *2021 5th International Conference on Information Systems and Computer Networks (ISCON)*, 1–10. https://doi.org/10.1109/iscon52037.2021.9702498

Kumar, D., Sarangi, P. K., & Verma, R. (2022). A systematic review of stock market prediction using machine learning and statistical techniques. *Materials Today: Proceedings*, *49*, 3187–3191. https://doi.org/10.1016/j.matpr.2020.11.399

Kurani, A., Doshi, P., Vakharia, A., & Shah, M. (2023). A comprehensive comparative study of artificial neural network (ANN) and support vector machines (SVM) on stock forecasting. *Annals of Data Science*, *10*(1), 183–208. https://doi.org/10.1007/s40745-021-00344-x

Panggabean, R., & Widyasari, Y. D. L. (2023). A comparison between super vector regression, random forest regressor, lstm, and gru in forecasting bitcoin price. *International ABEC*, 281–287.

Pashankar, S. S., Shendage, J. D., & Pawar, J. (2024). Machine Learning Techniques For Stock Price Prediction-A Comparative Analysis Of Linear Regression, Random Forest, And Support Vector Regression. *Journal of Advanced Zoology*, *45*(S-4), 118–127. https://doi.org/10.53555/jaz.v45is4.4164.

Patriya, E. (2020). Implementasi support vector machine pada prediksi harga saham gabungan (ihsg). *Jurnal Ilmiah Teknologi Dan Rekayasa*, *25*(1), 24–38. http://dx.doi.org/10.35760/tr.2020.25i1.2571

Prasetya, M. R. A., Priyatno, A. M., & Nurhaeni. (2023). Penanganan Imputasi Missing Values pada Data Time Series dengan Menggunakan Metode Data Mining. *Jurnal Informasi Dan Teknologi,* *5*(2), 52–62. https://doi.org/10.37034/jidt.v5i2.324.

Priyatno, A. M., Tanjung, L. S., Ramadhan, W. F., Cholidhazia, P., Jati, P. Z., & Firmananda, F. I. (2023). Comparison Random Forest Regression and Linear Regression For Forecasting BBCA Stock Price. *Jurnal Teknik Industri Terintegrasi*, *6*(3), 718–732. https://doi.org/10.31004/jutin.v6i3.16933

Putra, A. L., & Kurniawati, A. K. (2021). Analisis Prediksi Harga Saham PT. Astra International Tbk Menggunakan Metode Autoregressive Integrated Moving Average (ARIMA) dan Support Vector Regression (SVR): Array. *Jurnal Ilmiah Komputasi*, *20*(3), 417–424. Retrieved from https://ejournal.jak-stik.ac.id/index.php/komputasi/article/view/2732

Raya, M., Srinivasan, D., Vishnu, M., Adedoyin, A., & Sathiyanarayanan, M. (2022). Visualizing, Comparing and Forecasting Stock Market Prediction. *2022 IEEE Delhi Section Conference (DELCON)*, 1–7. https://doi.org/10.1109/delcon54057.2022.9753359

Saputra, C. B., & Koesrindartoto, D. P. (2024). Pemanfaatan Analisis Sentimen Youtube untuk Prediksi Harga Saham: Studi pada Investor Retail Indonesia. *Jurnal Manajemen,* *21*(1), 1–17. https://doi.org/10.25170/jm.v21i1.5184

Syafira, M. N., & Saepudin, D. (2023). Prediksi Return Saham dengan Metode Random Forest dan Penerapannya untuk Seleksi Portofolio. *EProceedings of Engineering*, *10*(5).

Thumu, S. R., Nellore, G., & others. (2024). Optimized Ensemble Support Vector Regression Models for Predicting Stock Prices with Multiple Kernels. *Acta Informatica Pragensia*, *13*(1), 24-37. doi: 10.18267/j.aip.226

Wahanani, H. E., Swari, M. H. P., & Akbar, F. A. (2020). Case based reasoning prediksi waktu studi mahasiswa menggunakan metode euclidean distance dan normalisasi min-max. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, *7*(6), 1279–1288. https://doi.org/10.25126/jtiik.2020763880

Yin, L., Li, B., Li, P., & Zhang, R. (2023). Research on stock trend prediction method based on optimized random forest. *CAAI Transactions on Intelligence Technology*, *8*(1), 274–284. https://doi.org/10.1049/cit2.12067

Zheng, J., Xin, D., Cheng, Q., Tian, M., & Yang, L. (2024). The Random Forest Model for Analyzing and Forecasting the US Stock Market in the Context of Smart Finance. *ArXiv Preprint ArXiv:2402.17194*.