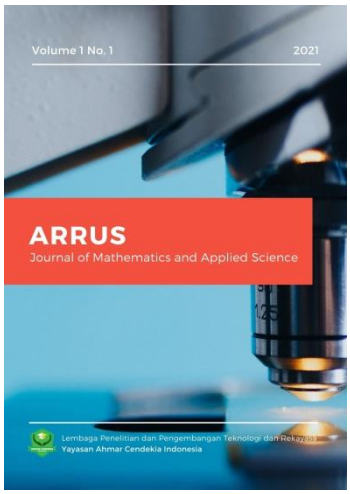
https://doi.org/10.35877/mathscience5529

OPEN ACCESS

ISSN [Print]: 2776-7922



*Corresponding author: Ansari Saleh Ahmar, Department Statistics, Faculty of Mathematics and Natural Sciences, Universitas Negeri Makassar, Makassar, South Sulawesi, Indonesia

E-mail: ansari@unm.ac.id

RESEARCH ARTICLE

The Effect of Mutual Information-Based Feature-Selection on Support Vector Regression for Goto Stock Price Prediction

Indy Pratiwi HR, Ansari Saleh Ahmar*, & Zulkifli Rais

Universitas Negeri Makassar, Makassar, South Sulawesi, Indonesia

Abstract: The dynamic movement of stock prices makes price prediction an important aspect in supporting investment decision-making. This study aims to apply Support Vector Regression (SVR) with Mutual Information (MI)-based feature selection to predict the stock price of PT. GoTo Gojek Tokopedia Tbk (GOTO). The data used are historical GOTO.JK data from 2 January 2023 to 30 December 2025, obtained from Yahoo Finance. After preprocessing, 688 observations were obtained and divided chronologically into 550 training data and 138 testing data. Next, 32 features were constructed, consisting of historical prices, trend indicators, momentum, volatility, volume, as well as lag and return features. Feature selection was performed using MI to measure the degree of association of each feature with the next period's closing price, and several scenarios were then formed based on the MI ranking. Modeling was carried out using SVR, while evaluation used RMSE, MAE, MAPE, and Directional Accuracy (DA). The results show that the Top 5 MI scenario, consisting of Close, Low, High, the 5-period EMA, and Open, yields an RMSE of 1.5774, an MAE of 1.1203, a MAPE of 1.8342%, and a DA of 48.5507%. These results indicate that MI-based feature selection helps SVR exploit more informative features and achieve effective predictive performance with fewer features. Accordingly, Top 5 MI was selected as the best feature configuration for SVR modeling in GOTO stock price prediction.

Keywords: Mutual Information, Support Vector Regression, Feature Selection, Stock Price Prediction, GOTO.JK

1. Introduction

Investment is an economic activity carried out by allocating a certain amount of funds or assets to obtain benefits and profits in the future. Along with the development of financial markets, various investment instruments are available to the public, one of which is stocks traded on the capital market. Stock investment has become a popular alternative because it offers investors the opportunity to earn returns through increases in stock prices as well as dividend payments (Fadilah et al., 2023).

In stock investment, the stock price is an important aspect that investors need to consider. A stock price is the value of a share formed through trading activity in the capital market and may change over time. Stock price movements can reflect a company's condition and performance and are one of the pieces of information investors consider when making investment decisions. Therefore, understanding stock price movements is important for investors in weighing investment decisions, given that stock prices may change from time to time (Fadila & Nuswandari, 2022). This is because stock prices are influenced by a number

This open access article is distributed under a Creative Commons Attribution (CC-BY-NC) 4.0 license.

https://doi.org/10.35877/mathscience5529

OPEN ACCESS

ISSN [Print]: 2776-7922

of factors, including rising interest rates and political dynamics (Ariyadi et al., 2022). Technology stocks such as that of PT. GoTo Gojek Tokopedia Tbk (GOTO) are no exception.

PT. GoTo Gojek Tokopedia Tbk (GOTO) is a company operating in Indonesia's digital economy ecosystem, with business activities that include on-demand transportation and delivery services, e-commerce, and financial services. As a publicly listed company whose shares are traded on the Indonesia Stock Exchange, the movement of GOTO's stock price is one aspect that can be observed in assessing market dynamics toward the company (Mulia & Gustian, 2025). The dynamic nature of GOTO's stock price, which can change over time, shows the importance of analysis that can provide an overview of the possible price movement in the next period. Therefore, an approach is needed that can utilize historical data to identify price movement patterns and produce predictions that can support investment decision-making (Kurniawan et al., 2026).

The ability to analyze stock movements is important for investors and traders in making investment decisions. Analysis is carried out to obtain an overview of possible changes in stock prices and to consider the level of return and risk that may be faced. In the decision-making process, investors can use various analytical approaches, including fundamental analysis, which focuses on the company's condition and performance, and technical analysis, which utilizes price movement patterns and historical stock data (Yuwono & Fajaryanto, 2025).

Advances in data analysis technology have encouraged the use of machine learning methods to predict stock price movements (Mahgfirah & Rahayu, 2025). One method that can be used is Support Vector Regression (SVR). SVR is a machine learning method used to make predictions on continuous data. SVR is capable of handling nonlinear relationships by utilizing kernel functions, so it can be applied to data with complex relationship patterns. The method works by seeking a regression function that gives predictions with a tolerable level of error while keeping model complexity under control so that it does not follow the characteristics of the training data too closely. This capability makes SVR a method that can be used for a variety of prediction problems, including stock price prediction, where price changes follow complex and dynamic patterns (Ishlah et al., 2023).

In SVR modeling, the use of relevant features is an aspect that needs attention because each feature can provide a different level of information about the target. One method that can be used to measure this relevance is Mutual Information (MI). Mutual Information is a measure used to determine how much information a feature provides about the target, based on the degree of informational dependence. A higher MI value indicates that a feature contains more information about the target. In addition, MI is able to identify nonlinear relationships, so it can be used to rank features according to their level of information. The resulting ranking can then be used as the basis for the feature selection process to determine the relevant features before they are applied to the SVR model (Xing et al., 2008).

Several previous studies exist. (Fahrezi et al., 2024) applied Support Vector Regression (SVR) to predict BBCA stock prices and obtained a MAPE of 1.07%. Another study by (Kurniawan et al., 2026) applied XGBoost to predict GOTO stock prices and obtained a MAPE of 1.96%. Meanwhile, (Xiong et al., 2022) developed a Mutual Information-based Support Vector Regression model to predict air passenger indices and obtained a MAPE of 8.04%.

These studies show that SVR can be used for financial data prediction, while Mutual Information can be used to support the feature selection process. However, the application of SVR with Mutual Information-based feature selection to GOTO stock price prediction still needs further study, particularly in determining the most informative features from historical data and technical indicators.

Based on these issues, this study applies Support Vector Regression (SVR) with Mutual Information-based feature selection to predict the stock price of PT GoTo Gojek Tokopedia

This open access article is distributed under a Creative Commons Attribution (CC-BY-NC) 4.0 license.

https://doi.org/10.35877/mathscience5529

OPEN ACCESS

ISSN [Print]: 2776-7922

Tbk (GOTO) using historical data. Mutual Information is used to identify and rank features according to their relevance to the stock price, and the features with the most relevant information are then used in SVR modeling. This approach is expected to produce a more effective model by utilizing informative features and reducing the use of less relevant ones. Model performance is then evaluated using prediction accuracy measures to determine the ability of SVR to predict the GOTO stock price. The results are expected to contribute to the application of Mutual Information-based feature selection in stock price prediction and to serve as a reference for future research on applying machine learning to capital market data.

2. Literature Review

2.1. Feature Engineering

In real-world classification and prediction tasks, datasets often contain a large number of attributes or features. Some of these features may contain similar information, be redundant, or even be irrelevant to the modeling process (Fayyad et al., 1996). Therefore, a feature engineering process is needed to construct and process features that can represent the characteristics of the data more informatively (Sulaiman & Labadin, 2015).

2.2. Mutual Information Feature Selection

In general, feature selection algorithms can be grouped into three main categories (Liu & Yu, 2005; Sulaiman & Labadin, 2015). The first category is filter methods, which use the statistical characteristics and probabilistic distributions of the attributes in a dataset to determine a relevant subset of features. The selection process in this method is carried out independently of any particular learning algorithm, so the selected features can be used by various learning algorithms. One example of a filter method is Mutual Information (MI)-based feature selection (Unler et al., 2011).

Mutual Information (MI), which is formulated based on the probability density function of the variables, is used as a filter method in feature selection. In the initial stage, the MI value is calculated between each feature in the training data and the output vector. Subsequently, the MI score obtained for each feature is used for feature ranking, so that features with a higher degree of association can be identified (Doquire & Verleysen, 2011).

In calculating Mutual Information (MI), information entropy is one of the basic concepts used to measure the level of uncertainty of a variable. Information entropy is an important concept in information theory introduced by Shannon in 1948 (Fan et al., 2013). This concept explains that the more orderly a system is, the smaller its information entropy. Conversely, the more disordered a system is, or the higher its uncertainty, the greater its entropy. Therefore, information entropy can be used to measure the level of uncertainty or orderliness of a system (Zhang et al., 2016). Mathematically, information entropy can be formulated as follows (Xiong et al., 2022):

$$H(X) = \sum_{i=1}^{n} P(x_i) log_2 P(x_i) \quad (1)$$

where:
$P(x_i)$ = probability of sample x
n = number of samples

Suppose the joint probability distribution of the random vector (X, Y) is denoted as p_{ij}. The two-dimensional joint entropy of the vector (X, Y) can be formulated as follows:

$$H(X,Y) = -\sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij} log_2 p_{ij} \quad (2)$$

This open access article is distributed under a Creative Commons Attribution (CC-BY-NC) 4.0 license.

https://doi.org/10.35877/mathscience5529

OPEN ACCESS

ISSN [Print]: 2776-7922

Assuming that the joint probability distributions of (X) and (Y) are denoted as p_{ig} and p_{gj}, respectively, the conditional entropy can be defined as follows:

$$H(X/Y) = -\sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij} log_2 \frac{p_{ij}}{p_{gj}} \quad (3)$$

$$H(Y/X) = -\sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij} log_2 \frac{p_{ij}}{p_{ig}} \quad (4)$$

Thus, Mutual Information (MI) can be expressed as an entropy value that indicates the reduction in uncertainty about variable (X) resulting from information about variable (Y), or conversely, the reduction in uncertainty about variable (Y) resulting from information about variable (X).

$$\begin{aligned} I(X;Y) &= H(X) - H(X|Y) \\ &= H(Y) - H(X|Y) \\ &= H(X/) + H(Y) - H(X,Y) \end{aligned} \quad (5)$$

By combining Equations (3)–(5), the complete formulation of Mutual Information (MI) can be simplified into the following form:

$$I(X;Y) = \sum_{i,j} p_{ij} log_2 \frac{p_{ij}}{p_{ig} p_{gj}} \quad (6)$$

2.3. *Support Vector Reggression*

Support Vector Regression (SVR) is an extension of the Support Vector Machine (SVM) method applied to solve regression problems. SVR aims to obtain a function $f(x)$ that acts as the optimal hyperplane or regression function, so that it can represent the relationship between input and output data with an error tolerance of ε. The value of ε is kept as small as possible so that the resulting estimates come closer to the actual values (Hyndman, 2014; Septinigrum et al., 2015; Smola et al., 2002).

Suppose there are l training data points (x_i, y_i), with $i = 1, \dots, l$. The variable x_i is an input vector expressed as $x = \{x_1, x_2, \dots, x_n\} \subseteq \mathbb{R}$, while y_i is a scalar output value with $y = \{y_1, y_2, \dots, y_l\} \subseteq \mathbb{R}$. Here, l denotes the number of data points used in the training process. Through SVR, a function $f(x)$ is constructed that produces predicted values with a maximum deviation of ε from the actual value y_i on the training data. When $\varepsilon = 0$, the model is expected to produce a regression function with very small deviation from the actual data (Santosa, 2007).

In the SVR process, the input vector can be mapped into a higher-dimensional feature space so that complex or nonlinear relationships in the data can be represented better (Abe, 2005). The regression function that acts as the optimal hyperplane can be expressed as:

$$f(x) = w^T \varphi(x) + b \quad (7)$$

where:

w = weight vector of the regression function
$\varphi(x)$ = function that maps the input data x into the feature space
b = bias value

2.4. *Forecast Accuracy Evaluation*

This open access article is distributed under a Creative Commons Attribution (CC-BY-NC) 4.0 license.

https://doi.org/10.35877/mathscience5529

ISSN [Print]: 2776-7922

Evaluation is carried out using forecast error measures as a means of assessing the accuracy of the predictions. Several measures commonly used to evaluate the performance of forecasting methods include:

Mean Absolute Error (MAE) is an evaluation metric used to measure the closeness between forecast values and actual values. MAE shows the magnitude of the prediction error regardless of the direction of the error, whether positive or negative (Amansyah et al., 2024). The MAE value is obtained from the average of the absolute differences between actual values and forecast values. This metric is widely used to measure prediction error in time series analysis (Pebralia, 2022) The MAE formula is as follows:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{8}$$

with:
y_i = actual value
$\hat{y}_i$ = forecast value
n = number of observations

Root Mean Squared Error (RMSE) is an evaluation measure used to determine the magnitude of prediction error by taking the square root of the Mean Squared Error (MSE). The advantage of RMSE is that its unit is the same as that of the data being analyzed, so the results are easier to interpret. Accordingly, an increase in RMSE indicates an increase in the magnitude of model error. Unlike MSE, RMSE has a more direct interpretation because it is on the same scale as the data. By analogy, MSE can be viewed as a measure related to variance, whereas RMSE has characteristics similar to standard deviation (Sinaga & Irawati, 2018). The RMSE formula is as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{9}$$

Mean Absolute Percentage Error (MAPE) is a relative accuracy measure used to determine the average absolute percentage error between actual values and forecasts (Sinaga & Irawati, 2018). The smaller the MAPE value, the smaller the forecast error. The MAPE formula is as follows:

$$MAPE = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \tag{10}$$

Table 1. MAPE Value Criteria

MAPE Value	Interpretation
MAPE < 10%	Very Good
10% < MAPE ≤ 20%	Good
20% < MAPE ≤ 50%	Fair
MAPE > 50%	Poor

Directional Accuracy (DA) is an evaluation measure used to determine the model's ability to predict the direction of price movement, that is, whether the price in the next period rises or falls compared with the closing price of the previous period. The DA formula is as follows:

$$DA = \frac{1}{n}\sum_{i=1}^{n} d_t \times 100\% \tag{11}$$

With:
d_t = indicator of directional prediction accuracy at time t

This open access article is distributed under a Creative Commons Attribution (CC-BY-NC) 4.0 license.

https://doi.org/10.35877/mathscience5529

ISSN [Print]: 2776-7922

3. Research Method and Materials

This study employed a quantitative approach to investigate the effect of Mutual Information (MI)-based feature selection on Support Vector Regression (SVR) for predicting the stock price of PT GoTo Gojek Tokopedia Tbk (GOTO). The historical stock data were obtained from Yahoo Finance for the period from January 2, 2023, to December 30, 2025. After preprocessing, 688 observations were retained and divided chronologically into training and testing datasets using an 80:20 ratio, consisting of 550 training observations and 138 testing observations, respectively.

The preprocessing stage involved constructing features from historical stock prices, technical indicators, trading volume, lagged values, and returns. A total of 32 features were generated to represent price levels, trends, momentum, volatility, and trading activity. The prediction target was defined as the next-period closing price, (Close_{t+1}), using information available at time (t). Feature scaling was performed using Min-Max Scaler, with the scaling parameters fitted exclusively to the training data and subsequently applied to the testing data.

Mutual Information (MI) was applied as a filter-based feature selection method to measure the informational association between each input feature and the target variable. The MI scores were calculated using the training data, and the features were ranked in descending order according to their scores. Based on this ranking, four feature selection scenarios were established, namely Top 5, Top 10, Top 15, and Top 20 MI. A model using all 32 features was also included as a reference to examine the effect of feature selection on prediction performance.

4. Results and Discussion

Descriptive analysis is used to provide a general overview of the characteristics of the closing price of PT GoTo Gojek Tokopedia Tbk stock, with the stock code GOTO.JK. The research data were obtained from Yahoo Finance and cover the closing stock price from January 2023 to December 2025.

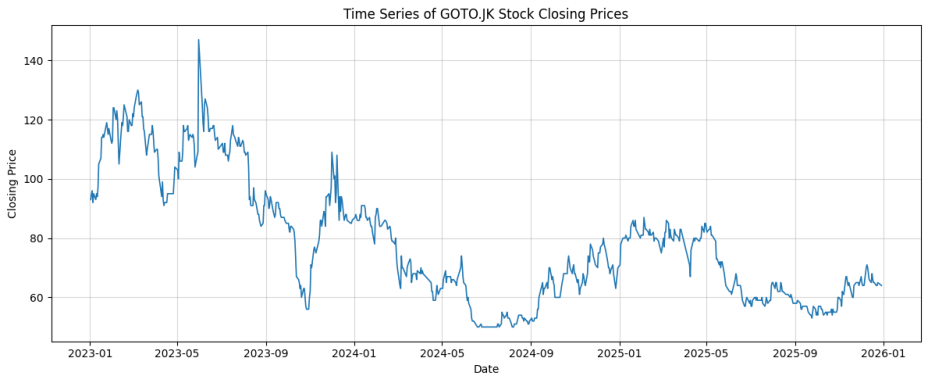


Figure 1. Time Series Plot OF GOTO JK Stock Closing Prices

Based on Figure 1, the closing price of GOTO.JK during the observation period shows fairly high fluctuation and changes in pattern over time. At the beginning of 2023, the stock price was around the 90s, then rose and reached a peak of 147 around mid-2023. After reaching this point, the stock price tended to decline, with several upward and downward movements. A significant decline can be seen from late 2023 to mid-2024, when the stock price fell to around 50. The stock price then rose again from late 2024 to early 2025, reaching the 80s, before fluctuating once more and moving within the range of 55–70 until the end of the observation period. Overall, the chart shows that the GOTO.JK stock price moves unstably, with a tendency to decline compared with its position at the beginning of the period. Based on the chart pattern, no strong seasonal pattern that repeats consistently every year is visible.

This open access article is distributed under a Creative Commons Attribution (CC-BY-NC) 4.0 license.

https://doi.org/10.35877/mathscience5529

OPEN ACCESS

ISSN [Print]: 2776-7922

Descriptive analysis was then carried out by calculating the minimum, maximum, mean, standard deviation, first quartile (25%), median (50%), and third quartile (75%). The results of the descriptive analysis of the GOTO.JK closing price are presented in Table 2.

Table 2. MAPE Value Criteria

Minimum	Maximum	Mean	Standard Deviation
50	147	78.25	20.9

Based on Table 2, the closing price of GOTO.JK during the observation period has a minimum value of 50.00 and a maximum value of 147.00. The mean closing price is 78.25, while the standard deviation is 20.90. A standard deviation that is fairly large relative to the mean indicates a fairly high dispersion of prices during the observation period. This is consistent with the chart, which shows that the GOTO.JK stock price fluctuates over time. Thus, descriptively, the GOTO.JK closing price shows fairly high variation, so this characteristic needs to be taken into account in time series modeling and forecasting.

The historical GOTO.JK stock data were split at a ratio of 80:20, with 550 observations as training data and 138 observations as testing data. Feature engineering was then carried out by constructing various technical features from the stock data. These features include trend indicators such as SMA and EMA, momentum indicators such as MACD, RSI, Stochastic, Williams %R, and ROC, volatility indicators such as Bollinger Bands and ATR, volume indicators such as OBV and average volume, as well as lag features of price and return.

In this study, the target to be predicted is the closing price in the next period, or Close(t+1), so that information at period t is used to predict the price at period t+1. After the features were constructed, scaling was performed using a Min-Max Scaler. The scaling was fitted only on the training data, and the resulting transformation was then applied to the testing data. Feature selection was then carried out using Mutual Information (MI) to determine the degree of association of each feature with the target. Based on the MI values, all features were ranked and several feature selection scenarios were created, namely the Top 5, Top 10, Top 15, and Top 20 features with the highest MI values. The next stage was modeling using Support Vector Regression (SVR).

After modeling and forecasting using the Support Vector Regression (SVR) method with Mutual Information (MI)-based feature selection, the next stage was evaluation to determine the performance of each model in producing forecasts. Model evaluation was carried out using several forecasting performance measures, namely MAE, RMSE, MAPE, and DA. Smaller MAE, RMSE, and MAPE values indicate better forecasting performance, whereas a larger DA value indicates better model performance. The evaluation results for each SVR model scenario are presented in Table 3.

Table 3. Comparison of Evaluation Results with the Best Model

Scenario	Number of Features	RMSE	MAE	MAPE	DA
All features	32	1.8318	1.4986	2.4774	27.5362
Top 5 MI	5	1.5774	1.1203	1.8342	48.5507
Top 10 MI	10	1.5859	1.1530	1.8911	41.3043
Top 15 MI	15	1.5773	1.1383	1.8640	43.4783
Top 20 MI	20	1.7389	1.3813	2.2667	28.9855
No features	0	1.6285	1.1304	1.8535	31.1594

Based on Table 1, Top 15 MI produced the lowest RMSE of 1.5773 and an R^2 of 0.8517, while Top 5 MI had the lowest MAE of 1.1203, the lowest MAPE of 1.8342, and the highest DA of 48.5507%. These results show that using selected features can improve forecasting performance compared with using all features.

The results show that the closing price of GOTO.JK during 2023–2025 exhibited fairly high fluctuation and an unstable movement pattern, with a lowest price of 50, a highest price of 147, and a standard deviation of 20.90. The price movement, which changed from around 90 at the beginning of 2023, rose to 147, then declined to around 50, and fluctuated again within the range of 55–70, shows that the stock price is influenced by market dynamics that change over time. The absence of a strong, repeating seasonal pattern indicates that GOTO.JK price

This open access article is distributed under a Creative Commons Attribution (CC-BY-NC) 4.0 license.

https://doi.org/10.35877/mathscience5529

ISSN [Print]: 2776-7922

changes during the study period were more dynamic than following a consistent annual seasonal pattern. This characteristic is the reason for using technical features that can capture information on past prices, trend, momentum, volatility, and volume in the forecasting process.

Methodologically, an important finding of this study is that using all 32 features does not always produce better forecasting performance. The model with all features produced an RMSE of 1.8318, an MAE of 1.4986, and a MAPE of 2.4774, whereas after feature selection using Mutual Information, model performance improved. Top 5 MI produced an MAE of 1.1203, a MAPE of 1.8342, an R^2 of 0.8517, and a DA of 48.5507%, while Top 15 MI produced the lowest RMSE of 1.5773. Scientifically, this finding can be explained by the fact that adding more features does not always provide additional information relevant to the target Close(t+1). Some features may have lower informational association with the target and therefore potentially add model complexity without improving predictive ability. Hence, MI feature selection can help SVR focus its learning on features that have a higher informational association with the next period's closing price. Indeed, the research process used MI on the training data to rank the features and form the Top 5, Top 10, Top 15, and Top 20 scenarios.

These findings also show that the optimal number of features need not be ever larger. Top 5 and Top 15 MI gave relatively better results than Top 20 MI. When the number of features increased to 20, the RMSE actually rose to 1.7389. This indicates that additional information from features with lower MI rankings does not necessarily improve SVR's predictive ability. Thus, the results support the notion that feature selection is an important part of stock price modeling, because a simpler model with more informative features can give better predictions than a model with a larger number of features.

Based on the test results, the Top 5 MI scenario was selected as the best scenario because it delivers good predictive performance using only the five features with the most relevant information about the target variable. This scenario produced an RMSE of 1.5774, an MAE of 1.1203, a MAPE of 1.8342%, and a DA of 48.5507%. The low MAE and MAPE values indicate that prediction errors are relatively small. In terms of directional prediction, Top 5 MI produced a DA of 48.5507%, the highest among all scenarios, whereas the model with all features produced a DA of only 27.5362%. This indicates that features selected based on Mutual Information (MI) contribute not only to the accuracy of predicted values but also help the model capture changes in the direction of stock price movements. Although a DA of 48.5507% has not yet reached 50%, the result still shows that features with high MI values can provide more relevant information to the model in predicting stock price movements. Thus, MI-based feature selection can help the SVR model utilize more informative features in the modeling process.

5. Conclusion

Based on the research results, selecting five features with the highest Mutual Information values (Top 5 MI) was able to improve the forecasting performance for the closing price of PT. GoTo Gojek Tokopedia Tbk stock using Support Vector Regression (SVR), compared with using all 32 features. The Top 5 MI model, consisting of Close, Low, High, the 5-period EMA, and Open, produced an MAE of 1.1203, a MAPE of 1.8342, a Directional Accuracy of 48.5507%, and an RMSE of 1.5774. These results show that using a small number of features that are highly informationally associated with the next period's closing price can produce a more effective model than using all features. Thus, the findings support the view that Mutual Information-based feature selection, particularly Top 5 MI, can improve SVR performance in forecasting the stock price of PT. GoTo Gojek Tokopedia Tbk.

References

Abe, S. (2005). Support Vector Machines For Pattern Classification. Springer.

This open access article is distributed under a Creative Commons Attribution (CC-BY-NC) 4.0 license.

ISSN [Print]: 2776-7922

Amansyah, I., Indra, J., Nurlaelasari, E., & Juwita, A. R. (2024). Prediksi Penjualan Kendaraan Menggunakan Regresi Linear : Studi Kasus pada Industri Otomotif di Indonesia. 4, 1199–1216.

Ariyadi, P., Effendi, M. M., & Raharjo, S. B. (2022). Analisa Prediksi Harga Saham Blue Chip LQ45 Dengan Metode Data Mining Backpropagation Neural Network (Studi Kasus di Bursa Efek Indonesia). Prosiding SAINTEK: Sains Dan Teknolog, 1(1), 68–76.

Doquire, G., & Verleysen, M. (2011). Feature Selection with Mutual Information for Uncertain Data. 330–341.

Fadila, A. N., & Nuswandari, C. (2022). Apa Saja Faktor – Faktor Yang Mempengaruhi Harga Saham? 15(2), 283–293.

Fadilah, A., Wiharno, H., & Nurfatimah, S. N. (2023). Pengaruh Harga Saham, Return Saham, Volatilitas Harga Saham, Ukuran Perusahaan Dan Volume Perdagangan Saham Terhadap Bid-Ask Spread Saham. Prosiding, 6, 212–226.

Fahrezi, R. A., Wijaya, M. Y., Fitriyati, N., Matematika, P. S., Islam, U., Syarif, N., & Jakarta, H. (2024). Prediksi Harga Penutupan Saham Bank Central Asia: Implementasi Algoritma Long Short-Term Memory Dan Perbandingannya Dengan Support Vector Regression. 5(1), 452–464.

Fan, X. L., Feng, H. H., & Yuan, M. (2013). Principal Component Analysis Feature Selection Algorithm Based On Mutual Information. Control and Decision, 28(6), 915–919.

Fayyad, U. M., Piatetsky-Shapiro, G., & Smyth, P. (1996). From Data Mining To Knowledge Discovery: An Overview.

Hyndman, R. J. (2014). Forecasting : Principles & Practice (Issue September).

Ishlah, A. W., Sudarno, & Kartikasari, P. (2023). Implementasi Gridsearchcv Pada Support Vector Regression (SVR) Untuk Peramalan Harga Saham (Studi Kasus: Harga Saham PT Anabatic Technologies Tbk). Jurnal Gaussian, 12(2), 276–286. https://doi.org/10.14710/j.gauss.12.2.276-286

Kurniawan, M. T., Pratama, I., Informasi, S., Mercu, U., & Yogyakarta, B. (2026). Implementasi XGBoost untuk Prediksi Saham. Jurnal Mahasiswa Teknik Informatika, 10(2), 3569–3574.

Liu, H., & Yu, L. (2005). Toward Integrating Feature Selection Algorithms for Classification and Clustering. IEEE Transactions on Knowledge and Data Engineering, 17(4).

Mahgfirah, A. A., & Rahayu, P. I. (2025). Analisis Support Vector Regression untuk Meramalkan Saham Perusahaan Dss di Indonesia. 7(1), 44–54. https://doi.org/10.35580/variansiunm356

Mulia, S., & Gustian, D. (2025). Analisis Kinerja Algoritma Prediksi Saham pada PT GoTo Gojek Tokopedia Tbk (GOTO). 11(1), 53–67.

Pebralia, J. (2022). (Jurnal Ilmu Fisika dan Pembelajarannya) Analisis Curah Hujan Menggunakan Machine Learning Metode Regresi Linier Berganda Berbasis Python dan Jupyter Notebook Rainfall Analysis using Machine Learning-Multiple Linear Regression Method Based on Python and Jupyter Notebook JIFP (Jurnal Ilmu Fisika dan Pembelajarannya) ISSN (online): ISSN (print): 6(2), 23–30.

Santosa, B. (2007). Data Mining Teknik Pemanfaatan Data untuk Keperluan Bisnis. Graha Ilmu.

Septinigrum, L., Yasin, H., & Sugito. (2015). Prediksi Indeks Harga Saham Gabungan Menggunakan Support Vector Regression (SVR) Dengan Algoritma Grid Search. Jurnal Gaussian, 4(2), 315–321.

Sinaga, H. D. E., & Irawati, N. (2018). Perbandingan Double Moving Average dengan Double Exponential Smoothing pada Peramalan Bahan Medis Habis Pakai. Jurnal Teknologi Dan Sistem Informasi, 4(2), 197–204.

Smola, A. J., Mangasarian, O. L., & Schölkopf, B. (2002). Sparse Kernel Feature Analysis. In: Gaul, W., Ritter, G. (eds) Classification, Automation, and New Media. Studies in Classification, Data Analysis, and Knowledge Organization. Springer.

Sulaiman, M. A., & Labadin, J. (2015). Feature Selection based on Mutual Information for Machine learning prediction of Petroleum reservoir properties. 2–7.

This open access article is distributed under a Creative Commons Attribution (CC-BY-NC) 4.0 license.

https://doi.org/10.35877/mathscience5529

ISSN [Print]: 2776-7922

Unler, A., Murat, A., & Babu, R. (2011). mr 2 PSO : A maximum relevance minimum redundancy feature selection method based on swarm intelligence for support vector machine classification. Information Sciences, 181, 4625–4641. https://doi.org/10.1016/j.ins.2010.05.037

Xing, H., Ha, M., Tian, D., & Hu, B. (2008). A Novel Support Vector Machine With Its Features Weighted by. 315–320.

Xiong, H., Fan, C., Chen, H., Yang, Y., Antwi, C. O., & Fan, X. (2022). A Novel Approach to Air Passenger Index Prediction : Based on Mutual Information Principle and Support Vector Regression Blended Model. 334. https://doi.org/10.1177/21582440211071102

Yuwono, K., & Fajaryanto, A. (2025). Sistem Prediksi Harga Saham Menggunakan Fuzzy Time Series Model Lee. 9(1), 92–103.

Zhang, C. T., Ma, Q. L., & Peng, H. (2016). Prediction of Chaotic Time Series Based On Information Entropy Optimization for Phase Space Reconstruction Parameters. Journal of Physics, 59, 7623–7629.

This open access article is distributed under a Creative Commons Attribution (CC-BY-NC) 4.0 license.