

# Google Trends and Technical Indicator-Based Machine Learning for Stock Market Prediction

Mamlatul Hani'ah , Moch Zawaruddin Abdullah , Wilda Imama Sabilla , Syafaat Akbar , Dikky Rahmad Shafara  
Politeknik Negeri Malang, Malang, Indonesia

---

## Article Info

### Article history:

Received August 15, 2022  
Revised February 23, 2023  
Accepted March 20, 2023

### Keywords:

Forecasting  
Google Trends  
Machine Learning  
Prediction  
Stock Price  
Technical Indicator

---

## ABSTRACT

The stock market often attracts investors to invest, but it is not uncommon for investors to experience losses when buying and selling shares. This causes investors to hesitate to determine when to sell or buy shares in the stock market. The accurate stock price prediction will help investors to decide when to buy or sell their shares. In this study, we proposed a new approach to predicting stocks using machine learning with a combination of features from stock price features, technical indicators, and Google Trends data. Combining features made it possible to gain a deeper understanding of the various factors that influence stock performance and make more accurate predictions. Three well-known machine learning algorithms, such as Support Vector Regression (SVR), Multilayer Perceptron (MLP), and Multiple Linear Regression, were used to predict future stock prices. The combination of stock price features, technical indicators, and Google Trends data can give highly accurate forecasting results. The test results showed that the SVR outperformed the MLP and Multiple Linear Regression to predict stock prices for Indonesian stocks, with an average MAPE is 0.50%. However, the Google Trends data only slightly affected the prediction outcome.

Copyright ©2022 The Authors.

This is an open access article under the [CC BY-SA](#) license.



---

## Corresponding Author:

Mamlatul Haniah,  
Department of Information Technology,  
Politeknik Negeri Malang, Malang, Indonesia,  
Email: [mamlatulhaniah@polinema.ac.id](mailto:mamlatulhaniah@polinema.ac.id)

---

How to Cite: M. Haniah, M. Abdullah, W. Sabilla, S. Akbar, and D. Shafara, "Google Trends and Technical Indicator based Machine Learning for Stock Market Prediction", *MATRIK : Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, vol. 22, no. 2, pp. 271-284, Mar. 2023.

This is an open access article under the CC BY-SA license (<https://creativecommons.org/licenses/by-sa/4.0/>)

## 1. INTRODUCTION

The increasing public investment in the capital market has an important role in economic growth and can affect economic resilience in the long term. The capital market, especially stocks, often attracts investors to invest. However, stocks are volatile, and stock prices can increase or decrease. For novice investors, the fluctuating value of stock prices is one thing to watch out for. This is because if investors make mistakes in selling or buying shares, it can result in losses for investors. It can even traumatize investors, so they do not dare to invest in the stock market [1].

The development of information technology in the economic sector has provided benefits such as buying, selling, investing, and banking that can be done remotely. In addition, with the development of information technology in the economic sector, we can quickly monitor the prices of staple foods [2]. Information technology in the stock market, especially machine learning techniques, can be used to predict stock prices [3, 4]. Stock price prediction is a classic problem. This problem often occurs in financial organizations, companies, and investors. The accurate stock price prediction will help investors to decide when to buy or sell their shares.

There has been much study on stock predictions, including technical indicators. One of the technical indicators that can be used in stock predictions is the moving average [5, 6]. Even nowadays, many investors use technical indicators to decide on buying and selling shares [7]. In addition, machine-learning approaches can be used in stock predictions [8, 9]. Many studies use technical indicators as a feature input processed by machine learning to make stock predictions [10]. [9] use technical indicators such as Simple Moving Average (SMA), Weighted Moving Average (WMA), Relative Strength Index (RSI), Accumulation/Distribution Oscillator (ADO), and Average True Range (ATR). Furthermore, this technical indicator is processed using machine learning algorithms such as Support Vector Regression (SVR) to predict stock prices. The study by [10] uses ten technical indicators as features. It is then processed with machine learning algorithms such as Artificial Neural Network (ANN), Support Vector Regression (SVR), and K-Nearest Neighbour (K-NN) for stock price predictions. Both studies only use technical indicator features as input for machine learning models without using stock price features such as stock price when opened (open), highest stock price in 1 day (high), lowest stock price in 1 day (low), and stock price when closed (close).

In stock predictions using machine learning, stock price data can be used as a feature, not only technical indicators but as feature input. The study [11] combines technical indicators with open, low, high, and close stock prices as feature input. The combination of technical indicator and stock price features is then processed using an integration of a Support Vector Machine (SVM) and K-Nearest Neighbour (KNN).

In addition to the technical indicators and stock price features, other studies [12] add Google Trends data as a feature to make stock predictions. Google provides a Trends service that can show how often a keyword is searched relative to the total volume of searches on Google Search. This service can provide a rough estimate of how many people are talking about a topic at any given moment. Before buying and selling shares, investors will look for more information about the stock market as decision support before buying and selling shares, affecting the increase in search volume on Google [13]. Therefore, Google Trends data can be used as a feature in stock predictions. Google Trends data can be used as an additional feature in stock prediction because it can help predict the direction of the stock market index [14]. The study [14] uses the Google Trends data feature and stock price features such as open, high, low, close, and volume to predict the stock's opening price direction for the S&P 500 index data and the Dow Jones average index. The use of Google Trends data provides better predictive results.

Another study that predicts stock price is a study by [15]. This research predicts the close price using Long Short-Term Memory (LSTM) technique. However, this study only relies on time series data on closing prices without using other stock price data. In addition, this study [15] does not use technical indicators as a feature. In contrast, in the world of investors, technical indicators are one of the considerations for investors in buying and selling shares [7].

This study proposed a new approach to predicting stocks using a combination of stock price features, technical indicators, and Google Trends data. The main contribution of this study is combining stock price features, technical indicators, and Google Trends data. Technical indicators can provide information about the stock market. Google Trends data provide information on the level of public interest in a stock. Using multiple features makes it possible to gain a deeper understanding of the factors that influence stock performance and make more accurate predictions. In this study, we look for the most relevant keywords by comparing the correlation coefficients of several keywords to determine the effect of a keyword on the close price. Keywords that have the highest correlation coefficient are used as input for the machine learning model. We use three well-known machine learning algorithms, Support Vector Regression (SVR), Multiple Linear Regression, and Multilayer Perceptron (MLP), to predict future stock prices. In this study, to analyze the model that had been created, we use data from six stocks in Indonesia. Furthermore, stock price prediction can help investors make strategic investment decisions, such as when to buy or sell their shares.

## 2. RESEARCH METHOD

This study is quantitative research that uses numerical data and statistical analysis to understand and analyze the results of stock market predictions. In this section, we will explain details of the data used in this study first, then details of the proposed method will be explained later.

### 2.1. Data

This study uses six stocks listed on the Indonesian stock exchange. These six stocks were chosen randomly from the top 10 in February 2022. The six stocks used were BBKA (Bank Central Asia), HMSP (PT Hanjaya Mandala Sampoerna), TLKM (PT Telkom Indonesia Tbk), BBRI (PT Bank Rakyat Indonesia Tbk), ASII (Astra International Tbk), UNVR (Unilever Indonesia Tbk). The historical stock price data for each stock is obtained from the yahoo finance website (<https://finance.yahoo.com/>), while Google Trends data is obtained from the Google site (<https://www.Google.com/Trends>). Data were collected from January 2017 to March 2022. Figure 1 is an example of data for all stocks. Furthermore, the data will be divided into training data, validation data, and testing data. Data from January 2017 to January 2022 is used as training data, data from February 2022 is used as validation data, and data from March 2022 as testing data.



Figure 1. Daily Close Price Dataset

### 2.2. Methods

This study proposes a new approach combining stock price data, technical indicators, and Google Trends data as input features for machine learning models. In this study, the Google Trends data will be shifted daily to determine the effect of Google Trends on stocks in the following days. The proposed method for stock market prediction is shown in Figure 2. The first stage is data preparation, combining input data in stock prices, technical indicators, and Google Trends data. The next step is data pre-processing. Several machine learning techniques, including Support Vector Regression (SVR), Multilayer Perceptron (MLP), and Multiple Linear regression, are used to predict stock market price. Prediction results were evaluated using MAPE to determine the performance of the proposed method.

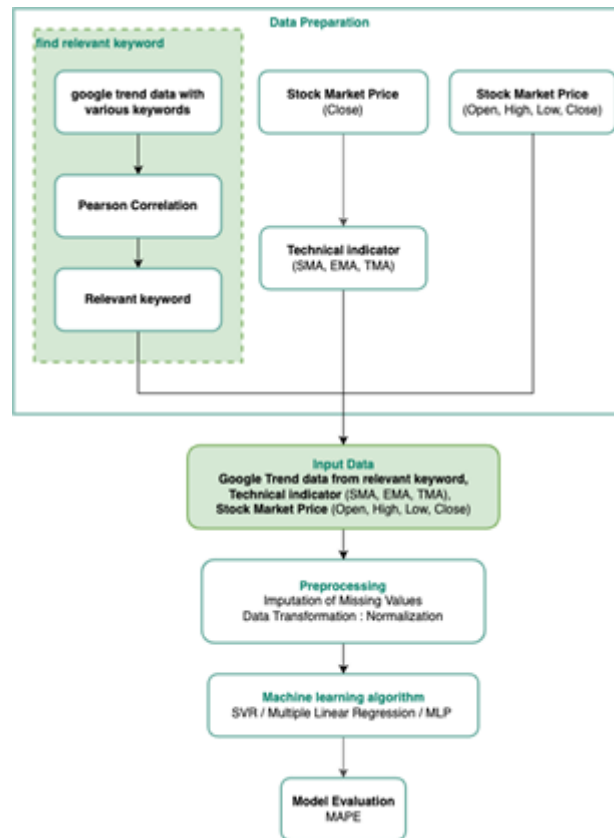


Figure 2. Proposed method

### 2.3. Data Preparation

At this stage, three processes must be carried out to obtain input data. The first process is to find the most relevant keywords in the Google Trends data, the second process is the calculation of technical indicators, and the third process is preparing stock price data.

#### 1. Find Relevant Keyword

Search keywords are needed to obtain Google Trends data. These keywords get the total search volume. Table 1 shows the search keywords used for each stock. The Pearson correlation coefficient determines the relevant keywords between Google Trends data and close prices. The correlation coefficient is a method to determine the relationship between two variables. The correlation value has a range between -1 to 1. A positive correlation coefficient means that the value of the two variables increases. If the correlation coefficient is negative, it means that the value of one variable increase while the other variable decrease. If the correlation coefficient is 0, then it is considered that there is no linear relationship between the two variables. Pearson correlation coefficient is shown in equation (1) [16].

$$p(x, y) = \frac{cov(x, y)}{\sigma_x \sigma_y} \quad (1)$$

In (1),  $cov(x, y)$  is the covariance between variables  $x$  and  $y$ ,  $\sigma_x$  is the standard deviation of variable  $x$ , and  $\sigma_y$  is the standard deviation of variable  $y$ . Correlation calculations are carried out for each keyword against the close price on the same date. It is to find the relationship between Google Trends and close price data. Google Trends keyword that has the highest correlation will be the feature input for machine learning models.

Table 1. List of keywords for each stock

Stock Symbols		Keywords			
UNVR.JK	saham UNVR	UNVR	saham unilever	harga UNVR	IDX:UNVR
BBRI.JK	saham BBRI	BBRI	saham BRI	harga BBRI	IDX:BBRI
TLKM.JK	saham TLKM	TLKM	saham TELKOM	harga TLKM	IDX:TLKM
HMSP.JK	saham HMSP	HMSP	saham Sampoerna	harga HMSP	IDX:HMSP
ASII.JK	saham ASII	ASII	saham astra	harga ASII	IDX:ASII
BBCA.JK	saham BBKA	BBCA	saham bca	harga BBKA	idx : BBKA

## 2. Calculate Technical Indicator

Technical indicators have proven to provide information about the stock market. In addition, it has been used to predict stock prices in the future [17]. In this study, technical indicators are calculated by processing the historical close price data. Three popular technical indicators used in this study are Simple Moving Average (SMA), Exponential Moving Average (EMA), and Triangular Moving Average (TMA).  $n$  periods (window length) must be determined beforehand to calculate the three technical indicators. The window length used in this study is five days.

### Simple Moving Average (SMA)

SMA is a calculation of the average stock price in  $n$  time periods. SMA aims to smooth the data by adding up the close price  $C_t$  of  $n$  data, which is then divided by  $n$ . This process will be repeated by shifting the window to the next data [10]. The SMA formula is in equation (2).

$$SMA = \frac{1}{n} \sum_{i=0}^{n-1} C_{t-i} \quad (2)$$

### Exponential Moving Average (EMA)

EMA is a type of Moving Average with a weight ( $\omega$ ) that will decrease exponentially so that the weight of the current price is greater than the previous price. Like SMA, the EMA also requires  $n$  periods (window length) to be determined beforehand. EMA can be calculated using equation (3) [10].

$$EMA_n = \sum_{i=0}^{n-1} \omega_i C_{t-i} \quad (3)$$

$$\sum_{i=0}^{n-1} \omega_i = 1 \quad (4)$$

### Triangular Moving Average (TMA)

TMA is the average of SMA with the same window length as SMA. TMA is calculated using equation (5), the sum of SMA divided by  $n$  periods (window length) [11].

$$TMA = \frac{SMA_1 + SMA_2 + \dots + SMA_n}{n} \quad (5)$$

## 3. Stock Price Data

Yahoo finance provides historical stock data. This study uses Open, High, Low, and Close data from January 2017 to March 2022. Open, High, and Low prices are used as a feature, along with technical indicators and Google Trends data, while the close price is the price that will be predicted.

### 2.4. Input Data

After the data preparation, Google trend data is generated from the most relevant keywords. Data input is also obtained from calculating SMA, EMA, and TMA data, as well as Open, High, and Low stock prices. Finally, we combine this data to become feature input and will be processed by machine learning algorithms.

## 2.5. Pre-processing

The pre-processing stage consists of data cleaning, missing value imputation, and data normalization using min-max normalization.

### 1. Data cleaning dan Missing Value Imputation

The missing stock price data will be replaced by using the mean imputation technique. This imputation works by calculating the average of the non-missing values in a column and then replacing the missing values.

### 2. Data Normalization

The addition of the Google Trends data feature causes a wide range of values between the Google Trends data and stock prices or technical indicators, so re-scaling of the input data is required. This study uses data normalization to generate a new data range between 0 and 1. The normalization formula can be seen in equation (6) [18].

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (6)$$

Where  $X'$  is the scale of the data in the new range,  $X$  is the value of the data to be normalized,  $X_{max}$  is the maximum value of the variable, and  $X_{min}$  is the minimum value of the variable.

## 2.6. Machine Learning Model

This stage describes the models used to conduct training, validation, and testing. We used Support Vector Regression (SVR), Multiple Linear Regression, and Multilayer Perceptron (MLP). The training process uses data from January 2017 to January 2022, then the tuning parameter uses validation data in February 2022, while testing used data in March 2022. When all models are trained and validated, it will predict market price data in March 2022.

### 1. Support Vector Regression (SVR)

SVR is a development of the SVM algorithm to solve regression problems. SVR tries to find the best hyperplane. The SVR algorithm has several kernels that can be used, and this study uses the Radial Basis Function (RBF) as the kernel. The RBF kernel functions are shown in equation (7) [19].

$$K(x, y) = \exp(-\gamma|x - y|^2) \quad (7)$$

In (7),  $x$  and  $y$  are data,  $|x - y|^2$  is the Euclidean distance of the two data, and  $\gamma$  is a free parameter to control the spread of the radial basis function. We use data validation to find the best  $\gamma$  parameter and, in this study, used  $\gamma = 0.1$ .

### 2. Multiple Linear Regression

Multiple Linear Regression is a derivative of simple linear regression where the independent variable in multiple linear has several features [18]. In (8), (9), and (10),  $y$  is the column vector of the response/dependent variable values.  $\alpha$  is the value of the intercept,  $\beta$  is the slope vector,  $x$  is a  $m \times n$  dimension matrix of the independent variable value in the training data,  $m$  is the number of training data, and  $n$  is the number of features.

$$y = \alpha + \beta_1x_1 + \beta_2x_2 + \dots + \beta_nx_n \quad (8)$$

$$Y = \beta X \quad (9)$$

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \dots \\ Y_n \end{bmatrix} = \begin{bmatrix} \alpha + \beta_1x_1 \\ \alpha + \beta_2x_2 \\ \dots \\ \alpha + \beta_nx_n \end{bmatrix} \quad (10)$$

### 3. Multilayer Perceptron (MLP)

Multilayer perceptron (MLP) is one of the Artificial Neural Network (ANN) algorithms that have several layers of neurons. Figure 3 is the architecture of the MLP used in this study. There are three layers in MLP, the input, hidden, and output layers. The hidden layer used in this study consists of 100 layers. The input layer and hidden layer consist of eight neurons, while the output layer consists of one neuron.

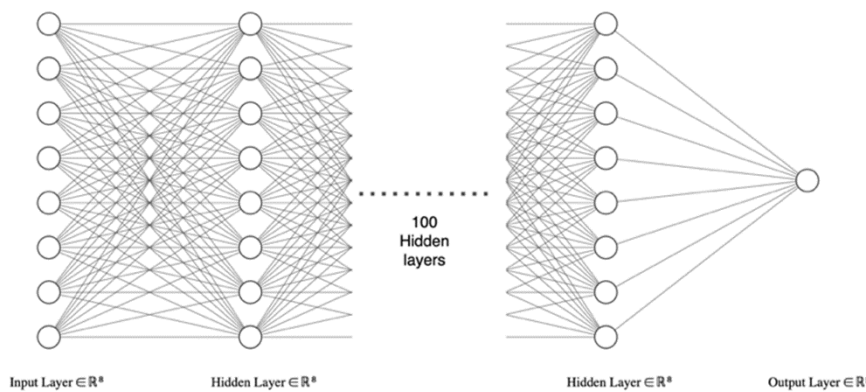


Figure 3. Effects of selecting different switching under dynamic condition

Furthermore, MLP requires an activation function. The study [20] compared the rectified linear unit (ReLU), sigmoid, and hyperbolic tangent (tanh) activation functions for stock price prediction and found that ReLU has a better performance compared to other activation functions. So, in this study, ReLU in equation (11) is used as an activation function.

$$ReLU(\alpha) = \max(0, \alpha) \tag{11}$$

### 2.7. Model Evaluation

When all models are trained and validated, prices are predicted from testing data in March 2022. Then the model will be evaluated using MAPE (Mean Absolute Percentage Error) to determine the prediction error. MAPE can be calculated as in (12) by finding the average of the absolute difference between the actual price and the predicted price divided by the actual price [21]. Actual is the close price stock, predicted is the corresponding value predicted by the model, and n is the number of data.

$$MAPE = \frac{1}{n} \sum \frac{|actual - prediction|}{actual} * 100 \tag{12}$$

Table 2 is the interpretation table of MAPE values to determine the model’s performance [22]. Based on Table 2, the smaller the MAPE value, the more accurate the prediction results. A low MAPE value indicates that the prediction model is accurate and the prediction error is small. On the other hand, a high MAPE indicates that the prediction model is less accurate and the prediction error is large.

Table 2. The interpretation of MAPE

MAPE	Interpretation
<10	Highly Accurate forecasting
10 - 20	Good forecasting
20 - 50	Reasonable forecasting
>50	Inaccurate forecasting

### 3. RESULT AND ANALYSIS

In this study, we used 63 months of data. Six company stocks were chosen randomly from the top 10 IDX stocks in February 2022. Experiments were carried out on each of the stocks of BBCA, HMSP, TLKM, BBRI, ASII, and UNVR. Each company stock has five keywords. We use Indonesian keywords because the stock data used is Indonesian stock data. Examples of keywords for UNVR: are saham UNVR (UNVR stock), UNVR, saham Unilever (Unilever stock), harga UNVR (UNVR price), and IDX:UNVR. The Google Trends data keyword with the highest correlation coefficient is used as a feature for the machine learning model. Table 3 is the result of the calculation of the Pearson correlation coefficient. If the correlation value is close to 1, then it indicates that when one variable increases, the other variable also increases at the same rate. If the correlation value is close to -1, then it indicates that when one variable increases, the other variable decreases at the same rate. A correlation of 0 indicates no linear relationship between the two variables. The data written in bold are the data selected as features. For all stocks (UNVR, BBRI, HMSP, ASII, BBCA), the keywords that have the highest correlation coefficient are keywords that match the company symbol on the stock exchange.

Table 3. Pearson correlation coefficient of search keywords

UNVR					
Keywords	<i>saham UNVR (UNVR stock)</i>	UNVR	<i>saham unilever (unilever stock)</i>	<i>harga UNVR (UNVR price)</i>	IDX:UNVR
Pearson correlation	-0.229173	<b>-0.28585</b>	0.12768	-0.023398	-0.26512
BBRI					
Keywords	<i>saham BBRI (BBRI stock)</i>	BBRI	<i>saham BRI (BRI stock)</i>	harga BBRI (BBRI price)	IDX:BBRI
Pearson correlation	0.156729	<b>0.22776</b>	-0.154502	0.074319	0.178292
TLKM					
Keywords	<i>saham TLKM (TLKM stock)</i>	TLKM	<i>saham TELKOM (TELKOM stock)</i>	<i>harga TLKM (TLKM price)</i>	IDX:TLKM
Pearson correlation	-0.185975	<b>-0.23331</b>	0.220197	-0.022358	-0.202742
HMSP					
Keywords	<i>saham HMSP (HMSP stock)</i>	HMSP	<i>saham Sampoerna (Sampoerna stock)</i>	<i>harga HMSP (HMSP price)</i>	IDX:HMSP
Pearson correlation	-0.307947	<b>-0.36766</b>	0.105158	-0.120991	-0.351156
ASII					
Keywords	<i>saham ASII (ASII stock)</i>	ASII	<i>saham astra (astra stock)</i>	<i>harga ASII (ASII price)</i>	IDX:ASII
Pearson correlation	-0.302238	<b>-0.36776</b>	0.037008	-0.031194	-0.321621
BBCA					
Keywords	<i>saham BBCA (BBCA stock)</i>	BBCA	<i>saham BCA (BCA stock)</i>	<i>harga BBCA (BBCA price)</i>	idx : BBCA
Pearson correlation	0.074452	<b>0.350086</b>	-0.095936	-0.014305	0.322234

We used Simple Moving Average (SMA), Exponential Moving Average (EMA), and Triangular Moving Average (TMA) to calculate technical indicator features. The moving average technical indicator is used to smooth out the price and identify Trends. Figure 4 is an example of a technical indicator plotted alongside the close price data on BBCA stock. The figure shows how close the technical indicator is to price data and follows the same trend of the stock's price. The black dashed line is close price stock data, the SMA is represented as a blue line, the EMA is represented as a green line, and the TMA is represented as a red line.





Figure 4. The technical indicator on BBCA stock

Google Trends data with selected keywords are combined with stock prices and technical indicators to become input for the machine learning models. Based on this combination, this study has seven features. We used the average imputation technique to fill in the missing values in the BBCA stock data on December 20, 2020. So, this data is replaced by the average value of the non-missing value data. In Figure 5, the chart shows the values of the seven features that we used. There is a big gap between the values of the Google Trends data and the other features. This means that the scale of Google Trends data is different from the other features, and this could affect the prediction result. Because of this big gap, we normalize the feature using min-max normalization. Figure 6 shows the result of applying min-max normalization to the seven features. All the features are in the range of 01, including Google Trends data, which was previously much higher than the other features. Normalizing the features ensures that all features are on the same scale, which allows for more accurate predictions.



Figure 5. Data before normalization on BBCA stock

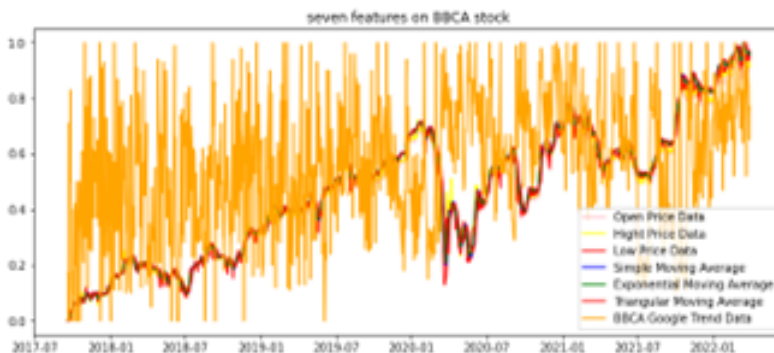


Figure 6. Data after normalization on BBCA stock

Stock price predictions are executed using three different models SVR, Multiple Linear Regression, and MLP. The graph of the test results for each stock is shown in Figure 7 to Figure 12. This graph is the prediction results of the proposed method to predict stock prices in February 2022. The black line is the actual stock price data; the blue line is the prediction result using the SVR model; the green line is the prediction result using Multiple Linear regression; and the red line is the prediction result using MLP.



Figure 7. ASII stock prediction result

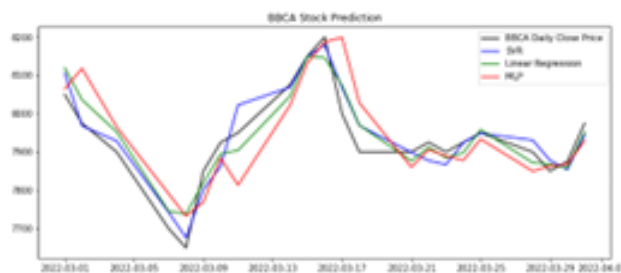


Figure 8. BBCA stock prediction result



Figure 9. BBRI stock prediction result

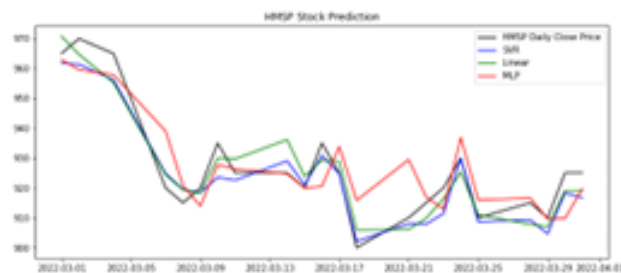


Figure 10. HMSP stock prediction result



Figure 11. TLKM stock prediction result

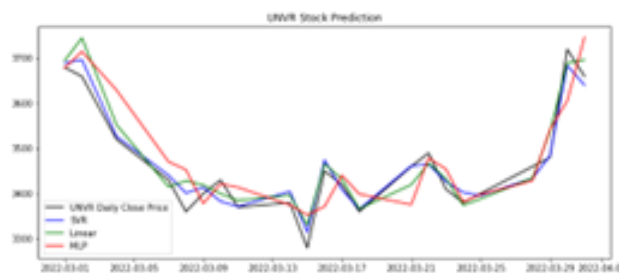


Figure 12. UNVR stock prediction result

From the picture of the prediction results, the SVR models give a prediction value that is close to the actual stock price data on the test data. It is also evident in Table 4. Table 4 is the MAPE result for each company stock using the proposed method. In all company stock, the use of SVR, Multiple Linear Regression, and MLP can produce MAPE values that are close to 0 and are in the range of less than 10. Based on the interpretation of the MAPE value, the proposed method gives highly accurate forecasting results. Furthermore, MAPE values in bold indicate the best MAPE values for each company stock. From the table, the SVR model has the best MAPE value for almost all company stocks except for BBRI. For BBRI, the MAPE value uses Multiple Linear Regression slightly better than the other two models. MLP shows lower performance compared to SVR, with an average of -0.4 lower than SVR.

Table 4. MAPE for each company stock using SVM, Multiple Linear Regression, and MLP model

Stock Symbols	SVR	Multiple Linear Regression	MLP
UNVR	0,5539	0,8289	1,3408
BBRI	0,5374	<b>0,5259</b>	0,6537
TLKM	0,4165	0,5812	0,7935
HMSP	0,4986	0,5478	0,7969
ASII	0,5362	0,7720	1,0797
BBCA	0,4603	0,4725	0,7556
Average	0,500483	0,621383	0,903367

The next experiment is to determine the performance of the proposed method, and a comparison is made using two scenarios. The first scenario uses Google Trends data as an additional feature, while the second scenario uses features without Google Trends data. The input features in the first scenario are Open, low, high, SMA, WMA, TMA, and Google Trends data. In contrast, the input features in the second scenario are Open, low, high, SMA, WMA, and TMA. Table 5 shows the comparison results using of first scenario and second scenarios. The first scenario using the additional Google Trends data feature can give MAPE values slightly better than the second scenario for TLKM and ASII. Meanwhile, for UNVR, BBRI, HMSP, and BBCA, a better MAPE value is obtained using the second scenario. If we look at the average for each model, the difference between the first and second scenarios is very small, around 0.01. This is because Google Trends data used in this study is data on the same day as stock prices. There is a possibility that Google Trends data can also affect stock prices in the following days. So that in the future, it is necessary to analyze using day shifts

Table 5. MAPE for first and second experiment scenario

Stock Symbols	Scenario 1 (with Google Trends data)			Scenario 2 (without Google Trends data)		
	SVR	Multiple Linear Regression	MLP	SVR	Multiple Linear Regression	MLP
UNVR	0,5539	0,8289	1,3408	0,5357	0,8253	1,1967
BBRI	0,5374	0,5259	0,6537	0,5494	0,5209	0,6677
TLKM	0,4165	0,5812	0,7935	0,4533	0,5846	0,8256
HMSP	0,4986	0,5478	0,7969	0,4093	0,5508	0,7170
ASII	0,5362	0,7720	1,0797	0,5567	0,7724	1,0803
BBCA	0,4603	0,4725	0,7556	0,4273	0,4670	0,7172
Average	0,5005	0,6214	0,9034	0,4886	0,6202	0,8674

Based on Table 5, All scenarios give highly accurate forecasting results in all stocks. In addition, SVR provides predictions with the smallest MAPE values for both the input feature using Google Trends and without using Google Trends. Therefore, we conclude that the SVR model outperformed the MLP and Multiple Linear Regression in predicting stock prices of Indonesian stocks.

The next experiment compares the MAPE value of the proposed method with that of previous research. This comparison is crucial to determining how superior the proposed method's performance is to previous research. The results of this comparison are presented in Table 6, where the proposed method for scenario 1 uses stock price, technical indicators, and the Google Trends feature. In contrast, scenario 2 uses stock price and the technical indicators feature. The proposed method is compared with the LSTM method from the previous research that only uses close price data without other stock price data [15]. To make the comparison fairer, the data used for comparison is the same data from this research. Based on Table 5, the proposed method outperforms the previous study with a smaller MAPE value. This finding highlights the importance of combination features in stock price prediction.

Table 6. Performance comparison

Stock Symbols	Scenario 1 (with Google Trends data)			Scenario 2 (without Google Trends data)			Previous research [15]
	SVR	Multiple Linear Regression	MLP	SVR	Multiple Linear Regression	MLP	LSTM
UNVR	0,5539	0,8289	13,408	0,5357	0,8253	11,967	7.9195
BBRI	0,5374	0,5259	0,6537	0,5494	0,5209	0,6677	14.5641
TLKM	0,4165	0,5812	0,7935	0,4533	0,5846	0,8256	22.1783
HMSP	0,4986	0,5478	0,7969	0,4093	0,5508	0,7170	2.2871
ASII	0,5362	0,7720	10,797	0,5567	0,7724	10,803	19.0425
BBCA	0,4603	0,4725	0,7556	0,4273	0,4670	0,7172	1.66
Average	0,5005	0,6214	0,9034	<b>0,4886</b>	0,6202	0,8674	11.2752

#### 4. CONCLUSION

In this study, a combination of stock price, technical indicators, and Google Trends data features have been used to predict stock prices using machine learning approaches. We compare Five keywords in the Google Trends data to show that the keywords with the highest correlation coefficient are those that match the stock exchange symbol. Technical indicators used in this study are SMA, WMA, and TMA. Close prices are used to calculate these three technical indicators. The technical indicator is close to actual price data and follows the same trend as the stock's price. Based on the experiment, the use of a combination of Google Trends data features, technical indicators, and stock prices provides highly accurate forecasting results in all algorithms. We compare SVR, Multiple Linear Regression, and MLP. The Multiple Linear Regression provides a 0.62% average MAPE, the MLP provides a 0.90% average MAPE, and the SVR provides the best MAPE with an average MAPE is 0.50%. Based on the experiment, we can conclude that the SVR outperformed the MLP and Multiple Linear Regression in predicting stock prices for Indonesian stocks. We also compare our proposed method with data without the Google Trends feature. The data without the Google Trends feature gives slightly better MAPE results because the Google Trends data used in this study is from the same day as stock prices.

In the future, it is necessary to conduct further research on the effect of Google Trends data on tomorrow's stock prices or the next few days. In addition, other technical indicators such as moving average convergence/divergence rules (MACD), relative strength index (RSI), and on-balance-volume (OBV) can also be used. The addition of the Google Trends data feature and other technical indicators is expected to produce a better MAPE value.

## 5. ACKNOWLEDGEMENTS

The researcher would like to express their sincere gratitude to Politeknik Negeri Malang for their invaluable support throughout the research project. The approval of the research proposal by the P2M reviewer team at Politeknik Negeri Malang is also appreciated. Additionally, the anonymous reviewers and editors are also appreciated for their insightful and constructive comments, which have contributed to the enhancement of the quality of the research work.

## 6. DECLARATIONS

### AUTHOR CONTRIBUTION

The first author was assigned to compile and design analysis, develop models, and write papers. The second author is tasked with helping analyze and helps compile research proposals. The third author is in charge of helping to compile research and write the proposal. The fourth and fifth author is assigned to collect data.

### FUNDING STATEMENT

This work was supported by Politeknik Negeri Malang under DIPA Fund Number: SP DIPA 023.18.2.677606/2022.

### COMPETING INTEREST

The authors state that there are no competing interests.

## REFERENCES

- [1] A. Pamela, "Fluktuatif dalam Investasi Saham dan Cara Menyikapinya [Fluctuations in Stock Investment and How to Respond to It]," dec 2021. [Online]. Available: <https://ajaib.co.id/fluktuatif-dalam-investasi-saham-dan-cara-menyikapinya/>
- [2] M. Mayadi and A. Anggrawan, "Pengembangan Sistem Informasi Pemantauan Harga Beras dan Gabah dengan Short Message Gateway [Development of Rice and Grain Price Monitoring Information System with Short Message Gateway]," *MATRIK : Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, vol. 21, no. 2, pp. 237–248, mar 2022. [Online]. Available: <https://journal.universitاسbumigora.ac.id/index.php/matrik/article/view/1546>
- [3] A. Santoso and S. Hansun, "Prediksi IHSG dengan Backpropagation Neural Network [JCI Prediction with Backpropagation Neural Network]," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 3, no. 2, pp. 313–318, aug 2019. [Online]. Available: <https://jurnal.iaii.or.id/index.php/RESTI/article/view/887>
- [4] M. M. Kumbure, C. Lohrmann, P. Luukka, and J. Porras, "Machine learning techniques and data for stock market forecasting: A literature review," *Expert Systems with Applications*, vol. 197, no. 1, p. 116659, jul 2022.
- [5] N. N. M. Cahyani and L. P. Mahyuni, "Akurasi Moving Average Dalam Prediksi Saham LQ45 di Bursa Efek Indonesia [Moving Average Accuracy In Prediction Of LQ45 Share In The Indonesia Stock Exchange]," *Jurnal Manajemen*, vol. 9, no. 7, pp. 2769–2789, jul 2020. [Online]. Available: <https://ojs.unud.ac.id/index.php/manajemen/article/view/60655>
- [6] A. Raudys and Ž. Pabarškait, "Optimising the smoothness and accuracy of moving average for stock price data," *Technological and Economic Development of Economy*, vol. 24, no. 3, pp. 984–1003, jan 2018.
- [7] M. C. Rivo and R. T. Ratnasari, "Faktor Yang Mempengaruhi Perilaku Investor Muslim Dalam Keputusan Berinvestasi Saham Syariah [Factors Affecting The Behavior Of Muslim Investors In Investment Decision Of Shariah Stock]," *Jurnal Ekonomi Syariah Teori dan Terapan*, vol. 7, no. 11, pp. 2202–2220, nov 2020. [Online]. Available: <https://e-journal.unair.ac.id/JESTT/article/view/22644>
- [8] B. M. Henrique, V. A. Sobreiro, and H. Kimura, "Literature review: Machine learning techniques applied to financial market prediction," *Expert Systems with Applications*, vol. 124, no. 1, pp. 226–251, jun 2019.
- [9] ———, "Stock price prediction using support vector regression on daily and up to the minute prices," *The Journal of Finance and Data Science*, vol. 4, no. 3, pp. 183–201, sep 2018.
- [10] Y. Shynkevich, T. M. McGinnity, S. A. Coleman, A. Belatreche, and Y. Li, "Forecasting price movements using technical indicators: Investigating the impact of varying input window length," *Neurocomputing*, vol. 264, no. 1, pp. 71–88, nov 2017.

- [11] R. K. Nayak, D. Mishra, and A. K. Rath, "A Naïve SVM-KNN based stock market trend reversal analysis for Indian benchmark indices," *Applied Soft Computing*, vol. 35, no. 1, pp. 670–680, oct 2015.
- [12] M. Y. Huang, R. R. Rojas, and P. D. Convery, "Forecasting stock market movements using Google Trend searches," *Empirical Economics*, vol. 59, no. 6, pp. 2821–2839, dec 2020. [Online]. Available: <https://link.springer.com/article/10.1007/s00181-019-01725-1>
- [13] T. Preis, H. S. Moat, and H. Eugene Stanley, "Quantifying Trading Behavior in Financial Markets Using Google Trends," *Scientific Reports 2013 3:1*, vol. 3, no. 1, pp. 1–6, apr 2013. [Online]. Available: <https://www.nature.com/articles/srep01684>
- [14] H. Hu, L. Tang, S. Zhang, and H. Wang, "Predicting the direction of stock markets using optimized neural networks with Google Trends," *Neurocomputing*, vol. 285, no. 1, pp. 188–195, apr 2018.
- [15] G. Budiprasetyo, M. Hani'ah, and D. Z. Alfah, "Prediksi Harga Saham Syariah Menggunakan Algoritma Long Short-Term Memory (LSTM) [Sharia Stock Price Prediction Using Long Short-Term Memory (LSTM) Algorithm]," *Jurnal Nasional Teknologi dan Sistem Informasi*, vol. 8, no. 3, pp. 164–172, jan 2022. [Online]. Available: <https://teknosi.fti.unand.ac.id/index.php/teknosi/article/view/2142>
- [16] H. Zhu, X. You, and S. Liu, "Multiple Ant Colony Optimization Based on Pearson Correlation Coefficient," *IEEE Access*, vol. 7, no. 1, pp. 61 628–61 638, 2019.
- [17] A. Hasan, O. Kalipsiz, and S. Akyoku, "Modeling Traders' Behavior with Deep Learning and Machine Learning Methods," *Complexity*, vol. 2020, no. 1, pp. 1–16, 2020. [Online]. Available: <https://dl.acm.org/doi/10.1155/2020/8285149>
- [18] A. D. W. Sumari, G. Budiprasetyo, M. Hani'ah, V. A. H. Firdaus, and R. Arianto, *Machine Learning*. Malang: Polinema Press, oct 2021.
- [19] M. Alida and M. Mustikasari, "Rupiah Exchange Prediction of US Dollar Using Linear, Polynomial, and Radial Basis Function Kernel in Support Vector Regression," *Jurnal Online Informatika*, vol. 5, no. 1, pp. 53–60, jul 2020. [Online]. Available: <http://join.if.uinsgd.ac.id/index.php/join/article/view/537>
- [20] A. Malar, P. Vincent, and H. Salleh, "An Investigation Into The Performance Of The Multilayer Perceptron Architecture Of Deep Learning In Forecasting Stock Prices," *Universiti Malaysia Terengganu Journal of Undergraduate Research*, vol. 3, no. 2, pp. 61–68, apr 2021. [Online]. Available: <https://journal.umt.edu.my/index.php/umtjur/article/view/205>
- [21] N. Wei, C. Li, X. Peng, F. Zeng, and X. Lu, "Conventional models and artificial intelligence-based models for energy consumption forecasting: A review," *Journal of Petroleum Science and Engineering*, vol. 181, no. 1, p. 106187, oct 2019.
- [22] A. A. Winoto and A. F. V. Roy, "Model of Predicting the Rating of Bridge Conditions in Indonesia with Regression and K-Fold Cross Validation," *International Journal of Sustainable Construction Engineering and Technology*, vol. 14, no. 1, pp. 249–259, feb 2023. [Online]. Available: <https://publisher.uthm.edu.my/ojs/index.php/IJSCET/article/view/12441>