Print ISSN: 2288-4637 / Online ISSN 2288-4645
doi:10.13106/jafeb.2021.vol8.no8.0399

A Prediction of Stock Price Movements Using Support Vector Machines in Indonesia

Ervandio Irzky ARDYANTA[1], Hasrini SARI[2]

Received: April 20, 2021 Revised: July 04, 2021 Accepted: July 15, 2021

Abstract

Stock movement is difficult to predict because it has dynamic characteristics and is influenced by many factors. Even so, there are some approaches to predict stock price movements, namely technical analysis, fundamental analysis, and sentiment analysis. Many researches have tried to predict stock price movement by utilizing these analysis techniques. However, the results obtained are varied and inconsistent depending on the variables and object used. This is because stock price movement is influenced by a variety of factors, and it is likely that those studies did not cover all of them. One of which is that no research considers the use of fundamental analysis in terms of currency exchange rates and the use of foreign stock price index movement related to the technical analysis. This research aims to predict stock price movements in Indonesia based on sentiment analysis, technical analysis, and fundamental analysis using Support Vector Machine. The result obtained has a prediction accuracy rate of 65,33% on an average. The inclusion of currency exchange rate and foreign stock price index movement as a predictor in this research which can increase average prediction accuracy rate by 11.78% compared to the prediction without using these two variables which only results in average prediction accuracy rate of 53.55%.

Keywords: Fundamental Analysis, Sentiment Analysis, Stock Prediction, Support Vector Machine, Technical Analysis

JEL Classification Code: E44, E47, G11, G17

1. Introduction

Stock is difficult to predict because its movement is affected by many factors and has dynamic characteristic. Even so, several techniques for predicting stock price movement have been developed (Masoud, 2017). There are two traditional techniques that are commonly used by investors. It forecasts stock price movements using past data such as opening and closing prices, transaction volume, average stock price, and so on. The second technique that is known as fundamental analysis uses qualitative and quantitative measurement based on the company's profile and financial condition, market condition, political, business, and economic climate (Hur et al., 2006). But with the development of informational technology and social media, there is also the third technique known as sentiment analysis (Derakhshan & Beigy, 2019; Sert et al., 2020). The sentiment is defined as perspective or opinion of a person – in this case, an investor – on information (Hu et al., 2012). Several researches have attempted to determine the relationship between sentiment and stock prices. Nguyen and Pham (2018) found that sentiment has significant effect on the stock market. Tetlock (2007) observed that adverse news reported in the Wall Street Journal might lead to a drop in the share price. Then, Tetlock et al. (2008) revealed that stock price movements in the United States are influenced by news sentiment which is provided by news media outlets such as Wall Street Journal and Dow Jones News Service.

Baker and Wurgler (2007) argued that, with the rapid development of computer capabilities, the concern now is how to assess investor sentiment and measure its impact through stock price prediction, rather than whether or not it impacts stock price. Artificial intelligence has been used by many studies to evaluate the effects of sentiment on the

[1]First Author and Corresponding Author. Researcher, Department of Industrial Engineering and Management, Faculty of Industrial Technology, Institut Teknologi Bandung, Indonesia [Postal Address: Jl. Ganesa 10, Coblong, Bandung, West Java, 40132, Indonesia] Email: ervandioirzky@student.itb.ac.id
[2]Researcher and Lecturer, Department of Industrial Engineering and Management, Faculty of Industrial Technology, Institut Teknologi Bandung, Indonesia. Email: hasrini@ti.itb.ac.id

stock price movements through prediction as computer capabilities have improved. Artificial intelligence-based stock price prediction can identify relationships and patterns in the variables, offering better results than traditional statistics (Maqsood, et al., 2020). The results of previous studies, on the other hand, varied depending on the object and variables analyzed. These inconsistencies are linked to the fact that each country's investor behaviour and capital market conditions are different (Corredor et al., 2015). Many factors can influence stock price changes, and the factors employed in those studies are thought to be insufficient to reflect all of them. As a result, to predict stock price movements, it is important to combine analysis techniques such as sentiment analysis with both technical and fundamental analysis.

According to Wu et al. (2012) combining multiple analysis techniques can enhance stock price prediction abilities. There is, however, no research that attempts to consider the use of fundamental analysis, such as currency exchange rates, and technical analysis in terms of stock price index movements in other countries to predict price fluctuations in the stock market, particularly in Indonesia. Several research on Indonesian stocks, such as by Afrianto et al. (2013), Rizkiana et al. (2017), as well as Yasin et al. (2014), only consider one aspect of the analysis into account, despite the fact that several factors might affect stock price movements. The use of just one aspect is considered to be insufficient to describe all available factors. Based on the preceding description, it is apparent that no research has been done to try to predict stock price movements in Indonesia using the three types of analysis discussed earlier, namely sentiment analysis, technical analysis, and fundamental analysis. So, the problem statement in this research is how to predict stock price movements in Indonesia based on sentiment analysis, technical analysis, and fundamental analysis using Support Vector Machine (SVM).

There are two objectives in this research. The first objective is to predict stock price movements in Indonesia based on sentiment analysis, technical analysis, and fundamental analysis using Support Vector Machine (SVM), and the second objective ought to measure the impact of sentiment analysis, technical analysis, and fundamental analysis on the stock price movement's prediction result.

2. Literature Review

According to Vijh et al. (2020), stock price prediction employing an artificial intelligence-based approach can enhance forecast accuracy by 60% to 86% when compared to traditional statistics. Li et al. (2020) used artificial intelligence approaches such as LSTM based on sentiment analysis to predict stock price movement in China, with prediction accuracy ranging from 40% to 80%. Even though in previous research it was found that investor sentiment affects stock prices, however Rizkiana et al. (2018) discovered that investor sentiment does not affect stock prices; rather, stock prices affect investor sentiment. This is because the study's data comes from an investor's forum, which contains discussion based on information as a type of reaction, rather than a response trigger that will generate sentiment, such as news. Rizkiana et al. (2019) investigated the effects of sentiment on stock price movements using news as one of their data sources. Nevertheless, this research yielded an insignificant result. This is due to the fact that the amount of news used in this study is very limited, and it only examines news linked to the object of study, thus it cannot represent all information or news on the stock market circulating in Indonesia. As a result, according to Rizkiana et al. (2019), the data sources utilized to examine the effect of sentiment must be carefully selected because not all data on the internet can be utilized to reflect investor sentiment.

Some studies, such as those conducted by Alshammari et al. (2020), rely only on historical closing prices to forecast stock market volatility. Meanwhile, another study such as by Picasso et al. (2019) attempted to predict stock price movement by combining investor sentiment analysis gathered from the news with technical analysis such as historical price, with forecast accuracy ranging from 50% to 68%. The result obtained by Picasso et al. (2019) is in line with a theory proposed by Wu et al. (2012) that stated combining various analysis techniques can enhance stock price prediction capabilities. So, to predict stock price movements, it is important to combine analysis techniques such as sentiment analysis with both technical and fundamental analysis. However, no research has been done to investigate the use of fundamental analysis in terms of USD-IDR exchange rates and technical analysis in terms of stock price index movements in other countries to predict the movement of stock prices, particularly in Indonesia.

According to Wong (2017), one of the most important variables impacting stock fluctuations is the currency exchange rate. According to Delgado et al. (2018), the strengthening of a country's currency exchange rate will affect the strengthening of that country's stock price index. Goh et al. (2021) mentioned that the exchange rate has a significant effect on the Indonesian stock market index. Dong and Yoon (2019) also mentioned that in emerging economies such as Indonesia, there is a linkage between stock price movement and exchange rate of currencies. Furthermore, besides currency exchange rate, previous studies have also not taken into account the use of stock price index movement in other countries, even though that Mensi et al. (2014) and Lee and Chou (2020) found that stock price index movement in one country, particularly in the United States, can influence the stock price movements in other countries, including Indonesia. There are several contributions in this research. First, unlike prior studies that used comments and opinions from social media to assess sentiment analysis, this study employs news about the economy, business, and politics that circulates in Indonesia as a data source. Furthermore, the amount of news data used in this study is larger than that used in Rizkiana et al. (2019), and it is viewed from a broader perspective, including both microeconomic and macroeconomic perspectives, so it should be able to represent overall investor sentiment toward news circulating in Indonesia. The next novelty is that this study attempts to use fundamental analysis in terms of USD-IDR exchange rate and technical analysis in terms of foreign stock price index movement as predictors in the prediction model. Finally, as far as we know, this is the first research to use Support Vector Machine (SVM) to forecast stock price changes in Indonesia based on sentiment analysis, technical analysis, and fundamental analysis.

3. Methodology

The dependent variable is a variable that is the primary focus of research (Situmorang, 2010). The stock price movements of nine firms listed on the Indonesia Stock Exchange with the largest market capitalization in each sector are the dependent variable in this study. These companies are Astra Agro Lestari (AALI) from the Agriculture sector, Astra International (ASII) from the Miscellaneous Industry sector, Bank Central Asia (BBCA) from the Finance sector, Merdeka Copper Gold (MDKA) from the Mining sector, Pakuwon Jati (PWON) from the Property, Real Estate, and Building Construction sector, Telekomunikasi Indonesia (TLKM) from the Infrastructure, Utilities, and Transportation sector, Chandra Asri Petrochemical (TPIA) from the Basic Industry and Chemicals sector, United Tractors (UNTR) from the Trade, Service, and Investment sector, as well as Unilever Indonesia (UNVR) from the Consumer Goods Industry sector. Meanwhile, predictor variable is defined as a variable that is used to predict another variable or the outcome. (Salkind, 2010; Williams & Levitas, 2019).

The first predictor variable in this study is sentiment analysis, which is news sentiment collected from online news media Twitter accounts like CNBC Indonesia. This news will be analyzed to determine whether it contains positive information that can increase stock prices or negative information that can decrease stock prices. The historical price of stocks, which includes the opening price, closing price, and transaction volume, with technical indicators such as the Moving Average 5-Period which is abbreviated as MA5, Money Flow Index which is abbreviated as MFI, and Relative Strength Index which is abbreviated as RSI is the second and third variable that associated with technical analysis. Technical indicators are calculated using historical price data with a certain formula. The fourth variable which is also related to the technical analysis consists of foreign stock price index movements comprised of DJI Index (Dow Jones Industrial Average from USA), FTSE (FTSE 100 Index from UK), GSPC (S&P 500 Index from USA), HSI (Hang Seng Index from Hong Kong), IXIC (NASDAQ from USA), N225 (Nikkei 225 Index from Japan), and SSEC (Shanghai Composite Index from Shanghai). The fifth variables cover aspects of fundamental analysis consisting of USD-IDR exchange rate data. Data regarding nine company's stock price movements, historical prices, currency exchange rates, and foreign stock price index movements are obtained from Yahoo Finance website. Furthermore, using the Python programming language, data processing is carried out to predict stock price movements in Indonesia based on sentiment analysis, technical analysis, and fundamental analysis.

3.1. Historical Stock Prices

The historical price data for each company consists of opening price, closing price, and transaction volume. The data is collected from 6 July 2020 to 11 January 2021 or equal to 124 Indonesia Stock Exchange's transaction days. The most recent historical price data, particularly historical prices at time $t-1$, is utilized to predict stock price movements at time t.

3.2. Foreign Stock Price Index Movements

Foreign stock price index movement data used in this research were taken from several countries. Foreign stock price index movements are defined as 1 if today's closing price (t) closes higher than the closing price on the previous day ($t-1$) or in other words the index price rises. Meanwhile, index movements are defined as 0 if the index's closing price today (t) is less than or equal to the previous day's closing price ($t-1$), or if the index has decreased or remained unchanged from the previous day's closing price. The most recent foreign stock price index movement data, specifically the index price at time $t-1$, is utilized to anticipate the movement of stock prices at time t.

3.3. Natural Language Processing (NLP)

One of the branches of Artificial Intelligence (AI) is Natural language processing (NLP) that is used in sentiment analysis to identify and process human language through computers. In this study, NLP is used to classify news related to economics, business, and politics taken from news media based on the sentiment category, whether the news contains positive information that can increase stock prices or contains negative information that can decrease stock prices. News is collected from CNBC Indonesia's

Twitter account from 6 July 2020 to 11 January 2021 or equal to 124 transaction days of the Indonesia Stock Exchange. 33.990 news items are collected over a period of 124 days. After the news is collected it is grouped per day to determine the daily sentiment. Indonesia Stock Exchange's trading hours start at 09.00 WIB until 15.00 WIB, so news published after 15.00 WIB will be categorized as news for the next day ($t + 1$) because sentiment towards news that is published after trading hours for that day is closed (t) will be responded during trading hours on the next day ($t + 1$). The same applies to the news that published on stock market holidays (Saturday, Sunday, and national holidays), where news that is published on holidays will be categorized as news for the next transaction days after the holiday ends. Furthermore, news that has been grouped per day then will be processed with an NLP algorithm using Python programming language to find out the sentiment for each news, whether it is positive or negative. Neutral sentiment will not be used because it does not affect stock price movements. After each news sentiment is known, the amount of news for each sentiment category will be calculated using percentage.

Table 1 is an example of news data that has been grouped per day and the percentage calculated for each sentiment category.

3.4. Technical Indicators

In addition to historical prices, technical analysis in this research also considers the use of technical indicators. The calculation of technical indicators consists of MA5, MFI, and RSI. The data used in the calculation of technical indicators come from historical prices such as opening prices, closing prices, transaction volumes, and so on which are then calculated using a certain formula to produce a number that shows the trend towards the movement of a stock. For example, an RSI value above 70 indicates overbought, so there is a tendency that the price will move down. The calculation of technical indicators is as follows.

Table 1: Example of Daily News Sentiment Calculation Results

Date	Negative Sentiment Percentage	Positive Sentiment Percentage
06/07/2020	0.52	0.48
07/07/2020	0.29	0.71
08/07/2020	0.33	0.67
09/07/2020	0.47	0.53
10/07/2020	0.72	0.28

3.4.1. MA5 (Moving Average 5-Period)

A moving average is a variety of technical indicator derived by averaging a stock's closing price across time. The 5-period used in this research shows that the average calculation is carried out on the closing price of a stock for the past 5 days. The Moving Average 5-Period formula is as follows (Chakrabarti, 2004).

$$\text{Moving Average 5-Period} = \frac{t_{-4} + t_{-3} + t_{-2} + t_{-1} + t}{5} \tag{1}$$

3.4.2. MFI (Money Flow Index)

The Money Flow Index, abbreviated as MFI, is derived using the closing price, transaction volume, highest price, and lowest price to assess the purchasing or selling pressure of a stock over a certain time period. The MFI calculation in this study uses a 14-days transaction period. The 14-day period is the period that is usually used in MFI calculations because it is needed to determine the tendency of buying or selling pressure on a stock, it is necessary to consider a longer transaction period. If the period is too short, then the signals obtained will be premature and on the other hand if the period is too long, then the signals obtained will be too slow/delayed. Therefore, the 14-days period is considered the most ideal period in the calculation of MFI. The MFI formula is as follows (Marek & Čadková, 2020).

$$\text{Typical Price} = \frac{\text{High} + \text{Low} + \text{Close}}{3} \tag{2}$$

$$\text{Raw Money Flow} = \text{Typical Price} \times \text{Volume} \tag{3}$$

$$\text{Ratio of Money Flow} = \frac{\text{14-period Positive Money Flow}}{\text{14-period Negative Money Flow}} \tag{4}$$

$$\text{Money Flow Index} = 100 - \frac{100}{1 + \text{Money Flow Ratio}} \tag{5}$$

3.4.3. RSI (Relative Strength Index)

The Relative Strength Index, abbreviated as RSI, is a technical indicator that assesses the momentum and trend direction of a stock's price movement by calculating its previous closing price. The calculation period that is generally used in RSI is 14 days, the reason is the same as in the MFI calculation, namely to get a more precise price movement signal, not premature or delayed. The following is the formula for RSI (Wilder, 1978).

$$\text{RS} = \frac{\text{Average Gain}}{\text{Average Loss}} \tag{6}$$

$$\mathrm{RSI} = 100 - \frac{100}{1 + \mathrm{RS}} \tag{7}$$

3.5. Data Pre-processing

Before it can be processed further, pre-processing of data is needed to be done as a form of data preparation. In this research, data pre-processing was divided into three primary steps: variable assignment, dataset splitting into training and testing sets, and feature scaling.

3.5.1. Variable Assignment

At this stage, the predictor variable (X) and the dependent variable (y) are determined. Predictor variables consist of sentiment analysis, historical stock prices, technical indicators, USD-IDR exchange rate, and foreign stock price index movements. The dependent variable consists of stock price movements of 9 companies that are the object of research.

3.5.2. Dataset Splitting into Training and Testing Sets

The data will be divided into two groups: training and testing. The training set is a collection of data that the algorithm will use to build a simulation model, whereas the test set is a collection of data that the algorithm will use to build a trial model. This prediction model was built using 70% of the total data as a training set and 30% as a test set.

3.5.3. Feature Scaling

Feature scaling is used to equalize the scale of the variable using the z-score standardization. For example, in sentiment analysis the measurement scale ranges from 0 to 1 because it is a percentage, while in historical prices the measurement scale is in thousands. Processing on data that has different measurement scales can cause bias, therefore a feature scaling is needed to equalize the measurement scale of all variables using the standardized z-score calculation. The feature scaling formula using the z-score standardization is as follows.

$$z = \frac{X - \mu}{\sigma} \tag{8}$$

By performing feature scaling, all variables' values will be normalized to have a standard deviation of 1 and an average of 0 (Buitinck et al., 2011). In addition, feature scaling also functions so that the data can meet the SVM assumption, which is independent and similarly distributed (Independent and Identically Distributed/IID).

3.6. Support Vector Machine (SVM) Model Building

The core notion of the Support Vector Machine (SVM), according to Cortes and Vapnik (1995), is to employ a hyperplane as a separator between data in an n-dimensional search space with distinct categories. The Support Vector Machine (SVM) was used to build the prediction model because it is not readily impacted by data outliers and is not prone to overfitting, which is common in historical stock price data due to substantial price increases or decreases at a certain period. The SVM method is also relatively free of assumptions so that there is no need for excessive manipulation and validation of the data. In addition, the SVM method was chosen because according to Bustos and Pomares-Quimbaya (2020), Nti et al. (2020), Gandhmal and Kumar (2019) the SVM method is the most widely used classification method in predicting stock price movements because of its ability to understand complex stock price movement data patterns and can provide better predictive results compared to other methods. The kernel function used in the construction of the SVM model is a linear kernel. The linear kernel is used because based on several experiments using other kernel functions such as RBF or sigmoid, the linear kernel is able to provide the highest level of prediction accuracy. The objective of constructing the Support Vector Machine algorithm model is to build a model that can identify patterns and categorize stock price movement direction categories, such as whether the stock will rise or decrease, based on the pattern generated between the predictor variables. For example, based on the data obtained, if the previous day's stock price, the USD-IDR exchange rate, and the stock index in the United States strengthen, the stock price would increase. The SVM algorithm will analyze this pattern so that if there is a strengthening of the previous day's stock price, the USD-IDR exchange rate, and the stock index in the United States in the future, the SVM algorithm will be able to anticipate the increase in stock prices.

4. Results and Discussion

4.1. Prediction Results

The accuracy rate gained from the prediction of stock price movements in each firm that is the subject of study is analyzed to assess the consistency of the prediction model performance in each industrial sector. Table 2 below summarizes the training set and test set accuracy rates acquired from stock price prediction in each firm.

According to Schumaker and Chen (2009), Si et al. (2013), and Tsibouris & Zeidenberg (1995), predictions can be said to be satisfactory if the accuracy rate is more than 56%. Based on the test results in this study, the training set

Table 2: Stock Price Prediction Accuracy Rate

Company Name	Stock Ticker	Accuracy Rate (Training Set)	Accuracy Rate (Test Set)
Astra Agro Lestari	AALI	65%	66%
Astra International	ASII	79%	71%
Bank Central Asia	BBCA	70%	65%
Merdeka Copper Gold	MDKA	71%	68%
Pakuwon Jati	PWON	81%	68%
Telkom Indonesia	TLKM	77%	63%
Chandra Asri Petrochemical	TPIA	72%	61%
United Tractors	UNTR	72%	63%
Unilever Indonesia	UNVR	79%	63%
Average Accuracy Rate		73.89%	65.33%

and the test set is acquired on an average prediction accuracy rate of 73.89% and 65.33%, respectively. This accuracy rate exceeds the accuracy limit of 56%, so it can be concluded that the average accuracy rate obtained in the SVM prediction model can be stated to be satisfactory.

Furthermore, a retest without the USD-IDR exchange rate and foreign stock index movement variables was performed to assess the increase in accuracy rate contributed by these two variables to the average accuracy rate of the prediction results. The accuracy rate obtained in the retest without using currency exchange rate and foreign stock price index variables is then compared with the level of accuracy in testing using these two variables to determine the increase in the accuracy rate. Table 3 compares the accuracy rate of prediction with and without the USD-IDR exchange rate and foreign stock index movement variables. See Table 3.

Based on the prediction of stock price movements without using currency exchange rates and foreign stock price index movements as predictors, the average prediction accuracy rate obtained is 53.55% as shown in Table 3. When compared to the average accuracy rate utilizing foreign stock index and USD-IDR exchange rates variables, this accuracy rate is lower by 11.78% . So, combining the foreign stock index movement and USD-IDR exchange rate variables in predicting stock price movements enhances prediction accuracy by 11.78% when compared to not using these two variables.

In comparison to the past studies appears that none of them have attempted to investigate the use of fundamental analysis in terms of the usage of the USD-IDR currency exchange rate variable as a predictor. Whereas according to Hur et al. (2006), fundamental analysis is one of the important analysis that must be considered in predicting stock price movements. In addition, previous studies have not considered the use of technical analysis aspects in terms of stock index movements in other countries, whereas Mensi et al. (2014) and Lee and Chou (2020) in their research stated that stock index movements in other countries, especially in the United States, can affect stock price movements in other countries, including Indonesia. As a result, this paper aimed to address a research gap by addressing the use of fundamental analysis, as well as technical analysis and sentiment analysis, as predictors of stock price movements.

Table 3: Accuracy Rate Comparison With and without using USD-IDR Exchange Rate and Foreign Stock Index Movements

Company Name	Stock Ticker	With Variables	Without Variables
Astra Agro Lestari	AALI	66%	61%
Astra International	ASII	71%	55%
Bank Central Asia	BBCA	65%	53%
Merdeka Copper Gold	MDKA	68%	61%
Pakuwon Jati	PWON	68%	55%
Telkom Indonesia	TLKM	63%	58%
Chandra Asri Petrochemical	TPIA	61%	61%
United Tractors	UNTR	63%	55%
Unilever Indonesia	UNVR	63%	53%
Average Accuracy Rate		65.33%	53.55%

4.2. Sensitivity Analysis

In this research, sensitivity analysis was carried out to determine changes in the accuracy rate when predictions

Table 4: Sensitivity Analysis using Cross-Validation

Company Name	Average Accuracy Rate	Standard Deviation	Minimum Accuracy Rate	Maximum Accuracy Rate
Astra Agro Lestari (AALI)	60%	+/– 9%	51%	69%
Astra International (ASII)	67%	+/– 6%	61%	73%
Bank Central Asia (BBCA)	66%	+/– 5%	61%	71%
Merdeka Copper Gold (MDKA)	71%	+/– 5%	66%	76%
Pakuwon Jati (PWON)	66%	+/– 2%	64%	68%
Telkom Indonesia (TLKM)	63%	+/– 5%	58%	68%
Chandra Asri Petrochemical (TPIA)	60%	+/– 9%	51%	69%
United Tractors (UNTR)	63%	+/– 7%	56%	70%
Unilever Indonesia (UNVR)	61%	+/– 5%	56%	66%

were made using different data sets. Sensitivity analysis was accomplished using a cross-validation technique which divides the data into several parts and tested each part as if it were new data. By performing a sensitivity analysis, it will be known the standard deviation of the accuracy rate if predictions are made on different new data sets. Table 4 below provides a summary of the sensitivity analysis for each company.

According to Table 4, the standard deviation of predictions differs by firm, with Pakuwon Jati (PWON) having the least standard deviation of 2% and Astra Agro Lestari (AALI) and Chandra Asri Petrochemical (TPIA) having the highest standard deviations of 9%. The average accuracy rate minus the standard deviation is the minimum accuracy rate, whereas the average accuracy rate plus the standard deviation is the maximum accuracy rate. A low standard deviation means that the predicted accuracy rate gained when testing new data will be more consistently close to the average, whereas a large standard deviation means that the accuracy rate achieved will be less consistent and further away from the average (Shane, 2008). Therefore, it can be concluded that the prediction accuracy rate in Pakuwon Jati (PWON) is more consistent than the prediction accuracy rate in Astra Agro Lestari (AALI) and Chandra Asri Petrochemical (TPIA).

4.3. Prediction Simulation and Validation

Prediction simulations are accomplished on new data outside of the datasets that have been used in making and testing prediction models. This simulation aims to validate and at the same time find out the consistency and performance of the SVM model when predictions are made on completely new data. The data used for the simulation is taken from 5 April 2021 to 7 May 2021 or equal to 25 Indonesia Stock Exchange's transaction days. Based on the simulation results, the prediction accuracy level is shown in Table 5.

Table 5: Simulation Results

Company Name	Stock Ticker	Accuracy Rate
Astra Agro Lestari	AALI	68%
Astra International	ASII	72%
Bank Central Asia	BBCA	64%
Merdeka Copper Gold	MDKA	72%
Pakuwon Jati	PWON	68%
Telkom Indonesia	TLKM	64%
Chandra Asri Petrochemical	TPIA	60%
United Tractors	UNTR	64%
Unilever Indonesia	UNVR	60%
Average Accuracy Rate		65.78%

Based on the results of the prediction simulation carried out for each company using new data as shown in Table 5, the average prediction accuracy rate is 65.78%. This value is not much different from the average prediction accuracy rate obtained from the prior SVM model using the original dataset, which is 65.33%, so it can be concluded that the SVM model in this study is able to provide consistent performance in predicting stock price movements.

5. Conclusion

This research predicts the stock price movements of nine companies with the largest market capitalization in each industrial sector listed on the Indonesia Stock Exchange using Support Vector Machine (SVM) and the study obtained the average prediction accuracy rate of 65.33%.

This accuracy rate of 65.33% is greater than the minimum accuracy rate set by Schumaker and Chen (2009), Si et al. (2013), and Tsibouris & Zeidenberg (1995) that is equal to 56%, so it can be concluded that the accuracy rate obtained using SVM prediction model in this research can be stated to be satisfactory. This study addresses a gap in prior research by taking into account the use of technical analysis in terms of foreign stock price index and fundamental analysis in terms of the USD-IDR currency exchange rate as well as sentiment analysis in predicting stock price movements. In addition, related to the sentiment analysis, this study also uses news both in the macro and microeconomic scope in large numbers compared to previous studies, so that it is able to represent the overall sentiment related to stocks circulating in Indonesia.

Furthermore, to determine the increase in the average accuracy rate provided by the use of USD-IDR exchange rate and foreign stock index variables, a comparison is made on the prediction results using and without using these two variables based on the same data. The accuracy rate in the test without the foreign stock price index and USD-IDR exchange rate variables is 53.55%. As a result, it is possible to conclude that using foreign stock price index and USD-IDR exchange rate variables can improve average prediction accuracy by 11.78%. Then to validate the performance of the prediction model, prediction simulation was also carried out using new data that had never been used before and obtained an average accuracy rate of 65.78%. This value is not much different from the average prediction accuracy rate obtained from the prior SVM model using the original dataset, which is 65.33%, so it can be concluded that the SVM model in this study is able to provide consistent performance in predicting stock price movements.

References

Afrianto, R. B., Tjandrasa, H., & Arieshanti, I. (2013). Stock Price Prediction using Back Propagation Neural Network Method. *Jurnal Simantec, 3*(3), 132–141. https://doi.org/10.21107/simantec.v4i1

Alshammari, T. S., Ismail, M. T., Al-Wadi, S., Saleh, M. H., & Jaber, J. J. (2020). Modeling and Forecasting Saudi Stock Market Volatility Using Wavelet Methods. *The Journal of Asian Finance, Economics and Business, 7*(11), 83–93. https://doi.org/10.13106/jafeb.2020.vol7.no11.083

Baker, M., & Wurgler, J. (2007). Investor sentiment in the stock market. *Journal of Economic Perspectives, 21*(2), 129–152. https://doi.org/10.1257/jep.21.2.129

Buitinck, L., Louppe, G., Blondel, M., Pedregosa, F., Mueller, A., Grisel, O., & Varoquaux, G. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research, 12*(85), 2825–2830. https://dl.acm.org/doi/10.5555/1953048.2078195

Bustos, O., & Pomares-Quimbaya, A. (2020). Stock market movement forecast: A Systematic review. *Expert Systems With Applications, 156*(15), 113464. https://doi.org/10.1016/j.eswa.2020.113464

Chakrabarti, J. (2004). *ISC Business Mathematics.* Mumbai: Allied Publishers Private Limited.

Corredor, P., Ferrer, E., & Santamaria, R. (2015). The Impact of Investor Sentiment on Stock Returns in Emerging Markets: The Case of Central European Markets. *Eastern European Economics, 53*(4), 328–355. https://doi.org/10.1080/00128775.2015.1079139

Cortes, & Vapnik. (1995). Support-vector networks. *Mach. Learn, 20*(1995), 273–297. https://doi.org/10.1007/BF00994018

Delgado, N. A., Delgado, E. B., & Saucedo, E. (2018). The relationship between oil prices, the stock market and the exchange rate: Evidence from Mexico. *North American Journal of Economics and Finance, 45*(C), 266–275. https://doi.org/10.1016/j.najef.2018.03.006

Derakhshan, A., & Beigy, H. (2019). Sentiment analysis on stock social media for stock price movement. *Engineering Applications of Artificial Intelligence, 85*(C), 569–578. https://doi.org/10.1016/j.engappai.2019.07.002

Dong, X., & Yoon, S.-M. (2019). What global economic factors drive emerging Asian stock market returns? Evidence from a dynamic model averaging approach. *Economic Modelling, 77*(C), 204–215. https://doi.org/10.1016/j.econmod.2018.09.003

Gandhmal, D. P., & Kumar, K. (2019). Systematic analysis and review of stock market prediction techniques. *Computer Science Review, 34*, 100190. https://doi.org/10.1016/j.cosrev.2019.08.001

Goh, T. S., Henry, H., & Albert, A. (2021). Determinants and Prediction of the Stock Market during COVID-19: Evidence from Indonesia. *The Journal of Asian Finance, Economics and Business, 8*(1), 1–6. https://doi.org/10.13106/jafeb.2021.vol8.no1.001

Hu, N., Bose, I., Koh, N. S., & Liu, L. (2012). Manipulation of online reviews: An analysis of ratings, readability, and sentiments. *Decision Support Systems, 52*(3), 674–684. https://doi.org/10.1016/j.dss.2011.11.002

Hur, J., Raj, M., & Riyanto, Y. E. (2006). Finance and trade: A cross-country empirical analysis on the impact of financial. *World Development, 34*(10), 1728–1741. https://doi.org/10.1016/j.worlddev.2006.02.003

Lee, C.-H., & Chou, P.-I. (2020). Structural breaks in the correlations between Asian and US stock markets. *The North American Journal of Economics and Finance, 51*(C), 101087. https://doi.org/10.1016/j.najef.2019.101087

Li, Y., Bu, H., Li, J., & Wu, J. (2020). The role of text-extracted investor sentiment in Chinese stock The role of text-extracted investor sentiment in Chinese stock. *International Journal of Forecasting, 36*(4), 1541–1562. https://doi.org/10.1016/j.ijforecast.2020.05.001

Maqsood, H., Mehmood, I., Maqsood, M., Yasir, M., Afzal, S., Aadil, F., & Muhammad, K. (2020). A local and global event sentiment based efficient stock exchange forecasting using deep learning. *International Journal of Information Management, 50*(C), 432–451. https://doi.org/10.1016/j.ijinfomgt.2019.07.011

Marek, P., & Čadková, V. (2020). Optimization and Testing of Money Flow Index. *APLIMAT 2020 - 19th Conference on Applied Mathematics*.

Masoud, N. M. (2017). The impact of stock market performance upon economic growth. *International Journal of Economics and Financial Issues, 3*(4), 788–798.

Mensi, W., Hammoudeh, S., Reboredo, J. C., & Nguyen, D. K. (2014). Do global factors impact BRICS stock markets? A quantile regression approach. *Emerging Markets Review, 19*(C), 1–17. https://doi.org/10.1016/j.ememar.2014.04.002

Nguyen, D. D., & Pham, M. C. (2018). Search-based Sentiment and Stock Market Reactions: An Empirical Evidence in Vietnam. *The Journal of Asian Finance, Economics and Business, 5*(4), 45–56. https://doi.org/10.13106/jafeb.2018.vol5.no4.45

Nti, I. K., Adekoya, A. F., & Weyori, B. A. (2020). A systematic review of fundamental and technical analysis of stock market predictions. *Artificial Intelligence Review, 53*, 3007–3057. https://doi.org/10.1007/s10462-019-09754-z

Picasso, A., Merello, S., Ma, Y., Oneto, L., & Cambria, E. (2019). Technical analysis and sentiment embeddings for market trend. *Expert Systems with Applications, 135*(C), 60–70. https://doi.org/10.1016/j.eswa.2019.06.014

Rizkiana, A., Sari, H., Hardjomidjojo, P., & Prihartono, B. (2019). The development of composite sentiment index in Indonesia based on the internet-available data. *Cogent Economics & Finance, 7*(1), 1669399. https://doi.org/10.1080/23322039.2019.1669399

Rizkiana, A., Sari, H., Hardjomidjojo, P., Prihartono, B., Sunaryo, I., & Prasetyo, I. R. (2018). Lead-Lag Relationship between Investor Sentiment in Social Media, Investor Attention in Google, and Stock Return. *Thirteenth International Conference on Digital Information Management (ICDIM 2018)*, 204–209. https://doi.org/10.1109/ICDIM.2018.8847094

Rizkiana, A., Sari, H., Hardjomijojo, P., Prihartono, B., & Yudhistira, T. (2017). Analyzing the impact of investor sentiment in social media to stock return: Survival analysis approach. *2017 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, 519–523. https://doi.org/10.1109/IEEM.2017.8289945.

Salkind, N. J. (2010). *Encyclopedia of Research Design.* Thousand Oaks: Sage Publications, Inc.

Schumaker, R. P., & Chen, H. (2009). A quantitative stock prediction system based on financial news. *Information Processing and Management, 45*(5), 571–583. https://doi.org/10.1016/j.ipm.2009.05.001

Sert, O. C., Şahin, S. D., Özyer, T., & Alhajj, R. (2020). Analysis and prediction in sparse and high dimensional text data: The case of Dow Jones stock market. *Physica A, 545*, 123752. https://doi.org/10.1016/j.physa.2019.123752

Shane, J. M. (2008). *Developing a Performance Management Model: Your Action Guide to What Every Chief Executive Should Know: Using Data to Measure Police Performance.* New York City: Looseleaf Law Publications, Inc.

Si, J., Mukherjee, A., Liu, B., Li, Q., Li, H., & Deng, X. (2013). Exploiting topic based twitter sentiment for stock prediction. *Proceedings of the 51st annual meeting of the association for computational linguistics, 2*, 24–29. https://doi.org/10.13140/2.1.3604.7043

Situmorang, S. H. (2010). *Data Analysis for Management and Business Research.* Medan: USU Press.

Tetlock, P. C. (2007). Giving Content to Investor Sentiment: The Role of Media in the Stock Market. *The Journal of Finance, 62*(3), 1139–1168. https://doi.org/10.1111/j.1540-6261.2007.01232.x

Tetlock, P., Saar-Tsechansky, M., & Macskassy, S. (2008). More Than Words: Quantifying Language to Measure Firms' Fundamentals. *Journal of Finance, 63*(3), 1437–1467. http://dx.doi.org/10.2139/ssrn.923911

Tsibouris, G., & Zeidenberg, M. (1995). Testing the efficient markets hypothesis with gradient descent algorithms. *Neural networks in the capital markets, 8*(10), 127–136. https://doi.org/10.1371/journal.pone.0078177

Vijh, M., Chandola, D., Tikkiwal, V. A., & Kumar, A. (2020). Stock Closing Price Prediction using Machine Learning Techniques. *Procedia Computer Science, 167*, 599–606. https://doi.org/10.1016/j.procs.2020.03.326

Wilder, J. W. (1978). *New Concepts in Technical Trading Systems.* Edmonton, Canada: Trend Research Inc.

Williams, Y., & Levitas, J. (2019). Predictor Variable: Definition & Example. *Study.com.* Retrieved May 24, 2021 from: https://study.com/academy/lesson/predictor-variable-definition-example.html

Wong, H. T. (2017). Real exchange rate returns and real stock price returns. *International Review of Economics and Finance, 49*, 340–352. https://doi.org/10.1016/j.iref.2017.02.004

Wu, J., Su, C., Yu, L., & Chang, P. (2012). Stock price predication using combinational features from sentimental analysis of stock news and technical analysis of trading information. *International Proceedings of Economics Development and Research, 55*, 39–43. https://doi.org/10.7763/IPEDR.2012.V55.8

Yasin, H., Prahutama, A., & Utami, T. W. (2014). Stock Price Prediction using Support Vector Regression with Grid Search Algorithm. *Media Statistika, 7*(1), 29–35. https://doi.org/10.14710/medstat.7.1.29-35