Optimization Selection on Deep Learning Algorithm for Stock Price Prediction in Indonesia Companies

Gunawan1*, Wresti Andriani2, Sawaviyya Anandianska3, Aang Alim Murtopo4, Bangkit Indarmawan Nugroho5, Naella Nabila Putri Wahyuning Naja6

1,2,3,4Department of Informatic Engineering, STMIK YMI Tegal, Indonesia
4Department of Information System, STMIK YMI Tegal, Indonesia
6Department of Management, Universitas Negeri Semarang, Indonesia

Abstract. Purpose: Share price movements after the COVID-19 pandemic experienced a decline in several sectors, especially in the share prices of the Aneka Tambang Company, which operates in the mining sector, the Wijaya Karya Company in the construction sector, and the Sinar Mas Company, which is a Holding Company. Several factors influence this, including investors' hesitation in investing their money. This research aims to predict stock price movements using a Deep Learning algorithm, which is optimized using Selection optimization at three large companies in Indonesia, namely PT. ANTIM, PT. WIKA, and PT. SINAR MAS. So that it can provide the correct information to investors to avoid losses.

Method: research through collecting data from the three companies, preprocessing, and then analyzing research data with several alternatives. The combination of inputs from the three companies using the deep learning method is then optimized using selection optimization to produce the best accuracy and use the results of the RMSE evaluation.

Results: The results of this research show that by using the Deep Learning method, the best evaluation results were obtained for the Company PT Wijaya Karya with an RMSE value of 0.432, an MAE value of 0.31505 and an MSE value of 1913.953. These results were then optimized using Selection optimization to obtain an RMSE increase of 0.022, namely 0.410.

Novelty: The contribution of this research is to get the best combination of input variables obtained using the windowing process from the three companies, which are then processed using the Deep Learning method to produce the most accurate evaluation results from the three companies, then the results are optimized again using Selection optimization to get the more optimal results.

Keywords: Deep learning, Input variable selection, Selection optimization, Stock price

Received October 2023 / Revised November 2023 / Accepted February 2024

This work is licensed under a Creative Commons Attribution 4.0 International License.

INTRODUCTION

Stock investment activities in the post-pandemic period are things that make investors hesitate because there are still many companies that are still recovering from the downturn [1]. Stock price movements fluctuate and are difficult to predict, making it challenging to predict ups and downs [2]. Several factors can make investors do activities on stocks, and some even make it as income [3]. These factors include external and internal factors; some even connect with weather factors and investor moods in increasing or decreasing stock prices [4].

These internal variables include operational expenses, while external factors include net interest margin, return on equity, and return on assets [5]. In manufacturing companies, internal factors are influenced by micro factors, namely marketing announcements, financing, management announcements of directors, diversification announcements, investment announcements, employment announcements, and financial announcements of the Company, which are generally affected due to the COVID-19 pandemic [6], [7]. At the same time, external factors are influenced by external factors [8], such as announcements of a country's government, legal announcements, industry announcements, the political turmoil of a country, fluctuations in currency values, and various issues from outside and within the country [9].

* Corresponding author.
Email addresses: gunawan.gayo@gmail.com (Gunawan)
DOI: 10.15294/sji.v11i1.47935
The movement of a company's stock price powerfully provokes investors to examine carefully and carefully so that the decision to buy and sell a form of security can generate profits and avoid losses, even though internal factors also affect stock price movements [10]. This is one of the reasons why share price movements cannot be predicted with certainty. So, it is essential to research to minimize losses.

Deep learning is a computational technology to change paradigms in data processing, pattern recognition, and understanding complex systems [11, 12, 13, 14]. This method is based on the architecture of deep artificial neural networks [15], inspired by the structure and function of human neural networks [16]. These artificial neural networks consist of many layers (called hidden layers”) [17, 18], which allows these algorithms to extract increasingly complex features from input data. This significantly contrasts conventional methods, where feature extraction is generally done manually.

Several studies predict the movement of a company's stock price [19], predictions are made using SVM and generate RMSE of 20.281, as well as research [20] which examines stock prices of shipping companies using linear regression and produces RMSE of 7.522, as well as research [1], [21], [22], [23] which predicts the sentiment of analysts on stock prices that use SVM which produces low data.

In the previous research above, there are weaknesses, for example, in the linear regression method to prediction, which is that this method is only limited to linear relationships but cannot show non-linear relationships without additional transformations that combine non-linear terms [24], [25]. At the same time, the Support Vector Machine has the disadvantage that it is unsuitable for large data sets because it requires a long training time [26], [27], [28], [29]. Some studies have proven that the evaluation results using the RMSE Support Vector Machine are still more significant than those using Linear Expression [30], [31].

This study will use Deep Learning algorithms that are optimized using Selection optimization to predict stock price movements in three major companies in Indonesia, namely PT. ANTAM [32], PT. WIKA [33], and PT. SINAR MAS [34] will be affected by the COVID-19 pandemic. This study aims to help investors predict stock price movements to avoid losses and determine whether this algorithm can predict stock price movements well, produce good accuracy, and be more optimal after being optimized using Selection Optimization.

**METHODS**

The research method that was carried out appears in Figure 1.

![Research methods](image)

Figure 1 shows that the first step of this study is data collection or data collection. This data is obtained online from the Yahoo.com page of stock price movements of three major companies in Indonesia. PT. ANTAM, PT. WIKA, and PT. SINAR MAS From Historical data from 2020 to [35]–[37]. The results obtained are in Excel form. This period was taken because, after the coronavirus pandemic in 2020, there must have been many companies that were affected and until now are still not healthy, bounce back. The data generated was 73 from each Company.

Data preprocessing is done after getting the dataset used [38], including filtering by changing the data arrangement from 2023 to 2020. The data variables obtained are seven: open, high, low, close, adj close,
and volume. Then, define the target or data label. This study focused on the Close variable. Close Price is the closing price of a stock [39], which measures several pieces of information that affect investor decisions at the close of trading and the Price of the following stock [40].

The next activity is windowing data [41], which converts univariate data from target data into multivariate data. The technique is sliding windows with daily or daily windowing size [42].

The next step is dividing training data and testing data [43]. This division is done with the 10-fold cross-validation model, a nested operator [44]. The validation concept divides the training data into ten equal parts and learns ten times [45]. Each time, it is selected on another part of the data set for learning, as many as nine, and the rest are used as testing. Then, the average and deviation values of 10 test results were calculated [46].

The next stage is to choose an evaluation model to measure the accuracy of the deep learning algorithm using Mean Absolute Error (MAE) and Root Mean Square Error (RMSE), where the lower the value, the better the level of accuracy. The accuracy of predictions is determined from the smallest value of each data accuracy method [47].

RESULTS AND DISCUSSIONS

The data to be used as a data set in this study is taken online from Yahoo.com on the stock data of three companies in Indonesia, namely PT. ANTAM, PT. WIKA, and PT. SINAR MAS, in Historical Data from 2020 to 2023, as many as 730. Data are obtained in Table 2. This research uses a computer with a Core-i7 processor specification, 512 GB SSD, 16 GB memory, and a 21-inch monitor. The software used is RapidMiner Studio version 10.

<table>
<thead>
<tr>
<th>Date</th>
<th>PT. ANTAM</th>
<th>PT. WIKA</th>
<th>PT. SINAR MAS</th>
</tr>
</thead>
<tbody>
<tr>
<td>8/22/23</td>
<td>376</td>
<td>378</td>
<td>370</td>
</tr>
<tr>
<td>8/31/20</td>
<td>1300</td>
<td>1305</td>
<td>1220</td>
</tr>
</tbody>
</table>

The dataset in Table 1 is obtained from the scribbling results on the page PT. ANTAM, PT. WIKA, and PT. SINAR MAS [35]–[37]. The data is then prepared to become data ready to be processed through data preprocessing. The variables obtained consist of 7 data: Open, High, Low, Close, Adj Close, and Volume. This study used close variables as targets or labels, as shown in Table 2.
By using the close variable as the target, a sliding window process is carried out, namely to convert univariate data into multivariate data by making data before the target data into event data the day before (H-1), two days back (H-2) and three days earlier (H-3), into data that will become training data, the results are as in Table 3.

### Table 3. Sliding window results

<table>
<thead>
<tr>
<th>Date</th>
<th>PT. ANTAM</th>
<th>PT. WIKA</th>
<th>PT. SINAR MAS</th>
</tr>
</thead>
<tbody>
<tr>
<td>8/30/23</td>
<td>1995</td>
<td>404</td>
<td>4400</td>
</tr>
<tr>
<td>8/29/23</td>
<td>1995</td>
<td>412</td>
<td>4410</td>
</tr>
<tr>
<td>8/28/23</td>
<td>1995</td>
<td>374</td>
<td>4380</td>
</tr>
<tr>
<td>8/25/23</td>
<td>1990</td>
<td>378</td>
<td>4400</td>
</tr>
<tr>
<td>8/24/23</td>
<td>1995</td>
<td>378</td>
<td>4430</td>
</tr>
<tr>
<td>8/23/23</td>
<td>1995</td>
<td>378</td>
<td>4410</td>
</tr>
<tr>
<td>8/22/23</td>
<td>1995</td>
<td>376</td>
<td>4450</td>
</tr>
<tr>
<td>8/21/23</td>
<td>1950</td>
<td>372</td>
<td>4400</td>
</tr>
<tr>
<td>8/18/23</td>
<td>1950</td>
<td>384</td>
<td>4400</td>
</tr>
<tr>
<td>8/16/23</td>
<td>1960</td>
<td>384</td>
<td>4400</td>
</tr>
</tbody>
</table>

After obtaining sliding window result data such as Table 3, the data is ready to be processed using Deep learning methods to find the most optimal comparison of accuracy results MAE and RMSE (Root Mean Square Error) (Mean Absolute Error). Graphic Image of PT. ANTAM, PT. WIKA and PT. SINAR MAS as shown in Figure 2.

![Figure 2. Stock price movements of (a) PT. ANTAM, (b) PT. WIKA and (c) PT. SINAR MAS](image)

Table 4 displays the findings of evaluating the three businesses’ deep learning techniques.

### Table 4. Comparison of deep learning methods

<table>
<thead>
<tr>
<th>Company</th>
<th>RMSE</th>
<th>MAE</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>PT. ANTAM</td>
<td>0.78973</td>
<td>0.54948</td>
<td>6355.132</td>
</tr>
<tr>
<td>PT. WIKA</td>
<td>0.43235</td>
<td>0.31505</td>
<td>1913.953</td>
</tr>
<tr>
<td>PT. SINAR MAS</td>
<td>0.98329</td>
<td>0.72337</td>
<td>9885.218</td>
</tr>
</tbody>
</table>

---

64 | *Scientific Journal of Informatics, Vol. 11, No. 1, Feb 2024*
From Table 4, the RMSE evaluation results at PT can be seen. WIKA has the best accuracy, 0.432235, MAE of 0.31505, and MSE of 1913.953, followed by PT. ANTAM with RMSE of 0.78973, MAE of 0.54948, MSE of 6355.132, and PT. SINAR MAS, RMSE of 0.98329, MAE of 0.72337, and MSE of 9885.218. The results of the Table Graph in Deep Learning Table 4 are shown in Figure 3.

Then, the results of PT. WIKA’s RMSE evaluation was optimized using selection optimization, and the evaluation results were increased to Table 5.

<table>
<thead>
<tr>
<th>Item</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE</td>
<td>0.41016</td>
</tr>
<tr>
<td>MAE</td>
<td>0.27990</td>
</tr>
<tr>
<td>MSE</td>
<td>1741.10200</td>
</tr>
</tbody>
</table>

From the research that has been carried out, it can be seen that using the Deep Learning method, the results obtained were the share price movements of PT. WIKA received the best evaluation results: RMSE of 0.43235, MAE of 0.31505, and MSE of 1913.95. After optimization using Optimization Selection, even better evaluation results were obtained. Namely, RMSE of 0.41016, MAE of 0.27990, and MSE of 1741.10, the difference obtained from the RMSE value is 0.022, MAE is 0.03515, and MSE is 172.85, this indicates that the share price movement of PT. WIKA is the safest to invest. So that investors can safely invest and avoid the risk of loss.

CONCLUSION
Based on the research that has been carried out, it can be concluded that the share price movements of the three companies in Indonesia, the share prices of PT. WIKA is the best and safest investment because the evaluation results are better than PT’s. SINAR MAS and PT. ANTAM has an RMSE of 0.4325, MAE of 0.315, and MSE of 1913.95. Optimized Optimization Selection produces an even better evaluation value with a value difference of 0.02, an RMSE value of 0.41016, an MAE value of 0.27990, and an MSE value of 1741.10. So, it can be used as information for investors to prefer PT. WIKA. Regarding the level of security in investing, this research aims to determine stock price movements after the coronavirus pandemic. The research results show that the Deep Learning approach has the best accuracy at PT. WIKA 0.02. To make it easier for investors planning to invest in this third business to decide whether to buy or sell PT. WIKA shares. Furthermore, this research only carries out experiments using Deep Learning algorithms, which are optimized using selection optimization, so further research can be carried out on applying other feature selection optimization algorithms to test whether this Deep Learning algorithm can increase its accuracy value.

REFERENCES


[36] “Stock Price Movement PT. WIKA.”

[37] “Stock Price Movements PT. Sinar Mas.”


