



Harnessing Machine Learning for Stock Price Prediction with Random Forest and Simple Moving Average Techniques

Arif Mudi Priyatno¹, Lidya Ningsih², Muhammad Noor³

¹University of Pahlawan Tuanku Tambusai, Riau, Indonesia

²Telkom University, Jawa Barat, Indonesia

³School of Computing, Univeriti Utara Malaysia, Kedah, Malaysia

Article Info

Keywords:

Machine Learning
Stock Price
Random Forest
Simple Moving Average
Prediction

ABSTRACT

This paper explores the application of machine learning in predicting stock price trends, specifically for PT Bank Central Asia Tbk (BBCA) shares, using the Random Forest Regression model and Simple Moving Average (SMA) techniques. The SMA parameters ranged from 3 to 200 days, aiding in forecasting the price trends as either rising, sideways, or declining. To achieve accurate and generalizable predictions, the data normalization process was implemented using the MinMax scaler. The methodological framework adopted a time series cross-validation (CV) approach, executed 10 times with a future test window of 40 days, ensuring the robustness and reliability of the predictive model. The model's performance was systematically evaluated based on metrics of accuracy, recall, precision, and F1-score. Results from the cross-validation series indicated varied performance, with the most notable achievements in the 9th and 10th iterations, where both demonstrated an F1-score surpassing 0.745 and 0.808 respectively, and similar levels of accuracy and recall at 0.825. These high F1-scores signify a strong harmonic balance between precision and recall, underscoring the model's capability to effectively predict the stock price movements of BBCA. The findings affirm the potential of utilizing advanced machine learning techniques like Random Forest in conjunction with SMA indicators to enhance the predictability of stock market trends, offering valuable insights for investors and financial analysts.

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



Corresponding Author:

Arif Mudi Priyatno
Department of Digital Business
University of Pahlawan Tuanku Tambusai
Jl. Tuanku Tambusai No.23, Bangkinang, Kec. Bangkinang, Kabupaten Kampar, Riau
Email: arifmudi@universitaspahlawan.ac.id

1. INTRODUCTION

The stock market plays a vital role in the global economy, acting as a barometer of economic health and a field for capital exchange [1]. Predicting stock price movements remains a perpetual challenge for

investors, analysts, and financial practitioners due to market complexity and volatility. In this context, the need for innovation in stock price prediction methodologies is crucial, given the rapidly changing market dynamics and the significant impact of accurate investment decisions [2].

Predictive technology, like machine learning algorithms, has revolutionized many sectors, including the financial market, by providing deeper insights and aiding more informed decision-making [3], [4]. Technical analysis, including the use of indicators like the Simple Moving Average (SMA), has become a vital tool in predicting stock price movements, giving investors the ability to identify market trends and patterns [5]–[7].

In the realm of financial market forecasting, the integration of machine learning techniques has emerged as a pivotal strategy for enhancing the accuracy of stock price predictions. Priyatno et al. [8] introduced a novel feature selection approach that leverages non-parametric correlations and important features within recursive feature elimination for stock price prediction. Their research underscores the complexities inherent in the vast array of potential features, proposing a method that surpasses standard recursive feature elimination by incorporating a strategic combination of important features and non-parametric correlations. This methodology demonstrated significant improvements in predicting the stock prices of PT Bank Rakyat Indonesia Tbk (BBRI), as evidenced by lower error metrics, highlighting the efficacy of their feature selection approach in enhancing predictive accuracy.

Further exploring feature selection, Priyatno et al. [9] presented an impurity-based method for identifying important features in recursive feature elimination aimed at stock price forecasting. By employing a mix of historical data and moving averages, this approach refined the prediction models, achieving optimal results with a reduced feature set in conjunction with linear regression. The success of this method is validated through its contribution to minimizing overfitting and ambiguity, thereby advancing the accuracy of stock market forecasts.

In a comparative study of machine learning algorithms for stock price forecasting, Priyatno et al. [10] examined the performance of Random Forest Regression and Linear Regression in predicting the stock prices of PT Bank Central Asia (BBCA). The study highlighted the challenges of stock price movements prediction and presented a comprehensive analysis of the effectiveness of these algorithms, revealing the superior forecasting capabilities of linear regression under certain conditions. Polamuri et al. [11], while delving into the prediction of stock market prices, applied both Random Forest and Extra Tree Regression. Their work elucidates the complexities of stock price prediction and underscores the potential of machine learning approaches to provide nuanced insights into market behaviors, thereby offering a more reliable tool for understanding stock exchange dynamics.

Yelne and Theng [12] along with Subasi et al. [13] have both contributed to the growing body of research on stock market prediction using supervised machine learning algorithms and a variety of machine learning techniques, respectively. These studies emphasize the diverse applicability of machine learning models across different stock exchanges and the importance of selecting appropriate classifiers to enhance prediction accuracy. Lastly, Khan et al. [14] explored the performance of machine learning models in stock market prediction, proposing a novel investment strategy aimed at improving the efficiency of these models. This strategy not only boosted the accuracy of predictions but also mitigated associated risks, underscoring the importance of innovative approaches in financial market forecasting.

Machine learning, with its adaptive capability to process and analyze large datasets, offers substantial potential in financial analysis [15]–[17]. The use of Random Forest Regression in this study aims to explore the extent to which this method can accurately predict stock price trends (upward, sideways, downward). The use of SMA as a technical indicator aims to add an analytical dimension in understanding the price movements of PT Bank Central Asia Tbk (BBCA), chosen as the case study due to its relevance and strong representation in the stock market. The selection of PT Bank Central Asia Tbk (BBCA) as the research subject is not only based on its significant position in the Indonesian financial market but also on its high volatility and liquidity, making it an ideal candidate for predictive stock price analysis.

In This paper, we predict the stock price trends of BBCA using Random Forest Regression based on historical stock prices and SMA analysis. The main contribution of this study is the development of a deeper understanding of how machine learning, particularly Random Forest, can be applied in stock market prediction, and providing new insights into the use of the SMA indicator. By measuring performance using accuracy, recall, precision, and F1-score, this study proposes a more nuanced understanding of the effectiveness of predictive algorithms in the dynamic stock market context.

This paper is organized into several main sections, starting with an introduction that provides context and motivation for the study, methodology explaining the data collection and analysis process, results and discussion presenting the research findings, and conclusion summarizing the key points and practical implications of the study.

2. RESEARCH METHOD

This research methodology discusses data, random forest, simple moving average, cross-validation, and performance evaluation.

2.1 Data

The data used is historical share data of PT Bank Central Asia Tbk (BBCA). Data was taken from 6 August 2004 to 9 October 2023. Stock history information used is open, close, high, low, adj close, and volume. Stock data obtained from Yahoo Finance.

2.2 Random Forest

Random Forest is an algorithm that integrates both bagging and multiple decision tree methods to produce effective solutions [18], [19]. Here, bagging is defined as a training method that develops basic models using subsets of data randomly selected from the entire training set. In the practice of the bagging method, the predictive solution is derived from the voting process conducted by each of these basic models.

Furthermore, the decision tree is recognized as an algorithm with a tree-like structure, where it primarily functions to address issues related to classification and regression [20]–[22]. Within the decision tree structure, there is a challenge in determining the most optimal attributes and thresholds, which will later be implemented as rules at the decision node. To overcome this challenge, two primary feature extraction methods are used: entropy and information gain.

Entropy is a method aimed at calculating the level of uncertainty, with the ultimate goal of obtaining the minimum uncertainty value from a predictor attribute, including its threshold [23]. On the other hand, information gain serves as a method to estimate the magnitude of information provided by a data attribute in explaining its class, where the calculation result of information gain is presented in probability form [24]. In this context, a higher probability value indicates a greater amount of information provided by that data attribute to facilitate the explanation of its class.

In implementing the Random Forest algorithm, the following steps are taken:

1. Produce a bootstrapped dataset by randomly selecting data samples in the same quantity as the training data set, using the replacement method.
2. Randomly select data features without replacement in the formed bootstrapped dataset to calculate the entropy and information gain values.
3. Utilize the feature and threshold with the highest information gain value to be implemented as the rule at the decision node at each depth level of the decision tree.
4. Repeat the aforementioned steps from 1 to 3 until the planned number of decision trees is fulfilled.
5. Execute the voting process on the test data to be processed by the trained number of decision trees.

2.3 Simple Moving Average (SMA) Techniques

Simple Moving Average (SMA) is a widely used technical indicator in financial markets for analyzing price movements and identifying trends [25], [26]. It calculates the average of prices over a specified number of periods, typically closing prices, to smooth out volatility and highlight price trends. The formula for SMA involves summing up prices over a specific period and dividing by the total number of periods. Short-term SMAs respond quickly to price changes, while long-term SMAs are slower to react. The formula for SMA involves calculating the average of prices over a specific number of days or periods. It is represented by the formula: $SMA = (A_1 + A_2 + \dots + A_n) / n$, where " $A_1 + A_2 + \dots + A_n$ " represents the sum of prices observed during a specific number of days or periods, and " n " is the total number of days or periods.

2.4 Data Normalization

Min-Max normalization is a common technique used in data preprocessing to scale the values of features within a specific range, typically between 0 and 1 [27]. This method involves performing a linear transformation on the original data to ensure that all values fall within the desired range. The formula for Min-Max normalization is as follows:

$$v' = \frac{v - \min(A)}{\max(A) - \min(A)}$$

Where: v' is the new normalized value. v is the original value of each entry in the data. $\min(A)$ and $\max(A)$ are the minimum and maximum absolute values of attribute A , respectively. By applying Min-Max normalization, all data points are transformed to a range between 0 and 1, preserving the relationships among

the original data values. However, it's important to note that this method may lead to smaller standard deviations, which can suppress the effect of outliers in the dataset.

2.5 Time Series Cross-Validation (CV) Approach

Time Series Cross-Validation is a crucial technique in machine learning, especially for time series data, as it extends traditional cross-validation methods to handle the temporal structure inherent in time series data [28]. Unlike standard cross-validation where data is randomly split into training and testing sets, Time Series Cross-Validation preserves the temporal order of observations. This ensures that the model is evaluated on past data and tested on future data, mimicking real-world scenarios.

One common approach in Time Series Cross-Validation is the TimeSeriesSplit method, which is a variation of KFold specifically designed for time series data. In each split, it returns the first k folds as the training set and the $(k+1)$ th fold as the test set. The key difference between TimeSeriesSplit and KFold is that in TimeSeriesSplit, the training dataset gradually increases in size, reflecting the chronological nature of time series data.

When applying Time Series Cross-Validation, it's essential to have sufficient historical data for meaningful evaluation and to choose an appropriate number of splits based on the length of the time series data. Experimenting with different models suitable for time series forecasting, selecting relevant evaluation metrics, and iteratively refining model configurations are also important steps in implementing Time Series Cross-Validation effectively.

2.6 Performance Evaluation Metrics

When evaluating the performance of a classification model, several key metrics are commonly used: accuracy, precision, recall, and F1-score. These metrics provide insights into different aspects of the model's performance in handling true positives, false positives, true negatives, and false negatives. The equation below is to calculate accuracy, precision, recall, and f1-score

$$\begin{aligned} \text{accuracy} &= \frac{TP+TN}{TP+TN+FP+FN} \\ \text{precision} &= \frac{TP}{TP+FP} \\ \text{recall} &= \frac{TP}{TP+FN} \\ \text{F1 - score} &= 2 * \frac{\text{precision} * \text{recall}}{\text{precision} + \text{recall}} \end{aligned}$$



Figure 1 Close movement of BBKA stock price

3. RESULTS AND DISCUSSION

The total number of historical data for BBKA stock is 4549 rows, while the number of features is 204 features. This feature consists of 6 price history features and 198 simple moving average features. The graph of the closing price movement of BBKA shares can be seen in Figure 1. Meanwhile, Table 1 is the BBKA share data used. Stock trends are up, down and sideways. This stock trend is made in individual numbers, namely 1 for up, 2 for down, and 3 for sideways.

Table 1 Stock data is used

Date	Open	High	...	SMA 199	SMA 200	Trend
5/23/2005	325.0	327.5	...	275.0	274.0	2
5/25/2005	325.0	327.5	...	276.0	275.0	2
5/26/2005	325.0	332.5	...	276.0	276.0	2
5/27/2005	330.0	330.0	...	277.0	277.0	1
5/30/2005	327.5	335.0	...	278.0	277.0	1

This stock data was divided into time series cross validation 10 times. Table 2 shows the distribution of training data and test data resulting from cross validation. The test data used was 40 days. The 1st to 10th cross validation training data is 4149, 4189, 4229, 4269, 4309, 4349, 4389, 4429, 4469, and 4509 days.

Table 2 Time Series Cross Validation Results

Cross-validation	Train Data	Test Data
1	4149	40
2	4189	40
3	4229	40
4	4269	40
5	4309	40
6	4349	40
7	4389	40
8	4429	40
9	4469	40
10	4509	40

The results of stock price trend prediction testing can be seen in Table 3. In the series of research carried out, Cross-Validation (CV) 1 shows the initial stage with an accuracy and recall of 15% each. Although this number is relatively low, it is important to understand that this is the first step in a series of experiments. With a precision of 2.3% and an F1-Score of 3.9%, there is significant room for improvement. However, it also shows that the machine learning process can start from any point, and every improvement, no matter how small, is progress.

Table 3 Cross validation test results

CV	Accuracy	Recall	Precision	F1-Score
1	15	15	2.3	3.9
2	32.5	32.5	62.5	27.8
3	50	50	52.4	50.4
4	27.5	27.5	32.8	23.7
5	40	40	36.6	35
6	25	25	17.7	14.4
7	35	35	42	33.4
8	40	40	68.7	42.7
9	82.5	82.5	68.1	74.6
10	82.5	82.5	80	80.8

Entering CV 2, there was a significant jump in model performance, with accuracy and recall both jumping to 32.5%. This shows that adjustments made based on previous results are starting to bear fruit. Dramatic improvements were seen in precision, which rose to 62.5%, and F1-Score which reached 27.8%. This indicates that the model is starting to understand the patterns better, leading to more accurate predictions. At CV 3, the model reaches a turning point with accuracy and recall both reaching 50%. This is an important milestone, showing that the model is now capable of classifying data with much higher

precision. Precision and F1-Score also showed healthy improvements, with figures of 52.4% and 50.4% respectively, confirming that the model is becoming more efficient in its predictions.

CV 4 provides important insight into the importance of iteration and adjustment in the model development process. With accuracy and recall at 27.5%, there is a visible decrease in performance. However, it also underscores the importance of continuous evaluation and adaptation of strategies. Precision and F1-Score, which are 32.8% and 23.7% respectively, confirm that each step in this process is important for deeper understanding. In CV 5, the model shows recovery with accuracy and recall of 40% each. This indicates that the adjustments made after observations on CV 4 were successful in producing results. With a precision of 36.6% and an F1-Score of 35%, there is further evidence that progress is a gradual process that requires persistence and dedication.

CV 6, although showing a small decrease in accuracy and recall to 25%, remains an important part of the learning process. Precision of 17.7% and F1-Score of 14.4% are reminders of the fact that not all experiments will be successful immediately, but each result provides a valuable lesson. Progress is again visible on CV 7, with the model reaching 35% for both accuracy and recall. This increase was accompanied by a precision of 42% and an F1-Score of 33.4%, showing that the courage to keep trying new strategies is the key to achieving success. CV 8 continues the positive trend with impressive performance, showing accuracy and recall of 40% each. The most striking improvements were seen in precision, which jumped to 68.7%, and F1-Score which rose to 42.7%. This underscores the importance of continuous optimization and the constant search for improvement.

CV 9 represents the culmination of this series of experiments, with outstanding results. Accuracy and recall both reached 82.5%, a figure that indicates extraordinary success. Precision and F1-Score are also impressive, with figures of 68.1% and 74.6% respectively. This is a clear testament to the team's ability to significantly improve and refine their model. At CV 10, the model achieved levels of success consistent with CV 9, confirming its reliability. With stable accuracy and recall figures at 82.5%, as well as precision of 80% and F1-Score of 80.8%, it shows that this model is not only effective but also consistent in its performance. This marks the culmination point of a rigorous iterative process, demonstrating the importance of persistence and innovation in model development.

Continuous improvement at each stage of Cross-Validation confirms the commitment to an iterative process in model development. This shows that the team is able to learn from each iteration and consistently apply that knowledge to make real improvements. The stability of the model at high performance levels, as shown in CVs 9 and 10, indicates reliable reliability in real applications. A stable model is critical to building user trust and ensuring consistency in decision making.

4. CONCLUSION

This research explores the application of machine learning in predicting stock price trends, specifically for PT Bank Central Asia Tbk (BBCA) shares, using the Random Forest Regression model and Simple Moving Average (SMA) techniques. The model's performance was systematically evaluated based on metrics of accuracy, recall, precision, and F1-score. Results from the cross-validation series indicated varied performance, with the most notable achievements in the 9th and 10th iterations, where both demonstrated an F1-score surpassing 0.745 and 0.808 respectively, and similar levels of accuracy and recall at 0.825. These high F1-scores signify a strong harmonic balance between precision and recall, underscoring the model's capability to effectively predict the stock price movements of BBCA. The findings affirm the potential of utilizing advanced machine learning techniques like Random Forest in conjunction with SMA indicators to enhance the predictability of stock market trends, offering valuable insights for investors and financial analysts.

REFERENCES

- [1] A. B and M. Deo, "The Influence of Macroeconomic Variables on the Stock Market Performance," *SSRN Electron. J.*, vol. 1, no. 1, pp. 1–15, Jun. 2021, doi: 10.2139/ssrn.4089305.
- [2] M.-P. Leonardo, C.-M. Carlos Eduardo, T.-H. Ana María, T.-G. José Luis, and L.-R. Campo Elias, "Formalization of a new stock trend prediction methodology based on the sector price book value for the Colombian market," *Heliyon*, vol. 8, no. 4, p. e09210, Apr. 2022, doi: 10.1016/j.heliyon.2022.e09210.
- [3] A. M. Rahmani, B. Rezazadeh, M. Haghparsat, W.-C. Chang, and S. G. Ting, "Applications of Artificial Intelligence in the Economy, Including Applications in Stock Trading, Market Analysis, and Risk Management," *IEEE Access*, vol. 11, pp. 80769–80793, 2023, doi: 10.1109/ACCESS.2023.3300036.
- [4] R. Regulagadda, V. Veeraiah, G. Muthugurunathan, L. N. K. S. Madupu, S. V Satyanarayana, and E. Muniyandy, "Predictive Analytics in Stock Markets: Unleashing the Power of IoT and Machine

- Learning,” *Int. J. Intell. Syst. Appl. Eng.*, vol. 12, no. 12s, pp. 414–422, 2024.
- [5] A. Thakkar and K. Chaudhari, “Fusion in stock market prediction: A decade survey on the necessity, recent developments, and potential future directions,” *Inf. Fusion*, vol. 65, no. July 2020, pp. 95–107, Jan. 2021, doi: 10.1016/j.inffus.2020.08.019.
- [6] I. K. Nti, A. F. Adekoya, and B. A. Weyori, “A systematic review of fundamental and technical analysis of stock market predictions,” *Artif. Intell. Rev.*, vol. 53, no. 4, pp. 3007–3057, Apr. 2020, doi: 10.1007/s10462-019-09754-z.
- [7] O. Bustos and A. Pomares-Quimbaya, “Stock market movement forecast: A Systematic review,” *Expert Syst. Appl.*, vol. 156, p. 113464, Oct. 2020, doi: 10.1016/j.eswa.2020.113464.
- [8] A. M. Priyatno, W. F. Ramadhan Sudirman, and R. J. Musridho, “Feature selection using non-parametric correlations and important features on recursive feature elimination for stock price prediction,” *Int. J. Electr. Comput. Eng.*, vol. 14, no. 2, p. 1906, Apr. 2024, doi: 10.11591/ijece.v14i2.pp1906-1915.
- [9] A. M. Priyatno, W. F. R. Sudirman, R. J. Musridho, and F. Amalia, “Impurity-Based Important Features for feature selection in Recursive Feature Elimination for Stock Price Forecasting,” *JUTIN J. Tek. Ind. Terintegrasi*, vol. 6, no. 4, pp. 1182–1194, 2023.
- [10] A. M. Priyatno, L. S. Tanjung, W. F. R. Sudirman, P. Cholidhazia, P. Z. Jati, and F. I. Firmananda, “Comparison Random Forest Regression and Linear Regression For Forecasting BBKA Stock Price,” *J. Tek. Ind. Terintegrasi*, vol. 6, no. 3, pp. 718–732, 2023, doi: 10.31004/jutin.v6i3.16933.
- [11] S. R. Polamuri, K. Srinivas, and A. Krishna Mohan, “Stock Market Prices Prediction using Random Forest and Extra Tree Regression,” *Int. J. Recent Technol. Eng.*, vol. 8, no. 3, pp. 1224–1228, Sep. 2019, doi: 10.35940/ijrte.C4314.098319.
- [12] A. Yelne and D. Theng, “Stock Prediction and analysis Using Supervised Machine Learning Algorithms,” in *2021 International Conference on Computational Intelligence and Computing Applications (ICCICA)*, Nov. 2021, no. ML, pp. 1–6. doi: 10.1109/ICCICA52458.2021.9697162.
- [13] A. Subasi, F. Amir, K. Bagedo, A. Shams, and A. Sarirete, “Stock Market Prediction Using Machine Learning,” *Procedia Comput. Sci.*, vol. 194, no. 4, pp. 173–179, 2021, doi: 10.1016/j.procs.2021.10.071.
- [14] A. H. Khan *et al.*, “A performance comparison of machine learning models for stock market prediction with novel investment strategy,” *PLoS One*, vol. 18, no. 9 September, pp. 1–19, 2023, doi: 10.1371/journal.pone.0286362.
- [15] J. W. Goodell, S. Kumar, W. M. Lim, and D. Pattnaik, “Artificial intelligence and machine learning in finance: Identifying foundations, themes, and research clusters from bibliometric analysis,” *J. Behav. Exp. Financ.*, vol. 32, p. 100577, 2021, doi: 10.1016/j.jbef.2021.100577.
- [16] N. Rouf *et al.*, “Stock Market Prediction Using Machine Learning Techniques: A Decade Survey on Methodologies, Recent Developments, and Future Directions,” *Electronics*, vol. 10, no. 21, p. 2717, Nov. 2021, doi: 10.3390/electronics10212717.
- [17] O. Gheibi, D. Weyns, and F. Quin, “Applying Machine Learning in Self-adaptive Systems,” *ACM Trans. Auton. Adapt. Syst.*, vol. 15, no. 3, pp. 1–37, Sep. 2020, doi: 10.1145/3469440.
- [18] P. Appiahene, Y. M. Missah, and U. Najim, “Predicting Bank Operational Efficiency Using Machine Learning Algorithm: Comparative Study of Decision Tree, Random Forest, and Neural Networks,” *Adv. Fuzzy Syst.*, vol. 2020, pp. 1–12, Jul. 2020, doi: 10.1155/2020/8581202.
- [19] H. Dabiri, V. Farhangi, M. J. Moradi, M. Zadehmohamad, and M. Karakouzian, “Applications of Decision Tree and Random Forest as Tree-Based Machine Learning Techniques for Analyzing the Ultimate Strain of Spliced and Non-Spliced Reinforcement Bars,” *Appl. Sci.*, vol. 12, no. 10, p. 4851, May 2022, doi: 10.3390/app12104851.
- [20] N. A. Priyanka and D. Kumar, “Decision tree classifier: a detailed survey,” *Int. J. Inf. Decis. Sci.*, vol. 12, no. 3, p. 246, 2020, doi: 10.1504/IJIDS.2020.108141.
- [21] R. Rivera-Lopez, J. Canul-Reich, E. Mezura-Montes, and M. A. Cruz-Chávez, “Induction of decision trees as classification models through metaheuristics,” *Swarm Evol. Comput.*, vol. 69, no. April 2021, p. 101006, Mar. 2022, doi: 10.1016/j.swevo.2021.101006.
- [22] M. Ahmad, N. A. Al-Shayea, X.-W. Tang, A. Jamal, H. M. Al-Ahmadi, and F. Ahmad, “Predicting the Pillar Stability of Underground Mines with Random Trees and C4.5 Decision Trees,” *Appl. Sci.*, vol. 10, no. 18, p. 6486, Sep. 2020, doi: 10.3390/app10186486.
- [23] M. S. T. T. Del and S. K. Tabrizi, “A methodological assessment of the importance of physical values in architectural conservation using Shannon entropy method,” *J. Cult. Herit.*, vol. 44, pp. 135–151, Jul. 2020, doi: 10.1016/j.culher.2019.12.012.
- [24] U. C. Bicici, T. H. S. Meral, and L. Akarun, “Conditional Information Gain Trellis,” *arXiv Prepr.*, vol. cs.cv, no. feb, pp. 1–20, Feb. 2024, doi: 10.48550/arXiv.2402.08345.
- [25] A. Kurani, P. Doshi, A. Vakharia, and M. Shah, “A Comprehensive Comparative Study of Artificial

- Neural Network (ANN) and Support Vector Machines (SVM) on Stock Forecasting,” *Ann. Data Sci.*, vol. 10, no. 1, pp. 183–208, Feb. 2023, doi: 10.1007/s40745-021-00344-x.
- [26] J. B. Pandya and P. (Dr. . U. K. Jaliya, “Opinion and Technical Indicator Based Optimized Deep Learning for Prediction of Stock Market,” *Indian J. Comput. Sci. Eng.*, vol. 12, no. 6, pp. 1860–1874, Dec. 2021, doi: 10.21817/indjcse/2021/v12i6/211206166.
- [27] D. Singh and B. Singh, “Investigating the impact of data normalization on classification performance,” *Appl. Soft Comput.*, vol. 97, p. 105524, 2020, doi: 10.1016/j.asoc.2019.105524.
- [28] S.-X. Lv, L. Peng, H. Hu, and L. Wang, “Effective machine learning model combination based on selective ensemble strategy for time series forecasting,” *Inf. Sci. (Ny).*, vol. 612, pp. 994–1023, Oct. 2022, doi: 10.1016/j.ins.2022.09.002.