IJSA (2025), Vol. 9, No. 2, 261-273
DOI: https://doi.org/10.29244/ijsa.v9i2p261-273

Indonesian Journal of Statistics and Its Applications
Available at https://journal-stats.ipb.ac.id/index.php/ijsa
e-ISSN: 2599-0802

Cluster Level Time Series Forecasting on Indonesian Banking Stock Prices Using the Gated Recurrent Unit Method*

Faisal Arkan[1], Budi Susetyo[2‡], and Rahma Anisa[3]

[1,2,3]School of Data Science, Mathematics, and Informatics - IPB University, Indonesia
‡corresponding author: budisu@apps.ipb.ac.id

Abstract

In recent years, there has been a significant increase in the number of Single Investor Identification registrations in the Indonesian capital market, as reported by the Financial Services Authority. Many investors favor stocks for their potential for high returns and liquidity. However, stock investments come with high risks due to their fluctuating prices, which are influenced by multiple factors. With 47 listed banking companies in the Indonesia Stock Exchange, clustering can help identify investor patterns. Forecasting stock prices is essential for anticipating future fluctuations. The large number of issuers and the tendency of stock prices to fluctuate increase the potential for outliers, requiring an appropriate clustering method. A study using the k-medoid method and dynamic time warping distance revealed 41 banking companies clustered into 5 clusters with a silhouette coefficient of 0.524. The Gated Recurrent Unit modeling, based on prototypes from the formed clusters, showed an excellent forecasting performance with root mean squared error and mean absolute percentage error ranging from 1-10%. The forecast for the next 8 weeks indicated varying price increases for each cluster. The first and third clusters are recommended for investors looking to maximize capital gains, due to their price increases and diverse cluster member characteristics. Additionally, investors should consider dividends provided by certain banking companies in their investment decision-making process.

Keywords: bank stock prices, dynamic time warping, gated recurrent unit, k-medoid swarm level forecasting.

* Received: Jan 2025; Reviewed: Jun 2025; Published: Des 2025

1. Pendahuluan

The Covid-19 pandemic, which had a major impact on the global economy, also had a positive effect, as it raised awareness among Indonesians about investing. This is supported by data from the Financial Services Authority (OJK), who reported that the number of single investor identification (SID) accounts in the capital market as of October 2023 grew by 14.04% year-on-year (YoY) to 11.75 million investors, and by 303% compared to 2020. Among the various investment instruments, stocks, or shares of ownership in a company, is a popular option for many people because they can provide high returns in a short period of time and have a high level of liquidity. However, stock investments also carry a high risk due to their fluctuating prices compared to other investments.

According to the Indonesia Stock Exchange (IDX), stock prices are influenced by the forces of supply and demand. If demand for stocks is high, prices will rise, and vice versa. Supply and demand are influenced by both external and internal factors. Stocks listed on the IDX consist of 852 issuers and are classified into 12 sectors. One sector with strong growth in recent years is the financial sector, particularly the banking sub-sector.

Banks are financial institutions whose main function is to collect funds and channel them in the form of credit to those who need it. In 2023, there were a total of 47 issuers listed in the banking sub-sector. The large number of issuers gives potential investors many options for investing. Clustering can be a method for selecting issuers that suit investor preferences by identifying the characteristics or patterns of these issuers. Stock price forecasting can also help investors predict future price fluctuations or changes.

Stock prices tend to fluctuate, increasing the potential for outliers, therefore requiring an appropriate clustering method. Luthfi & Wijayanto (2021) explained that the k-medoid method is better than the k-mean and hierarchical clustering methods when applied to Human Development Index (HDI) data containing outliers with an average standard deviation ratio of 0.517. One measure of dissimilarity that can be used for time series data clustering is dynamic time warping (DTW). DTW is a measure of dissimilarity that calculates the distance from the optimal warping path on two time series data. Sakinah et al. (2024) showed that the DTW measure performed better in clustering the prices of energy sector issuers than other distance measures.

One of the deep learning methods that can be used to capture complex patterns and long-term dependencies in stock data is the gated recurrent unit (GRU) method. The GRU method is part of a recurrent neural network (RNN) and an extension of long short-term memory (LSTM). Dey et al. (2021) mentioned that the GRU method was able to provide better results in predicting the stock prices of Honda Motor Company, Ltd. (HMC), Oracle Corporation (ORCL), and Intuit Inc. (INTU) compared to RNN and LSTM. Previous studies combining the use of the k-medoid clustering method with DTW distance measures and GRU forecasting are quite rare to found, thus study provides a novel contribution.

This study aims to cluster time series data on stock prices in the Indonesian banking sub-sector for the period from January 1, 2019 to April 2, 2024 using the k-medoid method and DTW distance measure, model cluster levels using the GRU method, and make forecasts based on the model obtained for the next 8 weeks.

2. Methodology

2.1 Data

This study uses secondary data obtained from the Yahoo Finance website at https://finance.yahoo.com/ . The data consist of weekly closing prices of publicly listed banking sub-sector companies (Tbk.) that were actively traded on the Indonesia Stock Exchange (IDX) during the period from January 1, 2019, to April 2, 2024.

2.2 Time Series Data Clustering

Cluster analysis is an approach in unsupervised learning involving grouping data objects into several clusters. Data objects in one cluster have high similarity in characteristics, while data objects from different clusters have low similarity in characteristics (Han et al., 2023). Clustering methods can be applied to time series data, enabling grouping based on observed patterns or trends over time.

Non-hierarchical clustering is a technique in data analysis aimed at grouping data objects into homogeneous clusters formed simultaneously without any hierarchy. A commonly used non-hierarchical clustering method is k-medoids. Unlike the k-means method, which uses the mean value as the cluster center (centroid), the k-medoids method uses medoids, objects that are centrally located within a cluster, making them robust against outliers. The partitioning around medoids (PAM) algorithm is the most commonly used k-medoid algorithm. This algorithm works by randomly determining the initial medoids, then iteratively updating the medoids to improve cluster quality. K-medoids work by minimizing an objective function defined in the following equation:

$$\underset{\{m_k\}, 1 \leq k \leq K}{min} \sum_{k=1}^{K} \sum_{x \in C_k} dist(x, m_k) \quad [1]$$

where x denotes the data object, C_k denotes the k-th cluster, m_k is the centroid of the k-th cluster, K denotes the number of clusters specified by the user, and "$dist$" denotes the distance function between the data object and the centroid.

The "dist" function calculates the distance between data objects using the dynamic time warping (DTW) method, a method for calculating the similarity between two time series data that may differ in time and speed. DTW distance calculates the optimal warping path of the two-time series (Berndt & Clifford, 1994). This allows DTW to capture patterns from a set of time series data. This distance has the advantage of calculating the distance between two time series that have different numbers of observations (Wan et al., 2021). DTW distance is more accurate when applied to time series data clustering than Euclidean distance because the measurement is based on pattern shape, not one-to-one mapping (Lee et al., 2020).

The DTW algorithm for calculating distance between two time series individuals, for example individuals $Q = q_1, q_2, \dots, q_n$ and $R = r_1, r_2, \dots, r_m$ can be performed as follows:

a. Creating a global cost matrix of size $n \times m$ by calculating the cumulative distance of each pair (i, j) and the minimum value of the three members adjacent to (i, j). The recursive formula of DTW cumulative distance is as follows:

$$d_{ij} = c_{ij} + min\{d_{(i-1)(j-1)}, d_{(i-1)j}, d_{i(j-1)}\} \quad [2]$$

where $c_{ij} = |q_i - r_j|$ represents the local distance between elements q_i and r_j. The terms d_{ij}, $d_{(i-1)(j-1)}, d_{(i-1)j}$ and $d_{i(j-1)}$ denote the cumulative distance at position (i, j), $(i-1, j-1)$, $(i-1, j)$ and $(i, j-1)$, respectively. This matrix is used to find the optimal warping path by providing the smallest cumulative distance. The first row and column of d_{ij} are initialized using the following equation:

$$d_{ij} = \begin{cases} c_{1j} + d_{1(j-1)} \\ c_{i1} + d_{(i-1)1} \end{cases}$$

b. Calculating the DTW distance using the following equation:

$$dist(Q, R) = min\{\Sigma_{l=1}^{L} g_L\} \quad [3]$$

where g_L indicates the value of the global cost matrix at the l-th position on the warping path, and L indicates the length of the warping path. Basically, the DTW distance is calculated by summing the values of the global cost matrix at each position on the warping path, then determining the minimum value of the total.

2.3 Gated Recurrent Unit

Gated Recurrent Unit (GRU) is a method developed from LSTM, which was first introduced in Cho et al. (2014) to overcome the vanishing gradient problem in the RNN method. Vanishing gradient occurs when the sequential data is too long, causing the information in the previous data to gradually disappear and leading to a decline in model performance. The GRU method has the advantage of a simpler architecture and fewer parameters compared to LSTM. In addition, this method also has faster computation time while still being able to capture long-term dependencies effectively.

The architecture/unit in GRU consists of two gates, namely the reset gate and the update gate. The reset gate (r_t) determines how much information from the previous process can be stored, while the update gate (z_t) determines how much new input (information) should be used in calculating the output. The update gate is a combination of the input gate and forget gate concepts in LSTM (Wang et al., 2018). The equations for the two gates can be written as follows (Liu et al., 2023):

$$r_t = \sigma(W_r \times x_t + U_r h_{(t-1)} + b_r) \quad [4]$$

$$z_t = \sigma(W_z \times x_{(t)} + U_z h_{(t-1)} + b_z) \quad [5]$$

where W_r and U_r are the weights of the previous input and output, while b_r is the bias.

2.4 Research Procedure

The research procedure in this study is as follows:

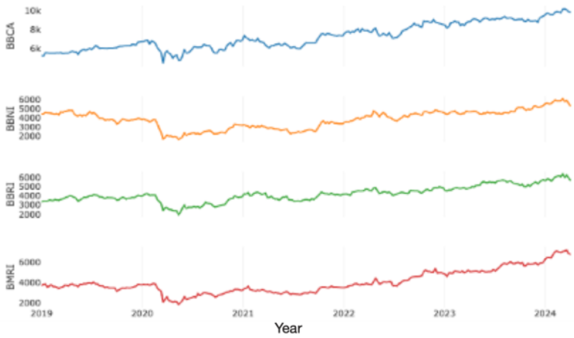
a. Preprocessing the data by selecting publicly listed banking issuers whose stocks were actively traded during the specified period.
b. Exploring weekly stock closing price data using time series plots to identify stock price movement patterns and boxplot to detect outliers.
c. Performing clustering of banking sub-sector issuers using the k-medoid method as follows:
   1) Selecting K objects as medoids based on silhouette coefficient values.
   2) Calculating the distance using the dynamic time warping (DTW) method for each object to the selected medoid.
   3) Assigning objects to the nearest cluster based on the minimum distance to the medoid and calculating the total distance from all objects to the nearest medoid (objective function).
   4) Redetermining the objects within each cluster as new medoid candidates.

5) Menghitung kembali jarak DTW pada masing-masing objek terhadap medoid baru.
6) Assigning each object to the nearest medoid based on the minimum dissimilarity and computing the total within-cluster dissimilarity as the updated objective function.
7) Calculating the total deviation (S) as the difference between the new objective function and the initial objective function. If $S < 0$ the initial medoid is replaced with the new medoid.
8) Repeating steps *d* through *g* until the minimum objective function is obtained.
9) Identifying the final cluster membership of each object.

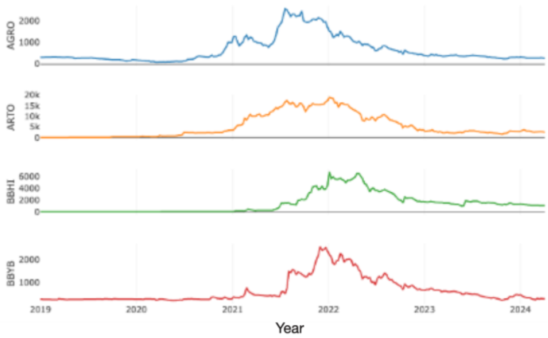
d. Performing *gated recurrent unit* (GRU) with the following steps:
1) Determining representative data (prototype) in each cluster using median values.
2) Investigating data stationarity in terms of the mean using autocorrelation function (ACF) plot and the Augmented Dicky-Fuller (ADF) test, and terms of the variance using Box-Cox approach. A non-stationary data is indicated by a tailing-off pattern on the ACF plot. Meanwhile, stationarity in variance is indicated then the confidence interval of the lambda parameter in the Box-Cox includes the value of one.
3) Performing feature scaling on each cluster prototype using standardization, which transforms the data to have a mean of zero and a variance of one. Standardization is conducted using the following equation (li & Liu, 2011):
$$p'_{kt} = \frac{p_{kt} - \bar{p}}{SD} \quad [6]$$
where $\bar{p}$ and is the prototype mean and SD is the standard deviation.
4) Dividing each group prototypes into training and testing data for evaluation process.
5) Constructing the GRU architecture mode through hyperparameter tuning with the grid search method and k-fold time series cross validation.
6) Evaluating the model on testing data using the mean absolute percentage error (MAPE) and root mean squared error (RMSE) defined as follows (Douglas C. Montgomery et al., 2015):
$$RMSE = \sqrt{\frac{1}{N}\Sigma_{i=1}^{N}(x_i - \hat{x_\iota})^2} \quad [7]$$
$$MAPE = \frac{1}{N}\Sigma_{i=1}^{N}\left|\frac{x_i - \hat{x_\iota}}{x_i}\right| \quad [8]$$
7) Forecasting banking stock prices at cluster levels for the next eight months.
8) Interpreting the resulting forecasts.

3. Results and Discussion

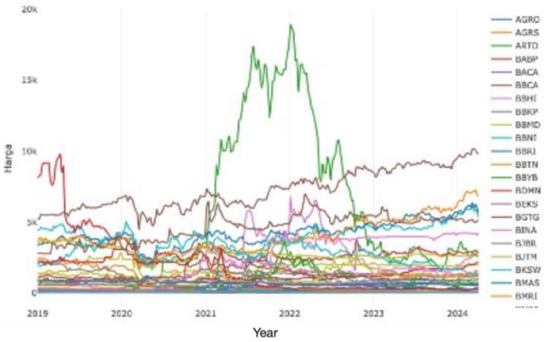
There are 41 issuers in the banking sub-sector out of a total of 47 issuers that meet the criteria in this study. Issuers that did not meet the criteria included Bank of India Indonesia (BSWD), Bank J Trust Indonesia (BCIC), Bank Amar Indonesia (AMAR), Bank Aladin Syariah (BANK), Krom Bank Indonesia (BBSI), and Bank Multiarta Sentosa (MASB) because they were suspended or had an initial public offering (IPO) after 2019. Data exploration was conducted on issuers that met the criteria using time series plots to observe stock price movement patterns during the period from January 1, 2019, to April 2, 2024, and box plots to detect outliers in each issuer. Visualization with time series plots shows that there are similarities in stock price movement patterns, such as in big banks issuers (BBCA, BBRI, BMRI, and BBNI) in Figure 1 (a) and digital banks (AGRO, ARTO, BBHI, and BBYB) in Figure 1 (b). The similarity in stock price movement patterns indicates that clustering can be performed.



(a)



(b)



(c)

Figure 1: Time series plots of (a) big banks, (b) digital banks, and (c) all issuers.

The Covid-19 pandemic that struck from March 2020 to June 2023 has put pressure on the global economy. However, banking issuers were still able to experience an increase in share prices, as shown in Figure 1 (c) for that period. One of the reasons for this was the credit restructuring policy issued by the OJK in March 2020. Credit restructuring is acountercyclical and landmark policy that supports the performance of debtors, banks, and the economy in general to get through the pandemic period. This policy was considered quite successful because during the pandemic, non-performing loans (NPLs) in the banking sector were generally in the range of and were still considered healthy. Ultimately, the policy was revoked by the OJK on March 31, 2024, after considering the improving condition of Indonesian banks post-pandemic.

Based on the exploration in Figure 2, 28 of the 41 issuers were identified as having outliers that indicated high volatility in these issuers. This level of volatility plays an important role in determining issuers and investment portfolios for investors (Hosseinioun, 2016). High volatility can be caused by investor perceptions of issuers, given that the banking sub-sector is sensitive to developments in financial regulations and market or corporate events. The eleven issuers that were not identified as having outliers were dominated by banks with large market capitalization in Indonesia, including Bank BCA (BBCA), Bank Mandiri (BMRI), Bank BNI (BBNI), Bank Syariah Indonesia (BRIS), Bank BTPN Syariah (BTPS), and Bank Mega (MEGA). This shows that banks with strong fundamentals tend to have low volatility.

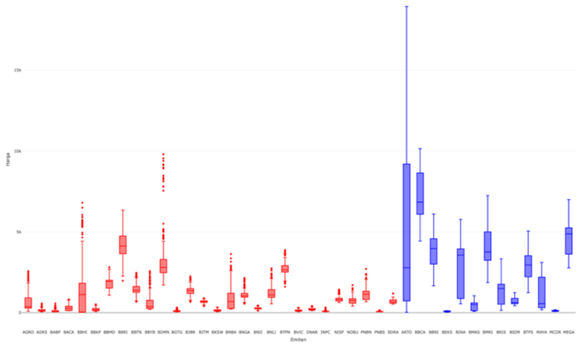


Figure 2: Box plot of stock prices of issuers in the banking sub-sector.

The clustering method that can be applied in this study is k-medoid because of its robustness to outliers compared to k-mean (Park and Jun 2009). The use of the DTW distance measure is also considered appropriate for use with the k-medoid method due to its ability to capture stock price movement patterns that tend to fluctuate over time (Nakagawa et al., 2019).

3.1 Time Series Clustering

The clustering of banking sub-sector stock prices was performed based on the proximity between issuers, measured using the Dynamic Time Warping (DTW) distance, where issuers with similar price movement patterns were grouped into the same cluster. The k-medoids clustering method requires the number of clusters (k) to be specified in advance; therefore, the silhouette coefficient was used to evaluate the optimal value of k. As shown in Figure 3, the highest silhouette coefficient (0.524) is obtained when $k = 2$; however, this configuration results in clusters with excessively large memberships, providing an overly coarse representation of the underlying data. Exploratory analysis indicates the presence of more than two distinct price patterns. Although the silhouette coefficient values for $k = 3$, $k = 4$, and $k = 5$ are relatively similar, a more substantial decline is observed beyond $k = 5$. Accordingly, $k = 5$was selected as the optimal number of clusters, as it offers a better balance between clustering quality and the ability to capture heterogeneous stock price dynamics.

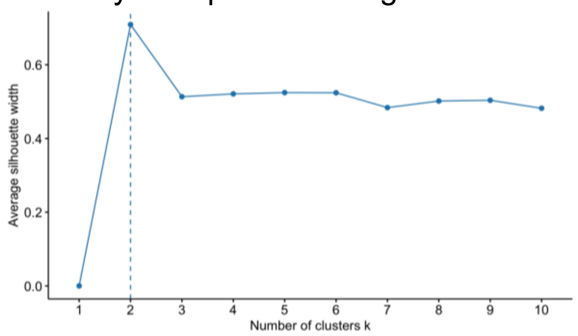


Figure 3: Determination of the optimum number of clusters.

Based on the time series plot in Figure 4, the five clusters formed have different movement patterns and characteristics. The first cluster in Figure 4 (a) consists of issuers owned by banks with strong fundamentals, which tend to experience fairly stable increases, such as BCA, BRI, Bank Mandiri, BNI, Bank Danamon, Bank BTPN, and Bank Mega. This fairly stable increase is due to these issuers having large market capitalization, which requires a large amount of investment to influence the increase or decrease in their prices. The share prices in this group range from IDR 2,000.00 to IDR 10,000.00 per share and have the highest average price compared to other issuers, which is IDR 4,285.00 per share.

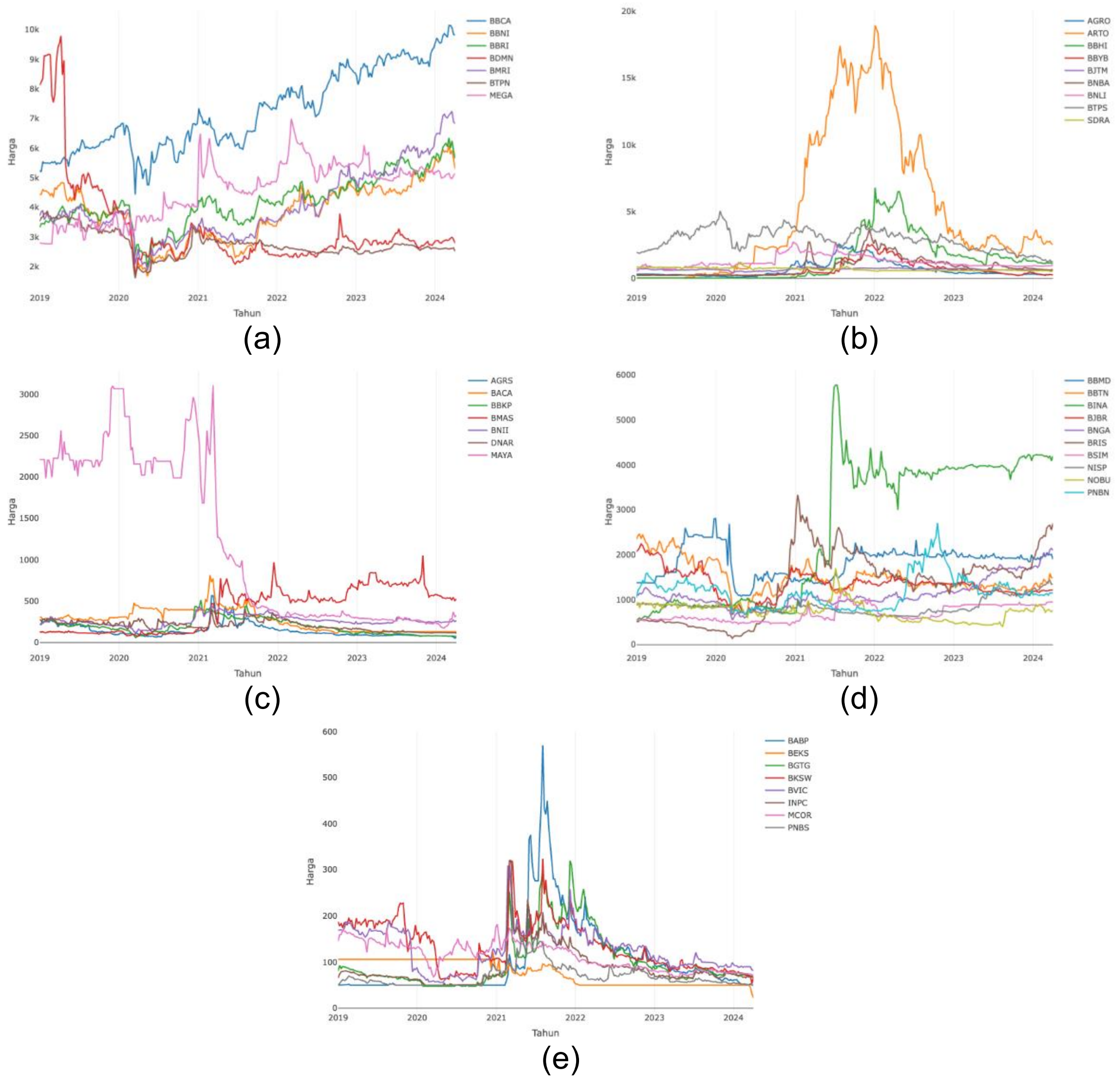


Figure 4: Comparison of time series plots (a) cluster 1 (b), cluster 2, (c) cluster 3, (d) cluster 4, and (e) cluster 5.

The second group in Figure 4 (b) consists of digital bank issuers such as Bank Jago, Bank Raya, Bank Neo Commerce, and Bank Allo. This group also consists of issuers Bank Jatim, Bank Bumi Arta, Bank Permata, Bank BTPN Syariah, and Bank Woori. Most members of this group tended to have an upward stock movement pattern in the period from 2021 to early 2022 and had the second highest average price after the first group, namely IDR 1,585.00 per share. The increase during this period, especially for digital bank issuers, was due to positive public sentiment that technological developments were accelerating and that digital banks were believed to have a bright future. However, in mid-2022, stock prices slowly declined due to inflation, causing investors to rebalance their portfolios towards sectors that benefited from rising commodity prices.

The third group in Figure 4 (c) consists mainly of issuers with a market capitalization of less than IDR 10 trillion, such as Bank Agris, Bank Capital, Bank Maspion, Maybank, Ok Bank Indonesia, and Bank Mayapada. There are also issuers with a market capitalization of more than 10 trillion rupiah, namely Bank Bukopin and Maybank. The pattern of stock price movements in this group tends to resemble the second group, occurring in the period from 2021 to 2022. This group has an average share price of Rp386.00 per share, and all issuers have a PBV below 1 (except for Bank Maspion), meaning the shares sold by these companies are relatively inexpensive compared to their book value.

The fourth group consists of 10 bank issuers, the majority of which have a market capitalization above 1 trillion rupiah, namely BTN Bank, BJB Bank, CIMB Niaga Bank, BSI Bank, Sinarmas Bank, OCBP Bank, and Panin Bank (Figure 4 (d)). The stock price movement pattern in the fourth group experienced a significant average increase in 2021 to 2022 after experiencing a decline due to Covid-19 in 2020 and a stable increase in 2023 to April 2024. This increase was due to positive public sentiment regarding mergers between several banks and positive financial performance reports. The average share price in this group ranged from IDR 600.00 to IDR 1,800.00 per share.

The fifth group consists of bank issuers with an average market capitalization of less than 2 trillion rupiah. These banks include MNC International Bank, Ganesha Bank, QNB Indonesia Bank, Victoria Bank, Artha Graha Bank, China Construction Bank, Panin Bank Syariah, and Banten Bank (Figure 4 (e)). Issuers in this group experienced a significant increase in 2021 and gradually declined as they entered 2022. This significant cumulative increase in share prices resulted in several issuers being suspended from the IDX in 2021.

3.2 Time Series Modeling

The time series modeling stage at the cluster level began with determining the prototype of each cluster formed. The prototype was determined using the median value because it is more robust when applied to data containing outliers. Then, a stationarity assumption check was performed on the five prototypes formed, and the results showed that all prototypes were non-stationary in mean and only the first cluster prototype was stationary in variance. Each prototype is then divided into training data and test data with a proportion of 76% or 209 observations as training data (January 1, 2019 to December 27, 2022) and 24% or 69 observations as test data (January 3, 2023 to April 2, 2024).

Table 1: Architecture and hyperparameter combinations of the GRU model.

Characteristics	Specifications	
Time step	8	
Layer	3 GRU; 1 *dropout*; 1 *dense*	
Unit	32; 64; 128	Hossain *et al.* 2021
Dropout rate	0,2; 0,4	Sonkavde *et al*. 2023
Batch size	16; 32; 64	Sonkavde *et al*. 2023
Epoch	40; 80; 100	
Optimizer	Adam	Hossain *et al.* 2021
Loss function	MSE	Hossain *et al.* 2021

The GRU model architecture used consists of 3 hidden layers, 1 dropout, and 1 dense layer (Table 1). The determination of the best hyperparameter combination for the model in each cluster was performed using the grid search method and k-fold time series cross validation with a value of 5. Grid search is an approach that is often used to train and optimize models obtained from each given hyperparameter combination (Ranjan et al., 2019). This study also used a time step value of 8. This value was determined based on the research by Arman and Zakaria (2019), which stated that certain months have an effect on stock issuers, especially in the financial sector. The best hyperparameter combination was obtained based on the smallest average MSE value from each combination set.

Table 2: Best hyperparameter combinations and evaluation values for each group.

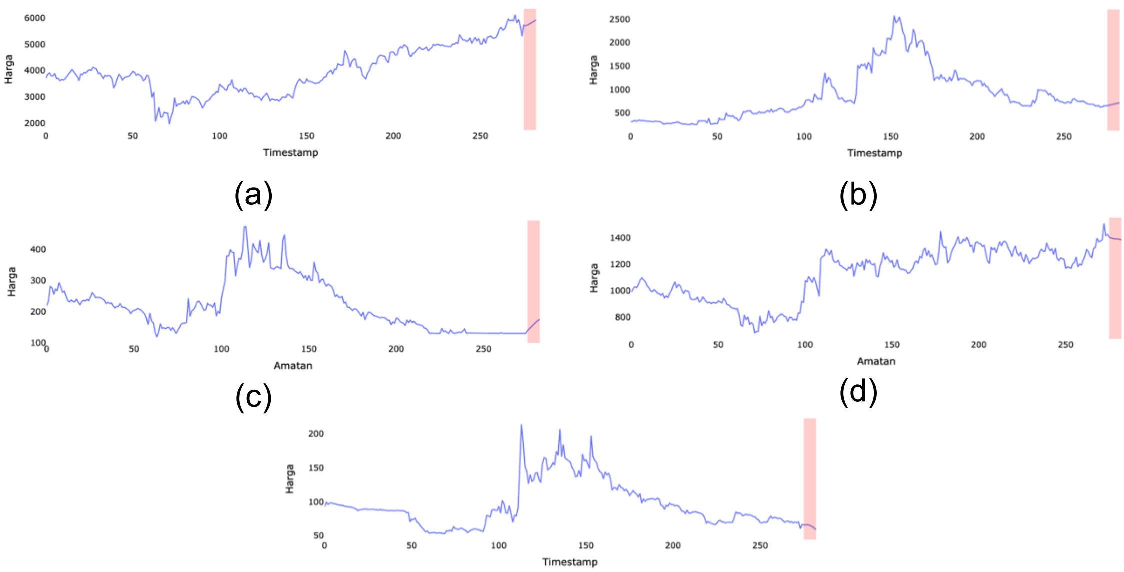
Cluster	Hyperparameter				Testing Data Evaluation	
	Units	Dropout	Batch Size	Epochs	RMSE	MAPE (%)
1	128	0.2	32	40	120.81	1.67
2	128	0.4	16	80	51.87	4.08
3	128	0.4	32	80	4.19	2.24
4	128	0.2	16	40	40.54	2.40
5	128	0.2	16	40	2.83	2.87

The evaluation values in Table 2 show that all models for each group have very good performance based on their MAPE values, which are less than 10%. In addition, the RMSE values obtained are also relatively low when compared to the range of each prototype. This indicates that the models obtained are good for forecasting.

3.3 Stock Price Forecasting

The forecasting results for the next 8 weeks (April 9, 2024 to May 28, 2024) for each prototype are shown in Figure 5 in red. The overall forecasting results show an increase in prices for each cluster except for the fourth and fifth clusters. The first and third groups show a more significant increase in price even though the members of each group have differences in their company characteristics (Figures 5 (a) and (c)). The first group consists of banks with the largest market capitalization in Indonesia, while the third group consists of small banks that still have the potential to grow.

The differences in the forecast results obtained can be taken into consideration by investors when diversifying their portfolios according to their risk profiles, given that the capital gains from each group tend to differ. In addition to capital gains, investors can also maximize their profits by investing in banks that pay regular dividends every year. Forecast results by the GRU model always contain uncertainty because increases and decreases in stock prices are influenced by demand and supply, which can continue to change in line with economic and geopolitical events.



(e)

Figure 5: Comparison of prototype prediction plots for (a) cluster 1, (b) cluster 2, (c) cluster 3, (d) cluster 4, and (e) cluster 5 over 8 weeks.

4. Conclusion and Recommendations

In this study, level cluster time series data forecasting analysis was applied to weekly closing prices of banking sub-sector stocks to help investors map their investments. Forty-one issuers in the banking sub-sector were clustered based on stock price movement patterns using the k-medoid method and DTW distance measure, which produced 5 clusters with a silhouette coefficient value of 0.524. The clustering results provide information about issuers with strong fundamentals that tend to have stable stock price movement patterns and are able to pay dividends regularly, while issuers with smaller market capitalization show higher volatility and tend to be unable to pay dividends.

The GRU modeling produced evaluation values in the form of MAPE below 10% and excellent RMSE. The forecasting results for eight periods show that issuers from the small market capitalization bank cluster tend to be more difficult to predict than more stable banks with large market capitalization. However, the technical analysis that has been carried out must be accompanied by the ability to monitor market conditions and the fundamentals of each bank in order to maximize profits.

This study focuses on the application of the GRU forecasting method at the group level with univariate data. Further research can be developed by considering the addition of variables that affect the stock prices of the banking sub-sector. In addition, the scope of the study can also be expanded through comparison with other deep learning methods to obtain a better model and understand the advantages and disadvantages of each method in stock price forecasting.

References

Berndt, D. J., & Clifford, J. (1994). Using dynamic time warping to find patterns in time series. *Proceedings of the 3rd International Conference on Knowledge Discovery and Data Mining*, 359–370.

Cho, K., van Merrienboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). *Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation*. http://arxiv.org/abs/1406.1078

Dey, P., Hossain, E., Hossain, Md. I., Chowdhury, M. A., Alam, Md. S., Hossain, M. S., & Andersson, K. (2021). Comparative Analysis of Recurrent Neural Networks in Stock Price Prediction for Different Frequency Domains. *Algorithms*, *14*(8), 251. https://doi.org/10.3390/a14080251

Douglas C. Montgomery, Cheryl L. Jennings, & Murat Kulahci. (2015). *Introduction to Time Series Analysis and Forecasting* (2nd ed.). Wiley.

Han, J., Pei, J., & Tong, H. (2023). *Data mining : concepts and techniques*. Morgan Kaufmann Publishers.

Hosseinioun, N. (2016). Forecasting Outlier Occurrence in Stock Market Time Series Based on Wavelet Transform and Adaptive ELM Algorithm. *Journal of Mathematical Finance*, *06*(01), 127–133. https://doi.org/10.4236/jmf.2016.61013

Lee, S., Kim, J., Hwang, J., Lee, E., Lee, K.-J., Oh, J., Park, J., & Heo, T.-Y. (2020). Clustering of Time Series Water Quality Data Using Dynamic Time Warping: A Case Study from the Bukhan River Water Quality Monitoring Network. *Water*, *12*(9), 2411. https://doi.org/10.3390/w12092411

li, W., & Liu, Z. (2011). A method of SVM with Normalization in Intrusion Detection. *Procedia Environmental Sciences*, *11*, 256–262. https://doi.org/10.1016/j.proenv.2011.12.040

Liu, C., Gao, F., Zhao, Q., & Zhang, M. (2023). The optimized gate recurrent unit based on improved evolutionary algorithm to predict stock market returns. *RAIRO - Operations Research*, *57*(2), 743–759. https://doi.org/10.1051/ro/2023029

Luthfi, E., & Wijayanto, A. W. (2021). Comparative analysis of hirearchical, k-means, and k-medoids clustering and methods in grouping Indonesia's human development index . *Inovasi: Jurnal Ekonomi, Keuangan, Dan Manajemen*, *17*(4), 761–773.

Nakagawa, K., Imamura, M., & Yoshida, K. (2019). Stock price prediction using *k* -medoids clustering with indexing dynamic time warping. *Electronics and Communications in Japan*, *102*(2), 3–8. https://doi.org/10.1002/ecj.12140

Ranjan, G. S. K., Kumar Verma, A., & Radhika, S. (2019). K-Nearest Neighbors and Grid Search CV Based Real Time Fault Monitoring System for Industries. *2019 IEEE 5th International Conference for Convergence in Technology (I2CT)*, 1–5. https://doi.org/10.1109/I2CT45611.2019.9033691

Sakinah, L., Anisa, R., & Sumertajaya, I. M. (2024). Energy Sector Stock Price Forecasting with Time Series Clustering Approach. *Indonesian Journal of Statistics and Its Applications*, *8*(2), 132–142. https://doi.org/10.29244/ijsa.v8i2p132-142

Wan, X., Li, H., Zhang, L., & Wu, Y. J. (2021). Multivariate Time Series Data Clustering Method Based on Dynamic Time Warping and Affinity Propagation. *Wireless Communications and Mobile Computing*, *2021*(1). https://doi.org/10.1155/2021/9915315

Wang, Y., Liu, M., Bao, Z., & Zhang, S. (2018). Short-Term Load Forecasting with Multi-Source Data Using Gated Recurrent Unit Neural Networks. *Energies*, *11*(5), 1138. https://doi.org/10.3390/en11051138