Jurnal Matematika dan Statistika serta Aplikasinya, Vol. 14, No. 2, 269-280, 2026
DOI: https://doi.org/10.24252/msa.v14i2.71030

Complexity-Invariant Distance-Based Clustering for Sharia Stocks Forecasting

Fauziah Roshafara[1]*, Fitri Amanah[2], Nafa Nurhanifah[3], Novianda Dwi Tanti Ramdani[4]
[1,2,3,4] *Department of Statistics, University of Islam Bandung, Indonesia*
***Corresponding author**: fauziahroshafara@unisba.ac.id

***Submission date**: 26 July 2026, **Revision**: 24 September 2026, **Accepted**: 22 October 2026

ABSTRACT
The rapid growth of sharia stock investment in Indonesia has increased the need for effective methods to analyze the movement patterns of multiple stocks. This study aims to cluster sharia stocks using Complexity-Invariant Distance (CID) combined with Average Linkage Hierarchical Clustering and to use the resulting clusters for cluster-level stock return forecasting using the Autoregressive Integrated Moving Average (ARIMA) model. Daily return data of 14 sharia stocks consistently listed in the IDX Sharia Growth Index from June 3, 2024, to May 29, 2026, were analyzed. The optimal cluster configurations were evaluated using the Silhouette Index and Dunn Index, while forecasting performance was assessed using Root Mean Square Error (RMSE). The Silhouette Index identifies the two-cluster configuration as having the highest value, while the Dunn Index reaches its highest value for the five- and six-cluster configurations. The two-cluster configuration was selected primarily because it achieved the highest Silhouette Index, while the five-cluster configuration was retained for a descriptive forecasting comparison.. The two-cluster configuration produces ARIMA(2,0,0) and ARIMA(2,0,2), with an average testing RMSE of 0.01587381, while the five-cluster configuration produces ARIMA(0,0,2), ARIMA(2,0,2), ARIMA(0,0,0), ARIMA(1,0,1), and ARIMA(0,0,1), with an average testing RMSE of 0.02740923. The two-cluster configuration yielded a lower average testing RMSE in this descriptive comparison. However, these errors are conditional on cluster membership determined from the full sample, and the configurations define different cluster-average forecasting targets. The findings demonstrate that CID-based clustering can serve as a preprocessing approach for cluster-level sharia stock return forecasting, where the forecasts represent the return dynamics of the identified clusters rather than individual stocks.

KEYWORDS
Complexity-invariant distance, Hierarchical clustering, Average linkage, ARIMA, Forecasting.

1. INTRODUCTION

The development of the Islamic capital market in Indonesia shows an increasing trend in recent years. The presence of various sharia investment instruments, especially sharia stocks that are members of sharia indices, in addition to offering potential profits for investors, also provides investment alternatives that are in accordance with Islamic principles due to the prohibition of Maghrib (Maysir, haram, gharar and riba) in practice [1]. The increasing number of investors and Islamic issuers has led to the need for accurate information on price movement patterns. This information can help investors understand the characteristics of stock movements and become the basis for investment decision-making in the coming period. Therefore, analysis of the historical pattern of sharia stock prices is one of the important aspects in supporting a more effective investment strategy.

However, the number of sharia stocks listed on the Indonesian capital market is relatively large and continues to grow, posing challenges in the analysis process. Investors and researchers often face difficulties in identifying stocks that have similar moving characteristics when each stock is analyzed individually. One approach that can be used to overcome this problem is

time series clustering, which is grouping stocks based on the similarity of their time series patterns. Wu et al. [2] shows that stocks can be grouped based on the similarity of trend characteristics (continuous trend characteristics), so that stocks with similar return movement patterns are in the same group and can be analyzed more effectively. Grouping based on the similarity of the return pattern not only simplifies the analysis process of a large number of stocks, but also results in a more homogeneous group that is able to support the investment decision-making process. Furthermore, Chen et al.[3] showed that the use of time series clustering results before the forecasting process can improve the model's ability to capture the characteristics of each time series group, resulting in better forecasting performance than building a model directly on all heterogeneous data. Therefore, the implementation of time series clustering is an important step before forecasting sharia stocks.

The success of the time series clustering process is greatly influenced by the similarity or distance measurement method used. One of the simplest measures of distance is Euclidean Distance (ED), which calculates similarity based on point-by-point differences between time series. Despite having fast and simple computation, ED assumes one-to-one alignment so it is sensitive to temporal shifts and less able to represent the similarity of two time series that have similar shapes but differ in phase and level of complexity [4, 5]. Alternatively, Dynamic Time Warping (DTW) was developed to overcome time shifts through a flexible alignment process so that it is able to match patterns that appear at different times [6, 7]. However, this flexibility causes DTW to have higher computational complexity than ED and has the potential to result in over-alignment in time series with different levels of fluctuation [8, 9]. To overcome these limitations, Batista et al. [8, 9] introduced Complexity-Invariant Distance (CID), which is a measure of distance that corrects Euclidean Distance by considering the level of complexity of each time series through a complexity correction factor. This method has also been proven to be able to produce higher silhouette values than the DTW distance [10].

CID calculates distances by adding a complexity correction factor to Euclidean Distance so that information about the rate of fluctuation and the dynamics of time series can be better considered. If two time series have a relatively similar shape but different levels of complexity, CID will produce a more appropriate distance value than Euclidean Distance, thereby improving the quality of similarity measurements and clustering results. This advantage makes CID one of the potential methods to be used in stock grouping, as stock price movements generally have varying volatility and complexity even though they show similar trend patterns. Some recent studies have also shown that CID is an effective distance measure for time series analysis because it can retain complex information that is often overlooked by ED and DTW [11].

The results of the stock grouping can then be used in the stock price forecasting process. Stocks that are in the same group are assumed to have similar movement characteristics so that group representation can be used to generate a more structured forecasting model. In this study, the Autoregressive Integrated Moving Average (ARIMA) method was used to conduct forecasting. The ARIMA model was chosen because it is one of the time series methods that has been proven to be effective in modeling data that has temporal dependence through autoregressive relationships, differencing, and moving averages, has a strong statistical theory basis, and can produce good forecasting for short- to medium-term financial data [12, 13]. Although various machine learning-based and deep learning-based approaches have developed rapidly, ARIMA is relatively easy to interpret compared to more complex approaches so it is still widely used as a benchmark model in time series forecasting research due to its simplicity, transparency, and ability to produce competitive predictions, especially in short- to medium-term financial data dominated by linear patterns [14, 15].

Based on this description, this study proposes the application of Complexity-Invariant Distance-Based Clustering on sharia stocks to group stocks based on the similarity of their return patterns, and subsequently uses the average return of each cluster as a representative time series for cluster-level forecasting using the ARIMA model. This approach is expected to produce representative groups of sharia stocks and provide cluster-level return forecasts that can support further analysis of stock return dynamics.

2. LITERATURE REVIEW

2.1 Complexity-Invariant Distance (CID)

The next stage is to calculate the degree of similarity between the return time series using the Complexity-Invariant Distance (CID) introduced by [9]. This method is an extension of Euclidean Distance by adding a complexity correction factor so that it is

able to distinguish two time series that have a similar shape but different levels of fluctuation. Euclidean Distance is defined as

$$ED(X,Y) = \sqrt{\left(\sum_{(i=1)}^{t} \left(x_i - y_i\right)^2\right)} \tag{1}$$

where $X = (x_1, x_2, \ldots, x_t)$ is the return of a certain stock and $Y = (y_1, y_2, \ldots, y_t)$ is another stock return. The complexity of each time series is calculated using

$$CE(X) = \sqrt{\left(\sum_{(i=2)}^{t} \left(x_i - x_{(i-1)}\right)^2\right)} \tag{2}$$

and

$$CE(Y) = \sqrt{\left(\sum_{(i=2)}^{t} \left(y_i - y_{(i-1)}\right)^2\right)} \tag{3}$$

Furthermore, the Complexity Correction Factor (CF) is calculated using

$$CF(X,Y) = \frac{(\max(CE(X), CE(Y)))}{(\min(CE(X), CE(Y)))} \tag{4}$$

so that the CID value is obtained through

$$CID(X,Y) = ED(X,Y) \times CF(X,Y) \tag{5}$$

If the two-time series have the same complexity, then the value $CF = 1$, so CID is identical to Euclidean Distance. On the other hand, if the level of complexity of the two series is different, the distance value will be adjusted to produce a more representative measure of similarity to the characteristics of the time series data.

2.2 Hierarchical Clustering using Average Linkage

The CID distance matrix is then used as an input in Hierarchical Agglomerative Clustering (HAC) with the Average Linkage method. This method uses an average of all object pairs to result in more stable cluster mergers than Single Linkage which is prone to chaining effects, or Complete Linkage which tends to result in clusters that are too compact. Therefore, Average Linkage is suitable for use in stock return data because it can accommodate variations in volatility between stocks without being too influenced by extreme values, so that the groups formed are more stable to the characteristics of return movements [16].

In this method each object is initially considered as a single cluster, then the two clusters that have the smallest average distance are gradually combined until the entire object is within a single dendrogram. The distance between the two clusters A and B is calculated using

$$d(A,B) = \frac{1}{|A|\,|B|} \sum_{i \in A} \sum_{j \in B} d(i,j) \tag{6}$$

with $|A|$: number of members of cluster A, $|B|$: number of members of cluster B, and $d(i,j)$: CID distance between objects.

2.3 Determining Number of Clusters

The optimal number of clusters is determined using two internal validation indices, namely the Silhouette Index and the Dunn Index. The silhouette value of each object is calculated using:

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \tag{7}$$

with $a(i)$: the average distance of the object to the same cluster member and $b(i)$: the average minimum distance to other clusters.

The Silhouette index is used to measure how close each object is to its own cluster compared to other clusters that produce values between -1 to 1. When the silhouette value gets closer to one indicates the object is on the right cluster [17].

Dunn Index is calculated using:

$$DI = \frac{\min\left(\delta\left(C_i, C_j\right)\right)}{\max\left(\Delta\left(C_k\right)\right)} \tag{8}$$

with $\delta\left(C_i, C_j\right)$: minimum distance between clusters and $\Delta(C_k)$: the diameter of the k–th cluster.

The larger the value of the Dunn Index indicates the more separate and more compact the cluster. The use of both indices simultaneously provides a more comprehensive evaluation of the quality of clustering results because it considers the aspects of compactness and separation. The combination of these two indices is also recommended in comparative studies of various cluster validation indices because it is able to provide a consistent assessment of the clustering results [18].

2.4 Autoregressive Integrated Moving Average (ARIMA)

After the candidate cluster configurations are determined, the forecasting process is carried out using the ARIMA model of the representation of each cluster through the mean value. The ARIMA model was developed through the Box–Jenkins approach [19], is a univariate time series model that combines autoregressive (AR) components, integrated (I) through a differencing process to achieve stationarity, and moving average (MA) to capture residual correlations. In general, the ARIMA model is denoted as ARIMA(p,d,q), with p expressing autoregressive order, d differencing order, and q moving average order can be written as

$$\left(1 - \phi_1 B - \cdots - \phi_p B^p\right)\left(1 - B\right)^d X_t = \left(1 + \theta_1 B + \cdots + \theta_q B^q\right)\varepsilon_t \tag{9}$$

Model formation is carried out through the stages of stationary identification, parameter estimation, residual diagnostic examination using the Ljung–Box test, and selection of the best model based on the Akaike Information Criterion (AIC) value.

2.5 Forecasting Evaluation

The performance of the forecasting model is evaluated using Root Mean Square Error (RMSE) which measures the average magnitude of prediction errors. RMSE is calculated using:

$$RMSE = \sqrt{\frac{1}{n}\sum\left(y_t - \hat{y}_t\right)^2} \tag{10}$$

Smaller RMSE values indicate that the forecasting results are closer to actual data so that the model has a better level of accuracy.

3. METHODOLOGY

3.1 Data

This research uses historical data in the form of daily closing prices of sharia stocks listed on IDX Sharia Growth consecutively from June 3, 2024 to May 29, 2026. Before the grouping process is carried out, stock prices are converted into daily return data to eliminate the influence of price scale differences between stocks and obtain a more representative time series on changes in investment value. The daily return is calculated using a simple return [20]as follows

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} \tag{11}$$

where P_t: stock closing price. The results of the daily returns that have been obtained will then be explored using data plots to see the characteristics of the movements of each stock.

3.2 Research Workflow

The overall methodology implemented in this study can be summarized as follows:

1. Collect daily closing prices of sharia stocks listed in the IDX Sharia Growth Index during the observation period.
2. Transform closing prices into daily returns to ensure comparability across different stock price scales.
3. Visualize the daily returns for each stock.
4. Calculate the Complexity-Invariant Distance (CID) matrix to measure similarity between stock return series.
5. Perform Hierarchical Agglomerative Clustering using Average Linkage based on the CID distance matrix.
6. Evaluate candidate cluster numbers from $k = 2$ to $k = 6$ using the Silhouette Index as the primary criterion for selecting the number of clusters, and the Dunn Index as a complementary measure. Retain the five-cluster configuration for an additional descriptive comparison of cluster-level forecasting performance.
7. Group the stocks into the selected optimal clusters and compute the average return of the stocks within each cluster to construct the cluster representative time series.
8. Divide each cluster representative time series chronologically into training and testing sets using an 80:20 proportion. The first 80% of observations are used as the training set for ARIMA model estimation, while the remaining 20% are reserved as the testing set for out-of-sample forecast evaluation. Although ARIMA models are estimated using only the training observations, cluster membership is determined from the complete observed return series, including the testing period. Therefore, the reported testing RMSE is conditional on the full-sample cluster membership and does not represent a strictly leakage-free evaluation of the entire clustering–forecasting pipeline.
9. Build ARIMA models for each cluster representative series using only the training data, with model selection based on appropriate criteria such as AIC and residual diagnostic checking.
10. Evaluate forecasting performance using Root Mean Square Error (RMSE) for both the training (in-sample) and testing (out-of-sample) data.
11. Forecast the cluster representative returns for the next five periods using the selected ARIMA models, with the final five-period forecasts reported in the Results section.

4. RESULT & DISCUSSION

The IDX Sharia Growth Index (IDXSHAGROW) measures the price performance of 30 Sharia-compliant stocks with relatively strong net income and revenue growth trends, transaction liquidity, and financial performance. However, only 14 stocks were consistently listed in the index throughout the study period from June 3, 2024, to May 29, 2026. For each selected stock, 477 daily closing price observations were recorded on trading days, excluding Saturdays, Sundays, and public holidays. These observations were then converted into daily returns using equation (11). The visualization of the return time series in Figure 4.1 shows that each stock has different movement characteristics, both in terms of volatility and the magnitude of return fluctuations. Some stocks, such as MAPA and DSNG, as well as PGAS and ITMG, visually show relatively similar fluctuations, while others such as ADRO and AKRA show more extreme spikes in returns in certain periods. The differences and similarities in these patterns indicate that not all stocks have the same movement characteristics, but there are some stocks that have the potential to form groups with similar return patterns. Therefore, it is necessary to analyze time series clustering using a similarity measure that is able to consider the shape as well as complexity of the time series to identify a more homogeneous group of stocks.

Before the clustering process, the distance matrix between stocks is first calculated using Complexity-Invariant Distance (CID) based on equation (5). The resulting distance matrix was then used in Hierarchical Agglomerative Clustering with Average Linkage. To determine the number of clusters, candidate values of $k = 2, 3, 4, 5$, and 6 were evaluated using the Silhouette Index and Dunn Index. This range was selected because the dataset contains 14 stocks, making solutions with a

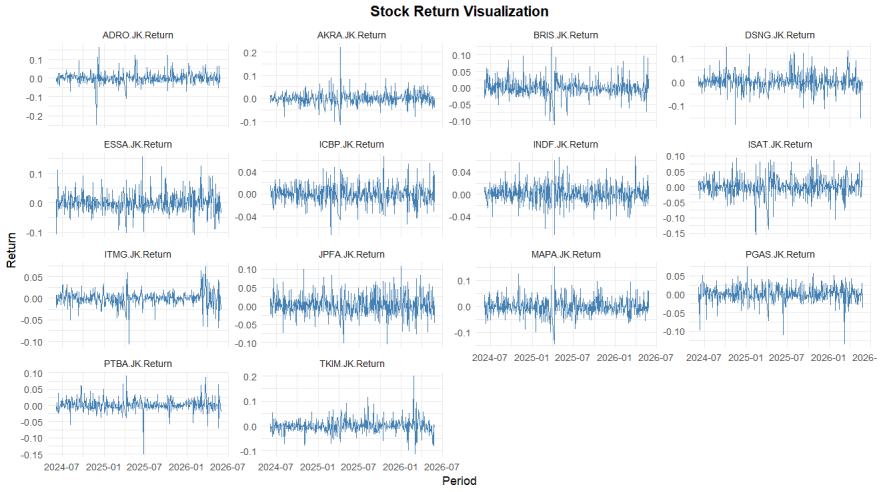


Figure 1. Stock Return Visualization

Table 1. Silhouette Score and Dunn Index Score

Number of Clusters	Silhouette Index	Dunn Index
2	**0.2929844**	0.6995234
3	0.2217084	0.7290604
4	0.2022643	0.7859783
5	**0.2067373**	**0.8671672**
6	0.1830745	0.8671672

larger number of clusters likely to produce very small clusters and less reliable cluster-level representative series. The lower bound of k=2 was used because a one-cluster solution does not provide a meaningful grouping, while the upper bound of $k = 6$ allows the analysis to examine finer partitioning without creating an excessive number of small groups.

As shown in **Table 1**, the highest Silhouette Index was obtained for the two-cluster configuration, with a value of 0.2930. Meanwhile, the Dunn Index increased as the number of clusters increased and reached its highest value of 0.8672 at $(k = 5)\, and\, (k = 6)$. The Silhouette Index and Dunn Index provide complementary information because the Silhouette Index evaluates the balance between within-cluster cohesion and separation from the nearest neighboring cluster, whereas the Dunn Index emphasizes the ratio between the minimum inter-cluster separation and the maximum within-cluster diameter.

Based primarily on the highest Silhouette Index, the two-cluster configuration was selected as the main partition for the subsequent analysis. The five-cluster configuration was retained as an additional comparison because it achieved the highest Dunn Index jointly with the six-cluster configuration, while having a higher Silhouette Index and fewer clusters. The comparison of testing RMSE between the two- and five-cluster configurations is treated as supporting descriptive evidence rather than the primary basis for selecting the number of clusters.

The Silhouette value of 0.2930 for $(k = 2)$ indicates that the cluster structure is relatively weak rather than strongly separated. Therefore, the resulting clusters should not be interpreted as completely distinct stock regimes. Instead, they represent groups of stocks that exhibit relatively similar return dynamics according to the CID-based similarity measure. The subsequent forecasting analysis was conducted for both $(k = 2)$ and $(k = 5)$ to further examine whether the different clustering configurations produce different characteristics in their cluster representative time series and forecasting performance.

Dendrograms of hierarchical clustering results using CID and Average Linkage are shown in **Figure 2**. In the formation of two clusters (**Figure 2a**), the stocks are divided into two large groups based on the similarity of the return pattern. When the number of clusters is increased to five (**Figure 2b**), some large groups split into more homogeneous subgroups so that

differences in characteristics between stocks can be identified in more detail. The merging process on the dendrogram also shows that stocks in the same branch have a relatively small CID value, thus showing a similar historical return pattern than stocks in different branches.

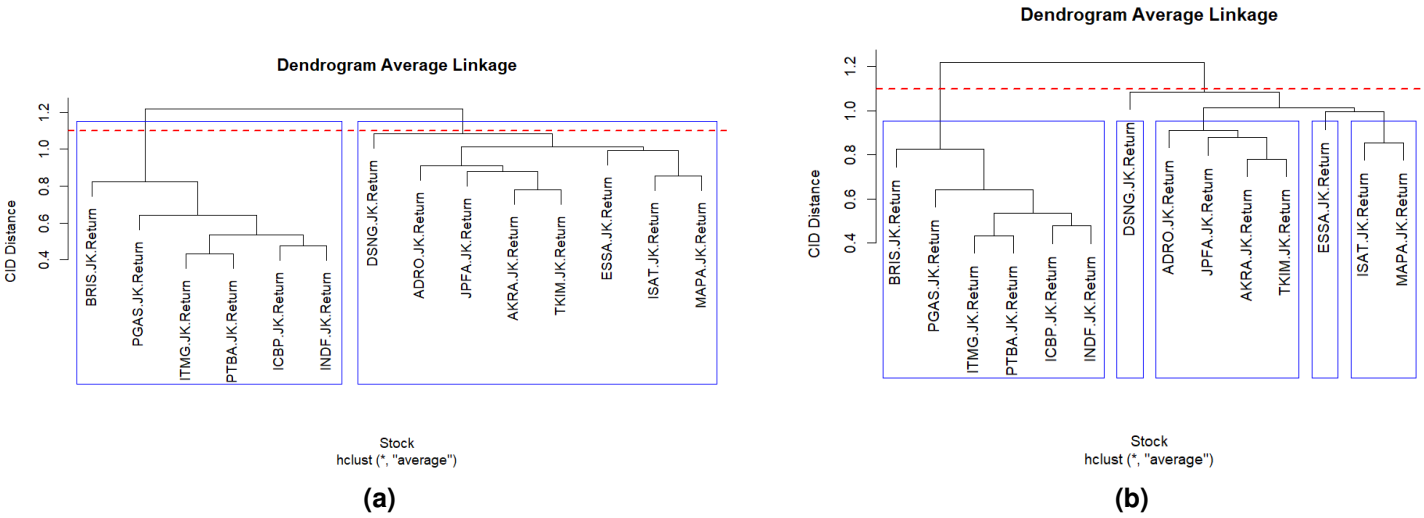


Figure 2. CID Distance Dendrogram with Average Linkage

Table 2 presents the stock membership for the selected two-cluster configuration and the five-cluster comparison configuration.

Table 2. Cluster Membership

Stock	$k = 2$	$k = 5$
ADRO	1	1
BRIS	2	2
JPFA	1	1
PGAS	2	2
AKRA	1	1
ITMG	2	2
TKIM	1	1
PTBA	2	2
ESSA	1	4
ICBP	2	2
ISAT	1	5
INDF	2	2
MAPA	1	5
DSNG	2	3

Furthermore, for each cluster configuration, the average return of all its members is calculated to represent the return characteristics of the corresponding group. These cluster representative series are then used as input for ARIMA modeling, so that the forecasting process focuses on the common movement pattern within each cluster rather than modeling each stock

individually. The identification of the ARIMA model for the two-cluster configuration was conducted by examining the stationarity test, ACF, and PACF plots shown in Figure 4.3. The Augmented Dickey-Fuller (ADF) test results in Table 4.3 show p-values of 0.01 for both cluster representative series, indicating that both series are stationary at the 5% significance level. Therefore, no differencing is required, resulting in an integration order of $(d = 0)$ for both cluster models.

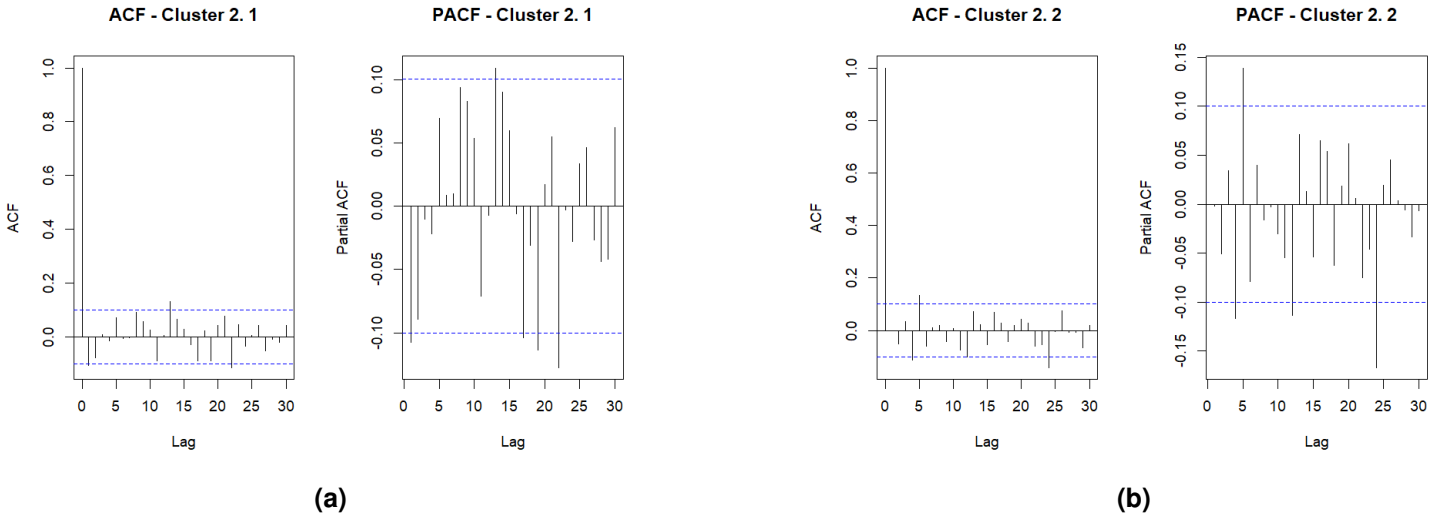


Figure 3. ACF and PACF of the 1^{st} Cluster (a) and 2^{nd} Cluster (b)

Figure 3 provides additional information for determining the autoregressive and moving-average orders. For Cluster 2.1, the ACF values are generally within the significance bounds, while several PACF spikes are observed at different lags. This pattern indicates that the dependence structure is more apparent in the partial autocorrelation function and provides a basis for considering autoregressive terms. For Cluster 2.2, the ACF also shows relatively small values, whereas the PACF contains significant spikes at several lags. These patterns are consistent with the consideration of ARIMA models with $(d = 0)$ and several combinations of autoregressive and moving-average orders.

Table 3. ARIMA Model Identification for Two-Cluster Configuration

Cluster 2.1			Cluster 2.2		
ADF p-value	**ARIMA Model**	**AIC**	**ADF p-value**	**ARIMA Model**	**AIC**
0.01	**(2,0,0)**	**-2116.947**	0.01	**(2,0,2)**	**-2358.830**
0.01	(0,0,2)	-2116.931	0.01	(2,0,1)	-2358.809
0.01	(0,0,1)	-2116.770	0.01	(1,0,2)	-2358.530
0.01	(1,0,1)	-2116.168	0.01	(1,0,1)	-2358.082
0.01	(1,0,0)	-2115.888	0.01	(0,0,0)	-2352.586
0.01	(1,0,2)	-2115.181	0.01	(0,0,1)	-2350.588
0.01	(2,0,1)	-2115.000	0.01	(1,0,0)	-2350.588
0.01	(0,0,0)	-2113.475	0.01	(0,0,2)	-2349.860
0.01	(2,0,2)	-2113.266	0.01	(2,0,0)	-2349.551

Several candidate ARIMA models were subsequently estimated and compared using the Akaike Information Criterion (AIC), as presented in **Table 3**. For Cluster 2.1, the smallest AIC value among the candidate models is obtained by ARIMA(2,0,0), with an AIC of -2116.947. For Cluster 2.2, the smallest AIC value is obtained by ARIMA(2,0,2), with an AIC of -2358.830.

Therefore, ARIMA(2,0,0) and ARIMA(2,0,2) were selected as the respective candidate models for the two cluster representative series based on the AIC criterion.

The same ARIMA model identification procedure was also applied to the five-cluster configuration. For each cluster representative series, the stationarity condition was first examined, followed by the identification of the autoregressive and moving-average orders using the ACF and PACF patterns. Several candidate ARIMA models were then estimated and compared based on their AIC values, with the model having the smallest AIC selected as the best model for each cluster. The resulting best ARIMA models for both the two-cluster and five-cluster configurations, together with their residual diagnostic results and forecasting performance, are summarized in **Table 4**.

Table 4. Model Selection

Clusters	ARIMA			Residual	RMSE		Average
	p	d	q	Diagnostic p-value	Training Set	Testing Set	Testing RMSE
Cluster 2.1	2	0	0	0.3810	0.01477295	0.01915998	**0.01587381**
Cluster 2.2	2	0	2	0.8037	0.01068737	0.01258764	
Cluster 5.1	0	0	2	0.0703	0.01602981	0.02156067	
Cluster 5.2	2	0	2	0.8037	0.01068775	0.01258807	
Cluster 5.3	0	0	0	0.7421	0.03508591	0.03542723	0.02740923
Cluster 5.4	1	0	1	0.3962	0.03054381	0.04121397	
Cluster 5.5	0	0	1	0.4408	0.02638246	0.02625621	

Table 4 presents the selected ARIMA models and their forecasting performance for the two- and five-cluster configurations. For the two-cluster configuration, Cluster 2.1 is modeled using ARIMA(2,0,0), while Cluster 2.2 is modeled using ARIMA(2,0,2), with residual diagnostic p-values of 0.3810 and 0.8037 from the Ljung-Box test, respectively, both exceeding 0.05. The corresponding testing RMSE values are 0.01915998 and 0.01258764, resulting in an average testing RMSE of 0.01587381. For the five-cluster configuration, the selected models are ARIMA(0,0,2), ARIMA(2,0,2), ARIMA(0,0,0), ARIMA(1,0,1), and ARIMA(0,0,1) for Clusters 5.1 to 5.5, respectively. All residual diagnostic p-values are greater than 0.05, while the testing RMSE values range from 0.01258807 to 0.04121397, resulting in an average testing RMSE of 0.02740923.

Cluster 5.3 produces an ARIMA(0,0,0) model, indicating that its representative series has limited autoregressive and moving-average dependence. Overall, the two-cluster configuration produces a lower average testing RMSE than the five-cluster configuration and provides ARIMA models with non-zero temporal components for both cluster representative series. Because the two-cluster configurations define different cluster-average target series, their average testing RMSE values are descriptive summaries and should not be interpreted as a controlled comparison of predictive accuracy for an identical target.The testing RMSE values should therefore be interpreted as forecast errors for cluster representative series conditional on the full-sample partition. A prospective evaluation would require clustering based only on the training period, with the resulting membership held fixed during testing.

For Cluster 2.1, the AR(1) coefficient was statistically significant at the 5% level ($\hat{\phi}_1 = -0.116996,\ p = 0.0222$),whereas the AR(2) coefficient was not significant ($\hat{\phi}_1 - 0.089814,\ p = 0.0797$). The estimated mean was 0.000490 and was not significantly different from zero ($p = 0.4374$). The ARIMA(2,0,0) specification was retained because it achieved the lowest AIC among the candidate models, although not all individual coefficients were statistically significant.

For Cluster 2.2, both AR coefficients and both MA coefficients were statistically significant at the 5% level. The estimated mean was 0.000091 and was not significantly different from zero ($p = 0.8672$). These results indicate evidence of temporal dependence in the fitted ARIMA(2,0,2) model, while neither cluster representative series had a mean significantly different from zero. **Figure 4** presents the five-period forecasts for the two cluster representative series using the selected ARIMA specifications. It also shows the five-period point forecasts for Clusters 2.1 and 2.2 under the selected two-cluster configuration. The forecasts remain close to zero and show no persistent upward or downward trend over the displayed horizon. This pattern is consistent with the small estimated means reported in Table 4.5. These forecasts describe the average return dynamics of the stocks within each cluster and should not be interpreted as forecasts for individual stocks.

Table 5. Parameter Estimation

Configuration	Cluster	ARIMA	Parameter	Estimate	p-value
$k = 2$	2.1	(2,0,0)	AR(1)	-0.116996	0.0222
			AR(2)	-0.089814	0.0797
			Const.	0.000490	0.4374
	2.2	(2,0,2)	AR(1)	-1.416572	< 0.001
			AR(2)	-0.596710	0.0072
			MA(1)	1.437649	< 0.001
			MA(2)	0.543439	0.0249
			Const.	0.000091	0.8672

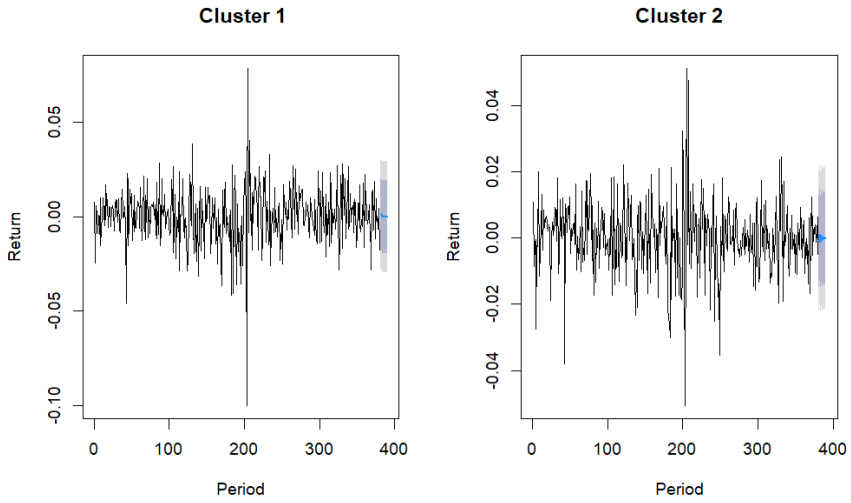


Figure 4. Forecasting Plot of the Upcoming 5 Periods

5. CONCLUSION

This study identified groups of sharia stocks with relatively similar return patterns using Complexity-Invariant Distance (CID) and Average Linkage hierarchical clustering. The forecasting process is conducted at the cluster level, where the average return of stocks within each cluster is used as the cluster representative time series for ARIMA modeling, rather than forecasting each stock individually. The two-cluster configuration was selected primarily on the basis of the highest Silhouette Index and yielded ARIMA(2,0,0) and ARIMA(2,0,2) for the two cluster representative series. In contrast, the five-cluster configuration produced an ARIMA(0,0,0) model for Cluster 5.3, indicating limited temporal dependence. The five-period forecasts show that the predicted returns of both cluster representatives fluctuate around their historical averages without a persistent upward or downward trend. These findings indicate that CID-based clustering can serve as a preprocessing approach for cluster-level sharia stock return forecasting, while the forecasting results should be interpreted as representing the return dynamics of the identified clusters rather than individual stocks. The reported testing RMSE is conditional on clusters identified from the complete sample. Future research should evaluate the framework using training-only clustering to assess its prospective forecasting performance.

REFERENCES

[1] B. Basrowi, P. Utami, E. Anggraeni, and M. Nasor, "Analisis swot pasar modal syariah sebagai sumber pembiayaan di indonesia," *El-Jizya: Jurnal Ekonomi Islam*, vol. 8, no. 2, pp. 210–227, 2020.

[2] D. Wu, X. Wang, and S. Wu, "Construction of stock portfolios based on k-means clustering of continuous trend features," *Knowledge-Based Systems*, vol. 252, Sep. 2022.

[3] J. Chen *et al.*, "From similarity to superiority: Channel clustering for time series forecasting," *Advances in Neural Information Processing Systems*, vol. 37, pp. 130 635–130 663, Dec. 2024.

[4] P. Esling and C. Agon, "Time-series data mining," *ACM Computing Surveys*, vol. 45, no. 1, Nov. 2012.

[5] A. Shifaz, C. Pelletier, F. Petitjean, and G. I. Webb, "Ts-chief: A scalable and accurate forest algorithm for time series classification," Feb. 2020.

[6] L. Sakinah, R. Anisa, and I. M. Sumertajaya, "Energy sector stock price forecasting with time series clustering approach," *Indonesian Journal of Statistics and Its Applications*, vol. 8, no. 2, pp. 132–142, Dec. 2024.

[7] F. Amanah, F. Roshafara, N. Nurhanifah, N. Dwi, and T. Ramdani, "Time series clustering and mean-variance portfolio modeling on idx sharia growth stocks."

[8] H. A. Dau *et al.*, "The ucr time series archive," *IEEE/CAA Journal of Automatica Sinica*, vol. 6, no. 6, pp. 1293–1305, Nov. 2019.

[9] G. E. A. P. A. Batista, X. Wang, and E. J. Keogh, "A complexity-invariant distance measure for time series," in *Proceedings of the 2011 SIAM International Conference on Data Mining*. Society for Industrial and Applied Mathematics, Apr. 2011, pp. 699–710.

[10] L. Fathiyaturrahmi, Andriano, H. A. Soleha, and D. D. Prastyo, "Perbandingan complexity invariant distance (cid) dan dynamic time warping (dtw) dalam analisis klaster deret waktu pada nilai tukar petani di indonesia," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 12, no. 4, pp. 771–778, 2025.

[11] P. Montero and J. A. Vilar, "Tsclust: An r package for time series clustering," *Journal of Statistical Software*, vol. 62, pp. 1–43, 2015.

[12] R. J. Hyndman and G. Athanasopoulos, *Forecasting: Principles and Practice*. Melbourne, Australia: OTexts, 2014, https://otexts.org/fpp2/.

[13] H. Wasserbacher and M. Spindler, "Machine learning for financial forecasting, planning and analysis: recent developments and pitfalls," *Digital Finance*, vol. 4, no. 1, pp. 63–88, Mar. 2022.

[14] V. I. Kontopoulou, A. D. Panagopoulos, I. Kakkos, and G. K. Matsopoulos, "A review of arima vs. machine learning approaches for time series forecasting in data driven networks," *Future Internet*, Aug. 2023.

[15] C. Zhang, N. N. A. Sjarif, and R. Ibrahim, "Deep learning models for price forecasting of financial time series: A review of recent advancements: 2020–2022," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, Jan. 2024.

[16] A. Arévalo, D. León, and G. Hernandez, "Portfolio selection based on hierarchical clustering and inverse-variance weighting," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Springer Verlag, 2019, pp. 315–325.

[17] P. J. Rousseeuw, "Silhouettes: a graphical aid to the interpretation and validation of cluster analysis," *Journal of Computational and Applied Mathematics*, vol. 20, pp. 53–65, 1987.

[18] O. Arbelaitz, I. Gurrutxaga, J. Muguerza, J. M. Pérez, and I. Perona, "An extensive comparative study of cluster validity indices," *Pattern Recognition*, vol. 46, no. 1, pp. 243–256, Jan. 2013.

[19] G. E. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time Series Analysis: Forecasting and Control*. New Jersey, USA: John Wiley & Sons, 2008.

[20] R. S. Tsay, *Analysis of Financial Time Series*, 2nd ed. New Jersey, USA: John Wiley & Sons, 2005.