



Analysis of Global Bank's Financial Performance with the Clustering K-Means Model

Suhari Santosa^{1*}, Jerry Heikal²

^{1,2} Master of Management, Faculty of Economic and Business, Universitas Bakrie, Jakarta, Indonesia

Abstract

Purpose: The purpose of this study is to find out the financial performance of global banks in each cluster for the years 2019 and 2023. In addition, this study is also to find out the position of Indonesia's banks compared to global and ASEAN banks in 2019 and 2023.

Methodology: The analysis model used is that the formation of clusters is based on the ratio of CAR, LDR, NIM, ROA and ROE. Testing was carried out with the K-Means model using SPSS.

Findings: The results of the study show that in general, global banking performance in 2023 is better than in 2019 in 4 clusters out of 5 clusters. However, the number of banks in the Very Good and Good cluster has decreased in 2023 compared to 2019. In addition, the number of banks in the Very Bad cluster also increased in 2023 compared to 2019.

Implication: The increase in the number of banks in the Very Bad cluster needs to be a concern, because the improvement in performance is not as good as other global banks. Local bank supervisory authorities, including the Financial Services Authority in Indonesia, need to pay attention to the performance of banks in the Very Poor cluster.

Originality: This study provides additional information about the condition of banks compared to their peers in 2019 and 2023 at the global, ASEAN and Indonesia levels for bank management, investors and also authorities.

Keywords: Clustering, K-Means, Ratio Analysis, Unsupervised Learning

Article Info

JEL Classification:
E50, F36, G21

Corresponding Author:
Suhari Santosa
(suhari.santosa@gmail.com)

Received: 01-08-2024
Revised: 04-09-2024
Accepted: 30-09-2024
Published: 02-10-2024



1. Introduction

The banking sector is the main driver of a country's economy, so it is natural that banking conditions, weaknesses and prospects of the banking sector are often the object of research for researchers, economists, and financial analysts (Vozniakovska et al., 2022). Various studies conducted in various countries in recent decades provide evidence that the growth of the financial system, driven by the banking sector, is a stimulus for economic growth known as the financial-growth nexus (Berger et al., 2020).

Similar circumstances also exist in Indonesia, where banks continue to dominate the financial industry. According to the OJK, the Master Plan for Indonesia's Financial Services Sector 2021 – 2025 states that banking assets (Rp 6,843 trillion) which accounted for 73,64% of the Indonesia's financial services sector in 2015. In 2019, the amount of Banking Assets reached Rp.8,713 trillion or 72.16% of Indonesia's total financial assets. And in the second quarter of 2020, Indonesia's banking assets reached Rp.8,817 trillion or 73.36% of Indonesia's total financial assets (Indonesia Financial Services Authority (OJK), 2020).

In general, bank supervision by the authority is carried out by grouping banks based on the nominal capital and nominal bank assets which are static. In Indonesia, the Financial Services Authority groups banks based on the nominal core capital (KBMI) which has not

changed much from year to year (static). The grouping of banks based on nominal capital and nominal assets is also carried out in various countries as well as the United States and the European Region.

On the other hand, grouping banks based on certain financial ratios will result in clusters whose members can change (dynamically) according to the proximity of their values. For this reason, in this study, the number of bank clusters will be determined based on 5 ratios, namely CAR, LDR, NIM, ROA and ROE. Furthermore, it is necessary to conduct an analysis to find out the performance of each cluster. Furthermore, an analysis of Indonesia's banking was carried out compared to global banking and banking in the ASEAN Region.

In this study, an analysis of global banking performance will be carried out using the K-Means clustering model using financial ratios in 2019 and 2023 with secondary data sourced from Bank Focus. Financial ratio analysis is a standard tool in assessing the company's financial performance (Özari & Can, 2023). The use of the clustering method in this study is motivated by the idea that processing large data with a variety of information will be better if homogeneous data grouping is carried out (Herman et al., 2022). The use of the K-mean method is based on the idea that the K-mean method is easy to use to partition a data set with a certain distance and does not cause overlapping clusters (Turkes, 2017).

2. Literature Review

One of the methods to find out the financial performance of a company is financial performance analysis. A company's financial performance reflects a company's ability to manage its finances efficiently and effectively. Operational activities, investments and funding over a given period can be measured using a universally applicable set of analyses. Because the analysis tool used is in the form of financial ratios, financial performance analysis is called Financial Ratio Analysis (Irfani, 2020).

According to Collier, the benefits of financial performance analysis are (1) helping to interpret the company's performance trends from year to year, (2) benchmarking the company's performance compared to the industry average performance, and (3) comparing the company's performance with the performance of competitors (Collier, 2003). Financial performance analysis can be used to interpret the performance of 5 main aspects of a company, namely liquidity, gearing, profitability, efficiency and rate of return to shareholders.

According to Aryanti (Ratnawati et al., 2022), banking performance can be assessed based on the bank's financial ratios, including the Capital Adequacy Ratio (CAR). CAR is used to measure the adequacy of a bank's capital to support potentially risk-exposed assets, such as loans disbursed by banks. The CAR formula is $\text{Equity} / \text{Risk Weighted Assets}$. If the bank's CAR is high, it can be interpreted that the bank's condition is good. In addition, a high CAR also indicates a high level of profit as well as a healthy bank condition. BCBS in Basel III requires global banks to have a minimum CAR ratio of 8% (Otoritas Jasa Keuangan, n.d.).

Another ratio used to assess a bank's performance is the profitability ratio. There are 3 ratios that can be used, namely Return On Assets (ROA), Return On Equity (ROE) and Net Interest Margin (NIM) (Ratnawati et al., 2022). The ROA formula is $\text{Net Profit After Tax (EAT)} / \text{Total Assets}$. The higher the ROA ratio value of bank, indicating the better the bank's performance.

Meanwhile, the ROE ratio formula is $\text{Net Profit After Tax (EAT)} / \text{Equity}$. The higher the value of a bank's ROE ratio, indicate the better the bank's performance. Adapun formula rasio NIM adalah $\text{Interest Receined} - \text{Interest Paid} / \text{Average Invested Assets}$. The higher the NIM ratio value of a bank, indicate the better the bank's performance.

Another ratio used to assess bank performance is the Loan to Deposit Ratio (LDR) ratio. LDR ratios can be grouped as liquidity ratios. The LDR ratio is used to measure the level of bank fund distribution to the level of third-party deposits. The higher the LDR ratio, the greater the opportunity for the bank to make a profit (Luh Shintya Anggari & Made Dana, 2020). Meanwhile, according to Hantono, LDR has a positive and insignificant impact on

profitability (Hantono, 2017). On the other hand, if the LDR exceeds 110%, it indicates the more liquidity risk and conversely, the lower the LDR, the more ineffective the bank's fund management will be (Carindri & Untara, 2019). The formula for the LDR ratio is Loans / Third Party Deposits.

Machine Learning is the science that studies computer algorithms that can recognize various patterns in data with the aim of transforming various types of data into real actions with minimal human intervention (Kurniawan, 2022). With machine learning, we can make machines study existing data and subsequently make independent decisions without the need for programming. In general, machine learning is in the scope of artificial intelligence.

There are 2 types of learning approaches used by machine learning, namely supervised learning and unsupervised learning. Supervised learning requires training data that contains significant information (Shalev-Shwartz & Ben-David, 2014). Some of the supervised learning algorithms include linear regression, logistic regression, decision trees, random forest, naïve bayes, and neural networks. Supervised learning algorithms are used as the basis for making predictions and generating predictive models (Kurniawan, 2022).

On the other hand, unsupervised learning does not distinguish between data training and test data (Shalev-Shwartz & Ben-David, 2014). Data in unsupervised learning does not require any prior labels/definitions. Unsupervised learning processes input data with the aim of generating a summary or compression of existing data. Examples of unsupervised learning algorithms are association and clustering.

Data mining is a data processing method used to find hidden patterns in a data set that can be used to aid decision-making (Kusumaningrum & Heikal, 2023). Data mining is a large-scale data processing that is very useful in industrial, financial, and technological analysis. In this study, the types of data mining that will be used are Clustering and K-Means.

The information database continues to grow, becoming more diverse, comprehensive and heterogeneous. So to get the right solution, the database needs to be systematically arranged and broken down into homogeneous groups. In a homogeneous group, evaluation and analysis of each group can be carried out better and more correctly, until finally the evaluation of the entire database can produce a more realistic picture (Herman et al., 2022). Clustering is a process of grouping data into homogeneous groups based on their attributes. Independently, machine learning will generate clusters based on data similarities automatically and without any human intervention (Kurniawan, 2022).

In general, according to Madhulatha (Madhulatha, 2012) there are 4 types of clustering algorithms, namely (1) Hierarchical Clustering which consists of agglomerative (bottom-up) and divisive (top-down), (2) Partitional Clustering which is based on determining the initial number of groups and then iteratively relocating objects between groups into convergence, for example Partitional Clustering are K-Mean and K-Medoids, (3) Density-Based Clustering which is designed to find arbitrarily shaped clusters and considers a cluster as a region whose density exceeds a threshold, examples are DBSCAN and SNN, and (4) Grid-Based Clustering which uses a multiresolution grid data structure approach and quantifies space into a finite number of cells that form the grid structure where all clustering operations are located done, examples are STING and CLIQUE.

K-Mean uses the mean/centroid group as the basis for data grouping (Herman et al., 2022). Each cluster is formed from the average value. The value or element in the cluster is the closest to the average. In general, according to Turkes, there are 5 stages in the formation of clusters (Turkes, 2017) which consists of: (1) determining the number of clusters, (2) determining the initial centroid, (3) grouping the data/elements based on the nearest centroid, (4) selecting a new centroid in each cluster, and (5) repeating stage 3 until the centroid does not change.

3. Methodology

Bank data withdrawal is carried out by querying Bankfocus with the main criteria, namely there is data on the ratio of CAR, NIM, LDR, ROA and ROE with the status of the bank active (still operating). The data that is pulled also includes BvD ID number data. The data

withdrawal was carried out for the period of 2019 and 2023. Furthermore, an in-depth evaluation of the data is carried out in order to produce a model that is in accordance with expectations. Some of the activities carried out are to evaluate whether there is a duplication of data based on the BvD ID number and then the duplication of data is removed/deleted. The next step is to evaluate whether there is outlier data. In this study, the data is considered outlier if it has a value of more than 3 standard deviations from the mean.

This quantitative research uses the unsupervised learning analysis method with the K-Means algorithm. The data of this study was processed using SPSS version 21. This study uses 5 (five) variables so that the cluster results can represent the bank's performance. The bank financial ratio data used in this study consists of CAR, NIM, LDR, ROA and ROE. In this study, the independent variables are 5 financial ratios CAR, NIM, LDR, ROA and ROE. The dependent variable in this study is the type of cluster type formed from the results of clustering using K-Means.

4. Results and Discussion

The query carried out on the data at Bankfocus uses the following criteria: the bank is still actively operating, has data on total assets, CAR, NIM, LDR, ROA and ROE in 2019-2023. There are 5,442 banks with a distribution of North America (2,784), South & Central America (312), Western Europe (692), Eastern Europe (281), Middle East (113), Africa (167), Far East & Central (1,016), and Oceania (77). Furthermore, data analysis was carried out to determine whether there was data duplication and identification of outlier data with a limit of 3 times the standard deviation.

Furthermore, based on using SPSS, data processing is carried out using K-means and produces clusters as shown in Table 1 and Table 2 below.

Table 1. Bank Cluster 2019 Result

	Cluster				
	1	2	3	4	5
CAR 2019	21,04	17,08	16,33	13,58	23,52
LDR 2019	122,54	72,70	94,10	79,72	41,47
NIM 2019	3,74	3,56	3,98	3,07	3,29
ROA 2019	1,04	1,03	1,17	-2,88	0,96
ROE 2019	8,47	9,45	10,18	-30,14	8,65

Source: Data Processed (2023)

Table 2. Bank Cluster 2023

	Cluster				
	1	2	3	4	5
CAR 2023	17,39	16,22	18,91	24,19	21,76
LDR 2023	72,72	93,96	62,97	38,91	119,89
NIM 2023	3,23	3,85	5,02	3,48	3,92
ROA 2023	0,80	1,10	2,46	1,00	1,02
ROE 2023	8,58	10,64	28,55	11,25	8,09

Source: Data Processed (2023)

To determine the persona *of the cluster*, weighting is carried out on the 5 ratios used. In this study, the CAR ratio is given a weight of 25% because the CAR ratio is quite dominant in the Bank's performance assessment. The NIM ratio is also given a weight of 25% because the ratio reflects the main activities of the main bank, namely interest income from loans and interest charges on third-party funds and has a major impact on the bank's main objective. Meanwhile, the ROA and ROE ratios are given a weight of 20% because *return* is the main goal of the bank's establishment. Meanwhile, LDR is given a weight of 10% because it does not directly impact bank revenue.

The bank's performance persona based on the ratio of CAR, LDR, NIM, ROA and ROE in 2019 is as follows: cluster 1 Very Good, cluster 2 Moderate, cluster 3 Good, cluster 4 Very

Bad and cluster 5 Poor. Meanwhile, the cluster persona in 2023 is cluster 1 Poor, cluster 2 Moderate, cluster 3 Good, cluster 4 Very Bad and cluster 5 Very Good.

The results of the clustering analysis with K-means globally in 2019 showed that there were 567 banks in the Very Good cluster, 1,894 banks in the Good cluster, 1,583 banks in the Medium cluster, 611 banks in the Bad cluster, and 60 banks in the Very Bad cluster. Meanwhile, in 2023 there are 519 banks in the Very Good cluster, 410 banks in the Good cluster, 1,673 banks in the Medium cluster, 1,509 banks in the Bad cluster, and 604 banks in the Very Bad cluster.

For the ASEAN region in 2019, there are 23 banks in the Very Good cluster (13 from Indonesia), the Good cluster has 79 banks (41 from Indonesia), the Medium cluster has 26 banks (18 banks from Indonesia), the Bad cluster has 6 banks and the Very Bad cluster has 3 banks, all of which are from Indonesia.

Meanwhile, the bank cluster in the ASEAN Region in 2023 is in the Very Good cluster there are 20 banks (12 from Indonesia), the Good cluster has 3 banks (2 from Indonesia), the Medium cluster has 62 banks (31 banks from Indonesia), the Bad cluster has 39 banks (23 from Indonesia) and the Very Bad cluster has 13 banks (7 from Indonesia).

5. Conclusion

Based on the results and discussion of the analysis in this study, it can be concluded that in general, global banking performance in 2023 is better than in 2019 in 4 clusters (Good, Medium, Poor, and Very Bad). However, the number of banks in the Very Good and Good cluster has decreased in 2023 compared to 2019. In addition, the number of banks in the Very Bad cluster increased sharply in 2023 compared to 2019. Although in general there has been an improvement in bank performance as shown in the increase in bank financial ratios in 2023 compared to 2019, bank managers, investors and bank supervisory authorities need to be aware of banks in the Very Bad cluster because their performance is not as good as other banks. With the limitations of the clustering analysis in this study which only uses data from 2019 and 2023 and uses 5 ratios (CAR, LDR, NIM, ROA and ROE), it is hoped that bank managers, investors and authorities can conduct periodic analysis to evaluate the performance of banks with their peers and serve as a reference for further research.

References

- Berger, A. N., Molyneux, P., & Wilson, J. O. (2020). Banks and the real economy: An assessment of the research. *Journal of Corporate Finance*, *62*, 101513. <https://doi.org/10.1016/j.jcorpfin.2019.101513>
- Carindri, F., & Untara. (2019). The Effect of Risk, Profitability, and liquidity on Capital Adequacy. *Journal of Business Economics*, *24*(1), 36–50. <https://doi.org/10.35760/eb.2019.v24i1.1854>
- Collier, P. M. (2003). *Accounting for Managers: Interpreting accounting information for decision-making*. John Wiley & Sons Ltd.
- Hantono. (2017). Effect of Capital Adequacy Ratio (CAR), Loan to Deposit Ratio (LDR) and Non Performing Loan (NPL) to Return On Assets (ROA) Listed in Banking in Indonesia Stock Exchange. In *International Journal of Education and Research* (Vol. 5, Issue 1). www.ijern.com
- Herman, E., Zsido, K. E., & Fenyves, V. (2022). Cluster Analysis with K-Mean versus K-Medoid in Financial Performance Evaluation. *Applied Sciences (Switzerland)*, *12*(16). <https://doi.org/10.3390/app12167985>
- Indonesia Financial Services Authority (OJK). (2020). *The Indonesian Financial Services Sector Master Plan 2021-2025*. <https://ojk.go.id/id/berita-dan-kegiatan/publikasi/Documents/Pages/Master-Plan-Sektor-Jasa-Kuangan-Indonesia-2021->

2025/The%20Indonesian%20Financial%20Services%20Sector%20Master%20Plan%202021-2025.pdf

- Irfani, A. S. (2020). *Manajemen Keuangan dan Bisnis - Teori dan Aplikasi* (Bernadine, Ed.; Pertama). PT Gramedia Pustaka Utama.
- Kurniawan, D. (2022). *Pengenalan Machine Learning dengan Python* (Edisi Digital). PT Elek Media Komputindo.
- Kusumaningrum, I., & Heikal, J. (2023). Evaluating The Prospects of Financial Performance After Merger at PT Pelabuhan Indonesia (Persero). *Adpebi International Journal of Multidisciplinary Sciences*, 2(2), 110–125. <https://doi.org/10.54099/aijms.v2i2.493>
- Luh Shintya Anggari, N., & Made Dana, I. (2020). The Effect of Capital Adequacy Ratio, Third Party Funds, Loan to Deposit Ratio, Bank Size on Profitability in Banking Companies on IDX. *American Journal of Humanities and Social Sciences Research (AJHSSR)*, 4(12), 334–338. www.ajhssr.com
- Madhulatha, T. S. (2012). An Overview On Clustering Methods. *IOSR Journal of Engineering*, 2(4), 719–725. www.iosrjen.org
- Otoritas Jasa Keuangan. (n.d.). *Implementasi Basel*. Retrieved April 11, 2023, from <https://www.ojk.go.id/id/kanal/perbankan/implementasi-basel/Pages/Road-Map.aspx>
- Özari, Ç., & Can, E. N. (2023). Financial Performance Evaluating and Ranking Approach for Banks in Bist Sustainability Index Using Topsis and K-Means Clustering Method. *Academic Journal of Interdisciplinary Studies*, 12(1), 34–50. <https://doi.org/10.36941/ajis-2023-0004>
- Ratnawati, A., Susanto, B., Saepudin, Rudianto, D., Herdiyanti, G., & Khalingga, M. A. (2022). The Effect Of Impaired Loan And Capital Adequacy Ratio (CAR) To Banking Performance At Private National Bank (Listed On Indonesia Stock Exchange 2015-2019). *Adpebi International Conference on Management, Education, Social Science, Economics and Technology (AICMEST)*. <http://adpebipublishing.com/index.php/AICMEST/article/view/95/82>
- Shalev-Shwartz, S., & Ben-David, S. (2014). *Understanding Machine Learning: From Theory to Algorithms* (First). Cambridge University Press. <http://www.cs.huji.ac.il/~shais/UnderstandingMachineLearning>
- Turkes, M. C. (2017). Cluster Analysis of Total Assets Provided By Banks from Four Continents. *Academic Journal of Economic Studies*, 4(4), 24–28. <https://www.ceeol.com/search/article-detail?id=592287>
- Vozniakovska, K., Tarasenko, I., Saienko, V., Kirizleyeva, A., Harashchenko, L., & Bodnar, O. (2022). Comparative Characteristics of the Banking Sector in Eastern Europe. *International Journal of Computer Science and Network Security (IJCSNS)*, 22(1), 1–22. <https://doi.org/10.22937/IJCSNS.2022.22.1.84>