Eur. J. Stat. 6 (2026) 17
doi: 10.28924/ada/stat.6.17

Integrating Density-Based Clustering and Sentiment Analysis for Portfolio Optimization: An Empirical Study on IDX ESG Leaders

Fausania Hibatullah*, Rico Dwi Firmansyah, Dwi Endah Kusrini

Department of Business Statistics, Faculty of Vocation, Institut Teknologi Sepuluh Nopember, Indonesia
fausania.hibatullah@its.ac.id, 2043221080@student.its.ac.id, dwi_endah@statistika.its.ac.id
**Correspondence: fausania.hibatullah@its.ac.id*

Abstract. The heightened uncertainty in Indonesia's capital market in early 2025, driven by capital outflows and macroeconomic instability, underscores the need for more robust data-driven portfolio construction frameworks beyond conventional fundamental analysis. This study develops an integrated statistical learning framework that combines transformer-based sentiment analysis and density-based clustering to improve portfolio optimization for stocks listed in the IDX ESG Leaders index. Market sentiment is quantified from financial news using the FinBERT model, producing structured sentiment scores that are incorporated as exogenous inputs. Subsequently, the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm is applied to identify latent structures within the asset space based on sentiment similarity, enabling non-parametric cluster formation without imposing distributional assumptions. The empirical findings indicate that IDX ESG Leaders stocks are characterized by predominantly positive sentiment distributions, with clustering results revealing two principal clusters and one noise component. To assess portfolio performance, the resulting clusters are integrated into a mean–variance optimization framework. The results demonstrate that portfolios constructed using sentiment-informed clustering achieve superior risk-adjusted performance, reflecting enhanced diversification and reduced estimation bias relative to conventional approaches. In contrast, portfolios exhibiting high concentration or sentiment homogeneity tend to yield lower efficiency, despite strong fundamental characteristics. Overall, this study demonstrates the statistical value of incorporating unstructured textual information and unsupervised learning techniques into portfolio optimization, providing empirical evidence that supports the integration of sentiment-driven features in data-driven asset allocation strategies.

1. Introduction

Since the beginning of 2025, the Indonesian capital market has experienced significant pressure due to the phenomenon of capital outflows. Based on net purchase by foreigners' data from the Indonesia Stock Exchange (IDX), the largest net sell occurred on 16 April 2025, amounting to IDR 8.211 billion, which triggered a 30-minute trading halt. Cumulatively (year-to-date),

Received: 1 May 2026.
Key words and phrases. DBSCAN; financial text mining; FinBERT; mean-variance model; risk-adjusted performance.

by the end of April 2025, foreign investors had recorded net sales totaling IDR 50.718 trillion. This volatility occurred amid rising domestic economic uncertainty, including weakening fiscal conditions, speculation over a change in the Minister of Finance, and sovereign rating downgrades by international agencies. At the same time, external pressures, particularly escalating trade tensions between the United States and China, further deteriorated global investor sentiment [1]. This heightened uncertainty encourages positive feedback trading behaviors, whereby investors overreact to short-term market information that does not reflect the intrinsic value of stocks [2].

Given the critical role of sentiment in influencing investor behaviors and stock price movements, a credible information source is required to accurately capture market perceptions. One Indonesian news platform that can serve as a sentiment source is *Bisnis.com*. In addition to providing up-to-date information, this news portal offers content that is both relevant and specific to the topics or listed company codes being searched. Such news coverage shapes public perceptions, which can be extracted through sentiment analysis, defined as a computational process for identifying positive, neutral, or negative opinions from textual data [3]. More specifically, positive sentiment can reduce the estimated expected return of a portfolio, thereby encouraging an increase in asset prices [2].

The decline in investor confidence regarding policy stability and governance in the Indonesian capital market has triggered capital outflows in early 2025, despite the continued presence of many fundamentally strong listed companies [4]. One such group is the IDX ESG Leaders index, which comprises stocks with high liquidity, strong financial performance, and adherence to ESG principles [5]. An optimal portfolio constructed from IDX ESG Leaders stocks has been shown to generate higher returns (11.18% per semester) with lower risk compared to the market index [6]. Nevertheless, the IDX ESG Leaders index does not fully capture the broader variation in financial ratios and market sentiment, thereby necessitating further analysis to evaluate individual constituent firms. Return on Assets (ROA) and Return on Equity (ROE) are employed to measure internal financial performance and a company's efficiency in utilising its assets and equity, while the Price-to-Earnings Ratio (PER) reflects market expectations regarding earnings growth [7]. In addition, liquidity ratios have a significant influence on stock prices and serve to assess a firm's ability to meet short-term obligations [8]. These financial ratios are subsequently combined with the percentage of positive sentiment and clustered using the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm. A comparative clustering study involving K-Means, Mean-Shift, Agglomerative, and DBSCAN methods finds that DBSCAN outperforms the alternatives due to its ability to handle non-spherical data structures and to generate a more diversified portfolio, with an expected return of 0.29 and a standard deviation of 0.39 [9]. The resulting clusters of listed companies are then allocated investment weights using the Markowitz portfolio optimization model, which has been shown to outperform other approaches by achieving the same level of return with 21.16% lower risk and a more diversified portfolio [10]. As a supporting metric, the Sharpe Ratio is employed to enable investors to compare portfolios along the efficient frontier and identify the optimal portfolio in terms of risk–return trade-off [11].

This study aims to examine the prevailing market sentiment toward IDX ESG Leaders stocks in early 2025 amid heightened macroeconomic uncertainty in Indonesia as a developing market, to analyse how the integration of firm-level financial ratios and positive news sentiment through DBSCAN clustering differentiates stock characteristics, and to comparatively evaluate how these distinctions are reflected in portfolio structure, risk–return, and risk-adjusted performance, with particular emphasis on the relative influence of sentiment and fundamental factors in determining portfolio efficiency.

2. Research Methods

2.1. **FinBERT Sentiment Analysis.** Bidirectional Encoder Representations from Transformers (BERT) is a machine learning model for Natural Language Processing (NLP) developed by Google [12]. However, as BERT is pre-trained on general-domain text, it may not perform optimally when processing financial texts written for professional investors. FinBERT addresses this limitation by constructing a vector space that is specific to the financial text context [13]. The FinBERT methodology begins with collecting textual data from the Bisnis.com news platform, followed by segmenting the text into chunks of up to 512 tokens per segment. Each chunk is then converted into tensor inputs and fed into the FinBERT model to produce logit outputs [14], as expressed in Eq. (1).

$$\mathbf{z}^{(q)} = \mathrm{FinBERT}\left(\mathbf{c}^{q}\right) = \left[z_{\text{pos}}^{(q)},\ z_{\text{neg}}^{(q)},\ z_{\text{neu}}^{(q)}\right] \tag{1}$$

In Eq. (1), $\mathbf{z}^{(q)}$ denotes the logit scores of text chunk q for the classes (positive, negative, and neutral). Subsequently, $\mathbf{z}^{(q)}$ is transformed into probabilities [14], as shown in Eq. (2).

$$p_j^{(q)} = \frac{e^{z_j^{(q)}}}{e^{z_{\text{pos}}^{(q)}} + e^{z_{\text{neg}}^{(q)}} + e^{z_{\text{neu}}^{(q)}}} \tag{2}$$

In Eq. (2), $p_j^{(q)}$ represents the probability of the j-th class for the q-th chunk, where j denotes the sentiment class, and $z_j^{(q)}$ corresponds to the logit value of the q-th chunk. All chunks are subsequently aggregated to produce a global sentiment measure, as presented in Eq. (3).

$$p_{\text{news}} = \frac{1}{N}\sum_{q=1}^{N} \mathbf{p}^{(q)} = \left[P_{\text{pos}},\ P_{\text{neg}},\ P_{\text{neu}}\right] \tag{3}$$

In Eq. (3), p_{news} represents the final article-level probability, computed as the average sentiment score across N chunks, with the highest value determining the article's sentiment. Firm-level positive sentiment percentages are then derived from all related news articles.

2.2. **DBSCAN Cluster Analysis.** Density-Based Spatial Clustering of Applications with Noise (DBSCAN) is a fundamental density-based clustering algorithm. DBSCAN is capable of identifying clusters of varying shapes and sizes within large datasets that contain noise and outliers. The algorithm relies on two key parameters, namely *minPts* and *Eps* (ε) [15]. Visualization of cluster formation using the DBSCAN method is presented in Figure 1.

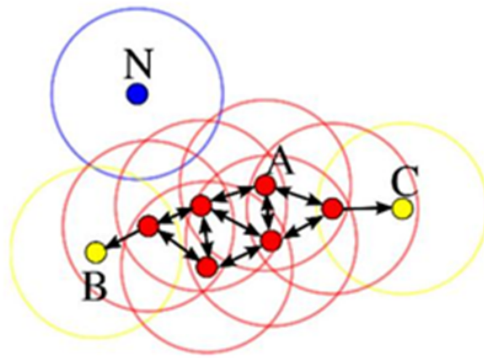

FIGURE 1. DBSCAN Process (Source : Han dkk., 2021)

DBSCAN operates by iteratively expanding clusters from each data point until all points have been visited. Points that do not belong to any cluster are labelled as outliers. In this context, point A represents a core point, while points B and C are border points, defined as points that do not have sufficient neighbors within the Eps radius to be considered core points but lie within the Eps radius of a core point. Point N denotes an outlier, which neither has sufficient neighboring points within the Eps radius nor falls within the neighborhood of any core [16]. DBSCAN parameters are selected through iterative trial-and-error and evaluated using the Silhouette score, defined by the minimum average distance to the nearest cluster $b(i)$ and the average intra-cluster distance $a(i)$, as specified in Eq. (4).

$$s(i) = \frac{b(i) - a(i)}{\max[b(i),\ a(i)]} \tag{4}$$

2.3. **Markowitz Portfolio Optimization.** Markowitz portfolio theory emphasises the maximisation of expected return and the minimisation of risk in constructing an optimal portfolio. Through the concept of diversification, investors can maximise the expected return of their investments for a given level of risk or, alternatively, minimise risk for a targeted level of return [17]. The Markowitz model is based on several assumptions, including the absence of transaction costs, investor preferences determined solely by expected return and portfolio risk, and the exclusion of risk-free borrowing and lending. The concept of diversification in the Markowitz framework is quantified using statistical measures, particularly covariance and correlation [18]. The stock return is calculated using the closing price at period t for each stock comprising the g-th portfolio ($g = 1, 2, \ldots, n$), as expressed in Eq. (5).

$$R_{g,t} = \frac{P_{g,t} - P_{g,t-1}}{P_{g,t-1}} \tag{5}$$

The expected return of each stock is calculated using the geometric mean, as expressed in Eq. (6).

$$E(R_g) = \prod_{t=1}^{T} (1 + R_{g,t})^{1/T} - 1 \tag{6}$$

After obtaining the expected returns, stock risk is measured by calculating the standard deviation and the covariance of stock returns to ensure adequate diversification. Consequently, the expected portfolio return is derived as presented in Eq. (7).

$$E(R_p) = \sum_{g=1}^{n} W_g \cdot E(R_g) \tag{7}$$

In Eq. (7), $E(R_p)$ denotes the expected return of the stock portfolio, where W_g represents the proportion of funds invested in the g-th stock and $E(R_g)$ denotes the expected return of the g-th stock. The portfolio risk is obtained using the portfolio standard deviation, as defined in Eq. (8).

$$S_p = \sqrt{\sum_{g=1}^{n} \sum_{h=1}^{n} W_g W_h S_{g,h}} \tag{8}$$

In Eq. (8), S_p represents the standard deviation of the stock portfolio, W_g denotes the weight of funds invested in the g-th stock, and $S_{g,h}$ represents the covariance between the g-th and the h-th stocks.

2.4. **Portfolio Performance (Sharpe Ratio).** Portfolio construction is closely related to the fundamental concepts of efficient and optimal portfolios. An efficient portfolio is a combination of assets that provides the highest return for a given level of risk, or the lowest risk for a given level of return. In contrast, an optimal portfolio is the portfolio selected by investors from the set of efficient portfolios [19]. The Sharpe Ratio is a widely used measure of portfolio performance, indicating the amount of excess return (return above the risk-free rate) earned per unit of total risk [20]. The calculation of the Sharpe Ratio as the difference between the expected portfolio return and the risk-free rate, divided by the portfolio risk, is presented in Eq. (9).

$$\text{Sharpe} = \frac{E(R_p) - R_{br}}{S_p} \tag{9}$$

3. Results and Discussion

This study proposes an integrated framework combining news-based sentiment, financial ratios, and portfolio optimization. News articles related to IDX ESG Leaders are scraped from Bisnis.com and analyzed using FinBERT to obtain positive sentiment percentages. Financial ratios, including ROA, ROE, PER, and liquidity, are then calculated and jointly used with sentiment measures for DBSCAN clustering. For each cluster, Markowitz optimization is applied to construct efficient and optimal portfolios, enabling a comparative evaluation of risk-adjusted performance across clusters. A comparison of portfolios based on financial ratios and the percentage of positive sentiment is discussed further in the subsequent section.

3.1. **Sentiment Analysis of IDX ESG Leaders.** This study analyses the period from January to June 2025 using 30 listed stocks based on the "Appendix to IDX Announcement No. Peng-00073/BEI.POP/04-2025 dated 24 April 2025" from the official Indonesia Stock Exchange website. A screening process is applied to exclude firms with returns below the risk-free rate daily (0.015%), resulting in 9 selected companies, with PGEO exhibiting the highest daily expected return of 0.38%. News data are collected from Bisnis.com using web scraping with the BeautifulSoup library in Python, excluding inaccessible premium content. The FinBERT model employed is `yiyanghkust/finbert-tone` by Huang et al. (2022), and the percentage of positive sentiment for each firm is reported in Table 1.

Table 1. Percentage of Positive Sentiment.

Company	Positive	Total Sentiment	Percentage
AKRA	46	55	83.636%
AVIA	13	20	65.000%
BNGA	18	29	62.069%
CMRY	9	12	75.000%
ERAA	25	35	71.429%
EXCL	34	47	72.340%
PGAS	70	85	82.353%
PGEO	36	48	75.000%
TLKM	98	119	82.353%

Table 1 shows the percentage of positive sentiment obtained from news data for the nine companies that successfully went through the risk-free return evaluation. AKRA shows the highest level of positive sentiment at 83.64%, whereas BNGA has the lowest level at 62.07%. In general, most companies exhibit a strong positive sentiment exceeding 74.353%, which suggests that there are relatively optimistic market views regarding these companies in the period from January to June 2025.

3.2. **Clustering of Listed Companies.** The listed companies are analysed using boxplot visualisations for each clustering variable, which are presented in Figure 2.

Figure 2 shows that ROA, PER, liquidity ratios, and positive sentiment exhibit relatively high variability across companies, whereas ROE displays lower variability with a more balanced distribution. ROA and liquidity ratios are skewed downwards, indicating that most companies operate at relatively lower levels, with a few positive outliers reflecting specific sector characteristics and financial performance. In contrast, PER and positive sentiment demonstrate more even distributions without significant outliers, suggesting a relatively uniform market perception of IDX ESG Leaders constituents.

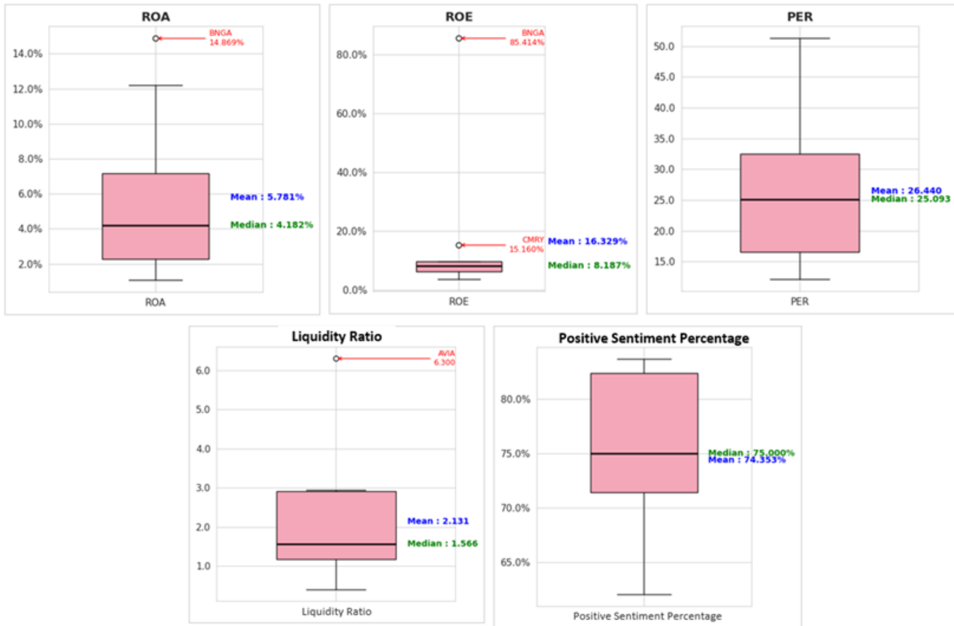

Figure 2. Boxplots of Positive Sentiment Percentage and Financial Ratios

The DBSCAN parameters are determined using grid search hyperparameter tuning. The optimal configuration yields the highest Silhouette Score of 0.2804 with Epsilon set to 2.5 and MinPts to 2. The resulting clusters are presented in Table 2.

Table 2. Results of Clustering Analysis.

Cluster	**Companies**
1	AKRA, ERAA, EXCL, PGAS, TLKM
2	CMRY, PGEO
Noise	AVIA, BNGA

Table 2 indicates that clustering the IDX ESG Leaders constituents using the DBSCAN method results in two clusters and one noise cluster. This outcome highlights DBSCAN's strength in identifying observations with extreme characteristics as noise, without forcing such companies into unsuitable clusters. The characteristics of the variables within each cluster are subsequently presented in Table 3.

Table 3 indicates differences in the characteristics of financial ratios, as well as the percentage of positive sentiment across clusters. Cluster 1 exhibits a lower average PER, suggesting relatively undervalued companies, along with higher positive sentiment, reflecting optimistic market perceptions and stronger reputational standing. In contrast, companies in the Noise cluster demonstrate higher ROA and ROE, indicating greater efficiency in utilizing assets and

TABLE 3. Characteristics of the Formed Company Clusters.

Cluster	ROA	ROE	PER	Liquidity Ratio	Positive Sentiment (%)
1	3.105%	6.930%	**20.647**	1.107	**78.422%**
2	7.226%	9.354%	45.077	2.922	75.000%
Noise	**11.027%**	**46.800%**	22.286	**3.901**	63.534%

equity to generate profits. This cluster also shows higher liquidity ratios, implying a stronger ability to meet short-term obligations.

3.3. **Portfolio Optimization Using the Markowitz Model.** Diversification is ensured by analysing covariance and correlation among returns, as shown in Figure 3.

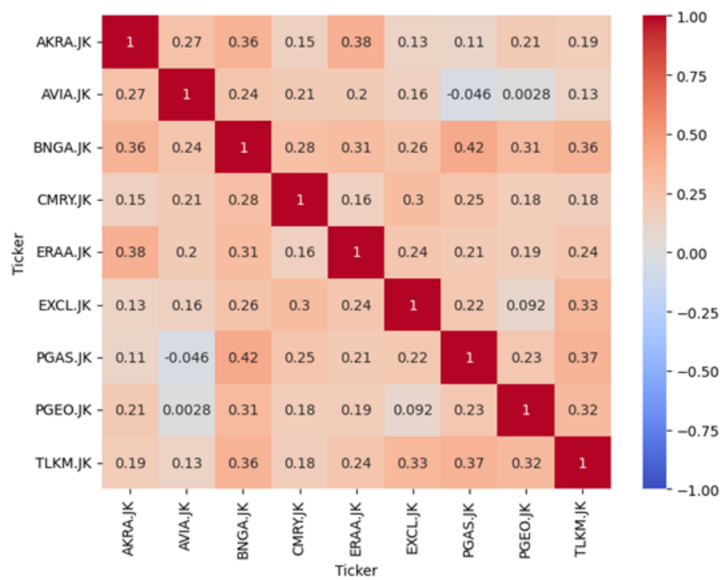

FIGURE 3. Correlation and Covariance Matrix of Portfolio Constituents.

Figure 3 shows that the candidate portfolio constituents exhibit relatively low correlations across diverse industries, indicating suitability for portfolio formation. Three clusters are identified and subsequently optimised using the Markowitz model and the Sharpe Ratio to obtain optimal and efficient portfolios for each cluster. The target returns are generated as 48 evenly spaced values between the minimum and maximum expected returns, where the maximum expected return represents the highest possible target among the candidate portfolios.

3.3.1. *Portfolio Cluster 1.* Cluster 1 consists of five candidate stocks that are combined to form a portfolio. Table 4 presents the resulting portfolio, which is evaluated using the Sharpe Ratio.

Table 4 shows that, within Cluster 1, Portfolio 2 represents the efficient portfolio (denoted by *), exhibiting the lowest risk of 1.110%, and is dominated by EXCL and PGAS, thereby reflecting a defensive strategy aimed at minimising volatility. Meanwhile, Portfolio 29 constitutes

Table 4. Portfolio Composition in Cluster 1

No.	AKRA	ERAA	EXCL	PGAS	TLKM	S_p	$E(R_p)$	Sharpe Ratio
1	4.37%	0.00%	88.33%	7.21%	0.08%	1.124%	0.041%	2.339%
2*	3.80%	0.00%	81.31%	11.25%	3.64%	1.110%	0.046%	2.779%
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
28	0.00%	41.38%	0.24%	35.68%	22.70%	2.285%	0.164%	6.535%
29**	0.00%	44.08%	0.00%	33.17%	22.75%	2.354%	0.169%	6.536%
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
47	0.00%	94.41%	0.00%	2.85%	2.74%	3.974%	0.251%	5.937%
48	0.00%	97.20%	0.00%	1.45%	1.35%	4.075%	0.256%	5.901%

* Efficient portfolio (minimum risk); ** Optimal portfolio (maximum Sharpe Ratio).

the optimal portfolio (denoted by **), achieving the highest Sharpe Ratio of 6.536%, indicating that each 1% of risk borne by investors generates an additional return of 6.536% above the risk-free rate. The optimal portfolio in Cluster 1 demonstrates a more evenly distributed allocation across ERAA, PGAS, and TLKM compared to the efficient portfolio. The visualisation of Cluster 1 portfolios is presented through the efficient frontier curve in Figure 4.

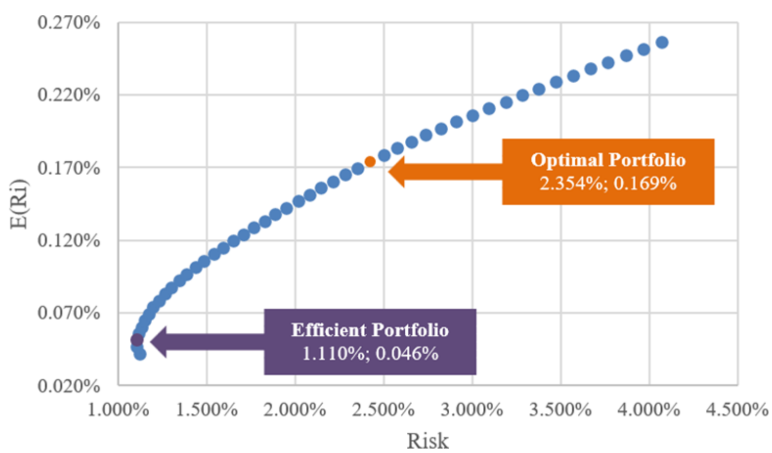

Figure 4. Efficient Frontier of the Cluster 1 Portfolio.

3.3.2. *Portfolio Cluster 2.* Cluster 2 consists of two candidate stocks (CMRY and PGEO) that are combined to form a portfolio. Table 5 presents the resulting portfolio, which is evaluated using the Sharpe Ratio.

Table 5 indicates that, within Cluster 2, improvements in the Sharpe Ratio are achieved through a very high concentration of weights in PGEO, particularly in the optimal portfolio. As a result, although risk-adjusted performance appears stronger, the portfolio structure becomes

Table 5. Portfolio Composition in Cluster 2

No.	CMRY	PGEO	S_p	$E(R_p)$	Sharpe Ratio
1	97.96%	2.041%	2.418%	0.03%	0.527%
2	95.92%	4.082%	2.388%	0.04%	0.871%
⋮	⋮	⋮	⋮	⋮	⋮
10	79.59%	20.408%	2.255%	0.10%	3.780%
11*	77.55%	22.449%	2.254%	0.11%	4.140%
⋮	⋮	⋮	⋮	⋮	⋮
47	4.08%	95.918%	3.900%	0.40%	9.827%
48**	2.04%	97.959%	3.973%	0.41%	9.851%

* Efficient portfolio; ** Optimal portfolio.

less diversified and is accompanied by relatively higher total risk. This condition suggests that portfolio efficiency in this cluster is driven more by the dominance of a single asset rather than by a well-balanced asset combination. These characteristics are consistent with the profile of Cluster 2, which exhibits relatively strong positive market sentiment, reflecting investor confidence in the firm's prospects, while simultaneously indicating a potential trade-off between market optimism and increased concentration risk in the optimal portfolio. The efficient frontier visualisation is further illustrated in Figure 5.

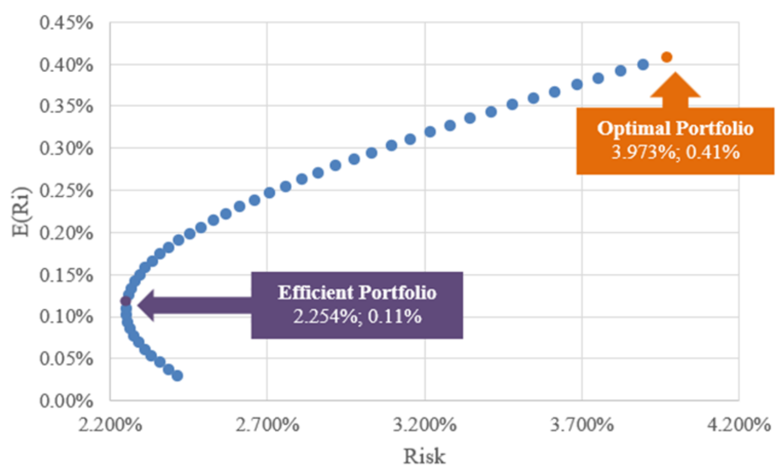

Figure 5. Efficient Frontier of the Cluster 2 Portfolio.

3.3.3. *Portfolio Cluster Noise.* The noise cluster consists of two candidate stocks (AVIA and BNGA) that are combined to form a portfolio. Table 6 presents the resulting portfolio, which is evaluated using the Sharpe Ratio.

Table 6. Portfolio Composition in Cluster Noise

No.	AVIA	BNGA	S_p	$E(R_p)$	Sharpe Ratio
1	2.041%	97.959%	1.384%	0.022%	0.529%
2	4.082%	95.918%	1.368%	0.023%	0.568%
⋮	⋮	⋮	⋮	⋮	⋮
12	24.490%	75.510%	1.281%	0.027%	0.949%
13*	26.531%	73.469%	1.281%	0.028%	0.984%
⋮	⋮	⋮	⋮	⋮	⋮
41**	83.673%	16.327%	1.806%	0.040%	1.380%
42	85.714%	14.286%	1.838%	0.040%	1.379%
⋮	⋮	⋮	⋮	⋮	⋮
48	97.959%	2.041%	2.039%	0.043%	1.373%

* Efficient portfolio; ** Optimal portfolio.

Table 6 shows that, within the noise cluster, the efficient portfolio is represented by Portfolio 13 with a risk level of 1.281%, while the optimal portfolio corresponds to Portfolio 41, achieving a Sharpe Ratio of 1.380%. Despite the constituent stocks exhibiting higher ROA, ROE, and liquidity ratios, the resulting portfolios generate the lowest Sharpe Ratios, even though their overall risk remains relatively low. This finding indicates that strong fundamental performance alone is insufficient to produce optimal risk-adjusted performance when investor sentiment is at its lowest level. Pessimistic market conditions constrain return potential and amplify investor reactions to short-term information, thereby reducing portfolio efficiency. The efficient frontier visualisation for the noise cluster is further illustrated in Figure 6.

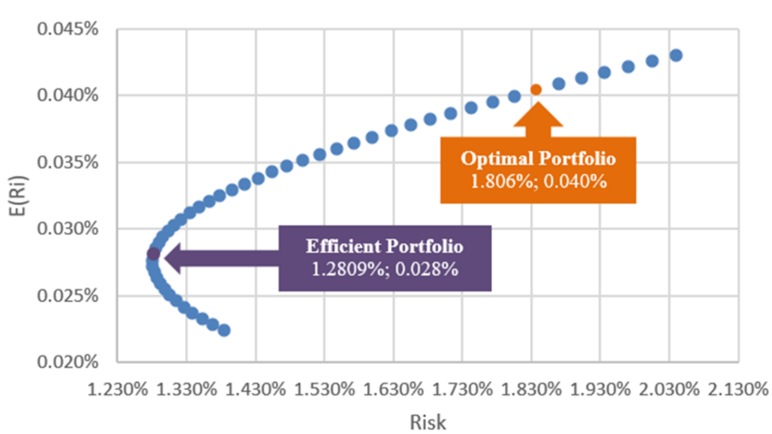

Figure 6. Efficient Frontier of the Cluster Noise Portfolio

4. Conclusion

The results obtained are consistent with prior findings indicating that positive news exerts a stronger influence on stock return volatility than negative news of equal magnitude in frontier markets [21]. Minor differences in portfolio performance arise from variations in asset composition and market conditions. The following conclusions are drawn from this study.

(1) Sentiment analysis of news articles from Bisnis.com reveals that IDX ESG Leaders stocks are characterized by a predominantly positive market sentiment, with an average positive sentiment of 74.35%, reflecting optimistic investor expectations in the Indonesian equity market.
(2) Stock clustering using the DBSCAN method based on financial ratios and positive sentiment percentage results in two main clusters and one noise cluster.
(3) A comparative analysis of Markowitz-optimized portfolios across clusters indicates that risk-adjusted performance, as measured by the Sharpe Ratio (Eq. (9)), is shaped by the interaction between portfolio construction and investor sentiment. Cluster 1 delivers superior Sharpe performance through relatively balanced diversification, Cluster 2 achieves higher Sharpe values driven by asset concentration at the cost of increased risk, while the Noise cluster demonstrates that strong financial fundamentals without supportive sentiment fail to translate into superior risk-adjusted returns. These findings highlight the critical role of market sentiment in frontier markets such as Indonesia, where investor behavior significantly influences the transmission of information into stock prices and portfolio efficiency.

Competing interests: The authors declare that there is no conflict of interest regarding the publication of this paper.

References

[1] N. Hafifa, S. Maisaroh, P.N. Umaira, W. Pangestoeti, Dampak Penurunan IHSG terhadap Stabilitas Ekonomi Publik Indonesia: Analisis Penyebab, Implikasi, dan Strategi Kebijakan, Socius J. Penelit. Ilmu-Ilmu Sos. 2 (2025), 630–634. https://doi.org/10.5281/zenodo.15703818.

[2] X. Liu, B. Han, L. Li, Impact of Investor Sentiment on Portfolio, Emerg. Mark. Financ. Trade 59 (2023), 880–894. https://doi.org/10.1080/1540496X.2022.2119805.

[3] A.G. Tando, M.I. Irawan, Analisis Dinamika Harga Saham yang Dipengaruhi oleh Analisis Sentimen di Media Sosial Menggunakan Algoritma Support Vector Machine, J. Sains Seni ITS 12 (2023), A47–A53. https://doi.org/10.12962/j23373520.v12i1.107080.

[4] R.N.A. Bastian, I. Permatasari, The Trend of Integrated Reporting in Emerging Markets, Int. J. Econ. Manag. Sci. 2 (2025), 282–296. https://doi.org/10.61132/ijems.v2i2.807.

[5] T.A. Singgih, E.L. Maretha, Financial Performance, Environmental, Social, Governance, and Stock Return of ESG Leaders Index, J. Ris. Akunt. Keuang. 13 (2025), 249–270. https://doi.org/10.17509/jrak.v13i1.77277.

[6] F.I.K. Irawan, I.P.S. Sanjaya, Portofolio Optimal Model Markowitz Saham IDX ESG Leaders dalam Terang Laudato Si', Krisna Kumpul. Ris. Akunt. 16 (2025), 198–208. https://doi.org/10.22225/kr.16.2.2025.198-208.

[7] F.N.R. Said, M.I. Musa, A. Anwar, A.M. Amin, A.P. Aslam, Analisis Kinerja Keuangan Perusahaan pada Sektor Transportasi dan Logistik yang Terdaftar di BEI Periode 2019-2023, Trend: Jurnal Manaj. Ekon. 3 (2025), 62–79. https://doi.org/10.30640/trending.v3i2.3916.

[8] N.A. Salsabila, T. Miranti, Faktor Pengaruh Rasio Keuangan terhadap Harga Saham Perusahaan Jakarta Islamic Index (JII), El Muhasaba Jurnal Akunt. 12 (2021), 42–55. https://doi.org/10.18860/em.v12i1.10123.

[9] I. Palupi, B.A. Wahyudi, I. Indwiarti, The Clustering Algorithms Approach for Decision Efficiency in Investment Portfolio Diversification, in: 2019 7th International Conference on Information and Communication Technology (ICoICT), IEEE, 2019, pp. 1–6. https://doi.org/10.1109/ICoICT.2019.8835314.

[10] I. Azis, M. Syazali, E.N. Manurung, The Optimal Portfolio of Islamic Stocks Using the Markowitz and Black-Litterman Models, Desimal: J. Mat. 6 (2023), 1–6. https://doi.org/10.24042/djm.v6i1.14499.

[11] X. Lin, X. Xu, Y. Xu, Investment Portfolio Establishment Based on Markowitz Model and Highest Sharp Ratio, in: Proceedings of the 4th International Conference on Economic Management and Model Engineering, ICEMME 2022, November 18-20, 2022, Nanjing, China, EAI, 2023. https://doi.org/10.4108/eai.18-11-2022.2326902.

[12] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, arXiv:1810.04805, 2018. https://doi.org/10.48550/arXiv.1810.04805.

[13] A.H. Huang, H. Wang, Y. Yang, FinBERT: A Large Language Model for Extracting Information from Financial Text, Contemp. Account. Res. 40 (2023), 806–841. https://doi.org/10.1111/1911-3846.12832.

[14] W. Cheng, S. Chen, Sentiment Analysis of Financial Texts Based on Attention Mechanism of FinBERT and BiLSTM, in: 2021 International Conference on Computer Engineering and Application (ICCEA), IEEE, 2021, pp. 73–78. https://doi.org/10.1109/ICCEA53728.2021.00022.

[15] T.A. Cinderatama, R.Z. Alhamri, Y. Yunhasnawa, Implementasi Metode K-Means, DBSCAN, dan MeanShift untuk Analisis Jenis Ancaman Jaringan pada Intrusion Detection System, Inovtek Polbeng - Seri Inform. 7 (2022), 169. https://doi.org/10.35314/isi.v7i1.2336.

[16] X. Han, C. Armenakis, M. Jadidi, Modeling Vessel Behaviours by Clustering AIS Data Using Optimized DBSCAN, Sustainability 13 (2021), 8162. https://doi.org/10.3390/su13158162.

[17] U. Muthohiroh, R. Rahmawati, D. Ispriyanti, Pendekatan Metode Markowitz untuk Optimalisasi Portofolio dengan Risiko Expected Shortfall (ES) pada Saham Syariah Dilengkapi GUI MATLAB, J. Gaussian 10 (2021), 445–454.

[18] A.R. Ramadhan, P. Iswanto, D.H. Perkasa, Analisis Portofolio Optimal Saham Indeks Infobank 15 dengan Model Markowitz untuk Pengambilan Keputusan Investasi Pasca Pandemi COVID-19, J. Akunt. Manaj. Bisnis 3 (2023), 53–60. https://doi.org/10.56127/jaman.v3i3.1093.

[19] M.F. Mingka, R.S. Lubis, Analisis Portofolio Saham Optimal dengan Metode Markowitz dan Model Indeks Tunggal pada Saham Perbankan Bursa Efek Indonesia, J. Lebesg.: J. Ilm. Pendidik. Mat. Mat. Stat. 4 (2023), 709–727. https://doi.org/10.46306/lb.v4i2.322.

[20] M.A. Wulan, R.S.D. Putra, Pengukuran Kinerja Portofolio Saham dengan Pendekatan Sharpe, Treynor, dan Jensen: Studi pada Jakarta Islamic Index (JII), Researchgate, 2025. https://doi.org/10.13140/RG.2.2.18838.25920.

[21] K.O. Emenike, O.N. Enock, How Does News Affect Stock Return Volatility in a Frontier Market?, Manag. Labour Stud. 45 (2020), 433–443. https://doi.org/10.1177/0258042X20939019.