

# Analisis Faktor – Faktor Keberlanjutan Sektor Perbankan di Indonesia dengan Unsupervised Learning

Nadila Putri Prihanita<sup>1\*</sup>, Dyah Sulistyowati Rahayu<sup>2</sup>, dan Kemala Putri Ayunda<sup>3</sup>

<sup>1</sup>Jurusan Rekayasa Perangkat Lunak, Universitas Pendidikan Indonesia, Bandung

<sup>2</sup>Jurusan Teknik Informatika, Fakultas Teknik, Universitas Pancasila, Jakarta

<sup>3</sup>The Indonesian Association of Islamic Economist, Jakarta

**Abstrak.** Bank berperan penting dalam menyalurkan dana dan pembiayaan sektor riil. Saat ini investor modern semakin mempertimbangkan aspek keberlanjutan dalam keputusan investasinya. Penelitian ini menganalisis pengelompokan bank di Indonesia berdasarkan variabel keberlanjutan dan keuangan menggunakan pendekatan unsupervised learning. Data berasal dari laporan keberlanjutan dan laporan tahunan bank konvensional. Preprocessing data meliputi proses cleaning, imputasi, feature engineering, serta penentuan tier Environment, Social, and Governance (ESG). Metode yang digunakan meliputi K-Means, Agglomerative Clustering, DBSCAN, dan PCA. Hasilnya menunjukkan adanya kluster bank dengan karakteristik berbeda berdasarkan ukuran aset, ekuitas, dan skor ESG. PCA dua komponen mampu menjelaskan sekitar 31% variasi data dan memperlihatkan pemisahan kluster yang cukup jelas. Temuan ini menunjukkan bahwa variabel keberlanjutan memiliki pola tertentu ketika dianalisis bersama indikator keuangan, sehingga klusterisasi dapat menjadi alat pemetaan profil keberlanjutan bank di Indonesia.

**Kata kunci**—*bank; clustering ESG; keberlanjutan; unsupervised learning.*

## 1. PENDAHULUAN

Perbankan merupakan salah satu komponen utama dalam sistem keuangan yang memiliki peran sebagai penghubung antara pemilik dana dan pihak yang memerlukan pembiayaan[1]. Pada umumnya, bank dinilai dari indikator keuangan seperti aset, laba dan rasio profitabilitas. Namun, kesadaran terhadap lingkungan dan sosial yang meningkat, menjadikan pendekatan penilaian tersebut kurang memadai. Perbankan merupakan komponen kunci dalam sistem keuangan yang berperan sebagai penghubung antara pihak yang memiliki surplus dana dan pihak yang membutuhkan pembiayaan. Selama ini, kinerja bank umumnya diukur melalui indikator finansial seperti total aset, laba, dan berbagai rasio profitabilitas. Namun, seiring meningkatnya perhatian publik terhadap isu lingkungan dan sosial, pendekatan penilaian yang hanya berfokus pada aspek keuangan dianggap tidak lagi memadai untuk menggambarkan kualitas kinerja sebuah bank secara menyeluruh. Investor modern, terutama investor institusional dan investor asing, semakin memperhatikan aspek keberlanjutan perusahaan melalui indikator Environmental, Social, and Governance (ESG)[2].

Dalam konteks ESG, bank tidak hanya dituntut menjaga kinerja finansial, tetapi juga menunjukkan tanggung jawab terhadap lingkungan, sosial, dan tata kelola. Lembaga pemeringkat kemudian menyusun skor ESG untuk mengukur ketiga aspek tersebut secara kuantitatif[3]. Bagi investor, skor ini membantu menilai risiko jangka panjang dan potensi pertumbuhan berkelanjutan, sementara bagi regulator dan manajemen, informasi ini menjadi dasar penyusunan kebijakan yang lebih ramah lingkungan dan berkelanjutan[4].

Penelitian ini menganalisis bank-bank di Indonesia dari sisi keberlanjutan dan melihat bagaimana faktor ESG berkaitan dengan kondisi keuangan menggunakan pendekatan *unsupervised learning* [5]. Karena penelitian dengan metode ini masih terbatas, studi ini berupaya mengelompokkan setiap institusi berdasarkan gabungan indikator keberlanjutan dan finansial, kemudian dilanjutkan dengan melihat ciri khas dari tiap kluster, terutama terkait profil ESG dan ukuran keuangannya [6].

\* Corresponding author: [nadilaputriprihanita@upi.edu](mailto:nadilaputriprihanita@upi.edu)



## 2. METODE PENELITIAN

### a. *Data dan Variabel*

Data yang digunakan dalam penelitian ini bersumber dari laporan keberlanjutan dan laporan tahunan bank umum konvensional yang terdaftar di Bursa Efek Indonesia. Informasi yang dikumpulkan mencakup tahun 2015 hingga 2023 dengan variabel sesuai dengan ketersediaan data ESG pada masing-masing institusi. Setelah seluruh data mentah dihimpun, dilakukan tabulasi terhadap keseluruhan variabel yang ada [7].

Variabel yang dianalisis mencakup variabel ESG, variabel keuangan, dan variabel struktural. Variabel ESG meliputi skor pilar lingkungan (*environment*), sosial (*social*), dan tata kelola (*governance*), skor ESG gabungan, serta indikator keberlanjutan seperti total emisi pada Scope 1, Scope 2, dan Scope 3, juga konsumsi energi, air, bahan bakar, dan kertas. Variabel keuangan mencakup pendapatan, laba, total aset, total utang, total ekuitas, dividen, serta berbagai rasio yang dibentuk melalui *feature engineering* seperti Return on Asset (ROA), Return on Equity (ROE), Debt to Equity Ratio (DER), profit margin, emisi per aset, emisi per pendapatan, dan *green financing ratio*. Penelitian ini juga menggunakan variabel struktural berupa kategori kepemilikan, penanda dummy untuk BUMN, dan variasi komposisi kepemilikan. Selain itu, skor ESG serta ketiga pilarnya dikelompokkan ke dalam tier rendah, menengah, dan tinggi untuk mempermudah interpretasi dalam proses pengelompokan [8].

### b. *Prapengolahan Data*

Tahapan prapengolahan dilakukan untuk memastikan seluruh variabel dapat dianalisis secara optimal. Proses ini dimulai dari pembersihan data yang mencakup penyesuaian nama kolom, penghapusan spasi berlebih, serta konversi seluruh nilai yang relevan ke dalam tipe numerik. Setelah data rapi, dilakukan identifikasi dan penandaan outlier menggunakan metode IQR (inter-quartil range) untuk menangkap nilai-nilai ekstrem yang berpotensi memengaruhi hasil pengelompokan. Penanganan *missing value* dilakukan dengan mengisi kekosongan pada variabel numerik menggunakan rata-rata per entitas melalui *groupby* berdasarkan nama institusi. Jika setelah proses tersebut masih ditemukan nilai yang kosong, dilakukan imputasi tambahan menggunakan rata-rata keseluruhan kolom.

Setelah data bersih dan lengkap, variabel numerik kemudian distandarisasi menggunakan *StandardScaler* sehingga setiap fitur memiliki rata-rata nol dan standar deviasi satu. Standardisasi penting untuk mencegah variabel berskala besar, seperti total aset atau total emisi, mendominasi proses pembentukan kluster. Selain itu, skor ESG dan ketiga pilarnya dikelompokkan ke dalam tiga tingkat keberlanjutan—rendah, menengah, dan tinggi—dengan menggunakan fungsi *score\_to\_tier* untuk mempermudah interpretasi pada tahap analisis.

### c. *Metode Unsupervised Learning*

Penelitian ini menerapkan beberapa metode *unsupervised learning* untuk mengidentifikasi pola dan struktur pengelompokan pada data keberlanjutan dan keuangan. Unsupervised learning telah terbukti mampu mengidentifikasi resiko bank [9]. Metode utama yang digunakan adalah K-Means Clustering, yang bertujuan mengelompokkan institusi berdasarkan fitur yang telah distandarisasi. Pada analisis awal, K-Means diterapkan pada tiga variabel utama, yaitu total aset, total ekuitas, dan skor ESG, dengan jumlah kluster yang ditetapkan sebanyak tiga. Penetapan tiga kluster mempertimbangkan jumlah sampel serta kemudahan interpretasi secara ekonomis, yaitu kategori tinggi, sedang, dan rendah. Pada analisis multivariat, K-Means juga digunakan setelah proses reduksi dimensi dengan PCA menggunakan lima komponen utama sebagai bentuk eksplorasi tambahan.

Sebagai pembanding, metode Agglomerative Clustering digunakan untuk melihat konsistensi hasil pengelompokan. Teknik hirarkis ini menggunakan pendekatan *bottom-up* yang secara bertahap menggabungkan observasi dengan kemiripan tertinggi hingga membentuk kluster dengan jumlah yang sama seperti model K-Means. Selain kedua metode tersebut, DBSCAN turut diterapkan untuk mengidentifikasi apakah terdapat observasi yang bersifat *noise* atau outlier dan tidak tergabung ke dalam kluster mana pun, sehingga memberikan perspektif alternatif dalam melihat struktur data [10].

Untuk mendukung proses visualisasi, penelitian ini menggunakan Principal Component Analysis atau PCA. Teknik ini mereduksi variabel numerik menjadi dua komponen utama sehingga memungkinkan representasi dua dimensi dari hasil pengelompokan. Komponen utama pertama (PC1) dan kedua (PC2)



kemudian digunakan untuk menggambarkan pemisahan kluster serta mengevaluasi sejauh mana hasil pengelompokan memberikan pemisahan yang jelas dalam ruang fitur.

Tabel 1 Daftar Variabel Penelitian.

<b>Kategori</b>	<b>Variabel</b>	<b>Definisi</b>
ESG	Environmental Pillar Score	Skor kinerja aspek lingkungan
Isi tabel	Social Pillar Score	Skor kinerja aspek sosial
Isi tabel	Governance Pillar Score	Skor kinerja tata kelola
	ESG Score	Skor gabungan pilar E, S, dan G
Emisi dan Sumber Daya	Scope 1	Emisi langsung dari operasional
	Scope 2	Emisi tidak langsung dari konsumsi listrik
	Scope 3	Emisi tidak langsung dari rantai nilai
	Total Emisi	Jumlah Scope 1 + Scope 2 + Scope 3
	Konsumsi energi	Penggunaan energi (listrik/BBM)
	Konsumsi Air	Penggunaan air
	Konsumsi BBM	Penggunaan bahan bakar
	Konsumsi Kertas	Penggunaan kertas
Keuangan	Pendapatan	Total revenue bank
	Laba	Laba bersih bank
	Deviden	Pembagian keuntungan kepada pemegang saham
	Total Aset	Nilai total aset
	Total Utang	Total liabilitas
	Total Ekuitas	Ekuitas pemegang saham
	Feature Engineering/Rasio	ROA
ROE		Laba/ekuitas
DER		Total utang/ekuitas
Profit Margin		Laba/Pendapatan
Emisi per Aset		Total emisi/total aset
Emisi per pendapatan		Total emisi/pendapatan
Green Financing ratio		Pembiayaan hijau/total aset
Listrik per aset		Konsumsi listrik/total aset
Air per aset		Konsumsi air/total aset
BBM per aset		Konsumsi BBM/total aset

	Kertas per aset	Konsumsi kertas/ total aset
Struktural	Kategori pemilik	BUMN atau non BUMN
	Ls_BUMN	Variabel dummy (1 = BUMN)
	Ownership diversity	Variasi komposisi kepemilikan

### 3. HASIL DAN PEMBAHASAN

#### a. Gambaran Umum Data

Hasil statistik deskriptif menunjukkan bahwa rata-rata ESG Score berada di kisaran 64 dengan rentang nilai sekitar 31 hingga 88. Secara umum, pilar sosial dan tata kelola memiliki skor yang lebih tinggi dibandingkan pilar lingkungan. Pola ini menggambarkan bahwa institusi perbankan cenderung lebih mapan dalam aspek hubungan sosial dan struktur tata kelola, sementara upaya transformasi lingkungan—seperti pengendalian emisi, efisiensi energi, dan peningkatan pembiayaan hijau—masih terus berkembang.

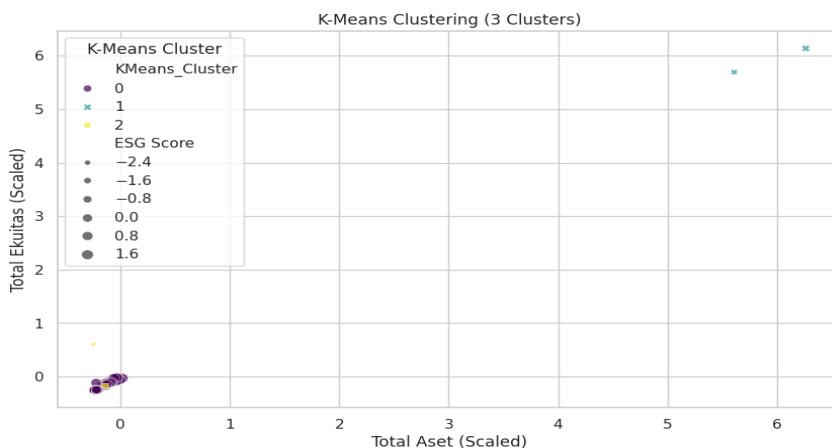
Sejumlah indikator intensitas, termasuk Emisi\_per\_Aset dan Listrik\_per\_Pendapatan, memperlihatkan variasi yang cukup lebar antar entitas. Hal ini mengindikasikan adanya perbedaan strategi operasional serta tingkat efisiensi energi yang diterapkan masing-masing institusi. Variasi yang sama juga terlihat pada rasio green financing, yang mencerminkan belum meratanya penyaluran pembiayaan hijau di sektor perbankan.

#### b. Klasterisasi Menggunakan K-Means dan Agglomerative

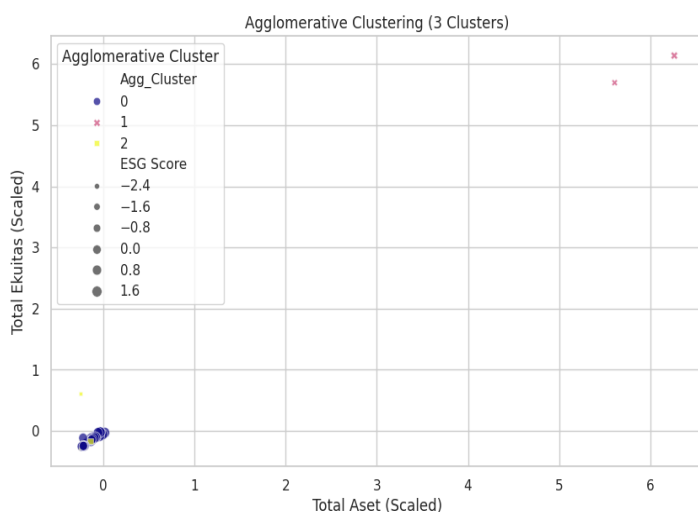
Pengelompokan pertama dilakukan menggunakan fitur total aset, total ekuitas, dan ESG Score yang telah distandarisasi. Dengan jumlah kluster ditetapkan sebanyak tiga, K-Means menghasilkan tiga kelompok utama. Kelompok pertama merupakan kluster dengan ESG tinggi dan ukuran aset besar. Kluster ini dicirikan oleh skor ESG terstandar yang positif dan mayoritas anggotanya berasal dari institusi besar yang cenderung menjadi acuan di sektor perbankan. Kelompok kedua menampilkan profil ESG terendah, dengan skor terstandar sekitar -1,42. Kluster ini tampak cukup terpisah dari kelompok lain dalam visualisasi sehingga mudah dikenali sebagai kelompok dengan performa keberlanjutan terendah. Sementara itu, kluster ketiga memiliki skor ESG yang berada pada rentang sedang hingga rendah, biasanya diisi oleh institusi dengan ukuran aset menengah atau kecil.

Visualisasi hasil K-Means dan Agglomerative Clustering pada sumbu total aset dan total ekuitas (keduanya dalam bentuk terstandarisasi) menunjukkan pola yang konsisten. Kelompok dengan skor ESG tinggi cenderung berada pada area dengan aset dan ekuitas lebih besar. Sebaliknya, institusi dengan ESG rendah lebih tersebar pada area aset yang lebih kecil, meskipun tidak berarti semua institusi kecil menunjukkan performa keberlanjutan rendah. Hasil Agglomerative Clustering memperkuat temuan tersebut karena struktur klasternya serupa dengan K-Means. Penerapan DBSCAN menghasilkan jumlah kluster yang

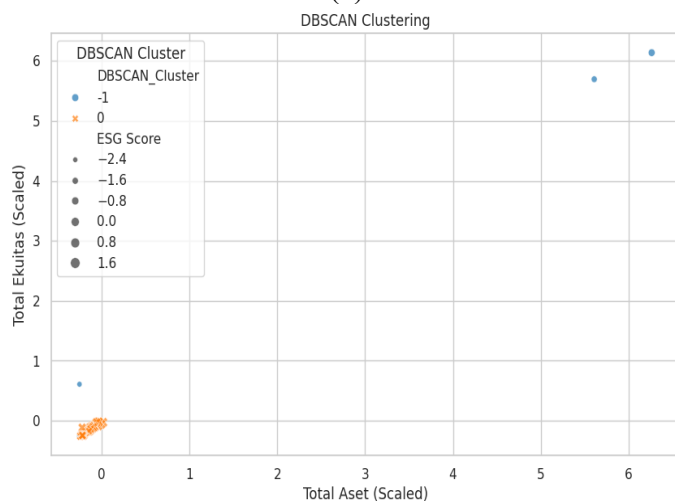
lebih sedikit dan sebagian observasi teridentifikasi sebagai *noise*. Kondisi ini wajar mengingat jumlah observasi yang terbatas dan parameter dasar DBSCAN yang cenderung konservatif. Pada penelitian ini, DBSCAN lebih berfungsi sebagai alat pendeteksi outlier daripada metode pengelompokan utama



(a)



(b)



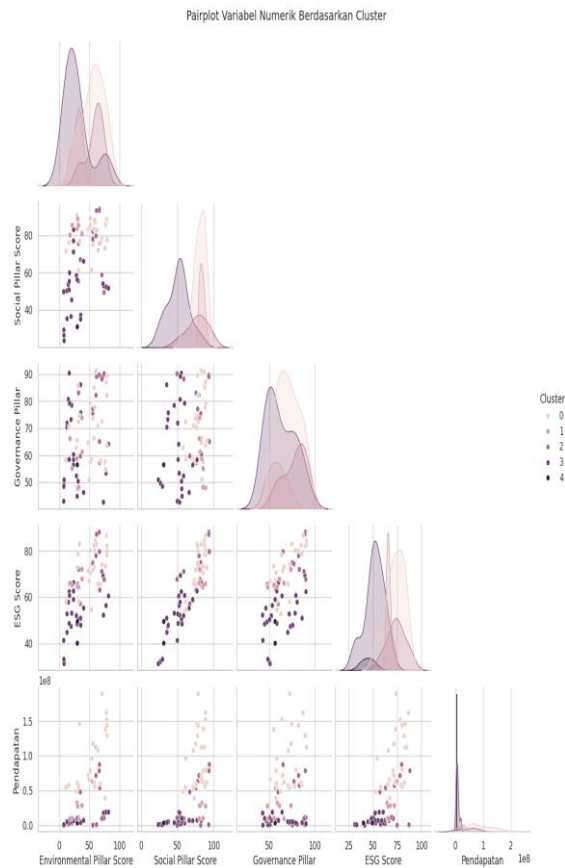
(c)

Gambar 1 : (a) K-means clustering k = 3; (b) Agglomerative Clustering 3; (c) DBSCAN Clustering.

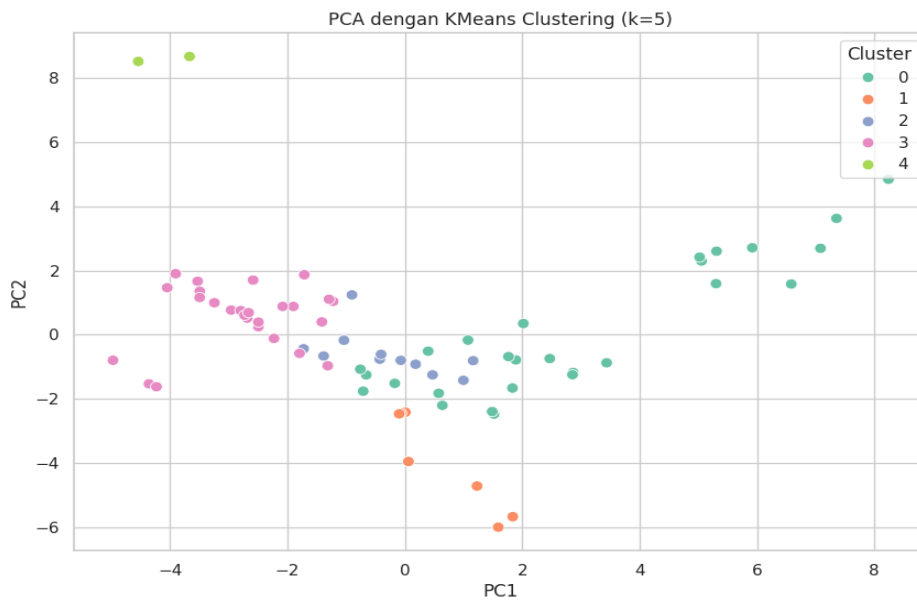
### c. Analisis PCA dan Klasterisasi Multivariat

Pada tahap berikutnya, seluruh variabel numerik yang digunakan dalam analisis multivariat direduksi dimensinya menggunakan PCA. Dua komponen utama (PC1 dan PC2) mampu menjelaskan sekitar 31 persen keragaman data. Visualisasi PCA yang dikombinasikan dengan K-Means ( $k = 5$ ) memperlihatkan pembentukan kelompok yang cukup jelas di ruang dua dimensi tersebut.

Salah satu klaster berada pada area PC1 dan PC2 yang tinggi, yang mencerminkan kombinasi ukuran keuangan besar, intensitas emisi yang relatif terkendali, dan skor ESG yang baik. Sementara itu, klaster lain terkonsentrasi pada area dengan performa ESG yang lebih rendah dan variasi efisiensi sumber daya yang kurang optimal. *Pairplot* dari seluruh variabel numerik berdasarkan label klaster menunjukkan bahwa pilar sosial dan tata kelola relatif tinggi di sebagian besar klaster, sedangkan pilar lingkungan dan indikator intensitas emisi menunjukkan variasi yang lebih besar [11]. Terdapat pula klaster kecil yang dapat dianggap sebagai outlier positif. Kelompok ini memiliki skor ESG tinggi meskipun ukuran institusinya tidak terlalu besar, sehingga dapat diinterpretasikan sebagai entitas dengan komitmen keberlanjutan kuat namun tidak didukung oleh ukuran aset yang besar. Mereka dapat berperan sebagai *niche players* dalam ekosistem keberlanjutan.



(a)



Gambar 2 : (a) Pairplot variable numerik berdasarkan cluster; (b) Pairplot variable numerik berdasarkan cluster; (c) PCA dengan KMeans Clustering (k=5)

**d. Implikasi**

Hasil analisis secara keseluruhan menunjukkan bahwa hubungan antara ukuran institusi dan performa ESG tidak sepenuhnya linear. Walaupun institusi besar cenderung memiliki sumber daya lebih besar untuk mengembangkan praktik keberlanjutan, terdapat pula institusi yang berukuran lebih kecil namun menunjukkan performa ESG yang baik. Sebaliknya, beberapa institusi berada pada posisi tertinggal dari sisi

keberlanjutan, yang berpotensi menghadapi tekanan lebih tinggi dari investor maupun regulator di masa mendatang.

Temuan ini menunjukkan bahwa klusterisasi dapat memberikan gambaran posisi relatif antar institusi dalam hal keberlanjutan. Informasi tersebut dapat dimanfaatkan untuk kebutuhan benchmarking, penyusunan kebijakan insentif atau regulasi berbasis profil keberlanjutan, serta strategi komunikasi kepada investor yang menonjolkan kinerja ESG.

#### 4. KESIMPULAN

Penelitian ini menganalisis faktor-faktor keberlanjutan pada sektor perbankan di Indonesia melalui pendekatan *unsupervised learning*, dengan memanfaatkan K-Means, Agglomerative Clustering, DBSCAN, dan PCA. Data yang digunakan mencakup skor ESG, pilar lingkungan–sosial–tata kelola, indikator emisi dan konsumsi sumber daya, serta sejumlah indikator keuangan yang kemudian diperkaya melalui proses *feature engineering*.

Hasil analisis memperlihatkan bahwa institusi perbankan dapat dikelompokkan ke dalam beberapa klaster dengan karakteristik keberlanjutan dan kondisi keuangan yang berbeda. Terdapat kelompok bank berukuran besar dengan skor ESG tinggi yang dapat dilihat sebagai pemimpin dalam praktik keberlanjutan. Di sisi lain, terdapat pula kelompok dengan skor ESG lebih rendah yang menunjukkan ruang perbaikan, terutama jika ingin meningkatkan daya tarik bagi investor yang sensitif terhadap isu keberlanjutan. Visualisasi melalui PCA membantu menunjukkan pemisahan klaster dan pola umum dalam data, meskipun kontribusinya terhadap keseluruhan variasi data masih terbatas sehingga interpretasinya perlu dilakukan secara hati-hati.

Secara keseluruhan, penelitian ini menunjukkan bahwa metode *unsupervised learning* dapat menjadi alat yang efektif untuk memahami pola keberlanjutan di sektor perbankan dan memberikan gambaran posisi relatif antar institusi. Ke depan, pendekatan ini dapat dikembangkan lebih jauh dengan menggabungkannya ke dalam *supervised learning* untuk memprediksi risiko atau kinerja keuangan berbasis klaster ESG, serta memperluas cakupan data agar hasil yang diperoleh semakin kuat dan representatif.

#### UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada pihak-pihak yang telah memberikan dukungan dan bantuan dalam pelaksanaan penelitian ini, baik dalam bentuk bimbingan, fasilitas, maupun dukungan moral sehingga penelitian ini dapat diselesaikan dengan baik.

#### DAFTAR PUSTAKA

1. M. Farnè and A. Vouldis, “Banks’ risk-taking within a banking union,” *Econ Lett*, vol. 204, Jul. 2021, doi: 10.1016/j.econlet.2021.109909.
2. [2] A. M. H. I. G. Sri Masdiana, “Bank Syariah Sebagai Pilar Sistem Keuangan Islam: Analisis Kedudukan Dan Peran Di Indonesia.,” *Journal of Islamic Banking and Finance Studies*, 2025.
3. [3] Y. K. Tumewang, K. P. Ayunda, M. R. Azzahra, and M. K. Hassan, “The effects of diversity and inclusion on ESG performance: A comparison between Islamic and conventional banks,” *Borsa Istanbul Review*, vol. 24, no. 6, pp. 1101–1110, Nov. 2024, doi: 10.1016/j.bir.2024.10.001.
4. [4] Moh. Farhan, “KESEIMBANGAN RISIKO DAN IMBAL HASIL DALAM STRATEGI INVESTASI BERKELANJUTAN: PENDEKATAN INTEGRATIF TERHADAP FAKTOR LINGKUNGAN, SOSIAL, DAN TATA KELOLA PERUSAHAAN (ESG),” *CURRENCY: Jurnal Ekonomi Dan Perbankan Syariah*, vol. 2, no. 2, pp. 243–264, Apr. 2024, doi: 10.32806/ccy.v2i2.241.
5. [5] W. C. Lin, C. F. Tsai, Y. H. Hu, and J. S. Jhang, “Clustering-based undersampling in class-imbalanced data,” *Inf Sci (N Y)*, vol. 409–410, pp. 17–26, 2017, doi: 10.1016/j.ins.2017.05.008.
6. [6] C. Ergenç and R. Aktaş, “Clustering S&P 500 companies by machine learning for sustainable decision-making,” *Economics and Business Review*, vol. 11, no. 3, pp. 91–117, Sep. 2025, doi: 10.18559/ebr.2025.3.1895
7. [7] D. Tirkaamiana and S. S. A. Basuki, “Enhancing ESG Insights Using Machine Learning: A Case Study of Top Performing Banks in Indonesia,” *Journal of Applied Informatics and Computing*, vol. 9, no. 3, pp. 810–818, Jun. 2025, doi: 10.30871/jaic.v9i3.9480



8. [8] Zulfa Eza Sabila, “PENGARUH GREEN BANKING DISCLOSURE, ISLAMIC SOCIAL REPORTING DISCLOSURE TERHADAP NILAI PERUSAHAAN DIMODERASI OLEH MEKANISME CORPORATE GOVERNANCE (Studi pada Bank Umum Syariah yang Terdaftar di Bursa Efek Indonesia Periode 2018-2022),” 2023.
9. [9] M. Mercadier, A. Tarazi, P. Armand, and J.-P. Lardy, “Monitoring bank risk around the world using unsupervised learning,” *Eur J Oper Res*, vol. 324, no. 2, pp. 590–615, Jul. 2025, doi: 10.1016/j.ejor.2025.01.036.
10. [10] T. Nyitrai and M. Virág, “The effects of handling outliers on the performance of bankruptcy prediction models,” *Socioecon Plann Sci*, no. July, pp. 1–9, 2018, doi: 10.1016/j.seps.2018.08.004.
11. [11] Azkia Husni Mubarak and Rini Widianingsih, “Pengaruh kinerja lingkungan, sosial, dan tata kelola (ESG) terhadap penilaian keberlanjutan perusahaan (CSA),” *Prosiding Seminar Nasional Penelitian dan Pengabdian Kepada Masyarakat*, vol. 3, no. 1, pp. 80–86, Sep. 2025, doi: 10.24002/senapas.v3i1.12602.