



Department of Digital Business

Journal of Artificial Intelligence and Digital Business (RIGGS)

Homepage: <https://journal.ilmudata.co.id/index.php/RIGGS>

Vol. 4 No. 2 (2025) pp: 5075-5080

P-ISSN: 2963-9298, e-ISSN: 2963-914X

Klasterisasi Bank di Indonesia Berdasarkan Kinerja Finansial Menggunakan Algoritma K-Means

Afiyah Nur Hafidzah¹, Hasrat Setiawan Gulo², Sevliana Prisca Yolanda³, Veliana Putri Ayu Saban⁴, Vella

Aprilia Sari⁵, Nanang⁶

¹Prodi Teknik Informatika, Universitas Pamulang, Pamulang, Indonesia

¹dosen02599@unpam.ac.id

Abstrak

Industri perbankan di Indonesia memiliki keragaman yang sangat luas, yang mencakup bank-bank milik negara, swasta domestik, asing, dan juga bank pembangunan daerah. Variasi ini menyebabkan kesulitan dalam mengevaluasi posisi kompetitif serta kesehatan seluruh pasar. Tujuan dari penelitian ini adalah untuk mengelompokkan bank-bank di Indonesia ke dalam kategori-kategori yang seragam berdasarkan kinerja mereka. Dengan memanfaatkan algoritma pembelajaran tanpa pengawasan K-Means, studi ini menganalisis kumpulan data yang terdiri dari 25 bank terkemuka di Indonesia. Variabel yang digunakan untuk pengelompokan meliputi Aset, Kredit yang Diberikan, Laba, dan proporsi Kepemilikan Asing. Sebelum analisis, data melalui proses pra-pemrosesan yang meliputi konversi data kategorikal menjadi numerik dan normalisasi dengan menggunakan StandardScaler. Analisis ini berhasil menemukan 3 kluster yang berbeda secara signifikan: (1) Kluster Bank Besar dengan Aset dan Laba yang dominan; (2) Kluster Bank yang Mayoritas Dimiliki Asing dengan ukuran menengah; dan (3) Kluster Bank Lokal dan Regional yang memiliki fokus pada pasar yang lebih khusus. Sebagai kesimpulan, pengelompokan ini memberikan gambaran yang jelas tentang struktur pasar perbankan dan bisa menjadi dasar bagi regulator, investor, serta pelaku industri dalam mengambil keputusan strategis.

Kata Kunci: K-Means, Klasterisasi, Perbankan Indonesia, Analisis Data, Kinerja Finansial.

1. Latar Belakang

Sektor perbankan merupakan pilar fundamental yang menopang stabilitas dan mendorong pertumbuhan ekonomi suatu negara. Sebagai lembaga intermediasi, bank memegang peran krusial dalam mengalirkan dana dari pihak surplus ke pihak defisit untuk kegiatan produktif dan konsumtif. Fungsi vital ini tidak hanya melancarkan peredaran uang, tetapi juga menjadi motor penggerak bagi sektor riil, investasi, dan penciptaan lapangan kerja. Kesehatan dan ketahanan industri perbankan, oleh karena itu, menjadi cerminan langsung dari kondisi makroekonomi bangsa. Dengan demikian, menjaga stabilitas dan efisiensi sektor ini merupakan prioritas utama bagi pemerintah dan otoritas moneter (Angelidis & Degiannakis, 2008) (Francis X. Diebold, 1996).

Di Indonesia, lanskap industri perbankan ditandai oleh tingkat persaingan yang sangat tinggi serta keragaman pelaku yang signifikan. Struktur pasar ini dihuni oleh pemain dengan skala dan fokus bisnis yang berbeda, mulai dari Bank Usaha Milik Negara (BUMN) yang memiliki aset masif dan jangkauan nasional, hingga bank swasta nasional yang dikenal lincah dalam berinovasi. Selain itu, kehadiran bank asing dengan jaringan global membawa standar internasional, sementara Bank Pembangunan Daerah (BPD) berkonsentrasi pada penguatan perekonomian regional. Keberagaman ini menciptakan sebuah ekosistem yang dinamis namun juga kompleks, di mana setiap jenis bank bersaing untuk memperebutkan pangsa pasar (D. D. Siregar, 2013) (OJK, 2021).

Memahami struktur pasar yang kompleks ini menjadi sebuah keharusan bagi berbagai pemangku kepentingan untuk mengambil keputusan yang strategis. Bagi regulator seperti Otoritas Jasa Keuangan (OJK), pemahaman mendalam tentang segmentasi pasar memungkinkan perumusan kebijakan yang efektif dan pengawasan yang lebih terarah. Bagi investor, analisis struktur pasar membantu mengidentifikasi peluang investasi yang menjanjikan serta menilai risiko secara lebih akurat. Sementara itu, bagi manajemen bank, pengetahuan mengenai posisi strategis

mereka di tengah persaingan adalah kunci untuk merancang model bisnis yang unggul. Pada akhirnya, wawasan ini sangat krusial untuk menjaga daya saing dan keberlanjutan setiap entitas perbankan.

Sayangnya, metode analisis tradisional seringkali cenderung menyederhanakan realitas dengan hanya mengelompokkan bank berdasarkan kategori kepemilikan. Pendekatan ini, meskipun mudah, memiliki kelemahan karena mengabaikan heterogenitas yang ada di dalam kategori yang sama; sebagai contoh, dua bank swasta bisa memiliki model bisnis dan skala operasi yang sangat berbeda. Akibatnya, wawasan yang dihasilkan menjadi kurang tajam dan tidak sepenuhnya merefleksikan dinamika persaingan yang sesungguhnya. Untuk mengatasi keterbatasan ini, diperlukan sebuah pendekatan analisis yang lebih objektif dan berbasis data. Oleh karena itu, dibutuhkan metode yang mampu mengidentifikasi kelompok-kelompok bank secara alami berdasarkan kemiripan karakteristik mereka (Mulyaningsih & Daly, 2012).

Salah satu solusi metodologis yang paling efektif untuk tugas ini adalah klusterisasi, sebuah teknik canggih dalam bidang data mining dan machine learning. Klusterisasi merupakan bagian dari unsupervised learning, yang artinya teknik ini bekerja tanpa memerlukan label yang telah ditentukan sebelumnya untuk menemukan pola tersembunyi. Tujuannya adalah untuk mengelompokkan objek data sedemikian rupa sehingga objek dalam satu kluster sangat mirip satu sama lain, namun sangat berbeda dengan objek di kluster lain. Pendekatan ini memungkinkan identifikasi segmen pasar yang lebih natural dan bermakna. Dengan demikian, klusterisasi menawarkan cara pandang yang lebih mendalam melampaui sekat-sekat klasifikasi konvensional.

Dalam penelitian ini, algoritma K-Means dipilih sebagai metode klusterisasi utama karena sejumlah keunggulannya yang telah terbukti. K-Means adalah salah satu algoritma klusterisasi partisi yang paling populer dan banyak digunakan karena konsepnya yang sederhana serta efisiensinya dalam menangani dataset besar. Algoritma ini bekerja dengan cara mengelompokkan data ke dalam k jumlah kluster yang telah ditentukan, dengan meminimalkan jarak antara setiap titik data dengan pusat klusternya. Popularitas dan efektivitas K-Means telah divalidasi dalam berbagai studi, menjadikannya alat yang andal untuk memetakan struktur pasar. Kesederhanaan komputasinya juga membuat metode ini sangat efisien untuk diterapkan (Han, 2012).

Dengan latar belakang tersebut, penelitian ini bertujuan untuk menerapkan algoritma K-Means guna mengelompokkan 25 bank terkemuka di Indonesia berdasarkan serangkaian variabel kunci. Variabel yang digunakan meliputi Total Aset, Jumlah Kredit yang Disalurkan, Laba Bersih, serta kategori Kepemilikan untuk memberikan gambaran yang komprehensif. Dengan demikian, rumusan masalah utama dari penelitian ini adalah: Bagaimana struktur segmentasi pasar perbankan di Indonesia jika dikelompokkan berdasarkan kinerja finansial dan kepemilikan menggunakan algoritma K-Means? Lebih lanjut, penelitian ini juga akan berupaya mengidentifikasi dan mendeskripsikan karakteristik utama yang membedakan setiap segmen yang terbentuk. Hasil analisis ini diharapkan dapat memberikan pemahaman yang lebih bernuansa tentang lanskap perbankan nasional.

2. Metode Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan menerapkan algoritma K-Means. Tahapan penelitian diuraikan sebagai berikut.

2.1. Sumber Data

Data yang digunakan adalah data sekunder yang terdiri dari 25 bank di Indonesia. Setiap entitas data memiliki atribut: Nama Bank, Aset, Kredit, Laba, dan Kepemilikan.

2.2. Pra – pemrosesan Data (Preprocessing)

1. Konversi Fitur Kepemilikan: Kolom "Kepemilikan" yang bersifat kategorikal diubah menjadi fitur numerik "Persen_Asing". Aturan konversi ditetapkan sebagai berikut: "Pemerintah" dan "Pemda" menjadi 0.0; "Swasta" diasumsikan 50.0; "Asing" tanpa persentase menjadi 100.0; dan "Asing X%" diekstrak nilai numerik X-nya.
2. Normalisasi Data: Fitur numerik yang digunakan (Aset, Kredit, Laba, Persen_Asing) memiliki rentang nilai yang sangat berbeda. Untuk memastikan setiap fitur memiliki bobot yang setara dalam perhitungan jarak K-Means, data dinormalisasi menggunakan StandardScaler dari library Scikit-learn. Proses ini mengubah distribusi setiap fitur sehingga memiliki rata-rata (mean) 0 dan standar deviasi 1 (Pedregosa, 2011).

2.3. Algoritma Klusterisasi

Algoritma K-Means digunakan untuk mempartisi data ke dalam 3 kluster ($k=3$). Jumlah kluster 3 dipilih untuk mengidentifikasi segmen tingkat atas (high-tier), menengah (mid-tier), dan penantang (challenger).

2.4. Evaluasi Model

Evaluasi keberhasilan klusterisasi dilakukan secara kualitatif dengan menganalisis keterpisahan visual antar kluster pada grafik scatter plot dan secara kuantitatif dengan memeriksa koherensi karakteristik anggota dalam setiap kluster yang terbentuk (Rousseeuw, 1987).

2.5. Tools

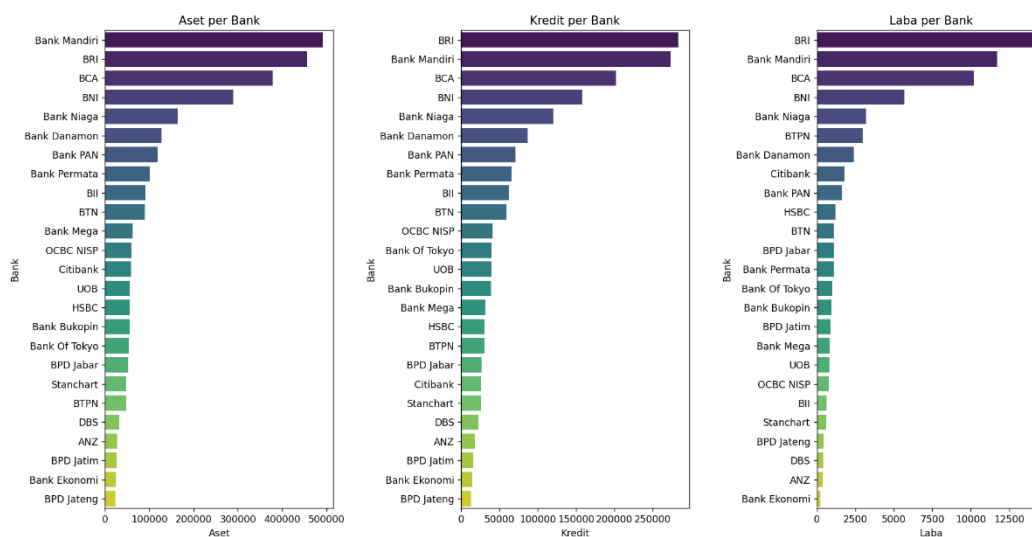
Seluruh proses analisis data dan pemodelan dilakukan menggunakan bahasa pemrograman Python (v3.8) dengan pustaka utama: Pandas untuk manipulasi data (McKinney, 2010), Scikit-learn untuk pemodelan K-Means dan normalisasi (Pedregosa, 2011), serta Matplotlib dan Seaborn untuk visualisasi data (J. VanderPlas, 2016). Berikut adalah pseudocode ringkas dari proses yang dilakukan:

Generated pseudocode

1. Mulai
2. Muat dataset bank
3. Buat kolom baru 'Persen_Asing' dari kolom 'Kepemilikan'
4. Pilih fitur = [Aset, Kredit, Laba, Persen_Asing]
5. Normalisasi fitur menggunakan StandardScaler
6. Inisialisasi model K-Means dengan $k=3$
7. Latih model pada data yang telah dinormalisasi untuk mendapatkan label kluster
8. Gabungkan label kluster ke dataset asli
9. Visualisasikan hasil klusterisasi
10. Analisis profil dan anggota dari setiap kluster
11. Selesai

3. Hasil dan Diskusi

Proses analisis dimulai dengan melakukan analisis data eksploratif untuk memahami distribusi dan karakteristik data mentah. Gambar 1 menampilkan peringkat 25 bank berdasarkan tiga metrik finansial utama: Aset, Kredit, dan Laba.



Gambar 1 Peringkat Bank Berdasarkan Aset, Kredit, dan Laba

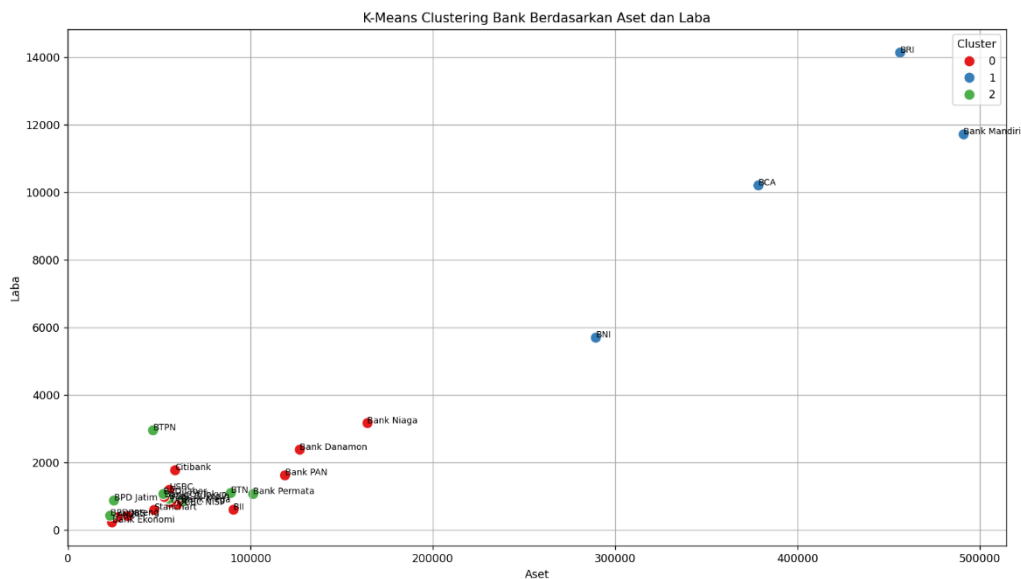
Dari Gambar 1, beberapa temuan awal dapat diidentifikasi:

DOI: <https://doi.org/10.31004/riggs.v4i2.1387>

Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

- Dominasi Pasar: Terdapat ketimpangan yang sangat jelas di pasar perbankan Indonesia. Empat bank teratas—Bank Mandiri, BRI, BCA, dan BNI—secara konsisten mendominasi di ketiga metrik. Aset, kredit, dan laba mereka jauh melampaui bank-bank lainnya.
- Distribusi Tidak Merata: Visualisasi ini mengkonfirmasi adanya "ekor panjang" (long tail), di mana sebagian besar bank memiliki skala operasi yang jauh lebih kecil dibandingkan dengan para pemimpin pasar.
- Justifikasi Klasterisasi: Adanya kelompok yang jelas di puncak dan sebaran bank lainnya di tingkat yang lebih rendah secara visual mengindikasikan bahwa data ini cocok untuk dianalisis menggunakan metode klasterisasi guna memformalkan segmen-segmen pasar tersebut.

Berdasarkan temuan awal ini, algoritma K-Means diterapkan untuk mengelompokkan bank-bank tersebut secara objektif. Setiap bank dikelompokkan ke dalam salah satu dari tiga klaster. Visualisasi hasil klasterisasi berdasarkan Aset dan Laba ditunjukkan pada Gambar 2.



Gambar 2. Visualisasi Hasil K-Means Clustering Berdasarkan Aset dan Laba

Dari Gambar 2, terlihat jelas adanya tiga pengelompokan yang berbeda. Klaster 1 (biru) terisolasi di pojok kanan atas, mengkonfirmasi temuan dari Gambar 1 mengenai adanya kelompok bank dengan Aset dan Laba yang sangat tinggi. Sementara itu, Klaster 0 (merah) dan Klaster 2 (hijau) berada di area Aset dan Laba yang lebih rendah, namun membentuk dua grup yang berbeda. Untuk memahami karakteristik setiap klaster secara mendalam, berikut adalah rincian anggota dan analisisnya. objektif.

3.1. Cluster 0: Pemain Mayoritas Asing (The Foreign-Owned Majors)

- Anggota: Bank Niaga, Bank Danamon, Bank PAN, BII, OCBC NISP, Citibank, UOB, HSBC, dll.
- Analisis: Klaster ini dicirikan oleh kepemilikan asing yang sangat dominan (rata-rata di atas 85%). Bank-bank ini memiliki skala aset dan laba di tingkat menengah, berada di bawah para raksasa tetapi tetap merupakan pemain signifikan di pasar. Mereka umumnya merupakan bagian dari grup keuangan global, yang memberi mereka akses ke modal dan keahlian internasional. Fokus mereka seringkali pada segmen korporat, komersial, dan ritel premium (Trivedi, 2018) (Berger et al., 2009).

	Bank	Aset	Kredit	Laba	Persen Asing
4	Bank Niaga	164239	120195	3173	97.90
5	Bank Danamon	127128	86700	2387	67.40
6	Bank PAN	118991	70818	1626	99.00
8	BII	90741	62574	608	97.50
11	OCBC NISP	59834	41077	753	85.10
12	Citibank	58737	26329	1777	100.00
13	UOB	56430	39460	819	98.99

14	HSBC	55551	30745	1196	100.00
16	Bank of Tokyo	52841	40168	988	100.00
18	Stanchart	47222	25680	602	100.00
20	DBS	32663	22063	422	100.00
21	ANZ	27145	18389	381	100.00
23	Bank Ekonomi	24155	14062	235	98.90

3.2. Cluster 1: Raksasa Finansial (The Giants)

- Anggota: Bank Mandiri, BRI, BCA, BNI.
- Analisis: Klaster ini terdiri dari empat bank terbesar di Indonesia yang secara signifikan mendominasi pasar dari segi Aset, penyaluran Kredit, dan perolehan Laba. Karakteristik utama klaster ini adalah skala operasi yang masif dan kepemilikan yang murni domestik (BUMN dan Swasta Nasional). Mereka adalah pilar utama sistem keuangan dan dapat dianggap sebagai bank *Tier-1* atau bank sistemik di Indonesia (Trivedi, 2018).

	Bank	Aset	Kredit	Laba	Persen Asing
0	Bank Mandiri	491225	273807	11718	0.0
1	BRI	456382	283877	14137	0.0
2	BCA	378652	202269	10207	50.0
3	BNI	289458	158165	5699	0.0

3.3. Cluster 2: Penantang Domestik & Regional (The Domestic & Regional Challengers)

- Anggota: BTN, Bank Permata, Bank Mega, Bank Bukopin, BTPN, BPD Jabar, BPD Jatim, BPD Jateng.
- Analisis: Klaster ini adalah kelompok yang paling heterogen dari segi model bisnis, namun serupa dalam skala operasinya (menengah ke bawah). Anggotanya terdiri dari BUMN dengan fokus spesifik (BTN di sektor perumahan), bank swasta nasional (Bank Mega, Bukopin), dan Bank Pembangunan Daerah (BPD) yang fokus pada ekonomi regionalnya masing-masing. Bank seperti BTPN menunjukkan profitabilitas yang tinggi relatif terhadap asetnya, menandakan keberhasilan dalam ceruk pasar pensiunan dan UMKM (Trivedi, 2018).

	Bank	Aset	Kredit	Laba	Persen Asing
7	Bank Permata	101538	65859	1074	44.5
9	BTN	89253	59338	1105	0.0
10	Bank Mega	62321	31874	832	50.0
15	Bank Bukopin	55430	38853	930	50.0
17	BPD Jabar	52264	26891	1074	0.0
19	BTPN	46638	30199	2959	50.0
22	BPD Jatim	25169	15921	881	0.0
24	BPD Jateng	23108	13108	434	0.0

Pembahasan hasil ini menunjukkan bahwa algoritma K-Means berhasil memetakan lanskap kompetitif perbankan Indonesia secara efektif. Segmentasi yang dihasilkan tidak hanya berdasarkan ukuran, tetapi juga secara implisit menangkap perbedaan model bisnis dan struktur kepemilikan.

4. Kesimpulan

Penelitian ini telah berhasil membuktikan bahwa algoritma K-Means dapat diterapkan secara efektif untuk mengelompokkan 25 bank di Indonesia, yang pada akhirnya memetakan struktur pasar perbankan secara fundamental ke dalam tiga segmen utama yang dapat diinterpretasikan. Segmen-segmen tersebut adalah Raksasa Finansial yang didominasi oleh BUMN dan swasta besar, Pemain Mayoritas Asing dengan skala menengah, serta Penantang Domestik & Regional yang memiliki fokus pasar lebih spesifik. Lebih jauh lagi, terbukti bahwa variabel Aset, Laba, dan Kepemilikan menjadi pembeda yang kuat dalam proses klusterisasi ini, mencerminkan bahwa skala ekonomi dan struktur modal merupakan penentu utama dari posisi kompetitif sebuah bank (Kou et al., 2014). Implikasi dari temuan ini bersifat strategis. Bagi regulator, pemetaan ini dapat membantu dalam merancang pengawasan yang lebih sesuai, sementara bagi investor, ini memberikan kerangka kerja untuk analisis *peer-group* yang lebih akurat. Meskipun demikian, perlu diakui keterbatasan penelitian yang menggunakan data *cross-sectional* dan jumlah variabel yang terbatas. Untuk pengembangan riset di masa mendatang, disarankan agar analisis diperdalam dengan menggunakan data *time-series* guna mengamati dinamika antar klaster seiring berjalannya waktu. Selain itu, penambahan variabel keuangan krusial lainnya seperti NPL dan ROA dapat memberikan pemahaman yang lebih holistik. Dari sisi metodologi, penentuan jumlah klaster sebaiknya divalidasi

secara kuantitatif melalui *Elbow Method* untuk memperoleh jumlah yang optimal. Terakhir, perbandingan hasil dengan algoritma alternatif seperti DBSCAN atau *Hierarchical Clustering* akan memperkuat validitas temuan penelitian.

Referensi

- Angelidis, T., & Degiannakis, S. (2008). Volatility forecasting: Intra-day versus inter-day models. *Journal of International Financial Markets, Institutions and Money*, 18(5), 449–465. <https://doi.org/10.1016/J.INTFIN.2007.07.001>
- Berger, A. N., Hasan, I., & Zhou, M. (2009). Bank ownership and efficiency in China: What will happen in the world's largest nation? *Journal of Banking & Finance*, 33(1), 113–130. <https://doi.org/10.1016/J.JBANKFIN.2007.05.016>
- Bholowalia, P., & Kumar, A. (2014). EBK-Means: A Clustering Technique based on Elbow Method and K-Means in WSN. *International Journal of Computer Applications*, 105(9), 975–8887.
- D. D. Siregar. (2013). *Kekayaan Negara: Siapa punya, siapa kuasa?*.
- Francis X. Diebold. (1996). *The Uncertain Unit Root in Real GNP: Comment on JSTOR*. <https://www.jstor.org/stable/2118292>
- Han, J. , K. M. , & P. J. (2012). *Data Mining: Concepts and Techniques | ScienceDirect*. <https://www.sciencedirect.com/book/9780123814791/data-mining-concepts-and-techniques>
- J. VanderPlas. (2016). *Python Data Science Handbook | Python Data Science Handbook*. <https://jakevdp.github.io/PythonDataScienceHandbook/>
- Kou, G., Peng, Y., & Wang, G. (2014). Evaluation of clustering algorithms for financial risk analysis using MCDM methods. *Information Sciences*, 275, 1–12. <https://doi.org/10.1016/J.INS.2014.02.137>
- McKinney, W. (2010). Data Structures for Statistical Computing in Python. *Scipy*, 56–61. <https://doi.org/10.25080/MAJORA-92BF1922-00A>
- Mulyaningsih, T., & Daly, A. (2012). COMPETITIVE CONDITIONS IN BANKING INDUSTRY: AN EMPIRICAL ANALYSIS OF THE CONSOLIDATION, COMPETITION AND CONCENTRATION IN THE INDONESIA BANKING INDUSTRY BETWEEN 2001 AND 2009. *Buletin Ekonomi Moneter Dan Perbankan*, 14(2), 151–186. <https://doi.org/10.21098/BEMP.V14I2.461>
- OJK. (2021). *id kanal perbankan data-dan-statistik statistik-perbankan-indonesia Default - ojk-go*. <https://www.ojk.go.id/id/kanal/perbankan/data-dan-statistik/statistik-perbankan-indonesia/Default.aspx>
- Pedregosa, F. , V. G. , G. A. , et al. (2011). *Scikit-learn: Machine Learning in Python*. <http://www.jmlr.org/papers/v12/pedregosa11a.html>
- Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20(C), 53–65. [https://doi.org/10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7)
- Tan, P. N. , S. M. , & K. V. (2019). *Introduction to Data Mining*. <https://www-users.cse.umn.edu/~kumar001/dmbook/index.php>
- Trivedi, S. , & S. M. K. (2018). Advances in lightning modeling, computation and measurement. *IEEE Transactions on Electromagnetic Compatibility*, 60(6), 2077–2077. <https://doi.org/10.1109/TEMC.2018.2818559>