

Penerapan Algoritma K-Means Clustering untuk Segmentasi Vendor Berdasarkan Data PPN Masukan pada PT Rifansi Dwi Putra

Sesy Ophelia Tampubolon^{1*}, Tri Andri Hutapea², Ines Monalisa Rumpea³, Septika Aulia Putri⁴,
Rehmuliana Niken Sagala⁵

^{1,2,3,4,5}Program Studi Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Medan,
Jl. William Iskandar Ps. V, Kenangan Baru, Kec. Percut Sei Tuan, Kabupaten Deli Serdang, Sumatera Utara
sesyophelia.4233230016@mhs.unimed.ac.id

Abstract

Vendor transaction data management plays an important role in supporting managerial decision-making within a company. PT Rifansi Dwi Putra maintains input Value Added Tax (VAT) transaction data originating from various vendors with diverse transaction characteristics. This study aims to segment vendors using the K-Means Clustering algorithm based on transaction frequency, invoice value, Value Added Tax (VAT), and Income Tax (PPH). The dataset used consists of input VAT data for the period 21 December 2025 to 20 January 2026, comprising 627 transactions from 183 vendors. The research stages include data preprocessing, vendor-level data aggregation, data normalization, determining the optimal number of clusters using the Silhouette Coefficient, and applying the K-Means algorithm. The results show that the optimal number of clusters is two, with a Silhouette Coefficient value of 0.8710. The first cluster consists of 176 vendors with relatively low to medium transaction values, while the second cluster consists of 7 vendors with significantly higher transaction values and tax contributions. The findings indicate that the K-Means algorithm is effective in identifying strategic vendor groups, thereby supporting better decision-making in vendor management.

Keywords: Data Mining, Clustering, K-Means, Vendor, Input VAT

Abstrak

Pengelolaan data transaksi vendor merupakan bagian penting dalam mendukung pengambilan keputusan perusahaan. PT Rifansi Dwi Putra memiliki data transaksi pajak masukan yang berasal dari berbagai vendor dengan karakteristik transaksi yang berbeda-beda. Penelitian ini bertujuan untuk melakukan segmentasi vendor menggunakan algoritma K-Means Clustering berdasarkan frekuensi transaksi, nilai invoice, Pajak Pertambahan Nilai (PPN), dan Pajak Penghasilan (PPH). Data yang digunakan merupakan data PPN Masukan periode 21 Desember 2025 sampai 20 Januari 2026 yang terdiri atas 627 transaksi dari 183 vendor. Tahapan penelitian meliputi preprocessing data, agregasi data vendor, normalisasi data, penentuan jumlah cluster menggunakan Silhouette Coefficient, serta penerapan algoritma K-Means. Hasil pengujian menunjukkan bahwa jumlah cluster optimal adalah dua cluster dengan nilai Silhouette Coefficient sebesar 0,8710. Cluster pertama terdiri atas 176 vendor dengan rata-rata nilai transaksi yang relatif rendah hingga menengah, sedangkan cluster kedua terdiri atas 7 vendor dengan nilai transaksi dan kontribusi pajak yang jauh lebih tinggi. Hasil penelitian menunjukkan bahwa algoritma K-Means mampu mengidentifikasi kelompok vendor strategis sehingga dapat mendukung pengambilan keputusan dalam pengelolaan vendor perusahaan.

Kata kunci: Data Mining, Clustering, K-Means, Vendor, PPN Masukan

Copyright (c) 2026 Sesy Ophelia Tampubolon, Tri Andri Hutapea, Ines Monalisa Rumpea, Septika Aulia Putri,
Rehmuliana Niken Sagala

✉Corresponding author: Sesy Ophelia Tampubolon

Email Address: sesyophelia.4233230016@mhs.unimed.ac.id (Jl. William Iskandar Ps. V, Kenangan Baru, Kec. Percut Sei Tuan, Kabupaten Deli Serdang, Sumatera Utara)

Received 21 May 2026, Accepted 27 May 2026, Published 02 June 2026

PENDAHULUAN

Perkembangan teknologi informasi dan digitalisasi proses bisnis menyebabkan perusahaan menghasilkan data dalam jumlah yang semakin besar dari berbagai aktivitas operasional. Data yang tersimpan tidak hanya berfungsi sebagai arsip perusahaan, tetapi juga dapat dimanfaatkan sebagai sumber informasi yang mendukung pengambilan keputusan secara lebih efektif dan berbasis data (Han et al., 2012; Tan et al., 2019). Pemanfaatan data yang tepat dapat membantu organisasi meningkatkan

efisiensi operasional, memahami pola aktivitas bisnis, serta mendukung penyusunan strategi yang lebih akurat (Tamriesfatno et al., 2025).

Salah satu pendekatan yang banyak digunakan untuk mengekstraksi informasi dan pola tersembunyi dari kumpulan data adalah data mining. Data mining merupakan proses menemukan pengetahuan, hubungan, dan pola yang bermanfaat dari data berukuran besar melalui berbagai teknik analisis (Han et al., 2012). Salah satu teknik yang banyak digunakan dalam data mining adalah clustering, yaitu metode pengelompokan objek berdasarkan tingkat kemiripan karakteristik sehingga objek dalam kelompok yang sama memiliki tingkat kesamaan yang tinggi dibandingkan dengan objek pada kelompok lainnya (Liantoni, 2022).

Di antara berbagai metode clustering, algoritma K-Means merupakan salah satu metode yang paling populer karena memiliki proses komputasi yang sederhana, efisien, mudah diimplementasikan, dan mampu menangani data berukuran besar dengan baik (Sinaga & Yang, 2020; Chong, 2021). K-Means bekerja dengan mengelompokkan data ke dalam sejumlah cluster berdasarkan kedekatan data terhadap titik pusat cluster (centroid). Untuk mengevaluasi kualitas hasil pengelompokan, salah satu metode yang sering digunakan adalah Silhouette Coefficient yang mampu mengukur tingkat kohesi dalam cluster dan separasi antarcluster sehingga dapat membantu menentukan jumlah cluster yang optimal (Guntara & Lutfi, 2023; Zurfani & Syahputra, 2024).

Penerapan algoritma K-Means telah banyak digunakan untuk membantu proses pengelompokan data dan pengambilan keputusan. Adji dan Dwilestari (2025) menerapkan K-Means pada data transaksi penjualan barang untuk menghasilkan kelompok data yang mendukung pengambilan keputusan bisnis. Kholila et al. (2023) menunjukkan bahwa K-Means mampu mengelompokkan data berdasarkan karakteristik tertentu secara efektif. Penelitian lain yang dilakukan oleh Jiménez-Preciado et al. (2024), Paramarta et al. (2025), Utami et al. (2024), dan Sari et al. (2025) juga menunjukkan bahwa algoritma K-Means memiliki kemampuan yang baik dalam mengidentifikasi pola serta karakteristik data sehingga dapat digunakan sebagai dasar dalam proses analisis dan pengambilan keputusan.

Meskipun demikian, penerapan algoritma K-Means untuk segmentasi vendor berdasarkan data Pajak Pertambahan Nilai (PPN) Masukan masih relatif terbatas. Padahal, data PPN Masukan mengandung informasi penting mengenai frekuensi transaksi, nilai invoice, PPN, dan PPh yang dapat menggambarkan karakteristik serta kontribusi masing-masing vendor terhadap aktivitas perusahaan (Jumingan, 2023). Informasi tersebut berpotensi dimanfaatkan sebagai dasar dalam mengidentifikasi vendor strategis dan mendukung pengambilan keputusan yang lebih efektif. Selain itu, penggabungan variabel frekuensi transaksi, nilai invoice, PPN, dan PPh dalam proses clustering diharapkan mampu menghasilkan segmentasi vendor yang lebih komprehensif dibandingkan penggunaan satu indikator transaksi saja (Adji & Dwilestari, 2025).

PT Rifansi Dwi Putra merupakan perusahaan yang memiliki aktivitas operasional dengan melibatkan banyak vendor dan pemasok. Setiap periode, perusahaan menghasilkan data transaksi yang berkaitan dengan invoice, PPN, dan PPh dari berbagai vendor. Data transaksi tersebut merupakan aset

informasi yang dapat dimanfaatkan untuk mendukung proses evaluasi dan pengelolaan vendor secara lebih efektif. Namun, data yang tersedia belum dimanfaatkan secara optimal untuk mengidentifikasi kelompok vendor berdasarkan karakteristik transaksinya, sehingga perusahaan belum memiliki dasar analitis yang kuat dalam menentukan vendor dengan kontribusi tinggi maupun vendor dengan kontribusi yang relatif rendah terhadap aktivitas operasional perusahaan (Putra, 2024; Sihombing, 2024).

Berdasarkan permasalahan tersebut, penelitian ini bertujuan untuk menerapkan algoritma K-Means Clustering dalam melakukan segmentasi vendor berdasarkan frekuensi transaksi, nilai invoice, nilai PPN, dan nilai PPh. Untuk memperoleh jumlah cluster yang optimal, penelitian ini menggunakan Silhouette Coefficient sebagai metode evaluasi. Hasil penelitian diharapkan dapat membantu perusahaan dalam mengidentifikasi vendor strategis, memahami karakteristik kelompok vendor, serta mendukung pengambilan keputusan yang lebih efektif dalam pengelolaan hubungan vendor.

METODE

Penelitian ini menggunakan metode data mining dengan algoritma K-Means Clustering untuk melakukan segmentasi vendor berdasarkan data PPN Masukan PT Rifansi Dwi Putra. Data yang digunakan merupakan data sekunder periode 21 Desember 2025 sampai dengan 20 Januari 2026 yang terdiri atas 627 transaksi dari 183 vendor. Variabel yang digunakan dalam penelitian meliputi frekuensi transaksi, total nilai invoice, total nilai Pajak Pertambahan Nilai (PPN), dan total nilai Pajak Penghasilan (PPh).

Tahapan penelitian terdiri atas beberapa langkah, yaitu:

1. Pengumpulan data PPN Masukan PT Rifansi Dwi Putra.
2. Data cleaning untuk menghapus data duplikat dan memperbaiki data yang tidak konsisten.
3. Agregasi data berdasarkan vendor sehingga setiap vendor direpresentasikan dalam satu baris data.
4. Normalisasi data menggunakan Standard Scaler agar seluruh variabel memiliki skala yang seimbang.
5. Penentuan jumlah cluster optimal menggunakan Silhouette Coefficient.
6. Proses clustering menggunakan algoritma K-Means.
7. Analisis karakteristik cluster yang terbentuk.

Algoritma K-Means digunakan untuk mengelompokkan vendor berdasarkan kemiripan karakteristik transaksi. Jarak antara data dan centroid dihitung menggunakan Euclidean Distance:

$$d(x, y) = \sqrt{\sum (x_i - y_i)^2}$$

dengan:

- $d(x,y)$ = jarak antara data dan centroid.
- x_i = nilai atribut ke-i pada data.
- y_i = nilai atribut ke-i pada centroid.

Proses K-Means diawali dengan menentukan jumlah cluster (k), kemudian memilih centroid awal secara acak. Selanjutnya, setiap data ditempatkan pada cluster yang memiliki jarak terdekat terhadap centroid. Setelah seluruh data dikelompokkan, centroid diperbarui menggunakan nilai rata-rata anggota cluster. Proses ini dilakukan secara berulang hingga tidak terjadi perubahan keanggotaan cluster.

Untuk menentukan jumlah cluster terbaik digunakan Silhouette Coefficient yang dirumuskan sebagai:

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

dengan:

- $s(i)$ = nilai silhouette untuk data ke- i .
- $a(i)$ = rata-rata jarak data ke seluruh anggota dalam cluster yang sama.
- $b(i)$ = rata-rata jarak data ke cluster terdekat lainnya.

Nilai Silhouette Coefficient berada pada rentang -1 sampai 1 . Semakin mendekati nilai 1 , hasil pengelompokan menunjukkan kualitas yang semakin baik karena memiliki tingkat kohesi yang tinggi serta pemisahan antarcluster yang jelas. Jumlah cluster yang menghasilkan nilai Silhouette Coefficient tertinggi dipilih sebagai jumlah cluster optimal.

HASIL DAN DISKUSI

Data yang digunakan dalam penelitian ini merupakan data PPN Masukan PT Rifansi Dwi Putra periode 21 Desember 2025 sampai dengan 20 Januari 2026. Dataset terdiri atas 627 transaksi yang berasal dari 183 vendor. Setiap transaksi memuat informasi vendor, nomor invoice, nilai invoice, nilai Pajak Pertambahan Nilai (PPN), nilai Pajak Penghasilan (PPh), dan deskripsi transaksi.

Sebelum dilakukan proses clustering, data terlebih dahulu melalui tahap pra-pemrosesan yang meliputi pembersihan data, validasi data, agregasi data berdasarkan vendor, dan normalisasi data. Pada tahap agregasi, seluruh transaksi yang dimiliki oleh vendor yang sama digabungkan sehingga setiap vendor direpresentasikan dalam satu baris data.

Variabel yang digunakan dalam proses clustering meliputi frekuensi transaksi, total nilai invoice, total nilai PPN, dan total nilai PPh. Keempat variabel tersebut dipilih karena mampu merepresentasikan tingkat aktivitas transaksi dan kontribusi finansial masing-masing vendor terhadap perusahaan.

Karena rentang nilai antar variabel berbeda cukup besar, dilakukan normalisasi menggunakan Standard Scaler agar setiap variabel memiliki skala yang seimbang dalam proses pembentukan cluster.

Penentuan Jumlah Cluster Optimal

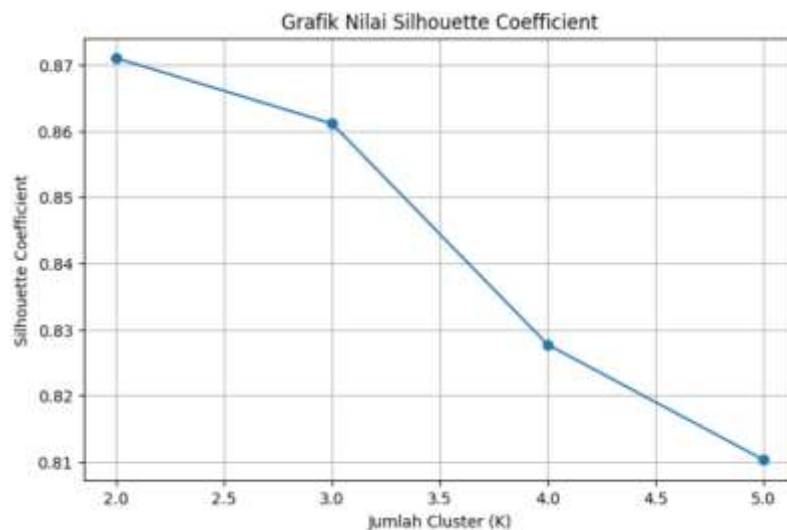
Penentuan jumlah cluster dilakukan menggunakan metode Silhouette Coefficient. Pengujian dilakukan untuk jumlah cluster $K = 2$ sampai $K = 5$. Nilai Silhouette Coefficient yang lebih tinggi menunjukkan kualitas clustering yang lebih baik karena memiliki tingkat kohesi dan separasi yang lebih baik.

Hasil pengujian nilai Silhouette Coefficient ditunjukkan pada Tabel 1.

Tabel 1. Nilai Silhouette Coefficient

Jumlah Cluster (K)	Nilai Silhouette
2	0,8710
3	0,8611
4	0,8277
5	0,8103

Berdasarkan hasil pengujian, nilai Silhouette Coefficient tertinggi diperoleh pada K = 2 yaitu sebesar 0,8710. Oleh karena itu, jumlah cluster optimal yang digunakan dalam penelitian ini adalah dua cluster.



Gambar 1. Grafik Nilai Silhouette Coefficient

Hasil Clustering Menggunakan K-Means

Algoritma K-Means diterapkan menggunakan jumlah cluster optimal yaitu dua cluster. Hasil clustering menunjukkan bahwa dari 183 vendor, sebanyak 176 vendor masuk ke Cluster 0 dan 7 vendor masuk ke Cluster 1.

Persentase masing-masing cluster dihitung menggunakan perbandingan jumlah anggota cluster terhadap total vendor. Cluster 0 memiliki persentase sebesar 96,17%, sedangkan Cluster 1 memiliki persentase sebesar 3,83%.

Tabel 2. Distribusi Vendor Berdasarkan Cluster

Cluster	Jumlah Vendor	Persentase
Cluster 0	176	96,17%
Cluster 1	7	3,83%
Total	183	100%

Karakteristik Cluster

Karakteristik masing-masing cluster dianalisis berdasarkan rata-rata frekuensi transaksi, total invoice, total PPN, dan total PPh.

Tabel 3. Karakteristik Cluster

Variabel	Cluster 0	Cluster 1
Frekuensi Transaksi	3,16	9,86
Total Invoice	Rp57,68 juta	Rp1,53 miliar
Total PPN	Rp4,42 juta	Rp168,82 juta

Total PPh	Rp215 ribu	Rp10,96 juta
-----------	------------	--------------

Hasil analisis menunjukkan bahwa vendor pada Cluster 1 memiliki aktivitas transaksi dan kontribusi pajak yang jauh lebih besar dibandingkan vendor pada Cluster 0. Frekuensi transaksi pada Cluster 1 sekitar 3,12 kali lebih tinggi dibandingkan Cluster 0. Selain itu, rata-rata total invoice 26,53 kali lebih besar, total PPN 38,19 kali lebih besar, dan total PPh 50,98 kali lebih besar.

Tabel 4. Rasio Perbandingan Antar Cluster

Variabel	Rasio (Cluster 1 / Cluster 0)
Frekuensi Transaksi	3,12 kali
Total Invoice	26,53 kali
Total PPN	38,19 kali
Total PPh	50,98 kali

Diskusi

Hasil penelitian menunjukkan bahwa algoritma K-Means mampu mengelompokkan vendor berdasarkan karakteristik transaksi secara efektif dengan kualitas cluster yang sangat baik, ditunjukkan oleh nilai Silhouette Coefficient sebesar 0,8710.

Cluster 0 terdiri atas 176 vendor (96,17%) dengan karakteristik transaksi rendah hingga menengah sehingga dapat dikategorikan sebagai vendor operasional reguler. Sementara itu, Cluster 1 terdiri atas 7 vendor (3,83%) dengan nilai transaksi, PPN, dan PPh yang jauh lebih tinggi sehingga dapat dikategorikan sebagai vendor strategis.

Hasil ini sejalan dengan penelitian Adji dan Dwilestari (2025) serta Chong (2021) yang menyatakan bahwa algoritma K-Means efektif dalam mengelompokkan data numerik berdasarkan pola kemiripan. Secara manajerial, hasil penelitian ini dapat digunakan sebagai dasar dalam evaluasi vendor, pengelolaan kontrak, pengendalian risiko pemasok, serta penyusunan strategi pengadaan yang lebih efektif dan berbasis data.

KESIMPULAN

Penelitian ini berhasil menerapkan algoritma K-Means Clustering untuk melakukan segmentasi vendor berdasarkan data PPN Masukan PT Rifansi Dwi Putra yang meliputi frekuensi transaksi, nilai invoice, nilai Pajak Pertambahan Nilai (PPN), dan nilai Pajak Penghasilan (PPh). Berdasarkan evaluasi menggunakan Silhouette Coefficient, jumlah cluster optimal yang diperoleh adalah dua cluster dengan nilai sebesar 0,8710, yang menunjukkan bahwa hasil pengelompokan memiliki kualitas yang sangat baik.

Hasil segmentasi menunjukkan bahwa Cluster 0 terdiri atas 176 vendor atau 96,17% dari total vendor dengan karakteristik transaksi rendah hingga menengah, sedangkan Cluster 1 terdiri atas 7 vendor atau 3,83% dari total vendor dengan nilai transaksi dan kontribusi pajak yang jauh lebih tinggi. Temuan ini menunjukkan bahwa sebagian besar aktivitas transaksi perusahaan terkonsentrasi pada sejumlah kecil vendor yang memiliki peran strategis dalam operasional perusahaan.

Penelitian ini memberikan kontribusi dalam pemanfaatan teknik data mining untuk mendukung pengelolaan vendor secara lebih objektif dan berbasis data. Hasil segmentasi yang diperoleh dapat dimanfaatkan sebagai dasar dalam evaluasi vendor, pengelolaan kontrak, pengendalian risiko pemasok, serta penyusunan strategi pengadaan yang lebih efektif.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Tuhan Yang Maha Esa atas segala rahmat dan karunia-Nya sehingga penelitian ini dapat diselesaikan dengan baik. Penulis juga menyampaikan terima kasih kepada Bapak Tri Andri Hutapea, S.Si., M.Si. selaku dosen pembimbing yang telah memberikan arahan, masukan, dan bimbingan selama proses penyusunan penelitian ini.

Ucapan terima kasih disampaikan kepada PT Rifansi Dwi Putra yang telah memberikan izin penggunaan data penelitian sehingga penelitian ini dapat dilaksanakan. Penulis juga mengucapkan terima kasih kepada Program Studi Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Medan atas dukungan akademik yang diberikan selama proses penelitian.

Selain itu, penulis menyampaikan apresiasi kepada seluruh pihak yang telah memberikan bantuan, dukungan, dan motivasi sehingga penelitian ini dapat terselesaikan dengan baik.

REFERENSI

- Adji, M. F. M., & Dwilestari, G. (2025). Analisis data transaksi penjualan barang menggunakan teknik K-means clustering. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 9(1), 619–625.
- Aprilliana, A., Adiwarmanto, M., Ramadhon, I. A., & Putra, P. (2023). Analisis produktivitas alat angkut pada kegiatan pengangkutan batubara dari temporary stockpile menuju dump hopper di PT Rifansi Dwi Putra Site Banko Barat PT Bukit Asam, Tanjung Enim, Sumatera Selatan. *Jurnal Ilmiah Teknik dan Sains*, 1(2), 106–112.
- Chong, B. (2021). K-means clustering algorithm: A brief review. *Academic Journal of Computing & Information Science*, 4(5), 29–35. DOI: <https://doi.org/10.25236/AJCIS.2021.040506>
- Guntara, M., & Lutfi, N. (2023). Cacah klaster pada klasterisasi dengan algoritma K-means menggunakan silhouette coefficient dan elbow method. *Jurnal Teknologi Informasi (JuTI)*, 2(1), 43–52. DOI: <https://doi.org/10.26798/juti.v2i1.944>
- Han, J., Kamber, M., & Pei, J. (2012). *Data Mining: Concepts and Techniques* (3rd ed.). Waltham: Morgan Kaufmann.
- Jiménez-Preciado, A. L., Cruz-Aké, S., & Venegas-Martínez, F. (2024). Identification of patterns in CO₂ emissions among 208 countries: K-means clustering combined with PCA and non-linear t-SNE visualization. *Mathematics*, 12(16), 2591. DOI: <https://doi.org/10.3390/math12162591>
- Jumingan, S. E. (2023). *Analisis Laporan Keuangan*. Jakarta: Bumi Aksara.

- Kholila, N. M., Mujiono, M., & Wahyudi, D. (2023). Pemetaan kondisi lingkungan tanam menggunakan K-means clustering. *JSITIK: Jurnal Sistem Informasi dan Teknologi Informasi Komputer*, 1(2), 137–147.
- Liantoni, F. (2022). *Data Mining dan Penerapan Metode*. Jawa Tengah: Eureka Media Aksara.
- Paramarta, V., Rahman, A., Priska, L., Gurning, R. R., & Purwati, W. A. (2025). Analisis komparatif K-means dan model campuran Gaussian dalam pengelompokan emisi CO₂ global. *bit-Tech*, 8(1), 998–1008.
- Putra, M. (2024). *Kerja Praktik pada PT Rifansi Dwi Putra Dumai-Riau. Laporan Kerja Praktik. Dumai: PT Rifansi Dwi Putra.*
- Sari, J. R., Fahira, F., Zahra, L., Fitrianto, A., Erfani, E., & Alifviansyah, K. (2025). Identification of latent dimensions of digital readiness and typology of districts/cities in Indonesia using PCA and K-means clustering. *Journal of Applied Informatics and Computing*, 9(6), 2937–2949.
- Sarmidi, S., Mases, Y., & Kesuma, A. P. (2023). Analisis efisiensi jam kerja excavator CAT 345 GC di area penambangan Pit E Banko Barat PT Rifansi Dwi Putra. *Jurnal Ilmiah Teknik dan Sains*, 1(1), 11–18.
- Schmitz, G. M. (2025). An unsupervised machine learning analysis of environmental sustainability indicators using the K-means clustering algorithm among 2,485 global corporations. *International Journal of Sustainability Policy and Practice*, 21(1).
- Sinaga, K. P., & Yang, M. S. (2020). Unsupervised K-means clustering algorithm. *IEEE Access*, 8, 80716–80727. DOI: <https://doi.org/10.1109/ACCESS.2020.2988796>
- Sihombing, D. I. (2024). *Pengaruh Kepuasan Kerja dan Motivasi Kerja terhadap Kinerja Karyawan pada PT Rifansi Dwi Putra. Disertasi tidak diterbitkan. Pekanbaru: Universitas Lancang Kuning.*
- Tamriesfatno, S., Armadi, L. O. M., Arapa, A., Sulkifli, S., Ramadani, A. R., HS, H., & Nugroho, C. W. (2025). *Data Mining: Teori dan Implementasi*. Yogyakarta: Digital Publishing Library.
- Tan, P. N., Steinbach, M., & Kumar, V. (2019). *Introduction to Data Mining (2nd ed.)*. London: Pearson Education.
- Utami, N., Kumala, S. A., & Handayani, S. D. (2024). Comparative study of earthquake clustering in Indonesia using K-medoids, K-means, DBSCAN, fuzzy C-means and K-AP algorithms. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 8(6), 768–778.
- Zurfani, F. A., & Syahputra, M. R. (2024). Analisis metode clustering K-means pada zonasi daerah terdampak banjir di Kota Medan dengan evaluasi silhouette coefficient. *Algoritma: Jurnal Matematika, Ilmu Pengetahuan Alam, Kebumihan dan Angkasa*, 2(6), 170–181.