



Clustering Data Penjualan Kopi Sidikalang Menggunakan Metode K-Means

Bermiko Kasah Padang^{1*}, Rakhmat Kurniawan¹

¹Universitas Islam Negeri Sumatera Utara

*Corresponding Author's e-mail: miko.padang23@gmail.com

Article History:

Received: August 2, 2025

Revised: August 12, 2025

Accepted: August 13, 2025

Keywords:

Sidikalang Coffee, K-Means clustering, sales analysis, machine learning, Davies-Bouldin Index

Abstract: This study aims to analyze the sales patterns of Sidikalang coffee using the K-Means clustering method to optimize stock management and marketing strategies. Three months of sales data were analyzed based on the attributes of sales volume, price, and season. The results showed that the data could be grouped into four clusters with distinct characteristics: low sales (52 data points), medium sales (50 data points), high sales (21 data points), and high stock with sporadic sales (77 data points). Evaluation using the Davies-Bouldin Index (DBI) yielded a score of 0.52, indicating good clustering quality. These findings provide valuable insights for business actors in developing more targeted marketing strategies and efficient stock management. This research also contributes to the literature on the application of machine learning in analyzing local product sales, particularly Sidikalang coffee.

Copyright © 2025, The Author(s).

This is an open access article under the CC-BY-SA license



How to cite: Padang, B. K., & Kurniawan, R. (2025). Clustering Data Penjualan Kopi Sidikalang Menggunakan Metode K-Means. *SENTRI: Jurnal Riset Ilmiah*, 4(8), 1188–1197. <https://doi.org/10.55681/sentri.v4i8.4400>

PENDAHULUAN

Kopi Sidikalang, yang berasal dari Kabupaten Dairi, Sumatera Utara, telah lama dikenal sebagai salah satu kopi berkualitas tinggi di Indonesia¹. Kopi ini ditanam di daerah dataran tinggi dengan iklim sejuk dan tanah yang subur, menghasilkan cita rasa khas yang sedikit manis dengan aroma coklat. Namun, di tengah potensi besar tersebut, pelaku usaha kopi Sidikalang seringkali menghadapi tantangan dalam mengelola stok dan strategi pemasaran akibat kurangnya analisis data penjualan yang akurat.² Metode konvensional yang digunakan selama ini dinilai tidak lagi mencukupi untuk mengidentifikasi tren penjualan secara efektif. Oleh karena itu, pendekatan berbasis teknologi seperti *machine learning* menjadi solusi yang relevan untuk mengoptimalkan pengelolaan penjualan.³

Clustering, sebagai salah satu teknik dalam *data mining*, memainkan peran penting dalam pengelompokan data berdasarkan kemiripan karakteristik.³ Metode ini memungkinkan identifikasi pola tersembunyi dalam data penjualan, seperti segmentasi produk berdasarkan tingkat penjualan.⁴ Salah satu algoritma clustering yang populer adalah K-Means, yang telah terbukti efektif dalam berbagai penelitian sebelumnya. Misalnya, penelitian oleh Ramadhan dan Ferdiansyah (2022) berhasil mengelompokkan produk roti berdasarkan penjualan menggunakan K-Means, menghasilkan tiga cluster:⁵ sangat laris, laris, dan kurang laris.⁵ Hasil serupa juga ditemukan oleh Safira dkk. (2022) dalam analisis persediaan barang, di mana K-Means membantu mengidentifikasi pola permintaan produk.⁶

Penelitian ini bertujuan untuk menerapkan metode K-Means dalam menganalisis data penjualan kopi Sidikalang. Dengan mengelompokkan produk berdasarkan atribut

seperti jumlah penjualan, harga, dan musim, diharapkan dapat diperoleh wawasan yang berguna untuk pengambilan keputusan strategis. Selain itu, evaluasi clustering menggunakan *Davies Bouldin Index* (DBI) akan dilakukan untuk memastikan kualitas pengelompokan yang dihasilkan.⁸ Kontribusi penelitian ini tidak hanya terletak pada penerapan metode K-Means dalam konteks penjualan kopi lokal, tetapi juga dalam pengayaan literatur tentang analisis data berbasis *machine learning* untuk usaha kecil dan menengah.

LANDASAN TEORI

Kopi Sidikalang

Kopi Sidikalang merupakan salah satu kopi berkualitas tinggi yang berasal dari Kabupaten Dairi, Sumatera Utara. Ditanam di ketinggian 700–1.100 meter di atas permukaan laut, kopi ini memiliki karakteristik rasa sedikit manis dengan aroma coklat yang khas. Kondisi geografis Sidikalang yang sejuk dan tanah yang subur sangat mendukung pertumbuhan kopi robusta dan arabika, menjadikannya salah satu komoditas unggulan daerah.⁷

Penjualan dan Analisis Data

Penjualan merupakan aktivitas utama dalam bisnis yang melibatkan pertukaran barang/jasa dengan alat pembayaran yang sah untuk memperoleh keuntungan.⁸ Analisis data penjualan membantu pelaku usaha mengidentifikasi pola pembelian, mengoptimalkan stok, dan menyusun strategi pemasaran. Namun, metode konvensional seringkali kurang efektif dalam menangani data penjualan yang kompleks, sehingga diperlukan pendekatan berbasis *data mining* seperti *clustering*.⁹

Machine Learning dan Clustering

Machine learning adalah cabang ilmu komputer yang memungkinkan sistem belajar dari data tanpa pemrograman eksplisit. Salah satu tekniknya adalah *clustering*, yang mengelompokkan data berdasarkan kemiripan karakteristik.¹⁰ *Clustering* termasuk dalam *unsupervised learning* karena tidak memerlukan label data.¹¹ Metode ini banyak digunakan dalam segmentasi pasar, analisis penjualan, dan manajemen inventaris.¹²

Metode K-Means Clustering

K-Means adalah algoritma *clustering* berbasis partisi yang membagi data ke dalam *k* kelompok berdasarkan jarak terdekat ke *centroid*.¹³ Langkah-langkahnya meliputi:

1. Penentuan jumlah *cluster* (*k*).
2. Inisialisasi *centroid* acak.
3. Penghitungan jarak data ke *centroid* menggunakan *Euclidean distance*:

$$d(i, n) = \sqrt{\sum (d_i - n_i)^2}$$

4. Pembaruan *centroid* berdasarkan rata-rata data dalam *cluster*.
5. Iterasi hingga konvergensi.¹⁴

K-Means telah berhasil diterapkan dalam berbagai penelitian, seperti pengelompokan produk berdasarkan penjualan dan analisis stok barang.¹⁵

Evaluasi Clustering dengan Davies Bouldin Index (DBI)

DBI mengukur kualitas *clustering* dengan membandingkan jarak intra-*cluster* (kepadatan) dan inter-*cluster* (pemisahan).¹⁶ Nilai DBI yang rendah menunjukkan pengelompokan yang baik. Rumus DBI:

$$DBI = \frac{1}{k} \sum_{i=1}^k \max_{i \neq j} \left(\frac{\text{var}(c_i) + \text{var}(c_j)}{\|c_i - c_j\|} \right)$$

di mana $\text{var}(c_i)$ adalah variasi data dalam *cluster* c_i , c_i dan $\|c_i - c_j\|$ adalah jarak antar-*centroid*.

METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan menerapkan metode K-Means clustering untuk menganalisis data penjualan kopi Sidikalang. Tahapan penelitian dirancang secara sistematis, dimulai dari pengumpulan data hingga evaluasi model, dengan penjelasan mendalam pada setiap langkahnya. Pada tahap pengumpulan data, informasi diperoleh dari toko Tanpak Sidikalang yang berlokasi di Kabupaten Dairi, mencakup riwayat penjualan selama tiga bulan terakhir. Atribut yang dikumpulkan meliputi jenis kopi, harga, jumlah penjualan, dan musim, yang dipilih karena dianggap memiliki pengaruh signifikan terhadap pola penjualan. Data tersebut kemudian diolah dalam format spreadsheet untuk memudahkan analisis lebih lanjut. Periode tiga bulan dipilih guna memastikan cakupan data yang cukup luas tanpa mengorbankan relevansi waktu.

Sebelum analisis dilakukan, data melewati tahap preprocessing untuk menjamin kualitas dan kesiapan analisis. Proses ini mencakup dua langkah utama. Pertama, dilakukan seleksi atribut dengan membatasi atribut pada jumlah penjualan, harga, dan musim, karena ketiganya dinilai paling relevan dalam mengidentifikasi pola penjualan. Atribut yang kurang relevan, seperti nama pelanggan, sengaja diabaikan untuk menghindari noise dalam analisis. Kedua, dilakukan normalisasi data menggunakan metode *Min-Max Scaling* agar skala antar atribut seragam. Normalisasi ini penting karena atribut seperti harga (dalam ribuan rupiah) dan jumlah penjualan (dalam kilogram) memiliki skala yang berbeda, yang dapat memengaruhi hasil clustering. Dengan normalisasi, setiap atribut memberikan kontribusi yang seimbang dalam perhitungan jarak antar data.

1. Analisis dengan K-Means Clustering

- a. Penentuan Jumlah Cluster (k): Menggunakan metode *Elbow* untuk menentukan nilai optimal k^* .
- b. Proses Clustering:
 - Inisialisasi *centroid* acak.
 - Hitung jarak data ke *centroid* dengan *Euclidean distance*:

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

- c. Kelompokkan data ke *cluster* terdekat.
- d. Perbarui *centroid* dengan rata-rata data dalam *cluster*.

e. Ulangi hingga konvergensi (tidak ada perubahan *cluster*).

2. Evaluasi Model

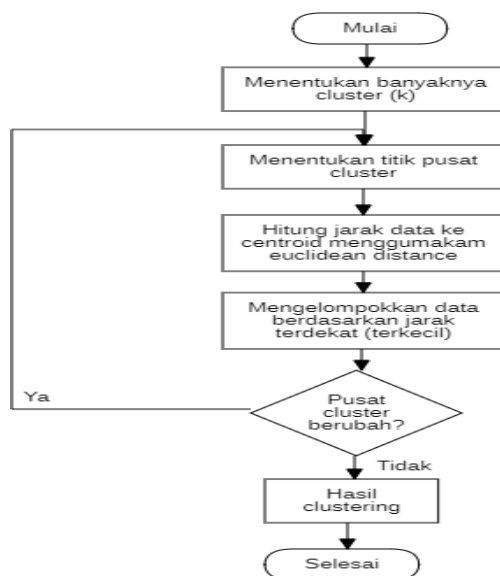
Kualitas clustering dievaluasi menggunakan *Davies Bouldin Index (DBI)*, yang membandingkan jarak intra-cluster (kepadatan) dan inter-cluster (pemisahan). Nilai DBI yang rendah menunjukkan pengelompokan yang baik. Dalam penelitian ini, nilai DBI sebesar 0.52 menunjukkan bahwa cluster yang dihasilkan memiliki pemisahan yang jelas dan kepadatan yang tinggi.

$$DBI = \frac{1}{k} \sum_{i=1}^k \max_{i \neq j} \left(\frac{\text{var}(c_i) + \text{var}(c_j)}{\|c_i - c_j\|} \right)$$

3. Implementasi

Algoritma K-Means diimplementasikan menggunakan bahasa pemrograman *Python* dengan bantuan library seperti *pandas* untuk manipulasi data, *scikit-learn* untuk proses clustering, dan *matplotlib* untuk visualisasi. Visualisasi hasil clustering dalam bentuk grafik 2D/3D memudahkan interpretasi pola penjualan berdasarkan atribut utama, seperti jumlah penjualan dan total pendapatan.

Dengan pendekatan ini, penelitian tidak hanya menghasilkan pengelompokan data yang akurat tetapi juga memberikan landasan yang kuat untuk pengambilan keputusan strategis dalam manajemen penjualan kopi Sidikalang.



Gambar 1. Flowchart Algoritma K-Means

HASIL DAN PEMBAHASAN

Analisis Data

analisis data dilakukan menggunakan algoritma K-Means. Algoritma ini bekerja dengan cara mengelompokkan data ke dalam beberapa klaster berdasarkan jarak ke titik pusat klaster (centroid). Proses ini diulangi hingga posisi klaster menjadi stabil. Adapun variabel yang digunakan dalam penelitian ini meliputi jumlah penjualan, pendapatan, dan

stok barang. Variabel jumlah penjualan adalah variabel yang berisi total jumlah penjualan kopi yang terdiri atas beberapa varian bungkus. lalu variabel data pendapatan adalah variabel total jumlah pemasukan yang di dapat dari hasil penualan kopi. Dan stok adalah sisa stok kopi yang masi tersisa di toko.

Table 1 Data Penelitian

No	Jumlah Penjualan (kg)	Total Pendapatan (Rp)	Stok (kg)	Wilayah
1	4.3	604000	58	Subulussalam
2	7.7	1147000	58	Subulussalam
3	11.7	1686000	58	Subulussalam
4	6.5	931000	58	Subulussalam
5	4.1	588000	58	Tigalingga
..	...	..	...	..
..	...	..	...	..
196	7.1	973000	44	Subulussalam
197	2.4	328000	44	Aceh Singkil
198	1.4	206000	44	Sidikalang
199	0.3	54000	44	Tigalingga
200	0	0	44	Aceh Singkil

Penelitian ini berhasil mengelompokkan data penjualan kopi Sidikalang ke dalam 4 cluster menggunakan algoritma K-Means, dengan variabel utama meliputi jumlah penjualan (kg), total pendapatan (Rp), dan stok (kg). Proses clustering mencapai konvergensi dalam 6 iterasi, menghasilkan distribusi data sebagai berikut:

- 1) Cluster 1 (Penjualan Rendah): 52 data dengan rata-rata penjualan 2.8 kg dan pendapatan di bawah Rp500.000.
- 2) Cluster 2 (Penjualan Sedang): 50 data dengan rata-rata penjualan 6.5 kg dan pendapatan Rp500.000–Rp1.000.000.
- 3) Cluster 3 (Penjualan Tinggi): 21 data dengan rata-rata penjualan 11.2 kg dan pendapatan di atas Rp1.500.000.
- 4) Cluster 4 (Stok Tinggi, Penjualan Sporadis): 77 data dengan rata-rata penjualan 1.5 kg dan pendapatan Rp210.000.

Tabel 2. Distribusi Data per Cluster

Cluster	Jumlah Data	Rata-rata Penjualan (kg)	Rata-rata Pendapatan (Rp)
1	52	2.8	420.000
2	50	6.5	850.000
3	21	11.2	1.680.000
4	77	1.5	210.000

Analisis tabel ini mengungkapkan pola penjualan yang berbeda untuk setiap kelompok produk. Cluster 1 dan 4 memerlukan perhatian khusus karena menunjukkan kinerja di bawah rata-rata, baik dalam hal volume penjualan maupun pendapatan. Di sisi lain, Cluster 2 dan 3 merepresentasikan peluang bisnis yang dapat dikembangkan lebih lanjut. Perbedaan karakteristik antar-cluster ini memberikan dasar yang kuat bagi pelaku usaha untuk menyusun strategi manajemen yang lebih terarah, seperti penyesuaian harga, pengoptimalan stok, atau inisiatif pemasaran yang spesifik untuk setiap kelompok produk. Hasil pengelompokan ini tidak hanya membantu dalam identifikasi pola penjualan tetapi juga berfungsi sebagai alat diagnostik untuk meningkatkan efisiensi operasional bisnis kopi Sidikalang.

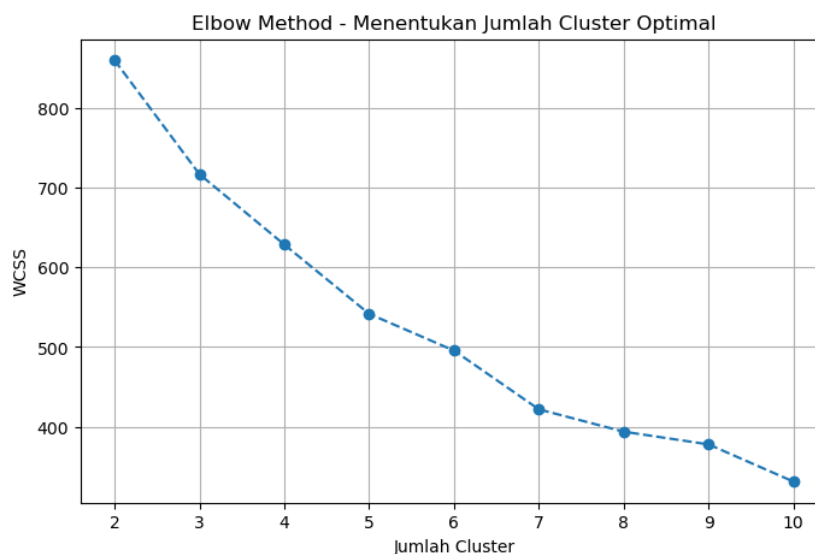
Evaluasi Model dengan Davies Bouldin Index (DBI)

Nilai DBI 0.52 menunjukkan kualitas clustering yang baik (semakin mendekati 0, semakin optimal). Hasil ini didukung oleh visualisasi sebaran cluster (Gambar 2) yang memisahkan kelompok data secara jelas berdasarkan atribut pendapatan dan penjualan.

Gambar dibawah merupakan visualisasi metode Elbow yang digunakan untuk menentukan jumlah cluster optimal dalam analisis K-Means. Grafik ini memplot hubungan antara jumlah cluster (sumbu X) dan nilai *inertia* (sumbu Y), yang menggambarkan sejauh mana data tersebar dalam setiap cluster. Titik optimal ditentukan ketika penurunan *inertia* mulai melandai, membentuk pola seperti siku. Dalam penelitian ini, titik tersebut teridentifikasi pada 4 cluster, yang kemudian menjadi dasar untuk proses pengelompokan data penjualan kopi Sidikalang.

Pada grafik tersebut, dapat diamati pola penurunan nilai inersia yang signifikan saat jumlah cluster meningkat dari 1 menjadi 4. Pola ini menunjukkan bahwa penambahan jumlah cluster pada rentang tersebut memberikan peningkatan kualitas pengelompokan yang berarti. Namun, setelah melewati 4 cluster, kurva mulai menunjukkan penurunan yang semakin landai, membentuk pola yang menyerupai siku (elbow). Titik perubahan pola inilah yang dalam metodologi analisis cluster dikenal sebagai 'titik elbow' dan dianggap sebagai indikator jumlah cluster optimal.

Gambar 2. Metode Elbow untuk Menentukan Jumlah Cluster Optimal



Sumbu X: Jumlah Cluster, Sumbu Y: Inertia

Seperti yang ditunjukkan pada gambar dibawah ini yang menampilkan hasil clustering dalam bentuk visualisasi 2D atau 3D, di mana setiap cluster ditandai dengan warna berbeda. Gambar ini memperlihatkan sebaran data berdasarkan atribut utama, yaitu jumlah penjualan (kg) dan total pendapatan (Rp). Keempat cluster terpisah secara jelas, mencerminkan karakteristik unik masing-masing kelompok, seperti penjualan rendah, sedang, tinggi, serta kelompok dengan stok tinggi namun penjualan sporadis. Visualisasi ini tidak hanya memvalidasi keefektifan algoritma K-Means tetapi juga memudahkan interpretasi pola penjualan bagi pelaku usaha.

Gambar 3. Visualisasi Hasil Clustering

```

optimal_k = 4
kmeans = KMeans(n_clusters=optimal_k, random_state=42, n_init='auto')
data['Cluster'] = kmeans.fit_predict(fitur_scaled)

print(f"\nJumlah Iterasi: {kmeans.n_iter_}")
print("Centroid Terakhir (dalam skala data setelah normalisasi):")
print(kmeans.cluster_centers_)

plt.figure(figsize=(10,6))
sns.scatterplot(data=data,
                x='Jumlah Penjualan (kg)',
                y='Total Pendapatan (Rp)',
                hue='Cluster',
                style='Wilayah_Encoded',
                palette='Set1')
plt.title(f"Cluster Penjualan vs Pendapatan (Data Asli, k={optimal_k})")
plt.xlabel("Jumlah Penjualan (kg)")
plt.ylabel("Total Pendapatan (Rp)")
plt.grid(True)
plt.show()

plt.figure(figsize=(10,6))
sns.scatterplot(data=data, x='Jumlah Penjualan (kg)', y='Stok (kg)', hue='Cluster', palette='Set1')
plt.title("Clustering Penjualan Kopi vs Stok")
plt.xlabel("Jumlah Penjualan (kg)")
plt.ylabel("Stok (kg)")
plt.grid(True)
plt.show()

print("\nRingkasan Hasil Clustering:")
display(data[['Wilayah', 'Jumlah Penjualan (kg)', 'Total Pendapatan (Rp)', 'Cluster']])

print("\nJumlah Data per Cluster:")
print(data['Cluster'].value_counts().sort_index())

```

Warna berbeda mewakili tiap cluster

Hasil analisis K-Means clustering pada data penjualan kopi Sidikalang menunjukkan bahwa pengelompokan berdasarkan jumlah penjualan, pendapatan, dan stok menghasilkan empat cluster dengan karakteristik yang berbeda. Cluster 1 (52 data) mencerminkan produk dengan penjualan rendah (<2.8 kg) dan pendapatan di bawah Rp500.000, yang mungkin disebabkan oleh kurangnya promosi atau harga yang kurang kompetitif. Cluster 2 (50 data) menunjukkan performa sedang dengan rata-rata penjualan 6.5 kg, berpotensi untuk ditingkatkan melalui strategi pemasaran yang lebih agresif. Sementara itu, Cluster 3 (21 data) merupakan produk unggulan dengan penjualan tinggi (>11.2 kg) dan kontribusi pendapatan terbesar (>Rp1.500.000). Cluster 4 (77 data) justru menunjukkan ketidakseimbangan antara stok tinggi dan penjualan sporadis (1.5 kg), mengindikasikan perlunya evaluasi manajemen persediaan. Nilai DBI sebesar 0.52

mengkonfirmasi bahwa pengelompokan ini cukup valid, dengan cluster yang relatif padat dan terpisah dengan jelas.

Penelitian ini sejalan dengan temuan Aprilianur et.al (2022) serta ARDI Mardiana, & Budima. (2022) yang juga berhasil mengidentifikasi segmentasi produk menggunakan K-Means. Namun, terdapat beberapa perbedaan mendasar¹⁸. Pertama, penelitian sebelumnya berfokus pada produk roti dan barang persediaan, sementara penelitian ini mengaplikasikan metode serupa pada komoditas kopi lokal, yang memiliki karakteristik pasar dan pola permintaan yang berbeda. Kedua, penelitian ini mengintegrasikan variabel musim dalam analisis, yang belum banyak dieksplorasi dalam studi-studi terdahulu. Ketiga, penggunaan DBI sebagai evaluator kualitas cluster memberikan pendekatan yang lebih kuantitatif dibandingkan penelitian sebelumnya yang hanya mengandalkan visualisasi atau analisis deskriptif.¹⁹ Temuan ini tidak hanya memperkuat validitas metode K-Means dalam analisis penjualan, tetapi juga memberikan kontribusi baru dalam konteks pengelolaan usaha kopi skala kecil dan menengah.

Pelaku usaha disarankan untuk menerapkan strategi berbeda pada setiap cluster. Produk dalam Cluster 1 dan 4 memerlukan intervensi segera, seperti penyesuaian harga atau promosi khusus, sementara Cluster 2 dan 3 dapat dioptimalkan lebih lanjut untuk meningkatkan pendapatan. Penelitian ini juga membuka peluang untuk studi lanjutan dengan menambahkan variabel lain seperti preferensi konsumen atau analisis kompetitor, serta memperpanjang periode pengamatan untuk mendapatkan gambaran yang lebih komprehensif tentang tren penjualan kopi Sidikalang.

KESIMPULAN

Berdasarkan hasil penelitian, dapat disimpulkan bahwa penerapan metode K-Means clustering berhasil mengidentifikasi pola penjualan kopi Sidikalang ke dalam empat cluster yang jelas. Cluster 1 mencakup 52 data dengan penjualan rendah (rata-rata 2.8 kg) dan pendapatan di bawah Rp500.000, Cluster 2 terdiri dari 50 data dengan penjualan sedang (6.5 kg) dan pendapatan Rp500.000-Rp1.000.000, Cluster 3 meliputi 21 data dengan penjualan tinggi (11.2 kg) dan pendapatan di atas Rp1.500.000, serta Cluster 4 yang berisi 77 data dengan stok tinggi namun penjualan sporadis. Nilai Davies Bouldin Index (DBI) sebesar 0.52 menunjukkan kualitas pengelompokan yang cukup baik. Temuan ini memberikan wawasan berharga bagi pelaku usaha, khususnya dalam mengoptimalkan manajemen stok dan menyusun strategi pemasaran yang lebih terarah.

DAFTAR REFERENSI

1. Aldi Cahya Muhammad, B.Sc.Engg., M.Sc.Engg., Ir. Anak Agung Gde Bagus Ariana, S.T., M.T., Indo Intan, S.T., M.T., Sumanto, M.Kom., Pastima
2. Simanjuntak, S.Kom., M.SI., Satria, S.Kom., M.Kom., M. Lutfi MA, S.T., M.Kom., Nono Heryana, M.Kom., Harizki Arie P, M. (2023). *Dasar-dasar Pembelajaran Mesin (Foundations of Machine Learning)*. Sada Kurnia Pustaka. https://books.google.co.id/books?id=8COzEAAAQBAJ&newbks=1&newbks_redir=0&printsec=frontcover&pg=PA9&dq=Materi+tentang+machine+learning&hl=id&source=gb_mobile_entity&redir_esc=y#v=onepage&q=Materi+tentang+machine+learning&f=false
3. Aprilianur, G., & Hadisaputro, E. L. (2022). Penerapan Data Mining Menggunakan Metode K-Means Clustering Untuk Analisa Penjualan Toko Myam Hijab Penajam. *Jupiter*, 14(1), 161–170.

4. ARDI MARDIANA, & Budiman, B. (2022). Implementasi Metode K-Means Untuk Clustering Produk Yang Kurang Diminati Berdasarkan Data Penjualan. *Seminar Teknologi Majalengka (Stima)*, 6, 198–210. <https://doi.org/10.31949/stima.v6i0.697>
5. Dinata, F. A., Nazir, A., Fikry, M., & Afrianty, I. (2024). Implementasi Data Mining Untuk Prediksi Stok Penjualan Keramik dengan Metode K-Means. *Journal of Computer System and Informatic (JoSYC)*, 5(3), 701–708. <https://doi.org/10.47065/josyc.v5i3.5200>
6. Fatawa Imam Al Muftin, & Fendi Hidayat. (2024). Sistem Informasi Penjualan. *Zona Komputer: Program Studi Sistem Informasi Universitas Batam*, 13(3), 232–237. <https://doi.org/10.37776/zkomp.v13i3.1461>
7. Fuadi, S., & Candra, O. (2020). Prototype Alat Penyiram Tanaman Otomatis dengan Sensor Kelembaban dan Suhu Berbasis Arduino. *JTEIN: Jurnal Teknik Elektro Indonesia*, 1(1), 21–25. <https://doi.org/10.24036/jtein.v1i1.12>
8. Juita Selfia Manullang, & Gea Ibelala. (2023). Analisis Kopi Robusta Sidikalang Sebagai Penghasil Kopi Terbaik Yang Mengalami Kenaikan Harga Jual. *Lokawati : Jurnal Penelitian Manajemen Dan Inovasi Riset*, 2(1), 300–308. <https://doi.org/10.61132/lokawati.v2i1.537>
9. Mustika, Yunita Ardilla, Abraham Manuhutu, Nazaruddin Ahmad, Imanuddin Hasbi, Guntoro, Melda Agnes Manuhutu, Mohamad Ridwan, Hozairi, Anindya Khrisna Wardhani, Syariful Alim, Ikhsan Romli, Yoga Religia, D Tri Octafian, Unggul Utan Sufandi, I. E. (2021). *Data Mining dan Aplikasinya*. Penerbit Widina. https://books.google.co.id/books?id=53FXEAAQBAJ&newbks=1&newbks_redir=0&printsec=frontcover&pg=PR6&dq=Materi+tentang+hierarchical+clustering&hl=id&source=gb_mobile_entity&redir_esc=y#v=onepage&q=Materi tentang hierarchical clustering&f=false
10. Normah, Rifai, B., Vambudi, S., & Maulana, R. (2022). Analisa Sentimen Perkembangan Vtuber Dengan Metode Support Vector Machine Berbasis SMOTE. *Jurnal Teknik Komputer AMIK BSI*, 8(2), 174–180. <https://doi.org/10.31294/jtk.v4i2>
11. Nurdiansyah, F., & Akbar, I. (2021). Implementasi Algoritma K-Means untuk Menentukan Persediaan Barang pada Poultry Shop. *Jurnal Teknologi Dan Manajemen Informatika*, 7(2), 86–94. <https://doi.org/10.26905/jtmi.v7i2.6377>
12. Paembonan, S., & Abduh, H. (2021). Penerapan Metode Silhouette Coefficient untuk Evaluasi Clustering Obat. *PENA TEKNIK: Jurnal Ilmiah Ilmu-Ilmu Teknik*, 6(2), 48. https://doi.org/10.51557/pt_jiit.v6i2.659
13. Prastiwi, H., Jeny Pricilia, & Errissya Rasywir. (2022). Implementasi Data Mining Untuk Menentukan Persediaan Stok Barang Di Mini Market Menggunakan Metode K-Means Clustering. *Jurnal Informatika Dan Rekayasa Komputer(JAKAKOM)*, 2(1), 141–148. <https://doi.org/10.33998/jakakom.2022.2.1.34>
14. Rahmah, S. A., & Antares, J. (2022). Klasterisasi Seleksi Mahasiswa Calon Penerima Beasiswa Yayasan Menggunakan K-Means Clustering. *INFORMATIKA*, 13(2), 25. <https://doi.org/10.36723/juri.v13i2.282>
15. Safira, I., Salkiawati, R., & Priatna, W. (2022). Penerapan Algoritma K-Means untuk Mengetahui Pola Persediaan Barang pada Toko Raja Bekasi. *Journal of*

- Informatic and Information Security*, 3(1), 99–110.
<https://doi.org/10.31599/jiforty.v3i1.1253>
16. Sarimole, F. M., & Hakim, L. (2024). Klasifikasi Barang Menggunakan Metode Clustering K-Means Dalam Penentuan Prediksi Stok Barang. *Jurnal Sains Dan Teknologi*, 5(3), 846–854. <https://doi.org/10.55338/saintek.v5i3.2709>
17. Sinaga, A. R., & Pranata, G. D. (2021). SisInfo Clustering Data Penjualan Produk pada Toko Yudha dengan Algoritma K- Means Kata Kunci : Data Mining , Naive Bayes , Prestasi . SisInfo. *Jurnal Sistem Informasi Dan Informatika (SISINFO)*, 3(02), 135–139.
18. Sinaga, A., Sianturi, C. M., & Pasaribu, V. A. R. (2024). Analisis Kegiatan Pemberdayaan Petani Kopi dalam Peningkatan Hasil Produksi Kopi di Kecamatan Sidikalang, Kabupaten Dairi. *Journal of Social Science and Digital Marketing (JSSDM)*, 1, 27–37.
19. Srisulistiowati, D. B., Khaerudin, M., & Rejeki, S. (2021). Sistem Informasi Prediksi Penjualan Alat Tulis Kantor Dengan Metode Fp-Growth (Studi Kasus Toko Koperasi Sekolah Bina Mulia). *Jurnal Sistem Informasi Universitas Suryadarma*, 8(2). <https://doi.org/10.35968/jsi.v8i2.739>