

CLUSTERING DATA PENJUALAN PRODUK KOPI MENGGUNAKAN K-MEANS UNTUK PERUMUSAN STRATEGI HARGA DAN PROMOSI

Deswita Hendiyani¹; Agus Bahtiar²;

Program Studi Komputerisasi Akuntansi¹
Program Studi Sistem Informasi²

STMIK IKMI Cirebon

<https://ikmi.ac.id/page/18/?lang=de>

¹deswitaaddress@gmail.com;

(*) Corresponding Author : deswitaaddress@gmail.com;

Published : 30 Desember 2025

Abstract—This study aims to cluster coffee product sales data based on profit margin and sales volume to formulate more effective pricing and promotion strategies. The method employed is the Knowledge Discovery in Databases (KDD) framework combined with the K-Means Clustering algorithm. The research stages include data selection and cleaning, transformation and normalization, determination of the optimal number of clusters, algorithm implementation, and result interpretation leading to strategic recommendations. The findings indicate that K-Means models with K=2, K=3, and K=6 produce well-separated and compact clusters, as reflected by low Davies-Bouldin Index values. The analysis identifies product segments with distinct characteristics, providing a basis for accurately targeted pricing and promotion strategies. These findings contribute practical insights for coffee business operators to enhance data-driven marketing effectiveness and serve as a reference for future research in sales analysis and data mining in the coffee sector.

Keywords: K-Means; Product Segmentation; Pricing Strategy; Promotion; Data Mining

Abstrak— Penelitian ini bertujuan untuk mengelompokkan data penjualan produk kopi berdasarkan margin laba dan volume penjualan guna merumuskan strategi harga dan promosi yang lebih efektif. Metode yang digunakan adalah Knowledge Discovery in Databases (KDD) dengan algoritma K-Means Clustering. Tahapan penelitian meliputi pemilihan dan pembersihan data, transformasi dan normalisasi, penentuan jumlah cluster optimal, penerapan algoritma, hingga interpretasi hasil dan penyusunan rekomendasi strategi. Hasil penelitian menunjukkan bahwa model K-Means dengan K=2, K=3, dan K=6 memberikan pemisahan dan kekompakan cluster yang baik berdasarkan nilai Davies-Bouldin Index yang rendah. Analisis ini mengidentifikasi segmen produk dengan karakteristik yang berbeda, yang dapat menjadi acuan dalam menentukan strategi harga dan promosi secara tepat sasaran. Temuan ini memberikan kontribusi nyata bagi pelaku usaha kopi dalam meningkatkan efektivitas pemasaran berbasis data, serta menjadi referensi bagi penelitian lanjutan dalam analisis penjualan dan data mining di sektor kopi.

Kata Kunci : K-Means; Segmentasi Produk; Strategi Harga; Promosi; Data Mining

INTRODUCTION

Perkembangan teknologi informasi dan komunikasi telah mendorong perubahan besar dalam cara perusahaan mengelola data penjualan dan menyusun strategi pemasaran. Transformasi digital memungkinkan perusahaan mengumpulkan, menyimpan, dan menganalisis data dalam jumlah besar secara lebih cepat dan efisien. Salah satu tantangan utama dalam era big data adalah bagaimana mengubah data yang melimpah menjadi informasi yang relevan dan dapat digunakan untuk pengambilan keputusan bisnis yang strategis [1], [2]. Di sektor ritel, khususnya industri kopi,

kemampuan untuk memahami pola penjualan, perilaku konsumen, dan tren pasar menjadi faktor kunci dalam meningkatkan daya saing dan profitabilitas. Penggunaan teknik analisis data, seperti data mining, memberikan peluang signifikan bagi perusahaan untuk melakukan segmentasi pasar yang lebih akurat sehingga strategi harga dan promosi dapat disesuaikan dengan karakteristik konsumen secara lebih tepat [3], [4].

Industri kopi di Indonesia mengalami pertumbuhan yang pesat dalam beberapa tahun terakhir, didorong oleh meningkatnya minat

masyarakat terhadap kopi sebagai bagian dari gaya hidup. Data dari International Coffee Organization (ICO) menunjukkan bahwa konsumsi kopi di Indonesia terus meningkat dari tahun ke tahun, baik untuk konsumsi domestik maupun untuk pasar ekspor. Persaingan yang semakin ketat antara produsen dan pengecer kopi menuntut setiap pelaku usaha untuk mampu mengidentifikasi segmen pasar secara lebih efektif. Dalam konteks ini, analisis berbasis data menjadi sangat penting untuk memahami perbedaan preferensi konsumen, pola pembelian, dan kontribusi setiap segmen terhadap total penjualan. Pendekatan berbasis segmentasi ini dapat memberikan dasar yang kuat dalam menentukan kebijakan harga yang kompetitif dan strategi promosi yang efektif [5], [6].

Permasalahan utama yang dihadapi pelaku usaha kopi adalah kesulitan dalam mengidentifikasi kelompok produk dan konsumen yang memiliki karakteristik serupa berdasarkan data penjualan yang kompleks. Banyak perusahaan masih mengandalkan pendekatan tradisional yang bersifat umum, seperti promosi massal atau penetapan harga tunggal, tanpa mempertimbangkan variasi margin laba dan volume penjualan antarproduk. Pendekatan seperti ini berpotensi menyebabkan ketidakefisienan dalam penggunaan sumber daya pemasaran dan hilangnya peluang untuk memaksimalkan keuntungan. Misalnya, produk dengan margin laba tinggi tetapi volumenya rendah mungkin memerlukan strategi promosi yang berbeda dibandingkan dengan produk dengan volume penjualan tinggi tetapi margin laba rendah [7], [8].

Salah satu solusi umum yang dapat digunakan untuk mengatasi masalah tersebut adalah dengan memanfaatkan teknik data mining untuk melakukan clustering atau pengelompokan data penjualan. Clustering memungkinkan pengelompokan produk atau pelanggan ke dalam segmen-segmen yang memiliki karakteristik serupa berdasarkan atribut tertentu. Dari berbagai algoritma clustering, metode K-Means menjadi salah satu yang paling populer dan banyak digunakan karena kesederhanaannya, efisiensinya, dan kemampuannya untuk mengelompokkan data dalam jumlah besar dengan cepat (Jain, 2010). Metode ini bekerja dengan meminimalkan jarak antara data dan pusat cluster, sehingga menghasilkan kelompok yang relatif homogen di dalamnya dan heterogen antar cluster [9].

Secara lebih spesifik, metode K-Means telah digunakan secara luas dalam penelitian sebelumnya untuk segmentasi pasar dan analisis

penjualan. Misalnya, Xu dan Wunsch (2005) menunjukkan bahwa K-Means mampu memberikan hasil clustering yang efektif dalam konteks pemasaran dengan data multi-dimensi. Penelitian lain oleh Ayesha et al. (2010) memanfaatkan K-Means untuk mengelompokkan data penjualan ritel dan berhasil mengidentifikasi segmen yang paling potensial untuk dilakukan promosi. Keunggulan metode ini terletak pada kemampuannya menangani data numerik dalam skala besar dan menghasilkan hasil yang mudah diinterpretasikan untuk pengambilan keputusan bisnis. Dalam konteks penjualan kopi, penerapan K-Means dapat membantu mengelompokkan produk berdasarkan kombinasi variabel margin laba dan volume penjualan, sehingga setiap kelompok dapat diperlakukan dengan strategi harga dan promosi yang berbeda sesuai potensinya.

Selain itu, literatur ilmiah juga menunjukkan bahwa pengelompokan berbasis K-Means dapat diintegrasikan dengan analisis strategis untuk mendukung optimalisasi pemasaran. Penelitian oleh Han et al. (2012) menekankan bahwa hasil clustering dapat digunakan untuk menyusun strategi yang lebih terarah, misalnya dengan mengalokasikan anggaran promosi yang lebih besar untuk cluster dengan potensi keuntungan tertinggi. Beberapa studi bahkan menggabungkan K-Means dengan metode evaluasi kinerja penjualan untuk memberikan rekomendasi yang lebih spesifik [10], [11], [12]. Hal ini menunjukkan bahwa pendekatan clustering bukan hanya sekadar proses teknis, tetapi juga memiliki implikasi langsung pada kebijakan pemasaran yang dapat meningkatkan daya saing perusahaan.

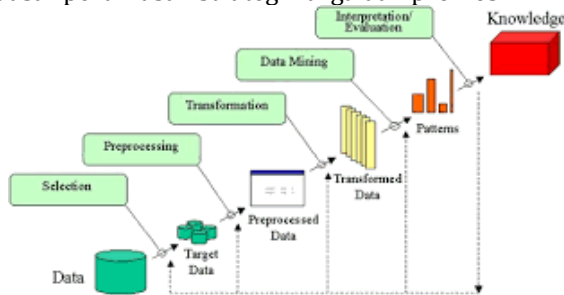
Namun demikian, tinjauan literatur menunjukkan adanya kesenjangan penelitian terkait penerapan K-Means dalam analisis penjualan produk kopi di pasar domestik. Sebagian besar penelitian sebelumnya lebih banyak difokuskan pada industri ritel secara umum atau pada komoditas pertanian tertentu selain kopi. Penelitian yang secara khusus mengkaji hubungan antara margin laba, volume penjualan, dan strategi harga serta promosi dalam konteks produk kopi masih relatif terbatas. Selain itu, masih jarang penelitian yang membahas implementasi praktis dari hasil clustering ke dalam strategi pemasaran yang dapat diadopsi langsung oleh pelaku usaha skala kecil hingga menengah di Indonesia [13], [14].

Berdasarkan uraian tersebut, penelitian ini bertujuan untuk menerapkan metode K-Means dalam mengelompokkan data penjualan produk

kopi berdasarkan margin laba dan volume penjualan guna merumuskan strategi harga dan promosi yang lebih efektif. Kebaruan penelitian ini terletak pada fokusnya yang spesifik pada industri kopi, penggunaan kombinasi variabel margin laba dan volume penjualan dalam proses clustering, serta orientasinya untuk menghasilkan rekomendasi strategis yang aplikatif bagi pelaku usaha. Dengan ruang lingkup penelitian yang mencakup analisis data penjualan Coffee Zeid, studi ini diharapkan dapat memberikan kontribusi bagi pengembangan strategi pemasaran berbasis data di sektor kopi dan menjadi referensi bagi penelitian lanjutan di bidang analisis penjualan dan data mining.

MATERIALS AND METHODS

Penelitian ini menggunakan pendekatan Knowledge Discovery in Databases (KDD) sebagai kerangka kerja utama dalam proses analisis data penjualan produk kopi. KDD dipilih karena mampu memberikan alur sistematis dalam menemukan pengetahuan yang berguna dari kumpulan data yang besar, melalui serangkaian tahap yang meliputi pemahaman data, pembersihan data, transformasi, penambangan data (data mining), hingga interpretasi hasil. Pendekatan ini sangat relevan dalam konteks penelitian ini, mengingat tujuan utamanya adalah untuk mengidentifikasi pola dan segmentasi penjualan yang dapat menjadi dasar perumusan strategi harga dan promosi.



Gambar 1. KDD

Tahap awal yang dilakukan adalah pemilihan data dari basis data penjualan Coffee Zeid. Data yang digunakan berisi informasi penjualan produk kopi yang mencakup variabel margin laba dan volume penjualan. Kedua variabel ini dipilih karena dinilai paling relevan dalam merepresentasikan tingkat profitabilitas dan popularitas penjualan setiap produk. Pemilihan variabel dilakukan secara selektif agar data yang dianalisis benar-benar mendukung tujuan penelitian dan mengurangi kompleksitas yang tidak perlu. Dengan demikian, data yang digunakan merupakan subset dari keseluruhan basis data

penjualan yang berfokus pada indikator utama keberhasilan penjualan.

Tahap berikutnya adalah pembersihan data untuk memastikan kualitas dan konsistensi informasi. Pada tahap ini dilakukan identifikasi dan penghapusan data duplikat agar tidak terjadi penggandaan informasi yang dapat mengganggu akurasi analisis. Nilai kosong atau missing values diperiksa secara menyeluruh, kemudian ditangani dengan metode penghapusan atau imputasi tergantung pada tingkat kelengkapan data yang tersedia. Selain itu, dilakukan deteksi terhadap outlier yang dapat memengaruhi hasil pengelompokan, dengan mempertimbangkan distribusi nilai pada setiap variabel. Penanganan outlier ini penting agar hasil clustering tidak terdistorsi oleh data yang menyimpang jauh dari mayoritas.

Data yang telah dibersihkan selanjutnya melalui tahap transformasi agar sesuai dengan kebutuhan algoritma clustering. Proses ini meliputi normalisasi data menggunakan metode min-max scaling untuk menyamakan skala nilai setiap variabel, sehingga perhitungan jarak antar data menjadi seimbang dan tidak dipengaruhi oleh perbedaan satuan atau rentang nilai. Normalisasi sangat penting pada metode K-Means karena algoritma ini sensitif terhadap skala data. Data yang telah dinormalisasi kemudian diformat agar kompatibel dengan perangkat lunak analisis yang digunakan, yaitu RapidMiner.

Proses inti dari penelitian ini adalah penambangan data (data mining) dengan menggunakan algoritma K-Means Clustering. Algoritma ini dipilih karena kesederhanaannya, efisiensinya, dan kemampuannya mengelompokkan data numerik dalam jumlah besar secara efektif. Proses clustering dimulai dengan menentukan jumlah cluster (K) yang optimal. Penentuan K dilakukan melalui percobaan bertahap dari $K=2$ hingga $K=10$, kemudian dievaluasi menggunakan metrik seperti Sum of Squared Errors (SSE) atau Davies-Bouldin Index untuk menemukan nilai K yang menghasilkan pemisahan cluster paling optimal. Selanjutnya, algoritma menginisialisasi pusat cluster dan secara iteratif mengelompokkan data hingga pusat cluster konvergen, yaitu ketika perubahan posisi pusat cluster antar iterasi menjadi sangat kecil.

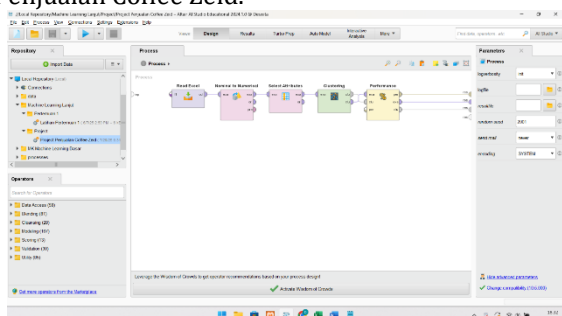
Hasil clustering kemudian diinterpretasikan dan dievaluasi untuk memahami karakteristik setiap cluster yang terbentuk. Interpretasi dilakukan dengan mengidentifikasi profil masing-masing cluster, misalnya kelompok produk dengan margin laba tinggi dan volume penjualan tinggi yang memiliki potensi besar untuk dipromosikan secara premium, atau kelompok

produk dengan volume penjualan tinggi namun margin laba rendah yang memerlukan evaluasi strategi harga. Visualisasi hasil pengelompokan dalam bentuk scatter plot digunakan untuk memudahkan pemahaman pola distribusi data antar cluster. Evaluasi tidak hanya berfokus pada validitas statistik hasil pengelompokan, tetapi juga pada relevansi bisnisnya terhadap tujuan penelitian.

Tahap akhir adalah penerapan hasil (deployment) dalam bentuk rekomendasi strategi harga dan promosi. Rekomendasi ini disusun berdasarkan profil cluster yang telah diidentifikasi. Misalnya, cluster dengan margin laba tinggi dan volume penjualan tinggi diprioritaskan untuk dipertahankan dan diperluas melalui promosi premium, sedangkan cluster dengan margin laba tinggi tetapi volume penjualan rendah dapat difokuskan pada peningkatan volume melalui promosi intensif. Sementara itu, cluster dengan margin laba rendah namun volume penjualan tinggi dapat dievaluasi strategi biayanya atau dipertimbangkan untuk strategi bundling, dan cluster dengan margin laba rendah serta volume penjualan rendah dapat dipertimbangkan untuk reposisi atau bahkan dihentikan dari lini produk. Dengan demikian, hasil dari tahapan KDD ini tidak hanya memberikan pengelompokan data yang terstruktur, tetapi juga memberikan dasar rekomendasi strategis yang dapat langsung diimplementasikan oleh manajemen Coffee Zeid untuk meningkatkan efektivitas strategi harga dan promosi.

RESULTS AND DISCUSSION

Desain proses analisis data penjualan kopi menggunakan perangkat lunak Altair AI Studio Educational 2024.1.0 dengan alur kerja (workflow) yang terdiri dari beberapa tahapan utama. Alur ini menggambarkan proses penerapan algoritma clustering (K-Means) pada dataset Project Penjualan Coffee Zeid.

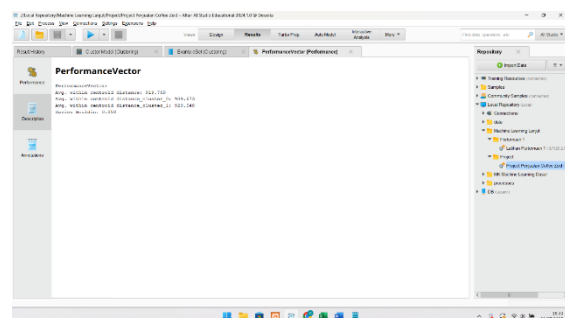


Gambar 2. Desain Algoritma K-Means

1. Read Excel Tahap ini berfungsi untuk membaca dataset yang disimpan dalam format Excel. Dataset yang diimpor berisi data penjualan kopi,

termasuk atribut yang akan dianalisis seperti margin laba dan volume penjualan. Operator ini mengubah data mentah dari file menjadi format internal yang dapat diproses oleh Altair AI Studio.

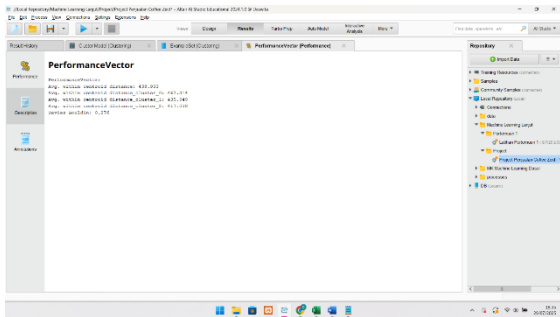
2. Nominal to Numerical Tahap ini digunakan untuk mengonversi data kategorikal (nominal) menjadi bentuk numerik. Proses ini penting karena algoritma K-Means membutuhkan input data dalam bentuk angka agar dapat menghitung jarak antar titik data. Misalnya, kategori produk atau jenis kopi yang awalnya berupa teks akan diubah menjadi representasi angka.
3. Select Attributes Pada tahap ini dilakukan pemilihan atribut (variabel) yang akan digunakan dalam analisis clustering. Peneliti memilih variabel yang relevan, seperti *margin laba* dan *volume penjualan*, untuk memastikan hasil clustering fokus pada indikator yang berkaitan langsung dengan tujuan penelitian.
4. Clustering Tahap inti dari proses, di mana algoritma K-Means digunakan untuk mengelompokkan data berdasarkan kesamaan karakteristik antar data. Proses ini melibatkan penentuan jumlah cluster (K) dan iterasi penghitungan jarak hingga pusat cluster stabil. Hasil dari tahap ini adalah pembagian data ke dalam beberapa kelompok dengan karakteristik yang serupa.
5. Performance Tahap ini digunakan untuk mengevaluasi hasil clustering. Evaluasi ini bisa dilakukan dengan berbagai metrik, seperti *Sum of Squared Errors (SSE)*, *Davies-Bouldin Index*, atau *Silhouette Coefficient*, guna menilai kualitas pemisahan antar cluster. Evaluasi ini membantu menentukan apakah jumlah cluster yang dipilih sudah optimal atau perlu disesuaikan.



Gambar 3. Hasil K2

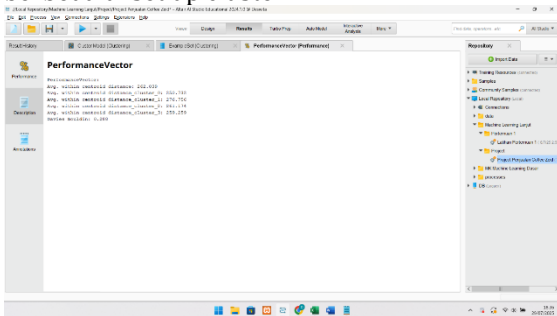
Hasil ini menunjukkan bahwa model K-Means yang diterapkan pada dataset penjualan kopi menghasilkan pengelompokan yang cukup rapat dan terpisah dengan baik. Cluster 0 sedikit lebih rapat daripada cluster 1, namun perbedaannya kecil. Nilai Davies-Bouldin Index yang rendah (0.258) mengindikasikan bahwa kedua cluster memiliki perbedaan karakteristik yang jelas dan

tidak banyak tumpang tindih. Hal ini berarti pengelompokan ini dapat digunakan sebagai dasar yang kuat untuk analisis lebih lanjut, seperti perumusan strategi harga dan promosi berdasarkan karakteristik masing-masing cluster.



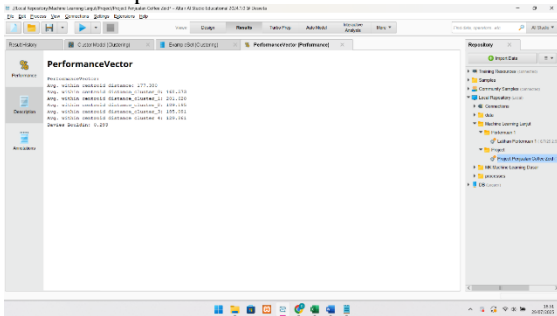
Gambar 4. Hasil K3

Model K-Means dengan K=3 ini menghasilkan pengelompokan data penjualan kopi yang rapat dan terpisah dengan baik. Cluster 2 adalah yang paling kompak, diikuti cluster 1, sedangkan cluster 0 sedikit lebih tersebar. Nilai Davies-Bouldin Index yang rendah (0.276) mengindikasikan bahwa ketiga cluster memiliki perbedaan karakteristik yang jelas, sehingga hasil ini dapat menjadi dasar yang kuat untuk perumusan strategi harga dan promosi yang berbeda di setiap cluster.

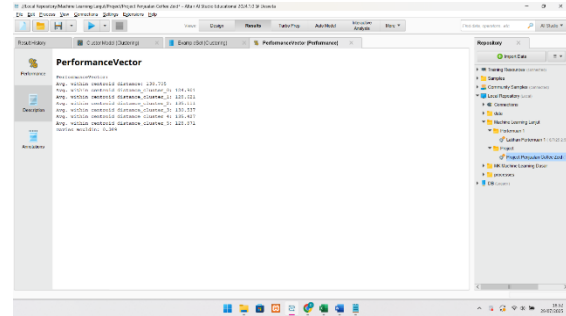


Gambar 5. Hasil K4

Model K-Means dengan K=4 menghasilkan kerapatan cluster yang cukup baik secara internal (ditunjukkan oleh nilai rata-rata jarak yang relatif kecil), namun pemisahan antar cluster tidak sebaik pada K=2 dan K=3 karena Davies-Bouldin Index cukup tinggi (1.048). Ini mengindikasikan bahwa meskipun tiap cluster cukup rapat, batas antar cluster tidak terlalu jelas sehingga potensi tumpang tindih cukup besar.

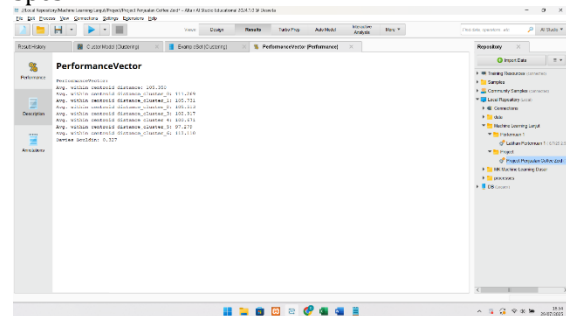


Gambar 6. Hasil K5
Model K-Means dengan K=5 ini menghasilkan kerapatan data yang cukup baik di setiap cluster, namun pemisahan antar cluster tidak optimal karena Davies-Bouldin Index masih di atas 1. Hal ini menunjukkan bahwa meskipun data dalam setiap cluster cukup rapat, batas antar cluster kurang tegas dan ada potensi kesamaan karakteristik antar cluster. Secara umum, dari segi kualitas pemisahan, hasil ini lebih lemah dibandingkan dengan model K=2 atau K=3.



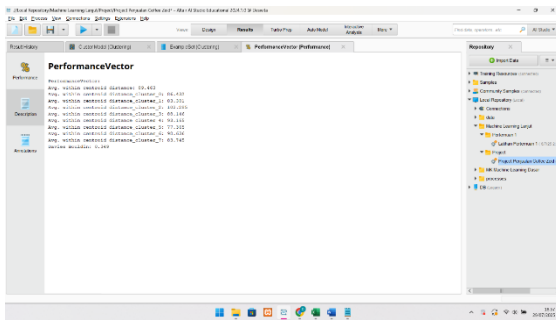
Gambar 7. Hasil K6

Model K-Means dengan K=6 menghasilkan cluster yang sangat rapat dan pemisahan antar cluster yang baik. Nilai rata-rata jarak ke pusat cluster sangat kecil, menandakan setiap cluster memiliki konsistensi internal yang tinggi. Nilai Davies-Bouldin Index yang rendah (0.309) menunjukkan bahwa meskipun jumlah cluster cukup banyak, batas antar cluster tetap jelas. Ini menandakan K=6 dapat menjadi kandidat yang baik jika tujuan analisis adalah membuat segmentasi yang lebih spesifik.



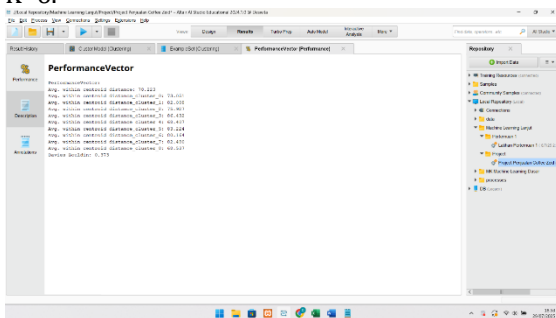
Gambar 8. Hasil K7

Model K-Means dengan K=7 menghasilkan cluster yang cukup rapat secara internal (ditunjukkan oleh nilai jarak rata-rata yang rendah), tetapi pemisahan antar cluster tidak sebaik model dengan jumlah cluster lebih sedikit seperti K=2, K=3, atau K=6. Nilai Davies-Bouldin Index yang berada di atas 1 menunjukkan adanya kemiripan antar cluster yang cukup besar, sehingga batas antar cluster kurang tegas.



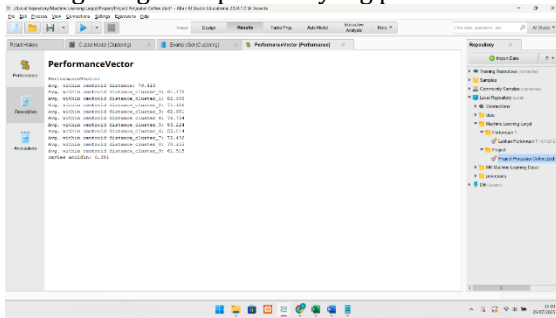
Gambar 9. Hasil K8

Model K-Means dengan K=8 menghasilkan cluster yang secara umum rapat secara internal dan memiliki pemisahan yang cukup baik antar cluster. Beberapa cluster seperti cluster_1 dan cluster_7 memiliki kekompakan sangat tinggi, sedangkan cluster_2 dan cluster_5 sedikit lebih menyebar. Nilai Davies-Bouldin Index yang rendah (0.349) menunjukkan hasil yang cukup baik, meskipun jumlah cluster yang banyak ini mungkin terlalu rinci untuk analisis strategis jika dibandingkan dengan jumlah cluster yang lebih sedikit seperti K=3 atau K=6.



Gambar 10 Hasil K9

Model K-Means dengan K=9 menghasilkan cluster yang rapat secara internal dan memiliki pemisahan antar cluster yang cukup baik. Nilai Davies-Bouldin Index yang rendah mengindikasikan kualitas pemisahan yang memadai. Namun, jumlah cluster yang cukup banyak ini berpotensi membuat interpretasi bisnis menjadi lebih kompleks, terutama jika tujuannya adalah menentukan strategi harga dan promosi yang praktis.



Gambar 11. Hasil K10

Model K-Means dengan K=10 menghasilkan cluster yang rapat secara internal dan pemisahan antar cluster yang cukup jelas. Beberapa cluster seperti

cluster_0, cluster_3, cluster_6, dan cluster_9 sangat rapat, sedangkan cluster_1 sedikit lebih menyebar dibanding lainnya. Meskipun Davies-Bouldin Index rendah, jumlah cluster yang terlalu banyak dapat membuat interpretasi strategis menjadi lebih rumit, terutama jika tujuannya adalah merumuskan strategi harga dan promosi yang praktis.

CONCLUSION

Penelitian ini berhasil menerapkan metode K-Means Clustering untuk mengelompokkan data penjualan produk kopi berdasarkan variabel margin laba dan volume penjualan, sehingga dapat digunakan sebagai dasar dalam perumusan strategi harga dan promosi yang lebih efektif. Hasil analisis menunjukkan bahwa model dengan jumlah cluster tertentu, khususnya K=2, K=3, dan K=6, memberikan pemisahan dan kekompakan cluster yang optimal sesuai indikator Davies-Bouldin Index. Temuan ini mengindikasikan adanya segmen produk yang berbeda secara signifikan, baik dari segi profitabilitas maupun popularitas penjualan, sehingga memerlukan pendekatan strategi pemasaran yang berbeda. Implikasi dari penelitian ini adalah memberikan panduan berbasis data bagi pelaku usaha kopi dalam menyusun strategi yang lebih tepat sasaran, efisien, dan berpotensi meningkatkan profitabilitas. Kontribusi penelitian ini terletak pada fokus penerapannya di sektor kopi dengan kombinasi variabel yang relevan serta orientasi pada rekomendasi strategis yang aplikatif. Penelitian selanjutnya disarankan untuk mengintegrasikan metode clustering dengan analisis perilaku konsumen atau model prediktif penjualan, sehingga hasilnya dapat memberikan wawasan yang lebih komprehensif dalam pengambilan keputusan pemasaran

REFERENCE

- [1] M. Yunianto, F. Anwar, D. Nur Septianingsih, T. Dwi Ardyanto, and R. Farits Pradana, "KLASIFIKASI KANKER PARU PARU MENGGUNAKAN NAÏVE BAYES DENGAN VARIASI FILTER DAN EKSTRAKSI CIRI GRAY LEVEL CO-OCCURANCE MATRIX (GLCM)," *Indonesian Journal of Applied Physics*, vol. 11, no. 2, 2021.
- [2] L. Listyalina, "PENINGKATAN KUALITAS CITRA FOTO RONTGEN SEBAGAI MEDIA DETEKSI KANKER PARU," *Jurnal Teknologi Informasi*, 2017.
- [3] N. Elsayed, Z. Saad Zaghloul, and M. Bayoumi, "Brain Computer Interface: EEG Signal Preprocessing Issues and Solutions," 2017.

- [4] D. Candra and R. Novitasari, "Klasifikasi Sinyal EEG Menggunakan Metode Fuzzy C-Means Clustering (FCM) Dan Adaptive Neighborhood Modified Backpropagation (ANMBP)," *Jurnal Online Fakultas Sains dan Teknologi*, 2022, [Online]. Available: http://sccn.ucsd.edu/~arno/fam2data/publicly_
- [5] R. Widadi, B. ArifWidodo, and D. Zulherman, "Klasifikasi Sinyal EEG pada Sistem BCI Pergerakan Jari Manusia Menggunakan Convolutional Neural Network," *Jurnal Techno.com*, vol. 19, no. 4, pp. 459–467, 2020.
- [6] T. M. M. Keumala, M. Melinda, and S. Syahrial, "Decision tree method to classify the electroencephalography-based emotion data," *JURNAL INFOTEL*, vol. 14, no. 1, pp. 37–49, Feb. 2022, doi: 10.20895/infotel.v14i1.750.
- [7] H. B. Kim, Y. S. Park, J. E. Lee, K. Do Han, and Y. H. Park, "Study on relationship between self-recognition of voice disorder and mental health status: Korea National Health and Nutrition Examination Survey," *J Affect Disord*, vol. 338, pp. 482–486, Oct. 2023, doi: 10.1016/j.jad.2023.05.082.
- [8] U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, and H. Adeli, "Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals," *Comput Biol Med*, vol. 100, pp. 270–278, Sep. 2018, doi: 10.1016/j.compbimed.2017.09.017.
- [9] A. Novianto and M. D. Anasanti, "Autism Spectrum Disorder (ASD) Identification Using Feature-Based Machine Learning Classification Model," *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, vol. 17, no. 3, p. 259, Jul. 2023, doi: 10.22146/ijccs.83585.
- [10] M. Subramanian, J. Cho, V. E. Sathishkumar, and O. S. Naren, "Multiple Types of Cancer Classification Using CT/MRI Images Based on Learning Without Forgetting Powered Deep Learning Models," vol. 11, pp. 10336–10354, 2023, doi: 10.1109/ACCESS.2023.3240443.
- [11] J. Yi, X. Liu, S. Cheng, L. Chen, and S. Zeng, "Multi-scale window transformer for cervical cytopathology image recognition," vol. 24, pp. 314–321, 2024, doi: 10.1016/j.csbj.2024.04.028.
- [12] S. M. Abd-Alhalem, H. S. Marie, W. El-Shafai, T. Altameem, R. S. Rathore, and T. M. Hassan, "Cervical cancer classification based on a bilinear convolutional neural network approach and random projection," vol. 127, 2024, doi: 10.1016/j.engappai.2023.107261.
- [13] P. Malathi and G. Charlyn Pushpa Latha, "Classification of Multi-view Digital Mammogram Images Using SMO-WkNN," vol. 46, no. 2, pp. 1741–1758, 2023, doi: 10.32604/csse.2023.035185.
- [14] M. Salehi, A. Vafaei Sadr, S. R. Mahdavi, H. Arabi, I. Shiri, and R. Reiazi, "Deep Learning-based Non-rigid Image Registration for High-dose Rate Brachytherapy in Inter-fraction Cervical Cancer," vol. 36, no. 2, pp. 574–587, 2023, doi: 10.1007/s10278-022-00732-6.