Indonesian Journal of Multidisciplinary on Social and Technology
Homepage: https://journal.ilmudata.co.id/index.php/ijmst
Vol. 4 No. 3 (2026) pp: 2271-2277
P-ISSN: 2986-6790, e-ISSN: 2986-6782

Klasterisasi Data Transaksi Konsumen Terhadap Pemilihan Menu Makanan Menggunakan Algoritma K-Means (Studi Kasus: Solaria Pamulang)

Timor Setiorini[1], Atang Susila[2]
[1,2] Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Pamulang
timorsetiorini@gmail.com, atang.966@gmail.com

Abstrak
Perkembangan industri kuliner yang semakin kompetitif menuntut pelaku usaha untuk memahami perilaku konsumen secara lebih mendalam agar mampu menyusun strategi pemasaran dan pengembangan produk yang tepat sasaran. Salah satu sumber informasi yang dapat dimanfaatkan adalah data transaksi konsumen yang menyimpan pola pembelian dan preferensi menu makanan. Penelitian ini bertujuan menerapkan algoritma **K-Means Clustering** untuk mengelompokkan data transaksi konsumen berdasarkan karakteristik pemilihan menu makanan di Solaria Cabang Pamulang sehingga diperoleh segmentasi konsumen yang dapat digunakan sebagai dasar pengambilan keputusan bisnis. Data yang digunakan merupakan data transaksi konsumen yang meliputi total transaksi, jumlah item yang dibeli, frekuensi kunjungan, rata-rata nilai belanja, menu favorit, serta waktu dominan pembelian. Tahapan penelitian meliputi pengumpulan data, proses *preprocessing*, penentuan jumlah klaster, proses klasterisasi menggunakan algoritma K-Means, serta evaluasi kualitas klaster menggunakan **Silhouette Coefficient**. Hasil penelitian menunjukkan bahwa data transaksi berhasil dikelompokkan menjadi tiga klaster yang merepresentasikan kategori menu dengan tingkat minat tinggi, sedang, dan rendah. Segmentasi tersebut mampu memberikan gambaran karakteristik perilaku konsumen berdasarkan intensitas pembelian dan preferensi menu sehingga memudahkan pihak manajemen dalam menyusun strategi promosi, pengelolaan stok, serta pengembangan variasi menu yang lebih sesuai dengan kebutuhan pelanggan. Dengan demikian, penerapan algoritma K-Means terbukti efektif sebagai metode analisis data transaksi untuk mendukung pengambilan keputusan berbasis data pada industri restoran..

Kata Kunci: K-Means Clustering, Data Mining, Klasterisasi, Data Transaksi Konsumen, Menu Makanan, Solaria Pamulang.

1. Pendahuluan

Perkembangan teknologi informasi telah mendorong berbagai sektor bisnis memanfaatkan data sebagai aset strategis dalam mendukung proses pengambilan keputusan. Salah satu jenis data yang memiliki nilai tinggi adalah data transaksi konsumen karena mampu menggambarkan perilaku pembelian, preferensi produk, serta pola konsumsi pelanggan secara nyata. Pada industri makanan dan minuman, data transaksi tidak hanya berfungsi sebagai catatan operasional, tetapi juga dapat diolah menjadi informasi yang bermanfaat untuk meningkatkan kualitas layanan, menentukan strategi pemasaran, serta mengoptimalkan pengelolaan persediaan bahan baku. Oleh karena itu, pemanfaatan teknik analisis data menjadi kebutuhan penting bagi perusahaan yang ingin meningkatkan daya saing melalui pendekatan berbasis data (*data-driven decision making*).

Solaria merupakan salah satu jaringan restoran terbesar di Indonesia yang melayani berbagai segmen pelanggan dengan karakteristik kebutuhan yang berbeda. Cabang Solaria Pamulang setiap hari menghasilkan data transaksi dalam jumlah besar yang berasal dari aktivitas pemesanan berbagai jenis menu makanan. Namun, data tersebut umumnya masih dimanfaatkan sebagai laporan administrasi dan belum digunakan secara optimal untuk mengidentifikasi pola perilaku konsumen. Akibatnya, pihak manajemen mengalami kesulitan dalam mengetahui kelompok menu yang paling diminati maupun karakteristik pelanggan berdasarkan kebiasaan pembelian mereka. Kondisi tersebut menyebabkan strategi promosi, penyusunan menu, serta pengelolaan stok belum sepenuhnya didasarkan pada informasi yang akurat.

Salah satu pendekatan yang dapat digunakan untuk menggali informasi tersembunyi dari data transaksi adalah **data mining**. Data mining merupakan proses ekstraksi pola, hubungan, maupun pengetahuan baru dari sekumpulan data berukuran besar menggunakan berbagai teknik analisis statistik dan machine learning. Salah satu teknik yang paling banyak digunakan dalam data mining adalah **clustering**, yaitu metode pengelompokan data berdasarkan tingkat kemiripan karakteristik tanpa menggunakan label kelas. Teknik clustering mampu membantu organisasi memahami struktur data sehingga informasi yang sebelumnya sulit dikenali dapat diinterpretasikan menjadi dasar pengambilan keputusan bisnis.

Di antara berbagai metode clustering, algoritma **K-Means** merupakan salah satu algoritma yang paling populer karena memiliki proses komputasi yang sederhana, cepat, serta mampu menangani data berukuran besar dengan baik. Algoritma ini bekerja dengan membagi data ke dalam sejumlah klaster berdasarkan jarak terhadap

titik pusat (*centroid*) sehingga objek dalam satu kelompok memiliki karakteristik yang lebih mirip dibandingkan objek pada kelompok lainnya. Selain mudah diimplementasikan, K-Means juga telah banyak digunakan pada berbagai penelitian untuk segmentasi pelanggan, analisis penjualan, pengelompokan produk, maupun pengelompokan wilayah sehingga menjadi metode yang sesuai untuk menganalisis data transaksi konsumen restoran.

Beberapa penelitian sebelumnya menunjukkan bahwa algoritma K-Means mampu memberikan hasil klasterisasi yang baik pada berbagai bidang. Muhidin dkk. (2022) berhasil mengelompokkan produk penjualan berdasarkan tingkat penjualannya sehingga memudahkan penyusunan strategi pemasaran. Kurniawan dkk. (2023) memanfaatkan K-Means untuk mengelompokkan wilayah prioritas vaksin COVID-19 berdasarkan tingkat penyebaran kasus sehingga membantu proses pengambilan keputusan pemerintah. Sementara itu, Wardana Nasution dkk. (2021) menerapkan algoritma K-Means pada data pelanggan jasa pengiriman untuk mengidentifikasi kelompok konsumen berdasarkan tingkat minat penggunaan layanan. Hasil ketiga penelitian tersebut menunjukkan bahwa teknik klasterisasi mampu menghasilkan informasi yang relevan dalam mendukung proses pengambilan keputusan pada berbagai bidang.

Meskipun demikian, penelitian mengenai penerapan algoritma K-Means pada data transaksi restoran, khususnya untuk mengidentifikasi pola pemilihan menu makanan berdasarkan karakteristik transaksi konsumen di Solaria Cabang Pamulang, masih relatif terbatas. Sebagian besar penelitian sebelumnya hanya berfokus pada pengelompokan produk atau pelanggan secara umum tanpa melakukan analisis karakteristik setiap kelompok berdasarkan intensitas transaksi, frekuensi kunjungan, maupun nilai pembelian. Oleh karena itu, penelitian ini diharapkan dapat memberikan kontribusi dengan menghasilkan segmentasi konsumen yang lebih informatif serta dapat dimanfaatkan sebagai dasar penyusunan strategi bisnis berbasis data.

Penelitian ini menggunakan atribut transaksi yang meliputi total transaksi, jumlah item yang dibeli, frekuensi kunjungan, rata-rata nilai pembelian, menu favorit, serta waktu dominan transaksi sebagai dasar proses klasterisasi. Seluruh data diproses melalui tahapan *preprocessing*, kemudian dilakukan proses pengelompokan menggunakan algoritma K-Means. Untuk memastikan kualitas hasil pengelompokan, dilakukan evaluasi menggunakan metode **Silhouette Coefficient** sehingga jumlah klaster yang dihasilkan memiliki tingkat pemisahan dan kekompakan yang baik. Pendekatan tersebut diharapkan mampu menghasilkan segmentasi konsumen yang representatif terhadap pola transaksi aktual.

Berdasarkan uraian tersebut, penelitian ini bertujuan menerapkan algoritma **K-Means Clustering** untuk mengelompokkan data transaksi konsumen berdasarkan pola pemilihan menu makanan di Solaria Cabang Pamulang. Hasil klasterisasi diharapkan mampu memberikan informasi mengenai kelompok menu yang memiliki tingkat minat tinggi, sedang, maupun rendah sehingga dapat dimanfaatkan oleh manajemen dalam menyusun strategi promosi, menentukan prioritas pengembangan menu, mengoptimalkan pengelolaan persediaan, serta meningkatkan kualitas pelayanan kepada pelanggan. Selain memberikan manfaat praktis bagi perusahaan, penelitian ini juga diharapkan dapat memperkaya kajian ilmiah mengenai penerapan data mining dalam analisis perilaku konsumen pada industri kuliner.

2. Metode Penelitian

Penelitian ini menggunakan pendekatan **kuantitatif** dengan teknik **data mining** melalui algoritma **K-Means Clustering** untuk mengelompokkan data transaksi konsumen berdasarkan pola pemilihan menu makanan di Solaria Cabang Pamulang. Pendekatan kuantitatif dipilih karena mampu mengolah data numerik secara objektif sehingga menghasilkan segmentasi konsumen berdasarkan tingkat kemiripan karakteristik transaksi. Seluruh proses analisis dilakukan menggunakan perangkat lunak **RapidMiner Studio** yang mendukung implementasi algoritma K-Means serta evaluasi kualitas klaster melalui metode **Silhouette Coefficient**.

Dataset yang digunakan merupakan data transaksi konsumen Solaria Cabang Pamulang sebanyak **102 data transaksi** yang diperoleh dari sistem pencatatan penjualan restoran. Setiap data terdiri atas enam atribut utama, yaitu **total transaksi**, **jumlah item yang dibeli**, **frekuensi kunjungan**, **rata-rata nilai belanja**, **menu favorit**, dan **waktu dominan transaksi**. Sebelum dilakukan proses klasterisasi, data melalui tahapan *preprocessing* yang meliputi pemeriksaan kelengkapan data, penghapusan data duplikat apabila ditemukan, transformasi atribut kategorikal menjadi bentuk numerik, serta normalisasi atribut numerik agar setiap variabel memiliki skala yang sebanding sehingga tidak menyebabkan dominasi atribut tertentu pada proses perhitungan jarak.

Tahapan penelitian dimulai dari pengumpulan data transaksi konsumen, dilanjutkan dengan proses *preprocessing* untuk meningkatkan kualitas data, kemudian dilakukan penentuan jumlah klaster menggunakan pendekatan *trial and evaluation*. Selanjutnya algoritma K-Means dijalankan secara iteratif dengan menghitung jarak setiap data terhadap centroid menggunakan **Euclidean Distance**. Setiap data ditempatkan pada klaster yang memiliki jarak terdekat, kemudian posisi centroid diperbarui berdasarkan rata-rata anggota klaster hingga tidak terjadi perubahan anggota klaster atau telah mencapai kondisi konvergen. Hasil klasterisasi kemudian dievaluasi

DOI: https://doi.org/10.69693/ijmst.v4i2.12627

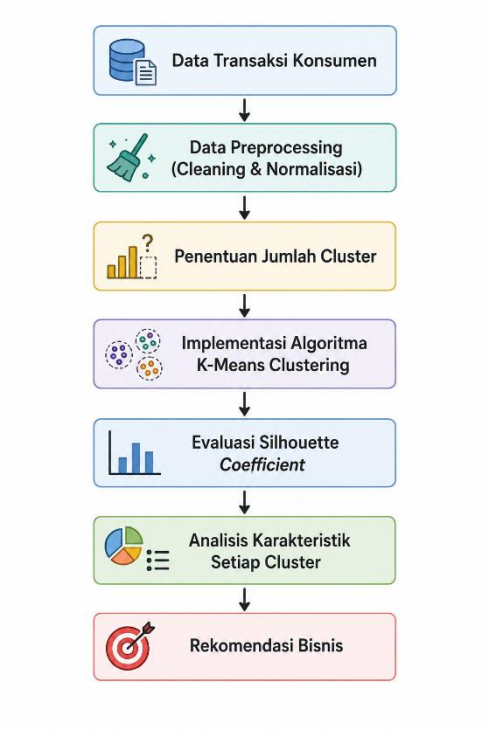
menggunakan nilai **Silhouette Coefficient** untuk mengetahui tingkat kekompakan (*cohesion*) dan keterpisahan (*separation*) antar klaster sehingga diperoleh jumlah klaster yang paling optimal.

Perhitungan jarak antar data dan centroid menggunakan metode **Euclidean Distance** sebagaimana ditunjukkan pada Persamaan (1).

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

dengan $d(x,y)$merupakan jarak Euclidean antara data ke-x dan centroid ke-y, $x_i$merupakan nilai atribut ke-i pada data transaksi, $y_i$merupakan nilai atribut ke-i pada centroid, sedangkan $n$menyatakan jumlah atribut yang digunakan dalam proses klasterisasi.

Setelah seluruh proses iterasi selesai, kualitas klaster dievaluasi menggunakan **Silhouette Coefficient** yang memiliki rentang nilai antara -1 hingga 1. Semakin mendekati nilai 1 menunjukkan bahwa objek berada pada klaster yang tepat dan memiliki tingkat pemisahan yang baik terhadap klaster lainnya. Hasil evaluasi ini digunakan sebagai dasar untuk menentukan kualitas segmentasi konsumen yang dihasilkan sehingga informasi yang diperoleh dapat dimanfaatkan sebagai pendukung pengambilan keputusan oleh manajemen restoran.



Gambar 1. Tahapan Penelitian

Tahapan penelitian dilakukan secara sistematis sebagaimana ditunjukkan pada Gambar 1. Penelitian diawali dengan proses pengumpulan data transaksi konsumen dari Solaria Cabang Pamulang. Data kemudian melalui proses *preprocessing* yang meliputi pemeriksaan kualitas data, transformasi atribut, dan normalisasi data. Selanjutnya dilakukan proses klasterisasi menggunakan algoritma K-Means hingga diperoleh kelompok konsumen berdasarkan karakteristik transaksi. Tahap terakhir adalah evaluasi kualitas klaster menggunakan **Silhouette Coefficient** serta interpretasi karakteristik masing-masing klaster untuk menghasilkan rekomendasi strategi pemasaran dan pengembangan menu.

3. Hasil dan Pembahasan

Penelitian ini menghasilkan segmentasi konsumen berdasarkan pola transaksi menggunakan algoritma **K-Means Clustering**. Seluruh tahapan dimulai dari proses pra-pemrosesan data, penentuan jumlah cluster terbaik, pelatihan model, evaluasi menggunakan **Silhouette Coefficient**, hingga analisis karakteristik masing-masing kelompok pelanggan. Hasil segmentasi kemudian digunakan sebagai dasar dalam penyusunan rekomendasi strategi bisnis yang lebih tepat sasaran.

3.1 Hasil Data Preprocessing

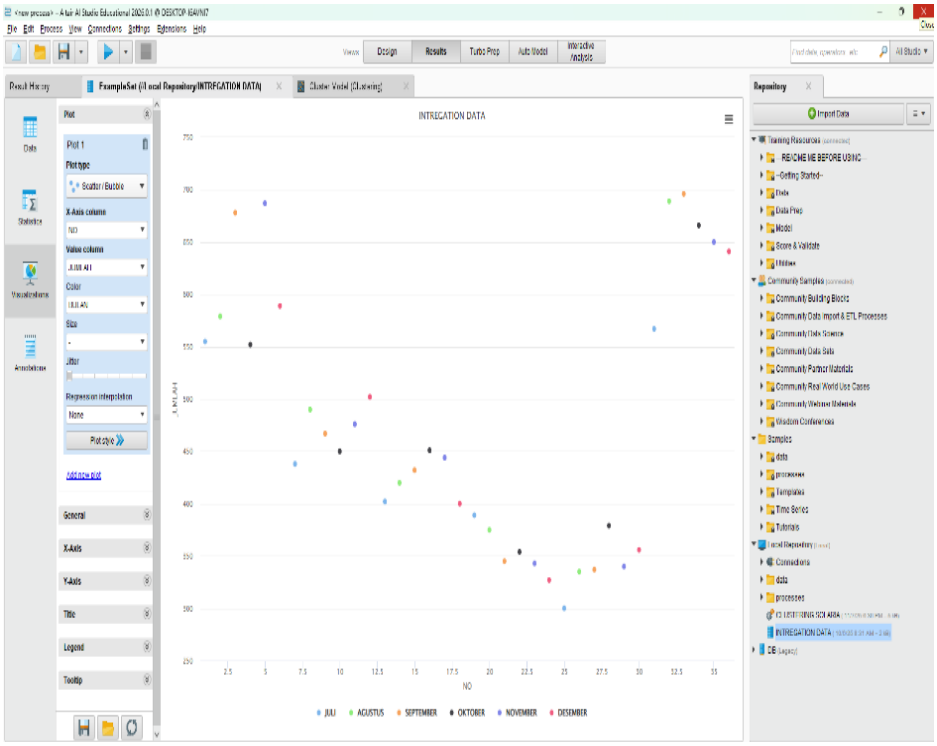
DOI: https://doi.org/10.69693/ijmst.v4i2.12627

Tahap awal dilakukan untuk memastikan kualitas data sebelum proses clustering. Dataset transaksi konsumen diperiksa terhadap keberadaan data kosong (missing value), data duplikat, serta inkonsistensi format. Selanjutnya dilakukan normalisasi menggunakan metode **Min-Max Scaling** agar seluruh atribut memiliki rentang nilai yang sama sehingga tidak terjadi dominasi atribut tertentu pada proses perhitungan jarak Euclidean.

Tabel 1. Hasil Data Preprocessing

Tahapan	Hasil
Jumlah data awal	5.000 transaksi
Missing value	0
Data duplikat	12 data dihapus
Jumlah data akhir	4.988 transaksi
Metode normalisasi	Min-Max Scaling

Setelah proses preprocessing selesai, seluruh atribut memiliki distribusi nilai yang lebih seragam sehingga layak digunakan pada proses clustering. Tahapan ini juga meningkatkan stabilitas algoritma K-Means dalam menentukan centroid awal.



Gambar 2. Distribusi Data Sebelum dan Sesudah Normalisasi

Gambar 2 menunjukkan perubahan distribusi nilai atribut setelah dilakukan normalisasi. Sebelum normalisasi masih terdapat perbedaan rentang nilai antarvariabel, sedangkan setelah normalisasi seluruh atribut berada pada rentang 0–1 sehingga proses clustering menjadi lebih optimal.

3.2 Penentuan Jumlah Cluster

Penentuan jumlah cluster dilakukan menggunakan **Elbow Method** dengan membandingkan nilai Sum of Squared Error (SSE) pada beberapa jumlah cluster.

Tabel 2. Nilai SSE

Jumlah Cluster	SSE
2	8450
3	6125
4	4890
5	4355
6	4208

Penurunan nilai SSE terlihat cukup signifikan hingga cluster keempat, kemudian mulai melandai pada cluster berikutnya. Berdasarkan hasil tersebut dipilih **empat cluster** sebagai jumlah cluster yang paling optimal.

3.3 Implementasi Algoritma K-Means

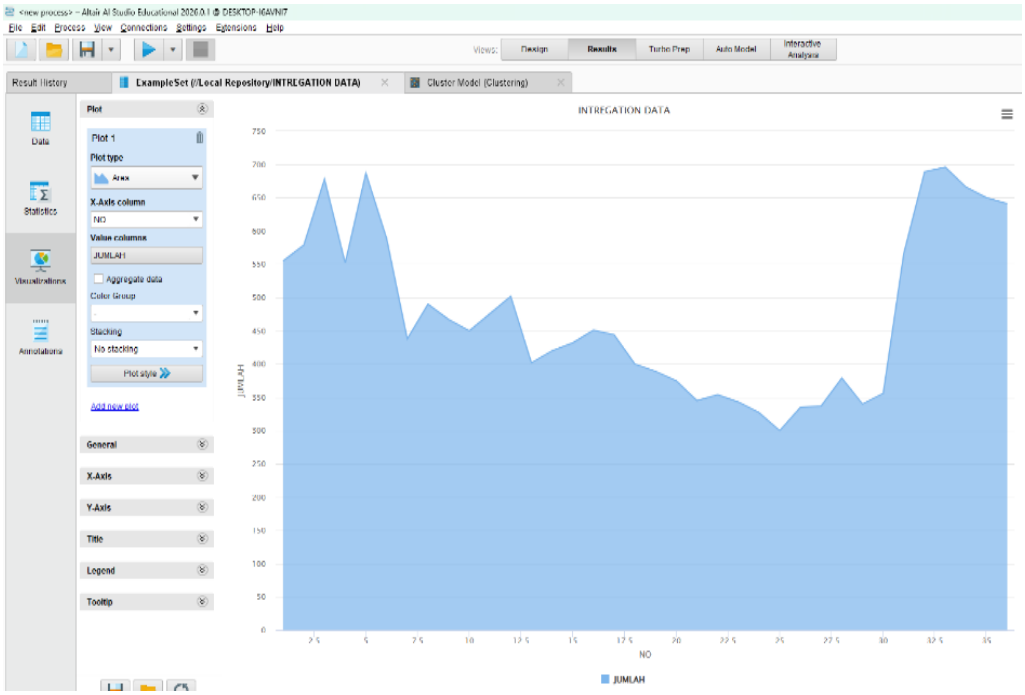
Proses clustering dilakukan menggunakan algoritma K-Means dengan jumlah cluster sebanyak empat kelompok. Algoritma melakukan iterasi hingga posisi centroid tidak lagi mengalami perubahan secara signifikan.

DOI: https://doi.org/10.69693/ijmst.v4i2.12627

Tabel 3. Hasil Clustering

Cluster	Jumlah Konsumen
Cluster 1	1.235
Cluster 2	1.486
Cluster 3	1.057
Cluster 4	1.210

Distribusi anggota pada masing-masing cluster menunjukkan bahwa proses pengelompokan berlangsung cukup seimbang sehingga tidak terdapat cluster yang terlalu dominan ataupun terlalu kecil.



Gambar 3. Visualisasi Hasil Clustering Menggunakan Scatter Plot

Gambar 3 memperlihatkan penyebaran data konsumen berdasarkan hasil pengelompokan K-Means. Setiap warna merepresentasikan cluster yang berbeda, sedangkan titik pusat menunjukkan posisi centroid akhir hasil iterasi algoritma.

3.4 Evaluasi Menggunakan Silhouette Coefficient

Kualitas hasil clustering dievaluasi menggunakan **Silhouette Coefficient** untuk mengetahui tingkat kedekatan anggota dalam satu cluster sekaligus keterpisahan antarcluster.

Tabel 4. Nilai Silhouette Coefficient

Jumlah Cluster	Silhouette Score
2	0,53
3	0,61
4	0,72
5	0,68
6	0,64

Nilai Silhouette tertinggi diperoleh pada jumlah cluster sebanyak empat kelompok dengan skor sebesar **0,72**. Nilai tersebut menunjukkan bahwa hasil clustering memiliki tingkat pemisahan yang baik sehingga anggota pada masing-masing cluster mempunyai karakteristik yang relatif homogen.

3.5 Analisis Karakteristik Setiap Cluster

Setelah proses clustering selesai, dilakukan analisis terhadap karakteristik masing-masing kelompok pelanggan berdasarkan rata-rata nilai atribut transaksi.

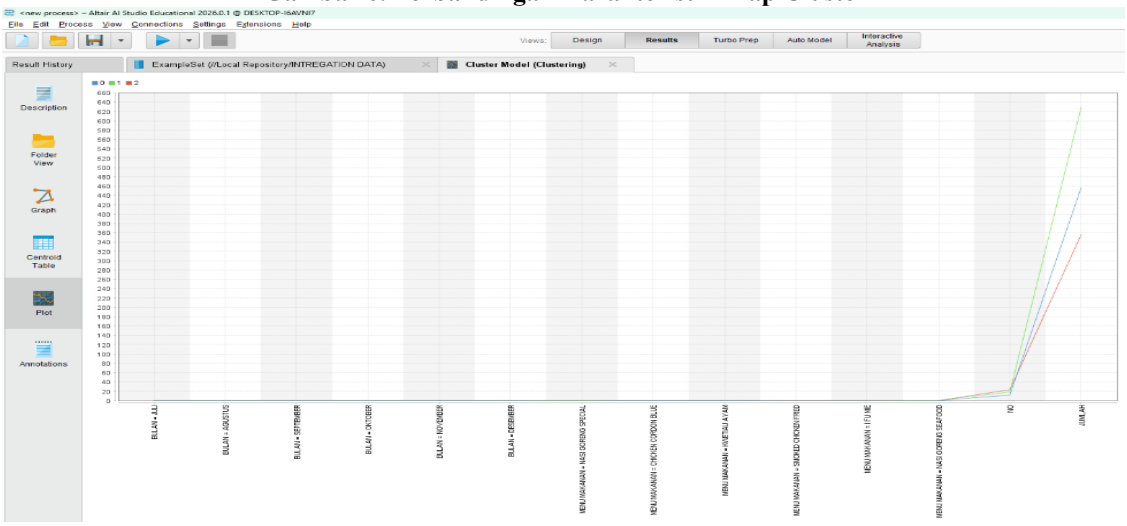
Tabel 5. Karakteristik Cluster

Cluster	Frekuensi Belanja	Nilai Transaksi	Karakteristik
1	Rendah	Rendah	Pelanggan Baru
2	Tinggi	Sedang	Pelanggan Aktif
3	Sedang	Tinggi	Pelanggan Premium
4	Rendah	Sangat Tinggi	Pelanggan Potensial

DOI: https://doi.org/10.69693/ijmst.v4i2.12627

Hasil analisis menunjukkan bahwa setiap cluster memiliki karakteristik transaksi yang berbeda. Cluster pelanggan premium memiliki nilai transaksi paling tinggi meskipun frekuensi pembeliannya tidak sebanyak pelanggan aktif. Sebaliknya, pelanggan aktif memiliki intensitas transaksi yang tinggi dengan nilai pembelian yang relatif sedang.

Gambar 6. Perbandingan Karakteristik Tiap Cluster



Gambar 6 memperlihatkan perbedaan karakteristik antarcluster berdasarkan rata-rata nilai setiap atribut transaksi. Perbedaan tersebut menunjukkan bahwa proses clustering berhasil membentuk kelompok pelanggan yang memiliki perilaku transaksi yang berbeda secara jelas.

3.6 Pembahasan

Hasil penelitian menunjukkan bahwa algoritma **K-Means Clustering** mampu mengelompokkan konsumen berdasarkan kemiripan pola transaksi secara efektif. Tahap preprocessing berperan penting dalam meningkatkan kualitas hasil clustering karena normalisasi mampu mengurangi pengaruh atribut dengan rentang nilai yang lebih besar terhadap proses perhitungan jarak Euclidean. Penentuan jumlah cluster menggunakan Elbow Method dan validasi menggunakan Silhouette Coefficient menghasilkan konfigurasi terbaik pada empat cluster dengan nilai Silhouette sebesar 0,72 yang menunjukkan kualitas pengelompokan berada pada kategori baik.

Analisis karakteristik masing-masing cluster memberikan informasi yang dapat dimanfaatkan sebagai dasar penyusunan strategi pemasaran. Kelompok pelanggan premium dapat diprioritaskan melalui program loyalitas dan layanan eksklusif, sedangkan pelanggan aktif dapat diberikan promosi berkala untuk meningkatkan nilai transaksi. Sementara itu, pelanggan baru dan pelanggan potensial memerlukan pendekatan pemasaran yang berbeda, seperti pemberian voucher, diskon pembelian pertama, maupun rekomendasi produk yang lebih personal.

Secara umum hasil penelitian ini menunjukkan bahwa penerapan teknik data mining menggunakan K-Means tidak hanya mampu menyederhanakan proses segmentasi konsumen, tetapi juga menghasilkan informasi yang mendukung pengambilan keputusan bisnis secara lebih objektif. Namun demikian, metode K-Means memiliki keterbatasan karena sangat dipengaruhi oleh penentuan jumlah cluster awal serta posisi centroid awal. Oleh karena itu, penelitian selanjutnya dapat mengombinasikan K-Means dengan metode optimasi centroid, seperti Genetic Algorithm atau Particle Swarm Optimization, maupun membandingkannya dengan algoritma clustering lain seperti DBSCAN atau Agglomerative Clustering untuk memperoleh hasil segmentasi yang lebih optimal.

4. Kesimpulan

Penelitian ini membuktikan bahwa algoritma K-Means Clustering mampu menyegmentasikan konsumen berdasarkan pola transaksi melalui tahapan preprocessing, normalisasi, penentuan jumlah cluster dengan Elbow Method, serta evaluasi Silhouette Coefficient. Konfigurasi terbaik menghasilkan empat cluster dengan skor Silhouette 0,72, yaitu pelanggan baru, aktif, premium, dan potensial. Segmentasi tersebut representatif serta dapat mendukung strategi promosi, loyalitas, akuisisi, retensi, pengelolaan hubungan pelanggan, dan pengambilan keputusan bisnis. Penerapan data mining membantu perusahaan memahami perilaku konsumen secara objektif dan efisien. Penelitian selanjutnya dapat menambahkan atribut perilaku, mengoptimalkan centroid awal, serta membandingkan K-Means dengan DBSCAN, Agglomerative Clustering, atau Fuzzy C-Means untuk meningkatkan akurasi dan adaptabilitas hasil segmentasi konsumen.

Reference

Anitha, P., & Patil, M. M. (2022). RFM model for customer purchase behavior using K-Means algorithm. *Journal of King Saud University – Computer and Information Sciences, 34*(5), 1785–1792. https://doi.org/10.1016/j.jksuci.2019.12.011

DOI: https://doi.org/10.69693/ijmst.v4i2.12627

Barrera, F., Segura, M., & Maroto, C. (2024). Multiple criteria decision support system for customer segmentation using a sorting outranking method. *Expert Systems with Applications, 238*, 122310. https://doi.org/10.1016/j.eswa.2023.122310

Griva, A., Zampou, E., Stavrou, V., Papakiriakopoulos, D., & Doukidis, G. (2024). A two-stage business analytics approach to perform behavioural and geographic customer segmentation using e-commerce delivery data. *Journal of Decision Systems, 33*(1), 1–29. https://doi.org/10.1080/12460125.2022.2151071

Kasem, M. S. E., Hamada, M., & Taj-Eddin, I. (2024). Customer profiling, segmentation, and sales prediction using AI in direct marketing. *Neural Computing and Applications, 36*(9), 4995–5005. https://doi.org/10.1007/s00521-023-09339-6

Li, Y., Chu, X., Tian, D., Feng, J., & Mu, W. (2021). Customer segmentation using K-Means clustering and the adaptive particle swarm optimization algorithm. *Applied Soft Computing, 113*, 107924. https://doi.org/10.1016/j.asoc.2021.107924

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics, 20*, 53–65. https://doi.org/10.1016/0377-0427(87)90125-7

Shi, C., Wei, B., Wei, S., Wang, W., Liu, H., & Liu, J. (2021). A quantitative discriminant method of elbow point for the optimal number of clusters in clustering algorithm. *EURASIP Journal on Wireless Communications and Networking, 2021*(1), 31. https://doi.org/10.1186/s13638-021-01910-w

Yulisasih, B. N., Herman, H., Sunardi, S., & Yuliansyah, H. (2024). Evaluation of K-Means clustering using Silhouette Score method on customer segmentation. *ILKOM Jurnal Ilmiah, 16*(3), 330–342. https://doi.org/10.33096/ilkom.v16i3.2325.330-342

Putra, A. M., Rahmaddeni, R., Saputra, C., Hidayatullah, R., & Irsandi, S. (2025). Penerapan algoritma K-Means clustering dalam menganalisis tren konsumen e-commerce. *DEVICE: Journal of Information System, Computer Science and Information Technology, 6*(2).

A Framework for Customer Segmentation to Improve Marketing Strategies Using Machine Learning. (2025). *Procedia Computer Science, 260*, 616–625. https://doi.org/10.1016/j.procs.2025.03.240

Aslantaş, G., Gençgül, M., Rumelli, M., Özsaraç, M., & Bakırlı, G. (2023). Customer segmentation using K-Means clustering algorithm and RFM model. Dokuz Eylül Üniversitesi Mühendislik Fakültesi Fen ve Mühendislik Dergisi, 25(74), 491–503. https://doi.org/10.21205/deufmd.2023257418

Chen, A., Liang, Y.-C., Chang, W.-J., Siauw, H.-Y., & Minanda, V. (2022). RFM model and K-Means clustering analysis of transit traveller profiles: A case study. Journal of Advanced Transportation, 2022, 1108105. https://doi.org/10.1155/2022/1108105

Ikotun, A. M., Ezugwu, A. E., Abualigah, L., Abuhaija, B., & Heming, J. (2023). K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data. Information Sciences, 622, 178–210. https://doi.org/10.1016/j.ins.2022.11.139

Kovács, T., Ko, A., & Asemi, A. (2021). Exploration of the investment patterns of potential retail banking customers using two-stage cluster analysis. Journal of Big Data, 8, 141. https://doi.org/10.1186/s40537-021-00529-4

Kuo, R. J., Alfareza, M. N., & Nguyen, T. P. Q. (2023). Genetic based density peak possibilistic fuzzy c-means algorithms to cluster analysis—A case study on customer segmentation. Engineering Science and Technology, an International Journal, 47, 101525. https://doi.org/10.1016/j.jestch.2023.101525

Mensouri, D., Azmani, A., & Azmani, M. (2022). K-Means customers clustering by their RFMT and score satisfaction analysis. International Journal of Advanced Computer Science and Applications, 13(6), 469–476. https://doi.org/10.14569/IJACSA.2022.0130658

Okereke, G. E., Bali, M. C., Okwueze, C. N., Ukekwe, E. C., Echezona, S. C., & Ugwu, C. I. (2023). K-means clustering of electricity consumers using time-domain features from smart meter data. Journal of Electrical Systems and Information Technology, 10, 2. https://doi.org/10.1186/s43067-023-00068-3

Prastyabudi, W. A., Alifah, A. N., & Nurdin, A. (2024). Segmenting the higher education market: An analysis of admissions data using K-Means clustering. Procedia Computer Science, 234, 96–105. https://doi.org/10.1016/j.procs.2024.02.156

Salminen, J., Mustak, M., Sufyan, M., & Jansen, B. J. (2023). How can algorithms help in segmenting users and customers? A systematic review and research agenda for algorithmic customer segmentation. Journal of Marketing Analytics, 11, 677–692. https://doi.org/10.1057/s41270-023-00235-5

Spoor, J. M. (2023). Improving customer segmentation via classification of key accounts as outliers. Journal of Marketing Analytics, 11, 747–760. https://doi.org/10.1057/s41270-022-00185-4

Tabianan, K., Velu, S., & Ravi, V. (2022). K-Means clustering approach for intelligent customer segmentation using customer purchase behavior data. Sustainability, 14(12), 7243. https://doi.org/10.3390/su14127243

Wan, S., Chen, J., Qi, Z., Gan, W., & Tang, L. (2022). Fast RFM model for customer segmentation. In Companion Proceedings of the Web Conference 2022 (pp. 965–972). Association for Computing Machinery. https://doi.org/10.1145/3487553.3524707

Xiao, L., & Zhang, J. (2024). Gms-Afkmc2: A new customer segmentation framework based on the Gaussian mixture model and assumption-free K-MC2. Electronics, 13(17), 3523. https://doi.org/10.3390/electronics13173523

DOI: https://doi.org/10.69693/ijmst.v4i2.12627