

Analisis Segmentasi Kebiasaan Belanja Berdasarkan Data Penjualan Menggunakan Metode K-Means

¹ Nurun Nihayatur Rifqiyah Aulia, ² Cindy Avitaselly Bambang Saputri,
³ Erna Daniati

¹⁻³ Sistem Informasi, Universitas Nusantara PGRI Kediri

E-mail: ¹nurunniha13@gmail.com, ²cindyavitaselly@gmail.com, ³ernadaniati@unpkediri.ac.id
Penulis Korespondensi : Erna Daniati

Abstrak—Segmentasi pelanggan merupakan strategi penting dalam meningkatkan efektivitas pemasaran ritel di era data-driven. Penelitian ini bertujuan mengelompokkan pelanggan berdasarkan kebiasaan belanja menggunakan algoritma *K-Means*. Data transaksi sekunder dianalisis melalui tahapan preprocessing, normalisasi, dan visualisasi *Elbow Method* untuk menentukan jumlah klaster. Hasilnya menunjukkan empat klaster utama dengan karakteristik unik, seperti pelanggan muda dengan frekuensi tinggi, hingga pelanggan senior dengan nilai transaksi tertinggi dan preferensi pembayaran digital. Analisis mendalam mengungkap hubungan antara usia, jumlah pembelian, dan metode pembayaran terhadap nilai ekonomi tiap segmen. Kebaruan penelitian ini terletak pada integrasi data perilaku dan demografis secara simultan dalam proses segmentasi, yang jarang dioptimalkan pada studi sebelumnya. Pendekatan ini mendukung penyusunan strategi pemasaran yang lebih personal dan tepat sasaran, serta mendorong pemanfaatan data transaksi sebagai aset strategis dalam pengambilan keputusan bisnis.

Kata Kunci— Analisis data transaksi, K-Means, Segmentasi Pelanggan

Abstract— Customer segmentation is a crucial strategy for enhancing retail marketing effectiveness in the data-driven era. This study aims to classify customers based on shopping behavior using the K-Means algorithm. Secondary transaction data is analyzed through preprocessing, normalization, and Elbow Method visualization to determine the optimal number of clusters. The results reveal four main clusters with distinct characteristics, ranging from young customers with high purchase frequency to senior customers with the highest transaction values and a preference for digital payment methods. In-depth analysis uncovers the relationship between age, purchase quantity, and payment method in shaping each segment's economic value. The novelty of this research lies in the simultaneous integration of behavioral and demographic data in the segmentation process, which is rarely optimized in previous studies. This approach supports the development of more personalized and targeted marketing strategies while promoting the use of transaction data as a strategic asset in business decision-making.

Keywords— transaction data analysis, K-Means, Customer Segmentation

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



I. PENDAHULUAN

Dalam lingkungan bisnis yang kompetitif dan dinamis, segmentasi pelanggan menjadi fondasi penting untuk meningkatkan efektivitas pemasaran dan membangun kedekatan dengan konsumen. Melalui pendekatan klasterisasi seperti *K-Means*, perusahaan dapat mengidentifikasi pola perilaku belanja pelanggan dan guna merancang rencana pemasaran

yang lebih personal. Sebagaimana dijelaskan oleh [1], keberhasilan bisnis bergantung pada kemampuannya dalam memenuhi kebutuhan pelanggan dan menjaga kinerja agar tetap kompetitif di pasar.

Penelitian ini bertujuan mengisi celah dalam strategi segmentasi dengan menganalisis data transaksi penjualan ritel, serta mempertimbangkan aspek demografis dan psikografis guna merumuskan pendekatan pemasaran yang lebih tepat sasaran. Menurut [2] segmentasi berbasis perilaku pelanggan mampu meningkatkan relevansi penawaran. Triayudi dan Sholihat juga menekankan bahwa pemanfaatan *data mining* mendukung akurasi pengambilan keputusan di sektor *ritel*. Berdasarkan latar belakang yang telah dipaparkan, penelitian ini berfokus pada beberapa permasalahan utama. Salah satunya adalah bagaimana penerapan metode *K-Means* dalam segmentasi pelanggan dapat meningkatkan efektivitas strategi pemasaran. Selain itu, penelitian ini juga akan membahas bagaimana analisis data transaksi dapat membantu memahami variasi pola konsumsi yang sebelumnya kurang dioptimalkan. Tantangan lain seperti tingginya variasi produk dan persaingan digital menuntut pendekatan yang lebih akurat untuk merancang strategi pemasaran yang sesuai dengan ekspektasi pelanggan [3] menjelaskan bahwa *K-Means* dapat membantu menghasilkan pengelompokan yang lebih objektif dalam analisis data. Senada dengan itu, [4] menyebutkan bahwa *K-Means* efektif dalam mengelompokkan data yang memiliki kemiripan.

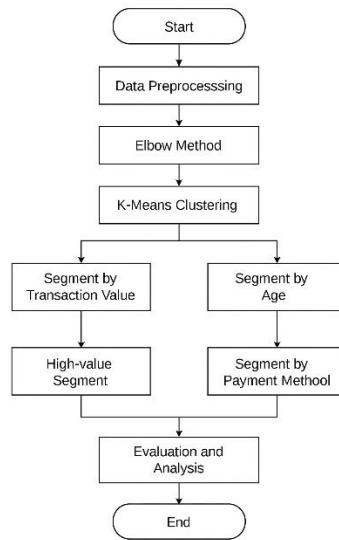
Namun, tantangan tetap ada, seperti tingginya variasi produk dan persaingan digital yang menuntut pemahaman lebih dalam terhadap pola konsumsi, serta ketidaksesuaian antara strategi promosi dan ekspektasi pelanggan [5]. Di sisi lain, variabel seperti usia, pendapatan, jenis kelamin, dan gaya hidup berpengaruh besar terhadap perilaku belanja, tetapi belum sepenuhnya dimanfaatkan karena keterbatasan integrasi data. Studi oleh [6] yang menggunakan metode *CHAID Exhaustive* menunjukkan signifikansi faktor demografis dan psikografis dalam keputusan belanja. Temuan [7] juga menggaris bawahi pentingnya strategi segmentasi, lalu *targeting*, dan *positioning* dalam memaksimalkan penjualan produk *ritel*. Sayangnya, sebagian besar riset belum mengoptimalkan data transaksi sebagai alat utama analisis perbedaan persepsi manajemen dan perilaku aktual konsumen—hal ini yang menjadi fokus penulis pada penelitian ini, sebagaimana juga ditelaah dalam studi oleh [8].

II. METODE

Penelitian ini menerapkan pendekatan kuantitatif dengan metode deskriptif yang bertujuan untuk menganalisis pola belanja pelanggan berdasarkan data transaksi penjualan. Pendekatan ini dipilih untuk menghasilkan analisis yang objektif dan berbasis data numerik, tanpa dipengaruhi bias klasifikasi awal. Dengan metode ini, pola konsumsi dapat diidentifikasi secara lebih akurat, memungkinkan segmentasi yang lebih tepat sasaran dalam strategi pemasaran [9].

Sumber data berasal dari data sekunder yang diperoleh melalui *dataset* toko *ritel* yang diperoleh dari *Kaggle*, mencakup detail seperti identitas pelanggan, tanggal dan jumlah pembelian, jenis produk yang dibeli, serta nilai transaksi. Data tersebut mencerminkan perilaku aktual pelanggan selama periode pengamatan tertentu dan digunakan sebagai dasar dalam proses segmentasi. [9]

A. Alur Pemrosesan Data Segmentasi



Gambar 1. Alur Pemrosesan

Alur Segmentasi Data Proses segmentasi data pelanggan dalam penelitian ini mengikuti beberapa tahapan utama, yaitu:

- 1) **Pengumpulan Data:** Tahap awal yang melibatkan pengumpulan data transaksi pelanggan dari sumber sekunder seperti *Kaggle*. *Dataset* ini harus mencakup informasi penting seperti ID pelanggan, tanggal transaksi, jenis produk, serta total nilai transaksi untuk memastikan analisis yang akurat.
- 2) **Pra-pemrosesan Data:** Setelah data dikumpulkan, langkah berikutnya adalah membersihkan *dataset* dari duplikasi, nilai yang hilang (*missing values*), dan data ekstrem (*outlier*) yang dapat mempengaruhi hasil analisis. Hal ini bertujuan untuk meningkatkan kualitas data sebelum masuk ke tahap analisis lebih lanjut.
- 3) **Pemilihan Variabel:** Memilih variabel yang paling relevan untuk analisis, seperti total transaksi, rata-rata nilai transaksi, dan frekuensi pembelian. Variabel ini harus mencerminkan karakteristik utama dari kebiasaan belanja konsumen selama periode pengamatan.
- 4) **Normalisasi Data:** Menyelaraskan skala variabel untuk menghindari dominasi fitur tertentu dalam proses klasterisasi. Normalisasi penting untuk memastikan setiap variabel memiliki bobot yang seimbang dalam menentukan hasil kluster.
- 5) **Penentuan Jumlah Klaster:** Menggunakan *Elbow Method* untuk menemukan jumlah klaster optimal. Metode ini melibatkan perhitungan nilai *Within-Cluster Sum of Squares* (WCSS) untuk berbagai jumlah klaster, dengan titik siku pada grafik WCSS digunakan sebagai indikator jumlah klaster terbaik.
- 6) **Segmentasi Berdasarkan Nilai Transaksi:** Mengelompokkan pelanggan berdasarkan total nilai transaksi mereka, yang dapat menunjukkan perbedaan dalam kontribusi pendapatan masing-masing segmen.
- 7) **Segmentasi Berdasarkan Usia:** Menganalisis perbedaan perilaku pelanggan berdasarkan kelompok usia untuk memahami preferensi yang mungkin bervariasi antar kelompok demografis.

- 8) **Segmentasi Berdasarkan Metode Pembayaran:** Mengidentifikasi preferensi metode pembayaran yang dominan di setiap klaster, seperti kartu kredit, *PayPal*, atau uang tunai, untuk menentukan strategi pembayaran yang lebih tepat.
- 9) **Interpretasi Hasil Klasterisasi:** Menginterpretasikan karakteristik masing-masing klaster untuk pengembangan strategi pemasaran yang lebih tepat sasaran. Hasil ini akan menjadi dasar untuk merancang strategi yang lebih personal dan relevan dengan setiap kelompok pelanggan.

B. Pra-pemrosesan Data

Tahapan awal yang dilakukan adalah pembersihan data, yang mencakup penghapusan data duplikat, nilai yang hilang (*missing values*), dan data ekstrem yang berpotensi mengganggu hasil analisis. Setelah itu, dipilih variabel-variabel yang dianggap relevan dan representatif terhadap kebiasaan belanja konsumen, seperti jumlah total transaksi, rata-rata nominal belanja per transaksi, dan frekuensi pembelian selama periode tertentu. Variabel-variabel ini kemudian dinormalisasi agar memiliki skala yang setara, sehingga tidak terjadi dominasi antar fitur dalam proses klasterisasi. [6]

Segmentasi pelanggan dilaksanakan dengan menerapkan algoritma *K-Means Clustering*. Untuk menghitung jumlah klaster yang maksimal, digunakan pendekatan *Elbow Method*, dengan memvisualisasikan nilai *Within-Cluster Sum of Squares* (WCSS) terhadap berbagai nilai klaster, dan memilih titik siku grafik sebagai jumlah klaster yang paling representatif.[10] [11][12]

Setelah jumlah klaster optimal diperoleh, proses pengelompokan data pelanggan ke dalam masing-masing segmen dilakukan menggunakan algoritma *K-Means*. Hasil dari proses ini berupa beberapa kelompok pelanggan yang memiliki ciri-ciri perilaku belanja yang serupa dalam hal frekuensi, volume, dan jenis pembelanjaan. Klaster-klaster tersebut kemudian dianalisis lebih lanjut untuk memberikan interpretasi deskriptif terhadap profil masing-masing kelompok. [2]

C. Alat dan Perangkat Lunak

Tahapan analisis dan pemrosesan data dilaksanakan dengan bantuan perangkat lunak *Python*, dengan memanfaatkan pustaka seperti *pandas* untuk manajemen data, *scikit-learn* untuk proses klasterisasi, serta *matplotlib* dan *seaborn* untuk visualisasi data dan hasil klasterisasi. Metode ini diharapkan mampu menyajikan wawasan yang bermanfaat dalam proses perumusan strategi pemasaran yang lebih tepat sasaran. Proses utama dalam algoritma ini melibatkan penghitungan jarak antara setiap data dengan *centroid*. Jarak yang digunakan adalah jarak *Euclidean*, yang dirumuskan sebagai berikut:

$$d(x_i, c_j) = \sqrt{\sum_{k=1}^n (x_{ik} - c_{jk})^2} \quad (1)$$

Notasi $d(x_i, c_j)$ digunakan untuk menggambarkan jarak antara data ke- i dengan pusat klaster (*centroid*) ke- j . Sementara itu, x_{ik} menunjukkan nilai dari atribut ke- k pada data ke- i , dan c_{jk} merepresentasikan nilai atribut ke- k pada centroid ke- j . Adapun superskrip n menunjukkan jumlah atribut yang digunakan dalam proses klasterisasi tersebut.

Tabel 1. Tabel Klasterisasi Dataset

Tabel Dataset			
No.	Nama Kolom	Jumlah	Keterangan
1	Customer ID	3900	ID Pelanggan
2	Purchase Amount (USD)	81	Jumlah pembelian (dollar)
3	Age	53	Usia pelanggan
4	Previous Purchases	50	Jumlah pembelian sebelumnya
5	Location	50	Lokasi pelanggan
6	Review Rating	26	Rating yang diberikan pelanggan
7	Item Purchased	25	Nama barang yang dibeli
8	Payment Method	7	Metode pembayaran

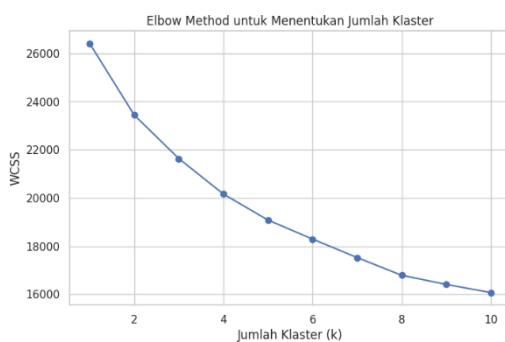
III. HASIL DAN PEMBAHASAN

A. Penentuan Jumlah Klaster

Penentuan jumlah klaster dalam algoritma *K-Means* dilakukan dengan pendekatan *Elbow Method*, yang memanfaatkan nilai *Within-Cluster Sum of Squares* (WCSS) untuk berbagai nilai k. Nilai WCSS dihitung dengan menjumlahkan jarak kuadrat *Euclidean* antara setiap titik data x_i dan *centroid* klaster c_j , yang dirumuskan sebagai berikut:

$$d(x_i, c_j) = \sqrt{\sum_{k=1}^n (x_{ik} - c_{jk})^2} \quad (2)$$

Rumus ini mengukur jarak antara titik data dengan pusat klaster berdasarkan seluruh atribut yang digunakan. Semakin kecil nilai WCSS, semakin baik representasi data oleh *centroid* klaster. Berdasarkan grafik *Elbow*, penurunan WCSS terlihat tajam hingga k = 4, kemudian mulai melandai, sehingga k = 4 dipilih sebagai jumlah klaster yang representatif.

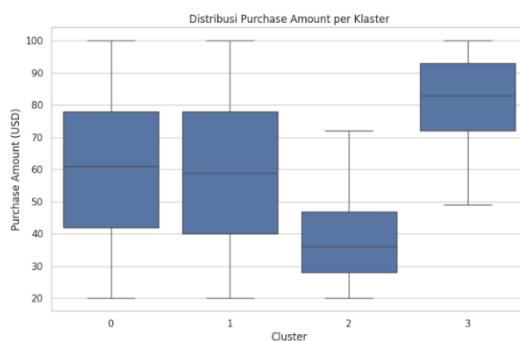


Gambar 2. Elbow Method

Langkah awal dalam proses segmentasi pelanggan dengan algoritma *K-Means* adalah menentukan jumlah klaster yang sesuai. Dalam penelitian ini digunakan pendekatan *Elbow Method*, yang mengevaluasi nilai *Within-Cluster Sum of Squares* (WCSS) dari k = 1 hingga

$k = 10$. Nilai WCSS dihitung berdasarkan rumus jarak *Euclidean* sebagaimana ditunjukkan pada Persamaan (1), yang mencerminkan total jarak kuadrat antara setiap titik data dengan *centroid* klaster. Berdasarkan Gambar 2, terlihat bahwa nilai WCSS mengalami penurunan tajam hingga $k = 4$, kemudian mulai melandai. Titik perubahan ini dikenal sebagai “*elbow*” dan dipilih sebagai jumlah klaster yang paling representatif. Temuan ini sejalan dengan studi oleh [13], yang menyatakan bahwa metode *Elbow* efektif dalam mengidentifikasi jumlah klaster yang tepat dalam konteks segmentasi pelanggan ritel daring.

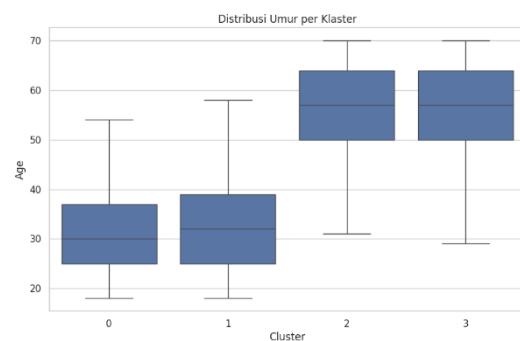
B. Karakteristik Klaster Berdasarkan Nilai Transaksi



Gambar 3. Distribusi Jumlah Pembayaran Per Klaster

Gambar 2 menampilkan distribusi nilai transaksi per klaster. Dapat diamati bahwa Klaster 3 menunjukkan nilai transaksi tertinggi, dengan median hampir USD 90, serta interkuartil yang sempit, menandakan bahwa pelanggan dalam klaster ini memiliki pola pembelian yang stabil dan bernilai tinggi. Di sisi lain, Klaster 2 memiliki penyebaran nilai transaksi yang luas dan median yang rendah, mencerminkan adanya variasi signifikan dalam kebiasaan belanja. Temuan ini selaras dengan penelitian oleh [14], yang menyimpulkan bahwa pelanggan dengan nilai transaksi tinggi cenderung menjadi target prioritas dalam strategi loyalitas dan retensi pelanggan.

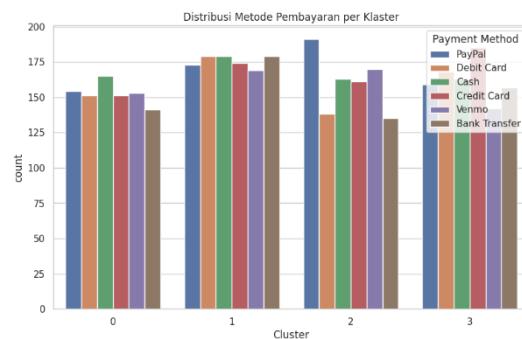
C. Segmentasi Berdasarkan Usia



Gambar 4. Distribusi Umum Per Klaster

Analisis segmentasi berdasarkan usia (Gambar 3) menunjukkan bahwa Klaster 2 dan 3 sebagian besar terdiri dari pelanggan berusia 50 tahun ke atas. Sebaliknya, Klaster 0 dan 1 didominasi oleh kelompok usia yang lebih muda (25–40 tahun). Hal ini mengindikasikan adanya hubungan antara usia dengan pola konsumsi, di mana konsumen usia lanjut cenderung memiliki daya beli lebih tinggi dan stabil. Hal ini sesuai dengan temuan [15] yang dalam studi kasus Superindo Tambun menyatakan bahwa usia menjadi variabel penting dalam memahami perilaku dan kebutuhan konsumen.

D. Preferensi Metode Pembayaran



Gambar 5. Distribusi Metode Pembayaran Per Klaster

Preferensi metode pembayaran antar klaster juga menunjukkan variasi signifikan. Klaster 3 yang didominasi oleh pelanggan dengan transaksi tinggi lebih cenderung menggunakan metode digital seperti kartu kredit dan *PayPal*. Sebaliknya, klaster lain menunjukkan keberagaman dalam metode pembayaran termasuk penggunaan uang tunai dan transfer bank. Studi oleh [16] menegaskan bahwa pemahaman terhadap metode pembayaran dominan di tiap segmen dapat membantu pengembangan sistem pembayaran yang lebih adaptif dan inklusif.

E. Interpretasi Hasil

Rata-rata Fitur Numerik per Klaster:					
Cluster	Purchase Amount (USD)	Age	Previous Purchases	Review Rating	
0	60.392350	31.184699	38.945355	3.811366	
1	58.956315	32.360874	11.349478	3.732289	
2	37.863257	56.541754	26.981211	3.507516	
3	81.589322	56.560575	26.116016	3.949795	

Distribusi Payment Method per Klaster:						
Cluster	Bank Transfer	Cash	Credit Card	Debit Card	PayPal	Venmo
0	141	165	151	151	154	153
1	179	179	174	179	173	169
2	135	163	161	138	191	170
3	157	163	185	168	159	142

Gambar 6. Interpretasi Hasil

Interpretasi Hasil Klasterisasi Berdasarkan hasil analisis klustering, terdapat empat klaster utama dengan karakteristik sebagai berikut:

- 1) **Klaster 0 (Pembelanja Muda dengan Pembelian Tinggi):** Rata-rata usia pelanggan adalah 31 tahun, dengan jumlah pembelian rata-rata 39 transaksi dan nilai transaksi rata-rata USD 60,39. Mereka cenderung menggunakan berbagai metode pembayaran dengan distribusi yang seimbang.
- 2) **Klaster 1 (Pembeli Kasual Muda):** Pelanggan di klaster ini memiliki usia rata-rata 32 tahun, namun dengan jumlah pembelian yang lebih sedikit (11 transaksi) dan nilai transaksi yang relatif rendah (USD 58,95). Metode pembayaran bervariasi, namun dominan pada *debit card* dan *PayPal*.
- 3) **Klaster 2 (Pelanggan Loyal dan Berusia Matang):** Klaster ini memiliki usia rata-rata 56 tahun, dengan jumlah pembelian 27 transaksi dan nilai transaksi lebih rendah (USD 37,86), namun mereka cenderung lebih loyal dengan frekuensi pembelian yang stabil.
- 4) **Klaster 3 (Pembelanja Premium Senior):** Pelanggan di klaster ini adalah yang tertua dengan rata-rata usia 56 tahun, namun dengan nilai transaksi tertinggi (USD 81,58). Mereka sering menggunakan kartu kredit untuk pembayaran dan cenderung memberikan ulasan yang lebih positif (3,95 dari 5).

F. Pembahasan

Secara umum, algoritma K-Means terbukti efektif dalam membedakan karakteristik pelanggan berdasarkan variabel seperti nilai pembelian, usia, dan metode pembayaran. Hasil ini konsisten dengan penelitian sebelumnya [13], yang mengungkapkan bahwa metode klasterisasi seperti K-Means sangat berguna untuk mengungkap pola perilaku konsumen secara kuantitatif serta membantu perusahaan dalam menyusun strategi pemasaran yang lebih tepat sasaran. Dengan membagi pelanggan ke dalam kelompok-kelompok yang homogen, perusahaan dapat mengoptimalkan alokasi sumber daya dan menyesuaikan produk atau layanan sesuai kebutuhan tiap segmen. Meski demikian, metode K-Means memiliki beberapa keterbatasan penting, terutama dalam hal pengolahan data temporal. Algoritma ini kurang mampu menangkap perubahan pola pembelian yang bersifat musiman atau dinamika perilaku konsumen dari waktu ke waktu. Padahal, dalam dunia nyata, preferensi dan kebiasaan konsumen sering berubah dipengaruhi oleh tren, promosi, atau kondisi ekonomi. Oleh sebab itu, jika hanya mengandalkan *K-Means*, hasil segmentasi bisa saja kurang mencerminkan perubahan yang sebenarnya terjadi. Untuk mengatasi keterbatasan ini, integrasi dengan metode lain yang mampu mengolah data berbasis waktu dan pola perilaku yang berubah-ubah menjadi solusi yang tepat. Misalnya, *Time Series Clustering* dapat digunakan untuk mengelompokkan pelanggan berdasarkan pola pembelian mereka yang berubah secara temporal, sehingga perusahaan bisa lebih responsif dalam menyesuaikan strategi pemasaran. Selain itu, algoritma berbasis kepadatan seperti DBSCAN juga memberikan kelebihan dengan kemampuannya mengenali klaster dengan bentuk tidak teratur serta mendeteksi data yang tidak konsisten atau *outlier*, yang sering muncul dalam data perilaku konsumen yang kompleks. Dengan menggabungkan metode-metode tersebut, analisis segmentasi akan menjadi lebih komprehensif dan adaptif, memberikan gambaran yang lebih akurat tentang karakteristik dan perubahan perilaku pelanggan. Pendekatan ini sangat direkomendasikan untuk bisnis dengan volume transaksi yang besar dan pola konsumsi yang beragam agar dapat mengikuti perkembangan perilaku konsumen secara lebih efektif..

IV. KESIMPULAN

Penerapan algoritma *K-Means* dalam segmentasi pelanggan terbukti efektif untuk mengidentifikasi perbedaan perilaku konsumen dan mendukung strategi pemasaran yang lebih tepat sasaran. Sebagai kontribusi terhadap bidang teknik industri, pendekatan ini memperkuat penerapan analitik dalam pengambilan keputusan berbasis data. Disarankan agar penelitian selanjutnya mengintegrasikan variabel temporal atau perilaku historis untuk menghasilkan segmentasi yang lebih dinamis dan adaptif terhadap perubahan perilaku pelanggan.

DAFTAR PUSTAKA

- [1] A. P. Nugroho, D. Karunia, and S. Suroto, “SEGMENTASI PASAR TERHADAP PENINGKATAN PENJUALAN KEJU MOZZARELLA DI KOTA MALANG,” 2023.
- [2] P. Indra Pangestu, T. Iman Hermanto, and D. Irmayanti, “ANALISIS SEGMENTASI PELANGGAN BERBASIS RECENCY FREQUENCY MONETARY (RFM) MENGGUNAKAN ALGORITMA K-MEANS,” *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 7, no. 3, pp. 1486–1492, Oct. 2023, doi: 10.36040/jati.v7i3.7171.
- [3] E. Daniati and H. Utama, *2019 4th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE)*. IEEE, 2019. doi: 10.1109/ICITISEE48480.2019.9003858.
- [4] E. Daniati and A. Nugroho, *Proceedings, 6th IEEE International Conference on Control System, Computing and Engineering (ICCSCE 2016) : Parkroyal Penang Resort, Batu Ferringhi, Penang, Malaysia, 25th-27th Nov 2016*. IEEE, 2016. doi: 10.1109/ICCSCE.2016.7893593.
- [5] A. Setiono, A. Triayudi, and E. T. Esti Handayani, “ANALISIS RECENCY FREQUENCY MONETARY DAN K-MEANS CLUSTERING PADA KLINIK GIGI UNTUK MENENTUKAN SEGMENTASI PASIEN,” *JSiI (Jurnal Sistem Informasi)*, vol. 10, no. 1, pp. 1–6, Mar. 2023, doi: 10.30656/jsii.v10i1.5999.
- [6] A. Febriani and S. A. Putri, “Segmentasi Konsumen Berdasarkan Model Recency, Frequency, Monetary dengan Metode K-Means,” *JIEMS (Journal of Industrial Engineering and Management Systems)*, vol. 13, no. 2, Sep. 2020, doi: 10.30813/jiems.v13i2.2274.
- [7] Y. Yuliana and A. Ardansyah, “Analisis Strategi Pemasaran Untuk Meningkatkan Penjualan Minuman Kedai Kopi Square di Kota Bandar Lampung,” *SINOMIKA Journal: Publikasi Ilmiah Bidang Ekonomi dan Akuntansi*, vol. 1, no. 3, pp. 309–320, Aug. 2022, doi: 10.54443/sinomika.v1i3.265.
- [8] N. R. Noperahila and F. Sulianta, “Analisis Pola Perilaku Konsumen dan Kebiasaan Belanja menggunakan Metode K-Means Clustering,” 2024. doi: 10.30813/jiems.v13i2.2274.
- [9] A. Pramudiansyah and H. Munte, “SEGMENTASI PELANGGAN MENGGUNAKAN ALGORITMA K-MEANS BERDASARKAN MODEL RECENCY FREQUENCY MONETARY,” vol. 7, no. 2, 2021, [Online]. Available: <http://ejournal.fikom-unasman.ac.id>

- [10] Andy Hermawan, Fachmi Aditama, Lintang Rizki Ramadhani, Nuur Muhammad Ilham, Aji Saputra, and Nila Rusiardi Jayanti, "Analisis Segmentasi Pelanggan Berbasis RFM dan Evaluasi Efektivitas Kampanye Pemasaran untuk Meningkatkan Retensi," *Neptunus: Jurnal Ilmu Komputer Dan Teknologi Informasi*, vol. 2, no. 4, pp. 28–42, Oct. 2024, doi: 10.61132/neptunus.v2i4.400.
- [11] A. Hermawan *et al.*, "Optimalisasi Strategi Pemasaran Melalui Analisis RFM pada Dataset Transaksi Ritel Menggunakan Python," *Jurnal Manajemen Riset Inovasi*, vol. 2, no. 4, pp. 254–267, doi: 10.55606/mri.v2i4.3246.
- [12] N. H. Harani, C. Prianto, and F. A. Nugraha, "Segmentasi Pelanggan Produk Digital Service Indihome Menggunakan Algoritma K-Means Berbasis Python," *Jurnal Manajemen Informatika (JAMIKA)*, 2020, doi: 10.34010/jamika.v10i2.
- [13] E. Febrianty, L. Awalina, and W. I. Rahayu, "Optimalisasi Strategi Pemasaran dengan Segmentasi Pelanggan Menggunakan Penerapan K-Means Clustering pada Transaksi Online Retail Optimizing Marketing Strategies with Customer Segmentation Using K-Means Clustering on Online Retail Transactions," *Jurnal Teknologi dan Informasi (JATI)*, vol. 13, 2023, doi: 10.34010/jati.v13i2.
- [14] N. Puspitasari, J. A. Widians, and N. B. Setiawan, "Customer segmentation using bisecting k-means algorithm based on recency, frequency, and monetary (RFM) model," *Jurnal Teknologi dan Sistem Komputer*, vol. 8, no. 2, pp. 78–83, Apr. 2020, doi: 10.14710/jtsiskom.8.2.2020.78-83.
- [15] R. G. Praditya, G. Sembodo, and J. Heikal, "Market segmentation analysis to find out products and services that suit customer needs using the python KMEANS clustering method (Case study: Superindo Tambun Area, Bekasi)," *Jurnal Teknik Industri Terintegrasi*, vol. 7, no. 4, pp. 2072–2081, Oct. 2024, doi: 10.31004/jutin.v7i4.35889.
- [16] P. H. Suharti, A. S. Suryandari, and R. N. Amalia, "ANALISIS KINERJA MODUL PENGENDALI TEKANAN UDARA PCT-14 BERBASIS PLC DENGAN BERBAGAI METODA TUNING," *Sebatik*, vol. 26, no. 2, pp. 420–427, Dec. 2022, doi: 10.46984/sebatik.v26i2.2134.