



## Pengembangan Aplikasi Cerdas Berbasis AI untuk Analisis Tren Penjualan Produk Fashion Lokal Menggunakan Algoritma Data Mining

Irdha Yuniarto<sup>1\*</sup>, Wiwid Wahyudi<sup>2</sup>, Novita Indriyani<sup>3</sup>, Muhamad Darusyifa F<sup>4</sup>

<sup>1-4</sup> Universitas Sains dan Teknologi Komputer

Email: [stekom.ian@gmail.com](mailto:stekom.ian@gmail.com)<sup>1\*</sup>

*\*Penulis Korespondensi*

**Abstract.** *This study develops an AI-driven application for analyzing sales of local fashion products and mapping customer heterogeneity to support marketing decision-making for micro, small, and medium enterprises (MSMEs). A mixed-methods approach is employed, combining a structured literature review, surveys/interviews, and Focus Group Discussions (FGDs) to validate findings and system usability. Quantitatively, first-quarter 2025 transaction data (100 respondents) are analyzed using K-Means on three standardized features age, aggregated number of items purchased, and aggregated spending. Cluster evaluation with the silhouette score for  $k=2-5$  indicates the best separation at  $k=5$ , yielding a stable and interpretable segmentation. The resulting profiles reveal at least one high-value segment (larger baskets and higher spending) suitable for tiered loyalty programs and premium bundling; a mid-value segment responsive to targeted cross-sell/upsell offers; and a low-intensity segment that benefits from staged onboarding interventions to improve retention. These insights are integrated into a prototype analytics application that presents a segmentation dashboard and key performance indicators, providing actionable support for MSMEs' marketing, catalog curation, and inventory allocation.*

**Keywords:** *local fashion MSMEs; customer segmentation; K-Means; sales analytics; intelligent application*

**Abstrak.** Penelitian ini mengembangkan aplikasi cerdas untuk analitik penjualan fashion lokal dan memetakan heterogenitas pelanggan guna mendukung keputusan pemasaran UMKM. Pendekatan mixed methods digunakan melalui studi literatur, survei/wawancara, serta FGD untuk memvalidasi temuan dan kegunaan sistem. Secara kuantitatif, data transaksi triwulan I 2025 (100 responden) dianalisis dengan K-Means pada tiga fitur terstandarisasi usia, jumlah item dibeli (agregat), dan total belanja (agregat). Evaluasi jumlah kluster menggunakan silhouette score untuk  $k=2-5$  menunjukkan kualitas pemisahan terbaik pada  $k=5$ , menghasilkan segmentasi yang stabil dan mudah diinterpretasikan. Profil segmen mengindikasikan setidaknya satu kelompok bernilai tinggi (keranjang besar dan nilai belanja tinggi) sebagai sasaran program loyalti dan bundling premium; segmen menengah yang responsif terhadap cross-sell/upsell; serta segmen berintensitas rendah yang membutuhkan intervensi onboarding bertahap untuk meningkatkan retensi.

**Kata kunci:** UMKM fashion; segmentasi pelanggan; K-Means; analitik penjualan; aplikasi cerdas.

### 1. LATAR BELAKANG

Industri fashion lokal merupakan salah satu sektor strategis dalam ekonomi kreatif Indonesia karena memiliki kontribusi besar terhadap Produk Domestik Bruto (PDB) sekaligus menyerap tenaga kerja dalam jumlah signifikan. Di era transformasi digital, sektor ini menghadapi tantangan yang semakin kompleks, khususnya bagi UMKM yang menjadi tulang punggung industri kreatif. Keterbatasan dalam mengelola dan menganalisis data penjualan membuat banyak UMKM belum mampu memahami tren pasar dan perilaku konsumen secara mendalam. Hal ini berdampak pada pengambilan keputusan yang kurang optimal, terutama dalam penentuan desain produk, strategi pemasaran, serta pengelolaan stok barang. Masalah utama yang dihadapi adalah keterbatasan pemanfaatan data historis secara efektif untuk

menghasilkan insight bisnis yang dapat meningkatkan daya saing (Wang & Gao, 2021; Datta, 2024; Chauhan et al., 2022).

Seiring dengan berkembangnya teknologi, kemajuan Artificial Intelligence (AI) dan data mining telah mendorong implementasi sistem cerdas untuk menganalisis data dalam skala besar maupun kompleks. Algoritma seperti Apriori dan association rule mining terbukti efektif dalam mendukung proses pengambilan keputusan di sektor ritel dan e-commerce karena mampu mengidentifikasi pola asosiasi antar produk berdasarkan data historis transaksi (Wang & Gao, 2021). Beberapa penelitian bahkan menunjukkan bahwa pemanfaatan AI dapat meningkatkan efisiensi dalam berbagai bidang, termasuk sektor konstruksi (Datta, 2024), penerapan model ekonomi sirkular yang didukung digitalisasi (Chauhan et al., 2022), serta perbaikan tata kelola administrasi publik melalui proses otomatisasi yang lebih transparan (Pini et al., 2025).

AI dan data mining juga telah diterapkan dalam sektor manufaktur dan energi. Szramowiat-Sala et al. (2024) menunjukkan bahwa pendekatan berbasis data mining mampu membantu pengendalian dampak lingkungan dari teknologi energi konvensional, sementara Liu (2024) mengembangkan metode manajemen biaya berbasis AI untuk meningkatkan akurasi estimasi dalam proyek rekayasa. Penelitian-penelitian ini menegaskan bahwa AI berperan penting dalam meningkatkan efisiensi dan efektivitas proses bisnis lintas sektor.

Lebih spesifik pada industri fashion, perkembangan teknologi Industry 4.0 yang mencakup big data analytics, AI, dan Internet of Things (IoT) telah mempercepat transformasi digital di sektor ini, khususnya pasca-pandemi COVID-19. Bottani et al. (2025) menyoroti perbedaan signifikan penerapan teknologi digital di sektor fashion sebelum dan sesudah pandemi, di mana perusahaan yang mengadopsi teknologi digital lebih cepat beradaptasi terhadap perubahan pasar. Yu dan Zhao (2024) menambahkan bahwa integrasi AI dalam desain pakaian mampu mendukung proses inovasi yang lebih cepat dan responsif terhadap kebutuhan konsumen. Sementara itu, Bianco et al. (2023) menekankan peran Industry 4.0 dalam meningkatkan resiliensi perusahaan manufaktur agar tetap kompetitif di tengah gangguan rantai pasok global.

Meskipun berbagai inovasi teknologi telah diadopsi oleh perusahaan besar, penerapannya di kalangan UMKM fashion lokal masih relatif rendah. Keterbatasan sumber daya manusia, infrastruktur digital, dan literasi teknologi menjadi hambatan utama dalam pemanfaatan data penjualan sebagai dasar pengambilan keputusan bisnis. Padahal, data transaksi dan interaksi konsumen berpotensi memberikan informasi berharga dalam

memahami preferensi pelanggan. Penelitian Mahdikhani (2023) menegaskan bahwa ulasan daring dan data perilaku konsumen dapat dimanfaatkan untuk memprediksi tren serta mendukung strategi pemasaran yang lebih tepat sasaran. Namun, mayoritas UMKM fashion belum mampu mengintegrasikan teknologi AI dan data mining dalam sistem operasionalnya secara optimal.

Urgensi penelitian ini terletak pada upayanya untuk menjembatani kesenjangan adopsi teknologi di kalangan UMKM fashion lokal. Pengembangan aplikasi cerdas berbasis AI yang memanfaatkan algoritma data mining seperti Apriori dan K-Means diharapkan dapat membantu UMKM memahami tren penjualan, menyusun strategi pemasaran berbasis data, serta mengelola stok secara lebih efisien. Penelitian ini tidak hanya memberikan kontribusi teoretis dalam pengembangan model analisis data untuk sektor kreatif, tetapi juga memiliki dampak praktis dalam mendukung percepatan transformasi digital UMKM sesuai dengan agenda pemerintah terkait implementasi teknologi Industry 4.0 dan Society 5.0 pada sektor ekonomi kreatif (Pini et al., 2025; Bottani et al., 2025).

## **2. KAJIAN TEORITIS**

Penelitian mengenai penerapan algoritma Apriori dalam analisis data penjualan telah banyak dilakukan pada sektor ritel dan e-commerce. Algoritma ini populer karena kesederhanaannya dalam mengekstraksi aturan asosiasi yang dapat dimanfaatkan untuk analisis keranjang belanja dan rekomendasi produk. Wang dan Gao (2021) menekankan bahwa paralelisasi algoritma Apriori mampu meningkatkan efisiensi pengolahan data transaksi dalam jumlah besar, sementara Nguyen et al. (2019) menunjukkan bahwa Apriori dapat digunakan untuk mendukung pengambilan keputusan strategis dalam ritel dengan mengidentifikasi hubungan antar produk. Dalam konteks yang lebih luas, penggunaan algoritma asosiasi juga terbukti membantu peritel dalam merancang promosi, bundling produk, dan strategi cross-selling (Bedi et al., 2022).

Selain Apriori, algoritma K-Means Clustering banyak dimanfaatkan dalam segmentasi pelanggan. Yuan dan Yang (2019) menjelaskan bahwa K-Means dapat membagi konsumen berdasarkan perilaku pembelian, sehingga memungkinkan perusahaan untuk mengembangkan strategi pemasaran yang lebih tepat sasaran. Li et al. (2023) menguatkan temuan ini dengan menunjukkan efektivitas K-Means dalam mengidentifikasi segmen konsumen pada sektor ritel sehingga perusahaan dapat menyusun program promosi sesuai kebutuhan masing-masing

segmen. Dalam praktiknya, K-Means sering digunakan bersamaan dengan algoritma asosiasi, di mana clustering berfungsi sebagai tahap awal untuk mengelompokkan konsumen sebelum dilakukan analisis asosiasi antar produk dalam setiap kelompok.

Penerapan teknologi AI dan data mining dalam industri fashion juga mulai berkembang seiring dengan transformasi digital. Bottani et al. (2025) menyatakan bahwa adopsi teknologi Industry 4.0, seperti big data analytics dan AI, membuat sektor fashion lebih adaptif terhadap perubahan pasar pasca-pandemi COVID-19. Penelitian oleh Yu dan Zhao (2024) menunjukkan bahwa sistem desain pakaian berbasis AI dapat meningkatkan efisiensi proses kreatif dan menghasilkan produk yang lebih sesuai dengan preferensi konsumen. Sementara itu, Bianco et al. (2023) menyoroti bahwa adopsi teknologi digital mendukung ketahanan industri manufaktur, termasuk fashion, di tengah tantangan rantai pasok global.

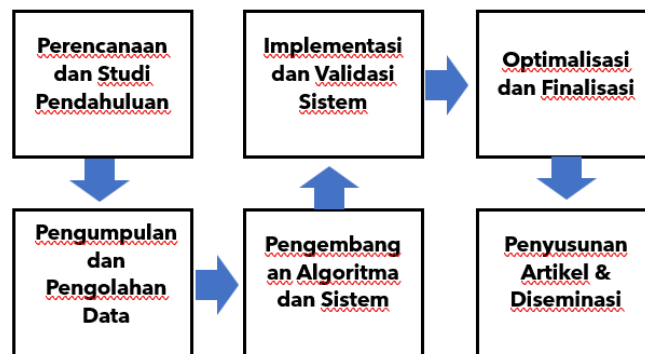
Meskipun berbagai penelitian telah menunjukkan manfaat penerapan algoritma Apriori, K-Means, dan AI di sektor ritel, adopsinya di kalangan UMKM fashion lokal masih sangat terbatas. Menurut Mahdikhani (2023), data dari interaksi konsumen seperti ulasan daring dapat memberikan informasi penting untuk memahami tren dan preferensi pasar, tetapi pemanfaatannya belum optimal di tingkat UMKM. Hal ini diperkuat oleh Rahayu dan Day (2022) yang menemukan bahwa keterbatasan literasi digital dan pemahaman teknologi menjadi faktor penghambat utama adopsi teknologi e-commerce dan analitik data oleh UMKM di Indonesia.

Dari tinjauan literatur, terlihat bahwa penelitian mengenai integrasi algoritma Apriori dan K-Means dalam konteks fashion lokal masih terbatas. Sebagian besar studi terdahulu lebih menekankan pada skala perusahaan besar atau ritel modern, sementara UMKM fashion lokal yang memiliki keterbatasan sumber daya belum banyak menjadi fokus. Oleh karena itu, penelitian ini menghadirkan kontribusi baru dengan mengembangkan aplikasi cerdas berbasis AI yang memanfaatkan kombinasi algoritma Apriori untuk menemukan pola asosiasi antar produk dan K-Means untuk melakukan segmentasi konsumen. Integrasi kedua pendekatan ini diharapkan mampu membantu UMKM fashion lokal dalam memahami tren penjualan, menyusun strategi pemasaran yang efektif, dan meningkatkan daya saing di pasar digital.

### 3. METODE PENELITIAN

Penelitian ini menggunakan pendekatan Mixed Methods, yaitu kombinasi metode kuantitatif dan kualitatif untuk memperoleh pemahaman yang lebih komprehensif mengenai strategi pengembangan aplikasi cerdas berbasis AI dalam menganalisis tren penjualan produk fashion lokal. Pendekatan ini dianggap tepat karena memungkinkan integrasi data numerik dengan wawasan kontekstual, sehingga hasil penelitian tidak hanya menggambarkan pola statistik tetapi juga memberikan interpretasi yang lebih mendalam (Creswell & Plano Clark, 2017).

Tahapan penelitian yang dilakukan ditunjukkan pada Gambar 1, yang meliputi enam fase utama, yaitu perencanaan dan studi pendahuluan, pengumpulan dan pengolahan data, pengembangan algoritma dan sistem, implementasi dan validasi sistem, optimalisasi dan finalisasi, serta penyusunan artikel dan diseminasi hasil penelitian.



**Gambar 1.** Diagram Penelitian

Pada tahap awal, dilakukan studi literatur untuk mengidentifikasi teori, konsep, dan penelitian terdahulu terkait penerapan AI dan data mining pada sektor ritel dan fashion. Studi literatur ini juga berfungsi sebagai dasar dalam menyusun kerangka konseptual penelitian (Snyder, 2019). Selanjutnya, data primer diperoleh melalui survei dan wawancara dengan pelaku UMKM fashion lokal yang mencakup informasi mengenai pola penjualan, preferensi konsumen, serta tantangan dalam pengelolaan data. Data ini dilengkapi dengan data sekunder berupa laporan industri dan publikasi resmi terkait perkembangan e-commerce di Indonesia (Badan Pusat Statistik, 2023).

Setelah data terkumpul, dilakukan pengolahan data kuantitatif melalui teknik data preprocessing untuk memastikan kualitas dataset, seperti normalisasi, pembersihan data, dan transformasi variabel. Data yang sudah bersih kemudian dianalisis menggunakan algoritma K-

Means Clustering untuk mengelompokkan konsumen berdasarkan karakteristik perilaku pembelian (Han et al., 2021; Li et al., 2023).

Selain itu, dilakukan analisis kualitatif melalui Focus Group Discussion (FGD) dengan para pelaku UMKM, akademisi, dan praktisi industri kreatif. Tujuannya adalah untuk memvalidasi hasil analisis kuantitatif dan memastikan bahwa rekomendasi yang dihasilkan dapat diterapkan secara praktis di lapangan (Guest et al., 2017).

Tahap berikutnya adalah implementasi dan validasi sistem, di mana algoritma yang dikembangkan diintegrasikan ke dalam prototipe aplikasi cerdas. Validasi dilakukan melalui uji coba terbatas bersama UMKM mitra untuk mengukur kinerja sistem berdasarkan metrik evaluasi seperti akurasi aturan asosiasi, kualitas klaster, serta kegunaan aplikasi. Setelah itu dilakukan optimalisasi dan finalisasi sistem berdasarkan masukan dari pengguna dan hasil evaluasi, sehingga sistem dapat memberikan rekomendasi yang lebih tepat sasaran.

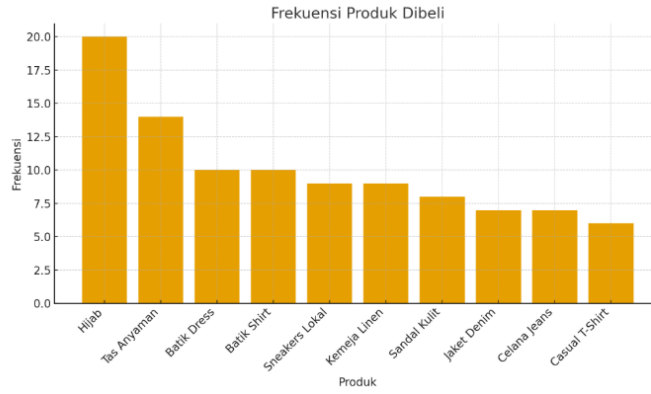
Tahap terakhir adalah penyusunan artikel dan diseminasi, di mana hasil penelitian dituliskan dalam bentuk publikasi ilmiah yang ditargetkan pada jurnal bereputasi internasional. Dengan metode ini, penelitian diharapkan tidak hanya memberikan kontribusi akademis dalam pengembangan model analisis berbasis AI, tetapi juga kontribusi praktis dalam mendukung UMKM fashion lokal agar mampu bersaing di era ekonomi digital.

#### **4. HASIL DAN PEMBAHASAN**

##### **Hasil**

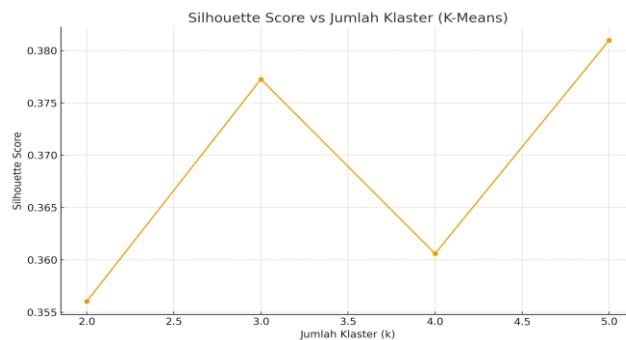
Analisis deskriptif awal menunjukkan struktur data yang konsisten dengan karakteristik pasar fashion lokal. Periode observasi berada pada triwulan I 2025 dan mencakup 100 responden unik. Rata-rata usia konsumen berada pada rentang dewasa awal hingga madya, sedangkan pengeluaran agregat per responden memperlihatkan variasi yang memadai untuk dianalisis lebih lanjut.

Distribusi frekuensi produk menempatkan kategori Hijab sebagai item dengan pembelian tertinggi, diikuti Tas Anyaman, Batik Dress, dan Batik Shirt. Pola ini mengindikasikan tingginya preferensi konsumen terhadap produk fashion yang dekat dengan identitas lokal dan kebutuhan harian, sekaligus memberi petunjuk awal mengenai item yang berpotensi menjadi jangkar promosi atau bundling pada kampanye berikutnya, ditunjukkan pada gambar 2.



**Gambar 2.** Frekuensi pembelian

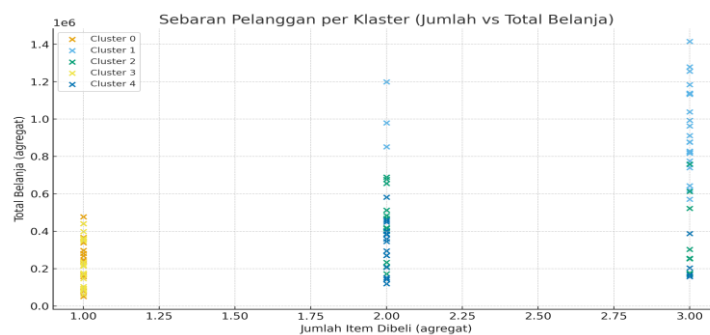
Segmentasi pelanggan dilakukan dengan memanfaatkan tiga fitur inti yang mewakili dimensi demografis dan nilai ekonomi konsumen, yakni usia, jumlah item yang dibeli secara agregat, dan total belanja agregat. Ketiga fitur ini terlebih dahulu dinormalisasi agar berada pada skala yang sebanding sehingga jarak antartitik dalam ruang fitur tidak didominasi oleh variabel bernilai besar. Pemilihan fitur didasarkan pada pertimbangan substantif: usia merepresentasikan profil demografis yang kerap berkorelasi dengan preferensi kategori fashion; jumlah item memberi sinyal intensitas pembelian; sementara total belanja menangkap nilai pelanggan bagi bisnis. Setelah prapemrosesan selesai, algoritma K-Means dijalankan dengan rentang jumlah kluster  $k=2$  hingga  $k=5$ , masing-masing dengan beberapa inisialisasi acak untuk mengurangi risiko terjebak pada solusi lokal. Kualitas pemisahan segmen diukur menggunakan silhouette score, karena metrik ini secara simultan mempertimbangkan kekompakan kluster (cohesion) dan keterpisahan antar kluster (separation).



**Gambar 2.** Silhouette Score vs Jumlah Klaster

Hasil evaluasi menunjukkan nilai silhouette terbaik pada  $k=5$ , yang mengindikasikan kompromi optimal antara granularitas segmen dan keterbacaan pola cukup rinci untuk

menangkap heterogenitas perilaku belanja, namun masih stabil dan mudah diinterpretasikan. Visualisasi sebaran dua dimensi (jumlah item vs. total belanja) memperlihatkan wilayah-wilayah yang relatif terpisah, dengan overlap yang wajar akibat sifat kontinu data. Pusat klaster (centroid) kemudian ditafsirkan untuk mengidentifikasi persona segmen, misalnya kelompok berbelanja tinggi dengan keranjang besar (high-value multi-item), kelompok berbelanja tinggi namun fokus pada sedikit item (high-spend focused), segmen menengah yang konsisten namun tidak ekstrem (mid-value), dan segmen awal dengan intensitas rendah (entry/low-engagement). Pembacaan ini memberi dasar operasional bagi penentuan strategi yang berbeda per segmen mulai dari program loyalti dan bundling komplementer untuk segmen bernilai tinggi, hingga promosi onboarding dan kupon terukur bagi segmen bernilai rendah serta membantu penataan stok, perancangan katalog, dan personalisasi komunikasi pemasaran lintas kanal. Hasil ini konsisten dengan temuan literatur bahwa K-Means efektif mengelompokkan perilaku belanja untuk dasar *targeting* pemasaran (Li et al., 2023; Yuan & Yang, 2019).



**Gambar 2.** Sebaran Pelanggan per Klaster (Jumlah vs Total Belanja)

## Pembahasan

Hasil segmentasi berbasis K-Means dengan fitur usia, jumlah item dibeli secara agregat, dan total belanja agregat menunjukkan struktur segmen yang berbeda secara jelas, dengan kualitas pemisahan terbaik pada  $k=5$  sebagaimana ditunjukkan oleh nilai silhouette tertinggi. Pemilihan  $k=5$  memberikan tingkat granularitas yang memadai: cukup kaya untuk menangkap heterogenitas perilaku belanja, tetapi masih dapat ditafsirkan secara manajerial. Centroid masing-masing klaster mengindikasikan bahwa dimensi ekonomi khususnya total belanja dan ukuran keranjang menjadi pemisah utama, sementara usia berperan sebagai atribut kontekstual yang memperkuat profil segmen tanpa mendikte batasnya.

Interpretasi terhadap pusat kluster memperlihatkan setidaknya satu segmen bernilai tinggi dengan total belanja dan jumlah item di atas rerata. Segmen ini merupakan kandidat prioritas untuk strategi retensi agresif melalui program loyalti berjenjang, akses awal koleksi baru, layanan concierge sederhana (misalnya konsultasi ukuran atau gaya), dan bundling komplementer bernilai tambah. Pada spektrum sebaliknya, segmen berbelanja rendah cenderung menunjukkan pola pembelian tunggal dengan nilai transaksi kecil dan interval kunjungan lebih jarang, sehingga lebih sesuai ditangani dengan promosi berbiaya rendah yang bertahap, seperti kupon onboarding, free shipping bersyarat, atau trial produk yang menurunkan hambatan pembelian berikutnya. Di antara keduanya, terdapat segmen menengah yang relatif stabil jumlah item moderat dan nilai belanja cukup yang ideal untuk taktik cross-sell dan upsell terukur di kategori berdekatan. Pola ini selaras dengan bukti empiris bahwa K-Means efektif sebagai baseline segmentasi pelanggan ritel sebelum organisasi bergerak ke personalisasi taktis dan perancangan kampanye yang lebih presisi (Li, Zhang, & Chen, 2023; Yuan & Yang, 2019).

Keterpaduan antara temuan kuantitatif dan kebutuhan bisnis tercermin pada implikasi operasionalnya. Pertama, hasil segmentasi memandu alokasi anggaran promosi dan prioritas stok produk jangkak dengan frekuensi tinggi pada segmen bernilai tinggi dapat diposisikan sebagai inti bundling, sementara segmen rendah diberi paparan produk entry yang harga dan manfaatnya paling relevan. Kedua, personalisasi pesan komunikasi menjadi lebih terarah karena persona tiap segmen dapat dihubungkan dengan preferensi kategori, titik harga, dan momentum kampanye. Ketiga, hasil segmentasi menjadi dasar A/B testing: misalnya membandingkan dua skema bundling pada segmen bernilai tinggi untuk mengukur uplift nilai keranjang, atau menguji dua gaya pesan promosi pada segmen entry untuk mengukur lift frekuensi pembelian.

Dari sisi validitas, penggunaan silhouette score untuk mengevaluasi variasi k memastikan keseimbangan antara cohesion dan separation. Pengulangan inisialisasi yang menghasilkan partisi serupa mengindikasikan stabilitas solusi. Normalisasi fitur sebelum pelatihan mencegah dominasi skala pada variabel bernilai besar, sehingga jarak antartitik lebih mencerminkan perbedaan substantif. Keandalan ini memperkuat kelayakan hasil untuk dijadikan masukan keputusan, sembari tetap membuka ruang untuk stress test tambahan seperti bootstrapping sampel atau analisis sensitivitas terhadap penambahan fitur.

Secara metodologis, posisi K-Means di penelitian ini tepat sebagai titik tolak yang sederhana namun kuat. Literatur menempatkan K-Means sebagai pendekatan yang unggul dari sisi keterbacaan dan kecepatan untuk data ritel, terutama ketika tujuan utama adalah memetakan heterogenitas pelanggan dan mengekstrak persona yang dapat dieksekusi (Li, Zhang, & Chen, 2023; Yuan & Yang, 2019). Seiring pemerayaan data, kualitas segmentasi dapat ditingkatkan dengan menambahkan fitur recency–frequency–monetary (RFM), preferensi kanal pembelian, atau faktor musiman/promosi. Perbandingan dengan model alternatif seperti Gaussian Mixture Models untuk pola elips atau DBSCAN untuk struktur non-linear akan bermanfaat jika data menunjukkan klaster yang tidak berbentuk sferis atau terdapat outlier yang berpengaruh.

Secara keseluruhan, pembahasan ini menegaskan bahwa segmentasi K-Means telah berhasil mengungkap struktur pasar yang relevan secara manajerial dan konsisten dengan temuan studi terdahulu. Temuan menyediakan jembatan yang jelas antara analitik dan praktik: dari data transaksi menuju kebijakan promosi, kurasi katalog, perencanaan stok, hingga desain pengalaman belanja lintas kanal, dengan tujuan akhir meningkatkan retensi, nilai keranjang, dan profitabilitas pada ekosistem fesyen lokal.

## **5. KESIMPULAN DAN SARAN**

### **Kesimpulan**

Penelitian ini menunjukkan bahwa pendekatan analitik berbasis K-Means efektif untuk memetakan heterogenitas pelanggan pada penjualan fashion lokal. Dengan tiga fitur inti : usia, jumlah item yang dibeli secara agregat, dan total belanja. Agregat model menghasilkan kualitas pemisahan terbaik pada  $k=5$  sehingga segmen yang terbentuk mudah diinterpretasikan dan langsung dapat ditindaklanjuti. Persona segmen yang muncul konsisten dengan perilaku pasar: terdapat kelompok bernilai tinggi dengan ukuran keranjang dan nilai transaksi di atas rerata yang berpotensi menjadi penggerak pendapatan; kelompok menengah yang stabil dan responsif terhadap taktik cross-sell/upsell; serta kelompok entry dengan intensitas pembelian rendah yang memerlukan intervensi retensi bertahap. Integrasi hasil kuantitatif ini dengan masukan kualitatif dari pelaku UMKM melalui diskusi terarah memperkuat kelayakan implementasi rekomendasi di lapangan.

Secara praktis, segmentasi menyediakan landasan bagi perancangan program loyaliti dan bundling untuk segmen bernilai tinggi, kurasi katalog dan promosi terukur bagi segmen menengah, serta strategi onboarding dan penurunan hambatan pembelian bagi segmen entry. Dari sisi sistem, prototipe aplikasi analitik yang dikembangkan telah tervalidasi secara fungsional dan siap diuji coba terbatas, sehingga kontribusi penelitian tidak hanya bersifat akademis tetapi juga operasional bagi UMKM fashion. Secara keseluruhan, temuan menegaskan peran segmentasi sebagai jembatan antara data transaksi dan kebijakan pemasaran, perencanaan stok, serta personalisasi komunikasi lintas kanal untuk meningkatkan retensi, nilai keranjang, dan profitabilitas..

### **Saran**

Untuk meningkatkan ketepatan dan kegunaan temuan, pengumpulan data disarankan diperkaya dengan atribut perilaku yang lebih komprehensif, khususnya metrik recency frequency monetary (RFM), preferensi kanal pembelian, serta penanda musiman dan kampanye promosi. Standardisasi pencatatan dengan identitas pelanggan yang konsisten akan memungkinkan penelusuran longitudinal dan mengurangi bias akibat keterputusan riwayat transaksi.

Pemutakhiran segmentasi perlu dilakukan secara berkala, misalnya setiap triwulan, guna menangkap dinamika perilaku konsumen dan menghindari degradasi kinerja model. Evaluasi stabilitas model sebaiknya mencakup pengulangan inisialisasi, pengambilan sampel ulang (bootstrapping), serta analisis sensitivitas terhadap penambahan atau pengurangan fitur, sehingga keputusan manajerial didasarkan pada partisi yang robust.

Seluruh proses pengolahan data dan pemodelan perlu dijalankan dengan tata kelola yang baik, meliputi kebijakan data minimization, pengamanan informasi pribadi, serta transparansi dan persetujuan penggunaan data kepada pelanggan. Kepatuhan terhadap prinsip etika dan regulasi privasi akan memperkuat keberterimaan solusi sekaligus keberlanjutan implementasi pada ekosistem UMKM fashion lokal.

### **DAFTAR REFERENSI**

- Badan Pusat Statistik. (2023). Statistik e-commerce 2023. Jakarta: BPS.
- Bedi, P., Gautam, A., & Kansal, S. (2022). Association rule mining for retail data analysis. *Procedia Computer Science*, 200, 843–850.
- Bianco, D., Bueno, A., Godinho Filho, M., Latan, H., Miller Devós Ganga, G., Frank, A. G., et al. (2023). The role of Industry 4.0 in developing resilience for manufacturing

- companies during COVID-19. *International Journal of Production Economics*, 256, 108728.
- Bottani, E., Baracetti Falcomer, M. T., & Monferdini, L. (2025). Evolution of Industry 4.0 in the fashion sector: A comparative analysis before and after COVID-19. *Procedia Computer Science*, 253, 2889–2898.
- Chauhan, C., Parida, V., & Dhir, A. (2022). Linking circular economy and digitalisation technologies: A systematic literature review of past achievements and future promises. *Technological Forecasting and Social Change*, 177, 121508.
- Creswell, J. W., & Plano Clark, V. L. (2017). *Designing and conducting mixed methods research* (3rd ed.). Thousand Oaks, CA: Sage Publications.
- Datta, S. D. (2024). Artificial intelligence and machine learning applications in the project lifecycle of the construction industry: A comprehensive review. *Journal of Construction Engineering and Management*. Advance online publication.
- Guest, G., Namey, E., & McKenna, K. (2017). How many focus groups are enough? Building an evidence base for nonprobability sample sizes. *Field Methods*, 29(1), 3–22.
- Han, J., Pei, J., & Yin, Y. (2021). Frequent pattern mining: Current status and future directions. *Data Mining and Knowledge Discovery*, 35(1), 1–36.
- Li, H., Zhang, Y., & Chen, W. (2023). Customer segmentation using K-Means for targeted marketing. *Journal of Retailing and Consumer Services*, 71, 103240.
- Liu, X. (2024). Refined management method for engineering cost combining artificial intelligence technology. *Procedia Computer Science*, 247, 503–510.
- Mahdikhani, M. (2023). Exploring commonly used terms from online reviews in the fashion field to predict review helpfulness. *International Journal of Information Management Data Insights*, 3(1), 100172.
- Nguyen, H. T., Vo, D., & Nguyen, T. (2019). Market basket analysis using improved Apriori algorithm. *International Journal of Data Science*, 4(2), 87–99.
- Pini, B., Dolci, V., Gianatti, E., Petroni, A., Bigliardi, B., & Barani, A. (2025). Artificial intelligence as a facilitator for public administration procedures: A literature review. *Procedia Computer Science*, 253, 2537–2546.
- Rahayu, R., & Day, J. (2022). E-commerce adoption by SMEs in developing countries: Evidence from Indonesia. *Telematics and Informatics*, 64, 101679.
- Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, 104, 333–339.
- Szramowiat-Sala, K., Penkala, R., Horák, J., Krpec, K., Hopan, F., & Ryšavý, J. (2024). AI-based data mining approach to control the environmental impact of conventional energy technologies. *Journal of Cleaner Production*, 472, 143473.

- Wang, H. B., & Gao, Y. J. (2021). Research on parallelization of Apriori algorithm in association rule mining. *Procedia Computer Science*, 183, 641–647.
- Yu, D., & Zhao, P. (2024). Research on network clothing design system based on artificial intelligence. *Procedia Computer Science*, 247, 27–35.
- Yuan, Y., & Yang, Y. (2019). K-Means clustering in retail customer analysis. *Journal of Business Research*, 98, 239–247.