

BAB II

TINJAUAN PUSTAKA

2.1 E-Wallet

E-Wallet (electronic wallet) merupakan salah satu bentuk teknologi finansial yang digunakan untuk menyimpan dan mengelola dana secara digital. E-wallet memungkinkan penggunanya melakukan berbagai jenis transaksi secara elektronik tanpa perlu menggunakan uang tunai secara langsung. Dalam operasionalnya, e-wallet menyimpan data keuangan pengguna melalui aplikasi digital yang terhubung dengan sistem pembayaran daring dan lembaga keuangan (9). Fungsi utama dari dompet digital adalah sebagai alat pembayaran yang dapat digunakan dalam berbagai keperluan transaksi harian. Melalui aplikasi e-wallet, pengguna dapat melakukan pembayaran barang dan jasa, pengisian pulsa dan paket data, pembayaran tagihan listrik dan air, pembelian tiket transportasi, hingga transfer saldo antar pengguna. Selain itu, e-wallet juga dimanfaatkan pada transaksi di platform e-commerce maupun merchant offline yang telah terintegrasi dengan sistem pembayaran berbasis *QR Code*.

Manfaat penggunaan dompet digital mencakup kemudahan akses, kecepatan transaksi, efisiensi biaya, serta peningkatan keamanan. Transaksi melalui e-wallet dapat dilakukan kapan saja dan di mana saja selama pengguna memiliki koneksi internet dan saldo yang mencukupi. Sistem ini umumnya dilengkapi fitur keamanan seperti PIN, verifikasi dua langkah, serta notifikasi transaksi secara real-time yang berfungsi melindungi akun dari penyalahgunaan. Kemudahan pelacakan histori transaksi serta integrasi dengan berbagai program promosi menjadikan e-wallet sebagai solusi praktis dalam menunjang kebutuhan keuangan sehari-hari (10). Beberapa layanan dompet digital yang populer digunakan di Indonesia antara lain GoPay, OVO, DANA, dan ShopeePay. GoPay terintegrasi dalam ekosistem Gojek, OVO digunakan secara luas pada platform Grab dan Tokopedia, DANA banyak terhubung dengan Bukalapak dan merchant lainnya, sementara ShopeePay menjadi metode utama pada platform Shopee. Masing-masing layanan menawarkan keunggulan yang berbeda, seperti *cashback*, potongan harga, loyalty

point, hingga integrasi dengan layanan transportasi, hiburan digital, dan logistik. Ragam fitur ini mendorong terbentuknya preferensi pengguna yang beragam terhadap e-wallet di Indonesia. Tren penggunaan dompet digital di Indonesia menunjukkan pertumbuhan yang signifikan dalam beberapa tahun terakhir.

Berdasarkan data dari (11), e-wallet merupakan metode pembayaran digital paling banyak digunakan masyarakat Indonesia, terutama untuk transaksi di platform *e-commerce*. Survei menunjukkan bahwa sekitar 84,3% pengguna *e-commerce* di Indonesia menggunakan e-wallet sebagai metode pembayaran utama, mengungguli transfer bank maupun kartu debit dan kredit. Peningkatan ini terjadi secara konsisten dari tahun ke tahun, mencerminkan tingginya adopsi sistem pembayaran digital di berbagai lapisan masyarakat. Pandemi COVID-19 turut mempercepat pertumbuhan penggunaan dompet digital, seiring meningkatnya kebutuhan transaksi tanpa kontak fisik. Kondisi ini mendorong lonjakan pengguna baru dari berbagai kalangan, termasuk pelajar, mahasiswa, ibu rumah tangga, hingga masyarakat perdesaan. Penggunaan e-wallet kini tidak hanya didominasi oleh masyarakat urban dengan tingkat literasi digital tinggi, namun juga meluas ke segmen pengguna yang sebelumnya lebih terbiasa dengan transaksi tunai.

Alasan utama masyarakat menggunakan e-wallet antara lain adalah kemudahan penggunaan, kecepatan transaksi, keamanan, serta beragam promosi menarik seperti cashback dan potongan harga. Preferensi ini menjadikan e-wallet bukan sekadar alat pembayaran alternatif, melainkan sarana utama dalam pengelolaan transaksi harian yang praktis dan efisien. Seiring dengan meningkatnya kebutuhan akan transaksi digital yang cepat dan aman, serta berkembangnya ekosistem layanan digital di berbagai sektor, dompet digital diperkirakan akan terus mengalami pertumbuhan baik dari sisi jumlah pengguna maupun volume transaksi. Peran e-wallet sebagai bagian dari sistem pembayaran modern menjadikannya instrumen penting dalam mendukung transformasi sistem keuangan nasional menuju arah yang lebih digital dan adaptif (12).

2.2 Perilaku Penggunaan E-Wallet

Perilaku konsumen merupakan serangkaian proses yang mencakup bagaimana individu mencari, memilih, menggunakan, dan mengevaluasi suatu produk atau layanan. Dalam konteks digital, perilaku konsumen mengalami pergeseran dari pola tradisional ke pola berbasis teknologi, seiring berkembangnya ekosistem digital yang memudahkan akses terhadap informasi dan layanan keuangan. Perilaku konsumen dalam penggunaan e-wallet mencerminkan pola preferensi konsumen dalam memanfaatkan layanan dompet digital untuk memenuhi kebutuhan transaksi secara praktis dan efisien (13).

Faktor-faktor yang memengaruhi perilaku pengguna e-wallet sangat beragam. Beberapa di antaranya adalah kemudahan penggunaan, kenyamanan, kecepatan transaksi, persepsi keamanan, tampilan antarmuka aplikasi, serta insentif promosi seperti cashback atau diskon. Selain itu, faktor eksternal seperti dukungan infrastruktur digital, kebiasaan sosial, dan keterpaparan terhadap media digital turut membentuk preferensi dan kebiasaan penggunaan dompet digital di kalangan masyarakat (14). Peran lingkungan, termasuk pengaruh teman, keluarga, dan komunitas, juga dapat memengaruhi keputusan seseorang dalam memilih e-wallet tertentu. Selain itu, faktor sosial budaya turut berkontribusi dalam membentuk preferensi pengguna.

Secara umum, perilaku pengguna e-wallet dapat diklasifikasikan ke dalam beberapa tipe berdasarkan intensitas dan motivasi penggunaannya. Tipe pertama adalah pengguna intensif (*frequent users*), yaitu individu yang hampir seluruh transaksinya dilakukan secara digital melalui e-wallet. Tipe kedua adalah pengguna selektif, yang hanya menggunakan e-wallet untuk kebutuhan tertentu seperti belanja daring atau pembayaran tagihan. Selanjutnya, terdapat pengguna oportunistik, yaitu pengguna yang tertarik memanfaatkan e-wallet hanya ketika terdapat insentif seperti promo atau *cashback*. Terakhir, pengguna pasif adalah individu yang menggunakan e-wallet dalam situasi terbatas dan masih mengandalkan metode pembayaran konvensional dalam aktivitas keuangannya sehari-hari.

Kelompok usia muda, khususnya generasi Z dan milenial, memiliki tingkat adopsi yang lebih tinggi terhadap layanan e-wallet dibandingkan kelompok usia lain. Hal ini disebabkan oleh tingginya literasi digital, gaya hidup praktis, serta orientasi terhadap kemudahan dan kecepatan. Survei (15) mengindikasikan bahwa mayoritas pengguna aktif e-wallet berasal dari segmen usia 18–35 tahun, yang mengutamakan efisiensi dalam bertransaksi dan seringkali memanfaatkan promo yang ditawarkan oleh *platform* digital.

Perilaku penggunaan e-wallet di Indonesia juga dipengaruhi oleh segmentasi geografis dan sosio-ekonomi. Di daerah perkotaan, adopsi e-wallet cenderung lebih tinggi karena dukungan infrastruktur teknologi dan aksesibilitas yang lebih baik. Sedangkan di daerah perdesaan, meskipun adopsinya lebih lambat, tren penggunaannya tetap menunjukkan pertumbuhan seiring dengan penyebaran internet dan smartphone yang semakin meluas. Perbedaan gaya hidup, tingkat pendapatan, serta keterpaparan terhadap layanan digital turut memengaruhi intensitas dan pola penggunaan dompet digital di setiap segmen masyarakat.

2.3 Data Mining

Data mining adalah proses untuk menemukan informasi tersembunyi dari kumpulan data berukuran besar, yang dikenal sebagai *knowledge discovery*. Metode ini memanfaatkan teknik analisis data untuk mengidentifikasi pola-pola tertentu dalam basis data. Secara umum, data mining merupakan proses eksplorasi dan ekstraksi informasi bernilai dari data berskala besar dengan bantuan metode statistik, pembelajaran mesin, dan kecerdasan buatan. Tujuannya adalah mengungkap pengetahuan tersembunyi yang dapat digunakan dalam mendukung pengambilan keputusan. Data mining juga merupakan bagian dari proses yang lebih luas, yaitu *Knowledge Discovery in Databases* (KDD), yang mencakup tahapan seperti pengumpulan data, prapemrosesan, analisis, hingga interpretasi hasil (16).

Pola dalam data mining dianggap bernilai jika memenuhi kriteria seperti tidak bersifat remeh, mengandung makna tersembunyi, belum diketahui sebelumnya, serta bermanfaat dalam proses pengambilan keputusan. Pola yang ditemukan sebaiknya mudah dipahami, mampu memprediksi data lain dengan

akurasi tertentu, dan dapat diterapkan dalam berbagai bidang. Proses pencarian pola ini juga dikenal dengan istilah lain seperti *Knowledge Discovery in Databases* (KDD), analisis pola, eksplorasi data, ekstraksi informasi, maupun business intelligence(17) .

Data mining dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan, seperti:

1. Deskripsi

Bertujuan untuk menggali dan memaparkan pola-pola yang ada di dalam data, serta memberikan pemahaman tentang tren atau kecenderungan yang muncul. Informasi yang diperoleh dari analisis ini dapat membantu dalam mengidentifikasi hubungan antara variabel-variabel dalam dataset.

2. Estimasi

Memiliki konsep yang serupa dengan klasifikasi, namun variabel target yang digunakan bersifat numerik, bukan kategorikal. Model estimasi dibangun dengan menggunakan data yang memiliki nilai lengkap, dan model tersebut kemudian digunakan untuk memperkirakan nilai target berdasarkan variabel prediktor lainnya.

3. Prediksi

Digunakan untuk meramalkan nilai atau kondisi di masa depan berdasarkan pola dalam data historis. Metode ini umum digunakan dalam peramalan penjualan, prediksi kebutuhan bahan baku, dan analisis tren jangka panjang.

4. Klasifikasi

Klasifikasi merupakan teknik dalam data mining yang digunakan untuk mengelompokkan data ke dalam kategori tertentu berdasarkan sifat atau karakteristiknya. Contoh penerapannya antara lain deteksi spam, diagnosis medis dan segmentasi pelanggan.

5. *Clustering*

Teknik yang digunakan untuk mengelompokkan data berdasarkan kesamaan karakteristik tanpa menggunakan variabel target. Objek yang mirip akan dikelompokkan ke dalam satu klaster. *Clustering* sering digunakan dalam

analisis pemasaran, pengelompokan dokumen, dan analisis perilaku pelanggan.

Proses data mining terdiri dari beberapa tahapan penting, yaitu:

1. Seleksi data: pemilihan data relevan dari berbagai sumber yang tersedia.
2. Prapemrosesan data: melibatkan proses pembersihan, penanganan nilai kosong, dan normalisasi agar data siap dianalisis.
3. Transformasi data: konversi data ke format atau bentuk yang sesuai untuk proses analitik.
4. Penerapan metode data mining: proses ini melibatkan teknik seperti klasifikasi, asosiasi, regresi, atau *clustering*.
5. Evaluasi hasil: untuk menilai validitas, akurasi, dan relevansi dari pola yang ditemukan (18).

Dengan perkembangan teknologi dan meningkatnya volume data, metode data mining telah menjadi pendekatan penting dalam berbagai bidang seperti pemasaran, keuangan, kesehatan, pendidikan, serta pengembangan sistem berbasis data. Metode ini mampu mengolah data dalam jumlah besar dan kompleks menjadi informasi yang berguna dan mudah dipahami.

2.4 Clustering

Clustering merupakan salah satu teknik dalam data mining yang termasuk dalam kategori analisis deskriptif, yaitu analisis yang bertujuan menggambarkan struktur atau pola dalam data. *Clustering* bertujuan untuk mengelompokkan data ke dalam beberapa kelompok atau klaster berdasarkan kemiripan atribut tertentu, tanpa adanya label atau variabel target. Teknik ini sering disebut sebagai *unsupervised learning* karena proses pengelompokan dilakukan tanpa proses pembelajaran dari data berlabel (19).

Tujuan utama dari *clustering* adalah untuk menemukan struktur tersembunyi dalam data dengan mengelompokkan objek-objek yang serupa ke dalam satu kelompok, dan memisahkan objek yang berbeda ke dalam kelompok lainnya. Hasil dari proses ini dapat memberikan wawasan mengenai distribusi data,

pola hubungan antar variabel, atau kecenderungan umum suatu populasi. Teknik ini bermanfaat untuk menyederhanakan data kompleks, melakukan segmentasi pasar, mendeteksi anomali, serta mengelompokkan dokumen atau perilaku pelanggan.

Clustering sering dibandingkan dengan klasifikasi, meskipun keduanya memiliki perbedaan mendasar. Pada klasifikasi, proses pembelajaran dilakukan menggunakan data berlabel (*supervised learning*), sedangkan *clustering* tidak menggunakan data berlabel (*unsupervised learning*). Klasifikasi bertujuan untuk memprediksi label dari data baru, sedangkan *clustering* digunakan untuk mengeksplorasi struktur dan pola dalam data.

Terdapat beberapa metode *clustering* yang umum digunakan, antara lain:

1. *Hierarchical Clustering*

Metode yang membentuk hierarki kluster secara bertahap melalui pendekatan *agglomerative* (penggabungan) atau *divisive* (pemecahan).

2. *Density-Based Clustering*

Contohnya DBSCAN, yaitu metode yang mengelompokkan data berdasarkan kerapatan distribusi data.

3. *Grid-Based Clustering*

Membagi ruang data ke dalam grid dan mengelompokkan berdasarkan kepadatan masing-masing sel grid.

4. *Partitioning Clustering*

Metode yang membagi dataset ke dalam sejumlah kluster yang telah ditentukan sebelumnya, seperti pada algoritma *K-Means*.

Teknik *clustering* telah banyak diterapkan dalam berbagai bidang, seperti pemasaran (segmentasi pelanggan), kesehatan (pengelompokan pasien berdasarkan gejala), pendidikan (pengelompokan gaya belajar siswa), dan sistem informasi (pengelompokan pengguna berdasarkan perilaku). Kemampuannya dalam mengungkap struktur tersembunyi dalam data menjadikannya salah satu pendekatan penting dalam analisis eksploratif.

2.5 K-Means Clustering

K-Means Clustering merupakan salah satu algoritma pengelompokan data (*clustering*) yang paling banyak digunakan dalam bidang data mining. Metode ini termasuk dalam *unsupervised learning* yang bertujuan untuk mengelompokkan data ke dalam sejumlah klaster berdasarkan kemiripan atribut. Setiap data akan dimasukkan ke dalam klaster dengan *centroid* atau titik pusat terdekat.

Menurut (20), *K-Means* merupakan algoritma partisi berbasis iterasi yang membagi data ke dalam klaster dengan meminimalkan jarak *Euclidean* antara data dan pusat klaster (*centroid*). Algoritma ini dimulai dengan menentukan jumlah klaster (k), memilih titik awal *centroid* secara acak, menghitung jarak setiap data ke masing-masing *centroid*, dan mengelompokkan data berdasarkan jarak terdekat. Proses ini diulang hingga posisi *centroid* stabil atau tidak mengalami perubahan signifikan. Menurut (21), menyatakan bahwa algoritma *K-Means* sangat cocok diterapkan dalam studi perilaku pengguna e-wallet karena data yang digunakan bersifat numerik, terukur, dan memiliki dimensi yang jelas. Hal ini memungkinkan terbentuknya klaster yang relevan dan mudah diinterpretasikan.

Metode ini dikenal efisien dalam menangani dataset berukuran sedang hingga besar, serta fleksibel untuk berbagai konteks analisis seperti segmentasi pasar, pengelompokan pengguna, dan pemetaan perilaku digital. *K-Means* banyak digunakan untuk mengelompokkan individu berdasarkan kecenderungan perilaku yang dapat diamati dari data numerik.

Tahapan kerja algoritma *K-Means* secara umum adalah sebagai berikut:

1. Menentukan jumlah *cluster* (k) yang diinginkan.
2. Menginisialisasi k *centroid* secara acak.
3. Menghitung jarak antara setiap data dan masing-masing *centroid*.
4. Menetapkan data ke *cluster* berdasarkan jarak terdekat.
5. Menghitung ulang posisi *centroid* berdasarkan rata-rata data dalam tiap *cluster*.

6. Mengulangi proses hingga posisi *centroid* konvergen atau jumlah iterasi maksimum tercapai.

Algoritma ini memiliki empat komponen penting, yaitu:

1. Rumus Jarak *Euclidean*

Dalam proses *K-Means*, perhitungan jarak antara data dan *centroid* sangat krusial untuk menentukan *cluster* yang tepat. Salah satu metode yang digunakan adalah jarak *Euclidean*, yang mengukur seberapa jauh satu titik dari titik lainnya dalam ruang berdimensi.

Rumus jarak *Euclidean* adalah sebagai berikut:

$$d(x, \mu) = \sqrt{\sum_{j=1}^n (x_j - \mu_j)^2} \dots \dots \dots (2.1)$$

Keterangan:

- $d(x, \mu)$: Jarak *Euclidean* antara titik data x dan *centroid* μ
- x_j : Nilai data pada dimensi ke- j
- μ_j : Nilai *centroid* pada dimensi ke- j
- n : Jumlah dimensi atau atribut

2. Rumus *Centroid* Baru

Dalam proses *K-Means*, setelah setiap data dikelompokkan berdasarkan jarak terdekat terhadap pusat *cluster* (*centroid*), langkah selanjutnya pada algoritma *K-Means* adalah menghitung ulang posisi *centroid* baru. Tujuan dari proses ini adalah untuk memperbarui posisi pusat *cluster* agar lebih representatif terhadap anggota *cluster*, sehingga hasil pengelompokan menjadi lebih akurat pada iterasi selanjutnya.

Rumus untuk menghitung *centroid* baru ditunjukkan sebagai berikut :

$$\mu_j = \frac{1}{n_i} \sum_{x \in C_i} x \dots \dots \dots (2.2)$$

Keterangan :

μ_j : *Centroid* (titik pusat) baru pada *cluster* ke- i

n_i : Jumlah data pada *cluster* ke- i

C_i : Himpunan data yang termasuk ke dalam *cluster* ke- i

x : Vektor data (nilai tiap variabel atau atribut) yang menjadi anggota *cluster* ke- i

3. *Sum of Squared Errors* (SSE)

Sum of Squared Errors merupakan ukuran kesalahan (*error*) yang digunakan untuk mengevaluasi kualitas klaster pada algoritma *K-Means*. SSE menghitung jumlah kuadrat jarak antara setiap data dengan *centroid cluster* tempat data tersebut tergabung. Nilai SSE digunakan untuk melihat tingkat kompaknya klaster.

Semakin kecil nilai SSE, maka semakin baik kualitas *cluster* yang dihasilkan, karena menunjukkan bahwa data dalam klaster tersebut semakin rapat terhadap *centroid*-nya. Sebaliknya nilai SSE yang besar menunjukkan bahwa data dalam *cluster* masih tersebar jauh dari *centroid*.

Rumus SSE dapat dituliskan sebagai berikut:

$$SSE = \sum_{i=1}^k \sum_{x \in C_i} \|x - C_i\|^2 \dots \dots \dots (2.3)$$

Keterangan:

k : Jumlah *cluster*

C_i : *Cluster* ke- i

x : Data yang termasuk dalam *cluster* C_i

c_i : *Centroid* dari *cluster* ke- i

$\|x - c_i\|^2$: kuadrat jarak *Euclidean* antara data x dengan *centroid* c_i

SSE sering digunakan untuk menentukan jumlah *cluster* yang optimal atau kompak data. Nilai SSE akan semakin menurun dengan bertambahnya jumlah *cluster* namun penurunan tersebut akan melambat setelah mencapai jumlah *cluster* yang tepat.

2.6 Davies-Bouldin Index

Davies-Bouldin Index diperkenalkan oleh David L. Davies and Donald W. Bouldin pada tahun 1979. Yang dimana salah satu metode evaluasi internal yang digunakan untuk mengukur kualitas hasil *clustering*. Evaluasi ini dilakukan dengan menghitung rasio antara kedekatan data dalam *cluster* (*compactness*) dengan jarak antar kluster (*separation*). Prinsip utama DBI adalah *cluster* yang baik yang dimana data dalam *cluster* tersebut saling berdekatan (kompak), dan *centroid* antar kluster saling berjauhan (terpisah dengan jelas). *Sum-of square within cluster* (SSW) sebagai metrik kohesi dalam sebuah *cluster*. Separasi dengan *Sum-of-square-between-cluster* (SSWB) dengan mengukur jarak antara *centroid* C_i dan C_j . $R_{i,j}$ adalah ukuran rasio seberapa baik nilai perbandingan antara *cluster* ke- i dan *cluster* ke- j . Dengan kata lain, DBI berusaha menilai keseimbangan antara homogenitas dalam *cluster* dan geterogenitas antar *cluster*.

Rumus DBI adalah :

$$DBI = \frac{1}{k} \sum_{i=1}^k \text{Max}_{j \neq i} \left(\frac{S_i + S_j}{M_{ij}} \right) \dots \dots \dots (2.4)$$

Keterangan:

k : Jumlah *cluster*

S_i : Rata-rata jarak antara setiap data dalam *cluster* i terhadap *centroid cluster* i (mengukur kompaknya *cluster*)

M_{ij} : Jarak antara *centroid* kluster i dan *centroid cluster* j (mengukur pemisahan antar *cluster*)

$\frac{S_i + S_j}{M_{ij}}$: Rasio antara gabungan jarak dalam *cluster* dengan jarak antar *cluster*.

Rumus SSW adalah :

$$\frac{1}{N} \sum_{i=1}^N \|X_i - C_{pi}\|^2 \dots \dots \dots (2.5)$$

Rumus SSB adalah :

$$\frac{2}{M(M-1)} \sum_{i=1}^M \sum_{j=1}^M \|C_i - C_j\|^2 \dots \dots \dots (2.6)$$

Rumus R dan DBI adalah :

$$R_{i,j} = \frac{SSW_i + SSW_j}{SSB_{i,j}} \dots \dots \dots (2.7)$$

2.7 Penelitian Terkait

Pencarian literatur berfungsi sebagai acuan dalam menyusun dasar teori dan arah penelitian. Tujuannya adalah untuk mengetahui kesenjangan, relevansi topik, dan sebagai bahan perbandingan dari penelitian sebelumnya.

Tabel 2. 1 Penelitian Terkait

No	Landasan Literatur	Metode	Masalah	Hasil Penelitian
1	“PENGELOMPOKAN TRANSAKSI KARTU DEBIT PERBANKAN MENGGUNAKAN ALGORITMA <i>K-MEANS</i> ” (22)	<i>K-Means</i>	Masalah yang ada: Strategi pemasaran pihak bank belum efektif karena masih menggunakan segmentasi generik— tidak didasarkan pada pola penggunaan kartu kredit aktual. Hal ini menyebabkan munculnya segmen yang tidak sepakat dengan karakteristik perilaku transaksi, sehingga ROI promosi menurun.	Hasil dari penelitian ini: Analisis menunjukkan bahwa <i>K-Means</i> dengan $k=3$ menghasilkan segmentasi paling sesuai dengan data nasabah. Tiga kluster yang terbentuk memiliki karakteristik berbeda pada frekuensi transaksi, batas kredit, dan jenis belanja. Bank dapat menggunakan ini sebagai dasar perumusan strategi promosi dan penawaran produk yang lebih spesifik ke tiap kelompok nasabah.

2	<p>“Penerapan K-Means Clustering untuk Segmentasi Konsumen E-Commerce Berdasarkan Pola Pembelian” (4)</p>	<p><i>K-Means + Elbow</i></p>	<p>Masalah yang ada: Platform <i>e-commerce</i> kesulitan merancang strategi pemasaran dan pengelolaan inventori karena belum memiliki segmentasi pelanggan berdasarkan kebiasaan belanja. Segmentasi yang ada terlalu umum—hanya mengandalkan data demografi—sehingga kampanye promosi tidak tepat sasaran.</p>	<p>Hasil dari penelitian ini: Analisis frekuensi, nilai pembelian, dan waktu transaksi menghasilkan 5 kluster. Misalnya, kluster pertama berisi “Pembeli VIP” dengan frekuensi tinggi dan total belanja besar; kluster kelima terdiri dari “Pembeli pasif” yang hanya membeli ketika ada diskon. Rekomendasi meliputi pemberian penawaran khusus promosi waktu untuk meningkatkan keterlibatan tiap kluster.</p>
3	<p>“Segmentasi Pelanggan Berdasarkan Perilaku Penggunaan Kartu Kredit Menggunakan Metode <i>K-Means Clustering</i>” (23)</p>	<p><i>K-Means</i></p>	<p>Masalah yang ada: Pemerintah Provinsi Jawa Tengah ingin meningkatkan kampanye pendidikan <i>cashless</i>. Namun, belum punya segmentasi pengguna e-wallet yang jelas berdasarkan frekuensi penggunaan dan</p>	<p>Hasil dari penelitian ini: Data pengguna (frekuensi penggunaan, jumlah e-wallet, reaksi promo) terbagi menjadi tiga kluster: 1) <i>heavy users & promo-savvy</i>; 2) <i>multi-wallet low-frequency users</i>; 3) <i>casual users</i>. Disarankan strategi</p>

			sensitivitas promosi, sehingga edukasi kebijakan belum tertarget.	edukasi berbeda, misalnya <i>heavy users</i> bisa jadi <i>ambassador cashless</i> , sedangkan <i>casual users</i> perlu pemberian insentif langsung agar mulai menggunakan e-wallet lebih sering.
4	“SEGMENTASI NASABAH BANK MENGGUNAKAN ALGORITMA <i>K-MEANS CLUSTERING</i> DAN VISUALISASI DINAMIS DENGAN <i>STREAMLIT</i> ” (24)	<i>K-Means</i> + <i>WCSS</i> + <i>Davies-Bouldin Index</i>	Masalah yang ada: Lembaga keuangan tidak memiliki segmentasi risiko nasabah berdasarkan profil demografi dan komposisi portofolio produk—sehingga penawaran produk tidak dipersonalisasi dan dapat meningkatkan risiko kredit.	Hasil dari penelitian ini: Ditemukan dua pilihan kluster optimal: $k=2$ (berisi nasabah aktif dan pasif) atau $k=4$ (memisahkan berdasarkan risiko tinggi/rendah dan keterlibatan produk). Rekomendasi: bank dapat menyesuaikan pendekatan pemasaran produk dan limit kredit sesuai karakter kluster.
5	“Analisis Segmentasi Pelanggan Berbasis Data Transaksi Menggunakan Metode <i>K-Means</i> ” (25)	<i>K-Means</i>	Masalah yang ada: Supermarket tidak memiliki cara sistematis mengidentifikasi pelanggan prioritas. Mereka belum bisa	Hasil dari penelitian ini: Dua kluster terbentuk “Konsumen Prioritas” dan “Konsumen Biasa”. Kluster prioritas (109 anggota) memiliki rata-

			<p>membedakan pelanggan potensial tinggi yang loyal dari pembeli biasa sehingga alokasi budget promosi dan reward program kurang efektif.</p>	<p>rata frekuensi belanja tinggi dan total pengeluaran signifikan, serta memiliki gaya belanja yang konsisten. Supermarket dapat menyiapkan program loyalty khusus untuk klaster ini dan iklan lebih aggressif untuk klaster biasa.</p>
6	<p>“OPTIMISASI STRATEGI PENJUALAN BERBASIS SEGMENTASI PELANGGAN DENGAN METODE <i>K-MEANS CLUSTERING</i> DI PT. SANSYU PRECISION INDONESIA” (26)</p>	<p><i>K-Means</i></p>	<p>Masalah yang ada: Distributor tidak memiliki segmentasi geografis dan nilai transaksi konsumen yang jelas. Hal tersebut membuat strategi distribusi dan pemasaran menjadi kurang terstruktur dan tidak efisien dalam pengiriman dan penawaran promosi.</p>	<p>Hasil dari penelitian ini: Tiga klaster utama terbentuk: klaster bisnis (transaksi besar bisnis), klaster regional (khusus konsumen di Jawa), dan klaster <i>high spenders</i> (tinggi belanja per tahun). Distributor bisa memfokuskan promosi wilayah, eksplorasi pasar bisnis, dan menyesuaikan program loyalty berdasarkan klaster.</p>
7	<p>“Analisis Preferensi Konsumen dan</p>	<p><i>K-Means + Elbow Method</i></p>	<p>Masalah yang ada: Perusahaan otomotif tidak memiliki strategi</p>	<p>Hasil dari penelitian ini: Dua klaster terbaik terbentuk (DBI=0.802)</p>

	Profil Pembeli di Industri Otomotif: Studi Kasus Pembelian Kendaraan di Wilayah Jabodetabek” (27)		pemasaran yang tertarget karena belum mengidentifikasi segmen pelanggan berdasarkan preferensi kendaraan, rentang umur, dan lokasi. Kampanye iklan menjadi terlalu umum	yang membedakan pelanggan berdasarkan jenis kendaraan dan demografi, memberikan strategi komunikasi dan promo yang dipersonalisasi menurut karakteristik kluster tersebut.
8	“SEGMENTASI DATA TRANSAKSI PENJUALAN TOKO ONLINE VASTYLE UNTUK MANAJEMEN STOK MENGGUNAKAN ALGORITMA <i>K-MEANS CLUSTERING</i> ” (28)	<i>K-Means + Elbow Method</i>	Masalah yang ada: Retail online tidak memiliki pemetaan kluster pelanggan yang mendukung strategi pengaturan stok dan segmentasi pelanggan— akibatnya, stok cepat habis di satu segmen tapi surplus di segmen lain.	Hasil dari penelitian ini: Empat kluster terbentuk berdasarkan kuantitas belanja dan rata-rata harga per item. Retail dapat menggunakan kluster ini untuk optimasi manajemen stok, menyesuaikan tawar diskon, dan menyiapkan prediksi permintaan per segmen.
9	“Segmentasi Pelanggan Retail Berbasis Perilaku menggunakan Algoritma <i>K Means Clustering</i> ” (29)	<i>K-Means + Elbow Method + Silhouette Score</i>	Masalah yang ada: Perusahaan ritel belum menerapkan segmentasi berbasis nilai pelanggan (RFM), membuat CRM tidak terarah	Hasil dari penelitian ini: Tiga kluster terbentuk— <i>General Development, Maintenance, dan Important Maintenance</i> . Dengan

			dan program loyalitas kurang efektif.	segmentasi ini, CRM perusahaan dapat mengimplementasikan program loyalitas jangka pendek (promo diskon) serta skema VIP untuk anggota klaster unggul.
10	“KLASTERISASI LOYALITAS PELANGGAN BERDASARKAN ANALISIS RFM MENGGUNAKAN GAUSSIAN MIXTURE MODEL (GMM)” (30)	<i>K-Means</i>	Masalah yang ada: Kedai kopi belum memiliki strategi promosi yang efektif karena tidak mengetahui segmentasi pelanggan berdasarkan budget ngopi, frekuensi kunjungan, dan respons terhadap promo <i>cashless</i> .	Hasil dari penelitian ini: Tiga klaster terbentuk: (1) pelanggan rutin GoPay + <i>cashback minded</i> ; (2) pelanggan pasif tanpa preferensi promo; (3) pelanggan menengah. Rekomendasi: kembangkan penawaran <i>cashback eksklusif</i> untuk klaster 1, buat program membership untuk klaster 3, dan edukasi penggunaan e-wallet untuk klaster 2.