

Implementasi Data Mining Untuk Analisis Prilaku Konsumen Dalam Meningkatkan Penjualan Di Shopee

¹Asyraf Muntasir Pratama, ²Siswanto, ³Ricky Zulfiandry

¹ Mahasiswa, Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Dehasen Bengkulu
Jalan Melati 3, No 497 Rt.15, Rw.03 Kelurahan Bumi Ayu. Kecamatan Selebar Kota Bengkulu
e-mail: asyrafpratama12@gmail.com

² Dosen Tetap, Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Dehasen Bengkulu
Jl. Meranti Raya No. 32 Kota Bengkulu 38228 Telp. (0736) 22027, 26957 Fax. (0736) 341139;
e-mail: sisunived2017@gmail.com

³ Dosen Tetap, Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Dehasen Bengkulu
Jl. Meranti Raya No. 32 Kota Bengkulu 38228 Telp. (0736) 22027, 26957 Fax. (0736) 341139;
e-mail: ricky.zulfiandry@unived.ac.id

(Received: Mei 2025, Revised: Agustus 2025, Accepted: Oktober 2025)

Abstract-The rapid growth of e-commerce requires an in-depth understanding of consumer behavior to increase sales. This research applies a data mining method using the K-Means clustering algorithm to analyze weekly revenue data from the Anjani_Store.id store on the Shopee platform, in the period from January to July 2023. The research uses a waterfall approach and is implemented through the Python programming language with the Flask framework. The clustering results produce three groups of consumer behavior based on income levels: low, medium, and high. Clusters are analyzed to support promotional strategies and business decision making. This system was tested and showed that the application of the K-Means method was effective in grouping consumers appropriately and supporting increased sales based on data analysis.

Keywords: Data Mining, K-Means, Clustering, Consumer Behavior, Shopee, Anjani_Store.id.

Intisari-Pertumbuhan e-commerce yang pesat menuntut pemahaman perilaku konsumen secara mendalam untuk meningkatkan penjualan. Penelitian ini menerapkan metode data mining menggunakan algoritma K-Means clustering guna menganalisis data pendapatan mingguan toko Anjani_Store.id di platform Shopee, dalam rentang waktu Januari hingga Juli 2023. Penelitian menggunakan pendekatan waterfall dan diimplementasikan melalui bahasa pemrograman Python dengan framework Flask. Hasil klusterisasi menghasilkan tiga kelompok perilaku konsumen berdasarkan tingkat pendapatan: rendah, sedang, dan tinggi. kluster dianalisis untuk mendukung strategi promosi dan pengambilan keputusan bisnis. Sistem ini diuji dan menunjukkan bahwa penerapan metode K-Means efektif dalam mengelompokkan konsumen secara tepat dan mendukung peningkatan penjualan berbasis analisis data.

Kata Kunci: Data Mining, K-Means, Clustering, Perilaku Konsumen, Shopee, Anjani_Store.id.

I. PENDAHULUAN

Perkembangan teknologi informasi dan komunikasi dalam dekade terakhir telah menghasilkan transformasi besar dalam berbagai sektor, salah satunya adalah sektor perdagangan. Digitalisasi yang pesat mendorong pergeseran dari transaksi konvensional ke platform daring (*E-commerce*). Dalam konteks Indonesia, penetrasi internet yang terus meningkat dan kemudahan aksesibilitas terhadap perangkat mobile telah menjadikan *E-commerce* sebagai salah satu sektor yang tumbuh

paling cepat. Berdasarkan laporan dari Google, Temasek, dan Bain & Company (2023), nilai pasar *E-commerce* Indonesia diproyeksikan akan mencapai lebih dari USD 90 miliar pada tahun 2025. Di antara berbagai platform *E-commerce*, Shopee muncul sebagai pemain dominan, menarik perhatian jutaan pengguna aktif dengan berbagai fitur seperti pengiriman cepat, promosi besar-besaran, serta antarmuka yang ramah pengguna. Meskipun pertumbuhan Shopee begitu pesat, persaingan yang ketat dengan platform lain seperti Tokopedia, Lazada, dan Bukalapak menuntut perusahaan untuk terus meningkatkan kualitas layanan dan pengalaman pengguna. Salah satu tantangan utama yang dihadapi oleh platform *E-commerce* saat ini adalah memahami perilaku konsumen secara mendalam guna menciptakan strategi pemasaran yang lebih personal dan tepat sasaran. Mengingat tingginya volume data transaksi yang dihasilkan setiap hari, manual tracking terhadap pola perilaku konsumen menjadi tidak efektif. Oleh karena itu, diperlukan teknologi yang dapat secara otomatis mengolah data besar (*Big data*) untuk menghasilkan *insight* yang relevan dan bernilai bagi pengambilan keputusan. *Data Mining* adalah salah satu metode yang efektif dalam menangani masalah pengolahan data besar di era digital. *Data mining* merupakan proses penggalian informasi berharga dari data yang kompleks dan berjumlah besar melalui teknik-teknik statistika, kecerdasan buatan, dan pembelajaran mesin (*Machine learning*). Dalam konteks *E-commerce*, data mining dapat digunakan untuk menganalisis perilaku konsumen melalui penggalian pola-pola tersembunyi yang dapat memberikan gambaran mendalam mengenai preferensi pembelian, waktu optimal untuk melakukan promosi, dan segmentasi pasar yang lebih tepat. Dengan pemahaman yang lebih baik mengenai perilaku konsumen, Shopee dapat merancang strategi pemasaran yang lebih efektif dan efisien,

meningkatkan loyalitas pelanggan, serta memaksimalkan penjualan. Di sisi lain, munculnya konsep personalisasi dalam pemasaran digital semakin mendorong pentingnya analisis perilaku konsumen. Konsumen modern mengharapkan pengalaman berbelanja yang disesuaikan dengan preferensi dan kebutuhan mereka. Shopee, dengan jumlah pengguna yang terus bertambah, dapat memanfaatkan data perilaku konsumen yang diperoleh melalui platformnya untuk memberikan rekomendasi produk yang lebih relevan, merancang kampanye promosi yang disesuaikan dengan kebiasaan belanja konsumen, serta mengoptimalkan pengalaman pengguna secara keseluruhan. Dengan demikian, personalisasi berbasis data menjadi faktor kunci dalam memenangkan persaingan di pasar *E-commerce*. Namun, tantangan dalam implementasi *data mining* tidak dapat diabaikan. Ketersediaan data yang sangat besar dan bervariasi (*Heterogen*) menimbulkan masalah dalam hal kualitas data, seperti data yang tidak lengkap atau data yang tidak konsisten. Selain itu, proses pemilihan metode data mining yang tepat juga menjadi isu penting, mengingat berbagai teknik yang tersedia seperti *clustering*, *association rule*, dan *decision tree*, yang masing-masing memiliki kelebihan dan keterbatasan dalam konteks analisis perilaku konsumen. Oleh karena itu, studi ini berfokus pada eksplorasi dan penerapan metode data *clustering* untuk mengidentifikasi pola perilaku konsumen di Shopee, serta bagaimana hasil analisis tersebut dapat digunakan untuk meningkatkan strategi pemasaran dan penjualan platform. Penelitian ini diharapkan mampu memberikan kontribusi baik secara teoretis maupun praktis. Dari segi akademis, penelitian ini memperkaya literatur mengenai penerapan data mining di industri *e-commerce*, khususnya dalam analisis perilaku konsumen. Dari segi praktis, hasil penelitian ini diharapkan dapat memberikan panduan bagi Shopee dalam menyusun strategi pemasaran berbasis data, sehingga dapat meningkatkan performa penjualan dan kepuasan pelanggan.

II. TINJAUAN PUSTAKA

A. Data Mining

Data mining adalah proses penggalian data besar untuk mengidentifikasi pola, anomali, dan korelasi yang dapat membantu dalam pengambilan keputusan strategis. Menurut Han, Pei, dan Kamber (2011), data mining menggunakan berbagai teknik analisis data, seperti klasifikasi, regresi, asosiasi, dan clustering, untuk menemukan informasi yang bermakna dan tersembunyi dalam kumpulan data besar. *Data mining* sangat berperan penting dalam dunia *E-commerce*, di mana jutaan transaksi dilakukan setiap hari dan data konsumen yang besar dan variatif tersedia. Dalam konteks analisis perilaku konsumen di Shopee, data

mining memfasilitasi proses penggalian data transaksi pengguna untuk memahami pola belanja, respons terhadap promosi, dan kebiasaan pembelian. Informasi ini menjadi dasar bagi pengembangan strategi pemasaran yang lebih efektif, personalisasi promosi, serta segmentasi konsumen.

Fungsi Data Mining

a. Deskripsi (*Description*)

Menurut (Han, Kamber and Pei, 2014) deskripsi bertujuan untuk mengidentifikasi pola yang muncul secara berulang pada suatu data dan mengubah pola tersebut menjadi aturan dan kriteria yang dapat mudah dimengerti oleh para ahli pada domain aplikasinya. Aturan yang dihasilkan harus mudah dimengerti agar dapat dengan efektif meningkatkan tingkat pengetahuan (*knowledge*) pada sistem.

b. Estimasi (*Estimation*)

Menurut (Han, Kamber and Pei, 2014) Estimasi hampir serupa dengan prediksi, kecuali bahwa variabel target untuk estimasi adalah numerik daripada kategori. Sebuah model dibangun menggunakan dataset lengkap yang menyediakan nilai-nilai variabel target sebagai prediksi. Pengecekan selanjutnya juga membuat estimasi variabel target berdasarkan nilai variabel prediktor. Misalnya, memperkirakan tekanan darah sistolik pasien rawat inap berdasarkan usia pasien, jenis kelamin, berat badan, dan kadar natrium darah. Hubungan antara tekanan darah sistolik dengan nilai prediktor dalam proses pembelajaran menghasilkan model estimasi.

c. Prediksi (*Prediction*)

Menurut (Han, Kamber and Pei, 2014) prediksi mirip dengan klasifikasi, tetapi data diklasifikasikan berdasarkan perilaku atau nilai masa depan yang diharapkan. Contoh tugas peramalan adalah memprediksi jumlah pelanggan yang berkurang dalam waktu dekat dan memprediksi harga saham selama tiga bulan ke depan.

d. Klasifikasi (*Classification*)

Menurut (Han, Kamber and Pei, 2014) Klasifikasi adalah suatu bentuk analisis data yang menciptakan model yang menggambarkan kelas-kelas penting dari data. Klasifikasi memprediksi kategori (diskrit, tidak berurutan) dari label kelas. Klasifikasi adalah proses menemukan model atau fungsi yang menggambarkan atau membedakan konsep atau kelas dalam data, dengan tujuan untuk dapat menyimpulkan kelas suatu objek yang peruntukannya tidak diketahui. Model itu sendiri dapat berupa aturan *if-then-rules*, *decision tree*, formula matematis atau *neural network*. Contohnya, membuat model klasifikasi untuk mengklasifikasikan aplikasi pinjaman bank sebagai aman atau berisiko. Analisis semacam itu dapat membantu lebih memahami data secara umum. Klasifikasi memiliki banyak kegunaan, termasuk

deteksi penipuan, pemasaran bertarget, prediksi kinerja, manufaktur, dan diagnostik medis.

e. Pengelompokan (*Clustering*)

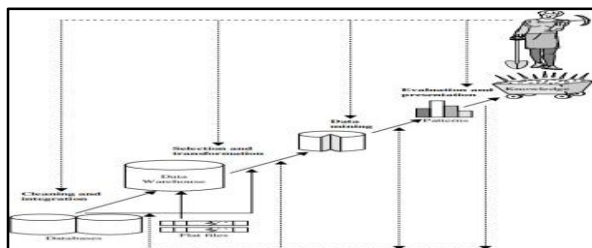
Menurut (Noviyanto, 2020) *Clustering* merujuk pada pengelompokan dokumen, observasi atau kasus pada kelas yang objeknya mirip. Klaster adalah kumpulan dokumen yang mirip satu sama lain dan berbeda dengan dokumen pada klaster lain. *Clustering* berbeda dengan *Clasification*, pada *Clustering* tidak ada target variabel untuk dikelompokkan. Algoritma *Clustering* mencoba untuk membagi kumpulan data menjadi klaster yang anggotanya relatif sama, dimana kemiripan dokumen di klaster yang sama tinggi, dan kemiripan dokumen di klaster lain kecil.

f. Asosiasi (*Association*)

Menurut (Han, Kamber and Pei, 2014) Fungsi dari *Association* adalah untuk mencari keterkaitan antara atribut atau item sel, berdasarkan item yang muncul dan rule association yang ada.

Knowledge Discovery From Data (KDD)

Menurut (Han, Kamber and Pei, 2014) pada metodologi KDD memiliki tahapan seperti pada gambar 1.



Gambar 1. Tahapan KDD (Han, Kamber and Pei, 2014)

Gambar 1 merupakan tahapan-tahapan yang terdapat pada KDD antara lain data cleaning, data integration, data selection, data transformation, data mining, pattern evaluation dan knowledge presentation dapat dijelaskan sebagai berikut:

Data Cleaning merupakan tahapan yang melakukan pemilihan data yang relevan dari database dengan melakukan pemisahan terhadap data yang tidak konsisten dan data yang tidak relevan. *Data Integration* merupakan tahapan yang dilakukan integrasi terhadap data yang ada dengan cara menggabungkan berbagai sumber data menjadi satu sumber. *Data Selection* merupakan tahapan yang melakukan pemilihan terhadap data yang relevan dengan analisa yang akan dilakukan pada database. *Data Transformation* merupakan tahapan yang dilakukan pengubahan terhadap format data yang ada menjadi format data yang sesuai untuk diproses dalam data mining. *Data Mining* merupakan tahapan yang dilakukan proses data mining, dengan menerapkan metode tertentu untuk mendapatkan informasi yang tersembunyi dari data yang ada. *Pattern Evaluation* merupakan tahapan yang

dilakukan identifikasi terhadap pola-pola yang menarik yang didapat dari hasil data mining, untuk kemudian direpresentasikan. *Knowledge Presentation* merupakan tahapan yang dilakukan visualisasi dan penyajian terhadap pengetahuan mengenai teknik yang digunakan untuk memperoleh pengetahuan yang diperoleh user.

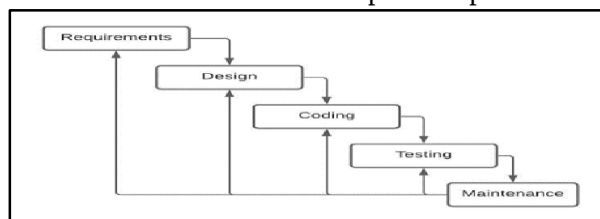
Clustering

Clustering adalah teknik dalam data mining yang berfungsi untuk mengelompokkan data menjadi beberapa cluster berdasarkan kesamaan antar-data. Menurut Kaufman dan Rousseeuw (2009), *clustering* membagi dataset tanpa label menjadi beberapa grup, dengan asumsi bahwa data dalam satu grup memiliki kesamaan yang lebih tinggi dibandingkan dengan data di grup lain. Teknik *clustering* umum digunakan dalam analisis eksploratif, seperti segmentasi pelanggan, karena tidak memerlukan label pada data.

III. METODOLOGI PENELITIAN

Metode Penelitian

Metode penelitian yang diterapkan pada penelitian ini adalah dengan pengembangan metode waterfall. Metode waterfall memiliki tahapan-tahapan



Gambar 2. Tahapan Metode Waterfall

Keterangan:

1) Requirement Analysis (Analisis Kebutuhan)

Tahap ini mencakup pengumpulan dan analisis kebutuhan pengguna yang kemudian didokumentasikan dalam *Software Requirement Specification* (SRS). Contohnya, dalam pengembangan aplikasi e-commerce, tahap ini menentukan fitur seperti pencarian produk dan metode pembayaran.

2) System Design (Perancangan Sistem)

Berdasarkan spesifikasi kebutuhan, sistem dirancang mencakup arsitektur, basis data, dan antarmuka pengguna. Hasilnya berupa dokumen desain yang menjadi panduan pengembang. Sebagai contoh, diagram *Entity-Relationship Diagram* (ERD) dibuat untuk mengelola data pelanggan dan produk.

3) Implementation (Implementasi/Koding)

Pada tahap ini, pengembang menulis kode sesuai desain yang telah dibuat. Setiap modul dikembangkan dan diuji sebelum digabungkan dalam sistem. Contohnya, pembuatan fitur pencarian produk agar pengguna dapat menemukan barang dengan mudah.

4) Testing (Pengujian)

Pengujian dilakukan untuk memastikan sistem berfungsi sesuai spesifikasi. Metode yang digunakan

meliputi *unit testing*, *integration testing*, dan *user acceptance testing* (UAT). Misalnya, pengujian fitur pembayaran dalam aplikasi e-commerce untuk memastikan transaksi berjalan lancar.

5) *Deployment* (Penerapan dan Implementasi)

Setelah pengujian berhasil, perangkat lunak diterapkan dan digunakan oleh pengguna. Proses ini dapat dilakukan secara bertahap (*pilot deployment*) atau sekaligus (*full deployment*). Contohnya, peluncuran aplikasi e-commerce ke server produksi atau platform digital.

6) *Maintenance* (Pemeliharaan)

Tahap ini bertujuan untuk memastikan sistem tetap optimal melalui perbaikan bug, pembaruan fitur, dan peningkatan performa. Contohnya, perbaikan sistem checkout jika ditemukan kendala setelah peluncuran aplikasi.

Metode Pengambilan Sampel

Penelitian ini menggunakan *purposive sampling* untuk memilih data transaksi berdasarkan periode tertentu yang dianggap representatif. Pengambilan sampel didasarkan pada rentang waktu mingguan dari Januari 2023 hingga Juli 2023, yang mencakup minggu-minggu promosi dan non-promosi agar dapat menggambarkan pola belanja konsumen sepanjang tahun.

Kriteria Pemilihan Sampel:

1. Kriteria Waktu: Data penghasilan diambil secara mingguan selama satu tahun penuh untuk mencakup berbagai musim belanja, termasuk masa promosi besar, hari-hari libur, dan periode non-promosi.
2. Kriteria Produk: Penelitian ini difokuskan pada data transaksi produk kesehatan hewan, khususnya produk untuk kucing dan anjing. Kategori produk ini terkait dengan kesehatan dan kebersihan hewan peliharaan. Pemilihan kategori ini didasarkan pada meningkatnya tren kepemilikan hewan peliharaan di masyarakat, yang diikuti dengan kebutuhan untuk menjaga kesehatan dan kebersihan hewan.
3. Kriteria Promosi: Menggunakan data transaksi yang mencakup periode promosi dan periode normal untuk menganalisis perbedaan perilaku konsumen. Data promosi terdiri dari transaksi yang menggunakan voucher diskon atau potongan harga.

Metode Pengumpulan Data

Penelitian ini melibatkan pengumpulan data primer dan sekunder yang bersumber dari sistem e-commerce Shopee, berfokus pada toko "Anjani_store.id." Data mencakup catatan transaksi mingguan seperti subtotal pesanan, potongan harga, biaya administrasi, biaya layanan, dan penghasilan bersih, yang penting untuk memahami pola penghasilan dan perilaku belanja konsumen dalam periode yang ditentukan.

Data Primer

Data primer pada penelitian ini diambil secara manual dari toko "Anjani_store.id" di platform Shopee. Proses pengumpulan data dilakukan dengan mengakses dashboard Shopee Seller Center dan mencatat data transaksi mingguan, yang kemudian diolah lebih lanjut untuk analisis. Data yang diambil meliputi:

Subtotal Pesanan: Total nilai pesanan mingguan sebelum potongan harga dan biaya, yang mencerminkan volume penjualan. Potongan Harga: Jumlah diskon dan voucher yang digunakan konsumen setiap minggu, yang menggambarkan dampak promosi. Biaya Lainnya: Biaya administrasi, layanan, dan pengiriman per transaksi. Total Penghasilan: Penghasilan bersih toko setelah semua potongan dan biaya, menjadi variabel utama untuk mengukur pendapatan mingguan.

Prosedur pengumpulan data primer secara manual dilakukan melalui langkah-langkah berikut:

1. Akses Data Transaksi Shopee Seller Center: Mengakses akun "Anjani_store.id" di Shopee Seller Center setiap minggu untuk mencatat transaksi terbaru dalam bentuk laporan penjualan.
2. Ekstraksi Data Transaksi: Data transaksi diunduh dalam format Excel atau CSV jika Shopee menyediakan opsi tersebut; jika tidak, data dicatat secara manual dan disimpan dalam perangkat lunak pengolah data seperti Microsoft Excel atau Google Sheets.
3. Format Data: Data yang diunduh atau dicatat secara manual kemudian disusun dalam format tabel (data frame) yang terstruktur di Python atau R untuk memastikan kemudahan analisis. Setiap variabel dipisahkan dalam kolom agar siap untuk diolah.
4. Pembersihan Data (Data Cleaning): Data diperiksa untuk menghapus ketidaksesuaian seperti data duplikat, data kosong, atau outliers. Missing values ditangani dengan metode imputasi (mengisi dengan nilai rata-rata atau nilai median), dan outliers diperiksa menggunakan z-score atau boxplot untuk mengidentifikasi nilai ekstrim.
5. Transformasi Data: Variabel-variabel yang menunjukkan jumlah dalam Rupiah ditransformasikan menggunakan teknik normalisasi seperti Min-Max Scaling agar skala data lebih konsisten dan mudah dibandingkan.
6. Pengkodean Data Promosi: Diskon atau voucher yang berlaku diidentifikasi secara manual dengan memberikan kode khusus untuk minggu-minggu promosi. Pengkodean ini membantu memisahkan data promosi dari non-promosi untuk analisis perilaku konsumen.

IV. HASIL DAN PEMBAHASAN

A. Hasil

Aplikasi pengelompokan data pendapatan mingguan konsumen berdasarkan hasil penjualan di platform Shopee telah berhasil dibangun menggunakan bahasa pemrograman Python dengan framework Flask serta database MySQL. Proses analisis dilakukan menggunakan algoritma data mining yaitu *K-Means Clustering*. Pengelompokan dilakukan berdasarkan data riil pendapatan mingguan selama tahun 2023 yang diperoleh dari toko Shopee. Data tersebut mencakup subtotal pesanan, voucher dan diskon, biaya lainnya, serta total penghasilan. Melalui aplikasi ini, sistem dapat mengelompokkan periode penjualan ke dalam klaster-klaster yang mewakili pendapatan rendah, sedang, dan tinggi. Tujuan dari pengelompokan ini adalah untuk membantu pemilik toko dalam menganalisis perilaku konsumen berdasarkan fluktuasi pendapatan, serta sebagai dasar evaluasi strategi pemasaran dan promosi. Klasterisasi data dilakukan dengan membagi data ke dalam 3 kelompok, yaitu:

1. Cluster 1: Pendapatan Rendah
2. Cluster 2: Pendapatan Sedang
3. Cluster 3: Pendapatan Tinggi

Proses ini diawali dengan normalisasi data dan perhitungan *elbow method* untuk menentukan jumlah klaster optimal. Selanjutnya, dilakukan pengolahan data menggunakan metode K-Means dengan iterasi maksimal tertentu yang ditentukan pengguna. Hasil klasterisasi ditampilkan dalam bentuk tabel dan visualisasi grafik pie chart. Setiap klaster memuat minggu-minggu dengan nilai pendapatan yang serupa. Informasi ini dapat digunakan sebagai dasar pengambilan keputusan strategi promosi atau diskon yang lebih tepat sasaran. Adapun hasil aplikasi pengelompokan data pasien berdasarkan data pendapatan toko Anjani_Store.id, antara lain:

1. Menu Login

Menu login merupakan form yang digunakan untuk membatasi akses penggunaan aplikasi melalui username dan password. Sehingga admin harus memasukkan username dan password yang benar agar dapat masuk ke menu utama agar dapat mengakses secara keseluruhan dari aplikasi pengelompokan data pendapatan di Toko Anjani_Store.Id.

Gambar 3. Menu Login

2. Menu Utama

Menu utama merupakan form menu yang memiliki beberapa sub menu yang dapat diakses oleh admin untuk melakukan pengolahan data pada aplikasi pengelompokan data pasien berdasarkan data pendapatan di toko Anjani_Store.id yang memiliki fungsi yang berbeda-beda.

Gambar 4. Menu Utama

3. Input Data Pendapatan

Merupakan form yang digunakan untuk mengolah data pendapatan toko Anjani_Store.id. Pengelolaan data dapat dilakukan dengan mengisi field yang telah disediakan dan dapat melakukan proses tambah, koreksi, hapus data pada form.

Gambar 5 Input Data Pendapatan

4. Form Klasterisasi K-Means

Klasterisasi Algoritma K-Means merupakan antarmuka aplikasi klasterisasi data pendapatan yang digunakan untuk menjalankan proses pengelompokan data pendapatan menjadi 3 kelompok yaitu Cluster C1, Cluster C2, dan Cluster C3. Dalam proses klasterisasi tersebut akan diberikan informasi jumlah iterasi dan waktu proses yang terjadi.

Gambar 6.. Klasterisasi Algoritma K-Means

5. Output Laporan Hasil Klasterisasi

Merupakan output yang menampilkan informasi hasil klasterisasi data pendapatan.



Gambar 7. Output Hasil Klastering

B. Pembahasan

Dalam pembahasan ini, akan dibahas analisis klasterisasi pendapatan perminggu berdasarkan Metode K-Means, Dimana diambil sampel data pendapatan sebanyak 30 minggu data yang diperoleh dari toko Anjani_Store.id.

Minggu	TOTAL PENGHASILAN
2 Januari - 8 Januari	2483515
9 Januari - 15 Januari	25475322
16 Januari - 22 Januari	25445967
23 Januari - 29 Januari	20911435
30 Januari - 5 Februari	20920100
6 Februari - 12 Februari	23241334
13 Februari - 19 Februari	24127921
20 Februari - 26 Februari	218427481
27 Februari - 5 Maret	20811435
6 Maret - 12 Maret	21209201
13 Maret - 19 Maret	17500030
20 Maret - 26 Maret	20487337
27 Maret - 2 April	22113422
3 April - 9 April	21812872
10 April - 16 April	21812872
17 April - 23 April	21812872
24 April - 30 April	21812872
1 Mei - 7 Mei	24011502
8 Mei - 14 Mei	26356415
15 Mei - 21 Mei	27761334
22 Mei - 28 Mei	26356415

Gambar 8. Data Pendapatan Toko

Berdasarkan data pendapatan toko pada Gambar 9. tersebut, dilakukan klasterisasi menggunakan metode K-Means.



Gambar 9. Hasil Klasterisasi Metode K-Means

Dari Gambar 9. tersebut hasil dari klasterisasi disimpan. Dapat dilihat masing masing data sudah dibagi sesuai dengan clusternya masing masing agar data pendapatan dapat lebih mudah untuk di analisis. Adapun hasil dari clusterisasi semua data menggunakan metode K-Means, Seperti Gambar 4.8., 4.9., 4.10.

No	Minggu	Total Penghasilan	Cluster
1	20 Maret - 26 Maret	17960836	Rendah
2	10 April - 16 April	19047897	Rendah
3	17 April - 23 April	17436618	Rendah
4	22 Mei - 28 Mei	16159779	Rendah
5	29 Mei - 4 Juni	19633700	Rendah
6	12 Juni - 18 Juni	15685567	Rendah
7	26 Juni - 2 Juli	19732307	Rendah
8	17 Juli - 23 Juli	15874669	Rendah
9	24 Juli - 30 Juli	13356570	Rendah

Gambar 10. Hasil Clustering Pendapatan “Rendah”

No	Minggu	Total Penghasilan	Cluster
1	2 Januari - 8 Januari	25483515	Sedang
2	9 Januari - 15 Januari	24475322	Sedang
3	16 Januari - 22 Januari	24445967	Sedang
4	23 Januari - 29 Januari	20911435	Sedang
5	30 Januari - 5 Februari	20920100	Sedang
6	6 Februari - 12 Februari	23241334	Sedang
7	13 Februari - 19 Februari	24127921	Sedang
8	20 Februari - 26 Februari	218427481	Sedang
9	27 Februari - 5 Maret	20811435	Sedang
10	6 Maret - 12 Maret	21209201	Sedang
11	13 Maret - 19 Maret	17500030	Sedang
12	20 Maret - 26 Maret	20487337	Sedang
13	27 Maret - 2 April	22113422	Sedang
14	3 April - 9 April	21812872	Sedang
15	10 April - 16 April	21812872	Sedang
16	17 April - 23 April	21812872	Sedang
17	24 April - 30 April	21812872	Sedang
18	1 Mei - 7 Mei	24011502	Sedang
19	8 Mei - 14 Mei	26356415	Sedang
20	15 Mei - 21 Mei	27761334	Sedang
21	22 Mei - 28 Mei	26356415	Sedang

Gambar 12. Hasil Clustering Pendapatan “Sedang”

No	Minggu	Total Penghasilan	Cluster
1	30 Januari - 5 Februari	28938100	Tinggi
2	27 Februari - 5 Maret	30466577	Tinggi
3	27 Maret - 2 April	30547337	Tinggi
4	3 April - 9 April	27112872	Tinggi
5	8 Mei - 14 Mei	26356415	Tinggi
6	3 Juli - 9 Juli	27973543	Tinggi
7	10 Juli - 16 Juli	26754759	Tinggi

Gambar 13. Hasil Clustering Pendapatan “Tinggi”

Disimpulkan bahwa berdasarkan hasil pengelompokkan terlihat bahwasannya pendapatan tinggi seringkali terjadi disaat awal bulan dan berlaku sama dengan pendapatan sedang. Begitu juga sebaliknya, pendapatan rendah seringkali terjadi di pertengahan dan menjelang akhir bulan.

Hal ini terjadi karena beberapa faktor antara mulai dari voucher yang disediakan oleh toko dan juga faktor ekonomi seperti kebanyakan pekerja baru menerima gaji di awal bulan.

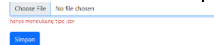
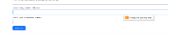
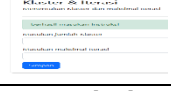
Maka dari itu dapat disimpulkan bahwa agar penjualan lebih optimal toko dapat meluncurkan promo di awal bulan agar alur pendapatan tinggi akan lebih meningkat. Dan begitu pula sebaliknya, agar pendapatan lebih stabil menjelang pertengahan dan akhir bulan toko dapat meluncurkan promo diskon agar lebih dapat menjangkau daya beli masyarakat saat pertengahan dan menjelang akhir bulan.

C. Hasil Pengujian

Pengujian blackbox (blackbox testing) adalah salah metode pengujian perangkat lunak yang berfokus pada sisi fungsionalitas, khususnya pada input dan output aplikasi pengelompokkan data pendapatan toko Anjani_Store.id. Adapun hasil pengujian blackbox yang telah dilakukan

Tabel 1. Tabel Pengujian

Form Uji	Komponen Yang Diuji	Hasil Pengujian BlackBox	Kesimpulan
Login	Field username dan password dikosongkan	System menolak akses login	Berhasil
	Mengisi username yang salah	System menolak akses login	Berhasil
	Mengisi password yang salah	System menolak akses login	Berhasil
	Mengisi username dan password yang benar	System menerima akses login	Berhasil

Input Data Penghasilan	Menyimpan data tanpa menginput file	System tidak dapat menampilkan data 	Berhasil
	Menyimpan data penghasilan setelah menginput file	System dapat menampilkan data 	Berhasil
Klaster dan iterasi	Klaster dan iterasi dikosongkan	System menolak instruksi 	Berhasil
	Klaster dan iterasi di isi	System menerima instruksi 	Berhasil
Proses Clustering	Melakukan proses klasterisasi data pendapatan	System berhasil melakukan klastering 	Berhasil
Hasil Klastering	Memvisualisasikan hasil klastering	System berhasil menampilkan hasil klasterisasi dan memberikan informasi pengelompokan pendapatan perminggu 	Berhasil

Berdasarkan hasil pengujian yang telah dilakukan, aplikasi pengelompokan data pendapatan berdasarkan pendapatan toko Anjani_Store.id berhasil dilakukan, dan dapat memberikan informasi berdasarkan 3 kelompok yaitu Cluster C1 (Pendapatan Rendah), Cluster C2 (Pendapatan Sedang). Dan Cluster C3 (Pendapatan Tinggi), serta fungsional aplikasi telah berjalan sesuai dengan yang diharapkan.

V. PENUTUP

A. Kesimpulan

Berdasarkan hasil dari pembahasan dan pengujian yang telah dilakukan, maka dapat disimpulkan bahwa:

1. Pengelompokan data pendapatan berdasarkan pendapatan toko Anjani_Store.id dapat digunakan untuk mengetahui pendapatan mana yang masuk ke dalam kelompok pendapatan rendah (Cluster C1), pendapatan sedang (Cluster C2), dan pendapatan tinggi (Cluster C3).

2. Berdasarkan hasil pengelompokan dapat diketahui bahwasannya pendapatan tinggi cenderung ada di awal bulan, serta pendapatan rendah dan sedang cenderung ada di akhir dan pertengahan bulan.

D. Saran

Berdasarkan penelitian yang penulis lakukan, maka penulis menyarankan perlu dilakukan pengembangan pada penelitian selanjutnya dengan menambahkan variable-variabel lainnya agar perbandingan dan hasil yang didapatkan jauh lebih kompleks.

DAFTAR PUSTAKA

- [1] Han, J., Pei, J., & Kamber, M. (2021). *Data Mining: Concepts and Techniques* (4th ed.). Morgan Kaufmann..
- [2] Putra, A. G., & Raharjo, B. (2021). Analisis Perilaku Konsumen Online Menggunakan Data Mining Metode Clustering. *Jurnal Manajemen Informatika (JMI)*, 12(1), 33–42.
- [3] Jaiswal, A., & Yadav, V. (2021). Customer Segmentation Using K-Means Clustering. *International Journal of Computer Applications*, 183(18), 10–14.
- [4] Kurniawan, R. A., & Siregar, H. (2023). Implementasi K-Means Clustering untuk Menentukan Tingkat Konsumsi Konsumen. *Jurnal Sains dan Teknologi Informasi*, 12(1), 45–51.
- [5] Saputra, I. H. (2020). Penerapan Algoritma K-Means untuk Analisis Pola Pembelian Konsumen pada Marketplace. *Jurnal Teknologi Informasi dan Komputer*, 6(2), 123–130.
- [6] Shopee Indonesia. (2023). *Laporan Penjualan dan Transaksi Shopee Seller Center Tahun 2023*. Jakarta: Shopee.
- [7] Shopee Indonesia. (2024). *Shopee API Documentation*. Diakses pada 5 Desember 2024, dari <https://developers.shopee.co.id>.
- [8] Sari, P. A., & Suhartono, D. (2020). Data Mining dalam E-Commerce: Review dan Trend Penelitian. *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)*, 7(1), 1–9.
- [9] Sugiyono. (2020). *Metode Penelitian Kuantitatif, Kualitatif, dan R&D*. Bandung: Alfabeta.
- [10] Suryani, T., Hidayatullah, S., & Anggraini, R. (2021). Analisis Perilaku Konsumen dalam Belanja Online pada Masa Pandemi COVID-19. *Jurnal Ekonomi dan Bisnis*, 12(3), 178–190.