

# Penerapan Metode K-Means Dalam Penjualan Kain Sasirangan Pada UMKM “Sasirangan Bunda Maburai”

Ade Putri Maharani <sup>1,\*</sup>, Finki Dona Marleny <sup>2</sup>, Kamarudin <sup>3</sup>, Rudy Ansari <sup>4</sup>

<sup>1,2,3,4</sup> Program Studi Informatika; Universitas Muhammadiyah Banjarmasin; Jl. Gubernur Syarkawi, Semangat Dalam, Kec. Alalak, Kabupaten Barito Kuala, Kalimantan Selatan 70581, (0511) 3363002; e-mail: [adeputrimaharani09@gmail.com](mailto:adeputrimaharani09@gmail.com), [finkidona@umbjm.ac.id](mailto:finkidona@umbjm.ac.id), [kamarudin.skom@gmail.com](mailto:kamarudin.skom@gmail.com), [rudy@umbjm.ac.id](mailto:rudy@umbjm.ac.id)

\* Korespondensi: e-mail: [adeputrimaharani09@gmail.com](mailto:adeputrimaharani09@gmail.com)

Diterima: 2 Maret 2025; Review: 20 Mei 2025; Disetujui: 27 Mei 2025

Cara sitasi: Maharani AP, Marleny FD, Kamarudin, Ansari R. 2025. Penerapan Metode K-Means Dalam Penjualan Kain Sasirangan Pada UMKM “Sasirangan Bunda Maburai”. *Informatics for Educators and Professionals : Journal of Informatics*. Vol.10 (1): 34-46.

**Abstrak:** Penelitian ini mengkaji penerapan algoritma *K-means clustering* pada data penjualan kain Sasirangan di UMKM "Sasirangan Bunda Maburai" dengan pendekatan analitik berbasis kombinasi variabel yang belum banyak diterapkan dalam konteks UMKM tradisional. Metode penelitian menggunakan pendekatan KDD (*Knowledge Discovery in Databases*) dengan melakukan pengolahan data penjualan dari Januari 2020 hingga Juli 2024, pembersihan data, transformasi, dan implementasi *K-means clustering*. Kebaruan dalam penelitian ini terletak pada eksplorasi enam kombinasi atribut kain (jenis kain, motif, warna dasar, dan warna motif) untuk menemukan pola preferensi konsumen yang dapat mendukung pengambilan keputusan manajemen stok berbasis data. Kombinasi jenis kain–motif menunjukkan performa terbaik dengan *silhouette score* 0,65 dan tiga cluster optimal, disusul warna dasar–warna motif (*score* 0,63). Hasil ini memberikan kontribusi bagi pendekatan *data-driven* dalam pengelolaan produksi dan stok pada UMKM berbasis kearifan lokal.

**Kata kunci:** K-means, *Clustering*, Data Mining, UMKM Kain Sasirangan, *silhouette score*

**Abstract:** This study examines the application of the *K-means clustering* algorithm on Sasirangan fabric sales data in UMKM "Sasirangan Bunda Maburai" with an analytical approach based on a combination of variables that has not been widely applied in the context of traditional UMKM. The research method uses the KDD (*Knowledge Discovery in Databases*) approach by processing sales data from January 2020 to July 2024, data cleaning, transformation, and implementation of *K-means clustering*. The novelty in this study lies in the exploration of six combinations of fabric attributes (fabric type, motif, base color, and motif color) to find patterns of consumer preferences that can support data-driven stock management decision-making. The combination of fabric-motif types showed the best performance with a *silhouette score* of 0.65 and three optimal clusters, followed by the base color of the motif (*score* 0.63). These results contribute to a *data-driven* approach in managing production and stock in UMKM based on local wisdom.

**Keywords:** *K-means*, *Clustering*, *Data Mining*, *UMKM Fabric Sasirangan*, *silhouette score*

## 1. Pendahuluan

Kain sasirangan merupakan warisan budaya dari Provinsi Kalimantan Selatan yang memiliki arti simbolis yang mendalam bagi masyarakat banjar. Pada awalnya sasirangan digunakan untuk acara adat istiadat dari beberapa kalangan, namun semakin berkembangnya zaman, kini sasirangan berubah menjadi cendramata atau produk khas daerah Kalimantan.

Pada tanggal 07 Januari 2016 pemerintah setempat berhasil mendirikan UMKM “Sasirangan Bunda Maburai” dengan melakukan pelatihan pembuatan kain sasirangan tradisional, meningkatkan keterampilan dan kualitas produk yang di olah, serta menjadi sumber ekonomi bagi ibu-ibu Desa Maburai sekaligus dapat melestarikan warisan budaya daerah Kalimantan [1].

Semakin berkembangnya zaman, semakin ketat pula persaingan yang dihadapi UMKM “Sasirangan Bunda Maburai” dalam memproduksi produk sasirangan. Dalam hal ini, UMKM memerlukan sistem manajemen persediaan yang efektif untuk mengelola stok produk kain sasirangan[2]. Penerapan metode analisis data menjadi pertolongan pertama untuk mengidentifikasi pola penjualan dan preferensi konsumen[3]. Pemilihan metode K-means clustering sebagai salah satu teknik data mining yang bersifat *unsupervised learning* menjadi solusi karena algoritma ini mampu mengelompokkan produk kain sasirangan berdasarkan beberapa karakteristik tanpa memerlukan label kategori terlebih dahulu [4].

Meskipun metode K-means telah banyak digunakan dalam berbagai studi terkait segmentasi produk, sebagian besar penerapannya masih terbatas pada sektor retail modern dan produk digital. Belum banyak penelitian yang secara spesifik menerapkan metode ini pada UMKM berbasis produk budaya lokal seperti kain sasirangan. Hal ini menunjukkan adanya *theoretical gap* dalam pemanfaatan data mining untuk optimalisasi manajemen stok di sektor ekonomi kreatif berbasis warisan budaya.

Tujuan dari penelitian ini adalah menganalisis efisiensi penggunaan algoritma K-means clustering pada data penjualan kain sasirangan dengan membandingkan berbagai macam kombinasi variabel seperti jenis kain, motif, warna dasar, dan warna motif. Penelitian ini akan mengidentifikasi jumlah optimal cluster untuk setiap kombinasi variabel yang diuji serta menganalisis performa masing-masing cluster menggunakan enam kombinasi variabel, yaitu: jenis kain - motif, jenis kain - warna dasar, motif - warna dasar, motif - warna motif, warna dasar - warna motif, dan jenis kain - warna motif. Selanjutnya, penelitian ini akan membandingkan kualitas cluster yang diperoleh dari masing-masing kombinasi dengan menggunakan metode *cross validation*.

Melalui penelitian ini Penelitian ini penting dilakukan karena dapat membantu UMKM dalam mengelola stok secara lebih efisien, mengurangi penumpukan barang yang tidak laku, serta menyusun strategi produksi yang sesuai dengan preferensi konsumen berdasarkan analisis data historis. Dalam jangka panjang, ini akan meningkatkan daya saing UMKM berbasis budaya di tengah persaingan industri tekstil modern [5].

## 2. Metode Penelitian

Dalam penelitian ini digunakan metode KDD (*Knowledge Discovery in Databases*) sebagai pendekatan analitis untuk mengolah dan menganalisis data penjualan kain sasirangan. Metode ini dipilih karena kemampuannya dalam mengekstrak informasi berharga dari kumpulan data besar melalui serangkaian tahapan sistematis. KDD melibatkan proses pengumpulan dataset, pembersihan data, integrasi data dari berbagai sumber, seleksi atribut yang relevan, transformasi data numerik, penerapan algoritma *k-means clustering*, dan validasi hasil menggunakan *cross validation*. Pendekatan ini memungkinkan peneliti mengidentifikasi pola penjualan dan preferensi konsumen terhadap produk kain sasirangan secara akurat dan terukur [6].

Penerapan KDD berkaitan erat dengan proses data mining, yang merupakan inti dari proses ekstraksi informasi dari kumpulan data bervolume besar. Tahapan data mining meliputi pengumpulan, ekstraksi, analisis, dan visualisasi statistik data untuk menghasilkan wawasan bermakna. Proses ini sering disebut juga sebagai penemuan pengetahuan dalam basis data, ekstraksi pengetahuan, analisis pola data, atau pengumpulan informasi, tergantung konteks penggunaannya [7].

Algoritma *k-means clustering* menjadi metode utama dalam penelitian ini karena karakteristiknya sebagai pembelajaran tanpa pengawasan (*unsupervised learning*), yang berarti pengelompokan data dilakukan tanpa memerlukan label atau kategori awal [8]. Teknik ini membagi dataset menjadi beberapa kelompok berdasarkan parameter seperti jumlah *cluster* dan nilai centroid [9]. Prosesnya melibatkan perhitungan jarak antar titik, pengelompokan objek berdasarkan kedekatan, dan iterasi berulang hingga mencapai stabilitas dimana anggota *cluster* tidak lagi berpindah [10].

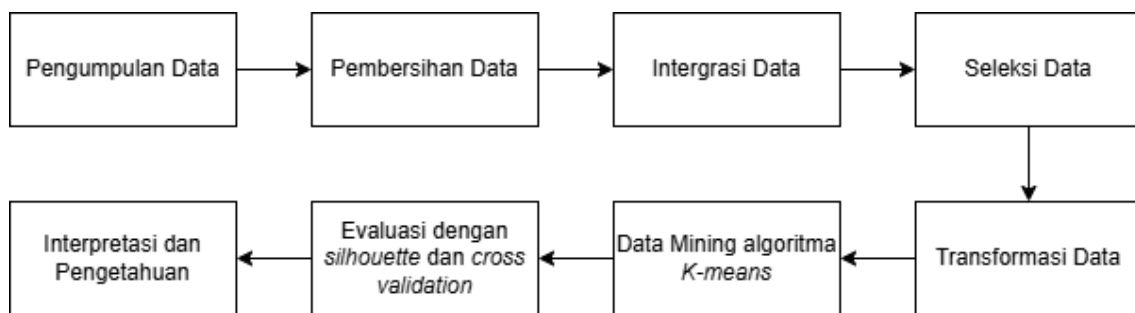
Untuk memastikan keandalan hasil, penelitian ini memanfaatkan model validasi silang (*cross validation*) yang membagi dataset menjadi beberapa bagian proporsional. Metode

validasi ini menjamin konsistensi dan ketahanan model terhadap data baru yang belum pernah diproses sebelumnya, sehingga meningkatkan kepercayaan terhadap hasil analisis [11].

Implementasi teknis dari penelitian ini menggunakan bahasa pemrograman Python yang dikenal dengan fleksibilitas dan kemudahan penggunaannya [12]. Python menyediakan berbagai pustaka untuk analisis data, visualisasi grafis, dan otomatisasi proses yang sangat membantu dalam mengolah data kompleks dari penjualan kain sasirangan pada UMKM "Sasirangan Bunda Maburai" [13].

Penelitian ini memanfaatkan dataset Riwayat transaksi penjualan kain sasirangan dari UMKM "Sasirangan Bunda Maburai" dari bulan Januari 2020 hingga bulan Juli 2024 dengan total 554 data. Berikut ini disajikan tampilan data mentah hasil pengumpulan catatan transaksi penjualan dari UMKM "Sasirangan Bunda Maburai" mencakup beberapa komponen informasi penting, termasuk tanggal produksi kain, spesifikasi produk (jenis kain, motif, dan variasi warna), volume produksi, tanggal penjualan, aspek finansial (harga jual, modal produksi, dan margin keuntungan), serta informasi pelanggan. Keberagaman data ini menyediakan landasan yang kuat untuk analisis pola penjualan dan preferensi konsumen yang menjadi fokus penelitian.

Pengumpulan data mentah ini merupakan langkah awal penting dalam proses KDD (Knowledge Discovery in Databases) yang diterapkan dalam penelitian. Data yang terkumpul selanjutnya akan melalui serangkaian tahapan pengolahan seperti pembersihan, integrasi, seleksi, dan transformasi sebelum diproses menggunakan algoritma *k-means clustering* untuk menghasilkan wawasan bernilai bagi manajemen persediaan UMKM. Adapun untuk tahapan metode penelitian dapat dilihat pada gambar 1 dibawah ini.



Sumber: Hasil Penelitian (2025)

Gambar 1. Tahapan metode penelitian

Proses KDD di mulai dari tahapan pengumpulan dataset dan pembersihan data yang akan digunakan, integrasi data, seleksi data, transformasi data, proses penambangan data dengan algoritma *k-means clustering*, evolusi pola dan validasi data dengan menggunakan cross validation, serta pengetahuan [14]. Berikut ini merupakan penjabaran proses KDD (Knowledge Discovery in Databases) [15]:

#### **Pengelolaan data**

Pengelolaan data merupakan Langkah pertama dari proses KDD. Pada tahapan ini dilakukan pengumpulan data yang akan digunakan untuk proses dataset analisa data.

#### **Pembersihan data (data cleaning)**

Pembersihan data adalah proses penghapusan data yang tidak sesuai dengan tujuan penelitian dan menghapus data yang tidak lengkap atau data yang salah saat penginputan nilai.

#### **Data integration (integrasi data)**

Integrasi data adalah proses penggabungan data dari berbagai sumber dan format seperti berupa gambar, buku, hingga online (excel). Dataset di jadikan satu format agar dataset menjadi utuh dan dapat diolah lebih lanjut.

#### **Seleksi data (selection)**

Seleksi data adalah proses pemilihan atribut atau data yang akan digunakan dalam proses data mining agar hasil analisis sesuai dengan tujuan utama penelitian.

#### **Transformasi data (data transformation)**

Transformasi data adalah proses normalisasi atau standarisasi data. Proses ini melihatkan tahap normalisasi data yang bertujuan untuk memastikan dan mengubah data yang

awalnya berupa teks menjadi numerik, agar data dapat digunakan pada tahapan perhitungan data mining.

**Proses penambangan data (data mining)**

Penambangan data adalah proses menerapkan algoritma atau metode yang telah di tentukan sebelumnya dengan dataset untuk menemukan pola atau informasi pada dataset.

**Validasi data atau evolusi data**

Evolusi pola adalah tahapan pada pola yang telah didapatkan saat proses data mining, yang kemudian pola di evaluasi dan divalidasi bertujuan untuk memastikan apakah hasil sesuai dengan tujuan awal peneliti. Pada penelitian ini menggunakan cross validation atau validasi silang antar dataset guna mengevaluasi performa pemodelan dan mengurangi resiko overfitting.

**Pengetahuan**

Memberikan informasi penting berdasarkan tujuan penelitian.

**3. Hasil dan Pembahasan**

Berikut ini hasil implementasi penerapan KDD pada penelitian ini, sebagai berikut:

**Pengolahan data**

Penelitian ini menggunakan riwayat penjualan kain sasirangan pada UMKM "Sasirangan Bunda Maburai" dari bulan Januari 2020 hingga Juli 2024 dengan total 554 data. Berikut ini disajikan tampilan data mentah hasil pengumpulan catatan transaksi penjualan dari UMKM "Sasirangan Bunda Maburai" pada Gambar 2.

Table 1 Data mentah penjualan UMKM

No	Tanggal	Produksi			Kain keluar		Harga		Modal		keuntungan	Nama pemesan		
		Jenis kain	Motif	Warna		Jumlah	Tanggal	Jumlah	Harga satuan	Total			Modal satuan	Total modal
				Dasar	Motif									
1	5/1/2020	Katun satin	langsat - obor	kuning	cecer hijau	313	15/01/2020	313	Rp135.000.00	Rp42.255.000.00	Rp75.000.00	Rp23.475.000.00	Rp18.780.000.00	ibu rina adaro
2	7/1/2020	Katun satin	kolak - kotak biasa	merah	kuning	2	8/1/2020	2	Rp125.000.00	Rp250.000.00	Rp75.000.00	Rp150.000.00	Rp100.000.00	roy taylor
3	7/1/2020	Katun satin	gigi haruan	hijau	coklat	1	9/1/2020	1	Rp125.000.00	Rp125.000.00	Rp70.000.00	Rp70.000.00	Rp55.000.00	mm dina
4	9/1/2020	Sutra katun	gegasas	cream	merah	1	10/1/2020	1	Rp200.000.00	Rp200.000.00	Rp107.000.00	Rp107.000.00	Rp93.000.00	mm dewi desa kapar
5	9/1/2020	Katun satin	Naga Balimbur	hitam	putih	4	13/01/2020	4	Rp130.000.00	Rp520.000.00	Rp75.000.00	Rp300.000.00	Rp220.000.00	joss taylor
6	13/1/2020	Katun sutra warna alam	bintang behambur	grey	putih	1	15/01/2020	1	Rp135.000.00	Rp135.000.00	Rp122.000.00	Rp122.000.00	Rp128.000.00	pak andi
7	13/1/2020	Katun sutra	Ombak Sinapur Karang - langsung	hijau	kuning	1	16/01/2020	1	Rp200.000.00	Rp200.000.00	Rp107.000.00	Rp107.000.00	Rp93.000.00	pak fuji
8	15/1/2020	Katun satin	Kangkung Kaombakan	hijau tua	biru - putih	1	16/01/2020	1	Rp130.000.00	Rp130.000.00	Rp75.000.00	Rp75.000.00	Rp55.000.00	mm rian
9	15/1/2020	Katun sutra	langsat - bulat-bulat	kuning	putih ceceer hijau	1	17/01/2020	1	Rp200.000.00	Rp200.000.00	Rp107.000.00	Rp107.000.00	Rp93.000.00	mm dody
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
554	16/07/2024	Katun polisima warna alam	langsat	hijau	kuning	1	12/7/2024	1	Rp300.000.00	Rp300.000.00	Rp85.000.00	Rp85.000.00	Rp215.000.00	bu aning

Sumber: Hasil Penelitian (2025)

Tampilan data tersebut masih memerlukan serangkaian pengolahan lebih lanjut berupa pembersihan dan pengubahan format agar siap untuk dianalisis pada langkah berikutnya. Pada gambar diatas menjelaskan isi data mentah penjualan kain sasirangan UMKM "Sasirangan Bunda Maburai" dengan isian kolom yakni tanggal produksi kain, jenis kain, motif, warna, jumlah, tanggal keluar kain, harga jual, modal produksi, keuntungan, dan keterangan pemesan atau pembeli kain sasirangan.

**Pembersihan data (data cleaning)**

Tahapan pembersihan data merupakan langkah pertama dalam proses KDD ntuk memastikan bahwa data yang digunakan valid dan layak dianalisis. Dalam penelitian ini, data yang digunakan berasal dari riwayat transaksi penjualan kain sasirangan UMKM "Sasirangan Bunda Maburai" dengan total 554 entri. Data mentah yang diperoleh terdiri atas kolom: tanggal produksi, jenis kain, motif, warna dasar, warna motif, jumlah produksi, harga jual, modal produksi, keuntungan, dan nama pemesan. Pembersihan data dilakukan secara manual menggunakan Microsoft Excel, dengan beberapa Langkah yakni: menghapus baris data yang kosong, mengidentifikasi dan menghapus data duplikat, memperbaiki kesalahan penulisan,

memastikan format data sesuai seperti tanggal dengan format dd-mm-yyyy, serta angka dalam format numerik tanpa karakter non-angka.

#### Data integration (integrasi data)

Setelah data dibersihkan, tahapan selanjutnya ialah proses menyatukan data atau pengintegrasian data dari berbagai sumber. Pada penelitian ini, data yang digunakan adalah data penjualan kain sasirangan dari berbagai jenis sumber, seperti laporan manual yang ditulisi tangan oleh pengelola UMKM, dan catatan digital seperti file excel yang disimpan dalam beberapa periode secara terpisah. Semua format data tersebut perlu digabungkan kedalam satu dataset disini peneliti menyatukan data dalam format excel agar mudah untuk dikelola.

#### Seleksi data (selection)

Setelah data disatukan dalam satu format excel yang sama, tahapan selanjutnya yaitu seleksi data, tahapan ini dilakukan untuk memilih atribut atau variabel yang relevan untuk penelitian yaitu mengelompokkan produk kain sasirangan berdasarkan pola penjualan untuk membantu manajemen persediaan dan menghindari penumpukan stok. Pada penelitian ini peneliti menentukan master data dengan empat variabel utama yang digunakan adalah jenis kain, motif, dan warna kain sasirangan. Alasan pemilihan keempat atribut ini adalah karena kombinasi jenis kain dan karakteristik visual (warna dan motif) diyakini sebagai faktor utama yang memengaruhi preferensi konsumen terhadap produk kain sasirangan.

Proses seleksi dilakukan dengan menghapus kolom yang tidak digunakan seperti "tanggal produksi", "modal", "keuntungan", dan "nama pemesan" karena tidak relevan dalam konteks analisis preferensi berbasis karakteristik produk. Berikut ini tampilan dataset hasil seleksi atribut yang digunakan.

Table 2. dataset setelah seleksi data

No	Jenis kain	Motif	Warna dasar	Warna motif
1	Katun satin	langsat - obor	kuning	cecer hijau
2	Katun satin	kotak-kotak biasa	merah	kuning
3	Katun satin	gigi haruan	hijau	coklat
4	Sutra katun	gegatas	cream	merah
5	Katun satin	Naga Balimbur	hitam	putih
6	Katun sutra warna alam	bintang behambur	grey	putih

Sumber: Hasil Penelitian (2025)

#### Transformasi data (data transformation)

Setelah data diseleksi, tahapan selanjutnya yaitu mentransformasikan data agar sesuai dengan algoritma data mining. Transformasi data melibatkan normalisasi atau standarisasi nilai pada atribut yang telah ditentukan sebelumnya sehingga setiap atribut memiliki bobot yang sama pada proses *clustering*. Pada atribut jenis kain, motif dan warna sasirangan, akan di normalisasikan menggunakan metode Min-Max Scaling agar mencegah atribut dengan nilai lebih besar mendominasi hasil *clustering*.

#### Proses penambangan data (data mining)

Pada tahapan ini merupakan proses penambangan data dengan menggunakan Bahasa pemrograman python, dimana proses algoritma *K-means clustering* di terapkan untuk mengelompokkan data penjualan kain sasirangan. Pada algoritma ini akan membagi data menjadi beberapa kelompok atau *cluster* berdasarkan atribut yang telah ditentukan sebelumnya yakni dengan enam kombinasi variabel seperti : jenis kain - motif, jenis kain - warna dasar, motif - warna dasar, motif - warna motif, warna dasar - warna motif, dan jenis kain - warna motif.

#### Validasi data atau evolusi data

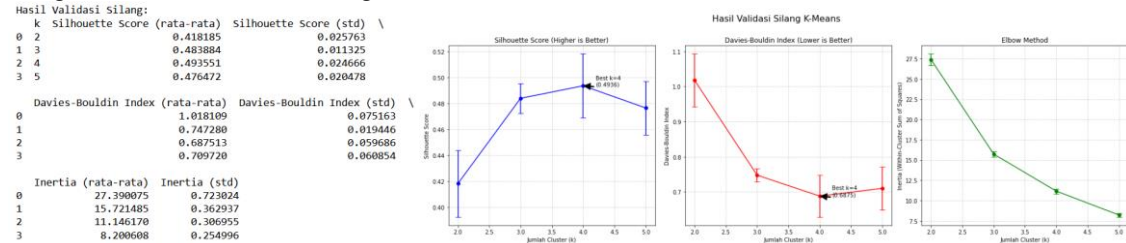
Pada tahapan evolusi atau validasi data merupakan proses untuk menguji efektivitas pemodelan machine learning dengan menggunakan data yang berbeda tapi tetap dengan dataset yang telah disediakan sebagai alat pengujinya. Dalam proses validasi data ada enam kombinasi variabel yang diuji secara terpisah untuk menganalisis kualitas *cluster* yang telah terbentuk. Kualitas *cluster* diukur menggunakan silhouette score yang bertujuan untuk mengevaluasi seberapa baik objek antar *cluster*. Nilai Silhouette score berkisar antara -1 hingga 1, di mana

nilai yang mendekati 1 menunjukkan bahwa objek ditempatkan dengan tepat pada *cluster* yang sesuai.

Adapun untuk hasil validasi dari enam kombinasi variabel sebagai berikut:

**Variabel jenis kain - motif**

Pengujian validitas data untuk pasangan variabel jenis kain dan motif dilaksanakan menggunakan teknik cross-validation, dengan hasil yang tersaji pada gambar 3 validasi data dengan *Cross-Validation* dan grafik Cross-Validation.

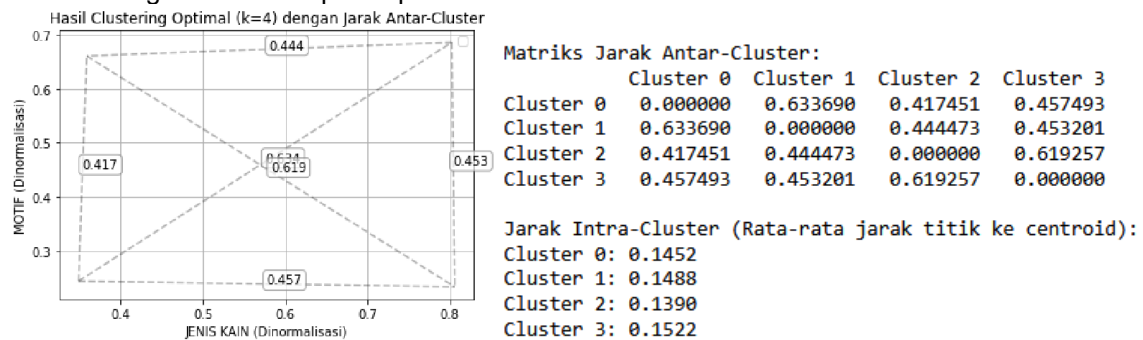


Sumber: Hasil Penelitian (2025)

Gambar 2. Validasi data dan grafik dengan Cross-Validation

Berdasarkan pengujian tersebut, terlihat bahwa model memiliki performa yang tetap dengan nilai silhouette score yang relatif serupa pada setiap lipatan data. Melalui tampilan grafik di atas, tergambar keselarasan kinerja model ketika diuji dengan berbagai variasi data dan nilai K yang berbeda-beda.

Penentuan jumlah kelompok yang paling tepat dilakukan dengan menerapkan metode Elbow sebagaimana ditampilkan pada Gambar

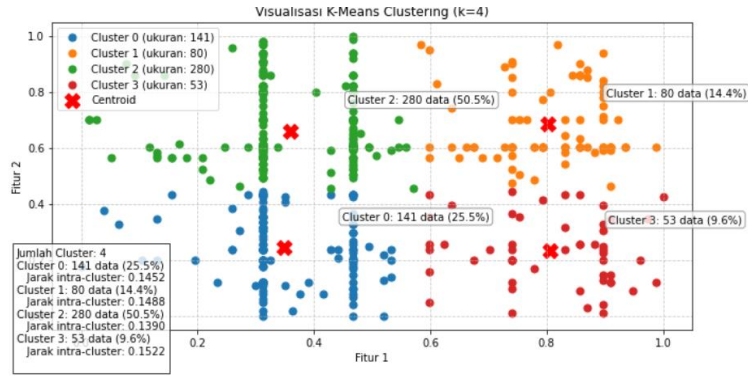


Sumber: Hasil Penelitian (2025)

Gambar 3. Hasil perhitungan jarak *cluster* untuk k optimal

Berdasarkan analisis grafik tersebut, terlihat adanya titik belok yang signifikan pada kurva, mengindikasikan bahwa pengelompokan menjadi 3 *cluster* merupakan pilihan paling optimal untuk kombinasi variabel yang sedang diuji. Hasil perhitungan tersebut menunjukkan adanya pemisahan yang baik antara kelompok-kelompok data, sehingga memperkuat kesahihan hasil pengelompokan yang telah dilakukan.

Berikut ini ditampilkan representasi visual hasil pengelompokan data menggunakan algoritma K-means untuk kombinasi jenis kain dan motif pada Gambar 5 grafik visualisasi k-means.



Sumber: Hasil Penelitian (2025)

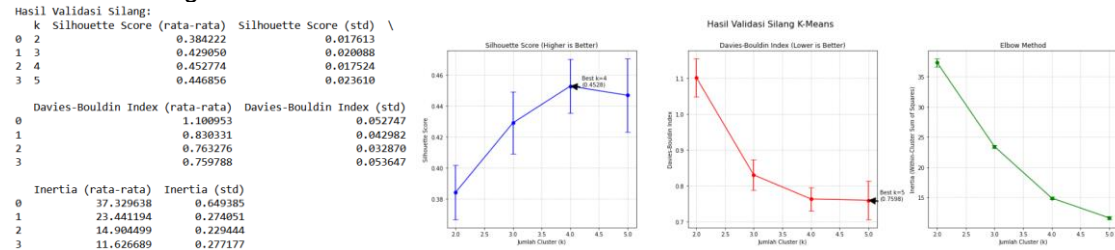
Gambar 4. Grafik visualisasi k-means

Visualisasi di atas memperlihatkan pembagian kelompok data yang terbentuk dengan jelas serta memiliki batas pemisah yang cukup tegas antara satu kelompok dengan kelompok lainnya.

Kombinasi variabel jenis kain – motif menunjukkan nilai silhouette score tertinggi pada nilai 0,65. Cluster 3 menunjukkan jumlah optimal cluser pada elbow method. Hal ini menunjukkan bahwa preferensi konsumen berdasarkan jenis kain dan motif dapat dikelompokkan menjadi 3 kategori utama. Visualisasi K-Means menunjukkan bahwa preferensi konsumen sangat dipengaruhi oleh jenis kain dan karakteristik pola motif. Validasi silang menunjukkan stabilitas model dan pemisahan antara kelompok yang jelas. Oleh karena itu, variasi ini sangat dianjurkan sebagai referensi utama untuk strategi segmentasi produk.

**Variabel warna dasar - warna motif**

Proses validasi data untuk kombinasi selanjutnya, yaitu warna dasar dan warna motif, dengan metode *Cross-Validation* ditampilkan pada Gambar 6 validasi data dengan *Cross-Validation* dan grafik *Cross-Validation*.

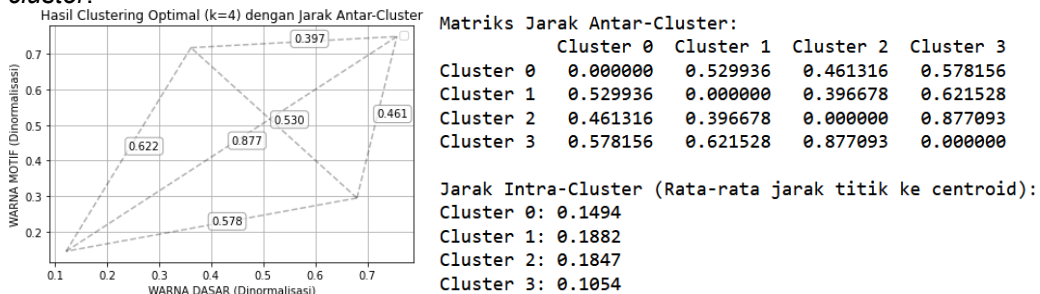


Sumber: Hasil Penelitian (2025)

Gambar 5. Validasi data dengan *Cross-Validation* dan grafik *Cross-Validation*.

Hasil validasi ini memberikan gambaran tentang kinerja model dalam mengelompokkan data berdasarkan kedua atribut warna tersebut. Dan grafik tersebut menunjukkan tingkat keandalan model yang konsisten pada berbagai pengujian dengan nilai K yang berbeda.

Analisis penentuan jumlah cluster optimal untuk kombinasi warna dasar dan warna motif menggunakan metode Elbow ditunjukkan pada Gambar 7 hasil perhitungan jarak antar cluster.

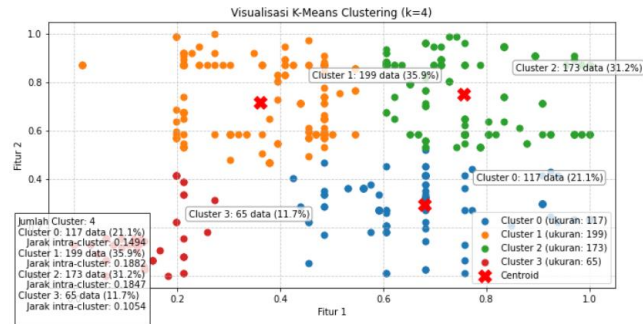


Sumber: Hasil Penelitian (2025)

Gambar 6. Hasil perhitungan jarak antar cluster.

Hasil analisis menunjukkan bahwa 3 *cluster* merupakan jumlah optimal untuk pengelompokan data dengan kombinasi kedua atribut warna ini. Perhitungan ini mengkonfirmasi pemisahan yang optimal antara ketiga *cluster*, menandakan bahwa kombinasi atribut warna memiliki pengaruh signifikan terhadap preferensi konsumen kain sasirangan.

Berikut ini ditampilkan representasi visual hasil pengelompokan data menggunakan algoritma K-means untuk kombinasi warna dasar dan warna motif pada Gambar 8 grafik visualisasi k-means.



Sumber: Hasil Penelitian (2025)

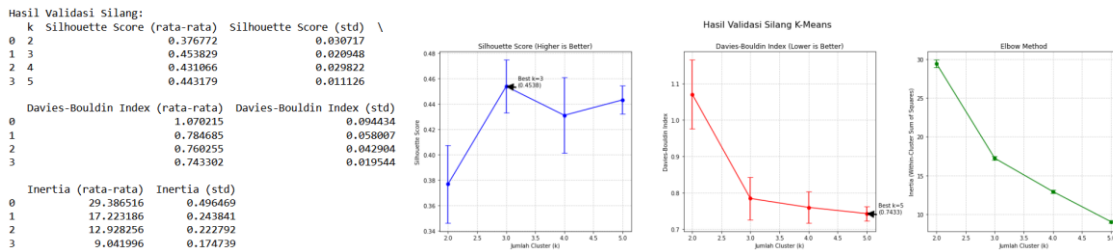
Gambar 7. Grafik visualisasi k-means.

Dari visualisasi tersebut, terlihat pembagian data menjadi tiga kelompok dengan pemisahan yang sangat baik dan tegas.

Kombinasi variabel warna dasar – warna motif menunjukkan nilai silhouette score tertinggi pada nilai 0,63 dengan 3 jumlah *cluster* optimal sebagai pemisah antar *cluster* yang tajam, hal ini menandakan jika warna menjadi indikator penting pada preferensi konsumen.

**Variabel jenis kain - warna dasar**

Pada bagian berikutnya, disajikan hasil validasi dan Visualisasi data menggunakan *Cross-Validation* untuk pasangan variabel jenis kain dan warna dasar seperti terlihat pada Gambar 9.

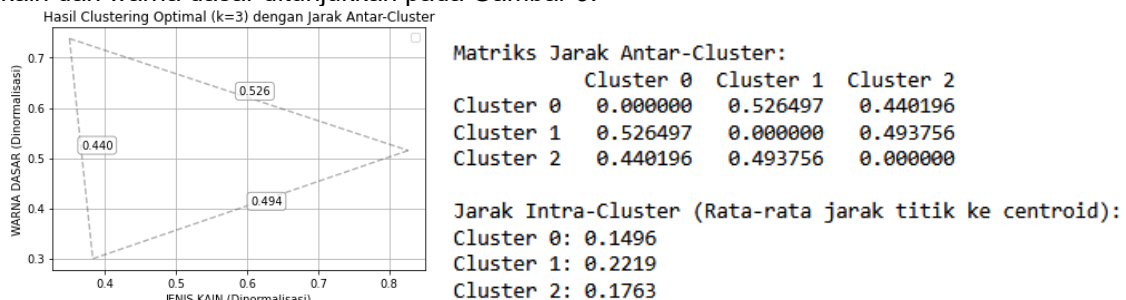


Sumber: Hasil Penelitian (2025)

Gambar 8. Validasi data dengan *Cross-Validation* dan grafik *Cross-Validation*.

Validasi ini memberikan gambaran tentang kestabilan model dalam mengelompokkan data berdasarkan jenis kain dan warna dasar. Dan Grafik tersebut memperlihatkan konsistensi kinerja model pada berbagai percobaan dengan nilai K yang bervariasi.

Analisis penentuan jumlah *cluster* optimal dengan metode Elbow untuk kombinasi jenis kain dan warna dasar ditunjukkan pada Gambar 9.

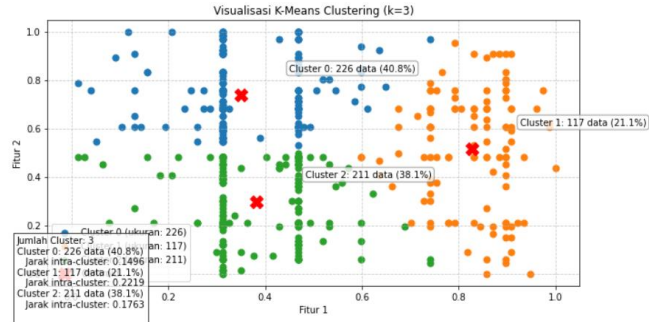


Sumber: Hasil Penelitian (2025)

Gambar 9. Hasil perhitungan jarak antar *cluster*.

Hasil analisis menunjukkan bahwa 3 *cluster* merupakan jumlah optimal untuk pengelompokan data dengan kombinasi variabel ini. Dan hasil Perhitungan ini menunjukkan bahwa terdapat pemisahan yang cukup antara *cluster-cluster* yang terbentuk.

Berikut ini ditampilkan representasi visual hasil pengelompokan data menggunakan algoritma K-means untuk kombinasi variabel jenis kain dan warna dasar pada Gambar 10 grafik visualisasi k-means.



Sumber: Hasil Penelitian (2025)

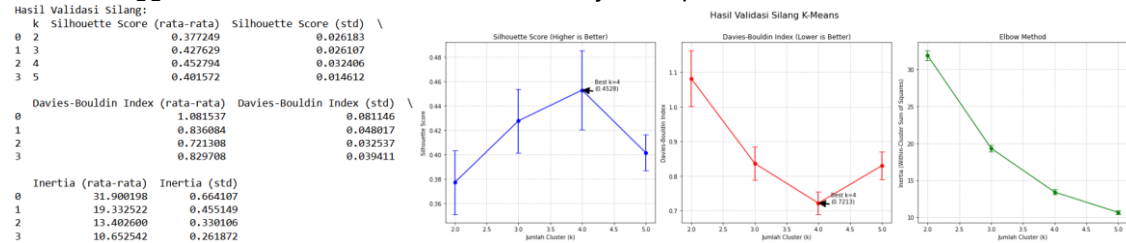
Gambar 10. Grafik visualisasi k-means.

Visualisasi tersebut menggambarkan distribusi data yang cukup terpisah meskipun tidak sebaik kombinasi variabel sebelumnya.

Kombinasi variabel jenis kain – warna dasar menunjukkan nilai silhouette score tertinggi pada nilai 0,58 dengan 3 jumlah *cluster* optimal. Hasil pengelompokan menunjukkan bahwa atribut jenis kain tetap variabel penting, tetapi dikombinasikan dengan warna dasar. Tiga *cluster* secara konsisten dibentuk oleh visualisasi yang menunjukkan distribusi data yang terlalu terbuka antar *cluster*.

**Variabel motif - warna dasar**

Proses validasi dan Representasi grafis untuk kombinasi variabel motif dan warna dasar menggunakan teknik *Cross-Validation* ditunjukkan pada Gambar 12.

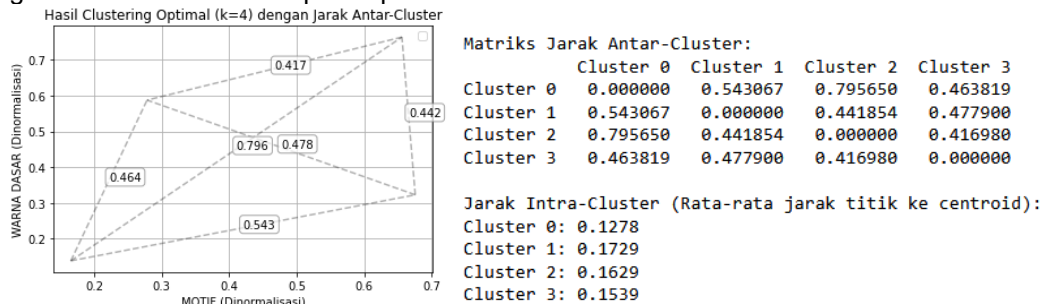


Sumber: Hasil Penelitian (2025)

Gambar 11. Validasi dan Visualisasi data menggunakan Cross-Validation.

Hasil validasi ini memberikan informasi tentang kemampuan model dalam mengelompokkan data berdasarkan motif dan warna dasar. Dan Grafik tersebut menunjukkan tingkat keandalan model pada berbagai pengujian dengan nilai K yang berbeda.

Analisis penentuan jumlah *cluster* optimal untuk kombinasi motif dan warna dasar dengan metode Elbow ditampilkan pada Gambar

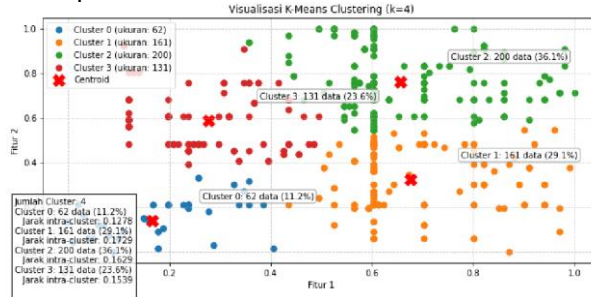


Sumber: Hasil Penelitian (2025)

Gambar 12. Hasil perhitungan jarak antar *cluster*.

Berdasarkan hasil analisis, dapat ditentukan bahwa 4 *cluster* merupakan jumlah optimal untuk kombinasi variabel ini. Perhitungan ini menunjukkan tingkat pemisahan yang cukup baik antara keempat *cluster* yang terbentuk.

Visualisasi hasil *clustering* dengan algoritma K-means untuk kombinasi variabel motif dan warna dasar ditunjukkan pada Gambar 14.



Sumber: Hasil Penelitian (2025)

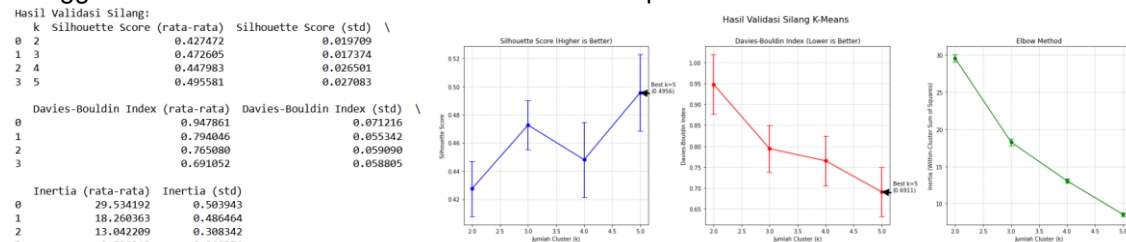
Gambar 13. Grafik visualisasi k-means.

Dari visualisasi tersebut, terlihat pembagian data menjadi empat kelompok yang cukup terpisah dengan jelas.

Kombinasi variabel motif – warna dasar menunjukkan nilai silhouette score tertinggi pada nilai 0,56 dengan 4 jumlah *cluster* optimal. Pada kombinasi variabel ini menunjukkan bahwa data dapat dikelompokkan dengan baik, berdasarkan visualisasi k-means motif kain sangat berpengaruh terhadap preferensi konsumen Ketika variabel digabungkan dengan warna dasar kain sasirangan.

**Variabel jenis kain - warna motif**

Proses validasi dan grafik data untuk kombinasi variabel jenis kain dan warna motif menggunakan metode *Cross-Validation* diilustrasikan pada Gambar 15.

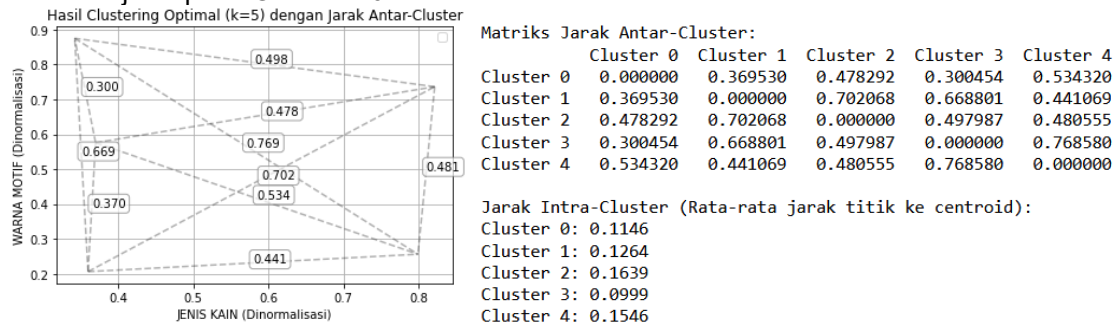


Sumber: Hasil Penelitian (2025)

Gambar 14. Validasi dan visualisasi data menggunakan *k-means clustering*.

Hasil validasi ini menggambarkan tingkat keandalan model dalam mengelompokkan data berdasarkan jenis kain dan warna motif. Dan Grafik ini memberikan gambaran tentang konsistensi kinerja model dengan beragam nilai K yang diuji.

Analisis untuk menentukan jumlah *cluster* optimal pada kombinasi jenis kain dan warna motif disajikan pada Gambar 16.

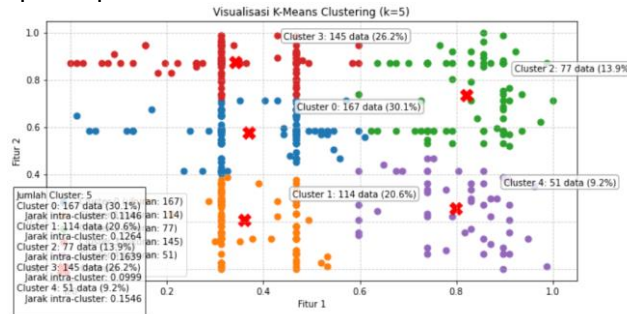


Sumber: Hasil Penelitian (2025)

Gambar 15. Hasil perhitungan jarak antar *cluster*

Berdasarkan analisis tersebut, dapat ditentukan bahwa 3 cluster juga menjadi jumlah optimal untuk kombinasi variabel ini. Perhitungan ini memperlihatkan bahwa pemisahan antar cluster tidak setajam pada kombinasi variabel sebelumnya.

Visualisasi hasil pengelompokan dengan algoritma K-means untuk kombinasi jenis kain dan warna motif ditampilkan pada Gambar 17.



Sumber: Hasil Penelitian (2025)

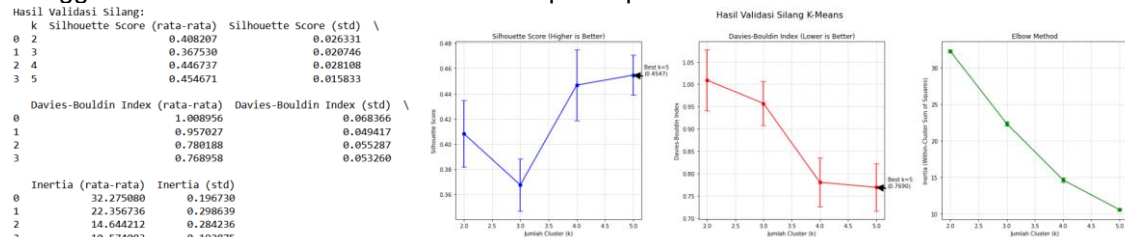
Gambar 16. Grafik visualisasi k-means.

Dari visualisasi tersebut, terlihat pembagian kelompok yang kurang tegas dibandingkan dengan kombinasi variabel lainnya.

Kombinasi variabel jenis kain – warna motif menunjukkan nilai silhouette score tertinggi pada nilai 0,54 dengan 3 jumlah cluster optimal. Hal ini menunjukkan performa rendah dengan K-fold Cross validation k=5.

**Variabel motif - warna motif**

Proses validasi dan visualisasi data untuk kombinasi variabel motif dan warna motif menggunakan metode Cross-Validation ditampilkan pada Gambar 18.

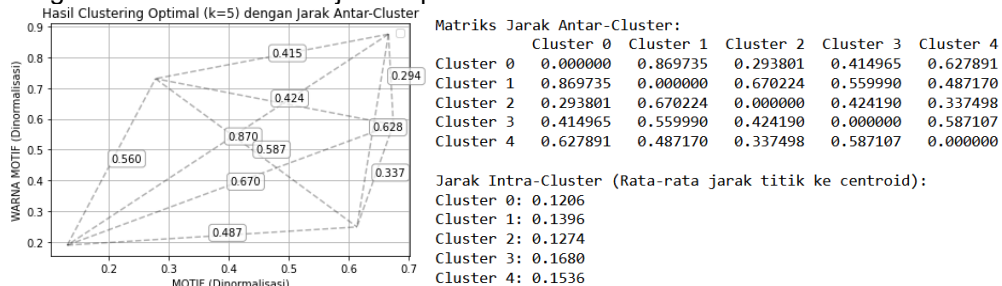


Sumber: Hasil Penelitian (2025)

Gambar 17. Grafik visualisasi k-means.

Hasil validasi ini menggambarkan stabilitas model dalam mengelompokkan data berdasarkan motif dan warna motif kain sasirangan. dan Grafik ini memperlihatkan konsistensi kinerja model pada berbagai pengujian dengan beragam nilai K.

Analisis untuk menentukan jumlah cluster optimal pada kombinasi motif dan warna motif dengan metode Elbow ditunjukkan pada Gambar 19.

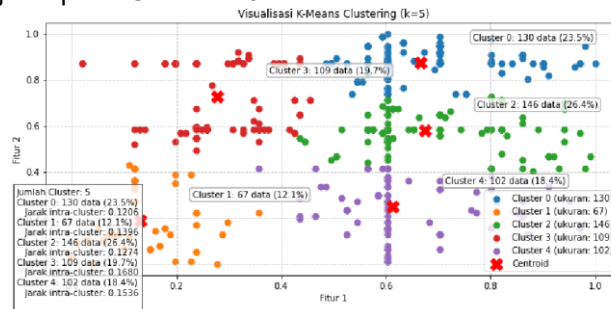


Sumber: Hasil Penelitian (2025)

Gambar 18. Grafik visualisasi k-means.

Berdasarkan analisis tersebut, diperoleh bahwa 5 cluster merupakan jumlah optimal untuk kombinasi variabel ini. Dan hasil validasi ini memberikan gambaran tentang kinerja model dalam mengelompokkan data berdasarkan kedua atribut warna tersebut.

Representasi grafis dari hasil *Cross-Validation* untuk kombinasi variabel warna dasar dan warna motif disajikan pada Gambar 20.



Sumber: Hasil Penelitian (2025)

Gambar 19. Grafik visualisasi k-means.

Dari visualisasi tersebut, terlihat pembagian data menjadi lima kelompok dengan tingkat pemisahan yang cukup baik.

Kombinasi variabel motif – warna motif menunjukkan nilai silhouette score tertinggi pada nilai 0,52 dengan 5 jumlah *cluster* optimal dengan nilai silhouette score yang cukup seimbang hal ini menunjukkan warna motif menjadi variabel penting saat dikombinasikan dengan motif untuk segmentasi produk.

Kebaruan penelitian ini terletak pada pendekatan kombinasi variabel yang disesuaikan dengan karakteristik kain Sasirangan. Pendekatan ini berbeda dari penelitian sebelumnya yang umumnya hanya menggunakan satu atau dua atribut. Pendekatan multidimensional ini memberikan pemahaman lebih dalam terhadap pola konsumsi produk budaya lokal.

#### 4. Kesimpulan

Berdasarkan hasil penelitian ini, dapat disimpulkan bahwa penerapan algoritma K-means clustering pada data penjualan kain Sasirangan UMKM "Sasirangan Bunda Maburai" efektif dalam mengidentifikasi pola preferensi konsumen. Dari enam kombinasi variabel yang diuji, kombinasi jenis kain – motif menunjukkan performa clustering terbaik dengan silhouette score sebesar 0,65 dan jumlah optimal sebanyak 3 cluster. Kombinasi ini diikuti oleh warna dasar – warna motif dengan nilai silhouette score 0,63, yang juga menunjukkan pembagian cluster yang baik.

Penelitian ini memberikan kontribusi nyata dalam pemanfaatan teknik data mining berbasis algoritma unsupervised learning untuk sektor UMKM berbasis warisan budaya lokal, yang sebelumnya belum banyak dibahas dalam literatur. Pendekatan yang digunakan terbukti mampu menghasilkan segmentasi produk secara objektif berdasarkan data penjualan historis, sehingga dapat digunakan untuk Menyusun strategi produksi yang lebih terarah, mengelola stok dengan lebih efisien, dan mengurangi resiko penumpukan produk di butik. Implikasi praktis dari hasil ini adalah bahwa UMKM dapat mengembangkan kebijakan manajemen persediaan yang berbasis data konsumen, bukan hanya intuisi atau pengalaman subjektif. Adapun harapan untuk peneliti selanjutnya disarankan menambahkan variabel waktu penjualan secara musiman dan demografi pembeli untuk menganalisis segmentasi lebih menyeluruh.

#### Ucapan Terima Kasih

Peneliti mengucapkan banyak terimakasih kepada pihak UMKM "Sasirangan Bunda Maburai" yang telah memberikan data transaksi penjualan kain sasirangan. Peneliti juga mengucapkan terimakasih kepada dosen pembimbing yang telah membimbing peneliti hingga dapat menyelesaikan penelitian ini. Tak lupa pula peneliti mengucapkan terimakasih kepada kedua orang tua serta teman-teman peneliti yang telah memberikan kasih sayang dan dukungan kepada peneliti dari awal penelitian hingga saat ini.

#### Referensi

- [1] Misbahwati, *Sasirangan Bunda Maburai*. Tanjung, Tabalong, 2018.
- [2] M. Effendi, S. Purnamasari, Kurniaty, P. Rozak, and A. Wahab, "Pengembangan Ekonomi Kreatif pada Pengrajin Kain Sasirangan dalam Perspektif Ekonomi Syariah (Studi Kasus Toko Luthfiah Sasirangan)," *Formosa J. Sci. Technol.*, vol. 1, no. 6, pp.

- 647–662, 2022, doi: 10.55927/fjst.v1i6.1367.
- [3] L. Widyawati and V. Lusiana, "Penerapan K-Means Clustering Untuk Mengelompokkan Data Transaksi Penjualan ( Studi Kasus pada Wijaya Hijab )".
- [4] A. Nugraha, O. Nurdiawan, and G. Dwilestari, "Penerapan Data Mining Metode K-Means Clustering Untuk Analisa Penjualan Pada Toko Yana Sport," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 6, no. 2, pp. 849–855, 2022, doi: 10.36040/jati.v6i2.5755.
- [5] T. Asy Aria, M. Julkarnain, and F. Hamdani, "KLIK: Kajian Ilmiah Informatika dan Komputer Penerapan Algoritma K-Means Clustering Untuk Data Obat," *Media Online*, vol. 4, no. 1, pp. 649–657, 2023, doi: 10.30865/klik.v4i1.1117.
- [6] I. Yati Beti and H. Juliansa, "KLIK: Kajian Ilmiah Informatika dan Komputer Penerapan Normalisasi Data Metode Decimal Scaling Dan Metode K-Means Dalam Mengelompokkan Kasus Demam Berdarah," *Media Online*, vol. 4, no. 6, pp. 2928–2936, 2024, doi: 10.30865/klik.v4i6.1925.
- [7] Amna, W. S, and I. G. I. Sudipa, *Data Mining*, vol. 2, no. January 2013. 2023. [Online]. Available: [https://www.cambridge.org/core/product/identifier/CBO9781139058452A007/type/book\\_part](https://www.cambridge.org/core/product/identifier/CBO9781139058452A007/type/book_part)
- [8] Trivusi, "K-Means Clustering: Pengertian, Cara Kerja, Kelebihan, dan Kekurangannya," *trivusi.com*. [Online]. Available: <https://www.trivusi.web.id/2022/06/algoritma-kmeans-clustering.html>
- [9] N. Nur Afidah, "Penerapan Metode Clustering dengan Algoritma K-means untuk Pengelompokkan Data Migrasi Penduduk Tiap Kecamatan di Kabupaten Rembang," *Prism. Pros. Semin. Nas. Mat.*, vol. 6, pp. 729–738, 2023, [Online]. Available: <https://journal.unnes.ac.id/sju/index.php/prisma/>
- [10] S. Putri R, "Kenalan dengan Metode Clustering," Unit Pengelola Jurnal Ilmiah. [Online]. Available: <https://uptjurnal.umsu.ac.id/kenalan-dengan-metode-clustering/>
- [11] Y. Khoeri, R. Kurniawan, and Y. Arie Wijaya, "Penerapan Teknik Smote Dan Cross Validation Pada Decision Tree Untuk Klasifikasi Tingkat Kemacetan Lalu Lintas," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 8, no. 2, pp. 2006–2012, 2024, doi: 10.36040/jati.v8i2.8879.
- [12] W. W. Kristianto and C. Rudianto, "Penerapan Data Mining Pada Penjualan Produk Menggunakan Metode K-Means Clustering (Studi Kasus Toko Sepatu Kakikaki)," *J. Pendidik. Teknol. Inf.*, no. 5, pp. 90–98, 2020.
- [13] Siradjuddin, *Algoritma Pemrograman dengan Menggunakan Python*, no. September. 2022.
- [14] P. Rahayu *et al.*, *Buku Ajar Data Mining*, vol. 1, no. January 2024. 2018.
- [15] F. Alghifari and D. Juardi, "Penerapan Data Mining Pada Penjualan Makanan Dan Minuman Menggunakan Metode Algoritma Naïve Bayes," *J. Ilm. Inform.*, vol. 9, no. 02, pp. 75–81, 2021, doi: 10.33884/jif.v9i02.3755.