

Klasterisasi Karakteristik Produk Handphone Menggunakan Algoritma K-Means Untuk Menentukan Segmentasi Pasar

Muda Wisnu Aji, Tia Tanjung

Program studi Informatika, Universitas Bandar Lampung

Program studi Informatika, Universitas Bandar Lampung

Corresponding author: muda.20421049@student.ubl.ac.id

Abstract:

The current digital era has made mobile phones an inseparable device from everyday life. Technological developments have resulted in various mobile phone products that continue to circulate rapidly. However, there is no system that is able to effectively group these mobile phone products based on their characteristics. Therefore, this study aims to group mobile phone products using the clustering method and the K-Means algorithm. The system developed will group mobile phone products into several clusters based on the characteristics of the product. The results of this clustering can then be used to determine market segmentation by conducting further analysis. In addition, this study also evaluates the clustering results using two metrics, namely Sum of Squared Errors (SSE) and Silhouette Score. The evaluation aims to assess the quality of the grouping produced by the system. Based on the results of the study, the system successfully grouped mobile phone products into clusters with low SSE values and high Silhouette values. This shows that the grouping carried out has good quality, with data in one cluster more compact and close to the centroid, and the distance between clusters is quite clear. These results can be used as a reference in determining the market segmentation strategy for mobile phone products in the future.

Keywords: K-means; Clustering; Market Segmentation; Mobile Phone

Abstrak:

Era digital saat ini telah menjadikan handphone sebagai perangkat yang tidak bisa dipisahkan dari kehidupan sehari-hari. Perkembangan teknologi menghasilkan berbagai produk handphone yang terus beredar dengan cepat. Namun, belum ada sistem yang mampu secara efektif mengelompokkan produk-produk handphone tersebut berdasarkan karakteristiknya. Oleh karena itu, penelitian ini bertujuan untuk mengelompokkan produk handphone dengan menggunakan metode clustering dan algoritma K-Means. Sistem yang dikembangkan akan mengelompokkan produk handphone ke dalam beberapa cluster berdasarkan karakteristik produk tersebut. Hasil clustering ini kemudian dapat digunakan untuk menentukan segmentasi pasar dengan melakukan analisis lebih lanjut. Selain itu, penelitian ini juga mengevaluasi hasil clustering menggunakan dua metrik, yaitu Sum of Squared Errors (SSE) dan Silhouette Score. Evaluasi tersebut bertujuan untuk menilai kualitas pengelompokan yang dihasilkan oleh sistem. Berdasarkan hasil penelitian, sistem berhasil mengelompokkan produk handphone ke dalam cluster dengan nilai SSE yang rendah dan nilai Silhouette yang tinggi. Hal ini menunjukkan bahwa pengelompokan yang dilakukan memiliki kualitas yang baik, dengan data dalam satu cluster lebih kompak dan dekat dengan centroid, serta jarak antar cluster yang cukup jelas. Hasil ini dapat menjadi acuan dalam menentukan strategi segmentasi pasar produk handphone di masa mendatang.

Kata Kunci: K-means; Clustering; Segmentasi Pasar; Handphone.

Pendahuluan

Di era digital saat ini, handphone telah menjadi perangkat yang tidak terpisahkan dari kehidupan sehari-hari masyarakat. Handphone bukan lagi hanya alat komunikasi, tetapi juga untuk berbagai kebutuhan seperti akses internet, hiburan, dan beragam aplikasi berbasis visual dan audio [1]. Penggunaan handphone telah meluas ke berbagai kalangan, mulai dari anak-anak hingga lansia, karena kemudahan penggunaan, harga yang beragam, serta fitur-fitur canggih yang terus diperbarui setiap tahunnya. Pesatnya perkembangan teknologi menciptakan berbagai macam produk handphone tetapi hal ini tidak diimbangi dengan adanya sistem pengelolaan data yang baik, terutama dalam hal pengelompokan karakteristik produk yang dapat digunakan untuk segmentasi pasar.

Segmentasi pasar adalah salah satu strategi penting dalam pemasaran, di mana produk atau layanan dikelompokkan berdasarkan kebutuhan konsumen yang beragam [2]. Dengan permasalahan tersebut, peneliti memiliki ide atau sebuah gagasan untuk membuat sebuah sistem website pengolahan data untuk mengelompokkan produk handphone berdasarkan karakteristiknya. Pengelompokan produk handphone berdasarkan karakteristik produk, seperti harga, spesifikasi, dan fitur, dapat memberikan informasi penting yang dapat digunakan dalam menentukan segmentasi pasar yang lebih efektif. Selain itu, sistem website pengolahan data ini dibangun menggunakan bahasa pemrograman php.

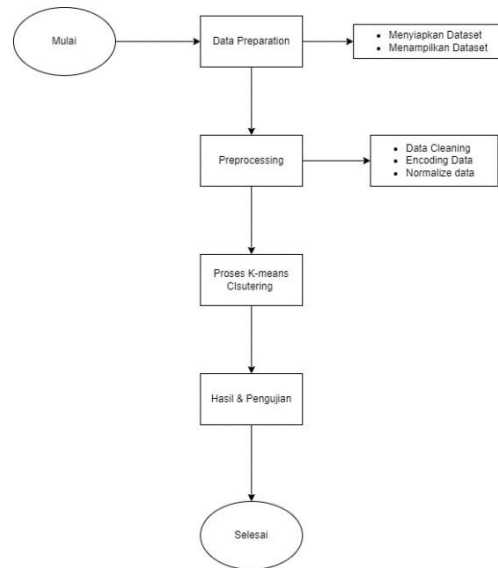
Penelitian ini akan menggunakan metode k-means clustering, salah satu teknik data mining atau machine learning yang populer dalam pengelompokan data. Clustering merupakan teknik yang membagi data menjadi beberapa kelompok berdasarkan kesamaan antar objek dalam satu kelompok, sehingga setiap kelompok memiliki kemiripan internal yang tinggi, sementara antar kelompok memiliki perbedaan yang signifikan [3]. Kelebihan dari metode k-means adalah kemudahannya dalam implementasi, efisiensi waktu pemrosesan, dan fleksibilitas dalam mengadaptasi data yang kompleks [4].

Berdasarkan studi literatur, penelitian sebelumnya telah berhasil menerapkan metode clustering dalam berbagai konteks, seperti pada segmentasi pasar penerbangan [5]. Perbandingan metode clustering untuk segmentasi pasar otomotif [6], serta analisis karakteristik pelanggan klub olahraga [7]. Hasil-hasil penelitian tersebut menunjukkan bahwa teknik clustering efektif dalam mengidentifikasi pola-pola preferensi konsumen dan segmentasi pasar yang berbeda-beda.

Penelitian ini berkontribusi pada klasterisasi produk handphone, khususnya dalam pengelompokan berdasarkan karakteristik produk untuk menentukan segmentasi pasar. Klasterisasi adalah proses mengelompokkan data ke dalam beberapa kelompok atau klaster dengan tujuan memaksimalkan kesamaan dalam klaster dan meminimalkan kesamaan antar [8]. Penelitian ini menggunakan metode k-means untuk mengelompokkan produk handphone berdasarkan atribut khas, dengan tujuan melakukan segmentasi pasar. Dataset sekunder dari Kaggle digunakan untuk mengklasifikasikan produk, sehingga hasilnya diharapkan dapat memberikan gambaran segmentasi produk dan membantu konsumen dalam memilih ponsel sesuai kebutuhan.

Metodologi

Alur penelitian ini melalui beberapa tahap seperti data preparation, preprocessing, proses k-means clustering dan visualisasi, hasil dan pengujian. Tahapan tersebut dijabarkan pada flowchart berikut:



Gambar 1 Flowchart Penelitian

A. Data

a) Data Preparation

1) Menyiapkan Dataset

Dataset adalah kumpulan data terstruktur, seperti tabel atau file, yang mencakup informasi dari berbagai sumber. Beragam jenis data, termasuk numerik, teks, dan gambar, memungkinkan penggunaannya di berbagai industri dan dapat dikategorikan ke dalam beberapa tipe [9]. Dataset yang digunakan dalam penelitian ini adalah dataset sekunder dari kaggle.com, yang berisi berbagai atribut produk handphone seperti name, price, rating, spec_score, sim, processor, ram (inbuilt), battery (in mAh), display size, display pixels, front_camera, rear_camera. Dataset ini kemudian akan diproses melalui tahap preprocessing menggunakan aplikasi rapidminer sebelum digunakan dalam pembuatan website dengan metode k-means clustering.

2) Menampilkan Dataset

Tabel 1 Sampel Dataset Mobile Phone

name	price	rating	specs_score	sim	processor	ram (inbuilt)
Apple iPhone 11	6999000	43	73	Dual Sim	A13 Bionic	64?GB inbuilt
Asus ROG Phone 6 Pro 5G	18999000	47	88	Dual Sim	Snapdragon 8+ Gen1	512?GB inbuilt
Google Pixel 5A	3450000	41	78	Single Sim	Snapdragon 765G	128?GB inbuilt
Honor X6	1999000	45	71	Dual Sim	Helio G25	64?GB inbuilt
Infinix Note 12 (G96)	2769000	46	79	Dual Sim	Helio G96	128?GB inbuilt

Tabel 2 Sampel Dataset Mobile Phone (lanjutan)

ram	battery (in mAh)	display size	display pixels	front_camera	rear_camera
4?GB RAM	3110?mAh	6.1 inches	828?x?1792?px Display with Large Notch	12?MP Front Camera	12?MP + 12?MP Dual Rear
18?GB RAM	6000?mAh	6.78 inches	1080?x?2448?px	12?MP Front Camera	50?MP + 13?MP + 5?MP Triple Rear
6?GB RAM	4620?mAh	6.34 inches	1080?x?2400?px Display with Punch Hole	8?MP Front Camera	16?MP + 12.2?MP Dual Rear
4?GB RAM	5000?mAh	6.5 inches	720?x?1600?px Display with Water Drop Notch	5?MP Front Camera	50?MP + 2?MP + 2?MP Triple Rear
8?GB RAM	5000?mAh	6.7 inches	1080?x?2400?px Display with Punch Hole	16?MP Front Camera	50?MP + 2?MP Triple Rear

Dataset yang digunakan dalam penelitian ini adalah sampel dataset dengan format "csv", berisi jenis-jenis handphone dan karakteristiknya. Dataset ini memiliki 1009 baris dan 13 atribut, yaitu name, price, rating, specs_score, sim, processor, ram (inbuilt), ram, battery (in mAh), display size, display pixels, front_camera, dan rear_camera. Dataset tersebut diambil dari kaggle.com dengan nama "mobile phone dataset" dan dapat diakses melalui halaman nomor 3 di bagian fitur dataset. Kaggle adalah platform online untuk data science yang memungkinkan kolaborasi, berbagi dataset, dan kompetisi dalam tantangan ilmu data. Didirikan pada 2010 dan diakuisisi Google pada 2017, platform ini menyediakan alat dan sumber daya untuk mendukung pengembangan ilmu data [10].

b) Preprocessing

Preprocessing data merupakan proses penting dalam analisis data mining yang bertujuan untuk membersihkan [11]. Preprocessing dilakukan menggunakan aplikasi rapidminer, Berikut tahapan preprocessing pada penelitian ini:

1) Data Cleaning

Dalam proses data cleaning, peneliti menghapus noise berupa tanda "?" pada atribut seperti ram, battery, display pixels, front_camera, dan rear_camera karena tidak relevan untuk clustering. Penghapusan ini memastikan dataset lebih bersih sehingga proses clustering berjalan optimal. Proses pembersihan data dilakukan menggunakan rapidminer.

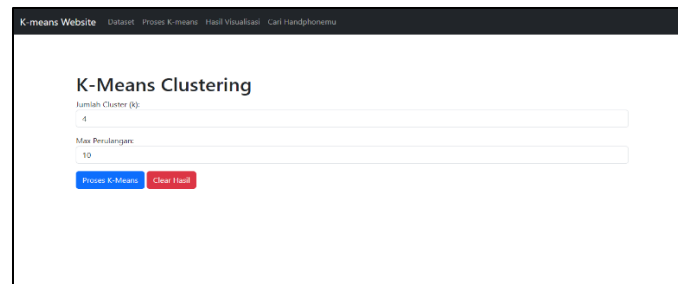
2) Encoding Data

Peneliti mengubah atribut kategorial menjadi numerik untuk mempermudah klasterisasi dengan algoritma k-means, yang membutuhkan data numerik agar dapat menghitung jarak dan mengelompokkan data secara efektif. Peneliti memilih beberapa atribut yaitu, battery (in mAh), display pixels, display size, front_camera, rear_camera, processor, ram, ram (inbuilt), dan sim.

3) Normalize Data

Peneliti menggunakan metode range transformation di rapidminer untuk menormalize data, mengubah skala atribut agar berada dalam rentang 0 hingga 1. Normalize data ini penting untuk mencegah perbedaan skala atribut mempengaruhi hasil clustering.

c) Proses K-means Clustering



Gambar 2 Proses Clustering

K-means adalah algoritma pembelajaran tanpa pengawasan yang digunakan untuk mengelompokkan data ke dalam kluster yang berbeda [12]. Clustering atau klasterisasi adalah metode pengelompokan data. Clustering merupakan proses partisi satu set objek data ke dalam himpunan bagian yang disebut dengan cluster [13]. Proses clustering menggunakan sistem website pengolahan data yang dibuat peneliti, pada gambar tersebut Peneliti menetapkan jumlah cluster (k) sebanyak 4 dan iterasi 10. Algoritma K-means kemudian menginisialisasi centroid awal, mengelompokkan data berdasarkan jarak terdekat dengan centroid menggunakan euclidean distance. Centroid diperbarui dengan menghitung rata-rata posisi data di setiap cluster hingga stabil atau iterasi maksimum tercapai.

Hasil dan Pembahasan

A. Preprocessing

a) Data Cleaning

Tabel 3 Sampel Dataset Sebelum Proses Data Cleaning

ram (inbuilt)	ram	battery (in mAh)
64?GB inbuilt	4?GB RAM	3110?mAh
128?GB inbuilt	4?GB RAM	3110?mAh
64?GB inbuilt	4?GB RAM	3190?mAh
64?GB inbuilt	4?GB RAM	3500?mAh

Tabel 4 Sampel Dataset Sesudah Proses Data Cleaning

ram (inbuilt)	ram	battery (in mAh)
64GB inbuilt	4GB RAM	3110mAh
128GB inbuilt	4GB RAM	3110mAh
64GB inbuilt	4GB RAM	3190mAh
64GB inbuilt	4GB RAM	3500mAh

b) Encoding Data

Tabel 5 Sampel Dataset Hasil Proses Encoding Data

processor	ram (inbuilt)	ram	battery (in mAh)	display size	display pixels	front_camera	rear_camera
0	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0
0	0	0	1	1	1	0	1
0	0	0	2	2	2	0	1
1	0	0	3	0	3	0	0

c) Normalize Data

Tabel 6 Sampel Dataset Hasil Proses Normalize Data

rating	specs_score	sim	processor	ram (inbuilt)
0.5833333333333334	0.7362637362637363	0.0	0.0	0.0
0.5	0.7582417582417582	0.0	0.0	0.07142857142857142
0.9166666666666666	0.7802197802197802	0.0	0.0	0.0
0.5833333333333334	0.7802197802197802	0.0	0.0	0.0

B. Hasil Clustering

K-means Website Dataset Proses K-means Hasil Visualisasi Cari Handphonemu

Hasil Clustering:

Cluster	Device	Processor	RAM	Display Size	Front
Cluster 1	- Apple iPhone 11 - Apple iPhone 11 (128GB) - Apple iPhone 11 Pro - Apple iPhone 11 Pro Max - Apple iPhone 12 - Apple iPhone 12 (128GB) - Apple iPhone 12 (256GB) - Apple iPhone 12 Mini - Apple iPhone 12 Mini (128GB) - Apple iPhone 12 Pro (512GB) - Apple iPhone 13 - Apple iPhone 13 (256GB) - Apple iPhone 13 (512GB) - Apple iPhone 13 Mini - Apple iPhone 13 Mini (256GB) - Apple iPhone 13 Pro	- A13 Bionic - A13 Bionic - A13 Bionic - A13 Bionic - Bionic A14 - Bionic A14 - Bionic A14 - Bionic A14 - Bionic A14 - Bionic A14 - Bionic A15 - Bionic A15 - Bionic A15 - Bionic A15 - Bionic A15 - Bionic A15	4GB RAM 4GB RAM 4GB RAM 4GB RAM 4GB RAM 4GB RAM 4GB RAM 4GB RAM 4GB RAM 6GB RAM 4GB RAM 4GB RAM 4GB RAM 4GB RAM 4GB RAM 6GB RAM	6.1 inches 6.1 inches 5.8 inches 6.5 inches 6.1 inches 6.1 inches 6.1 inches 5.4 inches 5.4 inches 6.1 inches 6.1 inches 6.1 inches 6.1 inches 5.4 inches 5.4 inches 6.1 inches	12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12

Gambar 3 Sampel Cluster 1

K-means Website Dataset Proses K-means Hasil Visualisasi Cari Handphonemu

Hasil Clustering:

Cluster	Device	Processor	RAM	Display Size	Front
Cluster 2	- Huawei Nova 10 Pro - Infinix Note 12 (6GB RAM + 128GB) - Infinix Note 12 5G - Infinix Note 12 Pro - Infinix Note 12 Pro 5G - Infinix Note 12 VIP - Infinix Note 12i (2022) - Infinix Zero 20 - Infinix Zero Ultra - IQOO 9 5G - IQOO Neo 7 (12GB RAM + 256GB) - IQOO Neo 7 5G - IQOO Neo 7 SE 5G - IQOO Neo 8 Pro - Itel P36	- Snapdragon 778G - Helio G88 - Dimensity 810 5G - Helio G99 - Dimensity 810 5G - Helio G96 - Helio G85 - Helio G99 - Dimensity 920 5G - Snapdragon 888+ - Dimensity 8200 - Dimensity 8200 - Dimensity 8200 - Dimensity 9200 Plus - Spreadtrum SC7731E	8GB RAM 6GB RAM 6GB RAM 8GB RAM 8GB RAM 8GB RAM 4GB RAM 8GB RAM 8GB RAM 8GB RAM 12GB RAM 8GB RAM 8GB RAM 12GB RAM 1GB RAM	6.78 inches 6.7 inches 6.7 inches 6.7 inches 6.7 inches 6.7 inches 6.7 inches 6.7 inches 6.8 inches 6.56 inches 6.78 inches 6.78 inches 6.78 inches 6.78 inches 6.5 inches	6C 16 16 16 16 16 8B 6C 32 16 16 16 16 32 8B

Gambar 4 Sampel Cluster 2

License Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License (CC BY 4.0).

Doi: <http://dx.doi.org/xx.xxxxxx>

Vol 1 No 1 (August 2026)

K-means Website Dataset Proses K-means Hasil Visualisasi Cari Handphonemu

Hasil Clustering:

Cluster	Device	Processor	RAM	Display Size	Front
Cluster 3	• Ite! A23s	• Spreadtrum SC9832E	• 2GB RAM	• 5 inches	• 16
	• Ite! A60	• Unisoc SC9832E	• 2GB RAM	• 6.6 inches	• 5A
	• Ite! R21635	• Unisoc SC9832E	• 4GB RAM	• 1.8 inches	• 5A
	• Ite! i5029	• 17MHz Processor	• 32MB RAM	• 2.4 inches	• 5A
	• Ite! Magic 2 4G	• T117	• 64MB RAM	• 2.4 inches	• 5A
	• Ite! Magic X	• T117	• 48MB RAM	• 2.4 inches	• 5A
	• Ite! MagicX Pro	• 17GHz Processor	• 64MB RAM	• 2.4 inches	• 5A
	• Ite! Power 440	• Spreadtrum SC7731E	• 32MB RAM	• 2.4 inches	• 8A
	• Ite! S16 Pro	• Quad Core	• 2GB RAM	• 6.6 inches	• 8A
	• Ite! Vision 3 (2GB RAM + 32GB)	• Unisoc SC9863a	• 2GB RAM	• 6.6 inches	• 5A
	• Ite! JioPhone 2	• Dual Core	• 512MB RAM	• 2.4 inches	• 0.
	• Ite! Phone 3	• Quad Core	• 2GB RAM	• 5 inches	• 2A
	• Karbonn Y11	• 17MHz Processor	• 32MB RAM	• 1.9 inches	• 2A
	• Kechaoda K101 New	• 17MHz Processor	• 32MB RAM	• 2.8 inches	• 2A
	• Kechaoda K55 Pro	• 17MHz Processor	• 32MB RAM	• 1.8 inches	• 2A

Gambar 5 Sampel Cluster 3

K-means Website

Dataset

Proses K-means

Hasil Visualisasi

Cari Handphonemu

Hasil Clustering:

Cluster	Device	Processor	RAM	Display Size	Front
Cluster 4	• Motorola Edge 30 Fusion 5G	• Snapdragon 888+	• 8GB RAM	• 6.55 inches	• 32
	• Motorola Moto G22	• Helio G37	• 4GB RAM	• 6.53 inches	• 16
	• Motorola Moto G32	• Snapdragon 680	• 4GB RAM	• 6.55 inches	• 16
	• Motorola Moto G40 Fusion	• Snapdragon 732G	• 4GB RAM	• 6.78 inches	• 16
	• Motorola Moto G40 Fusion (6GB RAM + 128GB)	• Snapdragon 732G	• 6GB RAM	• 6.78 inches	• 16
	• Motorola Moto G42	• Snapdragon 680	• 4GB RAM	• 6.47 inches	• 16
	• Motorola Moto G52	• Snapdragon 680	• 4GB RAM	• 6.6 inches	• 16
	• Motorola Moto G52 (6GB RAM + 128GB)	• Snapdragon 680	• 6GB RAM	• 6.6 inches	• 16
	• Motorola Moto G60	• Snapdragon 732G	• 6GB RAM	• 6.78 inches	• 32
	• Motorola Moto G62 (8GB RAM + 128GB)	• Snapdragon 695	• 8GB RAM	• 6.55 inches	• 16
	• Motorola Moto G62 5G	• Snapdragon 695	• 6GB RAM	• 6.55 inches	• 16
	• Nokia C12	• Unisoc SC9863A	• 2GB RAM	• 6.3 inches	• 5A
	• Nokia G11 Plus	• Unisoc T606	• 4GB RAM	• 6.52 inches	• 8A
	• Nokia G21	• Unisoc T606	• 4GB RAM	• 6.5 inches	• 8A
	• Nokia G21 (6GB RAM + 128GB)	• Unisoc T606	• 6GB RAM	• 6.5 inches	• 8A
• Nokia G22	• Unisoc T606	• 4GB RAM	• 6.52 inches	• 8A	

Gambar 6 Sampel Cluster 4

Hasil clustering menampilkan 4 cluster produk handphone berdasarkan karakteristiknya. Cluster 1 dan 2 berisi handphone kelas menengah (mid-range) yang cocok untuk konsumen dengan budget terbatas namun fitur cukup baik. Cluster 3 berisi handphone kelas atas (high-end), ideal untuk konsumen dengan anggaran lebih yang mencari kualitas unggul. Cluster 4 mencakup produk near-mid-range, menawarkan keseimbangan antara harga dan kualitas, cocok untuk konsumen yang menginginkan produk dengan performa baik di harga terjangkau.

C. Pengujian

Pengujian hasil clustering dilakukan dengan menggunakan matrix sse dan silhouette score. Matrix sse adalah total dari semua jarak antara setiap data dan pusat cluster. Nilai SSE yang lebih rendah menunjukkan bahwa data dalam setiap cluster lebih homogen, yang berarti kualitas cluster yang dihasilkan semakin baik [14]. Silhouette score adalah ukuran yang menilai seberapa baik posisi suatu titik data dalam klaster dan seberapa jauh dari klaster tetangga. Nilai berkisar antara -1 hingga 1, dimana skor mendekati 1 menunjukkan pengelompokan yang baik, sedangkan skor mendekati -1 mengindikasikan kesalahan klasifikasi. Skor 0 menunjukkan posisi di batas antara dua klaster. Nilai ini dihitung untuk setiap titik data dan dirata-ratakan untuk memperoleh skor siluet keseluruhan. Skor rata-rata yang lebih tinggi mencerminkan pengelompokan yang lebih baik [15].

a) Matrix Sse

Perhitungan SSE:

Cluster	SSE per Cluster	Selisih SSE
Cluster 4	66,89	146,28
Cluster 3	54,1	159,07
Cluster 1	51,58	161,58
Cluster 2	40,6	172,56

Gambar 7 Perhitungan Sse

Gambar di atas menunjukkan hasil perhitungan SSE untuk evaluasi cluster. Cluster 4 memiliki nilai SSE per cluster sebesar 66,89 dengan selisih 146,28. Cluster 3 memiliki nilai SSE 54,1 dan selisih 159,07, sedangkan Cluster 1 mencatat nilai SSE 51,58 dengan selisih 161,58. Nilai SSE yang tinggi menunjukkan jarak data yang cukup jauh dari centroid, mengindikasikan tingkat kekompakan yang rendah. Sebaliknya, Cluster 2 mencatat nilai SSE 40,6 dengan selisih 172,56, menandakan kekompakan data yang tinggi dan kualitas cluster yang lebih baik dibandingkan dengan 3 cluster lainnya.

b) Silhouette Score

Perhitungan Silhouette Score:

Cluster	Nilai Rata-Rata Silhouette
Cluster 1	0,25
Cluster 2	0,15
Cluster 3	-0,01
Cluster 4	0,14

Gambar 8 Perhitungan Silhouette Score

Gambar di atas menunjukkan pembagian data menjadi 4 cluster sesuai inputan peneliti. Cluster 1 memiliki nilai silhouette 0,25, menjadikannya yang terbaik dalam hal kualitas pengelompokan. Cluster 2 dan 4, dengan nilai 0,15 dan 0,14, menunjukkan kualitas pengelompokan yang kurang baik, meskipun tetap positif, menandakan bahwa sebagian besar data masih dekat dengan cluster tersebut. Cluster 3, dengan nilai negatif -0,01, menunjukkan kualitas pengelompokan yang sangat buruk.

Pengujian menggunakan matrix sse dan silhouette menunjukkan hasil yang berbeda terkait kualitas cluster, disebabkan oleh perbedaan parameter masing-masing. Dari analisis sse, cluster 4, 3, dan 1 memiliki nilai tinggi, menandakan bahwa titik data kurang kompak dan jauh dari centroid, sehingga kualitasnya dianggap rendah. Sebaliknya, hasil silhouette menunjukkan bahwa cluster 2 dan 4 memiliki nilai kecil, sementara cluster 3 memiliki nilai negatif, yang menunjukkan bahwa ketiganya memiliki kualitas buruk. Namun, cluster 2 dinyatakan terbaik berdasarkan sse karena memiliki nilai rendah, menunjukkan kompaknya titik data dekat dengan centroid. Sebaliknya, berdasarkan silhouette, cluster 1 dianggap terbaik dengan nilai tinggi, menunjukkan kesamaan data yang baik.

Kesimpulan

Kesimpulan dari penelitian ini adalah bahwa proses clustering dilakukan dengan menentukan jumlah cluster dan iterasi yang diinginkan. Algoritma K-Means berhasil mengelompokkan produk berdasarkan karakteristik, menunjukkan efektivitasnya dalam segmentasi pasar. Analisis karakteristik pada setiap cluster yang terbentuk memungkinkan identifikasi segmen pasar. Selain itu, evaluasi menggunakan metrik SSE dan silhouette memberikan hasil yang jelas mengenai kualitas cluster, sehingga menghasilkan cluster terbaik berdasarkan masing-masing metrik tersebut.

Daftar Pustaka

- [1] Mitzi, Yofanka Theola; Hermana, Cecep. The Pengaruh Keluarga Dan Kelompok Acuan Terhadap Keputusan Pembelian Produk Handphone Samsung (Studi Kasus Masyarakat Karawang). Jurnal Pendidikan Dan Konseling (Jpdk), 2023, 5.2: 379-386..
- [2] Kusuma, Erlin; Sijabat, Yacobo Perdamaian. Analisis Strategi Pemasaran Untuk Meningkatkan Daya Saing Pada Umkm Rumah Makan Ibu Lili. Jurnal Ekonomi, Manajemen Pariwisata Dan Perhotelan, 2023, 2.2: 137-144.
- [3] Muhima, Rani Rotul, Et Al. An Improved Clustering Based On K-Means For Hotspots Data. Indones. J. Electr. Eng. Comput. Sci, 2023, 31.2: 1109-1117.
- [4] Safira, Intan; Salkiawati, Ratna; Priatna, Wowon Priatna. Penerapan Algoritma K-Means Untuk Mengetahui Pola Persediaan Barang Pada Toko Raja Bekasi. Journal Of Informatic And Information Security, 2022, 3.1: 99-110.
- [5] Safira, Intan; Salkiawati, Ratna; Priatna, Wowon Priatna. Penerapan Algoritma K-Means Untuk Mengetahui Pola Persediaan Barang Pada Toko Raja Bekasi. Journal Of Informatic And Information Security, 2022, 3.1: 99-110.
- [6] Wardani, Syafrina Dyah Kusuma, Et Al. Perbandingan Hasil Metode Clustering K-Means, Db Scanner & Hierarchical Untuk Analisa Segmentasi Pasar. Jiko (Jurnal Informatika Dan Komputer), 2023, 7.2: 191-201.
- [7] Indonesia, Bakrie. Analisis Segmentasi Pasar Untuk Mengembangkan Fitur Layanan Strategis Klub Biathlon Dengan Metode Python Kmeans Clustering Penulis: 1) Agung Fitriansyah (2221021048), 2) Rachmat Adi Chandra (2221021049).
- [8] SUSILO, Purnomo Hadi, et al. Klasterisasi Virus Covid-19 Di Wilayah Kabupaten Lamongan Dengan Metode K-Means Clustering. JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika), 2021, 6.2: 325-335.
- [9] E. Effendi, R. Hajar, N. Mukaramah, A. Padang, and S. Alfriansyah, "Pengertian Artikel Dan Editorial," Jurnal Edukasi Nonformal, vol. 3, 2020.
- [10] Saidah, Chaira; Sari, Andini Puspita. Pengaruh Metode Pembelajaran Eksperimen Terhadap Hasil Belajar Ilmu Pengetahuan Alam Madrasah Ibtidaiyah Nurul Ihsan. Eduinovasi: Journal Of Basic Educational Studies, 2024, 4.1: 248-259.
- [11] Febriani, Elsa Selvia, Et Al. Analisis Data Dalam Penelitian Tindakan Kelas. Jurnal Kreativitas Mahasiswa, 2023, 1.2: 140-153.
- [12] Sumiati, Endah; Santoso, Bambang. Perancangan Alat Penyiraman Tanaman Krisan Otomatis Menggunakan Nodemcu Berbasis Internet Of Things (Iot)(Studi Kasus: Riki Flora). Oktal: Jurnal Ilmu Komputer Dan Sains, 2023, 2.02: 513-519.



- [13] Tuasamu, Zainab, Et Al. Analisis Sistem Informasi Akuntansi Siklus Pendapatan Menggunakan Dfd Dan Flowchart Pada Bisnis Porobico. Jurnal Bisnis Dan Manajemen (Jurbisman), 2023, 1.2: 495-510.
- [14] Putra, I. Nyoman Tri Anindia, Et Al. Deteksi Kesegaran Buah Apel, Pisang, Dan Jeruk Menggunakan Metode Transformasi Ruang Warna Hsi Dan K-Nearest Neighbor. Informatics For Educators And Professional: Journal Of Informatics, 2023, 7.2: 120-129.
- [15] Sukanto, Sukanto; Id, Ibnu Daqiqil; Angraini, T. Rahmilia. Penentuan Daerah Rawan Titik Api Di Provinsi Riau Menggunakan Clustering Algoritma K-Means. Juita: Jurnal Informatika, 2018, 6.2: 137-147.