

# Analisis *K-Means* dalam Segmentasi Pasar Penggunaan Handphone di Lingkungan Mahasiswa STMIK Royal

Ade Febriyanti<sup>1</sup>, Putri Vina Bancin<sup>2</sup>, Siska Amanda<sup>3</sup>

<sup>1,2,3</sup> Mahasiswa Program Studi Sistem Informasi, Sekolah Tinggi Manajemen Informatika dan Komputer Royal

<sup>1</sup> [febriyantiade76@gmail.com](mailto:febriyantiade76@gmail.com); <sup>2</sup> [putrivinabancin@gmail.com](mailto:putrivinabancin@gmail.com); <sup>3</sup> [siskaamanda1602@gmail.com](mailto:siskaamanda1602@gmail.com)

\* Email Koresponden

## INFO ARTIKEL

### Histori Artikel

Diterima: 25/Juli/2024

Ditinjau: 28/Juli/2024

Disetujui: 31/Juli/2024

## ABSTRAK

Penggunaan smartphone di Indonesia terus meningkat setiap tahunnya. Di era Revolusi Industri 4.0, smartphone menjadi sektor bisnis yang menguntungkan, sehingga persaingan di pasar semakin ketat. Oleh karena itu, perusahaan smartphone perlu lebih memperhatikan segmentasi pasar yang diinginkan konsumen. Data mining adalah proses menemukan hubungan dan pola signifikan dengan menganalisis kumpulan data besar menggunakan teknik statistik dan matematika. Penelitian ini bertujuan untuk mengetahui dan menganalisis segmen pasar pengguna ponsel Android di kalangan mahasiswa STMIK Royal. Data yang digunakan berasal dari 122 responden mahasiswa. Dalam penelitian ini, diterapkan metode clustering dengan algoritma K-means. Hasil pemodelan data akan mengelompokkan segmentasi pasar ke dalam beberapa cluster. Pengelompokan ini menghasilkan tiga cluster: Cluster 1 (fitur) dengan 36 responden yang mengutamakan harga, baterai, kamera, dan garansi; Cluster 2 (produk) dengan 49 responden yang mengutamakan semua atribut kecuali garansi; dan Cluster 3 (keunggulan) dengan 37 responden yang mengutamakan kamera, merek, dan RAM.

**Kata Kunci:** K-Means; Clustering; Data Mining; Segmentasi Pasar



This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

Copyright ©2024 by Author. Published by PT Beranda Teknologi Academia

## ABSTRACT

The use of smartphones in Indonesia has been steadily increasing each year. In the era of the Fourth Industrial Revolution, smartphones have become a lucrative business sector, leading to intense market competition. Consequently, smartphone companies must pay closer attention to the market segmentation desired by consumers. Data mining is the process of discovering significant relationships and patterns by analyzing large datasets using statistical and mathematical techniques. This study aims to identify and analyze the market segments of Android smartphone users among students at STMIK Royal. The data used in this research were collected from 122 student respondents. The study employs clustering using the K-means algorithm. The resulting data modeling will categorize market segments into several clusters. This segmentation yields three clusters: Cluster 1 (features), consisting of 36 respondents who prioritize price, battery, camera, and warranty; Cluster 2 (product), with 49 respondents who value all attributes except warranty; and Cluster 3 (superiority), comprising 37 respondents who prioritize camera, brand, and RAM.

**Keywords:** K-Means, Clustering, Data Mining, Market Segmentation

## PENDAHULUAN

Perkembangan pengguna smartphone di Indonesia semakin tinggi dari tahun ke tahun. Hal tersebut tak lepas dari maraknya pilihan versi smartphone. Di era Revolusi Industri 4.0, smartphone menjadi sektor bisnis yang patut diperhitungkan karena cukup menguntungkan. Sehingga persaingan di pasar smartphone semakin ketat. Jadi tidak mengherankan jika ada begitu banyak merek pintar di luar sana [1]. Beberapa produsen baru muncul untuk memamerkan produk terbaik mereka.

Telepon genggam, atau smartphone, adalah alat komunikasi jarak jauh yang portabel, mirip dengan telepon tradisional tetapi bisa dibawa ke mana saja. Awalnya, ponsel hanya berfungsi mengirim dan menerima pesan teks (SMS) serta melakukan dan menerima panggilan. Namun, dengan kemajuan teknologi digital, fungsi ponsel telah berkembang. Kini, ponsel dilengkapi dengan fitur-fitur canggih seperti pemutar audio dan video, kamera digital, bluetooth, game, dan akses internet (GPRS, 3G, 4G). Selain itu, ponsel juga menggunakan sistem operasi seperti iOS, Blackberry, Android, dan Windows Mobile, yang membuatnya bekerja hampir seperti komputer [2].

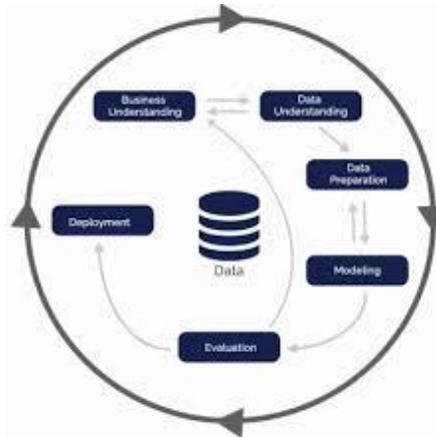
Segmentasi pasar adalah metode yang membagi pasar menjadi beberapa bagian berdasarkan variabel yaitu demografi, geografi, psikografi, perilaku, dan faktor terkait produk [3]. Segmentasi pasar, penargetan, dan positioning bertujuan untuk menempatkan produk atau merek dalam benak konsumen agar memiliki keistimewaan atau keunggulan kompetitif yang berkelanjutan [4]. Harga merupakan faktor penting dalam menarik minat konsumen. Jika harga yang ditetapkan terlalu tinggi atau terlalu rendah, pelanggan bisa menyimpulkan bahwa harga tersebut sesuai dengan ekspektasi mereka [5]. Merek adalah atribut grafis yang bisa berupa logo, kata, gambar, nama, huruf, angka, susunan warna dalam bentuk 2 atau 3 dimensi, suara, hologram, atau kombinasi dari beberapa elemen tersebut untuk membedakan barang atau jasa yang diproduksi oleh individu atau badan hukum dalam kegiatan perdagangan [6]. Random Access Memory (RAM) berbeda dengan ROM; RAM adalah jenis memori yang dapat menyimpan data sementara, di mana perubahan data juga akan tersimpan di dalamnya. [7]. Garansi adalah jaminan yang diberikan kepada individu, perusahaan, atau lembaga lain dalam bentuk surat jaminan [8]. Kamera adalah perangkat terintegrasi yang digunakan untuk menghasilkan gambar dan merekam video, sering kali terdapat pada HP atau smartphone sebagai fitur tambahan [9]. Penambangan data (Data Mining) adalah proses menemukan hubungan atau pola signifikan dengan menganalisis kumpulan data besar yang disimpan dalam memori menggunakan teknik pengenalan pola, seperti teknik statistik dan matematika, untuk mendukung pengambilan keputusan [10].

Penelitian ini bertujuan untuk mengidentifikasi dan menganalisis segmen pasar pengguna ponsel Android di kalangan mahasiswa STMIK Royal. Proses analisis melibatkan enam atribut, yaitu harga, RAM, kamera, baterai, merek ponsel, dan garansi. Dalam penelitian ini, digunakan metode clustering dengan algoritma K-means. K-means adalah algoritma pengelompokan iteratif yang membagi data ke dalam sejumlah cluster K yang telah ditentukan sebelumnya [11]. Cluster adalah kelompok objek yang serupa di antara anggotanya dan berbeda dari objek di cluster lain. K-means adalah algoritma yang sering digunakan dalam proses clustering [12]. Clustering adalah teknik dalam data mining yang berfungsi untuk mengelompokkan sejumlah data menjadi kelompok-kelompok tertentu (cluster). Algoritma clustering digunakan untuk melakukan pengelompokan ini [13]. Pengelompokan tersebut dapat menghasilkan informasi tentang segmentasi pasar dimana terdapat tiga kelompok yaitu *cluster 1* (fitur) dimana responden yang mementingkan atribut harga, baterai, kamera, dan garansi. *Cluster 2* (produk) dimana responden yang mementingkan semua atribut kecuali garansi. *Cluster 3* (keunggulan) dimana responden yang mementingkan atribut kamera, merk, dan ram.

## METODE

Metode penelitian ini mengikuti tahapan proses data mining yang dikenal sebagai Cross-Industry Standard Process for Data Mining (CRISP-DM) [14]. Metodologi CRISP-DM adalah standar data mining yang disusun oleh tiga pelopor pasar data mining: Daimler Chrysler (Daimler Benz), SPSS, dan NCR, serta dikembangkan melalui berbagai workshop antara tahun 1997-1999. CRISP-DM

memiliki enam tahapan analisis, yaitu pemahaman bisnis, pemahaman data, persiapan data, pemodelan, evaluasi, dan penyebaran [15].



Gambar 1. CRISP-DM

**Bussines Understanding**

Tahapan ini bertujuan menganalisis data segmentasi pasar handphone untuk memahami kebutuhan handphone mahasiswa STMIK Royal. Fokus utama adalah menganalisis cara membuat klasifikasi guna mengetahui segmentasi pasar handphone di kalangan mahasiswa STMIK Royal.

**Data Understanding**

Data sekunder dalam penelitian merupakan kuesioner yang diisi oleh 122 responden. Data ini mencakup jawaban mahasiswa terhadap pertanyaan kuesioner yang disusun berdasarkan atribut seperti harga, RAM, kamera, baterai, merek, dan garansi.

```
In [3]: #melihat info dataset
df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 122 entries, 0 to 121
Data columns (total 6 columns):
#   Column      Non-Null Count  Dtype
---  ---
0   harga       122 non-null    object
1   ram         122 non-null    object
2   kamera     122 non-null    object
3   baterai    122 non-null    object
4   merk_hp    122 non-null    object
5   garansi    122 non-null    object
dtypes: object(6)
memory usage: 5.8+ KB
```

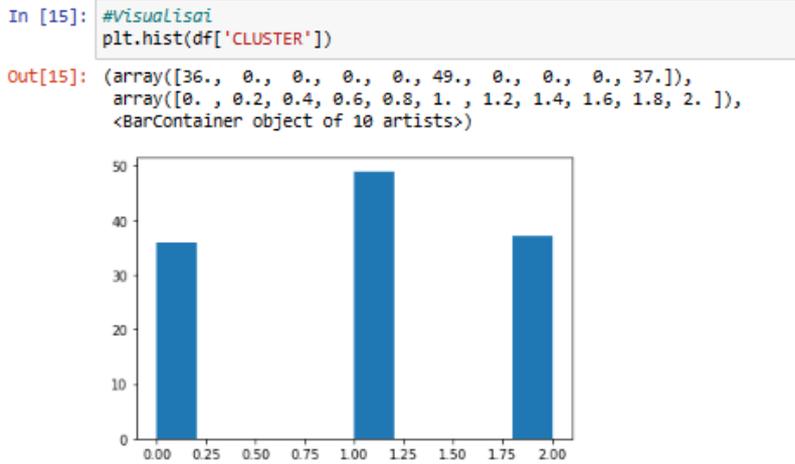
Gambar 2. Deskripsi Data

**Data Preparation**

Tahap ini melibatkan banyak persiapan, sering disebut sebagai proses pembersihan data, di mana data mentah dipilih dan disaring untuk menghilangkan data yang tidak diperlukan dan tidak akurat. Pemilihan field dan tabel juga terjadi pada langkah ini. Kegiatan pada langkah ini meliputi seleksi dan pembersihan data. Atribut yang digunakan untuk data mining dibagi menjadi enam kelompok: harga, RAM, kamera, merek, baterai, dan garansi.

**Modelling**

Pada fase pemodelan, digunakan aplikasi Jupyter Notebook dengan metode K-means. Data atribut yang sudah dipilih pada langkah persiapan data, yaitu harga, RAM, kamera, merek, baterai, dan garansi, digunakan sebagai parameter untuk clustering. Hasil akhirnya adalah prediksi yang bisa dievaluasi berdasarkan nilai akurasi.



Gambar 3. Cluster

**Evaluation**

Pada langkah ini, validasi hasil diukur menggunakan model clustering. Clustering merupakan proses pengelompokan data ke dalam beberapa kelompok sehingga data dalam satu kelompok memiliki kemiripan maksimum dan data antar cluster memiliki kemiripan minimum.

**HASIL DAN PEMBAHASAN**

Data dalam penelitian ini berasal dari kuesioner mahasiswa STMIK Royal. Atribut yang digunakan dalam algoritma ini disederhanakan menjadi enam variabel: harga, RAM, kamera, baterai, merek, dan garansi. Tahap awal analisis adalah mengimpor dataset segmentasi pasar.

```
In [1]: #menampilkan dataset
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

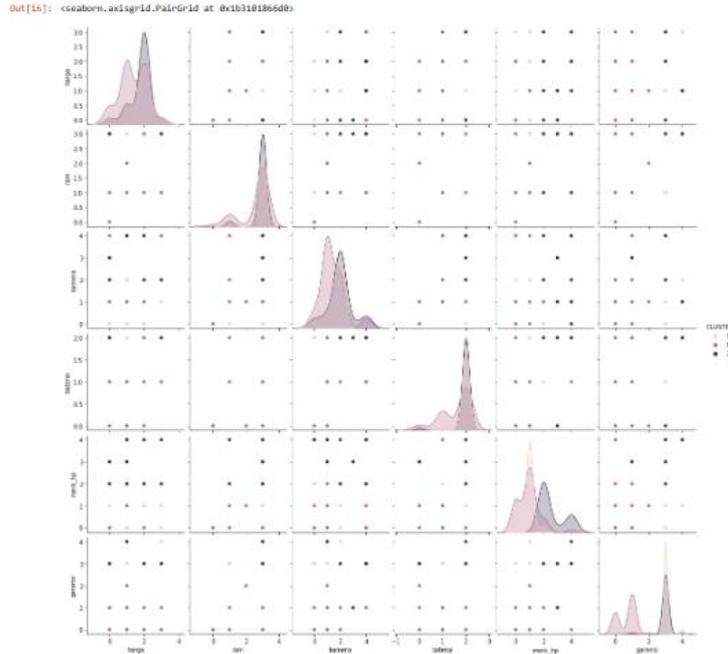
df = pd.read_excel('ML1_kelompok3_Dataset.xlsx')
df
```

```
Out[1]:
```

	Cap waktu	nama_lengkap	semester	harga	ram	kamera	baterai	merk_hp	garansi
0	2022/12/19 12:21:40 PM GMT+7	Ade Febriyanti	7	Penting	Sangat penting	Sangat penting	Sangat penting	Sangat penting	Cukup penting
1	2022/12/19 12:22:25 PM GMT+7	Siska Amanda	7	Penting	Sangat penting	Penting	Penting	Tidak penting	Cukup penting
2	2022/12/19 12:26:01 PM GMT+7	Deby Riyana Egi Ambarwati	7	Sangat penting	Sangat penting	Sangat penting	Sangat penting	Penting	Sangat penting
3	2022/12/19 12:27:16 PM GMT+7	Muhammad isra nurul iman	5	Penting	Sangat penting	Tidak penting	Penting	Cukup penting	Penting
4	2022/12/19 12:27:48 PM GMT+7	Ziki alfa	5	Tidak penting	Penting	Tidak penting	Sangat penting	Tidak penting	Sangat penting
...	...	...	...	...	...	...	...	...	...
117	2022/12/20 5:07:20 PM GMT+7	Herika Wulandari	7	Penting	Sangat penting	Penting	Sangat penting	Cukup penting	Sangat penting
118	2022/12/20 5:13:59 PM GMT+7	Rizki Rahmawati	5	Penting	Sangat penting	Sangat penting	Sangat penting	Cukup penting	Sangat penting
119	2022/12/20 5:57:32 PM GMT+7	Piya Revina Marpaung	1	Sangat penting					
120	2022/12/20 6:16:21 PM GMT+7	Dedi Setiawan	7	Sangat penting	Sangat penting	Penting	Sangat penting	Cukup penting	Penting
121	2022/12/20 6:16:21 PM GMT+7	Ade Irma Suriya Ningsih	7	Penting	Sangat Penting	Penting	Cukup penting	Penting	Sangat Tidak Penting

122 rows x 9 columns

Gambar 4. Dataset



Gambar 5. Hasil Visualisasi Cluster

```
In [18]: #Melihat Masing-masing cluster
df['CLUSTER'].value_counts()

Out[18]: Produk      49
Keunggulan    37
Fitur         36
Name: CLUSTER, dtype: int64
```

Gambar 6. Hasil Cluster

sesuai gambar di atas, hasil algoritma K-means menunjukkan 3 cluster yang terdiri dari 49 responden yang mementingkan produk, 37 responden yang mementingkan keunggulan dan 36 responden yang mementingkan fitur.

**KESIMPULAN**

Data dalam penelitian ini mencakup 122 tabel dengan 6 field, yaitu harga, RAM, kamera, baterai, merek, dan garansi, yang telah disortir sebelumnya. Data tersebut menggunakan algoritma K-means melalui tools Jupyter Notebook, menghasilkan item produk, fitur, dan keunggulan sehingga membantu pihak perusahaan produsen handphone dalam memperhatikan dan melakukan inovasi dan pengembangan produk handphone agar perusahaan dapat menciptakan handphone yang dibutuhkan oleh responden.

## DAFTAR PUSTAKA

- [1] F. Putri, A. Hasibuan, S. Sumarno, and I. Parlina, "Penerapan K-Means pada Pengelompokan Penjualan Produk Smartphone," vol. 1, no. 1, pp. 15–20, 2021.
- [2] Z. A. Gultom, M. K. Kamil Erwansyah S.Kom, and M. Drs. Ahmad Calam S.Kom, "Implementasi Data Mining untuk Memprediksi Penghasilan Penjualan Handphone di Toko Ponsel Takasimura dengan Metode Rough Set," *Jurnal*, pp. 1–14.
- [3] S. Early, "Penerapan K-Means pada Segmentasi Pasar untuk Riset Pemasaran pada," vol. 9, no. 4, pp. 966–973, 2022, doi: 10.30865/jurikom.v9i4.4486.
- [4] M. K. Putri, J. S. F. M. Rahman, F. A. Nursyifa, S. Alfarisi, T. G. S. Putro, and R. Agustin, "Analisis Segmentasi Pasar Dalam Penggunaan Produk Viefresh Di Wilayah Sekitar Kampus Universitas Muhammadiyah Surabaya," *Balanc. Econ. Business, Manag. Account. J.*, vol. 16, no. 2, pp. 156–161, 2019, doi: 10.30651/blc.v16i2.3133.
- [5] E. P. Erik, "Analisis Segmentation Customer Mall Menggunakan Clustering Algoritma K-Means," *J. Ilmu Data*, vol. 2, no. 11, pp. 1–11, 2022, [Online]. Available: <http://ilmudata.org/index.php/ilmudata/article/view/266>
- [6] G. Jotyka and I. G. Ketut Riski Suputra, "Prosedur Pendaftaran Dan Pengalihan Merek Serta Upaya Perlindungan Hukum Terhadap Merek Terkenal Menurut Undang-Undang Nomor 15 Tahun 2001," *Ganesha Law Rev.*, vol. 3, no. 2, pp. 125–139, 2021, doi: 10.23887/blr.v3i2.447.
- [7] M. F. Putra, A. H. Kridalaksana, and Z. Arifin, "Rancang Bangun Alat Pendeteksi Kebocoran Gas LPG Dengan Sensor Mq-6 Berbasis Mikrokontroler Melalui Smartphone Android Sebagai Media Informasi," *Inform. Mulawarman J. Ilm. Ilmu Komput.*, vol. 12, no. 1, p. 1, 2017, doi: 10.30872/jim.v12i1.215.
- [8] S. Kusnaedi, M. Zubair, Z. S. P. Syariah, and undefined 2020, "Penerapan Akad Kafalah pada Jasa Garansi Bank BTN Syariah Parepare," *Ejurnal.Iainpare.Ac.Id*, vol. 2, no. November, 2020, [Online]. Available: <https://ejurnal.iainpare.ac.id/index.php/banco/article/view/1810>
- [9] U. P. Utama, U. P. Utama, M. Fahzurrahman, and U. P. Utama, "PELATIHAN BEAUTY SHOT PHOTOGRAPHY DENGAN KAMERA HANDPHONE PADA MURID TATARIAS PENGANTIN," vol. 2, no. 3, pp. 141–150, 2022.
- [10] A. Darmawan, N. Kustian, and W. Rahayu, "Implementasi Data Mining Menggunakan Model SVM untuk Prediksi Kepuasan Pengunjung Taman Tabebuaya," *STRING (Satuan Tulisan Ris. dan Inov. Teknol.*, vol. 2, no. 3, p. 299, 2018, doi: 10.30998/string.v2i3.2439.
- [11] D. Juliawan, F. Amir, and E. Desi, "Penerapan Data Mining Metode Clustering Pada CV. Secom Infotech Menggunakan Algoritma K-Means," *Pros. Semin. Nas. Ris. Inf. Sci.*, vol. 1, no. September, p. 96, 2019, doi: 10.30645/senaris.v1i0.12.
- [12] P. Pt and S. Tunggal, "Penerapan metode k-means clustering dalam pengelompokan data penjualan barang pada pt. swasti tunggal mandiri," vol. 2, no. 3, pp. 205–210, 2022.
- [13] A. Aditya, I. Jovian, and B. N. Sari, "Implementasi K-Means Clustering Ujian Nasional Sekolah Menengah Pertama di Indonesia Tahun 2018/2019," *J. Media Inform. Budidarma*, vol. 4, no. 1, p. 51, 2020, doi: 10.30865/mib.v4i1.1784.
- [14] A. Prasatya, R. R. A. Siregar, and R. Arianto, "Penerapan Metode K-Means Dan C4.5 Untuk Prediksi Penderita Diabetes," *Petir*, vol. 13, no. 1, pp. 86–100, 2020, doi: 10.33322/petir.v13i1.925.
- [15] Y. P. Sari, A. Primajaya, and A. S. Y. Irawan, "Implementasi Algoritma K-Means untuk Clustering Penyebaran Tuberkulosis di Kabupaten Karawang," *INOVTEK Polbeng - Seri Inform.*, vol. 5, no. 2, p. 229, 2020, doi: 10.35314/isi.v5i2.1457.