

Penerapan Algoritma K-Means Clustering Dalam Mengelompokkan Smartphone Yang Rekomendasi Berdasarkan Spesifikasi

Amri Muliawan Nur^{1*}, Muhammad Saiful², Hariman Bahtiar³, Muhammad Taufik Hidayat⁴

^{1,2,3,4}Program Studi Sistem Informasi, Universitas Hamzanwadi

*muliaamriga@gmail.com

Abstrak

Berbagai jenis smartphone hadir dengan harga dan spesifikasi yang berbeda-beda, menyebabkan penjual kadang-kadang kesulitan dalam memberikan rekomendasi kepada konsumen yang ingin membeli smartphone yang sesuai dengan spesifikasi dan kisaran harga yang diinginkan. Tantangan ini muncul karena sulit bagi penjual untuk mengingat spesifikasi dari setiap smartphone yang dijual. K-Means Clustering bertujuan untuk mengelompokkan data spesifikasi yang ada ke dalam beberapa klaster, di mana data dalam satu klaster memiliki karakteristik yang serupa. Data yang digunakan adalah data private sebanyak 350 data, dengan melibatkan 4 atribut, antara lain Merek, Harga, Baterai, dan Penyimpanan Internal. Dengan membentuk kelompok-kelompok smartphone ini, penjual akan lebih mudah merekomendasikan smartphone yang sesuai kepada pelanggan. Hasil penelitian menunjukkan bahwa berbagai merek smartphone dikategorikan ke dalam tiga kelompok: Kelompok Rekomendasi, yang mencakup 225 item; Kelompok Sangat Direkomendasikan, yang mencakup 98 item; dan Kelompok Kurang Direkomendasikan, yang mencakup 27 item. Pengelompokan ini diharapkan dapat membantu penjual dengan mudah meningkatkan stok smartphone yang direkomendasikan sesuai dengan kebutuhan konsumen dalam hal harga dan spesifikasi.

Kata kunci: AI-Afgani Cellular, K-Means Clustering, Smartphone

Abstract

Various types of smartphones come with different prices and specifications, making it challenging for sellers to recommend the right smartphone to consumers based on their desired specifications and price range. This challenge arises because it's difficult for sellers to remember the specifications of every smartphone they sell. K-Means Clustering aims to group existing specification data into several clusters, where data within each cluster share similar characteristics. The dataset used consists of 350 private data points involving 4 attributes: Brand, Price, Battery, and Internal Storage. By grouping smartphones into clusters, sellers can more easily recommend suitable smartphones to customers. The research results indicate that various smartphone brands are categorized into three groups: the Recommended Group, which includes 225 items; the Highly Recommended Group, which includes 98 items; and the Less Recommended Group, which includes 27 items. This clustering is expected to assist sellers in effectively increasing their stock of recommended smartphones that match consumer needs in terms of price and specifications.

Keywords: AI-Afgani Cellular, K-Means Clustering, Smartphone

1. Pendahuluan

Perkembangan teknologi, khususnya dalam perangkat komunikasi, telah menjadi sangat signifikan di era saat ini. Salah satunya adalah melalui kemajuan teknologi dalam perangkat komunikasi yang menyesuaikan dengan

kebutuhan manusia. Penggunaan smartphone telah menjadi hal umum dalam berkomunikasi dan tidak hanya sebatas itu, tetapi juga digunakan untuk akses informasi, transaksi, hiburan seperti bermain game, dan banyak lagi. Perkembangan teknologi smartphone dari awal kemunculannya

hingga kini telah berlangsung dengan cepat, mendorong persaingan antara perusahaan pembuatnya untuk terus menghadirkan produk terbaru dengan fitur-fitur canggih yang sesuai dengan kebutuhan zaman. Hal ini menjadikan pasar smartphone menjadi semakin kompetitif dengan berbagai merek yang berlomba-lomba menarik perhatian konsumen.. Banyak sekali vendor smartphone yang mengeluarkan berbagai jenis smartphone dalam satu tahun yang membuat bingung konsumennya karena banyak pilihan[1].

Dari Banyaknya jenis smartphone yang beredar, pihak penjual terkadang bingung dalam memberikan rekomendasi pada konsumen utuk dapat membeli smartphone yang diinginkan sesuai dengan spesifikasi dan harga yang ada di dalam smartphone tersebut, hal ini dimungkinkan karena pemilik konter kesulitan untuk mengingat satu persatu spesifikasi smartphone yang dipajang ataupun yang pernah dan akan dijualnya.

Salah satunya yaitu pada Konter Al-Afgani Cellular yang berlokasi di kecamatan masbagik Lombok Timur. Dikarenakan pihak konter tidak pernah mencatat smartphone apa saja yang pernah dijual beserta spesifikasi dan harganya melainkan hanya dengan didokumentasikan melalui gambar atau foto maka dari itu perlu adanya sebuah analisa data yang tepat, yang dapat membantu pihak Konter Al-Afgani Cellular

untuk dapat mengelompokkan spesifikasi dan harga smartphone secara cepat yang menjadi rekomendasi, sehingga dapat memberikan layanan yang baik bagi konsumen.

Dalam penelitian ini penulis mencoba untuk melakukan analisa data dengan prinsip dan konsep dari data mining, dengan tujuan agar sistem rekomendasi penjualan smartphone pada Konter Al-Afgani Cellular lebih cepat sehingga pelayanan pada konsumen akan lebih baik. Data mining merupakan sebuah proses mengolah data menjadi sebuah informasi yang lebih bermakna[2][3].

Dalam pengolahan data dengan konsep data mining, penulis akan menggunakan algoritma K-Means Clustering yang merupakan metode data mining yang bertujuan mengelompokkan data yang ada ke dalam beberapa kelompok yang memiliki ciri identik. Dengan metode K-Means Clustering, setiap kategori smartphone tersebut akan dikelompokkan menjadi 3 cluster, yaitu rekomendasi, paling rekomendasi, dan kurang rekomendasi. Hasil yang diperoleh nanti, untuk data kategori smartphone yang berada pada cluster paling rekomendasi, inilah nantinya yang akan dijadikan sebagai bahan evaluasi bagi pihak konter dalam meningkatkan stok barang tersebut. Diharapkan dengan adanya sebuah sistem tersebut dapat membantu pihak konter dalam memecahkan masalah dengan cepat dan tepat.

Dari penjelasan yang disampaikan di atas maka penulis tertarik mengambil judul "Penerapan Algoritma K-Means Clustering Dalam Mengelompokkan Smartphone Yang Rekomendasi Berdasarkan Spesifikasinya

2. Tinjauan Pustaka

2.1. Penelitian Terkait

Adapun penelitian sebelumnya yang digunakan sebagai acuan dalam penulisan sesuai dengan yang akan di teliti sebagai pendukung penelitian ini, antara lain:

- Yahya dan Mahpuz telah melakukan penelitian dengan judul "Penggunaan Algoritma K-Means untuk Menganalisis Pelanggan Potensial pada Dealer SPS Motor Honda Lombok Timur Nusa Tenggara Barat". Dalam penelitian ini, kriteria pelanggan potensial di Dealer SPS Motor Honda Lombok Timur adalah mereka yang berusia antara 33-46 tahun dan bekerja sebagai wiraswasta atau petani. Hal ini karena mereka memiliki penghasilan di usia produktif. Terdapat dua kelompok pelanggan yang tidak dianggap potensial. Pertama, mereka yang berusia antara 17-32 tahun dan berprofesi sebagai pelajar atau wiraswasta, namun tidak memiliki penghasilan. Kedua, mereka yang berusia antara 47-73 tahun, tidak bekerja, dan berprofesi sebagai petani atau wiraswasta[4].

- Penelitian yang dilakukan oleh Haryani, Dicky Nofriansyah, dan Ita Mariami membahas "Implementasi Data Mining untuk Pengelompokan Buku di Perpustakaan Yayasan Nurul Islam Indonesia Baru dengan Metode K-Means Clustering". Dari hasil analisis penelitian tersebut, disimpulkan bahwa metode K-Means Clustering dapat berhasil diterapkan dalam mengatasi permasalahan pengelompokan buku di Perpustakaan Yayasan Nurul Islam Indonesia Baru. Selain itu, dalam merancang data mining berbasis desktop, penggunaan metode K-Means Clustering juga terbukti efektif dalam menyelesaikan masalah yang dihadapi oleh perpustakaan tersebut [5].

- Penelitian yang dilakukan oleh Lina Listiani, Yoga Handoko Agustin, dan Mochammad Zaenal Ramdhani berjudul "Implementasi Algoritma K-Means Cluster untuk Rekomendasi Pekerjaan Berdasarkan Pengelompokan Data Penduduk". Setelah melakukan analisis, perancangan, implementasi, dan pengujian, diperoleh kesimpulan mengenai rekomendasi jenis pekerjaan berdasarkan data penduduk. Penelitian ini menghasilkan 4 kelompok jenis pekerjaan berdasarkan jenis kelamin, pendidikan, umur, dan jenis pekerjaan dari data penduduk tahun 2017 untuk memberikan rekomendasi jenis pekerjaan kepada

penduduk Kelurahan Bungursari yang belum memiliki pekerjaan. Tingkat akurasi yang diperoleh sebesar 79%, yang diklasifikasikan sebagai fair atau cukup dalam klasifikasi nilai AUC[6].

- Penelitian yang dilakukan oleh Januardi Nasir berjudul "Penerapan Data Mining Clustering dalam Mengelompokkan Buku dengan Metode K-Means" menghasilkan kesimpulan bahwa jumlah buku yang sering dipinjam dan yang jarang dipinjam dapat diidentifikasi melalui pengelompokan menggunakan metode K-Means Clustering. Berdasarkan hasil analisis, ditemukan bahwa jumlah buku yang banyak dipinjam terdapat pada cluster 1 sebanyak 9 item, jumlah buku yang paling sedikit dipinjam terdapat pada cluster 2 sebanyak 15 item, dan jumlah buku yang cukup banyak dipinjam terdapat pada cluster 0 sebanyak 12 item. Penelitian ini membuktikan bahwa penerapan data mining dengan metode K-Means Clustering dapat membantu pihak perpustakaan dalam memahami pola peminjaman buku[7].
- Penelitian yang dilakukan oleh Rizky Adrianto dan Amiq Fahmi dengan judul "Penerapan Metode Clustering dengan Algoritma K-Means untuk Rekomendasi Pemilihan Jalur Peminatan Sesuai Kemampuan pada Program Studi Teknik Informatika - S1 Universitas Dian Nuswantoro" menghasilkan

5 kelompok mahasiswa berdasarkan kemampuan akademis mereka melalui penerapan data mining menggunakan metode Clustering dengan algoritma K-Means. Hasil analisis ini diharapkan dapat membantu program studi Teknik Informatika-S1 untuk memberikan rekomendasi yang sesuai kepada mahasiswa mengenai pemilihan jalur peminatan yang cocok dengan kemampuan akademis mereka di masa depan[8].

- Berdasarkan penelitian yang dilakukan oleh Putri Trisnawati dan Ade Irma Purnamasari dalam artikel mereka yang berjudul "Penerapan Pengelompokan Produktivitas Hasil Pertanian Menggunakan Algoritma K-Means", hasilnya adalah sebagai berikut : Cluster 1 memiliki 20 anggota, Cluster 2 memiliki 30 anggota, Cluster 3 memiliki 205 anggota, Cluster 4 memiliki 1 anggota, Cluster 5 memiliki 25 anggota, Cluster 6 memiliki 49 anggota, Cluster 7 memiliki 3 anggota, Cluster 8 memiliki 83 anggota, Cluster 9 memiliki nilai k terbaik dari Davies-Bouldin Index (DBI) sebesar 0,196, yang mendekati nilai 0. Hasil ini menggambarkan distribusi dan komposisi klaster yang dihasilkan oleh algoritma K-Means dalam analisis produktivitas hasil pertanian di kecamatan Bojonegoro berdasarkan dataset yang digunakan[9]

2.2. Landasan Teori

1. Data Mining

Data mining merujuk pada proses mengeksplorasi pengetahuan dari dalam basis data, dimana teknik statistik, matematika, kecerdasan buatan, dan pembelajaran mesin digunakan untuk mengekstraksi dan mengenali informasi berharga serta pengetahuan terkait dari berbagai basis data yang besar. Data mining memiliki beberapa kategori, seperti deskripsi, estimasi, prediksi, klasifikasi, pengelompokan, dan asosiasi[10][11].

2. Implementasi

Implementasi adalah proses pembentukan hubungan yang dijalankan untuk mencapai tujuan tertentu. Ini umumnya terjadi sebagai hasil dari suatu kegiatan di mana berbagai sarana telah direncanakan dan digunakan dengan harapan mencapai tujuan yang ditetapkan. Diharapkan menghasilkan ide, konsep, kebijakan, atau inovasi dalam tindakan yang dilakukan[12].

3. Algoritma K-means

K-Means merupakan algoritma analisis kluster non-hierarkis yang digunakan untuk mengelompokkan data. Analisis kluster sendiri adalah sebuah metode yang digunakan untuk mengelompokkan data berdasarkan variabel atau karakteristik tertentu. Tujuan dari K-Means, seperti halnya metode pengelompokan lainnya, adalah untuk mengelompokkan data dengan cara memaksimalkan kesamaan fitur di dalam kelompok dan memaksimalkan perbedaan antara

kelompok. Algoritma K-Means mengelompokkan data dengan cara menghitung jarak antara setiap data dengan centroid kluster, yang diperoleh melalui proses iteratif. Untuk menjalankan analisis, perlu ditentukan jumlah K sebagai input ke dalam algoritma tersebut[13].

Berikut adalah langkah-langkah untuk melakukan clustering dengan metode K-Means Clustering:

- 1) Tentukan jumlah cluster (k) yang di gunakan untuk menentukan jumlah cluster yang diinginkan.
- 2) Tentukan pusat cluster awal (centroid) sebanyak K.
- 3) Tentukn jarak terdekat setiap data pada pusat centroid dengan persamaan Euclidean Distance berikut

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

Keterangan :

- x = data x
- y = data y
- n = jumlah (atribut)
- d (x, y) = jarak antara data x dan y

- 4) Kelompokkan data berdasarkan titik centroid yang terdekat
- 5) Tentukan nilai *centroid* baru berdasarkan nilai cluster bersangkutan dengan Rumus :

$$\mu_j(t + 1) = \frac{1}{N_{sj}} \sum_{j \in S_j} x_j$$

Keterangan :

- $\mu_j(t + 1)$ = *centroid* iterasi ke (t+1)
- N_{sj} = jumlah data *cluster* S_j

- 6) Ulangi langkah 3 hingga, hingga anggota tiap *cluster* tidak ada yang berubah[14].

4. Metode *Clustering*

Clustering merupakan suatu Teknik dalam data mining yang digunakan dalam pengelompokan data berdasarkan pada kesamaan atribut data. Teknik ini dapat digunakan dalam menemukan karakteristik data yang memiliki kesamaan pola dari suatu data yang tersembunyi[15].

3. Metode Penelitian

3.1 Metode Pengumpulan Data

Data yang digunakan dalam penelitian ini adalah berupa data *private* yang diperoleh dari Konter Al-Afgani Cellular yang terletak di Jl. Masbagik-Pancor, Desa Masbagik Selatan Kecamatan Masbagik Kabupaten Lombok Timur. Dengan jumlah data sebanyak 350 data, dengan melibatkan 4 atribut, antara lain Merek, Harga, Baterai, dan Penyimpanan *Internal*. Untuk memperoleh data yang sesuai untuk keperluan penelitian, peneliti menggunakan berbagai metode pengumpulan data agar hasil yang didapatkan akurat dan dapat digunakan dalam penelitian selanjutnya. Berikut beberapa metode pengumpulan data yang digunakan antara lain:

1. Observasi

Teknik observasi yang diterapkan melibatkan pengumpulan data melalui pengamatan atau observasi langsung terhadap suatu peristiwa yang berhubungan dengan objek penelitian. Oleh karena itu, penelitian ini menggunakan metode observasi untuk pengumpulan data yang

dilakukan secara langsung di lapangan yakni di Konter Al-Afgani Cellular guna mendapatkan data yang sebenarnya.

2. Wawancara

Pengumpulan data dengan cara melakukan komunikasi dan wawancara secara langsung dengan pemilik Konter Al-Afgani Cellular, guna memperoleh data pengelompokan *Smartphone* terkait kondisi saat ini.

3. Studi Pustaka

Peneliti menggunakan metode pengumpulan data dengan membaca dan mempelajari buku-buku yang relevan serta mencari referensi di internet terkait artikel atau jurnal yang berkaitan dengan penelitian ini.

3.2 Lokasi Penelitian

Penelitian ini berlokasi di Konter Al-Afgani Cellular yang terletak di Jl. Masbagik-Pancor, Desa Masbagik Selatan Kecamatan Masbagik Kabupaten Lombok Timur Nusa Tenggara Barat.

4. Hasil dan Pembahasan

4.1 Hasil

Data diolah menggunakan tools rapidminer dengan menerapkan algoritma *K-Means*. Jumlah data yang digunakan adalah sebanyak 350 data, dengan melibatkan 4 atribut, antara lain Merek, Harga, Baterai, dan Penyimpanan *Internal*. Yang nantinya Data dikelompokkan menjadi 3 Cluster yaitu kelompok *Smartphone* rekomendasi, paling

rekomendasi, dan kurang rekomendasi. Berikut akan dijelaskan Langkah-langkah dalam pengolahannya:

1. Melakukan *selection* Data

Seleksi data dilakukan untuk membantu dalam pengolahan data dengan mengidentifikasi pola yang relevan. Oleh karena itu, hanya atribut-atribut yang dianggap penting yang akan digunakan. Berikut ini adalah tabel atribut yang dipilih untuk pengolahan data:

Tabel 1. Seleksi atribut yang digunakan

No	Atribut yang digunakan	Keterangan
1	No	Id
2	Merek	Atribut
3	Harga	Atribut
4	Baterai	Atribut
5	Penyimpanan <i>Internal</i>	Atribut

2. Transformasi Data

Pada Langkah ini, akan dilakukan inialisasi data dengan mengubah format data ke bentuk yang sesuai untuk proses pengolahan selanjutnya. Beberapa data akan diubah terlebih dahulu menjadi format numerik, agar dapat diolah menggunakan algoritma K-Means. Langkah inialisasi yang akan dilakukan adalah sebagai berikut:

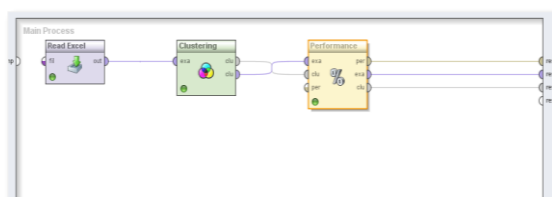
Tabel 2. Inialisasi Merek

No	Merek	Inialisasi
1	Asus	1
2	Infinix	2
3	Apple	3
4	Oppo	4
5	Poco	5
6	Realme	6
7	Xiaomi	7

8	Samsung	8
9	Vivo	9

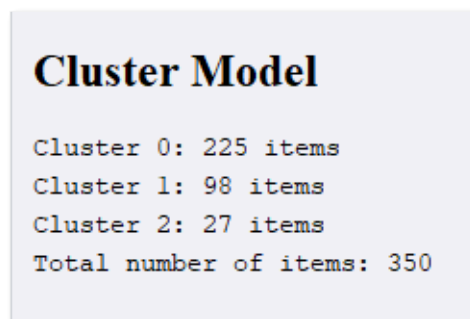
3. Eksperimen dan Pengujian Model

Dataset diuji dengan algoritma K-Means menggunakan tools RapidMiner. Berikut proses pengujian yang dilakukan.



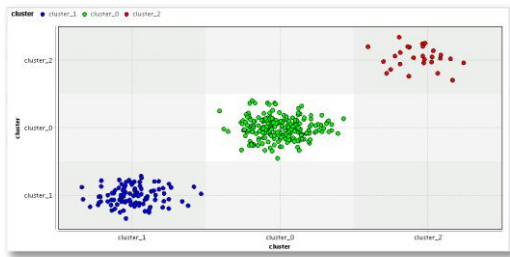
Gambar 1. Pengujian Menggunakan Algoritma K-Means Clustering

Berikut adalah hasil cluster yang diperoleh setelah diolah menggunakan algoritma K-Means menggunakan Tools *RaoidMiner*.



Gambar 2. Cluster Model

Gambar di atas merupakan hasil proses *clustering*, yang dihasilkan menggunakan algoritma K-Means dalam 3 cluster model. Hasil tersebut menunjukkan bahwa dimana hasil ini masih dalam bentuk model yang umum yang artinya belum dijelaskan secara detail mengenai hasil *clustering* itu sendiri, yang selanjutnya hasil tersebut dipetakan kedalam diagram scatter, seperti gambar di bawah ini



Gambar 3. Scatter Hasil Clustering

kelompok *cluster* 1 yang merupakan kelompok *Smartphone* rekomendasi di konter Al-Afgani Cellular berjumlah 225 items ditunjukkan pada gambar berwarna hijau pada diagram scater, yang dimana kelompok ini merupakan keterangan kelompok dengan rentang harga 2.000.000 sampai dengan 1.530.000 yang terdiri dari merek Infinix, Apple, Oppo, Poco, Realme, Xiaomi, Samsung, dan Vivo. kelompok *cluster* 2 paling rekomendasi berjumlah 98 items ditunjukkan pada gambar berwarna biru pada diagram scater, dimana kelompok ini merupakan kelompok dengan rentang harga 1.545.000 sampai dengan 3.100.000 yang terdiri dari merek Asus, Vivo, Samsung, Xiaomi, Realme, Poco, Oppo, Apple, Infinix, dan kelompok *cluster* 3 yang merupakan kelompok kurang rekomendasi berjumlah 27 items ditunjukkan pada gambar berwarna merah pada diagram scater, dimana kelompok ini merupakan kelompok dengan rentang harga 3.150.000 sampai dengan 7.500.000 yang terdiri dari merek Oppo, Apple, Poco, Samsung, dan Vivo. Pengelompokan ini didasarkan pada merek, harga, baterai, dan penyimpanan *internal* yang digunakan.

4.2 Pembahasan

Pada pembahasan ini penulis akan menjabarkan tahapan yang akan dilakukan dalam mengolah data menggunakan algoritma *K-Means*. Pengolahan data dihitung menggunakan pendekatan perhitungan jarak menggunakan korelasi Euclidean Distance.

Berikut adalah hasil yang diperoleh setelah dilakukan pengujian dan langkah-langkah dalam mengelompokkan data menggunakan algoritma *K-Means* :

1. Siapkan dataset *smartphone* yang sudah di *cleaning* sebelumnya.
2. Tentukan jumlah *cluster* (*k*). Adapun *cluster* yang digunakan pada penelitian ini ditentukan sebanyak 3 *cluster* (kelompok) data. Dari 3 *cluster* data tersebut dikelompokkan dalam kategori yaitu rekomendasi, paling rekomendasi, dan kurang rekomendasi
3. Tentukan titik pusat (*centroid*) awal secara acak yang diambil dari dataset. Titik pusat (*centroid*) yang digunakan bisa dilihat pada tabel dibawah ini :

Tabel 3. Centroid Awal

Cluster	Merek	Harga	Baterai	Penyimpanan Internal
C1	2	750000	4000	32
C2	4	2400000	4000	256
C3	7	3100000	5000	128

Berikutnya hitung jarak terdekat dari titik *centroid* terhadap seluruh data menggunakan persamaan di bawah ini dengan menggunakan rumus korelasi *Euclidean Distance*.

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

Berikut adalah hasil perhitungan jarak pada data 1 ke masing-masing *centroid* awal. Langkah ini selanjutnya disebut sebagai iterasi 1 :

$$\begin{aligned} d_1 &= \sqrt{(1-2)^2 + (1.650.000-750.000)^2 +} \\ &\quad \sqrt{(3.300-4.000)^2 + (64-32)^2} \\ &= \sqrt{(-1)^2 + (900.000)^2 +} \\ &\quad \sqrt{(-700)^2 + (32)^2} \\ &= \sqrt{(1) + (810.000.000.000) +} \\ &\quad \sqrt{(490.000) + (1.024)} \\ &= \sqrt{810.000.491.025} \\ &= 900.000,273 \end{aligned}$$

Dari hasil perhitungan diatas didapatkan hasil jarak data pertama dengan *centroid* pertama adalah 900.000,273.

$$\begin{aligned} d_2 &= \sqrt{(1-4)^2 + (1.650.000-2.400.000)^2 +} \\ &\quad \sqrt{(3.300-4.000)^2 + (64-256)^2} \\ &= \sqrt{(-3)^2 + (-750.000)^2 +} \\ &\quad \sqrt{(-700)^2 + (-192)^2} \\ &= \sqrt{(9) + (562.500.000.000) +} \\ &\quad \sqrt{(490.000) + (36.864)} \\ &= \sqrt{562.500.526.873} \\ &= 750.000,351 \end{aligned}$$

Hasil perhitungan jarak data pertama dengan *centroid* kedua adalah 750.000,351.

$$d_3 = \sqrt{(1-7)^2 + (1.650.000-3.100.000)^2 +}$$

$$\begin{aligned} &\sqrt{(3.300-5.000)^2 + (64-128)^2} \\ &= \sqrt{(-6)^2 + (-1.450.000)^2 +} \\ &\quad \sqrt{(-700)^2 + (-64)^2} \\ &= \sqrt{(36) + (2.102.500.000.000) +} \\ &\quad \sqrt{(2.890.000) + (4095)} \\ &= \sqrt{2.102.502.894.132} \\ &= 1.450.000,998 \end{aligned}$$

Hasil perhitungan jarak data pertama dengan *centroid* ketiga adalah 1.450.000,998.

Selanjutnya dengan cara yang sama hitung jarak data yang lain dengan titik *centroid*.

4. Setelah menghitung jarak antara data dengan titik pusat *centroid*, selanjutnya tentukan jarak minimum setiap data dengan pusat *centroid*.
5. Apabila hasil Analisa yang dilakukan secara matematis belum sama dengan hasil rapid miner maka dilakukan proses itrasi selanjutnya, dengan menentukan titik *centroid* baru.
6. Setelah mendapatkan titik *centroid* yang baru di setiap *cluster*, ulangi langkah ketiga hingga tidak ada *cluster* yang berubah dan tidak ada data yang berpindah dari satu *cluster* ke *cluster* lainnya.

Hasi pengujian ini dilakukan sebanyak 5 kali iterasi, hal ini disebabkan karena hasil akhir yang peroleh dengan perhitungan matematis menggunakan excel sama dengan peroleh yang dilakukan menggunakan rapidminer.

Berikut merupakan hasil perbandingan perhitungan matematis menggunakan excel dan pengolahan menggunakan rapidminer setelah dilakukan 5 kali *itrasi*.

Tabel 4. Perbandingan Hasil

Cluster	Rapid Miner	Excel	Persentase
C0	225	225	100%
C1	98	98	100%
C2	27	27	100%

Berdasarkan hasil perhitungan data yang dilakukan menggunakan RapidMiner dan *Microsoft Excel* untuk persentase kemiripan yang dihasilkan 100% sama. Hasil ditunjukkan pada gambar di atas

5. Kesimpulan

Berdasarkan hasil analisa yang dilakukan maka dapat ditarik kesimpulan bahwa algoritma K-Means Clustering dapat memudahkan dalam mengelompokkan data penjualan smartphone di Konter Al-Afgani Cellular dengan baik, sehingga berdasarkan hasil pengelompokan smartphone tersebut dapat menjadi fokus pihak konter dalam menekan angka peningkatan stok barang yang paling rekomendasi. Untuk smartphone yang masuk ke dalam pengelompokan produk Paling Rekomendasi, maka pihak konter akan lebih memperhatikan ketersediaannya, sehingga tidak terjadi kekurangan stok barang. Dan untuk barang yang Kurang Rekomendasi maka pihak konter tidak perlu melakukan stok barang yang berlebih,

yang akan mengakibatkan jumlah stok barang menumpuk

6. Daftar Pustaka

- [1] Merangin, "Sistem informasi geografis pencarian konter servis handphone di wilayah Yogyakarta," *Galang Tanjung*, no. 2504, pp. 1–9, 2018.
- [2] A. M. Nur, M. F. Wazdi, B. Harianto, and M. F. Zaini, "Implementation of Naive Bayes Algorithm in Analyzing Acceptance of Poor Student Assistance," *J. Phys. Conf. Ser.*, vol. 1539, no. 1, 2020, doi: 10.1088/1742-6596/1539/1/012018.
- [3] A. M. Nur, N. Nurhidayati, and I. Fathurrahman, "Penerapan Metode Naive Bayes Untuk Penentuan Penerima Beasiswa Program Indonesia Pintar (PIP).," *Infotek J. Inform. dan Teknol.*, vol. 7, no. 1, pp. 93–102, 2024.
- [4] Saverus, "Penggunaan Algoritma K-Means Untuk Menganalisis Pelanggan Potensial Pada Dealer SPS Motor Honda Lombok Timur Nusa Tenggara Barat," *J. Kaji. Pendidik. Ekon. dan Ilmu Ekon.*, vol. 2, no. 1, pp. 1–19, 2019.
- [5] I. M. Haryani, Dicky Nofriansyah, "Implementasi Data Mining Untuk Pengelompokan Buku Di Perpustakaan Yayasan Nurul Islam Indonesia Baru Dengan Metode K-Means Clustering," *J. Cyber TechTech*, vol. 1, no. 1, pp. 1–12, 2021.
- [6] M. Yahya, "Penggunaan Algoritma K-Means Untuk Menganalisis Pelanggan Potensial Pada Dealer SPS Motor Honda Lombok Timur Nusa Tenggara Barat," *J. Inform. dan Teknol.*, vol. 2, no. 2, pp. 109–118, 2019.
- [7] J. Nasir, "Penerapan Data Mining Clustering Dalam Mengelompokan Buku Dengan Metode K-Means," *Simetris J. Tek. Mesin, Elektro dan Ilmu Komput.*, vol. 11, no. 2, pp. 690–703, 2021, doi:

- 10.24176/simet.v11i2.5482.
- [8] R. Adrianto and A. Fahmi, "Penerapan Metode Clustering Dengan Algoritma K-Means Untuk Rekomendasi Pemilihan Jalur Peminatan Sesuai Kemampuan Pada Progam Studi Teknik Informatika - S1 Universitas Dian Nuswantoro," *JOINS (Journal Inf. Syst.*, vol. 1, no. 2, pp. 101–116, 2019.
- [9] P. Trisnawati and A. I. Purnamasari, "Penerapan Pengelompokan Produktivitas Hasil Pertanian Menggunakan Algoritma K-Means," *Infotek J. Inform. dan Teknol.*, vol. 6, no. 2, pp. 249–257, 2023.
- [10] D. Jollyta, W. Ramdhan, and M. Zarlis, *Konsep data mining dan penerapan*. Deepublish, 2020.
- [11] P. W. Rahayu *et al.*, *Buku Ajar Data Mining*. PT. Sonpedia Publishing Indonesia, 2024.
- [12] B. T. Haji, "Pengertian Implementasi," *Lap. AKHIR*, vol. 31.
- [13] A. Wicaksana and T. Rachman, "Implementasi Algoritma K-Means dalam Analisis Klasterisasi Penyebaran Penyakit Hiv/Aids," *Angew. Chemie Int. Ed.* 6(11), 951–952., vol. 3, no. 1, pp. 10–27, 2018.
- [14] S. Handoko, F. Fauziah, and E. T. E. Handayani, "Implementasi Data Mining Untuk Menentukan Tingkat Penjualan Paket Data Telkomsel Menggunakan Metode K-Means Clustering," *J. Ilm. Teknol. dan Rekayasa*, vol. 25, no. 1, pp. 76–88, 2020, doi: 10.35760/tr.2020.v25i1.2677.
- [15] Z. Setiawan *et al.*, *BUKU AJAR DATA MINING*. PT. Sonpedia Publishing Indonesia, 2023.