

Pengelompokan Minat Konsumen Pengguna Jasa Pengiriman pada CV. Lima Benua Nusa Indonesia Berdasarkan Algoritma K-Means

Hendra Jaya Zega¹, Peber Epenetus Halawa², Taronisokhi Zebua³

¹Universitas Budi Darma Medan, Indonesia, email: hendrazega09@gmail.com

²Universitas Budi Darma Medan, Indonesia, email: epenetushalawa@gmail.com

³Universitas Budi Darma Medan, Indonesia, email: taronizeb@gmail.com

*coresponding author)

Info Artikel

Diajukan: 12 Agustus 2024

Diterima: 13 Agustus 2024

Diterbitkan: 14 Agustus 2024

Kata Kunci:

*Pengelompokan;
Minat Konsumen;
Jasa Pengiriman;
CV. Lima Benua Nusa
Indonesia;
Algoritma K-Means.*

Keywords:

Keywords:
Grouping;
Consumer Interests;
Delivery Service;
CV. Lima Benua Nusa
Indonesia;
K-Means Algorithm.



Lisensi: *cc-by-sa*

Copyright © 2024 by Author. Published by
Faatuatua Media Karva

Abstrak

Salah satu metode yang dapat digunakan untuk menemukan pengetahuan dalam suatu tumpukan data yang cukup besar dengan proses menggali dan menganalisa sejumlah data yang sangat besar dalam memperoleh sesuatu yang benar, baru dan bermanfaat adalah teknik data mining sehingga dapat ditemukan suatu corak atau pola dalam data tersebut pelayanan secara umum adalah setiap kegiatan yang diperuntukkan atau ditujukan untuk memberikan kepuasan kepada pelanggan, melalui pelayanan ini keinginan dan kebutuhan pelanggan dapat terpenuhi. Pelayanan adalah sebagai usaha melayani kebutuhan orang lain, sedangkan melayani yaitu membantu menyiapkan (membantu apa yang diperlukan seseorang). Pelayanan pada hakekat adalah konsep atau praktik memberikan bantuan atau kebaikan kepada orang lain. Secara umum, pelayanan melibatkan tindakan atau upaya untuk membantu, melayani, atau memenuhi kebutuhan orang lain dengan cara yang bertanggung jawab dan peduli. Proses pelayanan berlangsung secara rutin dan berkesinambungan meliputi seluruh kehidupan orang dalam masyarakat, proses pemenuhan kebutuhan melalui aktivitas orang lain termasuk juga CV. Lima Benua Nusa Indonesia merupakan perusahaan yang menawarkan pelayanan yang bergerak dalam bidang pengiriman dan logistik yang bermarkas di Medan, Indonesia. Algoritma K-Means merupakan Algoritma pengelompokan iteratif yang melakukan partisi set data ke dalam sejumlah K cluster yang sudah ditetapkan di awal.

Abstract

Data mining is a method for finding knowledge in a fairly large pile of data by the process of digging and analyzing a very large amount of data in order to obtain something true, new and useful so that a style or pattern can be found in the data. Service in general is every activity which is intended or intended to provide satisfaction to customers, through this service customer desires and needs can be fulfilled. Service is an effort to serve other people's needs, while serving is helping to prepare (helping someone with what they need). Service is essentially the concept or practice of providing assistance or kindness to others. In general, service involves actions or efforts to help, serve, or meet the needs of others in a responsible and caring manner. As a service process that takes place routinely and continuously, covering all the lives of people in society, the process of fulfilling needs through the activities of other people, including CV. Lima Benua Nusa Indonesia is a company that offers services in the field of shipping and logistics, headquartered in Medan, Indonesia. The K-Means algorithm is an iterative clustering algorithm that partitions a data set into a number of K clusters that have been determined at the beginning.

1. PENDAHULUAN

Keberadaan Jasa pengiriman barang memiliki peran yang sangat strategis untuk meningkatkan kinerja arus barang dari dan tujuan suatu wilayah, Peran Utama barang dari tempat asal pengiriman menuju ketempat tujuan, Peran nya yaitu sebagai sarana transportasi barang yang meliputi penyediaan

akses transportasi barang baik dalam jarak dekat (dalam kota) maupun jarak jauh (luar kota). CV. Lima Benua Nusa Indonesia merupakan salah satu perusahaan yang berkembang jasa pengiriman barang yang berdomisili di kota Medan, selanjutnya disebut JNE.

Masalah dalam pengiriman barang selal terlambat yang menyebabkan kerugian dan berdampak pada mina konsumen di setiap provinsi. Masalah pengiriman ini selalu dicatat dan diolah untuk menemukan pola tertentu dalam pengiriman barang, CV. Lima Benua Nusa Indonesia yaitu salah satu pelayanan pengiriman barang, dimana memiliki aktivitas logistik pengiriman barang yang begitu banyak. sehingga data pengiriman barang dapat diketahui pola yang sangat membantu dalam memajukan CV. Lima Benua Nusa Indonesia.

Dalam hal ini penulis berencana mengelompokkan kembali data minat konsumen terhadap jasa pengiriman yang ada pada CV. Lima Benua Nusa Indoensia agar dapat menarik kembali minat masyarakat untuk Jasa pengiriman pada CV. Nusa Lima Nusantara tersebut. Untuk mempermudah dalam pengelompokan minat konsumen terhadap jasa pengiriman maka penelitian ini menggunakan metode data mining. Salah satu tugas utama data mining adalah Clustering. Clustering adalah proses pengorganisasian objek ke dalam class atau kelompok dengan cara mencari kemiripan setiap objek[1].

Solusi yang dapat digunakan untuk menyelesaikan masalah yang sedang dialami oleh CV. Lima Benua Nusa Indonesia adalah data mining. Data mining adalah proses ekstraksi pengetahuan atau pola yang bermanfaat dari jumlah data yang besar atau kompleks. Tujuan utama dari data mining adalah untuk mengidentifikasi pola tersembunyi, hubungan, atau tren yang tidak terlihat secara langsung melalui analisis konvensional. Teknik data mining melibatkan penggunaan berbagai metode statistik, matematika, kecerdasan buatan, dan teknik komputasi lainnya untuk menganalisis dataset besar dengan tujuan mendapatkan wawasan yang berharga. Tujuan utama dari data mining adalah untuk mengekstraksi pengetahuan yang berharga atau pola yang tersembunyi dari dataset besar atau kompleks.

Penelitian ini menguraikan penerapan teknik data mining untuk menyelesaikan permasalahan yang telah dijabarkan di atas dengan algoritma K- Means Clustering. Metode yang digunakan adalah metode clustering. Metode clustering adalah suatu proses pengelompokan objek data yang mirip satu sama lain kedalam kluster yang sama dan berbeda dengan objek yang ada di kluster lain. Algoritma yang digunakan dalam penelitian ini untuk pengelompokan minat konsumen terhadap pengguna jasa pengggiriman adalah algoritma K-Means[2].

Algoritma K-Means merupakan algoritma yang menjadi salah satu algoritma yang paling penting dalam bidang data mining, karena memiliki kelebihan sebagai algoritma yang mudah diimplementasikan, relatif cepat ditinjau dari waktu komputasi dan telah digunakan secara luas untuk menyelesaikan berbagai persoalan komputasi dan Algoritma K-Means bisa dibilang adalah metode pengelompokan yang paling populer, dan jika menggunakan metode K-means, harus melakukan transmisi data agar dapat diolah[3][4].

2. METODE PENELITIAN

Melalui metode penelitian, maka akan terjabarkan bagaimana penelitian ini diselesaikan serta teori-teori yang digunakan sebagai dasar pemikiran atau pustaka dalam menyelesaikannya.

2.1 Tahap Penelitian

Tahapan peneilitian ini dibuat dengan tujuan untuk menyajikan proses yang dilakukan dalam menyelesaikan penelitian ini secara terstruktur dan sistematis. Lebih jelas dapat dilihat pada gambar berikut :



Gambar 1. Tahapan Penelitian

Berdasarkan diagamam di atas, maka dapat diuraikan bawah:

1. Identifikasi Masalah

Tahap ini dilakukan untuk mengidentifikasi masalah untuk mengetahui permasalahan yang terjadi pada saat melakukan pengolompokan minat konsumen terhadap jasa pengiriman pada CV. Lima Benua Nusa Indonesia. Sehingga dapat diberikan solusi yang tepat untuk memecahkan permasalahan yang ada.

2. Pengumpulan Data

Pada tahap ini dilakukan dengan dua cara yaitu:

- a. Observasi
Tahapan ini merupakan kegiatan pengumpulan data melalui pengamatan secara langsung
 - b. Wawancara
Tahapan ini merupakan kegiatan pengumpulan data melalui interaksi langsung dengan pihak CV. Lima Benua Nusa Indonesia mengenai jasa pengiriman.
 - c. Studi Pustaka
Tahap Studi pustaka merupakan suatu cara untuk mengumpulkan referensi dan penelitian dan buku dari penelitian yang sebelumnya sehingga dapat dijadikan sebagai acuan dalam mempertanggung jawabkan metode yang digunakan dalam penelitian yang akan dibuat. Seperti jurnal atau artikel yang berhubungan dengan metode atau Algoritma K- Means Clustering.
3. Analisa Data
Analisa Data merupakan tahap yang dilakukan untuk menguraikan jawaban pertanyaan pada perumusan masalah dan juga merupakan hasil yang telah didapat pada pengamatan langsung di kantoh CV. Lima Benua Nusa Indonesia tentang pengelompokan jasa pengiriman
 4. Penerapan algoritma K-Means Clustering .
Penerapan algoritma K-Means Clustering adalah merupakan tahap untuk melakukan teknik penyelesaian masalah dengan menyesuaikan dengan aturan atau langkah-langkah yang ada pada Clustering. Sampai mendapatkan hasil perankingan untuk dapat mengambil keputusan.
 5. Perancangan Sistem
Perancangan system merupakan suatu cara untuk membuat sistem aplikasi untuk menguji hasil penerapan algoritma K-Means Clustering berbasis komputer, dengan tujuan untuk mengefesienkan pekerjaan dalam pengelompokan minat konsumen terhadap pengguna jasa pengiriman dan dapat mengambil keputusan dengan cepat.
 6. Pengujian
Pengujian merupakan tahap akhir dalam penyelesaian masalah dengan menguji system yang sudah dirancang dan juga analisa masalah, penerapan algoritma nearset neighbor dan juga menguji efesiensi penyelesaian masalah

2.2 Data Mining

Data mining adalah suatu proses penambangan informasi penting dari suatu data. Informasi penting ini didapat dari suatu proses yang amat rumit seperti menggunakan artificial intelligence, teknik statistik, ilmu matematika, machine learning, dan lain sebagainya. Teknik-teknik rumit tersebut nantinya akan mengidentifikasi dan mengekstraksi informasi yang bermanfaat dari suatu database besar. Data adalah catatan atas kumpulan fakta. Data merupakan bentuk jamak dari datum, berasal dari bahasa Latin yang berarti "sesuatu yang diberikan"[5]. Dalam penggunaan sehari-hari data berarti suatu pernyataan yang diterima secara apa adanya. Pernyataan ini adalah hasil pengukuran atau pengamatan suatu variabel yang bentuknya dapat berupa angka, katakata, atau citra. Dalam keilmuan fakta dikumpulkan untuk menjadi data. Data kemudian diolah sehingga dapat diutarakan secara jelas dan tepat sehingga dapat dimengerti oleh orang lain yang tidak langsung mengalaminya sendiri, hal ini dinamakan deskripsi. Pemilahan banyak data sesuai dengan persamaan atau perbedaan yang dikandungnya dinamakan klasifikasi[6].

2.3 Clustering

Analisis klaster atau clustering dipilih dalam menyelesaikan ini karena clustering bertujuan menemukan kelompok (cluster) objek yang berguna, dimana digunakannya tergantung dari tujuan analisa data[7]. Analisis cluster dapat dibagi menjadi teknik pengelompokan hierarchica (harariki) dan teknik pengelompokan nin-heirarchical (non hararki). Teknik klaster termasuk teknik yang sudah cukup banyak dikenal dan banyak dipakai dalam data mining. Sampai sekarang para ilmuwan dalam bidang data mining masih melakukan perbaikan model cluster, karena metode yang dikembangkan sekarang masih bersifat heuristik[8].

2.4 K-Means

K-Means merupakan salah satu metode data clustering data non hierarkis yang bertujuan untuk membagi data yang menjadi satu atau lebih cluster atau kelompok, sehingga data dengan karakteristik yang sama berada ke dalam satu cluster yang sama dan data dengan karakteristik yang berbeda dikelompokkan menjadi beberapa kelompok[9].

K-Means adalah metode clustering berbasis jarak yang membagi data menjadi beberapa cluster. Algoritma ini hanya bekerja dengan atribut numerik. Algoritma K-Means melibatkan pengelompokan partisi, yang membagi data menjadi daerah yang berbeda. Algoritma K-Means sangat populer karena kesederhanaan dan kemampuannya untuk mengelompokkan sejumlah besar data yang menyimpang

dengan sangat cepat. Pada algoritma K-Means, semua data harus dimiliki oleh cluster tertentu dan tidak menutup kemungkinan semua data yang dimiliki oleh cluster tertentu pada satu langkah proses akan dipindahkan ke cluster lain pada langkah berikutnya[10].

Algoritma K-Means adalah metode non-hierarkis yang pertama kali menggunakan banyak komponen populasi untuk dijadikan sebagai pusat cluster awal. Pada langkah ini, pusat cluster dipilih secara acak dari dataset populasi. Selanjutnya, K-Means menguji setiap komponen dalam populasi data dan menugaskan komponen tersebut ke salah satu pusat klaster yang ditentukan oleh jarak minimum antara komponen dan setiap klaster[11]. Lokasi pusat klaster dihitung ulang hingga semua komponen data di setiap pusat klaster terklasifikasi dan akhirnya terbentuk lokasi pusat klaster baru[12]. K-Means Clustering merupakan metode yang termasuk ke dalam golongan algoritma partitioning clustering. Langkah-langkah dari metode K-Means adalah sebagai berikut[13][14][15]:

1. Tentukan nilai k sebagai jumlah cluster yang ingin dibentuk.
2. Inisialisasi k pusat cluster ini bisa dilakukan dengan berbagai cara. Namun yang paling sering dilakukan adalah dengan cara random. Pusat-pusat cluster diberiduberi nilai awal dengan angka-angka random.
3. Alokasikan semua data ke *cluster* terdekat. Kedekatan dua objek ditentukan berdasarkan jarak kedua objek tersebut. Demikian juga kedekatan suatu data ke *cluster* tertentu ditentukan jarak antara data dengan pusat *cluster*. Dalam tahapan ini perlu dihitng jarak antara tiap data ke tiap pusat *cluster*. Jarak paling antara satu data dengan satu *cluster* tertentu akan menentukan suatu data masuk dalam *cluster* mana. Untuk menghitung jarak semua data ke setiao titik pusat *cluster* dapat menggunakan teori jarak *Euclidean* yang dirumuskan sebagai berikut:

$$D(i, j) = \sqrt{(X_{1i} - X_{1j})^2 + (X_{2i} - X_{1j})^2 + \dots + (X_{ki} - X_{kj})^2} \dots \dots \dots (1)$$

Keterangan variabel formula :

$D(i, j)$ = Jarak data ke i ke pusat *cluster* j

X_{ki} = Data ke i pada atribut data ke

X_{kj} = Data ke i pada atribut data ke k

4. Hitung kembali pusat *cluster* dengan keanggotaan *cluster* yang sekarang. Pusat *cluster* adalah rata-rata dari semua data dalam *cluster* tertentu. Jika dikehendaki bisa juga menggunakan median dari *cluster* tersebut. Jika rata-rata (mean) bukan satu-satunya ukuran yang bisa dipakai.
5. Tugaskan lagi setiap objek memakai pusat *cluster* yang baru. Jika pusat *cluster* tidak berubah lagi maka proses *clustering* selesai. Atau, kembali ke langkah nomor 3 sampai pusat *cluster* tidak berubah lagi.

2.5 Jasa Pengiriman

Jasa pengiriman adalah layanan yang menyediakan pengangkutan barang dari satu tempat ke tempat lainnya. Mereka bertanggung jawab untuk mengambil barang dari pengirim, memprosesnya, dan mengirimkannya kepada penerima sesuai dengan instruksi yang diberikan. Jasa pengiriman dapat beroperasi secara lokal, nasional, atau internasional, dan mereka biasanya menawarkan berbagai jenis layanan, termasuk pengiriman ekspres, pengiriman reguler, dan pengiriman kargo[16].

3. HASIL DAN ANALISIS

Data yang digunakan adalah data tahun 2023 dengan kurun waktu 3 bulan yaitu September, Oktober, dan November dengan cabang sebanyak 11 Provinsi. Pengumpulan data dalam penelitian ini dilakukan dengan mengambil data pengiriman barang di setiap provinsi yang telah ada di CV. Lima Benua Nusa Indonesia dan melakukan wawancara dengan pihak karyawan.

3.1 Penerapan Algoritma K-Means Clustering

Algoritma *K-Means Clustering* adalah salah satu teknik paling umum dalam analisis data untuk mengelompokkan data ke dalam kelompok atau *cluster* berdasarkan kesamaan fitur.

Tabel 1. Data Distribusi Barang CV Lima Benua Nusa Indonesia

No	Provinsi	Jumlah Pengiriman		
		September	Oktober	November
1.	Sumatera Utara	3550	3600	6275
2.	Sumatera Barat	6480	5400	5525
3.	Sumatera Selatan	2640	6300	6330
4.	Aceh	5400	5600	7250
5.	Riau	5160	6300	7340
6.	Kepulauan Riau	2880	3550	6220
7.	Jambi	6730	5650	6225
8.	Bengkulu	6320	5230	7450

9.	Aceh Singkil	4240	5440	5550
10.	Bakal Belitung	5550	5030	4910
11.	Lampung	3110	6450	6510

1. Iterasi 1

Untuk menentukan pusat *centroid* dari sebuah himpunan data yang akan dibagi menjadi tiga kelompok (*cluster*), langkah-langkah umumnya adalah sebagai berikut:

- Inisialisasi pusat centroid awal**
Pilih tiga titik awal sebagai pusat centroid untuk masing-masing cluster. Titik-titik bisa dipilih secara acak dari data atau menggunakan metode *k-means*.
- Assignment cluster**
Tentukan setiap *cluster* yang memiliki pusat *centroid* terdekat. Jarak umum digunakan adalah jarak *eclidean*, meskipun ada metrik lain yang dapat digunakan tergantung pada kasusnya.
- Perbarui pusat centroid**
Hitung ulang pusat centroid untuk setiap cluster sebagai rata-rata dari semua data yang termasuk dalam cluster tersebut.
- Ulangi Langkah 2 dan 3**
Ulangi proses penugasan dan pembaharuan pusat centroid sampai tidak ada perubahan yang signifikan dalam penempatan data atau maksimal iterasi yang ditentukan tercapai.
Proses di atas merupakan inti dari algoritma K-means clustering. Dengan menggunakan metode ini, data akan dikelompokkan menjadi tiga kelompok berdasarkan jarak terdekat ke pusat centroid yang sesuai.

Tabel 2. Centroid Awal

Keterangan	September	Oktober	November
C1	5160	6300	7340
C2	2880	3550	6220
C3	6730	5650	6225

Tabel 2 adalah menentukan pusat awal cluster secara acak diambil dari data pada tabel 4.1 data yang dipilih untuk cluster awal adalah produk 5, 6, dan 7. Jarak centroid data ke-1 pada cluster 1 (C1) yaitu: Cluster 1

$$\begin{aligned}
 D_{1,5} &= \sqrt{(3550 - 5160)^2 + (3600 - 6300)^2 + (6275 - 7340)^2} \\
 &= \sqrt{2592100 + 729000 + 1134225} \\
 &= \sqrt{11016325} \\
 &= \mathbf{3319.085} \\
 D_{2,5} &= \sqrt{(6480 - 5160)^2 + (5400 - 6300)^2 + (5525 - 7340)^2} \\
 &= \sqrt{1742400 + 810000 + 3294225} \\
 &= \sqrt{5846625} \\
 &= 2417.98 \\
 D_{3,5} &= \sqrt{(2640 - 5160)^2 + (6300 - 6300)^2 + (6330 - 7340)^2} \\
 &= \sqrt{6350400 + 0 + 1020100} \\
 &= \sqrt{7370500} \\
 &= 2714.86648 \\
 D_{4,5} &= \sqrt{5.400 - 5160^2 + (5.600 - 6.300)^2 + (7.250 - 7340)^2} \\
 &= \sqrt{57600 + 490000 + 8100} \\
 &= \sqrt{555700} \\
 &= 745.45288
 \end{aligned}$$

Proses untuk mendapatkan nilai *cluster* (6 dan 7) didapatkan dengan cara yang sesuai pada cluser 1 Berdasarkan hasil perhitungan iterasi pertama maka kita pengelompokkan data ke 1, seperti tabel 3

Tabel 3. Pengelompokan Data ke 1

No	Provinsi	C1	C2	C3	Dekat	Cluster
1	Sumatera Utara	3319,08496	674,111	3783,83	674,111	C2
2	Sumatera Barat	2417,97953	4106,77	784,219	784,219	C3
3	Smatera Selatan	2714,86648	2762,64	4142,66	2714,86648	C1
4	Aceh	745,452882	3407,9	1679,89	745,452882	C1
5	Riau	0	3743,7	2032,39	0	C1
6	Kepulauan Riau	3743,70128	0	4385,49	0	C2

7	Jambi	2032,39391	4385,49	0	0	C3
8	Bengkulu	1581,96081	4021,06	1358,35	1358,35	C3
9	Aceh Singkil	2188,62971	2422,93	2588,4	2188,62971	C1
10	Bakal Belitung	2769,45843	3321,96	1872,44	1872,44	C3
11	Lampung	2216,73183	2923,53	3718,28	2216,73183	C3

2. Iterasi 2

Iterasi kedua merupakan proses penghitungan kembali dikarenakan masih ada data yang dipindah cluster. Melalui pengumpulan kembali terhadap *cluster* ada yang pada tabel 3

Tabel 4. Pengelompokkan cluster iterasi 2

No	Provinsi	Bulan			Cluster
		September	Oktober	November	
3	Sumatera Selatan	2640	6300	6330	C1
4	Aceh	5400	5600	7250	C1
5	Riau	5160	6300	7340	C1
9	Aceh Singkil	4240	5440	5550	C1
11	Lampung	3110	6450	6510	C1
1	Sumatera Utara	3550	3600	6275	C2
6	Kepulauan Riau	2880	3550	6220	C2
2	Sumatera Barat	6480	5400	5525	C3
7	Jambi	6730	5650	6225	C3
8	Bengkulu	6320	5230	7450	C3
10	Bakal Belitung	5550	5030	4910	C3

Hasil Pengelompokan cluster iterasi 2, maka melakukan perhitungan dengan pengelompokan cluster yang sama dan membagi dengan banyaknya jumlah cluster yang sejenis.

Tabel 5. Centroid Baru iterasi 2

Keterangan	September	Oktober	November
C1	4110	6018	6596
C2	3215	3575	6247.5
C3	6270	5327.5	6027.5

Berdasarkan tabel di atas adalah centroid baru yang dihasilkan dari rata-rata dari 3 cluster.

- Perhitungan jarak dengan pusat cluster perhitungan jarak dari data ke *i* terhadap pusat centroid adalah:

Cluster 1

$$\begin{aligned}
 D_{1,5} &= \sqrt{(3.550 - 4110)^2 + (3.600 - 6018)^2 + (6275 - 6596)^2} \\
 &= \sqrt{313600 + 5846724 + 103041} \\
 &= \sqrt{62633665} \\
 &= \mathbf{2502.671573}
 \end{aligned}$$

$$\begin{aligned}
 D_{2,5} &= \sqrt{(6.480 - 4110)^2 + (5.400 - 6018)^2 + (5.525 - 6596)^2} \\
 &= \sqrt{5616900 + 381924 + 1147041} \\
 &= \sqrt{7145865} \\
 &= 2673.175078
 \end{aligned}$$

$$\begin{aligned}
 D_{3,5} &= \sqrt{(2.640 - 4110)^2 + (6.300 - 6018)^2 + (6.330 - 6596)^2} \\
 &= \sqrt{2160900 + 79524 + 70756} \\
 &= \sqrt{2311180} \\
 &= 1520.256557
 \end{aligned}$$

$$\begin{aligned}
 D_{4,5} &= \sqrt{(5.400 - 4110)^2 + (5.600 - 6018)^2 + (7.250 - 6596)^2} \\
 &= \sqrt{1664100 + 174724 + 427716} \\
 &= \sqrt{2266540} \\
 &= 1505.503238
 \end{aligned}$$

$$\begin{aligned}
 D_{5,5} &= \sqrt{(5.160 - 4110)^2 + (6.300 - 6018)^2 + (7.340 - 6596)^2} \\
 &= \sqrt{1102500 + 79524 + 553536} \\
 &= \sqrt{1735560} \\
 &= 1317.406543
 \end{aligned}$$

Centroid baru dihitung menggunakan rumus seperti iterasi pertama dan kedua karena ada data yang berpindah cluster dan cluster ke-1 dan cluster ke-2 hasilnya berbeda, maka proses perhitungan centroid yang baru dilanjutkan dengan iterasi ke-3.

Tabel 6. Pengelompokan data ke 2

No	Provinsi	C1	C2	C3	Dekat	Cluster
1	Sumatera Utara	2502,67157	337,0552625	3231,704272	337,0552625	C2
2	Sumatera Barat	2673,17508	3809,574287	549,4201489	549,4201489	C3
3	Smatera Selatan	1520,25656	2786,226166	3770,167437	1530,25656	C1
4	Aceh	1505,50324	3143,223863	1505,012295	1505,012295	C3
5	Riau	1317,40654	3521,676625	1974,971519	1317,40654	C1
6	Kepulauan Riau	2783,03791	337,0552625	3832,579092	337,0552625	C2
7	Jambi	2671,60345	4081,832462	595,493492	595,493492	C3
8	Bengkulu	2496,87004	3718,340524	1426,712882	1426,712882	C3
9	Aceh Singkil	1202,12312	2239,499107	2088,435419	1202,12312	C1
10	Bakal Belitung	2427,41426	3059,110369	1362,245389	1362,245389	C3
11	Lampung	1092,71222	2888,867641	3387,989593	1092,71222	C1

Hasilnya berbeda dari *cluster* 1 dengan *cluster* 2 maka iterasi centroid yang baru dilanjutkan dengan iterasi ke -3.

3. Iterasi 3

Merupakan proses penghitungan kembali dikarenakan masih ada data yang berpindah cluster.

Tabel 7. Pengelompokan Cluster iterasi 3

No	Provinsi	Bulan			Cluster
		September	Oktober	November	
3	Sumatera Selatan	2640	6300	6330	C1
5	Riau	5160	6300	7340	C1
9	Aceh Singkil	4240	5440	5550	C1
11	Lampung	3110	6450	6510	C1
1	Sumatera Utara	3550	3600	6275	C2
6	Kepulauan Riau	2880	3550	6220	C2
2	Sumatera Utara	6480	5400	5525	C3
4	Aceh	5400	5600	7250	C3
7	Jambi	6730	5650	6225	C3
8	Bengkuku	6320	5230	7450	C3
10	Bakal Belitung	5550	5030	4910	C3

Lakukan proses perhitungan iterasi 3 sama seperti iterasi 2 dan 1.

Tabel 8. Centroid Terbaru Iterasi 3

Keterangan	September	Oktober	November
C1	3787.5	6112.5	6432.5
C2	3215	3575	6247.5
C3	6096	5382	6272

Tabel 5 adalah centroid baru yang dihasilkan dari rata-rata dari 3 cluster dengan iterasi 3.

Cluster 1

$$\begin{aligned}
 D1,5 &= \sqrt{(3.550 - 3787.5)^2 + (3.600 - 6112.5)^2 + (6275 - 6432.5)^2} \\
 &= \sqrt{56406.25 + 6312656.25 + 24806.25} \\
 &= \sqrt{6393868.75} \\
 &= 2528.610043 \\
 D2,5 &= \sqrt{(6.480 - 3787.5)^2 + (5.400 - 6112.5)^2 + (5.525 - 6432.5)^2} \\
 &= \sqrt{7249556.25 + 507656.25 + 823556.25} \\
 &= \sqrt{8580768.75} \\
 &= 2929.294924 \\
 D3,5 &= \sqrt{(2.640 - 3787.5)^2 + (6.300 - 6112.5)^2 + (6.330 - 6432.5)^2} \\
 &= \sqrt{1316756.25 + 35156.25 + 10506.25} \\
 &= \sqrt{1362418.75} \\
 &= 1167.226949
 \end{aligned}$$

$$\begin{aligned}
 D_{4,5} &= \sqrt{(5.400 - 3787.5)^2 + (5.600 - 6112.5)^2 + (7.250 - 6432.5)^2} \\
 &= \sqrt{2600156.25 + 262656.25 + 668306.25} \\
 &= \sqrt{3553118.75} \\
 &= 1879.127124 \\
 D_{5,5} &= \sqrt{(5.160 - 3787.5)^2 + (6.300 - 6112.5)^2 + (7.340 - 6432.5)^2} \\
 &= \sqrt{1883756.25 + 35156.25 + 823556.25} \\
 &= \sqrt{2742468.75} \\
 &= 1656.040081
 \end{aligned}$$

Centroid baru dihitung menggunakan rumus seperti iterasi pertama dan kedua, karena tidak ada data yang berpindah cluster dan cluster ke satu, dua, dan tiga hasilnya sama, maka proses perhitungan centroid yang baru dihentikan dan berakhir pada iterasi ke tiga. Berdasarkan hasil iterasi 2 dapat dilihat hasil yang diperoleh sama dengan iterasi 2 maka perhitungan dihentikan sampai iterasi 3. Provinsi yang ada pada cluster C1 (tertinggi) sumatera selatan, Riau, Aceh Singkil, dan Lampung. Provinsi yang ada pada cluster C2 (sedang) Sumatera Utara, Kepulauan Riua, sedangkan untuk culster C3 (rendah) Sumatera barat, Aceh, Jambi, dan Bakal Belitung.

3.2 Pengujian Menggunakan Rapidminer

Setelah melakukan proses sistem dan pengolahan data, maka akan muncul hasil dari *rapidminer* sebagai berikut:

Cluster Model

```

Cluster 0: 2 items
Cluster 1: 5 items
Cluster 2: 4 items
Total number of items: 11
  
```

Gambar 1. Cluster Model

Berdasarkan dari penjelasan diatas mengenai tahap-tahap penggunaan serta hasil yang telah ditampilkan maka selanjutnya membahas mengenai keterkaitan dan hasil yang didapat antara perhitungan manual algortima dengan hasil yang ditampilkan oleh *tools rapidminer*. Data di atas akan diuji dengan sampel sebanyak 11 data dan diselesaikan dengan menggunakan *tools rapidminer* sehingga diperoleh hasil *cluster* seperti pada gambar 4 berikut:

Example 1

This dialog shows detailed information about the example with ID 1.

Attribute	Value
September	3550
October	3600
November	6275
Provinsi	Sumatera Utara
id	1
cluster	cluster_1

Example 2

This dialog shows detailed information about the example with ID 2.

Attribute	Value
September	6480
Oktober	5400
November	5525
Provinsi	Sumatera Barat
id	2
cluster	cluster_1

Example 7

This dialog shows detailed information about the example with ID 7.

Attribute	Value
September	6730
Oktober	5650
November	6225
Provinsi	Jambi
id	7

Example 3

This dialog shows detailed information about the example with ID 3.

Attribute	Value
September	2640
Oktober	6300
November	6330
Provinsi	Sumatera Selatan
id	3

Example 6

This dialog shows detailed information about the example with ID 6.

Attribute	Value
September	2880
Oktober	3550
November	6220
Provinsi	Kepulauan Riau
id	6

Example 11

This dialog shows detailed information about the example with ID 11.

Attribute	Value
September	3110
Oktober	6450
November	6510
Provinsi	Lampung
id	11
cluster	cluster_1

Example 10

This dialog shows detailed information about the example with ID 10.

Attribute	Value
September	5550
Oktober	5030
November	4910
Provinsi	Bakal Belitung
id	10
cluster	cluster_2

Example 4

This dialog shows detailed information about the example with ID 4.

Attribute	Value
September	5400
Oktober	5600
November	7250
Provinsi	Aceh
id	4
cluster	cluster_0

Example 9

This dialog shows detailed information about the example with ID 9.

Attribute	Value
September	4240
Oktober	5440
November	5550
Provinsi	Aceh Singkil
id	9
cluster	cluster_1

Example 8

This dialog shows detailed information about the example with ID 8.

Attribute	Value
September	6320
Oktober	5230
November	7450
Provinsi	Bengkulu
id	8
cluster	cluster_2

Berdasarkan gambar diatas dapat dilihat bahwa hasil dari perhitungan manual adalah sama dengan hasil perhitungan menggunakan *tools rapidminer*. Dalam melakukan validasi data perhitungan algoritma harus menghasilkan hasil akhir berupa pengelompokan dengan 3 *cluster*, serta data yang digunakan merupakan data yang *valid* dan sama dengan yang dipakai.

4. KESIMPULAN

Penerapan data mining untuk pengelompokan minat konsumen terhadap pengguna jasa pengiriman pada CV. Lima Benua Nusa Indonesia menjadi pertimbangan keputusan (decision maker) dalam menentukan minat konsumen pada CV. Lima Benua Nusa Indonesia. clustering adalah suatu proses pengelompokan objek data yang mirip satu sama lain kedalam kluster yang sama dan berbeda dengan objek yang ada di kluster lain. Metode penelitian ini menggunakan 3 cluster lokasi dari 3 iterasi.

Setelah didapatkan hasil dari setiap cluster kemudian pusat cluster baru dihitung berdasarkan data member tiap-tiap cluster yang sudah didapatkan. Iterasi selanjutnya dilakukan dengan cara yang sama hingga tidak ada perubahan data dalam suatu cluster.

Berdasarkan hasil dari setiap iterasi dapat dilihat bahwa hasil yang diperoleh sama dengan cluster tinggi sebanyak 11 alternatif..

REFERENSI

- [1] R. W. Nasution, I. O. Kirana, I. Gunawan, and I. P. Sari, "Penerapan Data Mining Untuk Pengelompokan Minat Konsumen Terhadap Pengguna Jasa Pengiriman Pada PT . Jalur Nugraha Ekakurir (JNE) Pematangsiantar," *Resolusi*, vol. 1, no. 4, pp. 274–281, 2021.
- [2] A. Karim, S. Esabella, M. Hidayatullah, and S. Suryadi, "Penerapan Data Mining Untuk Pengelompokan Terhadap Kualitas Kinerja Karyawan Dengan Menggunakan Algoritma K-Medoids Clustering," vol. 8, no. April, pp. 1001–1009, 2024, doi: 10.30865/mib.v8i2.7445.
- [3] N. Hendrastuty, "Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Dalam Evaluasi Hasil Pembelajaran Siswa," 2024, doi: 10.58602/jima-ilkom.v3i1.26. "Repository+PIM121".
- [4] U. R. Amanda and D. P. Utomo, "Penerapan Data Mining Algoritma Hash Based Pada Data Pemesanan Buah Impor Cv. Green Uni Fruit," *KOMIK (Konferensi Nasional Teknologi Informasi dan Komputer)*, vol. 5, no. 1, pp. 86–93, 2021, doi: 10.30865/komik.v5i1.3653.
- [5] S. Larno, M. Razi, and P. Anggraini, "Implementasi Website Promosi Dan Penjualan Pada Asosiasi Pedagang Sepatu Dan Tas Kota Padang," *Jurnal Teknologi Dan Sistem Informasi Bisnis*, vol. 1, no. 1, pp. 38–46, 2019, doi: 10.47233/jteksis.v1i1.5.
- [6] D. Ramdhan, G. Dwilestari, R. Danar Dana, and A. Ajiz, "Clustering Data Persediaan Barang dengan Menggunakan Metode K-Means. Clustering Data Persediaan Barang dengan Menggunakan Metode K-Means," vol. 7, no. 1, [Online]. Available: http://ejournal.ust.ac.id/index.php/Jurnal_Means/
- [7] I. Ferdiansyah, B. Huda, A. Hananto, and U. Buana Perjuangan Karawang, "Analisis Clustering Menggunakan Metode K-Means Pada Kemiskinan Di Jawa Timur Tahun 2020".
- [8] S. Aulia, "Klasterisasi Pola Penjualan Pestisida Menggunakan Metode K-Means Clustering (Studi Kasus Di Toko Juanda Tani Kecamatan Hutabayu Raja)," *Djtechno: Jurnal Teknologi Informasi*, vol. 1, no. 1, pp. 1–5, 2021, doi: 10.46576/djtechno.v1i1.964.
- [9] P. P. Tjaya, R. Rino, and H. Wijaya, "Implementasi Metode Clustering K-Means Untuk Rekomendasi Pengadaan Stock Lampu Pada Pt Global Lighting Indonesia," *Algor*, vol. 3, no. 1, pp. 50–59, 2021, doi: 10.31253/algor.v3i1.628.
- [10] F. Irmahna Rahman, S. Mujadilah, T. Wahyuni, and L. Anas, "Prediksi Tingkat Kelulusan Menggunakan K-Means Pada Program Studi Informatika Unismuh Makassar".
- [11] F. Handayani, "Aplikasi Data Mining Menggunakan Algoritma K-Means Clustering untuk Mengelompokkan Mahasiswa Berdasarkan Gaya Belajar," *Jurnal Teknologi dan Informasi*, doi: 10.34010/jati.v12i1.
- [12] E. Febrianty, L. Awalina, and W. I. Rahayu, "Optimalisasi Strategi Pemasaran dengan Segmentasi Pelanggan Menggunakan Penerapan K-Means Clustering pada Transaksi Online Retail Optimizing Marketing Strategies with Customer Segmentation Using K-Means Clustering on Online Retail Transactions," *Jurnal Teknologi dan Informasi (JATI)*, vol. 13, 2023, doi: 10.34010/jati.v13i2.
- [13] D. Kurniadi et al., "1) , yoga.handoko@itg.ac.id 2) , 1806129@itg.ac.id 3) , idafarida@itg.ac," *AITI: Jurnal Teknologi Informasi*, vol. 20, no. 1, pp. 64–77, 2023.
- [14] P. Apriyani, A. R. Dikananda, and I. Ali, "Penerapan Algoritma K-Means dalam Klasterisasi Kasus Stunting Balita Desa Tegalwangi," *Hello World Jurnal Ilmu Komputer*, vol. 2, no. 1, pp. 20–33, Mar. 2023, doi: 10.56211/helloworld.v2i1.230.
- [15] R. Mahfiroh, "Perlindungan Hukum terhadap Barang Kiriman Konsumen Pengguna Jasa Go-Send Instant Courier Melalui Tokopedia," *Jurnal Lex Renaissance*, vol. 5, no. 1, pp. 235–249, 2020, doi: 10.20885/jlr.vol5.iss1.art15.