



Implementasi Algoritma K-Means Clustering Pada Analisa Impor Beras

Nur Iman Sarumaha

Fakultas Ilmu Komputer Dan Teknologi Informasi, Tenik Informatika, Universitas Budi Darma, Medan, Indonesia
Email: nurimansarumaha@gmail.com

Abstrak—Pengadaan impor beras merupakan kebijakan pemerintah untuk memenuhi kebutuhan pangan masyarakat dan menjaga stok pangan nasional. Pemasukan jumlah komoditi beras impor yang meningkat menyebabkan terjadinya penyewaan gudang penyimpanan beras. Untuk menyesuaikan jumlah pemasukan komoditi beras impor dan untuk mengetahui kualitas beras yang bermutu maka dilakukan pengelompokan jumlah kuantum beras impor berdasarkan negara yang mengekspor terdiri dari 3 (tiga) kelompok yaitu, kelompok jumlah impor tertinggi, sedang dan kelompok jumlah impor terendah. Thailand, India, Vietnam dan Pakistan merupakan negara yang mengekspor beras pada Perum Bulog Subdivre Medan. Sebelum data beras impor diolah menggunakan Algoritma K-Means Clustering, data beras diinisialisasikan ke dalam bentuk angka. Data yang diinisialisasikan yaitu bulan/tahun dan negara pengekspor. Dari hasil pengujian menggunakan Tanagra versi 1.4.41 menunjukkan bahwa pada bulan Juli-Oktober 2018 terjadi pemasukan jumlah kuantum terbanyak dan negara Vietnam merupakan negara pengekspor dengan jumlah kuantum impor tertinggi sebesar 60.000 ton.

Kata Kunci: Data Mining, Algoritma K-Means, Impor.

Abstract—Procurement of rice imports is a government policy to meet people's food needs and maintain national food stocks. The increase in the number of imported rice commodities has led to the rental of rice storage warehouses. In order to adjust the amount of imports of rice commodities and to determine the quality of quality rice, the quantum quantity of imported rice is grouped based on the exporting country, consisting of 3 (three) groups, namely, the highest import number group, the medium import group and the lowest import group group. Thailand, India, Vietnam and Pakistan are countries that export rice to Perum Bulog Subdivre Medan. Before the imported rice data is processed using the K-Means Clustering Algorithm, the rice data is initialized in the form of numbers. Initialized data are month/year and exporting country. From the test results using Tanagra version 1.4.41, it shows that in July-October 2018 the largest number of quantum imports occurred and Vietnam was the exporting country with the highest import quantum amount of 60,000 tons.

Keywords: Data Mining, K-Means Algorithm, Import

1. PENDAHULUAN

Ketahanan pangan menjadi persoalan penting dengan perubahan iklim global, pertumbuhan jumlah penduduk serta kurangnya lahan untuk pertanian. Setiap warga negara berhak atas tercukupinya pangan dengan harga yang terjangkau, oleh karena itu menjadi tugas pemerintah untuk menetapkan kebijakan yang dapat menjamin kecukupan dan keterjangkauan pangan bagi seluruh masyarakat. Sehingga pemerintah melakukan kebijakan impor beras untuk memenuhi kebutuhan pangan masyarakat. Kebijakan impor beras tersebut merupakan cara pemerintah untuk menjaga stok ketahanan pangan nasional.

Pemasukan jumlah komoditi beras impor yang tidak sesuai dengan jumlah kuantum yang telah ditentukan dapat menyebabkan terjadinya penambahan atau pengurangan jumlah beras. Dengan adanya penambahan jumlah komoditi yang meningkat menyebabkan terjadinya penyewaan gudang penyimpanan beras. Sarana dan prasarana pelabuhan, kecepatan bongkar, jenis kapal, gelombang laut dan hal lainnya yang dapat mempengaruhi percepatan pemasukan komoditi beras ke gudang. Untuk menyesuaikan jumlah pemasukan komoditi beras impor dan untuk mengetahui kualitas beras yang bermutu maka dilakukan pengelompokan jumlah kuantum impor beras berdasarkan negara yang mengekspor terdiri dari 3 (tiga) kelompok yaitu, kelompok jumlah impor tertinggi, sedang dan kelompok jumlah impor terendah. Teknik penyelesaian yang penulis gunakan dalam penelitian ini adalah dengan analisis kelompok (*cluster analysis*) yang lebih dikenal dengan *clustering*. *Clustering* adalah metode yang digunakan untuk membagi rangkaian data menjadi beberapa group berdasarkan kesamaan-kesamaan yang telah ditentukan sebelumnya [1].

Pada penelitian sebelumnya, Johan Oscar Ong “Implementasi Algoritma *K-Means Clustering* Untuk Menentukan Strategi Marketing President University” adalah berdasarkan hasil pengelompokan data menggunakan Algoritma *K-Means Clustering*, promosi dilakukan dengan mengirim tim marketing yang sesuai dengan jurusan yang paling banyak diminati dan promosi dilakukan pada kota-kota di Indonesia yang didasarkan pada tingkat kemampuan akademik dari calon mahasiswa [2]. Koko Handoko “Penerapan Data Mining Dalam Meningkatkan Mutu Pembelajaran Pada Instansi Perguruan Tinggi Menggunakan Metode *K-Means Clustering* (Studi Kasus Di Program Studi TKJ Akademi Komunitas Solok Selatan)” adalah penggunaan Algoritma *K-Means Clustering* sangat membantu dalam pengelompokan data mutu pembelajaran serta penentuan *centroid* (titik pusat) pada tahap awal algoritma *K-Means* yang berpengaruh pada hasil *cluster* data mahasiswa TKJ Akademi Komunitas Solok Selatan [3].

2. METODOLOGI PENELITIAN

2.1 Data Mining





Data Mining adalah serangkaian proses untuk menggali nilai tambah berupa informasi yang selama ini tidak diketahui secara manual dari suatu basis data. Data mining sering disebut *Knowledge Discovery Databases (KDD)* karena digunakan untuk mencari pengetahuan yang terdapat dalam basis data yang besar [4]. Penggunaan teknik statistik, matematika, kecerdasan buatan dan *machine learning* merupakan proses dari data mining untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari *database* yang besar [5].

2.2 Algoritma K-Means

Algoritma *K-Means* merupakan algoritma klusterisasi yang paling banyak digunakan dalam berbagai aplikasi kecil hingga menengah karena kemudahannya implementasinya [6]. *K-Means* merupakan algoritma *clustering* yang berulang-ulang. Algoritma *K-Means* menetapkan nilai-nilai *cluster* (*K*) secara random, untuk sementara nilai tersebut menjadi pusat dari *cluster* atau biasa disebut dengan *centroid*, mean atau “means”. Kemudian menghitung jarak setiap data yang ada terhadap masing-masing *centroid* menggunakan rumus *Euclidian* hingga ditemukan jarak yang paling dekat dari setiap data dengan *centroid*. Klasifikasikan setiap data berdasarkan kedekatannya dengan *centroid* [4]. Langkah-langkah pengelompokan data dengan algoritma *K-Means* adalah sebagai berikut:

a. Menentukan jumlah kelompok/*cluster*

Dalam menentukan jumlah cluster yang akan dibuat, dapat dilakukan beberapa pertimbangan teoritis dan konseptual yang mungkin diusulkan untuk menentukan berapa banyak *cluster* yang harus dibentuk.

b. Menentukan nilai titik pusat *cluster* secara acak pada data awal.

Untuk menentukan titik pusat awal *cluster* dilakukan dengan cara memilih sejumlah *k* objek secara acak dari himpunan data *D* sebagai *centroid* awal.

Tabel 1. Titik Pusat Awal Setiap *Cluster*

Data ke-	Bulan ke-	Negara	Jumlah Kuantum (Ton Netto)
18a	6	3	35.000
23d	11	4	2.150
21d	9	4	6.962,5

b. Menghitung jarak pusat *cluster*

Untuk menghitung jarak setiap data ke pusat *cluster* menggunakan rumus *Euclidean Distance* sebagai berikut :

$$D(i,j) = \sqrt{(X1i - X1j)^2 + (X2i - X2j)^2 + \dots + (Xki - Xkj)^2} \tag{1}$$

dimana :

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

X_{ki} : Data ke *i* pada atribut data ke *k*

X_{kj} : Titik pusat ke *j* pada atribut ke *k*

c. Mengelompokan data dengan memasukan setiap objek ke dalam *cluster* (grup) berdasarkan jarak minimumnya.

Untuk mengukur jarak minimum setiap data ke pusat *cluster* menggunakan rumus jarak terpendek.

$$MIN (X_{1j} : X_{2j} : X_{kj}) \tag{2}$$

Pengalokasian kembali data ke dalam masing-masing kelompok dalam metode *K-Means* didasarkan pada perbandingan jarak antara data dengan sentroid setiap kelompok yang ada menggunakan persamaan 3.

$$a_{i1} = \begin{cases} 1 & \text{d} = \min\{D(x_i, C_1)\} \\ 0 & \text{lainnya} \end{cases} \tag{3}$$

a_{i1} adalah nilai keanggotaan titik x_i ke pusat kelompok C_1 , d adalah jarak terpendek dari data x ke K kelompok setelah dibandingkan, dan C_1 adalah sentroid (pusat kelompok) ke-1.

Fungsi objektif yang digunakan untuk *K-Means* ditentukan berdasarkan jarak dan nilai keanggotaan data dalam kelompok. Fungsi objektif yang digunakan adalah :

$$J = \sum_{i=1}^n \sum_{c=1}^k a_{ic} D(x_i, C_c)^2 \tag{4}$$

n adalah jumlah data, k adalah jumlah kelompok, a_{i1} adalah nilai keanggotaan titik data x ke pusat kelompok C_1 , C_1 adalah pusat kelompok ke-1, dan $D(x_1, C_1)$ adalah jarak titik x ke kelompok C_1 yang diikuti. a mempunyai nilai 0 atau 1. Apabila suatu data merupakan anggota suatu kelompok, nilai $a_{i1} = 1$, jika tidak maka nilai $a_{i1} = 0$.

d. Jika ada data yang harus dipindah, maka langkah selanjutnya adalah menghitung pusat *cluster* baru. Pusat *cluster* yang baru ditentukan berdasarkan pengelompokan anggota masing-masing *cluster* baru. Pengulangan dihentikan sampai hasil perhitungan menunjukkan adanya angka pusat *cluster* yang sama [7] [8].

2.2 Beras





Impor merupakan kegiatan memasukkan barang dari suatu negara (luar negeri) ke dalam wilayah pabean negara lain. Melibatkan dua negara yang bisa diwakili oleh kepentingan dua perusahaan antar dua negara tersebut yang berbeda peraturan serta perundang-undangan. Negara yang satu bertindak sebagai eksportir (supplier) dan yang lainnya bertindak sebagai negara penerima/importir [9]. Beras adalah bagian butir (gabah) yang telah dipisah dari sekam. Sekam secara anatomi disebut palea (bagian yang ditutupi) dan lemma (bagian yang menutupi). Pada salah satu tahap pemrosesan hasil panen padi, gabah digiling sehingga bagian luarnya (kulit gabah) terlepas dari isinya. Bagian isi inilah yang berwarna putih, kemerahan, ungu, atau bahkan hitam, yang disebut beras. Sebelum menjadi beras, gabah mengalami berbagai perlakuan penanganan pasca panen yang meliputi pemanenan, pra-pengeringan, pengeringan, penyimpanan, penggilingan dan penyosohan. Selanjutnya beras diolah menjadi aneka produk pangan dari yang sederhana seperti nasi dan tepung sampai ke produk-produk pangan industri olahan seperti bubur instan, anea kue, produk ekstruksi dan sebagainya[10].

3. HASIL DAN PEMBAHASAN

Pemasukan jumlah komoditi beras impor yang tidak sesuai dengan jumlah kuantum yang telah ditentukan dapat menyebabkan terjadinya penambahan atau pengurangan jumlah beras. Dengan adanya penambahan jumlah komoditi yang meningkat menyebabkan terjadinya penyewaan gudang penyimpanan beras. Thailand, India, Vietnam dan Pakistan merupakan negara yang mengekspor beras pada Perum Bulog Subdivre Medan. Pada penelitian ini penulis menggunakan Algoritma *K-Means Clustering* untuk mengelompokan jumlah kuantum impor beras berdasarkan negara yang mengekspor. Untuk menyesuaikan jumlah pemasukan komoditi beras impor maka dikelompokkan jumlah kuantum impor beras berdasarkan negara pengekspor terdiri dari 3 (tiga) kelompok yaitu, kelompok jumlah impor tertinggi, sedang dan kelompok jumlah impor terendah. Langkah-langkah pengelompokan data impor beras dengan Algoritma *K-Means* adalah sebagai berikut :

- a. Menentukan jumlah kelompok.
- b. Menentukan nilai titik pusat *cluster* secara acak pada data awal.
- c. Menghitung data impor beras ke *centroid* dengan menggunakan rumus *Euclidean Distance*.
- d. Mengelompokan data dengan memasukkan setiap objek ke dalam *cluster* (grup) berdasarkan jarak minimumnya,
- e. Jika ada data yang harus dipindah, maka langkah selanjutnya adalah menghitung pusat *cluster* baru. Pusat *cluster* yang baru ditentukan berdasarkan pengelompokan anggota masing-masing *cluster* baru. Pengulangan dihentikan sampai hasil perhitungan menunjukkan adanya angka pusat *cluster* yang sama.

Data yang digunakan dalam penelitian ini adalah data impor beras tahun 2016 dan 2018, atribut data yang digunakan yaitu bulan dan tahun, negara pengekspor dan jumlah kuantum beras yang diimpor. Berikut adalah data impor beras pada tahun 2016 dan 2018.

Tabel 2. Data Impor Beras Tahun 2016 dan 2018

No.	Bulan & Tahun	Negara	Jumlah Kuantum (Ton Netto)
1.	Januari 2016	Thailand, Vietnam	12.300
2.	Februari 2016	Vietnam	17.200
3.	Maret 2016	Vietnam	17.100
4.	April 2016	-	-
5.	Mei 2016	-	-
6.	Juni 2016	-	-
7.	Juli 2016	-	-
8.	Agustus 2016	-	-
9.	September 2016	-	-
10.	Oktober 2016	-	-
11.	Nopember 2016	-	-
12.	Desember 2016	-	-
13.	Januari 2018	-	-
14.	Februari 2018	Thailand	10.000
15.	Maret 2018	India, Thailand	19.984
16.	April 2018	India	19.578
17.	Mei 2018	India, Vietnam	54.188
18.	Juni 2018	India, Vietnam	50.688
19.	Juli 2018	Thailand, India, Pakistan dan Vietnam	109.660,5
20.	Agustus 2018	Thailand, India, Vietnam dan Pakistan	111.702,5
21.	September 2018	Thailand, India, Vietnam dan Pakistan	144.205,1
22.	Oktober 2018	Thailand, India, Vietnam dan Pakistan	149.275,1
23.	Nopember 2018	Thailand, India, Vietnam dan Pakistan	171.065,1
24.	Desember 2018	-	-

Sumber: Perum Bulog Subdivre Medan

Berikut adalah hasil dari inialisasi bulan & tahun dan nama-nama negara yang mengekspor :



Tabel 3. Hasil Inisialisasi Data Impor Beras Tahun 2016 dan 2018

No.	Bulan ke-	Negara	Jumlah Kuantum (Ton Netto)
1.	1	1	6.100
	1	3	6.200
2.	2	3	17.200
3.	3	3	17.100
4.	4	-	0
5.	5	-	0
6.	6	-	0
7.	7	-	0
8.	8	-	0
9.	9	-	0
10.	10	-	0
11.	11	-	0
12.	12	-	0
13.	1	-	0
14.	2	1	10.000
15.	3	2	9.984
	3	1	10.000
16.	4	2	19.578
17.	5	2	7.898
	5	3	12.000
18.	6	2	15.688
	6	3	35.000
	7	1	31.200
19.	7	2	14.648
	7	3	61.000
	7	4	2.812,5
	8	1	31.200
20.	8	2	23.540
	8	3	50.000
	8	4	6.962,5
21.	9	1	41.500
	9	2	45.742,6
	9	3	50.000
	9	4	6.962,5
22.	10	1	41.500
	10	2	50.162,6
	10	3	48.000
	10	4	6.462,5
23.	11	1	16.200
	11	2	12.150
	11	3	11.500
	11	4	2.150
24.	12	-	0

Setelah data impor beras yang telah diinisialisasikan ke dalam bentuk angka kemudian dikelompokkan menggunakan algoritma *K-Means Clustering*. Langkah-langkah pengelompokan data impor beras menggunakan algoritma *K-Means Clustering* :

a. Menentukan jumlah *cluster*.

Dalam menentukan jumlah *cluster* yang akan dibuat, dapat dilakukan beberapa pertimbangan teoritis dan konseptual yang mungkin diusulkan untuk menentukan berapa banyak *cluster* yang harus dibentuk.

b. Menentukan titik pusat awal *cluster*.

Untuk menentukan titik pusat awal *cluster* dilakukan dengan cara memilih sejumlah *k* objek secara acak dari himpunan data *D* sebagai *centroid* awal.

Tabel 4. Titik Pusat Awal Setiap *Cluster*

Data ke-	Bulan ke-	Negara	Jumlah Kuantum (Ton Netto)
18a	6	3	35.000
23d	11	4	2.150
21d	9	4	6.962,5



c. Menghitung jarak pusat cluster

Iterasi 1

$$D(1a,1)= \sqrt{(1 - 6)^2 + (1 - 3)^2 + (6.100 - 35.000)^2}$$

$$= \sqrt{25 + 4 + 835.210.000}$$

$$= 28.900$$

Sampai pada data D(24,3)

$$D(24,3)= \sqrt{(12 - 9)^2 + (0 - 4)^2 + (0 - 6.962,5)^2}$$

$$= \sqrt{9 + 16 + 48.476.406,25}$$

$$= 6.962,5$$

d. Menentukan jarak terpendek data terhadap pusat cluster awal

Untuk mengukur jarak terpendek setiap data ke pusat cluster menggunakan rumus jarak terpendek seperti berikut :

$$\text{Jarak D1a} = \text{MIN} (X_{1j} : X_{2j} : X_{3j})$$

$$= \text{MIN} (28.900 : 3.950,014 : 862,5423)$$

$$= 862,5423$$

Sampai data D24

$$\text{Jarak D24} = \text{MIN} (X_{1j} : X_{2j} : X_{3j})$$

$$= \text{MIN} (35.000 : 2.150,004 : 6.962,502)$$

$$= 2.150,004$$

Berikut hasil perhitungan iterasi 1 menggunakan rumus Euclidean Distance.

Tabel 5. Hasil Perhitungan Iterasi 1

Data ke	C1	C2	C3	Cluster
1a	28.900	3.950,014	862,5423	C3
1b	28.800	4.050,012	762,5426	C3
2	17.800	15.050	10.237,5	C3
3	17.900	14.950	10.137,5	C3
4	35.000	2.150,015	6.962,503	C2
5	35.000	2.150,012	6.962,502	C2
6	35.000	2.150,01	6.962,502	C2
7	35.000	2.150,007	6.962,501	C2
8	35.000	2.150,006	6.962,501	C2
9	35.000	2.150,005	6.962,501	C2
10	35.000	2.150,004	6.962,501	C2
11	35.000	2.150,004	6.962,501	C2
12	35.000	2.150,004	6.962,502	C2
13	35.000	2.150,027	6.962,506	C2
14	25.000	7.850,006	3.037,51	C3
15b	25.000	7.850,005	3.037,507	C3
16	15.422	17.428	12.615,5	C3
17a	27.102	5.748,003	935,5107	C3
17b	23.000	9.850,002	5.037,502	C3
18a	19.312	13.538	8.725,501	C3
18b	0	32.850	28.037,5	C1
19a	3.800,001	29.050	24.237,5	C1
19b	20.352	12.498	7.685,501	C3
19c	26.000	58.850	54.037,5	C1
19d	32.187,5	662,5121	4.150	C2
20a	3.800,001	29.050	24.237,5	C1
20b	11.460	21.390	16.577,5	C1
20c	15.000	47.850	43.037,5	C1
20d	28.037,5	4.812,501	1	C3
21a	6.500,001	39.350	34.537,5	C1
21b	10.742,6	43.592,6	38.780,1	C1
21c	15.000	47.850	43.037,5	C1
21d	28.037,5	4.812,5	0	C3
22a	6.500,002	39.350	34.537,5	C1
22b	15.812,6	48.662,6	43.850,1	C1
22c	15.000	47.850	43.037,5	C1
22d	28.037,5	4.812,5	1	C3
23a	18.800	14.050	9.237,501	C3
23b	22.850	10.000	5.187,501	C3



23c	23.500	9.350	4.537,501	C3
23d	32.850	0	4.812,5	C2
24	35.000	2.150,004	6.962,502	C2

Tahap selanjutnya adalah menentukan jarak terpendek data terhadap pusat *cluster* awal yang dipilih secara acak sebelumnya. Untuk menentukan jarak terpendek dengan pusat *cluster* data menggunakan rumus $MIN (X_{1j} : X_{2j} : X_{3j})$. Berikut hasil penentuan jarak terpendek data terhadap pusat *cluster* awal yang telah ditabulasikan.

Tabel 6. Hasil Pengelompokan Iterasi I

Data ke	C1	C2	C3	Cluster
1a	0	0	1	C3
1b	0	0	1	C3
2	0	0	1	C3
4	0	1	0	C2
5	0	1	0	C2
6	0	1	0	C2
7	0	1	0	C2
8	0	1	0	C2
9	0	1	0	C2
10	0	1	0	C2
11	0	1	0	C2
12	0	1	0	C2
13	0	1	0	C2
14	0	0	1	C3
15a	0	0	1	C3
15b	0	0	1	C3
16	0	0	1	C3
17a	0	0	1	C3
17b	0	0	1	C3
18a	0	0	1	C3
18b	1	0	0	C1
19a	1	0	0	C1
19b	0	0	1	C3
19c	1	0	0	C1
19d	0	1	0	C2
20a	1	0	0	C1
20b	1	0	0	C1
20c	1	0	0	C1
20d	0	0	1	C3
21a	1	0	0	C1
21b	1	0	0	C1
21c	1	0	0	C1
21d	0	0	1	C3
22a	1	0	0	C1
22b	1	0	0	C1
22c	1	0	0	C1
22d	0	0	1	C3
23a	0	0	1	C3
23b	0	0	1	C3
23c	0	0	1	C3
23d	0	1	0	C2
24	0	1	0	C2

Berdasarkan nilai minimum yang telah ditentukan menggunakan rumus $MIN (X_{1j} : X_{2j} : X_{3j})$, maka diperoleh hasil pengelompokan seperti pada tabel berikut ini :

Tabel 7. Hasil Pengelompokan *Centroid* Baru

Kelompok	Anggota Kelompok	Jumlah Anggota
1	18b, 19a, 19c, 20a, 20b, 20c, 21a, 21b, 21c, 22a, 22b, 22c	12
2	4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 19d, 23d, 24	13
3	1a, 1b, 2, 3, 14, 15a, 15b, 16, 17a, 17b, 18a, 19b, 20d, 21d, 22d, 23a, 23b, 23c	18



Selanjutnya menentukan titik pusat cluster baru dengan mencari nilai rata-rata semua anggota cluster dengan cara menjumlahkan seluruh value anggota kelompok dibagi jumlah anggota. Berikut adalah titik pusat cluster baru yaitu :

$$C_1(X1) = \frac{\text{Value Bulan ke} - (18b, 19a, 19c, 20a, 20b, 20c, 21a, 21b, 21c, 22a, 22b, 22c)}{12}$$

$$= \frac{(6+7+7+8+8+8+9+9+9+10+10+10)}{12} = \frac{101}{12} = 8,4167$$

$$C_1(X2) = \frac{\text{Value Negara} (18b, 19a, 19c, 20a, 20b, 20c, 21a, 21b, 21c, 22a, 22b, 22c)}{12}$$

$$= \frac{(3+1+3+1+2+3+1+2+3+1+2+3)}{12} = \frac{25}{12} = 2,0833$$

$$C_1(X3) = \frac{\text{Value Jumlah Kuantum} (18b, 19a, 19c, 20a, 20b, 20c, 21a, 21b, 21c, 22a, 22b, 22c)}{12}$$

$$= \frac{(35.000+31.200+61.000+31.200+23.540+50.000+41.500+45.742,6+50.000+41.500+50.812,6+50.000)}{12} = \frac{511.495,2}{12} = 42,624,6$$

Berdasarkan proses perhitungan di atas, maka diperoleh centroid baru sebagai tolak ukur pada iterasi selanjutnya dengan nilai sebagai berikut :

C₁ = [8,4167 ; 2,0833 ; 42.624,6]

C₁ = [7,9231 ; 0,6154 ; 381,7308]

C₁ = [5,6667 ; 2,3889 ; 11.507,4167]

Iterasi 2

Berikut ini hasil perhitungan iterasi II menggunakan rumus Euclidean Distance

Tabel 8. Hasil Perhitungan Iterasi II

Data ke	C1	C2	C3	Cluster
1a	36.542,6008	5.718,2734	5.407,4189	C3
1b	36.424,6008	5.818,2738	5.307,4188	C3
2	25.424,6009	16.818,2704	5.692,5845	C3
3	25.524,6006	16.718,2701	5.592,5840	C3
4	42.624,6003	381,7514	11.507,4171	C2
5	42.624,6003	381,7514	11.507,4171	C2
6	42.624,6003	381,7514	11.507,4171	C2
7	42.624,6003	381,7514	11.507,4171	C2
8	42.624,6003	381,7514	11.507,4171	C2
9	42.624,6003	381,7514	11.507,4171	C2
10	42.624,6003	381,7514	11.507,4171	C2
11	42.624,6003	381,7514	11.507,4171	C2
12	42.624,6003	381,7514	11.507,4171	C2
13	42.624,6003	381,7514	11.507,4171	C2
14	32.624,6006	9.618,2710	1.507,4218	C3
15a	32.640,6004	9.602,2706	1.523,4191	C3
15b	32.624,6005	9.618,2705	1.507,4197	C3
16	23.046,6004	19.196,2696	8.070,5835	C3
17a	34.726,6002	7.516,2699	3.609,4168	C3
17b	30.624,6002	11.618,2698	492,5841	C3
18a	26.936,6001	15.306,2694	4.180,5833	C3
18b	7.624,6004	34.618,2693	23.492,5833	C1
19a	11.424,6001	30.818,2692	19.692,5834	C1
19b	27.979,6	14.266,2693	3.140,5836	C3
19c	18.375,4	60.618,2692	49.492,5833	C1
19d	39.812,1	2.430,7717	8.694,9169	C2
20a	11.424,6	30.818,2692	19.692,5835	C1
20b	19.084,6	23.158,2692	12.032,5835	C3
20c	7.375,4	49.618,2692	38.492,5834	C1
20d	35.662,1	6.580,7701	4.544,9176	C3
21a	1.124,6007	41.118,2692	29.992,5835	C1
21b	3.118	45.360,8692	34.235,1835	C1
21c	7.375,4	49.618,2693	38.492,5834	C1

21d	35.662,1	6.580,7701	4.544,9182	C3
22a	1.124,6016	41.118,2692	29.992,5836	C1
22b	8.188	50.430,8693	39.305,1835	C1
22c	7.375,4	49.618,2693	38.492,5835	C1
22d	35.662,1	6.580,7704	4.544,9190	C3
23a	26.424,6	15.818,2695	4.692,5865	C3
23b	30.474,6	11.768,2697	642,6055	C3
23c	31.124,6	11.118,2699	9,1556	C3
23d	40.474,6	1.768,2751	9.357,4183	C2

Tahap selanjutnya adalah menentukan jarak terpendek data terhadap pusat *cluster* awal yang dipilih secara acak sebelumnya. Untuk menentukan jarak terpendek dengan pusat *cluster* data menggunakan rumus $MIN (X_{1j} : X_{2j} : X_{3j})$. Berikut hasil penentuan jarak terpendek data terhadap pusat *cluster* awal yang telah ditabulasikan.

Tabel 9. Hasil Pengelompokan Iterasi II

Data ke_	C1	C2	C3	Cluster
1a	0	0	1	C3
1b	0	0	1	C3
2	0	0	1	C3
3	0	0	1	C3
4	0	1	0	C2
5	0	1	0	C2
6	0	1	0	C2
7	0	1	0	C2
8	0	1	0	C2
9	0	1	0	C2
10	0	1	0	C2
11	0	1	0	C2
12	0	1	0	C2
13	0	1	0	C2
14	0	0	1	C3
15b	0	0	1	C3
16	0	0	1	C3
17a	0	0	1	C3
17b	0	0	1	C3
18a	0	0	1	C3
18b	1	0	0	C1
19a	1	0	0	C1
19b	0	0	1	C3
19c	1	0	0	C1
19d	0	1	0	C2
20a	1	0	0	C1
20b	1	0	0	C1
20c	1	0	0	C1
20d	0	0	1	C3
21a	1	0	0	C1
21b	1	0	0	C1
21c	1	0	0	C1
21d	0	0	1	C3
22a	1	0	0	C1
22b	1	0	0	C1
22c	1	0	0	C1
22d	0	0	1	C3
23a	0	0	1	C3
23b	0	0	1	C3
23c	0	0	1	C3
23d	0	1	0	C2
24	0	1	0	C2

Perbandingan hasil pengelompokan data pada iterasi ke 2 sama dengan hasil pengelompokan data sebelumnya (iterasi 1) sehingga pencarian iterasi dihentikan



4. KESIMPULAN

Berdasarkan analisa dan pembahasan maka penulis dapat menarik kesimpulan bahwa data impor beras tahun 2016 dan 2018 pada Perum Bulog dapat diimplementasikan menggunakan Algoritma K-Means Clustering. Berdasarkan hasil pengelompokan jumlah kuantum impor beras berdasarkan negara pengekspor bahwa jumlah kuantum impor tertinggi diekspor oleh negara Vietnam dan pemasukan jumlah kuantum impor terbanyak terjadi pada bulan Juli-September tahun 2018.

REFERENCES

- [1] Y. D. d. A. Setiawan, "Penerapan Metode Clustering K-Means Dalam Pengelompokan Penjualan Produk," *Jurnal Media Infotama*, vol. 12, no. 2, pp. 148-156, 2016.
- [2] J. O. Ong, "Implementasi Algoritma K-Means Clustering Untuk Menentukan Strategi Marketing President University," *Jurnal Ilmiah Teknik Industri*, vol. 12, no. 1, pp. 10-20, 2013.
- [3] K. Handoko, "Penerapan Data Mining Dalam Meningkatkan Mutu Pembelajaran Pada Instansi Perguruan Tinggi Menggunakan Metode K-Means Clustering (Studi Kasus Di Program Studi TKJ Akademi Komunitas Solok Selatan)," *Teknosi*, vol. 02, no. 03, pp. 31-40, 2016.
- [4] R. T. Vulandari, *DATA MINING Teori dan Aplikasi Rapidminer*, Yogyakarta: GAVA MEDIA, 2017.
- [5] P. Eko, *Data Mining - Konsep dan Aplikasi Menggunakan Matlab*, Yogyakarta: C.V Andi Offset, 2012.
- [6] Suyanto, *Data Mining Untuk Klasifikasi dan Klasterisasi Data*, Bandung: Penerbit INFORMATIKA, 2017.
- [7] Suarga, *Algoritma dan Pemrograman*, Yogyakarta: C.V Andi Offset, 2012.
- [8] W. M. P. Dhuhita, "Clustering Menggunakan Metode K-Means Untuk Menentukan Status Gizi Balita," *Jurnal Informatika*, vol. 15, no. 2, pp. 160-174, 2015.