

## **Pengelompokan Penyakit Pada Pasien Berdasarkan Usia Dengan Metode K-Means Clustering (Studi Kasus : Puskesmas Bahorok)**

**Tanty<sup>1)</sup>, Budi Serasi Ginting<sup>2)</sup>, Magdalena Simanjuntak<sup>3)</sup>**

**Teknik Informatika Stmik Kaputama Binjai**

Jl. Vetran No. 4A-9A, Binjai, Sumatra Utara, Telp (061)8828840, Fax: (061)8828845

E-mail : [tantyalt17@gmail.com](mailto:tantyalt17@gmail.com)<sup>1)</sup>, [budiserasiginting910@gmail.com](mailto:budiserasiginting910@gmail.com)<sup>2)</sup>,  
[magdalena.simanjuntak84@gmail.com](mailto:magdalena.simanjuntak84@gmail.com)<sup>3)</sup>

### **ABSTRACT**

*In this study, the process of applying the K-Means method will be carried out in classifying diseases in bahorok health center patients based on age. To simplify the process of managing a lot of data, the Bahorok Health Center needs a system in making decisions to find out the grouping of diseases based on the age of patients who are often affected by the disease at the Bahorok Health Center. The application of the K-Means Clustering method is a method used in data mining which works by finding and classifying data that have similar characteristics between one data and other data that has been obtained. As for the results. The results of the pattern analysis above From 20 data obtained 3 groups, it can be concluded as follows: Group 1 Centroid 1: (2.2 1.4 2.2) there are 5 data. Based on the above calculations, it can be seen that in cluster 1 group 1, the patients are children. Group 2 Centroid 2: (1 1,6 1,6) there are 3 data. Based on the above calculations, it can be concluded that in cluster 2, group 2 is an adult patient. Group 3 Centroid 3: (1,75 1,58 6) there are 12 data. Based on the above calculations, it can be concluded that in cluster 3, group 3 is the patient is the elderly.*

**Keywords : K-Means, Diseases in Patients.**

### **ABSTRAK**

Dalam penelitian ini akan dilakukan proses menerapkan metode K-Means dalam pengelompokan penyakit pada pasien puskesmas bahorok berdasarkan usia. Untuk mempermudah proses pengelolaan data yang banyak, Puskesmas Bahorok memerlukan suatu sistem dalam mengambil keputusan untuk mengetahui pengelompokan penyakit berdasarkan usia pasien yang sering terkena penyakit pada Puskesmas Bahorok. Penerapan metode K-Means Clustering merupakan metode yang digunakan dalam data mining yang cara kerjanya mencari dan mengelompokkan data yang mempunyai kemiripan karakteristik antara data satu dengan data lainnya yang telah diperoleh. Adapun hasilnya Hasil analisis pola diatas Dari 20 data diperoleh 3 grup, maka dapat disimpulkan sebagai berikut: Grup 1 Centroid 1 : (2,2 1,4 2,2) terdapat 5 data. Berdasarkan perhitungan diatas dapat diketahui bahwa pada claster 1 kelompok 1 merupakan pasiennya adalah Anak- Anak. Grup 2 Centroid 2 : ( 1 1,6 1,6) terdapat 3 data. Berdasarkan perhitungan diatas dapat disimpulkan pada claster 2 kelompok 2 merupakan pasiennya adalah orang Dewasa. Grup 3 Centroid 3 : (1,75 1,58 6) terdapat 12 data. Berdasarkan perhitungan diatas dapat disimpulkan pada claster 3 kelompok 3 merupakan pasiennya adalah orang Lansia.

**Kata Kunci : K-Means, Penyakit Pada Pasien**

## **1. PENDAHULUAN**

Perkembangan teknologi saat ini sudah sangat pesat, tidak hanya pada bidang informasi, industri, pendidikan, budidaya tetapi pada bidang kesehatan. Para ahli terus mengembangkan teknologi yang ada, sehingga dengan adanya teknologi tersebut para ahli merasa terbantu dalam menyelesaikan pekerjaannya. Dengan adanya berkembang ini diharapkan dapat mempermudah seseorang memperoleh pengetahuan tentang kesehatannya.

Puskesmas merupakan salah satu instansi kesehatan yang berada ditingkat kecamatan, termasuk Puskesmas Bahorok yang berada di kecamatan Bahorok.

Dalam rangka menciptakan pelayanan Instansi Kesehatan yang baik diperlukan tata kerja yang tertib, rapi, dan teliti sehingga menghasilkan informasi yang cepat, akurat, dan tepat waktu sesuai dengan kebutuhan. Seperti instansi kesehatan pada umumnya, Puskesmas Bahorok juga memberikan pelayanan kesehatan bagi masyarakat setiap harinya. Setiap harinya, Puskesmas Bahorok melayani pasien sekitar 55 orang dari berbagai wilayah yang ada di daerah Kecamatan Bahorok. Seiring dengan bertambahnya jumlah pasien tersebut, maka bertambah pula data pasien setiap harinya, sehingga data yang banyak tersebut dapat dimanfaatkan secara optimal oleh Puskesmas Bahorok untuk dijadikan arsip setiap tahunnya dan hasilnya dapat dijadikan bahan atau dasar penyuluhan kesehatan oleh Puskesmas Bahorok.

Inilah yang membuat para tenaga medis banyak mengeluh dikarenakan terlalu banyak pasien yang tidak merata usianya. Salah satu data yang perlu diperhatikan yaitu mengenai penyakit - penyakit pada pasien berdasarkan usia mereka, sehingga nantinya dapat membantu pihak pemerintah khususnya Dinas Kesehatan untuk mengetahui pasien yang sering terkena penyakit berdasarkan usianya. Untuk mempermudah proses pengelolaan data yang banyak, Puskesmas Bahorok memerlukan suatu sistem dalam mengambil keputusan untuk mengetahui pengelompokan penyakit berdasarkan usia pasien yang sering terkena penyakit pada Puskesmas Bahorok.

Penerapan metode *K-Means Clustering* merupakan metode yang digunakan dalam data mining yang cara kerjanya mencari dan mengelompokan data yang mempunyai kemiripan karakteristik antara data satu dengan data lain yang telah diperoleh data yang memiliki kesamaan bukan data yang sama tetapi memiliki karakteristik yang sama.

## 2. LANDASAN TEORI

Pada dasarnya setiap orang sangat mengidam-idamkan kondisi tubuh yang sehat, karena kesehatan sangatlah penting dalam kehidupan setiap manusia sehat merupakan kebutuhan dasar dan hak setiap orang. Kondisi sehat merupakan salah satu faktor penentu kualitas sumber daya manusia, dan sebagai suatu syarat untuk mewujudkan perkembangan jasmani, rohani (mental), sosial, dan ekonomi yang seimbang.

Kesehatan pribadi adalah segala usaha dan tindakan seseorang untuk menjaga, memelihara, dan meningkatkan derajat kesehatannya sendiri dalam batas-batas kemampuannya, agar mendapatkan kesenangan hidup dan mempunyai tenaga kerja yang sebaik – baiknya. Pemeliharaan kesehatan adalah upaya penanggulangan dan pencegahan gangguan kesehatan yang memerlukan pemeriksaan, pengobatan dan perawatan termasuk kehamilan dan persalinan.

Data terakhir menunjukkan bahwa saat ini lebih dari 80 persen rakyat Indonesia tidak mampu mendapat jaminan kesehatan dari lembaga atau perusahaan di bidang pemeliharaan kesehatan, seperti Akses, Taspen, dan Jamsostek. Golongan masyarakat yang dianggap ‘teranaktirikan’ dalam hal jaminan kesehatan adalah mereka dari golongan masyarakat kecil dan pedagang. Pasien adalah orang yang memiliki kelemahan fisik atau mentalnya menyerahkan pengawasan, perhatian dan perawatannya, menerima dan mengikuti pengobatan yang ditetapkan oleh tenaga kesehatan. Pasien sangatlah berpengaruh terhadap puskesmas atau pun rumah sakit, sebab pasien berperan penting dalam berjalannya suatu organisasi puskesmas ataupun rumah sakit. Jika tidak ada pasien siapa yang akan berkunjung dan berobat ke puskesmas dan rumah sakit.

Data Mining merupakan proses pencarian pola dan relasi-relasi yang tersembunyi dalam sejumlah data yang besar dengan tujuan untuk melakukan klasifikasi, estimasi, prediksi, association rule, *Clustering* dan visualisasi. Data Mining bisa dikatakan sebagai

teknik statistik, matematika, kecerdasan buatan, machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai *database* besar. Data mining (penambangan data), sesuai dengan namanya, berkonotasi sebagai pencarian informasi bisnis yang berharga dari basis data yang sangat besar. Didalam data mining terdapat beberapa teknik yang dapat digunakan yaitu sebagai berikut:

1. Deskripsi (*Description*)

Terkadang penelitian analisis secara sederhana ingin mencoba mencari cara deskripsi (*Description*) untuk menggambarkan pola dan kecenderungan yang terdapat dalam data. Sebagai contoh, petugas mengumpulkan suara mungkin tidak dapat menemukan keterangan atau fakta bahwa siapa yang tidak cukup profesional akan sedikit didukung dalam pemilihan presiden. Deskripsi dari pola dan kecenderungan sering memberikan kemungkinan penjelasan untuk suatu pola atau kecenderungan.

2. Estimasi (*Estimation*)

Estimasi hampir sama dengan klasifikasi, kecuali variabel target estimasi lebih ke arah numerik daripada ke arah kategori. Model dibangun menggunakan record lengkap yang menyediakan nilai dari variabel target sebagai nilai prediksi.

3. Prediksi (*Prediction*)

Prediksi hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil yang akan ada di masa datang.

4. Klasifikasi (*Classification*)

Dalam klasifikasi, terdapat target variabel kategori. Sebagai contoh, penggolongan pendapatan dapat dipisahkan dalam tiga kategori. yaitu pendapatan tinggi, pendapatan sedang dan pendapatan rendah. Berdasarkan atas nilai – nilai suatu variabel yang ada untuk pengklasifikasian suatu data.

5. Pengclusteran (*Clustering*)

Pengclusteran merupakan pengelompok record, pengamatan atau memperhatikan dan membentuk kelas objek-objek yang memiliki kemiripan. *Cluster* adalah kumpulan record yang memiliki kemiripan satu dengan yang lainnya dan memiliki ketidak miripan dengan record - record dalam *cluster* lain. Pengklusteran berbeda dengan klasifikasi yaitu tidak adanya variabel target dalam pengklusteran. Pengklusteran tidak mencoba untuk melakukan klasifikasi, mengestimasi, atau memprediksi nilai dari variabel target.

6. Asosiasi (*Assosiation*)

Tugas asosiasi dalam data mining adalah menemukan atribut yang muncul dalam satu waktu. Dalam dunia bisnis lebih umum disebut analisis keranjang belanja.

*Clustering* merupakan salah satu bagian dari teknik data mining yaitu sekumpulan objek yang mempunyai karakteristik diantara anggotanya dan memiliki “ketidaksamaan” dengan objek lain pada *cluster* lainnya, dengan kata lain sebuah *cluster* adalah sekumpulan objek yang digabung bersama karena persamaan atau kedekatannya. *Clustering* dapat dibedakan menjadi dua tujuan, yaitu *Clustering* untuk pemahaman dan *Clustering* untuk penggunaan. Jika tujuan untuk pemahaman maka *cluster* yang terbentuk harus menangkap struktur alami data. Biasanya proses *Clustering* dalam tujuan ini hanya sebagai proses awal untuk kemudian dilanjutkan dengan pekerjaan ini seperti summarization (rata-rata, standar deviasi), pelabelan kelas pada setiap kelompok untuk kemudian digunakan sebagai data latih klasifikasi, dan sebagainya. Sementara jika tujuannya untuk penggunaan, biasanya tujuan utama untuk mencari prototype *cluster* yang paling representative terhadap data dan memberikan abstraksi dan setiap objek data dalam *cluster* dimana sebuah data terletak didalamnya.

*K-Means* clustering merupakan salah satu metode data *Clustering* nonhirarki yang mengelompok data dalam bentuk satu atau lebih *cluster* atau kelompok. Data – data yang memiliki karakteristik yang sama dikelompokkan dalam satu *cluster* atau kelompok

dan data yang memiliki karakteristik yang berbeda dikelompokkan dengan *cluster* metode pengelompokan data/kelompok yang lain sehingga data berada dalam satu *cluster* atau kelompok memiliki tingkat variasi yang kecil.

Algoritma *K-Means* adalah metode *Clustering* non hierarchical berbasis jarak yang membagi data kedalam *cluster* dan algoritma ini bekerja pada atribut numerik. Algoritma *K-Means* termasuk dalam *partitioning Clustering* yang memisahkan data ke daerah bagian yang terpisah. Algoritma *K-Means* sangat terkenal karena kemudahannya dan kemampuannya untuk mengklaster data besar dan outlier dengan sangat cepat. Algoritma *K-Means* pada dasarnya melakukan 2 proses yakni proses pendeteksian lokasi pusat *cluster* dan proses pencarian anggota dari tiap-tiap *cluster* dan proses pencarian anggota dari tiap-tiap *cluster*.

### 3. METODE PENELITIAN

Penelitian ini menggunakan teknik data mining dengan metode *K-Means Clustering* yang bertujuan untuk mengelompokkan penyakit pada pasien berdasarkan usianya. Proses Algoritma *K-Means* sebagai berikut :

1. Tentukan K sebagai jumlah *cluster* yang ingin dibentuk.
2. Bangkitkan K centroids (titik pusat *cluster*) awal secara random..
3. Hitung jarak setiap data ke masing-masing centroids.
4. Setiap data memilih centroids yang terdekat.
5. Tentukan posisi centroids baru dengan cara menghitung nilai rata-rata dari data-data yang terletak pada centroids yang sama. Kembali ke langkah 3 jika posisi centroids baru dengan centroids lama tidak sama.

### 4. HASIL DAN PEMBAHASAN

#### 4.1 Hasil Analisa

Tujuan dari penelitian ini adalah bagaimana mengelompokkan penyakit pada pasien berdasarkan usianya dengan metode *K-Means Clustering* pada Puskesmas Bahorok. Jumlah data yang akan digunakan sebanyak 20 data penyakit pasien. Dalam menggunakan metode *Clustering*, proses yang dilakukan untuk pembentukan *cluster* data yang akan ditransformasikan kedalam bentuk numerik dengan kode – kode yang akan ditentukan, lalu tentukan jumlah grup (K), hitung centroid, hitung jarak objek ke centroid dan kemudian grupkan berdasarkan jarak terdekat, jika tidak ada objek yang berpindah grup maka iterasi selesai. Untuk menentukan grup dari suatu objek, pertama yang harus dilakukan adalah dengan mengukur jarak *Euclidean* antara dua titik objek ( X dan Y) yang didefinisikan sebagai berikut:

$$Euclidean Distance: D(ij) = \sqrt{(X1i - X1j)^2 + (X2i - X2j)^2 + (X3i - X3j)^2}$$

**Tabel 1. Sampel Data Pasien**

No	Nama	Usia	Jenis Kelamin	Diagnosa
1	Suhartono	53	Laki – laki	Stroke Ringan
2	Mykayla	7	Perempuan	Demam
3	Suhartini	55	Perempuan	Stroke Ringan
4	Suhartanto	57	Laki – laki	Sesak Nafas
5	Ranika	17	Perempuan	Batuk
6	Rudi	15	Laki – laki	Muntah-Muntah

No	Nama	Usia	Jenis Kelamin	Diagnosa
7	Nining	22	Perempuan	Perut Kembung
8	Hidayat	25	Laki – laki	Flu
9	Annisa	27	Perempuan	Perut Kembung
10	Joko Sutanto	67	Laki – laki	Gatal - Gatal
11	Rudiansyah	10	Laki – laki	Demam
12	Anggun	12	Perempuan	Batuk
13	Wawan	33	Laki – laki	Flu
14	Endang	38	Perempuan	Flu
15	Mega	75	Perempuan	Gatal - Gatal
16	Fahmi	72	Laki – laki	Sesak Nafas
17	Agung	15	Laki – laki	Muntah - Muntah
18	Ayu Annisa	18	Perempuan	Muntah - Muntah
19	Sundari	70	Perempuan	Stroke Ringan
20	Wanda	5	Laki – laki	Demam

Data yang diatas harus diubah nilainya ke dalam bentuk angka agar memudahkan untuk menghitung jaraknya. Nilai – nilai tersebut ditransformasikan terlebih dahulu seperti pada tabel 2. dibawah ini.

**Tabel 2. Nilai Transformasi Data Pada Variabel**

No	Variabel	Transformasi	Nilai Transformasi
1	Usia	5 s/d 16 Tahun	1
		17 s/d 50 Tahun	2
		51 s/d 75 Tahun	3
2	Jeis Kelamin	Laki-laki	1
		Perempuan	2
2	Diagnosa	Stroke Ringan	1
		Demam	2
		Sesak Nafas	3
		Batuk	4
		Muntah - Muntah	5
		Perut Kembung	6
		Flu	7
		Gatal - Gatal	8

Agar data diatas Agar data dapat diolah dengan menggunakan algoritma *K-Means* maka data yang berjenis non *numeric* seperti usia, jenis kelamin dan diagnosa harus diinisialisasikan terlebih dahulu kedalam bentuk *numeric* (angka). Data ini dapat dinyatakan dalam suatu variabel – variabel yang memudahkan untuk dipahami yaitu Usia (X), Jenis Kelamin (Y), Diagnosa (Z) adalah sebagai berikut:

**Tabel 3. Transformasi Data**

No	Nama	Usia (X)	Jenis Kelamin (Y)	Diagnosa (Z)
1	A	3	1	1
2	B	1	2	2
3	C	3	2	1
4	D	3	1	3
5	E	1	2	4
6	F	1	1	5
7	G	2	2	6
8	H	2	1	7
9	I	2	2	6
10	J	3	1	8
11	K	1	1	2
12	L	1	2	4
13	M	2	1	7
14	N	2	2	7
15	O	3	2	8
16	P	3	1	3
17	Q	1	1	5
18	R	1	2	5
19	S	3	2	1
20	T	1	1	2

Adapun berikut merupakan proses perhitungan metode *K-Means*. Pada contoh ini akan dibentuk 3 *cluster*, Nilai Centroid yang diambil secara acak yaitu pusat *cluster* 1 dan pusat *cluster* 3, inialisasi *cluster* dipilih secara random dengan range data antara nilai terendah sampai nilai tertinggi. Adapun nilai untuk setiap pusat *cluster* ditunjukkan sebagai berikut:

**Tabel 4. Pusat Sample Cluster Random**

No	Nama	X	Y	Z
1	A	3	1	1
3	C	3	2	1
7	G	2	2	6

Hitung jarak terhadap semua titik data dengan rumus *Euclidean distance*. Adapun perhitungan sebagai berikut:

1. A (3,1,1)  
 $C1(3,1,1) = \sqrt{(3-3)^2 + (1-1)^2 + (1-1)^2} = 0$   
 $C2(3,2,1) = \sqrt{(3-3)^2 + (1-2)^2 + (1-1)^2} = 1$   
 $C3(2,2,6) = \sqrt{(3-2)^2 + (1-2)^2 + (1-6)^2} = 5,196$
2. B (1,2,2)

- $C1(3,1,1) = \sqrt{(1-3)^2 + (2-1)^2 + (2-1)^2} = 2,449$   
 $C2(3,2,1) = \sqrt{(1-3)^2 + (2-2)^2 + (2-1)^2} = 2,236$   
 $C3(2,2,6) = \sqrt{(1-2)^2 + (2-2)^2 + (2-6)^2} = 4,123$
3. C (3,2,1)
- $C1(3,1,1) = \sqrt{(3-3)^2 + (2-1)^2 + (1-1)^2} = 1$   
 $C2(3,2,1) = \sqrt{(3-3)^2 + (2-2)^2 + (1-1)^2} = 0$   
 $C3(2,2,6) = \sqrt{(3-2)^2 + (2-2)^2 + (1-6)^2} = 5,099$
4. D (3,1,3)
- $C1(3,1,1) = \sqrt{(3-3)^2 + (1-1)^2 + (3-1)^2} = 2$   
 $C2(3,2,1) = \sqrt{(3-3)^2 + (1-2)^2 + (3-1)^2} = 2,236$   
 $C3(2,2,6) = \sqrt{(3-2)^2 + (1-2)^2 + (3-6)^2} = 3,316$
5. E (1,2,4)
- $C1(3,1,1) = \sqrt{(1-3)^2 + (2-1)^2 + (4-1)^2} = 3,7416$   
 $C2(3,2,1) = \sqrt{(1-3)^2 + (2-2)^2 + (4-1)^2} = 3,605$   
 $C3(2,2,6) = \sqrt{(1-2)^2 + (2-2)^2 + (4-6)^2} = 2,236$
6. F (1,1,5)
- $C1(3,1,1) = \sqrt{(1-3)^2 + (1-1)^2 + (5-1)^2} = 4,472$   
 $C2(3,2,1) = \sqrt{(1-3)^2 + (1-2)^2 + (5-1)^2} = 4,582$   
 $C3(2,2,6) = \sqrt{(1-2)^2 + (1-2)^2 + (5-6)^2} = 1,732$
7. G (2,2,6)
- $C1(3,1,1) = \sqrt{(2-3)^2 + (2-1)^2 + (6-1)^2} = 5,196$   
 $C2(3,2,1) = \sqrt{(2-3)^2 + (2-2)^2 + (6-1)^2} = 5,099$   
 $C3(2,2,6) = \sqrt{(2-2)^2 + (2-2)^2 + (6-6)^2} = 0$
8. H (2,1,7)
- $C1(3,1,1) = \sqrt{(2-3)^2 + (1-1)^2 + (7-1)^2} = 6,082$   
 $C2(3,2,1) = \sqrt{(2-3)^2 + (1-2)^2 + (7-1)^2} = 6,164$   
 $C3(2,2,6) = \sqrt{(2-2)^2 + (1-2)^2 + (6-6)^2} = 1,414$
9. I (2,2,6)
- $C1(3,1,1) = \sqrt{(2-3)^2 + (2-1)^2 + (6-1)^2} = 5,196$   
 $C2(3,2,1) = \sqrt{(2-3)^2 + (2-2)^2 + (6-1)^2} = 5,099$   
 $C3(2,2,6) = \sqrt{(2-2)^2 + (2-2)^2 + (6-6)^2} = 0$
10. J (3,1,8)
- $C1(3,1,1) = \sqrt{(3-3)^2 + (1-1)^2 + (8-1)^2} = 7$   
 $C2(3,2,1) = \sqrt{(3-3)^2 + (1-2)^2 + (8-1)^2} = 7,071$   
 $C3(2,2,6) = \sqrt{(3-2)^2 + (1-2)^2 + (8-6)^2} = 2,449$
11. K (1,1,2)
- $C1(3,1,1) = \sqrt{(1-3)^2 + (1-1)^2 + (2-1)^2} = 2,236$   
 $C2(3,2,1) = \sqrt{(1-3)^2 + (1-2)^2 + (2-1)^2} = 2,449$   
 $C3(2,2,6) = \sqrt{(1-2)^2 + (1-2)^2 + (2-6)^2} = 4,242$
12. L (1,2,4)
- $C1(3,1,1) = \sqrt{(1-3)^2 + (2-1)^2 + (4-1)^2} = 3,741$   
 $C2(3,2,1) = \sqrt{(1-3)^2 + (2-2)^2 + (4-1)^2} = 3,605$   
 $C3(2,2,6) = \sqrt{(1-2)^2 + (2-2)^2 + (4-6)^2} = 2,236$
13. M (2,1,7)
- $C1(3,1,1) = \sqrt{(2-3)^2 + (1-1)^2 + (7-1)^2} = 6,082$   
 $C2(3,2,1) = \sqrt{(2-3)^2 + (1-2)^2 + (7-1)^2} = 6,164$   
 $C3(2,2,6) = \sqrt{(2-2)^2 + (1-2)^2 + (7-6)^2} = 1,414$
14. N (2,2,7)
- $C1(3,1,1) = \sqrt{(2-3)^2 + (2-1)^2 + (7-1)^2} = 6,164$   
 $C2(3,2,1) = \sqrt{(2-3)^2 + (2-2)^2 + (7-1)^2} = 6,082$

$$C3(2,2,6) = \sqrt{(2-2)^2 + (2-2)^2 + (7-6)^2} = 1$$

15. O (3,2,8)

$$C1(3,1,1) = \sqrt{(3-3)^2 + (2-1)^2 + (8-1)^2} = 7,071$$

$$C2(3,2,1) = \sqrt{(3-3)^2 + (2-2)^2 + (8-1)^2} = 7$$

$$C3(2,2,6) = \sqrt{(3-2)^2 + (2-2)^2 + (8-6)^2} = 2,236$$

16. P (3,1,3)

$$C1(3,1,1) = \sqrt{(3-3)^2 + (1-1)^2 + (3-1)^2} = 2$$

$$C2(3,2,1) = \sqrt{(3-3)^2 + (1-2)^2 + (3-1)^2} = 2,236$$

$$C3(2,2,6) = \sqrt{(3-2)^2 + (1-2)^2 + (3-6)^2} = 3,316$$

17. Q (1,1,5)

$$C1(3,1,1) = \sqrt{(1-3)^2 + (1-1)^2 + (5-1)^2} = 4,472$$

$$C2(3,2,1) = \sqrt{(1-3)^2 + (1-2)^2 + (5-1)^2} = 4,582$$

$$C3(2,2,6) = \sqrt{(1-2)^2 + (1-2)^2 + (5-6)^2} = 1,732$$

18. R (1,2,5)

$$C1(3,1,1) = \sqrt{(1-3)^2 + (2-1)^2 + (5-1)^2} = 4,582$$

$$C2(3,2,1) = \sqrt{(1-3)^2 + (2-2)^2 + (5-1)^2} = 4,472$$

$$C3(2,2,6) = \sqrt{(1-2)^2 + (2-2)^2 + (5-6)^2} = 1,414$$

19. S (3,2,1)

$$C1(3,1,1) = \sqrt{(3-3)^2 + (2-1)^2 + (1-1)^2} = 1$$

$$C2(3,2,1) = \sqrt{(3-3)^2 + (2-2)^2 + (1-1)^2} = 0$$

$$C3(2,2,6) = \sqrt{(3-2)^2 + (2-2)^2 + (1-6)^2} = 5,099$$

20. T (1,1,2)

$$C1(3,1,1) = \sqrt{(1-3)^2 + (1-1)^2 + (2-1)^2} = 2,236$$

$$C2(3,2,1) = \sqrt{(1-3)^2 + (1-2)^2 + (2-1)^2} = 2,449$$

$$C3(2,2,6) = \sqrt{(1-2)^2 + (1-2)^2 + (2-6)^2} = 4,242$$

Perhitungan yang sama dapat dilakukan hingga data ke 20. Dari perhitungan diatas dapat diperoleh angka atau hasil pada setiap grup yang paling dekat jaraknya. Berikut ini dapat dilihat hasil perhitungannya pada tabel 5. dibawah ini.

Setelah masing-masing data dihitung jaraknya untuk tiap *cluster*, langkah selanjutnya adalah mengelompokkan data sesuai *clusternya*. Kelompok *cluster* suatu data diambil dari jarak terpendek data tersebut terhadap suatu *cluster*. Misalnya untuk data pasien A memiliki jarak 0 terhadap *cluster* 1. Pada *cluster* 2 memiliki jarak 1. Pada *cluster* 3 memiliki jarak 5,196 Oleh karena itu data pasien A masuk ke dalam *cluster* 1. Langkah yang sama diterapkan di ke-20 data untuk melakukan pengelompokan di iterasi 1. Hasil pengelompokan dapat dilihat pada tabel 5. dibawah ini:

**Tabel 5. Hasil Perhitungan Iterasi 1**

No	Nama	C1	C2	C3	Grup
1	A	0	1	5,196	1
2	B	2,449	2,236	4,123	2
3	C	1	0	5,099	2
4	D	2	2,236	3,316	1
5	E	3,741	3,605	2,236	3
6	F	4,472	4,582	1,732	3
7	G	5,196	5,099	0	3
8	H	6,082	6,164	1,414	3



No	Nama	C1	C2	C3	Grup
9	I	5,196	5,099	0	3
10	J	7	7,071	2,449	3
11	K	2,236	2,449	4,242	1
12	L	3,741	3,605	2,236	3
13	M	6,082	6,164	1,414	3
14	N	6,164	6,082	1	3
15	O	7,071	7	2,236	3
16	P	2	2,236	3,316	1
17	Q	4,472	4,582	1,732	3
18	R	4,582	4,472	1,414	3
19	S	1	0	5,099	2
20	T	2,236	2,449	4,242	1

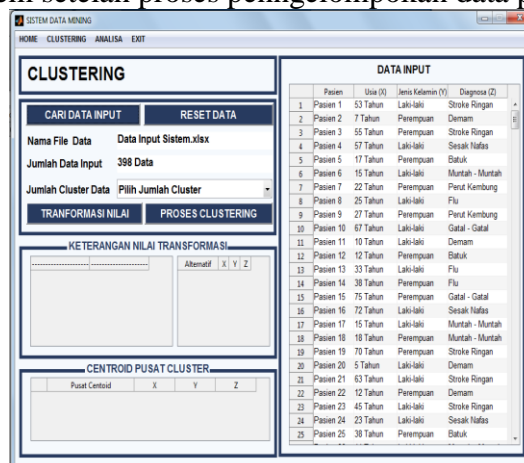
Keterangan dari tabel diatas adalah terdapat 5 data pada *cluster* pertama (1), terdapat 3 data pada *cluster* kedua (2), dan terdapat 12 data pada *cluster* ketiga (3).

#### 4.2 Uji Coba Sistem

Setelah melakukan implementasi metode pada rancangan sistem terhadap sistem pengelompokan data, maka untuk mengetahui hasil dari implementasi sistem tersebut, perlu dilakukan uji coba terhadap sistem yang telah selesai dirancang. Pengujian akan dilakukan dengan memproses data input sistem, proses pengelompokan data yang akan dilakukan adalah sebagai berikut:

##### 1. Proses Input Data

Pada tahap ini pengguna akan melakukan proses pengelompokan data yang telah diinputkan pada sistem. Proses pengelompokan data pada sistem dapat dilakukan dengan menekan tombol “PROSES CLUSTERING” pada halaman *Clustering* data. Proses akan memanfaatkan data yang telah diinputkan dan jumlah *cluster* yang telah di tentukan. . Berikut ini tampilan sistem setelah proses pengelompokan data pada sistem:

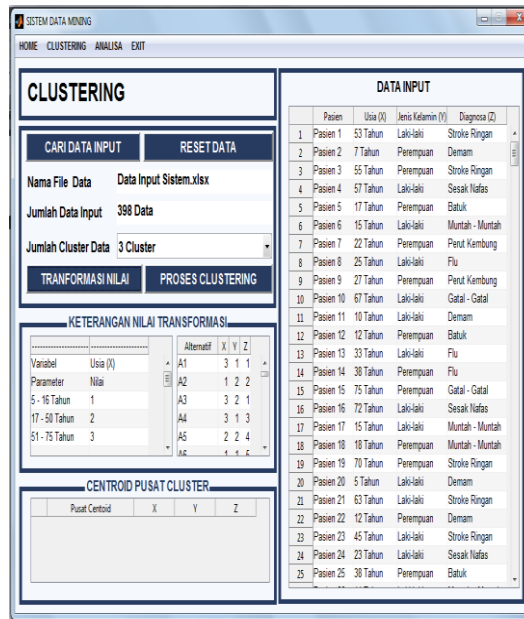


**Gambar 1. Input Data**

##### 2. Proses Transformasi Nilai

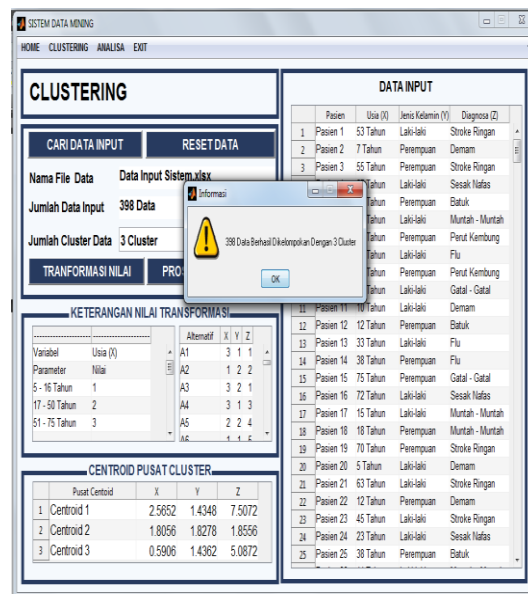
Data dapat diinputkan pada sistem dengan menggunakan tombol “CARI DATA INPUT” pada halaman proses *Clustering* di sistem. Setelah data di inputkan pada sistem, data akan

di tampilkan pada tabel data input dan sistem akan menampilkan jumlah data yang telah di inputkan. Proses nilai transformasi data dapat dilakukan dengan menekan tombol “TRANSFORMASI NILAI”. Berikut ini adalah tampilan setelah proses transformasi nilai pada data:



**Gambar 2. Proses Transformasi**

### 3. Proses Pengelompokan Data



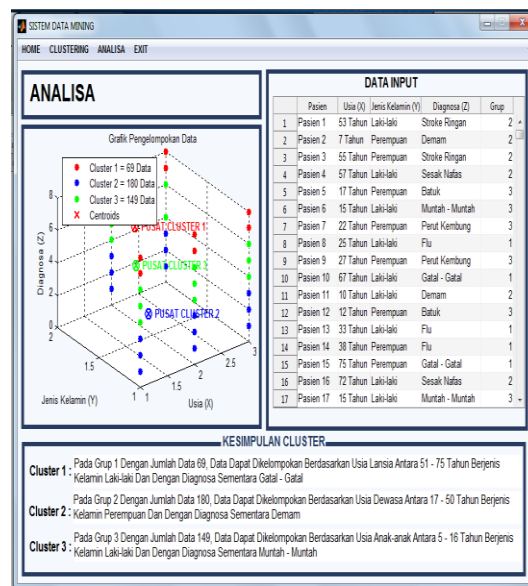
**Gambar 3. Pengelompokan Data**

Dapat dilihat pada gambar diatas dalam proses pengelompokan data penyakit pasien pada Puskesmas Bahorok yang telah dilakukan, dihasilkan pusat *cluster* data sebagai berikut:

- Centroid 1 : (2,5652; 1,4348; 7,5072)
- Centroid 2 : (1,8056; 1,8278; 1,8556)
- Centroid 3 : (0,5906; 1,4362; 5,0872)

Tampilan ini adalah tampilan hasil proses pengelompokan data yang telah dilakukan pada sistem pengelompokan data dengan metode *Clustering* yang telah di proses, halaman ini terdapat pada menu “ANALISA” pada sistem, pada tampilan ini sistem akan menampilkan data input yang telah dimasukkan beserta grup pengelompokan hasil proses pada sistem yang telah dilakukan, pada tampilan ini juga sistem akan menampilkan keterangan hasil *cluster* data yang menjadi pusat *cluster* pengelompokan data dan grafik data hasil pengelompokan yang memberikan keterangan berupa titik koordinat pengelompokan data serta pusat *cluster* data, tampilan hasil sebagai pengelompokan data tersebut adalah sebagai berikut:

#### 4. Grafik Proses Pengelompokan Data



**Gambar 4. Grafik Pengelompokan Data**

## 5. KESIMPULAN DAN SARAN

### 5.1 Kesimpulan

Dari penulisan skripsi yang berjudul Pengelompokan Penyakit Pada Pasien Berdasarkan Usia Dengan Metode *K-Means Clustering* (Studi Kasus : Puskesmas Bahorok), kesimpulan yang di dapat sebagai berikut:

1. Pengelompokan dengan metode *K-Means* dapat menghasilkan jumlah *cluster* yang sama dengan jumlah data yang berbeda – beda tanpa harus memiliki data yang sama.
2. Dengan dibangunnya sistem ini untuk mempermudah user dalam mengelompokkan penyakit pada pasien berdasarkan usia secara efektif dan efisien khususnya untuk Staff Pegawai dan Administrasi.
3. Dengan metode *K-Means* sangatlah mempermudah user dalam mengelompokkan suatu data hanya dengan memiliki karakteristik yang sama.

### 5.2 Saran

Dari penulisan skripsi ini adapun penulis memberikan saran untuk dapat dikembangkan untuk masa yang akan datang sebagai berikut:

1. Diharapkan kedepannya perlu dilakukan perbandingan terhadap algoritma lain, untuk menguji sejauh mana algoritma apriori masih dapat diandalkan untuk memproses dan menemukan pola hubungan asosiasi antar item pada *database* skala besar.

2. Diharapkan kedepannya adanya pengembangan baik berupa penambahan fitur-fitur aplikasi yang sesuai dibutuhkan dan anti virus yang berguna mengamankan sistem untuk melindungi data-data yang berada di dalam *database*.
3. Diharapkan sebaiknya ditambahkan fasilitas untuk backup data. Jadi, jika terjadi kerusakan pada server data tidak akan terhapus.

#### **DAFTAR PUSTAKA**

- [1]. Amir, Tjolleng. 2017. Pengantar Pemrograman Matlab. PT Elex Media Komputido. Jakarta.
- [2]. Agus, N. Khomaruddin. 2016. Teknik Data Mining : Algoritma *K-Means Clustering*. Komunitas eLearning IlmuKomputer.Com.
- [3]. Agus, Sugianto, dkk. 2020. Algoritma *K-Means* Untuk Pengelompokan Penyakit Pasien Pada Puskesmas Cigugur Tengah. Jurnal Gusman Politeknik TEDC. Bandung.
- [4]. Castaka, A. Sugianto, dkk. 2020. Algoritma *K-Means* Untuk Pengelompokan Penyakit Pasien Pada Puskesmas Cigugur Tengah. Journal of Information Technology. Vol. 02 No 02 Agustus 2020, pp. 39-44. (p)ISSN: 2527-9467/(e)ISSN: 2656-7539.
- [5]. Fitra, Kurnia, dkk. 2019. Penerapan Algoritma *K-Means* Untuk Pengelompokan Diagnosa Penyakit M.. Jurnal SPEKTRO / Vol. 2 No. 1 Mei 2019 : ISSN: 2655-5778.
- [6]. Mentari, T. Indah Rahmayani. 2018. Analisis *Clustering* Tingkat Keparahan Penyakit Pasien Menggunakan Algoritma *K-Means* (Studi Kasus Di Puskesmas Bandar Seikijang). JITI, Vol.1, No. 2, September 2018.
- [7]. Ulil, Ma'rifatin. 2020. Implementasi Algoritma *K-Means* Untuk Pengelompokan Penyakit Pasien Pada Puskesmas Warujayeng. Seminar Nasional Inovasi Teknologi : e-ISSN: 2549-7952.
- [8]. Windha, M. P. Duhita. 2015. *Clustering* Menggunakan Metode *K-Means* Untuk Menentukan Status Gizi Balita. Jurnal Informatika, Vol. 15, No. 2, Bulan Desember 2015.