



**CLUSTERISASI POLA PENJUALAN OBAT MENGGUNAKAN
METODE K-MEANS CLUSTERING
(STUDI KASUS DI APOTEK PELANGI SEI. JERING TELUK KUANTAN)**

Sri Yudarwati

Program Studi Teknik Informatika,
Fakultas Teknik,
Universitas Islam Kuantan Singingi, Indonesia
Jl. Gatot Subroto KM. 7 Kebun Nenas, Desa Jake, Kab. Kuantan Singingi

ABSTRAK

Data Mining adalah suatu proses penggalian sebuah informasi data yang tersembunyi yang ada di dalam suatu database, dimana data tersebut disimpan di dalam suatu tempat ukuran yang sangat besar seperti data warehouse dan tempat penyimpanan data lainnya. Data Mining juga sering dikenal dengan istilah Knowledge Discovery in Database (KDD). Data-data yang telah dikelompokkan nantinya dapat diolah kembali menjadi suatu pengetahuan yang baru sehingga dapat digunakan lagi apabila dibutuhkan. Apotek Pelangi Sei. Jering Teluk Kuantan merupakan salah satu toko yang terletak di Kecamatan Kuantan Tengah di Kabupaten Kuantan Singingi Apotek ini menjual berbagai macam jenis obat-obatan. Namun di dalam melakukan proses pencatatan data-data, Apotek masih menggunakan sistem secara manual sehingga apotek tidak dapat mengetahui dan tidak dapat mengelompokkan jenis obat apa saja yang paling tinggi penjualannya. Sehingga kesulitan yang dialami oleh apotek adalah seringnya terjadi kekurangan stok barang yang laku dikarenakan penjualan yang cukup tinggi. Data penjualan yang terjadi pada apotek pelangi setiap satu musimnya juga berbeda-beda sehingga perlu adanya pengelompokan (klasterisasi/clustering). Permasalahan yang terjadi adalah kesulitan dalam pengelompokkan data dalam jumlah besar. Hasil akhir dari 15 items penjualan sebagai data contoh pada Bulan Juli – Desember 2017 adalah penjualan Obat Laris sebanyak 6 items dan Penjualan Obat Kurang Laris sebanyak 9 items. Menghasilkan aplikasi sistem Data Mining metode K-Means clustering dalam pola penjualan yaitu dapat membandingkan hasil yang secara manual dengan aplikasi sistem yang dibangun.

Kata Kunci : Data Mining, K-Means Clustering, Aplikasi

1. PENDAHULUAN

Data Mining adalah suatu proses penggalian sebuah informasi data yang tersembunyi yang ada di dalam suatu database, dimana data tersebut disimpan di dalam suatu tempat ukuran yang sangat besar seperti data warehouse dan tempat penyimpanan data lainnya. Data Mining juga sering dikenal dengan istilah Knowledge Discovery in Database (KDD). Data-data yang telah dikelompokkan nantinya dapat diolah kembali menjadi suatu pengetahuan yang baru sehingga dapat digunakan lagi apabila dibutuhkan.

Salah satu metode yang terjadi di dalam Data Mining adalah clustering. Clustering merupakan salah satu metode analisa data yang dapat digunakan dalam memecahkan masalah dalam suatu pengelompokan data. Salah satu metode yang ada di dalam metode clustering adalah metode K-means. Metode K-means merupakan suatu metode yang dapat



melakukan pengelompokan data dalam jumlah yang cukup besar dengan perhitungan waktu yang relatif cepat dan efisien.

Apotek Pelangi Sei. Jering Teluk Kuantan merupakan salah satu toko yang terletak di Kecamatan Kuantan Tengah di Kabupaten Kuantan Singingi Apotek ini menjual berbagai macam jenis obat-obatan. Namun di dalam melakukan proses pencatatan data-data, Apotek masih menggunakan sistem secara manual sehingga apotek tidak dapat mengetahui dan tidak dapat mengelompokkan jenis obat apa saja yang paling tinggi penjualannya. Sehingga kesulitan yang dialami oleh apotek adalah seringnya terjadi kekurangan stok barang yang laku dikarenakan penjualan yang cukup tinggi. Data penjualan yang terjadi pada apotek pelangi setiap satu musimnya juga berbeda-beda sehingga perlu adanya pengelompokan (klasterisasi/ clustering).

2. METODE PENELITIAN

2.1 Teknik Pengumpulan Data

Berdasarkan sumber data yang digunakan pada penelitian ini, maka metode pengumpulan data yang penulis gunakan adalah sebagai berikut:

1. Wawancara

Wawancara adalah metode pengumpulan data melalui kegiatan tanya jawab langsung dengan person yang ada pada objek penelitian. Wawancara penulis lakukan dengan pimpinan Apotek Pelangi Sei. Jering Teluk Kuantan.

2. Dokumentasi

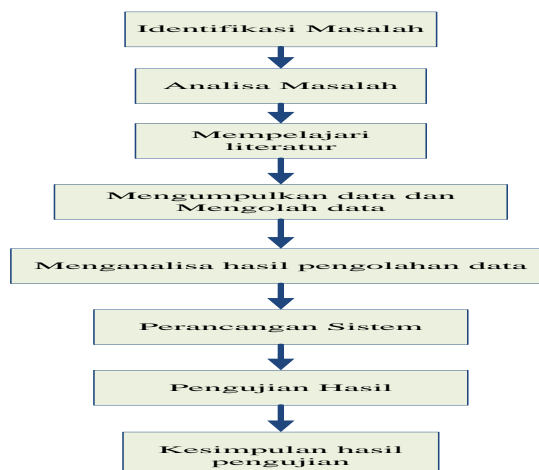
Dokumentasi adalah metode pengumpulan data dengan mengumpulkan catatan-catatan atau dokumen-dokumen. Data yang telah penulis dapatkan dari metode wawancara, penulis kumpulkan menjadi satu sehingga menjadi sebuah dokumen yang siap digunakan untuk kepentingan penelitian.

3. Studi Pustaka

Studi Pustaka adalah mempelajari karya ilmiah, buku ilmiah, dan sumber ilmiah lainnya yang sesuai dengan penelitian dan memiliki hubungan dengan masalah yang diteliti. Referensi ilmiah yang penulis gunakan adalah sumber-sumber yang terdapat dalam daftar kepustakaan.

2.2 Rancangan Penelitian

Ada beberapa tahapan dari penelitian yang akan dilaksanakan yaitu tergambar pada gambar sebagai berikut:



Gambar 1. Rancangan Penelitian



3 HASIL DAN PEMBAHASAN

3.1 Tahapan Metode *K-Means Clustering*

Tahapan-tahapan yang dijalankan dalam proses perhitungan menggunakan metode *k-means* adalah:

1. Penentuan pusat *Cluster* baru

Untuk mendapatkan pusat *Cluster* baru bisa dihitung dari rata-rata nilai anggota *Cluster* dan pusat *Cluster*. Pusat *Cluster* yang baru digunakan untuk melakukan iterasi selanjutnya, jika hasil yang didapatkan belum konvergen. Proses iterasi akan berhenti jika telah memenuhi maksimum iterasi yang dimasukkan oleh user atau hasil yang dicapai sudah konvergen (pusat *Cluster* baru sama dengan pusat *Cluster* lama). Algoritma penentuan pusat *Cluster*, yaitu :

- a. Cari jumlah anggota tiap *Cluster*
- b. Hitung pusat baru dengan rumus

Dari aturan-aturan asosiasi tersebut, penulis akan merancang sistem dengan melakukan perhitungan terhadap data yang telah didapat. Data yang masuk dalam perhitungan hanya beberapa data transaksi saja sebagai sampel dari perancangan sistem.

Tabel 1. Transaksi Penjualan dan Harga Obat

No	Nama Obat	Satuan	Total Penjualan	Harga
1	Remco Cought	118 ml	30	8000
2	Anacomidine	60 ml	80	10000
3	Pacdin Cought	60 ml	20	8000
4	Unibebe Cought Syrup	60 ml	3	10000
5	Vipcol Sirup	60 ml	6	8000
6	Snow Playe	Botol	4	30000
7	Termorex	60 ml	90	16000
8	Proxion	60 ml	10	29000
9	Tempra Sirup	60 ml	15	24000
10	Afibremol	60 ml	2	8000
11	Paracetamol Sirup	60 ml	10	5000
12	Pamtusin	60 ml	25	30000
13	Dyrexin	60 ml	45	14000
14	Batugin	300 ml	16	27000
15	Combatriin	125 mg	37	20000

Tahap selanjutnya adalah melakukan proses utama yaitu segmentasi atau Pengelompokan data penjualan obat yang diakses dari database, yaitu sebuah metode *Clustering* algoritma *K-Means*. Dari banyak data penjualan yang diperoleh, Percobaan dilakukan dengan menggunakan parameter-parameter berikut :

Jumlah *Cluster* : 2
 Jumlah data : 15
 Jumlah atribut : 2



2. Proses iterasi 1

1. Penentuan pusat awal Cluster

Pusat awal Cluster atau centroid didapatkan secara random (acak), untuk penentuan awal Cluster ini dapat diasumsikan sebagai berikut:

Pusat Cluster 1, diambil dari data (hasil penjumlahan total penjualan) / (jumlah banyak sampel yang digunakan) = $375 : 15 = 25$ dan (jumlah total harga) / (jumlah banyak sampel yang digunakan) = $311000 : 15 = 20733.33$

Jadi Pusat Cluster 1 adalah (25, 20733.33)

Pusat Cluster 2, diambil dari data diatas dengan posisi di ganti : (20733.33, 25)

C1 = tingkat penjualan obat Laris

C2 = tingkat penjualan obat Kurang Laris

2. Perhitungan jarak pusat Cluster

Agar bisa mengukur jarak antara data dengan pusat Cluster digunakan *Euclidean distance*. Dari 15 data yang dijadikan sampel telah dipilih pusat awal Cluster yaitu **C1 (25, 20733)** dan **C2 (20733, 25)**. Lalu dilakukan perhitungan jarak dari sisa sampel data dengan pusat Cluster yang dimisalkan dengan **Y(a,b)**, di mana **a** merupakan jumlah penjualan yang terjual dan **b** adalah harga obat.

1. Y1 (30, 8000)
2. Y2 (80, 10000)
3. Y3 (20, 8000)
4. Y4 (3, 10000)
5. Y5 (6, 8000)
6. Y6 (4, 30000)
7. Y7 (90, 16000)
8. Y8 (10, 29000)
9. Y9 (15, 24000)
10. Y10 (2, 8000)
11. Y11 (10,5000)
12. Y12 (25,30000)
13. Y13 (27,6000)
14. Y14 (16,27000)
15. Y15 (37,20000)

Langkah selanjutnya adalah hitung nilai *Euclidean distance* dari semua data setiap titik pusat pertama, pusat kedua, dan pusat ketiga, maka akan mendapatkan hasil sebagai berikut:

$$\begin{aligned}
 D11 &= \sqrt{(Y1x - C1x)^2 + (Y1y - C1y)^2} \\
 &= \sqrt{(30 - 25)^2 + (8000 - 20733)^2} \\
 &= 12733.33
 \end{aligned}$$

$$\begin{aligned}
 D12 &= \sqrt{(Y2x - C1x)^2 + (Y2y - C1y)^2} \\
 &= \sqrt{(80 - 25)^2 + (10000 - 20733)^2} \\
 &= 10733.47
 \end{aligned}$$

$$D13 = \sqrt{(Y3x - C1x)^2 + (Y3y - C1y)^2}$$



$$\begin{aligned}
 &= \sqrt{(20 - 25)^2 + (8000 - 20733)^2} \\
 &= 12733.33 \\
 \text{D21} &= \sqrt{(Y1x - C2x)^2 + (Y1y - C2y)^2} \\
 &= \sqrt{(30 - 20733)^2 + (8000 - 25)^2} \\
 &= 22186.23 \\
 \text{D22} &= \sqrt{(Y2x - C2x)^2 + (Y2y - C2y)^2} \\
 &= \sqrt{80 - 20733)^2 + (10000 - 25)^2} \\
 &= 22936.02 \\
 \text{D23} &= \sqrt{(Y3x - C2x)^2 + (Y3y - C2y)^2} \\
 &= \sqrt{(20 - 20733)^2 + (8000 - 25)^2} \\
 &= 22196.56
 \end{aligned}$$

Langkah selanjutnya adalah menghitung jarak terdekat, dengan cara membandingkan nilai yang terkecil dari setiap *Cluster*, maka dapat dihasilkan.

Tabel 2. Jarak Terpendek dari *Cluster* dan Hasil Iterasi 1

No	C1	C2	Cen 1	Cen 2
1	12733.33	22186.23	1	
2	10733.47	22936.02	1	
3	12733.33	22195.56	1	
4	10733.36	23005.38	1	
5	12733.35	22208.62	1	
6	9266.69	3644.56		1
7	4733.78	2610.64		1
8	8266.68	3562.13		1
9	3266.68	31686.75	1	
10	12733.35	22212.36	1	
11	15733.34	21312.14	1	
12	9266.67	3643.62		1
13	14733.33	21551.17	1	
14	6266.67	3401.62		1
15	733.43	287.50		1

Kesimpulan dari hasil perhitungan dan tabel di atas adalah **Y1, Y2, Y3, Y4, Y5, Y9, Y10, Y11, Y13** adalah anggota *Cluster* 1, **Y6, Y7, Y8, Y12, Y14, Y15** adalah anggota *Cluster* 2.



3. Proses iterasi ke 2

1. Hitung titik pusat baru

Tentukan posisi *centroid* baru (C_k) dengan cara menghitung nilai rata-rata dari data-data yang ada pada *centroid* yang sama, maka hasilnya adalah:

$$\begin{aligned} C1 &= (30 + 80 + 20 + 3 + 6 + 15 + 2 + 10 + 27) / 9 = 21.44 \\ &= (80000 + 10000 + 8000 + 10000 + 8000 + 24000 + 8000 + 5000 + 6000) / 9 \\ &= 17666.67 \end{aligned}$$

$$\begin{aligned} C2 &= (4 + 90 + 10 + 25 + 16 + 37) / 6 = 30.33 \\ &= (30000 + 16000 + 29000 + 30000 + 27000 + 30000) / 6 = 25333.33 \end{aligned}$$

Dari perhitungan di atas maka didapatkan hasil titik pusat atau *centroid* yang baru yaitu: C1 (21.44, 17666.67), C2 (30.33, 25333.33)

2. Perhitungan jarak pusat *Cluster*

Hitung *Euclidean distance* dari semua data ketitik pusat yang baru (C1, C2, C3) seperti yang telah dilakukan pada tahap 1.

$$\begin{aligned} D11 &= \sqrt{(Y1x - C1x)^2 + (Y1y - C1y)^2} \\ &= \sqrt{(30 - 21.44)^2 + (8000 - 17666.67)^2} \\ &= 9666.67 \end{aligned}$$

$$\begin{aligned} D12 &= \sqrt{(Y2x - C1x)^2 + (Y2y - C1y)^2} \\ &= \sqrt{(80 - 21.44)^2 + (10000 - 17666.67)^2} \\ &= 7666.89 \end{aligned}$$

$$\begin{aligned} D13 &= \sqrt{(Y3x - C1x)^2 + (Y3y - C1y)^2} \\ &= \sqrt{(20 - 21.44)^2 + (8000 - 17666.67)^2} \\ &= 9666.67 \end{aligned}$$

$$\begin{aligned} D21 &= \sqrt{(Y1x - C2x)^2 + (Y1y - C2y)^2} \\ &= \sqrt{(30 - 30.33)^2 + (8000 - 25333.33)^2} \\ &= 17333.33 \end{aligned}$$

$$\begin{aligned} D22 &= \sqrt{(Y2x - C2x)^2 + (Y2y - C2y)^2} \\ &= \sqrt{(80 - 30.33)^2 + (10000 - 25333.33)^2} \\ &= 15333.41 \end{aligned}$$

$$\begin{aligned} D23 &= \sqrt{(Y3x - C2x)^2 + (Y3y - C2y)^2} \\ &= \sqrt{(20 - 30.33)^2 + (8000 - 25333.33)^2} \\ &= 17333.34 \end{aligned}$$

Langkah selanjutnya adalah menghitung jarak terpendek, dengan cara membandingkan nilai yang terkecil dari setiap *Cluster*, maka dapat dihasilkan:

**Tabel 3. Jarak Terpendek dari Cluster dan Hasil Iterasi 2**

No	C1	C2	Cen 1	Cen 2
1	9666.67	1733.33		1
2	7666.89	1533.41		1
3	9666.67	1733.33		1
4	7666.69	1533.35		1
5	9666.68	1733.35		1
6	1233.34	4666.74	1	
7	1668.08	9333.52	1	
8	1133.34	3666.73	1	
9	6333.33	1333.42		1
10	9666.69	1733.35		1
11	12666.68	2033.34	1	
12	12333.33	4666.67	1	
13	11666.67	1933.33	1	
14	9333.33	1666.73		1
15	2333.38	5333.33	1	

Kesimpulan dari hasil perhitungan dan tabel di atas adalah **Y6, Y7, Y8, Y11, Y12, Y13, Y15** adalah anggota *Cluster 1*, **Y1, Y2, Y3, Y4, Y5, Y9, Y10, Y14** adalah anggota *Cluster 2*. Dikarenakan hasil dari iterasi 1 dan iterasi 2 dengan posisi *Cluster* pada iterasi 1 berbeda maka perhitungan dilanjutkan ke iterasi ke 3.

4. Proses iterasi ke 3

1. Hitung titik pusat baru

Tentukan posisi *centroid* baru (C_k) dengan cara menghitung nilai rata-rata dari data-data yang ada pada *centroid* yang sama, maka hasilnya adalah:

$$\begin{aligned} C1 &= (4 + 90 + 10 + 10 + 25 + 27 + 37) / 7 = 29 \\ &= (30000 + 16000 + 29000 + 5000 + 30000 + 6000 + 20000) / 7 \\ &= 19428.57 \end{aligned}$$

$$\begin{aligned} C2 &= (30 + 80 + 20 + 3 + 6 + 15 + 2 + 16) / 8 = 21.50 \\ &= (80000 + 10000 + 8000 + 10000 + 8000 + 24000 + 8000 + 27000) / 8 \\ &= 21875 \end{aligned}$$

Dari perhitungan di atas maka didapatkan hasil titik pusat atau *centroid* yang baru yaitu: C1 (29, 19428.57), C2 (21.50, 21875)

2. Perhitungan jarak pusat Cluster

Hitung *Euclidean distance* dari semua data ke titik pusat yang baru (C1, C2, C3) seperti yang telah dilakukan pada tahap 1.



$$\begin{aligned}
 D11 &= \sqrt{(Y1x - C1x)^2 + (Y1y - C1y)^2} \\
 &= \sqrt{(30 - 29)^2 + (8000 - 19428.57)^2} \\
 &= 11428.57 \\
 D12 &= \sqrt{(Y2x - C1x)^2 + (Y2y - C1y)^2} \\
 &= \sqrt{(80 - 29)^2 + (10000 - 19428.57)^2} \\
 &= 7666.89 \\
 D13 &= \sqrt{(Y3x - C1x)^2 + (Y3y - C1y)^2} \\
 &= \sqrt{(20 - 29)^2 + (8000 - 19428.57)^2} \\
 &= 9666.67 \\
 D21 &= \sqrt{(Y1x - C2x)^2 + (Y1y - C2y)^2} \\
 &= \sqrt{(30 - 21.50)^2 + (8000 - 21875)^2} \\
 &= 13875.00 \\
 D22 &= \sqrt{(Y2x - C2x)^2 + (Y2y - C2y)^2} \\
 &= \sqrt{(80 - 21.50)^2 + (10000 - 21875)^2} \\
 &= 15333.41 \\
 D23 &= \sqrt{(Y3x - C2x)^2 + (Y3y - C2y)^2} \\
 &= \sqrt{(20 - 21.50)^2 + (8000 - 21875)^2} \\
 &= 17333.41
 \end{aligned}$$

Langkah selanjutnya adalah menghitung jarak terpendek, dengan cara membandingkan nilai yang terkecil dari setiap *Cluster*, maka dapat dihasilkan:

Tabel 4. Jarak Terpendek dari *Cluster* dan Hasil Iterasi 3

No	C1	C2	Cen 1	Cen 2
1	11428.57	1387.00		1
2	9428.71	1187.14		1
3	11428.57	13875.00	1	
4	9428.61	1187.01		1
5	11428.59	13875.01	1	
6	1057.46	8125.02	1	
7	3429.11	5875.40	1	
8	9571.45	7125.01		1
9	4571.45	2125.01		1
10	11428.60	13875.01	1	
11	14428.58	16875.00	1	
12	1057.43	8125.00	1	
13	1342.57	15875.00	1	
14	7571.44	5125.00		1
15	571.49	187.06		1



Kesimpulan dari hasil perhitungan dan tabel di atas adalah **Y3, Y5, Y6, Y7, Y10, Y11, Y12, Y13** adalah anggota *Cluster* 1, **Y1, Y2, Y4, Y8, Y9, Y14, Y15** adalah anggota *Cluster* 2. Dikarenakan hasil dari iterasi 1, 2 dan iterasi 3 dengan posisi *Cluster* pada iterasi 1 berbeda maka perhitungan dilanjutkan ke iterasi ke 4.

5. Proses iterasi ke 4

1. Hitung titik pusat baru

Tentukan posisi *centroid* baru (C_k) dengan cara menghitung nilai rata-rata dari data-data yang ada pada *centroid* yang sama, maka hasilnya adalah:

$$\begin{aligned} C1 &= (20 + 6 + 4 + 90 + 2 + 10 + 25 + 27) / 8 = 23 \\ &= (8000 + 8000 + 30000 + 16000 + 8000 + 5000 + 30000 + 6000) / 8 \\ &= 13875 \end{aligned}$$

$$\begin{aligned} C2 &= (30 + 80 + 3 + 10 + 15 + 16 + 37) / 7 = 27.29 \\ &= (80000 + 10000 + 10000 + 29000 + 24000 + 27000 + 20000) / 7 \\ &= 28571.43 \end{aligned}$$

Dari perhitungan di atas maka didapatkan hasil titik pusat atau *centroid* yang baru yaitu: $C1 (23, 13875)$, $C2 (27.29, 28571.43)$

2. Perhitungan jarak pusat *Cluster*

Hitung *Euclidean distance* dari semua data ke titik pusat yang baru ($C1, C2, C3$) seperti yang telah dilakukan pada tahap 1.

$$\begin{aligned} D11 &= \sqrt{(Y1x - C1x)^2 + (Y1y - C1y)^2} \\ &= \sqrt{(30 - 23)^2 + (8000 - 13875)^2} \\ &= 11428.57 \end{aligned}$$

$$\begin{aligned} D12 &= \sqrt{(Y2x - C1x)^2 + (Y2y - C1y)^2} \\ &= \sqrt{(80 - 23)^2 + (10000 - 13875)^2} \\ &= 3875.42 \end{aligned}$$

$$\begin{aligned} D13 &= \sqrt{(Y3x - C1x)^2 + (Y3y - C1y)^2} \\ &= \sqrt{(20 - 23)^2 + (8000 - 13875)^2} \\ &= 5875 \end{aligned}$$

$$\begin{aligned} D21 &= \sqrt{(Y1x - C2x)^2 + (Y1y - C2y)^2} \\ &= \sqrt{(30 - 27.29)^2 + (8000 - 28571)^2} \\ &= 13875.00 \end{aligned}$$

$$\begin{aligned} D22 &= \sqrt{(Y2x - C2x)^2 + (Y2y - C2y)^2} \\ &= \sqrt{(80 - 27.29)^2 + (10000 - 28571)^2} \\ &= 18571.50 \end{aligned}$$

$$D23 = \sqrt{(Y3x - C2x)^2 + (Y3y - C2y)^2}$$



$$= \sqrt{(20 - 27.29)^2 + (8000 - 28571)^2}$$

$$= 20571.43$$

Langkah selanjutnya adalah menghitung jarak terpendek, dengan cara membandingkan nilai yang terkecil dari setiap *Cluster*, maka dapat dihasilkan:

Tabel 5. Jarak Terpendek dari *Cluster* dan Hasil Iterasi 4

No	C1	C2	Cen 1	Cen 2
1	11428.57	1387.00		1
2	9428.71	1187.14		1
3	11428.57	13875.00		1
4	9428.61	11875.01		1
5	11428.59	13875.01		1
6	10571.46	8125.02		1
7	3429.11	5875.40		1
8	9571.45	7125.01	1	
9	4571.45	2125.01	1	
10	11428.60	13875.01		1
11	14428.58	16875.00		1
12	10571.43	8125.00		1
13	13428.57	15875.00		1
14	7571.44	5125.00		1
15	571.49	1875.06	1	

Kesimpulan dari hasil perhitungan dan tabel di atas adalah **Y8, Y9, Y15**, adalah anggota *Cluster* 1, **Y1, Y2, Y3, Y4, Y5, Y6, Y7, Y10, Y12, Y13, Y14** adalah anggota *Cluster* 2. Dikarenakan hasil dari iterasi 1, 2, 3 dan iterasi 4 dengan posisi *Cluster* pada iterasi 1 berbeda maka perhitungan dilanjutkan ke iterasi ke 5.

6. Proses iterasi ke 5

1. Hitung titik pusat baru

Tentukan posisi *centroid* baru (C_k) dengan cara menghitung nilai rata-rata dari data-data yang ada pada *centroid* yang sama, maka hasilnya adalah:

$$C1 = (10 + 15 + 37) / 3 = 20.67$$

$$= (29000 + 24000 + 20000) / 3 = 24333.33$$

$$C2 = (30 + 80 + 20 + 3 + 6 + 4 + 90 + 2 + 10 + 25 + 27 + 16) / 12 = 26.08$$

$$= (80000 + 10000 + 8000 + 10000 + 8000 + 30000 + 16000 + 8000 + 5000 + 30000 + 6000 + 27000) / 12$$

$$= 19833.33$$

Dari perhitungan di atas maka didapatkan hasil titik pusat atau *centroid* yang baru yaitu: C1 (20, 24333.33), C2 (26.08, 19833.33)

2. Perhitungan jarak pusat *Cluster*



Hitung *Euclidean distance* dari semua data ke titik pusat yang baru (C1, C2, C3) seperti yang telah dilakukan pada tahap 1.

$$\begin{aligned} D11 &= \sqrt{(Y1x - C1x)^2 + (Y1y - C1y)^2} \\ &= \sqrt{(30 - 20)^2 + (8000 - 24333.33)^2} \\ &= 55666.67 \end{aligned}$$

$$\begin{aligned} D12 &= \sqrt{(Y2x - C1x)^2 + (Y2y - C1y)^2} \\ &= \sqrt{(80 - 20)^2 + (10000 - 24333.33)^2} \\ &= 14333.46 \end{aligned}$$

$$\begin{aligned} D13 &= \sqrt{(Y3x - C1x)^2 + (Y3y - C1y)^2} \\ &= \sqrt{(20 - 20)^2 + (8000 - 24333.33)^2} \\ &= 16333.33 \end{aligned}$$

$$\begin{aligned} D21 &= \sqrt{(Y1x - C2x)^2 + (Y1y - C2y)^2} \\ &= \sqrt{(30 - 26.08)^2 + (8000 - 19833.33)^2} \\ &= 60166.67 \end{aligned}$$

$$\begin{aligned} D22 &= \sqrt{(Y2x - C2x)^2 + (Y2y - C2y)^2} \\ &= \sqrt{(80 - 26.08)^2 + (10000 - 19833.33)^2} \\ &= 9833.48 \end{aligned}$$

$$\begin{aligned} D23 &= \sqrt{(Y3x - C2x)^2 + (Y3y - C2y)^2} \\ &= \sqrt{(20 - 26.08)^2 + (8000 - 19833.33)^2} \\ &= 11833.33 \end{aligned}$$

Langkah selanjutnya adalah menghitung jarak terdekat, dengan cara membandingkan nilai yang terkecil dari setiap *Cluster*, maka dapat dihasilkan:

Tabel 6. Jarak Terpendek dari *Cluster* dan Hasil Iterasi 5

No	C1	C2	Cen 1	Cen 2
1	11428.57	13875.00	1	
2	9428.71	11875.14	1	
3	11428.57	1387.00		1
4	9428.61	11875.01	1	
5	11428.59	1387.01		1
6	10571.46	8125.02		1
7	3429.11	5875.40		1
8	9571.45	7125.01	1	
9	4571.45	2125.01	1	
10	11428.60	13875.01		1
11	14428.58	1687.00		1

12	10571.43	8125.00		1
13	13428.57	15875.00		1
14	7571.44	5125.00	1	
15	571.49	1875.06		1

Kesimpulan dari hasil perhitungan dan tabel di atas adalah **Y1, Y2, Y4, Y8, Y9, Y14** adalah anggota *Cluster* 1, **Y3, Y5, Y6, Y7, Y10, Y11, Y12, Y13, Y15** adalah anggota *Cluster*

2. Dikarenakan hasil dari iterasi 5 dan iterasi 1 dengan posisi *Cluster* pada iterasi 1 sama/tidak berubah maka proses pencarian dan perhitungan iterasi dihentikan.

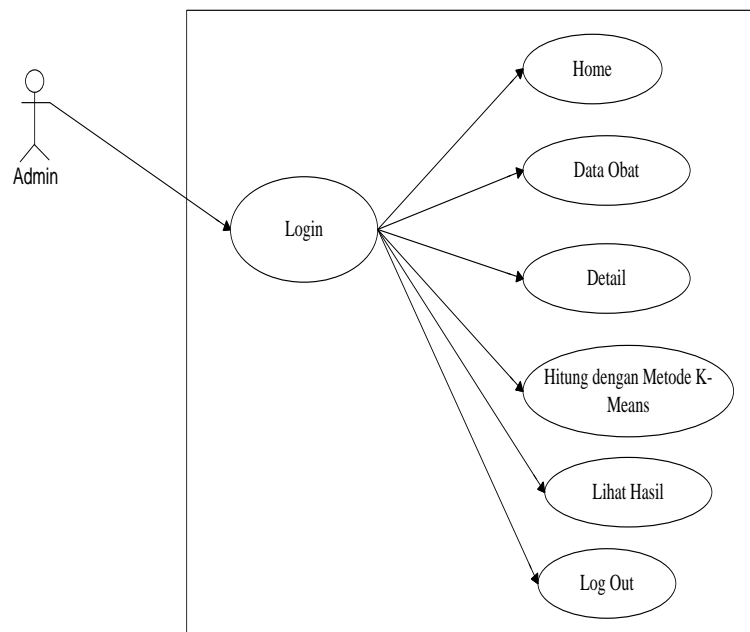
Berdasarkan perhitungan dari iterasi 1 hingga iterasi 5 dengan atribut total penjualan dan harga maka dapat digolongkan jenis obat yang Laris adalah **Remco Cought, Anacomidine, Unibebe Cought Syrup, Proxion, Tempra Sirup** dan **Batugin**. Golongan jenis obat yang Kurang Laris adalah **Pacdin Cought, Vipcol Sirup, Snow Playe, Termorex, Afibremol, Paracetamol Sirup, Pamtusin** dan **Dyrexin**.

3.2 Perancangan Sistem

Berikut gambaran umum sistem klasterisasi pola penjualan obat menggunakan metode *K-Means Clustering* (Studi Kasus di Apotek Pelangi Sei. Jering Teluk Kuantan) yang dirancang untuk mempermudah *user* dalam mengetahui bagaimana suatu sistem berjalan dan nantinya berguna juga untuk mengevaluasi sistem kerja yang dilaksanakan. Pada perancangan sistem ini digambarkan bagaimana model sistem yang akan digunakan untuk mengetahui klusterisasi pola penjualan obat.

3.3 Use Case Diagram

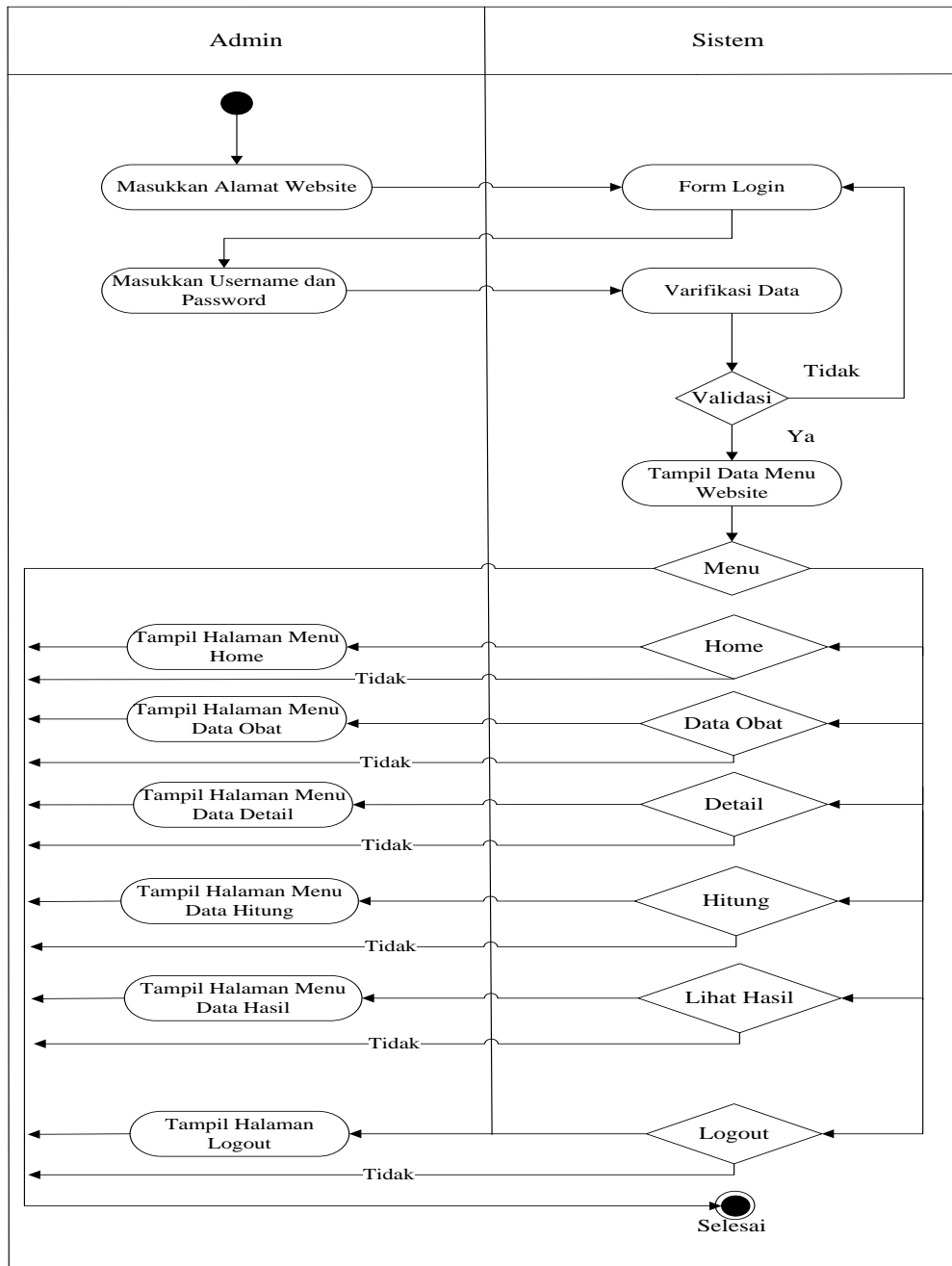
Use case diagram digunakan untuk mendefinisikan akses yang dapat dilakukan oleh aktor. *Use case diagram* pada sistem Klasterisasi Pola Penjualan Obat Menggunakan Metode *K-Means Clustering* (Studi Kasus di Apotek Pelangi Sei. Jering Teluk Kuantan) adalah sebagai berikut.



Gambar 2. Use Case Diagram

3.4 Activity Diagram

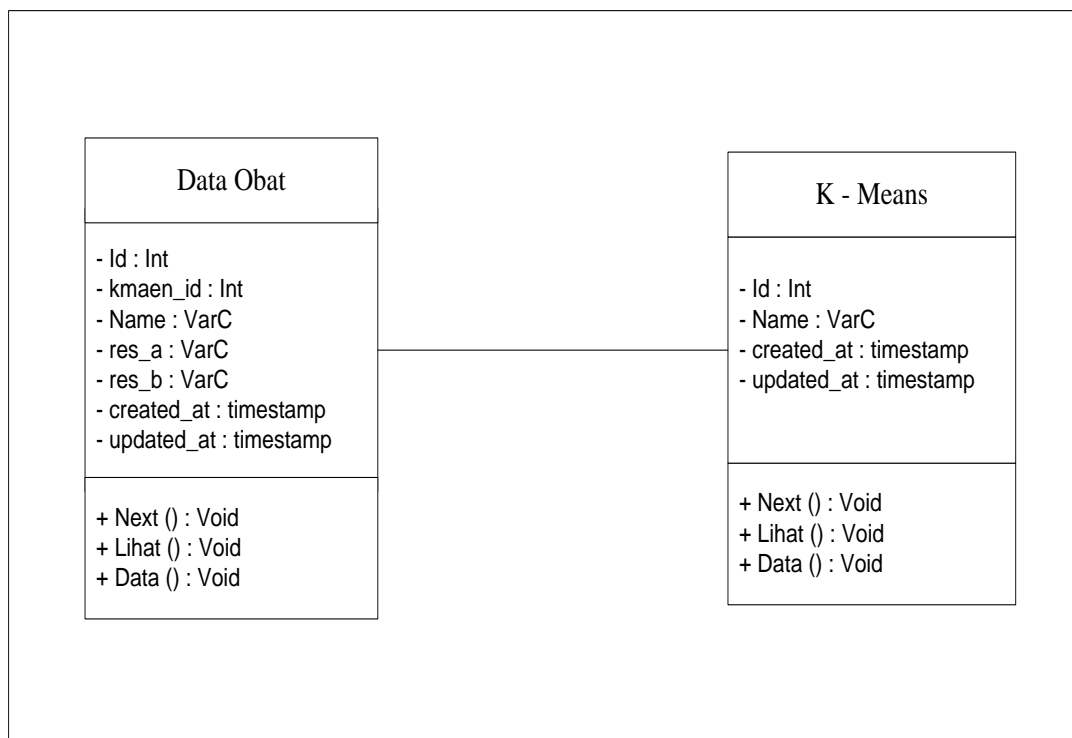
Activity diagram yang merupakan alat aktifitas sistem yang sedang dirancang, bagaimana masing-masing alir berawal dan bagaimana alir berakhir. Berikut akan digambarkan *activity diagram* sistem klusterisasi pola penjualan obat menggunakan metode *K-Means Clustering* (Studi Kasus di Apotek Pelangi Sei. Jering Teluk Kuantan).



Gambar 3. Activity Diagram

3.5 Class Diagram

Class diagram digunakan untuk merancang data base yang digunakan dan untuk menghubungkan antar tabel yang mempunyai relasi. Untuk lebih jelasnya dapat dilihat pada *class diagram* sebagai berikut.

**Gambar 4. Class Diagram**

4 PENUTUP

4.1 Kesimpulan

Setelah melakukan penelitian dan merancang sistem clusterisasi pola penjualan obat menggunakan metode *K-Means Clustering* (Studi Kasus di Apotek Pelangi Sei. Jering Teluk Kuantan), maka penulis dapat mengemukakan beberapa kesimpulan sebagai berikut.

1. Metode *K-Means clustering* dapat diterapkan pada pola penjualan obat pada Apotek Pelangi Sei. Jering, sehingga metode ini sangat membantu dalam mengelompokkan pola penjualan selama satu musim.
2. Dengan metode *K-Means clustering* maka dapat mengelompokkan data penjualan produk yang laku dijual dan yang tidak laku dijual pada apotek.
3. Hasil akhir dari 15 *items* penjualan sebagai data contoh pada Bulan Juli – Desember 2017 adalah penjualan Obat Laris sebanyak 6 *items* dan Penjualan Obat Kurang Laris sebanyak 9 *items*.
4. Dapat menghasilkan aplikasi sistem *Data Mining* metode *K-Means clustering* dalam pola penjualan yaitu dapat membandingkan hasil yang secara manual dengan aplikasi sistem yang dibangun

4.2. Saran

Berdasarkan kesimpulan di atas maka penulis mengemukakan beberapa saran yang diharapkan dapat bermanfaat bagi peneliti berikutnya dalam pengembangan kasus yang hampir sama dengan penelitian ini, sehingga hasilnya akan lebih efektif dari sistem yang telah dibangun ini. Berikut beberapa saran yang penulis kemukakan.

1. Diharapkan pada peneliti selanjutnya agar dapat menambah atribut yang lain sehingga hasil pengelompokan lebih jelas dan akurat.



2. Hasil pengujian dengan sistem yang telah dibangun menggunakan *website* agar dibuat perbandingan dengan sistem lain yaitu bisa dengan menggunakan rapid miner dan aplikasi lainnya.
3. Dari penelitian yang sudah dibangun ini penulis menyarankan agar dapat membandingkan dengan algoritma *Data Mining* lainnya seperti algoritma *Hierarchical clustering* dan *Partitional CLUSTERING*.

DAFTAR PUSTAKA

- Ahmad Ramadhani, Andi Farmadi Dan Irwan Budiman (2014), Clustering Data Cuaca Untuk Pengenalan Pola Perioditas Iklim Wilayah Pelayari Dengan Metode Fuzzy C-Means, Jurnal Teknologi Vol. 3, No. 1, ISSN 2087-6920
- Benri Melpa Metisen dan Herlina Latipa Sari. (2015). "*Analisis Clustering Menggunakan Metode K-Means Dalam Pengelompokan Penjualan Produk Pada Swalayan Fadhila*" Jurnal Media Infotama, Vol. 11, No. 2, ISSN 1858–2680
- David Hartanto Kamagi dan Seng Hansun (2014), Implementasi Data Mining dengan Algoritma C4.5 untuk Memprediksi Tingkat Kelulusan Mahasiswa, Jurnal Ultimatics, Vol. VI, No. 1, ISSN 2085-4552
- Elly Muningsih dan Sri Kiswati. (2015). "*Penerapan Metode K-Means Untuk Clustering Produk Online Shop Dalam Penentuan Stok Barang*" Jurnal Bianglala Informatika Vol 3 No 1 ISSN : 2338-9761.
- Fina Nasari dan Surya Darma (2015), Penerapan K-Means Clustering Pada Data Penerimaan Mahasiswa Baru (Studi Kasus : Universitas Potensi Utama), Jurnal Seminar Nasional Teknologi Informasi dan Multimedia, ISSN : 2302-3805
- Jasri, J., & Nazli, R. (2018). Penerapan Metode Mamdani Untuk Sistem Pendukung Keputusan Penentuan Golongan Obat Sesuai Dengan Penyakit Diabetes. Jurnal Teknologi Dan Open Source, 1(2), 67-74.
- Lianna Felicia. (2014). "*Penerapan Metode Clustering Dengan K-Means Untuk Memetakan Potensi Tanaman Padi Di Kota Semarang*" Jurnal Ilmu computer
- Syam, E. (2018). Rancang Bangun Sistem Informasi Manajemen Data Mahasiswa Dan Dosen Terintegrasi. IT Journal Research and Development, 2(2), 45-51.