

Penerapan Data Mining Dalam Analisa Data Pemakaian Obat Dengan Menerapkan Algoritma K-Means

Mei Pilida Tambunan

Teknik Informatika, Fakultas Ilmu Komputer & Teknologi Informasi, Universitas Budi Darma, Medan, Indonesia
Email: mei.tambunan97@gmail.com

Abstrak

Pelaksanaan pengobatan yang belum rasional selama ini telah memberikan dampak negatif terdiri dari pemborosan dana masyarakat, efek samping yang terdiri dari resistensi, interaksi obat yang berbahaya yang menurunkan mutu pengobatan dan mutu pelayanan kesehatan. Penelitian ini mengharapkan untuk menganalisis gambaran beroperasi umum, penggunaan obat pada pasien puskesmas teladan. Pada Puskesmas teladan medan dilakukan penentuan pemakaian obat setiap dua tahun sekali. Jumlah obat yang dipakai di puskesmas teladan medan. Masalah yang sering dihadapi adalah sulitnya Tata Usaha Puskesmas Teladan Medan dalam menentukan pemakaian obat karena data yang tersimpan dalam Microsoft Excel masih kurang sehingga memerlukan waktu yang lama apabila sewaktu-waktu ruang tata usaha puskesmas lain. Untuk itu diperlukan suatu sistem yang dapat mempermudah pihak tata usaha dalam menganalisa data pemakaian obat. Algoritma ini merupakan kelompok metode penganalisaan data atau metode data mining yang melakukan proses pemodelan tanpa supervisi (Unsupervised) dan merupakan salah satu metode yang melakukan pengelompokan data dengan sistem partisi. Terdapat dua jenis data Clusterasi yang sering digunakan dalam proses pengelompokan data yaitu hierarhical dan Non-Hierarchical, dan K-Means merupakan salah satu metode data Clusterasi Hierarhical atau Partitional Clustering.

Kata Kunci: *Pemakaian Obat, Puskesmas Teladan Medan, Algoritma K- Means*

1. PENDAHULUAN

Obat merupakan salah satu komponen yang penting dalam hal kesehatan baik untuk mencegah, mengurangi, menghilangkan atau menyembuhkan suatu penyakit atau gejala penyakit. Untuk itu obat perlu dikelola dengan baik, efektif dan efisien. Perencanaan kebutuhan obat merupakan hal yang penting dilakukan untuk menjamin ketersediaan dan pemerataan obat dengan jenis dan jumlah yang mencukupi agar obat dapat diperoleh dengan cepat pada tempat dan waktu yang tepat pada instansi yang terkait dengan pelayanan kesehatan puskesmas. Perencanaan kebutuhan obat akan memengaruhi pengadaan pendistribusian dan pemakaian obat di tempat pelayanan kesehatan.

Clusterasi kebutuhan obat diharapkan dapat menjadi salah satu sumber pengambilan keputusan untuk menjamin ketersediaan obat pada Dinas Kesehatan Puskesmas Teladan Medan. Clustering Data merupakan salah satu metode dalam data mining yang dapat digunakan untuk memetakan data kedalam kelompok yang lebih kecil berdasarkan kesamaan karakteristik yang dimilikinya. Dengan hasil clustering ini distribusi obat pada instansi layanan kesehatan dapat dikelompokkan sesuai kebutuhan berdasarkan data distribusi obat pada tahun sebelumnya dan dapat digunakan sebagai acuan perencanaan obat untuk tahun berikutnya. Dengan begitu diharapkan ketersediaan obat untuk tahun berikutnya dapat lebih terjamin dan dapat memenuhi permintaan obat dan instansi layanan kesehatan.

Pada penelitian ini digunakan data yang berasal dari laporan pemakaian obat dan lembar permintaan obat (LPLPO) Puskesmas teladan Medan. berdasarkan dari uraian diatas, maka penelitian yang akan dilakukan bertujuan untuk mengetahui clusterasi data obat dengan teknik data mining dimana pengolahan data obat menggunakan algoritma k-means. Batasan masalah dari penelitian yang dilakukan adalah data yang akan digunakan yaitu data obat dari Puskesmas Teladan Medan. Analisis digunakan dengan menggunakan Algoritma K-Means. Output yang dihasilkan adalah kelompok obat yang

pemakaian rendah, sedang dan tinggi. Tujuan penelitian ini untuk mengelompokkan data obat pada Puskesmas Teladan Medan berdasarkan pemakaian obat yang rendah, sedang dan tinggi. Hasil penelitian ini bermanfaat bagi bagian yang terkait dengan perencanaan kebutuhan obat pada Puskesmas Teladan Medan yang membutuhkan data jenis-jenis kelompok obat-obatan mana sajakah yang pemakaiannya rendah, sedang dan tinggi, setiap tahunnya, sehingga dapat digunakan sebagai acuan perencanaan kebutuhan obat ke depannya.

Penelitian sebelumnya pada tahun 2017 yang dilakukan oleh Fitri Yunita, yang melakukan penelitian penerimaan mahasiswa baru di Universitas Islam Indragiri dengan algoritma yang sama. Penelitian ini menghasilkan data mahasiswa yang sangat berlimpah berupa data profil mahasiswa dan data lainnya. fitri yunita.

Penelitian yang sama juga dilakukan oleh Gustientiedina, M. Hasmil Adiya, Yenny Desnelita pada tahun 2018 dalam jurnal ilmiahnya yang berjudul penerapan Algoritma K-Means untuk clustering Data obat-obatan pada RSUD Ppekanbaru, yang melakukan penelitian menentukan data obat dengan algoritma K-Means.

Alasan penggunaan algoritma K-Means diantaranya karena algoritma ini memiliki ketelitian yang cukup tinggi terhadap ukuran objek, sehingga algoritma ini relatif lebih terukur dan efisien untuk pengolahan objek dalam jumlah besar. Selain itu algoritma K-Means ini tidak terpengaruh oleh urutan objek. Untuk mendapatkan informasi yang berupa analisa melalui proses data mining maka dalam penelitian ini menggunakan proses algoritma k-means dimana dalam proses ini menggunakan tahapan data cleaning, data integration, data selection, data transformation, data mining, evaluation dan presentation dan dalam proses data mining ini nantinya akan menggunakan Software weka. Selain itu software Weka juga memiliki kemampuan untuk mengelola data dengan metode Clustering (pengelompokkan) yang dipakai dalam penelitian ini.

2. TEORITIS

2.1 Data Mining

Data mining adalah proses yang memperkerjakan satu atau lebih teknik pembelajaran komputer (machine learning) untuk menganalisis dan mengekstraksi pengetahuan komputer (knowledge) secara otomatis. Definisi lain diantaranya adalah pembelajaran berbasis induksi (induction-based learning) adalah proses pembentukan definisi-definisi konsep umum yang dilakukan dengan cara mengobsevasi contoh-contoh spesifik dari konsep konsep yang akan dipelajari. Knowledge Discovery in Database (KDD) adalah penerapan metode saintifik pada data mining. Dalam konteks ini data mining merupakan satu langkah dari proses KDD[1].

Data mining adalah bidang yang sepenuhnya menggunakan apa yang dihasilkan oleh data warehouse; bersama dengan bidang yang menangani masalah pelaporan dan manajemen data. Sementara, data warehouse sendiri bertugas untuk menarik/meng-quary data dari basis data mentah untuk memberikan hasil data yang nantinya digunakan oleh bidang yang menangani manajemen, pelaporan dan data mining[1].

2.2 Clustering

Salah satu teknik yang dikenal dalam data mining yaitu clustering. Pengertian clustering keilmuan dalam data mining adalah pengelompokan sejumlah data atau objek ke dalam cluster (group) sehingga setiap dalam cluster tersebut akan berisi data yang semirip mungkin dan berbeda dengan objek dalam cluster yang lainnya. Sampai saat ini, para ilmuwan masih terus melakukan berbagai usaha untuk melakukan perbaikan model cluster dan menghitung jumlah cluster yang optimal sehingga dapat dihasilkan cluster yang paling baik. Ada dua metode clustering yang kita kenal, yaitu hierarchical clustering dan partitioning. Metode hierarchical clustering sendiri terdiri dari complete linkage clustering, single linkage clustering, average linkage clustering dan centroid linkage clustering. Sedangkan metode partitioning sendiri terdiri dari k-means dan fuzzy k-means[6].

Pengelompokan (clustering) merupakan bagian dari ilmu data mining yang bersifat tanpa arahan (unsupervised). Clustering adalah proses pembagian data ke dalam kelas atau cluster berdasarkan tingkat kesamaannya. Dalam clustering, data yang memiliki kesamaan dimasukkan ke dalam cluster yang sama, sedangkan data yang tidak memiliki kesamaan dimasukkan dalam cluster yang berbeda[7].

2.3 Algoritma K-Means

K-Means Clustering adalah, K dimaksudkan sebagai konstanta jumlah cluster yang diinginkan, Means dalam hal ini berarti nilai suatu rata-rata dari suatu grup data yang dalam hal ini didefinisikan sebagai cluster, sehingga K-Means Clustering adalah suatu metode penganalisaan data atau metode data mining yang melakukan proses pemodelan tanpa supervisi (unsupervised) dan merupakan salah satu metode yang melakukan pengelompokan data dengan sistem partisi. Metode K-Means berusaha mengelompokkan data yang ada kedalam beberapa

kelompok, dimana data dalam satu kelompok mempunyai karakteristik yang sama satu sama lainnya dan mempunyai karakteristik yang berbeda dengan data yang ada didalam kelompok yang lain. Algoritma K-means merupakan algoritma yang membutuhkan parameter input sebanyak k dan membagi sekumpulan n objek kedalam k cluster sehingga tingkat kemiripan antar anggota dalam suatu cluster tinggi sedangkan tingkat kemiripan dengan anggota pada cluster lain sangat rendah. Kemiripan anggota terhadap cluster diukur dengan kedekatan objek terhadap nilai mean pada cluster atau dapat disebut sebagai centroid cluster atau pusat massa[7].

3. ANALISA

Analisa merupakan salah satu langkah awal dalam penyelesaian dan mengidentifikasi sebuah permasalahan yang terjadi. Analisa masalah memiliki peranan penting dalam proses analisis untuk mencapai dan memperoleh hasil yang akurat. Dinas kesehatan UPT Puskesmas Teladan Medan menganalisa data pemakaian obat dilakukan penggabungan data-data obat yang terdapat pada laporan setiap bulannya kedalam satu tahun menjadi total pemakaian obat. Administrasi Puskesmas Teladan Medan mengumpulkan data dalam bentuk kertas lewat atau lewat email dan kemudian melakukan penginputan data ke dalam Microsoft Excel 2007 dan flash disk sebagai media penyimpanannya.

Pada penelitian ini penulis menggunakan Algoritma K-Means Clustering untuk mengelompokkan jumlah berdasarkan yang terdiri dari 3 kelompok, yaitu kelompok jumlah kuantum, rendah, sedang dan tinggi. Langkah-langkah pengelompokan data dengan Algoritma K-Means adalah sebagai berikut.

- Menentukan jumlah kelompok.
- Menentukan nilai titik pusat cluster secara acak pada data awal.
- Menghitung data pemakaian obat ke centroid dengan menggunakan rumus Euclidean Distance.
- Mengelompokkan data dengan memasukkan setiap objek ke dalam cluster (grup) berdasarkan jarak minimumnya,
- 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.

Adapun data yang digunakan dalam penelitian ini merupakan data pilihan yang diberikan oleh pihak Tata Usaha Puskesmas Teladan. Data yang digunakan adalah data dua tahun terakhir (2017 dan 2018) untuk membandingkan Pemakaian obat. Tahap analisa dan pembahasan ini, algoritma yang dipakai dalam menganalisa data pemakaian obat menggunakan algoritma K-Means Clustering dan Weka sebagai alat bantu penerapan data mining, terutama membantu menyelesaikan masalah klasifikasi.

Setelah data pemakaian obat yang telah diinisialisasikan ke dalam bentuk angka kemudian dikelompokkan menggunakan algoritma K-Means Clustering. Langkah-langkah pengelompokan data

pemakaian obat menggunakan algoritma K-Means Clustering:

- 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.
- Menentukan titik pusat awal cluster.
Titik pusat awal cluster diambil berdasarkan jumlah pemakaian obat paling rendah, sedang dan tinggi.

Tabel 1. Titik Pusat Awal Setiap Cluster

Nama Obat	Jan- Mei 2017	Jun- Okt 2017	Nov- Mar2018	Apr- Agu2018	Sep- Des2018
Ampisilina kaplet 500 ml	29059	2857	2082	0	23
Anti Hemmoroid suppositoria	1937	302	1677	372	320
Asetosal tablet 500 mg	16	0	0	0	0

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

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

Di mana:

$D(i,j)$: 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

Iterasi I

Tentukan K jumlah pusat cluster secara acak (random) pada percobaan pertama ini akan di tentukan 3 data secara acak sebagai titik pusat awal untuk perhitungan jarak dari seluruh kelompok cluster yang akan dibentuk.

Jumlah cluster = 3 (sedikit, sedang, tinggi)

Jumlah data = 20

Jumlah atribut = 1

Data acak tersebut akan dapat dilihat pada tabel 2.

Tabel 2. Data Pusat *Cluster* Iterasi 1

Nama Obat	Jan- Mei 2017	Jun- Okt 2017	Nov- Mar20	Apr- Agu20	Sep- Des20	Clust er
Ampisilin a kaplet 500 ml	2905 9	285 7	2082	0	23	C1
Anti Hemmoroid id supposito ria	1937	302	1677	372	320	C2
Asetosal tablet 500 mg	16	0	0	0	0	C3

Dari tabel 2 dapat dilihat bahwa masing-masing data memiliki nilai pusat cluster yang terlihat pada tabel 3.

Hitung Euclidean Distance dari semua data ke tiap titik pusat pertama.

Hitung jarak tiap data dengan masing-masing cluster pusat dengan menggunakan persamaan (3) yaitu persamaan Euclidean Distance Hitung Euclidean Distance dari semua data ke tiap titik pusat pertama:

C1:

$$\begin{aligned}
 &= \sqrt{(a_n - c_{1a})^2 + (b_n - c_{1b})^2 + (c_n - c_{1c})^2 + (d_n - c_{1d})^2 + (e_n - c_{1e})^2} \\
 &= \sqrt{(0 - 29059)^2 + (328 - 2857)^2 + (4235 - 2082)^2 + (478 - 0)^2 + (0 - 23)^2} \\
 &= 29252,1 \\
 &= \sqrt{(0 - 29059)^2 + (3664 - 2857)^2 + (24333 - 2082)^2 + (4343 - 0)^2 + (0 - 23)^2} \\
 &= 36865,23 \\
 &= \sqrt{(0 - 29059)^2 + (10 - 2857)^2 + (160 - 2082)^2 + (30 - 0)^2 + (10 - 23)^2} \\
 &= 29261,34 \\
 &= \sqrt{(249 - 29059)^2 + (13 - 2857)^2 + (17 - 2082)^2 + (1346 - 0)^2 + (905 - 23)^2} \\
 &= 29054,8 \\
 &= \sqrt{(11 - 29059)^2 + (11046 - 2857)^2 + (150 - 2082)^2 + (0 - 0)^2 + (0 - 23)^2} \\
 &= 30242 \\
 &= \sqrt{(0 - 29059)^2 + (513 - 2857)^2 + (1013 - 2082)^2 + (233 - 0)^2 + (0 - 23)^2} \\
 &= 29173,91 \\
 &= \sqrt{(29059 - 29059)^2 + (2857 - 2857)^2 + (2082 - 2082)^2 + (0 - 0)^2 + (23 - 23)^2} \\
 &= 0 \\
 &= \sqrt{(0 - 29059)^2 + (589 - 2857)^2 + (45 - 2082)^2 + (164 - 0)^2 + (56 - 23)^2} \\
 &= 29218,93 \\
 &= \sqrt{(0 - 29059)^2 + (1500 - 2857)^2 + (23 - 2082)^2 + (112 - 0)^2 + (25 - 23)^2} \\
 &= 29163,66 \\
 &= \sqrt{(510 - 29059)^2 + (5394 - 2857)^2 + (2753 - 2082)^2 + (1400 - 0)^2 + (488 - 23)^2} \\
 &= 28703,53 \\
 &= \sqrt{(1937 - 29059)^2 + (302 - 2857)^2 + (1677 - 2082)^2 + (372 - 0)^2 + (320 - 23)^2} \\
 &= 27247,64 \\
 &= \sqrt{(0 - 29059)^2 + (12 - 2857)^2 + (70 - 2082)^2 + (1303 - 0)^2 + (851 - 23)^2} \\
 &= 29296,18 \\
 &= \sqrt{(0 - 29059)^2 + (9 - 2857)^2 + (21 - 2082)^2 + (263 - 0)^2 + (0 - 23)^2} \\
 &= 29272,06 \\
 &= \sqrt{(5000 - 29059)^2 + (0 - 2857)^2 + (0 - 2082)^2 + (7500 - 0)^2 + (0 - 23)^2} \\
 &= 25447,64 \\
 &= \sqrt{(51 - 29059)^2 + (148 - 2857)^2 + (85 - 2082)^2 + (1076 - 0)^2 + (1731 - 23)^2} \\
 &= 29222,43 \\
 &= \sqrt{(0 - 29059)^2 + (2262 - 2857)^2 + (3081 - 2082)^2 + (1022 - 0)^2 + (204 - 23)^2} \\
 &= 29100,21 \\
 &= \sqrt{(23 - 29059)^2 + (0 - 2857)^2 + (0 - 2082)^2 + (0 - 0)^2 + (1431 - 23)^2} \\
 &= 29250,43 \\
 &= \sqrt{(0 - 29059)^2 + (10625 - 2857)^2 + (59 - 2082)^2 + (0 - 0)^2 + (0 - 23)^2} \\
 &= 30147,3 \\
 &= \sqrt{(16 - 29059)^2 + (0 - 2857)^2 + (0 - 2082)^2 + (0 - 0)^2 + (0 - 23)^2} \\
 &= 29257,36 \\
 &= \sqrt{(0 - 29059)^2 + (0 - 2857)^2 + (0 - 2082)^2 + (0 - 0)^2 + (0 - 23)^2} \\
 &= 29273,24
 \end{aligned}$$

Hitung Euclidean Distance dari semua data ke tiap titik pusat kedua :

C2:

$$\begin{aligned}
 &= \sqrt{(a_n - c_{2a})^2 + (b_n - c_{2b})^2 + (c_n - c_{2c})^2 + (d_n - c_{2d})^2 + (e_n - c_{2e})^2} \\
 &= \sqrt{(0 - 1937)^2 + (328 - 302)^2 + (4235 - 1677)^2 + (478 - 372)^2 + (0 - 320)^2} \\
 &= 3210,44 \\
 &= \sqrt{(0 - 1937)^2 + (3664 - 302)^2 + (24333 - 1677)^2 + (4343 - 372)^2 + (0 - 320)^2} \\
 &= 23326,33 \\
 &= \sqrt{(0 - 1937)^2 + (10 - 302)^2 + (160 - 1677)^2 + (30 - 372)^2 + (10 - 320)^2} \\
 &= 2501,035
 \end{aligned}$$

$$\begin{aligned}
 &= \sqrt{(249-1937)^2 + (13-302)^2 + (17-1677)^2 + (1346-372)^2 + (905-320)^2} = 7943,146 \\
 &= 2576,378 \\
 &= \sqrt{(11-1937)^2 + (11046-302)^2 + (150-1677)^2 + (0-372)^2 + (0-320)^2} = 3160,248 \\
 &= 11027,82 \\
 &= \sqrt{(0-1937)^2 + (513-302)^2 + (1013-1677)^2 + (133-372)^2 + (0-320)^2} = 38,47077 \\
 &= 2063,101 \\
 &= \sqrt{(2905-1937)^2 + (2857-302)^2 + (2082-1677)^2 + (0-372)^2 + (23-320)^2} = 10625,18 \\
 &= 27247,62 \\
 &= \sqrt{(0-1937)^2 + (589-302)^2 + (45-1677)^2 + (164-372)^2 + (56-320)^2} = 16 \\
 &= 2557,491 \\
 &= \sqrt{(0-1937)^2 + (1500-302)^2 + (23-1677)^2 + (112-372)^2 + (25-320)^2} \\
 &= 2826,693 \\
 &= \sqrt{(510-1937)^2 + (5394-302)^2 + (2753-1677)^2 + (1400-372)^2 + (488-320)^2} \\
 &= 5493,589 \\
 &= \sqrt{(1937-1937)^2 + (302-302)^2 + (1677-1677)^2 + (372-372)^2 + (320-320)^2} = 0 \\
 &= \sqrt{(0-1937)^2 + (12-302)^2 + (70-1677)^2 + (1303-372)^2 + (851-320)^2} = 2699,224 \\
 &= 2699,224 \\
 &= \sqrt{(0-1937)^2 + (9-302)^2 + (21-1677)^2 + (263-372)^2 + (0-320)^2} = 2567,434 \\
 &= 2567,434 \\
 &= \sqrt{(5000-1937)^2 + (0-302)^2 + (0-1677)^2 + (7500-372)^2 + (0-320)^2} = 7943,146 \\
 &= 7943,146 \\
 &= \sqrt{(51-1937)^2 + (148-302)^2 + (85-1677)^2 + (1076-372)^2 + (1731-320)^2} = 2571,42 \\
 &= 2571,42 \\
 &= \sqrt{(0-1937)^2 + (2262-302)^2 + (3081-1677)^2 + (1022-372)^2 + (204-320)^2} = 3160,248 \\
 &= 3160,248 \\
 &= \sqrt{(23-1937)^2 + (0-302)^2 + (0-1677)^2 + (0-372)^2 + (1431-320)^2} = 38,47077 \\
 &= 38,47077 \\
 &= \sqrt{(0-1937)^2 + (10625-302)^2 + (59-1677)^2 + (0-372)^2 + (0-320)^2} = 10625,18 \\
 &= 10625,18 \\
 &= \sqrt{(0-1937)^2 + (16-302)^2 + (0-302)^2 + (0-1677)^2 + (0-372)^2 + (0-320)^2} = 0 \\
 &= 0 \\
 &= \sqrt{(0-1937)^2 + (0-302)^2 + (0-1677)^2 + (0-372)^2 + (0-320)^2} = 16 \\
 &= 16
 \end{aligned}$$

Tabel 4.5. Perhitungan Euclidean Distance Iterasi 1

No	Nama Obat	c1	c2	c3	jarak min	Clust er
1	Air Raksa Dental use	29252,1	3210,44	4274,523	3210,44	2
2	Aminofilina injeksi 24 mg/ml-10 ml	36865,23	23326,33	24987,63	23326,33	2
3	Aminofilina tablet 200 mg	29261,34	2501,035	163,9085	163,9085	3
4	Amitriplina hcl tablet salut 25 mg	29054,8	2576,378	1366,517	1366,517	3
5	Amoksisilina kapsul 250mg	30242	11027,82	11047,02	11027,82	2
6	Amoksisilina sirup kering 125 mg/5 ml	29173,91	2063,101	1159,26	1159,26	3
7	Ampisilina kaplet 500 ml	0	27247,62	29257,36	0	1
8	Ampisilina sirup kering 125 mg/5 ml	29218,93	2557,491	613,3139	613,3139	3
9	Antalgin (metampriron) tablet 500 mg	29163,66	2826,693	1504,445	1504,445	3
10	Antasida doen tablet kombinasi	28703,53	5493,589	6235,284	5493,589	2
11	Anti Hemmoroid suppositoria	27247,64	0	2594,702	0	2
12	Aqua pro injeksi steril bebas pirogen 20 ml	29296,18	2699,224	1305,358	1305,358	3
13	Aqudest steril 500 ml	29272,06	2567,434	264,475	264,475	3
14	Asam Askorbat (vit, c) tablet 50 mg	25447,64	7943,146	9005,013	7943,146	2
15	As, benzoat 3%	29222,43	2571,42	1090,807	1090,807	3
16	AS, Salisilat 6% salep	29100,21	3160,248	3956,532	3160,248	2
17	Asam, Salisilat 0,1 N	29250,43	2589,676	38,47077	38,47077	3
18	As, Salisilat 2% + belerang endap 4% salep	30147,3	10633,55	10625,18	10625,18	3
19	Asam sulfosalisilat 20%	29257,36	2594,579	0	0	3
20	Asetosal tablet 500 mg	29273,24	2606,477	16	16	3
	Atropina sulfat injeksi 0,25 mg/ml - 1 ml					

Hitung Euclidean Distance dari semua data ke tiap titik pusat ketiga :

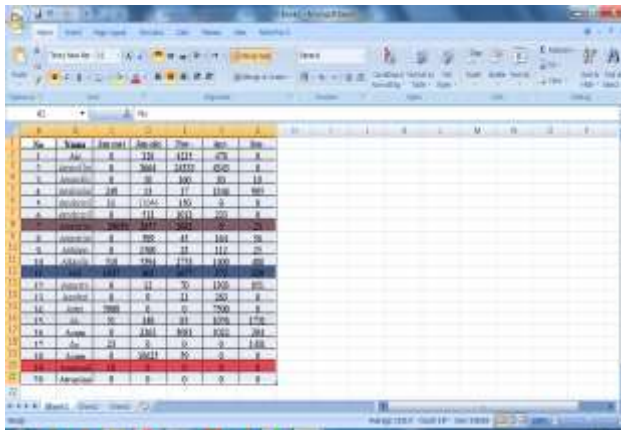
C3:

$$\begin{aligned}
 &= \sqrt{(a_n - c_{2a})^2 + (b_n - c_{2b})^2 + (c_n - c_{2c})^2 + (d_n - c_{2d})^2 + (e_n - c_{2e})^2} \\
 &= \sqrt{(0-16)^2 + (328-0)^2 + (4235-0)^2 + (478-0)^2 + (0-0)^2} = 3210,44 \\
 &= 3210,44 \\
 &= \sqrt{(0-16)^2 + (3664-0)^2 + (24333-0)^2 + (4343-0)^2 + (0-0)^2} = 23326,33 \\
 &= 23326,33 \\
 &= \sqrt{(0-16)^2 + (10-0)^2 + (160-0)^2 + (30-0)^2 + (10-0)^2} = 163,9085 \\
 &= 163,9085 \\
 &= \sqrt{(249-16)^2 + (13-0)^2 + (17-0)^2 + (1346-0)^2 + (905-0)^2} = 1366,517 \\
 &= 1366,517 \\
 &= \sqrt{(11-19316)^2 + (11046-0)^2 + (150-0)^2 + (0-0)^2 + (0-0)^2} = 11027,82 \\
 &= 11027,82 \\
 &= \sqrt{(0-16)^2 + (513-0)^2 + (1013-0)^2 + (133-0)^2 + (0-0)^2} = 1159,26 \\
 &= 1159,26 \\
 &= \sqrt{(2905-16)^2 + (2857-0)^2 + (2082-0)^2 + (0-0)^2 + (23-0)^2} = 0 \\
 &= 0 \\
 &= \sqrt{(0-16)^2 + (589-0)^2 + (45-0)^2 + (164-0)^2 + (56-0)^2} = 613,3139 \\
 &= 613,3139 \\
 &= \sqrt{(0-16)^2 + (1500-0)^2 + (23-0)^2 + (112-0)^2 + (25-0)^2} = 1504,445 \\
 &= 1504,445 \\
 &= \sqrt{(510-16)^2 + (5394-0)^2 + (2753-0)^2 + (1400-0)^2 + (488-0)^2} = 5493,589 \\
 &= 5493,589 \\
 &= \sqrt{(1937-16)^2 + (302-0)^2 + (1677-0)^2 + (372-0)^2 + (320-0)^2} = 0 \\
 &= 0 \\
 &= \sqrt{(0-16)^2 + (12-0)^2 + (70-0)^2 + (1303-0)^2 + (851-0)^2} = 1305,358 \\
 &= 1305,358 \\
 &= \sqrt{(0-16)^2 + (9-0)^2 + (21-0)^2 + (263-0)^2 + (0-0)^2} = 264,475 \\
 &= 264,475
 \end{aligned}$$

4. IMPLEMENTASI

Input data sangat berpengaruh di dalam pengisian data sehingga data yang dikehendaki dapat diterima oleh sistem dan juga diproses untuk pembuatan laporan yang dikehendaki. Adapun bentuk input tersebut adalah sebagai berikut.

Proses data mining secara aplikatif dimana proses data mining yang akan diuji dilakukan menggunakan software data mining Weka. Langkah yang dilakukan adalah input data kelompok obat, pemakaian rendah dan pemakaian sedang dan tinggi yang disimpan dengan type Microsoft Excel Comma Separated Values File (.csv). Tampilan datapemakaian obat tahun 2017 dan tahun 2018t, dapat dilihat pada gambar 1.



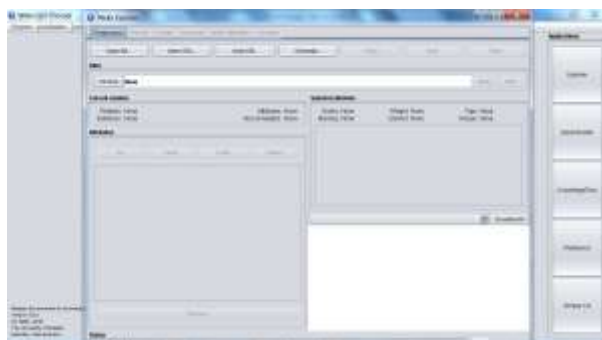
Tahun	Grup	Pemakaian	2017	2018
1	1	1	1	1
2	1	2	1	1
3	1	3	1	1
4	1	4	1	1
5	1	5	1	1
6	1	6	1	1
7	1	7	1	1
8	1	8	1	1
9	1	9	1	1
10	1	10	1	1
11	1	11	1	1
12	1	12	1	1
13	1	13	1	1
14	1	14	1	1
15	1	15	1	1
16	1	16	1	1
17	1	17	1	1
18	1	18	1	1
19	1	19	1	1
20	1	20	1	1
21	1	21	1	1
22	1	22	1	1
23	1	23	1	1
24	1	24	1	1
25	1	25	1	1
26	1	26	1	1
27	1	27	1	1
28	1	28	1	1
29	1	29	1	1
30	1	30	1	1
31	1	31	1	1
32	1	32	1	1
33	1	33	1	1
34	1	34	1	1
35	1	35	1	1
36	1	36	1	1
37	1	37	1	1
38	1	38	1	1
39	1	39	1	1
40	1	40	1	1
41	1	41	1	1
42	1	42	1	1
43	1	43	1	1
44	1	44	1	1
45	1	45	1	1
46	1	46	1	1
47	1	47	1	1
48	1	48	1	1
49	1	49	1	1
50	1	50	1	1
51	1	51	1	1
52	1	52	1	1
53	1	53	1	1
54	1	54	1	1
55	1	55	1	1
56	1	56	1	1
57	1	57	1	1
58	1	58	1	1
59	1	59	1	1
60	1	60	1	1
61	1	61	1	1
62	1	62	1	1
63	1	63	1	1
64	1	64	1	1
65	1	65	1	1
66	1	66	1	1
67	1	67	1	1
68	1	68	1	1
69	1	69	1	1
70	1	70	1	1
71	1	71	1	1
72	1	72	1	1
73	1	73	1	1
74	1	74	1	1
75	1	75	1	1
76	1	76	1	1
77	1	77	1	1
78	1	78	1	1
79	1	79	1	1
80	1	80	1	1
81	1	81	1	1
82	1	82	1	1
83	1	83	1	1
84	1	84	1	1
85	1	85	1	1
86	1	86	1	1
87	1	87	1	1
88	1	88	1	1
89	1	89	1	1
90	1	90	1	1
91	1	91	1	1
92	1	92	1	1
93	1	93	1	1
94	1	94	1	1
95	1	95	1	1
96	1	96	1	1
97	1	97	1	1
98	1	98	1	1
99	1	99	1	1
100	1	100	1	1

Gambar 1. Data Pemakaian Obat 2017 dan 2018

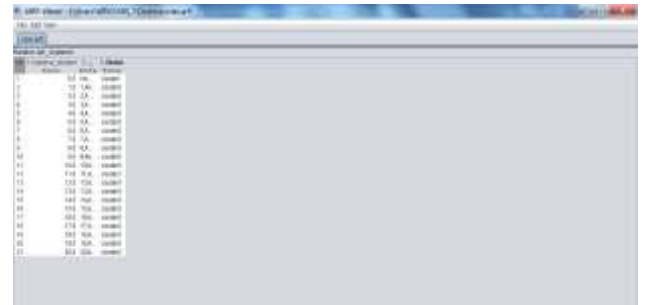
Buka Program weka yang telah diinstal akan muncul seperti gambar 2.



Gambar 2. Tampilan awal aplikasi



Gambar 3. Weka explore



Gambar 4. Tampilan hasil

5. KESIMPULAN

Dari hasil analisa data pemakaian obat pada Dinas puskesmas Teladan medan dapat di ambil kesimpulan sebagai berikut:

- Penentuan centroid (titik pusat) pada tahap awal Algoritma K-Means sangat berpengaruh pada hasil cluster seperti pada hasil pengujian yang dilakukan dengan menggunakan 20 dataset dengan centroid yang berbeda menghasilkan hasil cluster yang berbeda juga.
- Setelah dilakukan pengelompokan data pemakaian obat menggunakan metode k-means Clustering terbentuk tiga cluster yaitu cluser satu dengan jumlah 1 items, cluster dua 1 dengan jumlah 18 items, dan cluster tiga dengan jumlah 20 items.
- Jumlah data yang digunakan dalam melakukan analisa data pemakaian obat dengan menggunakan metode K-Means sangat penting. Penggunaan Weka Version 3.8 dapatdilakukan sebagai bahasa pemograman untuk merancang aplikasi data pemakaian obat.

DAFTAR PUSTAKA

- [1] E. Prasetyo, Data Mining. Yogyakarta: C.V ANDI OFFSET, 2012.
- [2] K. E. T. Luthfi, "Algoritma Data Mining," 2009.
- [3] F. A. Hermawati, Data Mining, Andi. Yogyakarta, 2013.
- [4] john willey&Sons. Inc, "Discovering Knowledge in Data : an Introduction to Data Mining.," 2005.
- [5] Syamsuni, Farmasetika Dasar & Hitungan Farmasi. Jakarta, 2005.
- [6] R. Alfina, T., Santosa, B.,& Barakbah, "Analisa Perbandingan Metode Hierarchical Clustering Data," 2012.
- [7] T. Khotimah, "Pengelompokan Surat Dalam AL QUR'AN Menggunakan Algoritma K-Means," 2014.
- [8] M. H. Adiya and Y. Desnelita, "Jurnal Nasional Teknologi dan Sistem Informasi Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan Pada RSUD Pekanbaru," vol. 01, pp. 17–24, 2019.
- [9] R.Setiawan, Teknik Pemecahan Masalah dengan Algoritma dan Flowchart(Basic dan C). jakarta, 2009.
- [10] A.-B. i. Ladjamudin, "Analisis dan Desain Sistem Informasi," Tangerang, 2012.