

# PEMANFAATAN ALGORITMA K-MEANS UNTUK STOK OBAT DI PUSKESMAS SILO

<sup>1</sup> Dwi Wahyu Aprillianto (1110651157)

<sup>2</sup> Daryanto, S.Kom, M.Kom, <sup>3</sup> Hardian Oktavianto, S.Si

Jurusan Teknik Informatika Fakultas Teknik Universitas Muhammadiyah Jember

Email : coboywahyu313@gmail.com

## ABSTRAK

Puskesmas Silo merupakan salah satu puskesmas yang ada di kabupaten Jember. Puskesmas Silo memiliki bermacam-macam obat seperti, paracetamol, amoxilin, dan lain-lain. Setelah melakukan pengamatan dalam jangka waktu tertentu, di dapatkan Puskesmas Silo memiliki kelemahan didalam manajemen puskesmas, dari permasalahan tersebut disebuah Puskesmas Silo tidak bisa menyimpulkan obat yang sering laku, laku, cukup laku dan tidak laku. Berdasarkan analisa yang telah dijabarkan diatas, penulis akan membangun sebuah aplikasi yang dapat mengelompokkan data dalam suatu stok jumlah besar maupun kecil. *Euclidean distance space* sering digunakan dalam perhitungan jarak, hal ini dikarenakan hasil yang diperoleh merupakan jarak terpendek antara dua titik yang diperhitungkan.

Selama ini stok obat di Puskesmas Silo masih menggunakan perhitungan secara manual menggunakan microsoft excel untuk mencatat setiap stok obat baru. Maka dari itu saya akan memecahkan masalah tersebut dengan otomatis, jadi nantinya akan diketahui stok obat yang sering laku, laku, cukup laku dan tidak laku.

Kata kunci : *Euclidean distance space, K-Means*

### 1. Pendahuluan

Puskesmas Silo memiliki bermacam-macam obat seperti, paracetamol, amoxilin, dan lain-lain. Setelah melakukan pengamatan dalam jangka waktu tertentu, di dapatkan Puskesmas Silo memiliki kelemahan didalam manajemen puskesmas, dari permasalahan tersebut disebuah Puskesmas Silo tidak bisa menyimpulkan obat yang sering laku, laku, cukup laku dan tidak laku. Dalam melakukan penyediaan stok obat, Puskesmas Silo menggunakan catatan dalam bentuk buku. Untuk mendapatkan hasil stok obat yang lebih baik maka dibuat aplikasi yang dapat digunakan untuk melakukan stok obat dalam melakukan pencariannya. Puskesmas Silo melakukan rekapitulasi data transaksi obat secara periodik, yaitu data transaksi harian, mingguan, bulanan ataupun tahunan. Dari data tersebut bisa dibuat alternatif dengan cara mengelompokkan, antara obat yang laku, cukup laku dan tidak laku. Metode *clustering* adalah metode yang tepat untuk pengumpulan data, sehingga menghasilkan informasi yang penting. Kemudian *algoritma k-means* merupakan salah satu metode *clustering* non hirarki yang berusaha mempartisi data yang ada

kedalam bentuk satu atau lebih *cluster*. Metode ini mempartisi data kedalam *cluster* sehingga data yang memiliki karakteristik yang sama dikelompokkan kedalam satu *cluster* yang sama dan data yang mempunyai karakteristik yang berbeda di kelompokkan kedalam *cluster* yang lain. Selama ini stok obat di Puskesmas Silo masih menggunakan perhitungan secara manual menggunakan microsoft excel untuk mencatat setiap stok obat baru. Maka dari itu saya akan memecahkan masalah tersebut dengan otomatis, jadi nantinya akan diketahui stok obat yang sering laku, laku, cukup laku dan tidak laku.

### 2. Metodologi Penelitian

#### 2.1 Data Mining

Data Mining adalah serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual. Dengan semakin pesat perkembangan teknologi, semakin berkembang pula pola pikir seseorang dalam mengumpulkan dan mengolah data. Penggunaan sistem komputerisasi semakin banyak digemari dalam berbagai bidang baik itu dalam transaksi-transaksi, jasa, bisnis,

maupun kalangan pemerintah dan sosial, dengan menggunakan sistem komputerisasi maka telah banyak menghasilkan data yang berukuran sangat besar. Data-data yang terkumpul ini merupakan suatu bahan yang dapat digunakan sebagai informasi dalam dunia jasa maupun bisnis.

Didalam ilmu komputerisasi telah banyak aplikasi-aplikasi seperti basis data yang telah banyak diterapkan diberbagai bidang antara lain bidang ilmu pengetahuan, administrasi pemerintah, manajemen dan bidang-bidang lainnya. sehingga data yang dihasilkan sangatlah banyak dan besar. Untuk itu maka dilakukan teknik-teknik untuk melakukan pengolahan data sehingga dari data-data yang diperoleh data dijadikan suatu informasi yang penting yang dapat digunakan untuk perkembangan dalam masing-masing bidang bisnis maupun jasa.

## 2.2 Clustering

*Clustering* adalah proses mengelompokkan objek berdasarkan informasi yang diperoleh dari data yang menjelaskan hubungan antar objek dengan prinsip untuk memaksimalkan kesamaan antar anggota satu kelas dan meminimumkan kesamaan antar kelas/cluster. Tujuannya menemukan *cluster* yang berkualitas dalam waktu yang layak. *Clustering* dalam data mining berguna untuk menemukan pola distribusi di dalam sebuah data set yang berguna untuk proses analisis data. Kesamaan objek biasanya diperoleh dari kedekatan nilai-nilai atribut yang menjelaskan objek-objek data, sedangkan objek-objek data biasanya direpresentasikan sebagai sebuah titik dalam ruang multidimensi. Dengan menggunakan *clustering*, dapat diidentifikasi daerah yang padat, pola-pola distribusi secara keseluruhan dan keterkaitan yang menarik antara atribut-atribut data. Dalam data mining usaha difokuskan pada metode-metode penemuan untuk *cluster* pada basis data berukuran besar secara efektif dan efisien. Banyaknya pendekatan *clustering* menyulitkan dalam menentukan ukuran kualitas yang *universal*. Namun, beberapa hal yang perlu diperhatikan adalah input parameter yang tidak menyulitkan *user*, *cluster* hasil yang dapat dianalisis, dan skalabilitas terhadap penambahan ukuran dimensi dan *record* dataset.

## 2.3 K-Means Clustering

Metode *K-Means* adalah metode pengelompokan data dengan mengambil

parameter sejumlah *k cluster*, dan mempartisi data kedalam *cluster* tersebut, dengan berpatokan pada kemiripan antar data dalam satu *cluster* dan ketidakmiripan di antar *cluster* yang berbeda, pusat dari *cluster* adalah rata-rata dari nilai anggota *cluster* yang disebut *centroid* atau *center of gravity* (Kamber, 2007,4). Selain itu *K-Means* melakukan pengelompokan dengan meminimalkan jumlah kuadrat dari jarak (*disance*) antara data dengan *centroid cluster* yang cocok. (Teknomo, 2006,25).

Algoritma *k-Means* adalah algoritma partitional (*non hierarchical*) *clustering* yang mempartisi atau membagi sekumpulan data ke dalam sejumlah *cluster*. Setiap *cluster* mempunyai titik pusat *centroid*. *Centroid* adalah rata-rata (*mean*) dari setiap titik anggota *cluster*.

## 2.4 Euclidean Distance

*Euclidean Distance* adalah metrika yang paling sering digunakan untuk menghitung kesamaan 2 vektor. *Euclidean distance* menghitung akar dari kuadrat perbedaan 2 vektor (*root of square differences between 2 vectors*). digunakan untuk menghitung jarak antara data dan *centroid*. Adapun persamaan yang dapat digunakan salah satunya yaitu *Euclidean Distance Space*. *Euclidean distance space* sering digunakan dalam perhitungan jarak, hal ini dikarenakan hasil yang diperoleh merupakan jarak terpendek antara dua titik yang diperhitungkan.

## 3. Hasil dan Pembahasan

Algoritma *k-means* dipilih karena metode ini sederhana, mudah diimplementasikan, memiliki perhitungan untuk mengklaster data yang besar, mampu menangani data banyak, dan kompleksitas waktunya linier  $O(nKT)$  dengan jumlah dokumen, *K* adalah jumlah kluster dan *T* adalah jumlah iterasi. Untuk pengertiannya Algoritma *k-means* adalah metode pengklasteran secara data yang akan dipilih kedalam kelompok yang berbeda. Selama ini stok obat di Puskesmas Silo masih menggunakan perhitungan secara manual menggunakan microsoft excel untuk mencatat setiap stok obat baru. Maka dari itu saya akan memecahkan masalah tersebut dengan otomatis, jadi nantinya akan diketahui stok obat yang sering laku, laku, cukup laku dan tidak laku.

No	Nama barang	B1	B2	B3	B4	B5	B6
1.	Paracetamol	187	213	107	69	165	191
2.	Amoxilin	131	164	41	127	154	89
3.	Alinamin F	40	41	42	55	36	50
4.	Blesifen	217	189	336	217	211	268
5.	Exluton	261	246	240	226	321	246
6.	Hemorid	137	145	176	149	177	143
7.	Laxing	32	28	32	22	48	41

Selanjutnya kita mencoba mengelompokkan data tersebut di atas menjadi 3 kelompok. Dengan menggunakan algoritma *K-Means*, berikut langkah-langkah penyelesaiannya: Menentukan jumlah cluster, dimana jumlah cluster = 3 Menentukan pusat cluster secara acak. Disini kita tentukan kita tentukan c1, c2 dan c3 Menghitung jarak setiap data yang ada terhadap setiap pusat cluster. Berikut perhitungannya dengan menggunakan persamaan *Euclidean Distance Space* :

Langkah pertama menentukan *Centroid* dengan nilai *random* ataupun nilai *value* dari data tersebut.

Nilai *Centroid* secara *Random* :

C1=(85, 75, 32, 35, 98, 37)

C2=(47, 63, 42, 3, 21, 22)

C3=(18, 28, 73, 86, 100, 58)

langkah ke dua

Perhitungan Menggunakan rumus *Euclidean Distance Space*.

1.Paracetmol

187	213	107	69	165	191
-----	-----	-----	----	-----	-----

$$C11 = \sqrt{(187 - 85)^2 + (213 - 75)^2 + (107 - 32)^2 + (69 - 35)^2 + (165 - 98)^2 + (191 - 37)^2}$$

$$= \sqrt{10.404 + 19.044 + 5.625 + 1.156 + 4.489 + 23.716}$$

$$= \sqrt{64.434}$$

$$= 253$$

$$C12 = \sqrt{(187 - 47)^2 + (213 - 63)^2 + (107 - 42)^2 + (69 - 3)^2 + (165 - 21)^2 + (191 - 22)^2}$$

$$= \sqrt{19.600 + 22.500 + 4.225 + 4.356 + 20.736 + 23.716}$$

$$= \sqrt{99.978}$$

$$= 316$$

$$C13 = \sqrt{(187 - 18)^2 + (213 - 28)^2 + (107 - 73)^2 + (69 - 86)^2 + (165 - 100)^2 + (191 - 58)^2}$$

$$= \sqrt{28.561 + 34.225 + 1.156 + 289 + 4.225 + 17.689}$$

$$= \sqrt{86.145}$$

$$= 293,5$$

2.Amoxilin

131	164	41	127	154	89
-----	-----	----	-----	-----	----

$$C11 = \sqrt{(131 - 85)^2 + (164 - 75)^2 + (41 - 32)^2 + (127 - 35)^2 + (154 - 98)^2 + (89 - 37)^2}$$

$$= \sqrt{2.116 + 7.921 + 81 + 8.464 + 3.136 + 2.704}$$

$$= \sqrt{24.422}$$

$$= 156,2$$

$$C12 = \sqrt{(131 - 47)^2 + (164 - 63)^2 + (41 - 42)^2 + (127 - 3)^2 + (154 - 21)^2 + (89 - 22)^2}$$

$$= \sqrt{7.056 + 10.201 + 1 + 15.376 + 17.689 + 4.489}$$

$$= \sqrt{54.812}$$

$$= 234,1$$

$$C13 = \sqrt{(131 - 18)^2 + (164 - 28)^2 + (41 - 73)^2 + (127 - 86)^2 + (154 - 100)^2 + (89 - 58)^2}$$

$$= \sqrt{12.769 + 18.496 + 676 + 1.681 + 2.916 + 961}$$

$$= \sqrt{37.847}$$

$$= 194,5$$

3. Alinamin F

40	41	42	55	36	50
----	----	----	----	----	----

$$C11 = \sqrt{(40 - 85)^2 + (41 - 75)^2 + (42 - 32)^2 + (55 - 35)^2 + (36 - 98)^2 + (50 - 37)^2}$$

$$= \sqrt{2.025 + 1.156 + 100 + 400 + 3.844 + 169}$$

$$= \sqrt{7.694}$$

$$= 87,7$$

$$C12 = \sqrt{(40 - 47)^2 + (41 - 63)^2 + (42 - 42)^2 + (55 - 3)^2 + (36 - 21)^2 + (50 - 22)^2}$$

$$= \sqrt{49 + 484 + 0 + 2.704 + 225 + 784}$$

$$= \sqrt{484 + 169 + 961 + 961 + 4.096 + 64}$$

$$= \sqrt{6.735}$$

$$= 82$$

#### 4. Implementasi Progam

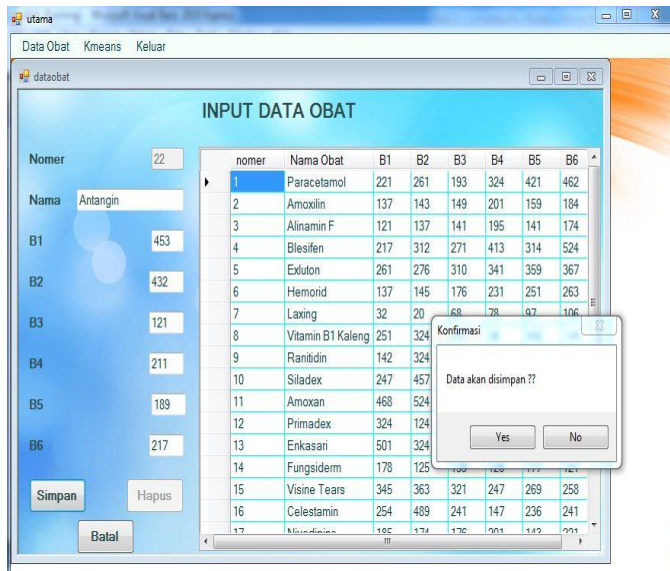
Proses Perhitungan Pada kasus ini pengelompokan menggunakan metode *K-Means*, terdapat beberapa tahapan. Tahapan dari metode *K-Means* yaitu set data dan

menghitung bobot jarak dari tiap data dengan cluster dengan rumus .

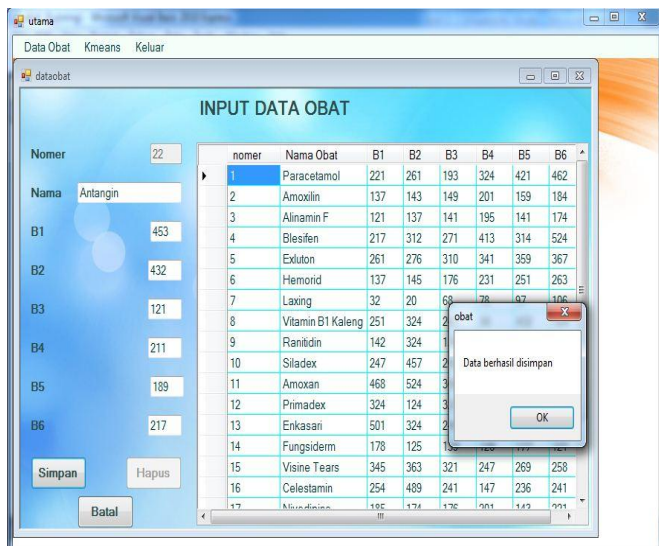
Untuk menghitung bobot jarak pada *K-Means* menggunakan persamaan berikut:

**Bobot**

$$ED = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2}$$



Untuk input data stok obat nantinya akan menginputkan nama obat, Bulan pertama, bulan kedua, sampai bulan keenam, semisal data ada yang belum diisi akan *error*.



Untuk pengimputan data berhasil disimpan, menyimpan data stok obat yang baru diinputkan oleh petugas Puskesmas Silo.

Nama Obat	b1	b2	b3	b4	b5	b6	cluster 1	cluster 2	cluster 3	cluster 4	posisi	iter
Paracetamol	221	261	193	324	421	462	260.894	351.148	410.986	628.686	1	5
Amoxilin	137	143	149	201	159	184	396.044	210.011	119.090	232.196	3	5
Alinamin F	121	137	141	195	141	174	421.196	239.483	140.343	209.273	3	5
Blesifen	217	312	271	413	314	524	253.791	401.961	467.974	696.559	1	5
Exluton	261	276	310	341	359	367	141.614	256.642	348.004	607.684	1	5
Hemorid	137	145	176	231	251	263	333.350	189.474	165.323	323.520	3	5
Laxing	32	20	68	78	97	106	631.934	430.788	324.590	29.000	4	5
Vitamin B1 Kaleng	251	324	247	56	432	124	368.905	261.937	346.698	525.271	2	5
Ranitidin	142	324	156	324	124	236	294.691	227.404	244.916	431.442	2	5
Siladex	247	457	265	368	245	336	129.715	291.858	392.200	651.400	1	5
Amoxan	468	524	369	348	241	214	257.879	416.641	493.204	789.554	1	5
Primadex	324	124	321	245	169	112	358.948	245.153	162.942	423.515	3	5
Enkasari	501	324	245	361	257	169	257.149	353.338	355.141	666.267	1	5
Fungsiderm	178	125	190	174	177	121	447.415	290.467	127.967	207.446	1	5

Uji perhitungan menggunakan 4 *Cluster*, Dari percobaan *cluster*, *Centroid C1,C2,C3,C4* dengan menggunakan nilai:

*Centroid 1*=(215,202,275,185,276,223).

*Centroid 2*=(169,159,214,144,219,177).

*Centroid 3*=(124,115,154,104,162,132).

*Centroid 4*=(78,72,93,63,105,86).

Setelah titik *centroid* sudah ditentukan, maka data stock obat di Puskesmas Silo akan diproses dengan menggunakan metode *K-Means* dengan menggunakan program *Visual Basic.NET*

## 5. Kesimpulan

Dalam penelitian tugas akhir diperoleh sebuah kesimpulan diantaranya. Hasil dari yang telah dilakukan mendapatkan parameter dari yang terkecil sampai yang terbesar. Dari 21 Data digunakan 4 *Cluster*, Data yang paling kecil berjumlah 1 dan yang paling besar berjumlah 8. Bahwa obat yang dikategorikan yang sering laku berjumlah 8 obat, katagori laku berjumlah 4 obat, katagori cukup laku berjumlah 8 obat, dan katagori tidak laku berjumlah 1 yaitu obat Laxing dengan pengujian data 21 obat mulai bulan pertama sampai bulan keenam.

## DAFTAR PUSTAKA

- Afif, A., Prayudhi, Reksa A., (2011), *Clustering data penjualan dan persediaan barang pada P.T sayap mas utama dengan metode K-means*. STMIK GI MDP, Denpasar, Bali.
- Agusta, Y., (2007), *K-Means – Penerapan, Permasalahan dan Metode Terkait* STMIK STIKOM BALI, Denpasar, Bali.
- Amri, MK., (2013), Penerapan data mining untuk menentukan strategi penjualan pada toko buku Gramedia Palembang. Tugas Akhir, Universitas Palembang, Palembang.
- Ngurah, I.G.S,(2014), Cara merancang aplikasi Visual Basic, Penerbit elex media komputido. Bali.
- Oscar, J.O, (2013), Implementasi algoritma k-means clustering untuk menentukan strategi marketingpresident university, President University, Bekasi.
- Wulandari, R.,(2012), Pengenalan rumput laut menggunakan Euclidean distance berbasis Ekstra fitur,Tugas Akhir, Universitas Nusantara PGRI Kediri, Kediri.
- Agusta, Yudi. 2007. “*K-Means – Penerapan, Permasalahan dan Metode Terkait*”.STMIK STIKOM BALI..
- Bachtiar.2010. “*Rancang Bangun Sistem Pendukung Keputusan(SPK) Penjurusan Program Studi*”.Fakultas Sains dan Teknologi,Universitas Islam Negeri Syarif Hidayatullah. Jakarta
- Eric Budiman,Isye Arieshanti dan Rully Soelaiman. 2013. “*Implementasi KD-Tree-K-Means Clustering untuk Klasterisasi Dokumen*”. Fakultas Teknologi Informasi, Institut Teknologi Sepuluh Nopember (ITS).
- Gautama, Mohammad Glesung. 2010. “*Penentuan Jurusan Di SMA N 8 Surakarta Dengan Fuzzy Inference System (FIS) Mamdani*”. Fakultas Matematik dan Ilmu Pengetahuan Alam, Universitas Sebelas Maret. Surakarta
- Kusrini, Emha Taufik Luthfi. 2009. “*Algoritma Data Mining*”. Yogyakarta: Andi.
- Setiawan,Deni.2015 .”*Perancangan Aplikasi K-Means Sebagai Penentu Konsentrasi Bagi Mahasiswa Informatika UMS*. Fakultas komunikasi dan Informatika, Universitas Muhammadiyah. Surakarta.
- Tedy dan Sri Kusumadewi. 2008. “*Aplikasi Untuk Mengelompokkan Mahasiswa Berdasarkan Nilai Body Mass Index (Bmi) & Ukuran Kerangka*”. Jurusan Teknik Informatika FTI, UII. Yogyakarta.
- Teguh Hariyadi, Suprayogi. 2012. “*Penerapan Algoritma K-Means Untuk Pengelompokan Data Nilai Siswa*”. Fakultas Ilmu Komputer, Universitas Dian Nuswantoro. Semarang
- Teknomo, Kardi. 2006. “*K-means Clustering Tutorial.*” *Medicine*. 100.4 (2006): 3.
- Turban ,Efraim, Kelly Rainer,Richard E Potter. 2005 .”*Pengantar Teknologi Informasi*”. Jakarta : Salemba .
- Wulandari, Retno . 2012. “*Pengenalan rumput laut menggunakan Euclidean distance berbasis Ekstra fitur*”.Universitas Nusantara PGRI Kediri.
- Yakub. 2012. “*Pengantar Sistem Informasi*”. Yogyakarta: Graha Ilmu.