

DATA MINING DENGAN METODE CLUSTERING UNTUK PENGOLAHAN INFORMASI PERSEDIAAN OBAT PADA PUSKESMAS PANDANARAN SEMARANG

Joanna Ardhyanti Mita Nugraha¹, Yupie Kusumawati²

^{1,2}Sistem Informasi, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro

Jalan Nakula I No. 5-11, Semarang, 50131, (024) 3517261

E-mail : joanna.nugraha@gmail.com¹, yupieku@gmail.com²

Abstrak

Persediaan obat merupakan salah satu faktor penting untuk kelangsungan proses pada Puskesmas. Belum adanya metode baku yang diterapkan merupakan masalah yang sering terjadi. Karena pada umumnya persediaan obat pada Puskesmas hanya berdasarkan pada pengalaman sebelumnya. Yaitu hanya dengan memeriksa persediaan obat yang hampir habis, kemudian membelinya. Hal ini kurang maksimal jika suatu waktu membutuhkan obat dalam jumlah yang besar dan ternyata stok habis. Penerapan data mining dapat membantu untuk menganalisa pemakaian obat, untuk mendapatkan informasi yang dapat digunakan untuk perencanaan dan pengendalian persediaan obat. Metode clustering merupakan salah satu fungsional dari data mining yang merupakan pengelompokan item data ke dalam sejumlah kecil grup sedemikian sehingga masing-masing grup mempunyai sesuatu persamaan yang esensial. Algoritma K-Means merupakan algoritma yang paling populer dan banyak digunakan dalam penggunaan metode clustering data mining. Hasil yang diperoleh berupa informasi jenis obat dengan pemakaian cepat setiap bulannya yang diambil dari data 3 tahun terakhir, yang nantinya dapat digunakan sebagai metode pengendalian persediaan obat yang lebih efektif dan efisien pada Puskesmas Pandanaran.

Kata Kunci: persediaan, obat, data mining, clustering, algoritma K-Means

Abstract

Inventories of drugs is one of the important factors for the sustainability of the process at the health center. The absence of standard methods applied to a common problem. Because the supply of drugs to health centers generally only based on previous experience. That is just by checking drug supplies running low, then buy it. This is less than the maximum if a drug requires time in large numbers and it turns out the stock runs out. Application of data mining can help to analyze the use of drugs, to obtain information that can be used for planning and control of drug supply. The method is one of the functional clustering of data mining which is a grouping of data items into a number of small groups so that each group has something essential equations. K-Means algorithm is an algorithm which is the most popular and widely used in the use of clustering method of data mining. The result will be a kind of medicine with the use of information quickly every month which is taken from the last 3 years of data, which can later be used as a method of inventory control medication more effective and efficient at PHC Pandanaran.

Keywords: inventories, drug, data mining, clustering, K-Means algorithm

1. PENDAHULUAN

Kesehatan merupakan hak asasi manusia dan sekaligus merupakan investasi untuk keberhasilan pembangunan Bangsa Indonesia. Oleh

karena itu perlu dilakukan pembangunan kesehatan secara menyeluruh dan berkesinambungan, dengan tujuan guna meningkatkan kesadaran, kemauan dan kemampuan hidup sehat bagi setiap orang agar

terwujud derajat kesehatan masyarakat yang setinggi-tingginya, salah satu caranya adalah dengan meningkatkan akses dan mutu pelayanan kesehatan.

Salah satu unsur penting dalam memelihara dan meningkatkan mutu pelayanan kesehatan adalah ketersediaan obat yang memadai, dimana diantara berbagai alternatif yang ada, intervensi dengan obat merupakan intervensi yang paling banyak digunakan dalam penyelenggaraan upaya kesehatan dan biaya yang digunakan untuk obat merupakan bagian yang cukup besar dari seluruh biaya kesehatan [1]. Di beberapa negara maju biaya obat berkisar antara 10-15 % dari anggaran kesehatan, sementara di negara berkembang biaya ini lebih besar lagi yaitu antara 35- 66 %, sebagai contoh di negara Thailand sebesar 35 %, Cina sebesar 45 %, Mali sebesar 66 % dan Indonesia sebesar 39 %. Mengingat besarnya biaya obat, maka diperlukan pengelolaan obat secara benar, efisien dan efektif secara berkesinambungan serta dengan koordinasi yang baik dan terbuka antara pihak yang terkait [2].

Perencanaan kebutuhan obat merupakan salah satu aspek penting dan menentukan dalam pengelolaan obat, karena perencanaan kebutuhan obat akan mempengaruhi pengadaan, pendistribusian dan pemakaian obat di unit pelayanan kesehatan, dimana dengan perencanaan kebutuhan obat yang tepat akan membuat pengadaan menjadi efektif dan efisien sehingga tersedia obat dengan jenis dan jumlah yang cukup sesuai dengan kebutuhan pelayanan kesehatan dengan mutu yang terjamin serta dapat diperoleh pada saat yang diperlukan.

Saat ini kenyataannya sebagian besar Puskesmas di Indonesia belum melakukan kegiatan pelayanan farmasi seperti yang diharapkan, mengingat beberapa kendala antara lain kemampuan tenaga farmasi, terbatasnya pengetahuan manajemen Puskesmas akan fungsi farmasi obat, kebijakan manajemen institusi yang membawahi suatu Puskesmas, terbatasnya pengetahuan pihak-pihak terkait tentang pelayanan farmasi. Akibat kondisi ini maka pelayanan farmasi Puskesmas masih bersifat konvensional yang hanya berorientasi pada produk yaitu sebatas penyediaan dan pendistribusian.

Salah satu metode *data mining* yang dapat digunakan dari analisa diatas adalah metode *clustering*, yang merupakan pengelompokan item data ke dalam sejumlah kecil grup sedemikian sehingga masing-masing grup mempunyai sesuatu persamaan yang esensial. Algoritma K-Means merupakan algoritma yang paling populer dan banyak digunakan. Algoritma ini disusun atas dasar ide sederhana. Ada awalnya ditentukan berapa *cluster* yang akan dibentuk. Sembarang objek atau elemen pertama dalam *cluster* dapat dipilih untuk dijadikan sebagai titik tengah (*control point*) *cluster*. Algoritma K-Means selanjutnya akan melakukan pengulangan langkah-langkah sampai terjadi kestabilan (tidak ada objek yang dapat dipindahkan). Dengan hasil *clustering data mining* akan dihasilkan jenis obat dengan pemakaian cepat setiap bulannya yang dapat digunakan acuan perencanaan persediaan pada tahun berikutnya. Selain itu informasi yang dihasilkan dari *data mining* dapat digunakan sebagai rekomendasi pada Puskesmas untuk meningkatkan pelayanan masyarakat.

2. DATA MINING

2.1 Pengertian Data Mining

Data mining merupakan proses pencarian pola-pola yang menarik dan tersembunyi (*hidden pattern*) dari suatu kumpulan data yang berukuran besar yang tersimpan dalam suatu basis data, *data warehouse*, atau tempat penyimpanan data lainnya. [3]

2.2 Metodologi CRISP-DM

CRISP-DM (*Cross Industry Standart Process for Data Mining*) adalah standarisasi data mining yang disusun oleh tiga penggagas *data mining market*. Yaitu Daimler Chrysler (Daimler-Benz), SPSS (ISL), NCR. Kemudian dikembangkan pada berbagai *workshops* (antara 1997 – 1999). Lebih dari 300 organisasi yang berkontribusi dalam proses modeling ini dan akhirnya CRISP-DM 1.0 dipublikasikan pada 1999. Proses data mining berdasarkan CRISP-DM terdiri dari 6 fase, yaitu : [4]

1.Fase Pemahaman Bisnis (*Business Understanding Phase*)

- a.Penentuan tujuan proyek dan kebutuhan secara detail dalam lingkup bisnis atau unit penelitian secara keseluruhan
- b.Menerjemahkan tujuan dan batasan menjadi formula dari permasalahan data mining.
- c.Menyiapkan strategi awal untuk mencapai tujuan.

2.Fase Pemahaman Data (*Data Understanding Phase*)

- a.Mengumpulkan data
- b.Menggunakan analisis penyelidikan data untuk mengenali lebih lanjut data dan pencarian pengetahuan awal.
- c.Mengevaluasi kualitas data

3.Fase Pengolahan Data (*Data Preparation Phase*)

- a.Siapkan dari data awal, kumpulan data yang akan digunakan untuk keseluruhan fase berikutnya. Fase ini merupakan pekerjaan berat yang perlu dilaksanakan secara intensif.
 - b.Pilih kasus dan variabel yang ingin di analisis dan yang sesuai analisa yang akan dilakukan.
 - c.Siapkan data awal sehingga siap untuk perangkat pemodelan.
- #### 4.Fase Pemodelan (*Modeling Phase*)
- a.Pilih dan aplikasikan teknik pemodelan yang sesuai
 - b.Kalibrasi aturan model untuk mengoptimalkan hasil
 - c.Perlu diperhatikan bahwa beberapa teknik mungkin untuk digunakan pada permasalahan data mining yang sama
 - d.Jika diperlukan proses dapat kembali ke fase pengolahan data untuk menjadikan data ke dalam bentuk yang sesuai dengan spesifikasi kebutuhan teknik data mining tertentu.
- #### 5.Fase Evaluasi (*Evaluation Phase*)
- a.Mengevaluasi satu atau lebih model yang digunakan dalam fase pemodelan untuk mendapatkan kualitas dan efektivitas sebelum disebarkan untuk digunakan.
 - b.Menetapkan apakah terdapat model yang memenuhi tujuan pada fase awal.
 - c.Menentukan apakah terdapat permasalahan penting dari bisnis atau penelitian yang tidak tertangani dengan baik.
 - d.Mengambil keputusan berkaitan dengan penggunaan hasil dari data mining.
- #### 6.Fase Penyebaran (*Deployment Phase*)
- a.Menggunakan model yang dihasilkan. Terbentuknya model tidak menandakan telah terselesaikannya proyek.

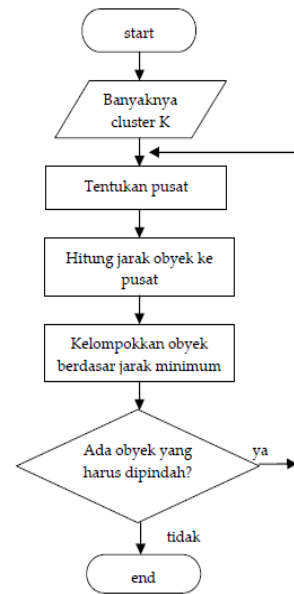
2.3 Clustering

Garcia - Molina et al. [5] menyatakan *Clustering* adalah mengelompokkan item data ke dalam sejumlah kecil grup sedemikian sehingga masing - masing grup mempunyai sesuatu persamaan yang esensial.

Pengklusteran merupakan pengelompokan *record*, pengamatan, atau memperhatikan dan membentuk kelas objek-objek yang memiliki kemiripan. Kluster adalah kumpulan *record* yang memiliki kemiripan suatu dengan yang lainnya dan memiliki ketidakmiripan dengan *record* dalam kluster lain.

2.4 Algoritma K-Means

Pada algoritma ini, pusat *cluster* atau *centroid* dipilih pada tahap awal secara acak dari sekumpulan koleksi (populasi) data. Kemudian *K-Means* menguji masing-masing komponen didalam populasi data dan menandai komponen tersebut ke salah satu *centroid* yang telah didefinisikan sebelumnya berdasarkan jarak minimum antara komponen (data) dengan masing-masing *centroid*. Posisi *centroid* akan dihitung kembali sampai semua komponen data dikelompokkan ke setiap *centroid* dan terakhir akan terbentuk posisi *centroid* baru. Iterasi ini akan terus dilakukan sampai tercipta kondisi konvergen. Secara lebih detail, algoritma *K-Means Clustering* adalah sebagai berikut :



Gambar 1. Flowchart K-Means Clustering

3. HASIL DAN PEMBAHASAN

3.1 Fase Pemahaman Bisnis

Pemahaman bisnis dilakukan dengan mempelajari objek penelitian yaitu Puskesmas Pandanaran. Dalam fase ini dilakukan penentuan tujuan bisnis yaitu untuk mengetahui jenis obat dengan pemakaian tertinggi setiap bulannya selama 3 tahun, sehingga nantinya jika akan melakukan persediaan, petugas dapat melihat apakah jenis obat yang akan di-stok memang layak untuk disediakan dalam jumlah besar atau disediakan dalam jumlah kecil sesuai dengan pemakaian yang telah terjadi setiap bulannya selama 3 tahun. Setelah itu dilakukan penerapan tujuan, penerapan tujuan sama dengan menentukan tujuan data mining yaitu dengan menggali pengetahuan tentang jenis obat yang pemakaiannya cepat dan lama yang didapatkan dari proses *clustering*. Selanjutnya adalah menyiapkan strategi awal penelitian, yang dilakukan dengan permintaan data pada Puskesmas Pandanaran. Data yang

akan diminta merupakan data LPLPO dari tahun 2011 – 2013.

3.2 Fase Pemahaman Bisnis

Data yang dikumpulkan merupakan data LPLPO dari tahun 2011 – 2013. Data LPLPO didapatkan dari bagian yang mengurus Laporan Data Obat dan data LPLPO disimpan dalam bentuk excel. Setelah data didapatkan, proses selanjutnya adalah memahami data tersebut. Data yang didapatkan sejumlah 9697 data selama 3 tahun. Selanjutnya dilakukan pengevaluasian kualitas data dan kelengkapan data. Nilai-nilai yang hilang sering terjadi, terutama jika data yang dikumpulkan di jangka waktu yang lama. Memeriksa atribut yang hilang atau kosong, ejaan nilai-nilai, dan apakah atribut dengan nilai yang berbeda memiliki arti yang sama.

3.3 Fase Pemahaman Data

Persiapan data dilakukan dengan memilih atribut yang nantinya akan digunakan untuk proses *modeling* yang diambil dari LPLPO. Atribut yang telah dipilih nantinya akan disimpan kembali kedalam dataset yang baru yang siap untuk diproses kedalam proses *modeling*. Atribut yang dipilih merupakan atribut nama obat, stock awal, penerimaan, pemakaian, persediaan, pemakaian, stok akhir dan permintaan. Kemudian dilakukan transformasi data dengan memindahkan atribut yang telah dipilih pada proses pemilihan data untuk dipindahkan pada dataset yang baru yang nantinya siap diproses pada fase pemodelan.

3.4 Fase Pemodelan

Pada tahap pemodelan akan menggunakan metode *clustering* dengan

algoritma k-means. Dalam penerapan metode *clustering*, akan dibagi menjadi 2 *cluster* yang nantinya akan meng-*cluster* jenis obat dengan pemakaian cepat dan jenis obat dengan pemakaian lama. Alat pemodelan yang digunakan adalah Rapidminer 5.3. Rapidminer dapat digunakan untuk memudahkan dalam perhitungan dengan algoritma k-means. Hasil yang didapatkan dari proses *clustering* bulan Januari 2011 telah membagi 2 *cluster* yaitu *cluster0* dan *cluster1* yang masing-masing bernilai 11 dan 471. Setelah didapatkan hasilnya, data dianalisa berdasarkan atribut yang dipakai. Sehingga didapatkan hasil bahwa *cluster0* merupakan obat dengan pemakaian cepat dan *cluster1* merupakan obat dengan pemakaian lama. Data akan diproses secara *clustering* setiap bulannya sehingga masing-masing data akan diketahui apakah obat tersebut termasuk *cluster* dengan pemakaian cepat atau *cluster* dengan pemakaian lama. Kemudian masing-masing *cluster* yang telah dibagi diberi penamaan sebagai label, yaitu label cepat dan lama. Setelah label ditambahkan, data dengan pemakaian cepat setiap bulannya, akan dipisahkan untuk dipakai pada metode persediaan obat nantinya. Pemisahan dilakukan per bulan selama 3 tahun, sehingga akan diketahui obat dengan pemakaian cepat pada bulan Januari 2011, Januari 2012 dan Januari 2013. Dari proses tersebut akan diketahui obat yang selalu mengalami *fast moving* setiap tahunnya pada bulan yang sama, dan mengetahui obat tersebut digunakan untuk penyakit apa, sehingga nantinya dapat digunakan untuk rekomendasi Puskesmas untuk melakukan tindakan seperti sosialisasi pada masyarakat tentang penyakit tersebut.

3.5 Fase Evaluasi

Pada tahap evaluasi model akan dinilai apakah hasil yang didapatkan dari proses *clustering* telah memenuhi tujuan yang telah ditentukan dalam tahap *business understanding*. Pada tahapan *business understanding* telah ditentukan tujuan yaitu untuk mengetahui obat dengan pemakaian cepat agar dapat digunakan untuk metode pengendalian persediaan pada Puskesmas Pandanaran. Dan hasil dari proses data mining dapat meng-*cluster* obat menjadi 2 *cluster* yaitu pemakaian cepat dan pemakaian lama. Setelah hasil yang didapatkan sama dengan tujuan pada *business understanding*, maka akan dilakukan proses pengecekan yang berfungsi untuk memastikan bahwa semua tahap telah dilakukan dalam proses pengolahan data atau tidak ada faktor penting yang terlewatkan. Kemudian dipastikan bahwa semua tahapan / faktor penting yang telah dilakukan dengan pengolahan data tidak ada yang terlewatkan. Dengan demikian dapat dilakukan proses selanjutnya pada proses pengolahan data karena telah memenuhi tujuan dari *data mining*. Selanjutnya pada tahap ini mempunyai 2 pilihan, yaitu kembali pada tahap awal atau melanjutkan ke tahap akhir. Karena pada tahap sebelumnya telah memenuhi tujuan dan tidak ada tahap yang terlewatkan, maka tahap selanjutnya adalah menuju tahapan akhir dengan menentukan persebaran dari hasil yang telah diperoleh dengan cara melakukan analisis.

3.6 Fase Penyebaran

Merupakan fase penerapan teknik klastering sesuai dengan tujuan / sasaran yang ingin dicapai pada fase pertama, yaitu untuk mengetahui obat dengan pemakaian cepat setiap bulannya yang diambil dari data 3 tahun. Sehingga nantinya dapat

digunakan sebagai metode persediaan obat pada Puskesmas Pandanaran. Pada fase penyebaran, akan digunakan model pengendalian persediaan yaitu Model Deterministik. Model deterministik ditandai oleh karakteristik permintaan dan periode kedatangan yang dapat diketahui secara pasti sebelumnya. Dalam model persediaan deterministik parameter-parameter yang berpengaruh terhadap sistem persediaan dapat diketahui dengan pasti. Rata-rata kebutuhan dan biaya-biaya persediaan diasumsi diketahui dengan pasti. Lamanya lead time juga diasumsikan selalu tetap. Karena semua parameter bersifat deterministik maka tidak dimungkinkan adanya kekurangan persediaan. Dalam dunia nyata, akan sangat jarang ditemukan situasi dimana seluruh parameter dapat diketahui dengan pasti. Karena itu, akan lebih masuk akal jika digunakan model-model probabilistik yang mempertimbangkan ketidakpastian pada parameter - parameter-nya. Namun, model deterministik terkadang merupakan pendekatan yang sangat baik, atau paling tidak merupakan langkah awal yang baik untuk menggambarkan fenomena persediaan. Dengan hasil *clustering* dapat digunakan Model Deterministik untuk pengendalian persediaan, yaitu dengan mengetahui karakteristik permintaan obat yang sering dipakai.

4. KESIMPULAN

Dari hasil penelitian dapat ditarik kesimpulan bahwa hasil dari metode *clustering data mining* dapat digunakan untuk metode pengendalian persediaan pada Puskesmas Pandanaran, sehingga apabila akan dilakukan pengadaan persediaan obat pada tahun 2014, petugas dapat melihat daftar obat dengan pemakaian cepat yang

merupakan hasil dari 3 tahun sebelumnya.

Dari hasil analisis data didapatkan hasil berupa jenis obat yang sering digunakan untuk resep atau pengobatan dari setiap bulannya dari sampel selama 3 tahun terakhir yaitu tahun 2011, 2012 dan 2013 yang dapat digunakan sebagai acuan persediaan obat pada tahun berikutnya. Selain itu terdapat pula obat yang setiap tahunnya menjadi obat yang sering digunakan untuk resep atau pengobatan, hal ini dapat digunakan untuk rekomendasi pada Puskesmas Pandanaran dalam persediaan obat untuk tahun atau bulan berikutnya.

Table 1: Hasil Clustering K-Means

Bulan	Tahun					
	2011		2012		2013	
	c0	c1	c0	c1	c0	c1
1	461	12	14	243	258	5
2	11	471	15	253	238	1
3	6	476	10	262	5	228
4	473	9	262	10	222	7
5	10	278	213	5	7	215
6	284	10	208	4	13	201
7	8	284	12	180	202	13
8	9	284	7	190	208	11
9	287	8	12	187	227	3
10	4	290	183	11	192	9
11	270	8	198	12	176	13
12	11	246	1	224	4	216

Bulan	Obat dengan pemakaian cepat			Obat yang ada di ketiga tahun
	Tahun 2011	Tahun 2012	Tahun 2013	
Mei	1. Amoksilin kapsul 500mg 2. Antalgin tablet 500 mg 3. Deksametason tablet 0,5 mg 4. Gliseril guayakolat tab 100mg 5. Klorafenikol kapsul 250mg 6. Klorfeniramin maleat tb 4 mg 7. Parasetamol tablet 500mg 8. Pridoksin HCL tablet 10mg 9. Tiamin HCL (Vit. B1) tab 50mg 10. Vitamin B kompleks tablet	1. Amoksilin kapsul 500mg 2. Antalgin tablet 500 mg 3. Klorfeniramin maleat tb 4 mg 4. Tiamin HCL (Vit. B1) tab 50mg 5. Vitamin B kompleks tablet	1. Eritromisin 250mg tb 4 mg 2. Klorfeniramin maleat tb 4 mg 3. Klorfeniramin maleat tb 4 mg 4. Pridoksin HCL tablet 10mg 5. Propranolol 10mg 6. Vitamin B Kompleks tablet	1. Klorfeniramin maleat tb 4 mg 2. Vitamin B kompleks tablet 3. Deksametason tablet 0,5 mg 4. Eritromisin 250mg 5. Gliseril guayakolat tab 100mg 6. Klorafenikol kapsul 250mg 7. Klorfeniramin maleat tb 4 mg 8. Parasetamol tablet 500mg 9. Pridoksin HCL tablet 10mg 10. Propranolol 10mg 11. Tiamin HCL (Vit. B1) tab 50mg 12. Vitamin B kompleks tablet

Gambar 2. Contoh pemakaian obat yang fast moving

DAFTAR PUSTAKA

[1] Notoatmodjo, Soekidjo. 2003. Metodologi Penelitian Kesehatan, Jakarta: Rineka Cipta.

[2] Depkes RI. 2004. Keputusan Menteri Kesehatan RI Nomor 128/MENKES/SK/II/

[3] Larose , Daniel T, 2005, *Discovering Knowledge in Data: An Introduction to Data Mining*, John Willey & Sons. Inc

[4] Larose D, T., 2006, *Data Mining Methods and Models*, Jhon Wiley & Sons, Inc. Hoboken New Jersey.

[5] Garcia - Molina, Hector; Ullman, JD., & Widom, Jennifer. 2002. *Database systems the complete book, International edition*. New Jersey, Prentice Hall.