

## Penerapan Algoritma K-Means Untuk Klasterisasi Data Obat Pada Rumah Sakit ASRI

Muhamad Rizki Nugroho<sup>1</sup>, Iwansyah Edo Hendrawan<sup>2</sup>, Puwantoro<sup>3</sup>

<sup>1,2,3</sup>Teknik Informatika Universitas Singa Perbangsa Karawang

E-mail: <sup>1</sup>[muhamad.rizki18224@student.unsika.ac.id](mailto:muhamad.rizki18224@student.unsika.ac.id),

<sup>2</sup>[iwansyah.edo18106@student.unsika.ac.id](mailto:iwansyah.edo18106@student.unsika.ac.id), <sup>3</sup>[purwantoro.masbro@staff.unsika.ac.id](mailto:purwantoro.masbro@staff.unsika.ac.id)

### Abstrak

*Pengelolaan obat merupakan salah satu hal yang sangat dibutuhkan dalam tujuan mengelola stok obat. Obat perlu dikelola secara baik, efektif dan efisien. Dengan pengelolaan obat yang baik maka obat dapat diperoleh dengan cepat dan tepat serta mengurangi hal-hal seperti kehabisan stock obat pada layanan kesehatan seperti Puskesmas, Rumah Sakit dan lain-lain. Hasil wawancara dengan salah satu pegawai yang mengurus data obat di Rumah Sakit Asri Purwakarta di rumah sakit tersebut sering terjadi kekurangan atau kelebihan obat walaupun jumlahnya tidak terlalu banyak. Clustering merupakan salah satu opsi yang dapat digunakan dalam pengelolaan obat sebab nantinya sistem cluster dapat membuat pengelompokan pada obat dengan pemakaian tinggi dan kurang sehingga nantinya dapat menjadi acuan atau knowledge base dalam pengambilan keputusan untuk mengatur stok obat. Algoritma K-means merupakan salah satu algoritma dalam clustering. Penggunaan Algoritma K-means dalam penelitian adalah karena kesederhanaan dan efisiensinya sehingga mudah diterapkan disegala bidang contohnya pada klasterisasi data obat. Hasil penelitian ini membagi data obat menjadi 2 cluster yaitu cluster pertama dengan pemakaian tinggi dengan beranggotakan 6 obat dan cluster kedua dengan pemakaian rendah yang beranggotakan 933 obat.*

**Kata Kunci**—Data Mining, Clustering, K-Means, Data Obat

### Abstract

*Drug management is needed to manage drug stocks. Drugs need to be managed properly, effectively, and efficiently. Through good drug management, the drugs can be obtained quickly and accurately and reduce bad possibilities such as running out of drug stock in health services such as Puskesmas, Hospitals, and others. The results of an interview with one of the employees who manage drug data at the Asri Purwakarta Hospital shows that at the hospital often have drug shortages or excess even though the drug amounts are not too many. Grouping or clustering is one of the best options that can be used in drug management system because this cluster system can classify the most frequently used drugs and it can become a reference or knowledge based in making decisions to manage the drugs. K-means algorithm is one of the algorithms in clustering that is used in this drug classification research. K-means algorithm is used in the research because of its simplicity and efficiency so it is easy to apply in all fields, especially drug data classification. The results of this study divided the drug data into 2 clusters, the first cluster with high usage there are 6 drugs and the second cluster with low usage with 933 drugs.*

**Keywords**—Data Mining, Clustering, K-Means, Drugs Data

## 1. PENDAHULUAN

Kesehatan merupakan hal yang penting bagi keberlangsungan hidup manusia. Kesehatan juga menjadi salah satu hak asasi manusia. Dengan berkembangnya

ilmu kesehatan maka keberlangsungan hidup manusia akan lebih terjamin. Kesehatan juga digunakan sebagai salah satu indikator apakah sebuah negara sudah maju atau masi berkembang. Kesehatan merupakan investasi untuk membangun

suatu bangsa. Oleh karena itu kesehatan perlu ditingkatkan lagi baik dari segi masyarakat atau dari layanan kesehatan agar terwujudnya negara dengan tingkat kesehatan yang tinggi. Salah satu cara untuk meningkatkan kesehatan disuatu negara yaitu dengan meningkatkan pelayanan kesehatan itu sendiri. Karena disitulah masyarakat akan berkonsultasi mengenai kesehatan, dengan meningkatnya layanan kesehatan di suatu daerah maka kesehatan masyarakat di daerah tersebut juga akan meningkat. Salah satu cara tersebut adalah dengan melakukan manajemen obat yang baik.

Obat merupakan salah satu hal yang penting didunia kesehatan. Obat digunakan untuk mencegah penyakit, mengurangi rasa sakit dan mengobati penyakit. Maka oleh sebab itu obat perlu dikelola secara baik, efisien dan efektif. Dengan pengelolaan obat yang baik maka obat akan dapat diperoleh secara cepat dan juga tepat pada saat obat tersebut dibutuhkan. Dengan pengelolaan yang baik juga dapat mengurangi hal-hal seperti kehabisan stok obat pada layanan kesehatan seperti rumah sakit, puskesmas dan lain-lain, serta pendistribusian, pembelian, dan pemakaian obat pada layanan kesehatan tersebut akan lebih efisien lagi daripada sebelumnya sehingga perekonomian pada layanan kesehatan tersebut juga akan ikut meningkat.

RSIA Asri Purwakarta adalah Rumah Sakit Ibu dan Anak yang terletak di wilayah Babakancikao, Kabupaten Purwakarta. RSIA Asri Purwakarta memiliki beberapa poliklinik seperti Poliklinik Gizi, Kandungan & Kebidanan, Anak, Penyakit Dalam dan lain sebagainya guna menunjang kebutuhan akan layanan kesehatan yang baik bagi masyarakat Purwakarta. Rumah Sakit Ibu dan Anak Asri Purwakarta ini sudah lama beroperasi dan juga sudah menerapkan sistem-sistem yang saling terintegrasi untuk mengelola data-data pada rumah sakit tersebut termasuk data-data obat. Namun berdasarkan hasil wawancara

dengan salah satu pegawai di rumah sakit tersebut yang bertugas untuk mengelola data-data obat disana mengatakan bahwa terkadang persediaan atau stok obat pada Rumah Sakit Ibu dan Anak Asri Purwakarta sering kali terjadi kekurangan atau kelebihan stok obat walaupun jumlahnya tidak terlalu banyak. Dikarenakan pada rumah sakit tersebut belum menerapkan Data mining maka pada penelitian ini akan dilakukan Data mining pada data-data obat yang telah dikelola tersebut. Data mining adalah suatu proses untuk mencari suatu pola atau suatu model yang sempurna, bermanfaat serta dapat dimengerti dari sebuah *database*. Pola atau informasi yang didapatkan dari *database* dapat digunakan untuk membantu dalam mengambil suatu keputusan di waktu yang akan datang[1].

Berdasarkan penelitian yang dilakukan oleh Gustientiedina, Hasmil dan Yenny bahwa teknik *clustering* bisa digunakan untuk mengelompokan data obat sehingga dapat diperoleh informasi mengenai obat mana yang memiliki pemakaian yang tinggi dan obat mana yang memiliki pemakaian yang rendah dan sedang. Hasil penelitian ini didapatkan 3 kelompok obat yaitu obat dengan pemakaian rendah dimana rata-rata permintaan obat kurang dari 1800 buah setiap tahunnya, kemudian obat dengan pemakaian sedang yaitu dengan rata-rata permintaan obat diantara 1800-7000 buah setiap tahunnya, dan obat dengan pemakaian tinggi yaitu dengan rata-rata permintaan diatas 7000 buah setiap tahunnya [2]. Penelitian yang sama juga dilakukan oleh Mei didalam penelitian tersebut Mei menggunakan teknik *clustering* dan algoritma K-means untuk Klasterisasi data pemakaian obat pada Dinas Puskesmas Teladan Medan. Hasil penelitian tersebut mengelompokan data obat menjadi 3 *cluster* yaitu *cluster* 1 dimana didalamnya terdapat 1 obat, kemudian *cluster* 2 dimana didalamnya terdapat 18 obat, dan *cluster* 3 atau *cluster* terakhir dimana didalamnya terdapat 20 obat[3]. Berdasarkan penelitian-penelitian tersebut maka dalam penelitian ini akan dilakukan penerapan salah satu teknik data mining yaitu *clustering* terhadap data

obat Rumah Sakit Asri Purwakarta dengan menggunakan algoritma K-means.

Klasterisasi penggunaan obat pada Rumah Sakit Asri purwakarta dapat menjadi salah satu sumber dalam mengambil keputusan dan mengatur persediaan obat. *Clustering* adalah salah satu metode dalam data mining yang dimana objek data yang mempunyai kemiripan atau karakteristik yang sama akan dikelompokkan menjadi satu kelompok dan yang berbeda di kelompokkan pada kelompok yang lainnya[4]. Dengan Hasil Klasterisasi data penggunaan obat ini pada Rumah Sakit Asri Purwakarta dapat dikelompokkan sesuai dengan kebutuhan pada bulan-bulan sebelumnya dan bisa dijadikan acuan untuk pengelolaan obat pada bulan-bulan berikutnya. Sehingga dengan hasil klasterisasi ini pengelolaan obat untuk bulan-bulan berikutnya menjadi lebih terjamin dan lebih baik.

Algoritma yang terkenal dalam metode *clustering* adalah K-means karena algoritmanya yang cukup sederhana dan efisien[5]. Algoritma K-means merupakan algoritma yang masuk kedalam Algoritma *unsupervised learning* sehingga didalam algoritma ini tidak perlu dilakukan proses *training* atau dengan kata lain didalam Algoritma K-means ini tidak ada fase *learning*. Algoritma K-means ini digunakan karena kesederhanaan algoritma tersebut dapat diaplikasikan pada berbagai bidang. Algoritma ini juga diakui sebagai salah satu algoritma teratas oleh IEE[6] dan berdasarkan review yang ditulis amit saxena dan kawan-kawan Algoritma ini adalah algoritma yang paling sering digunakan untuk menyelesaikan masalah klasterisasi[7].

## 2. METODE PENELITIAN

Objek Penelitian ini menggunakan data yang diperoleh dari Rumah Sakit Ibu dan Anak Asri Purwakarta yaitu berbentuk laporan stok obat dari mulai bulan januari sampai bulan juni 2021. Namun data bulan

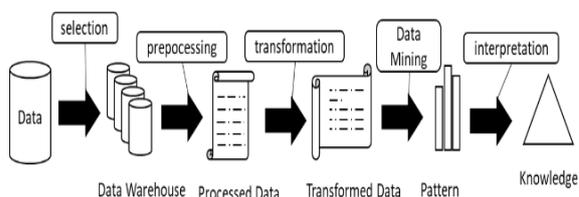
mei tidak bisa diakses hal tersebut dikarenakan adanya kesalahan pada sistem Rumah Sakit Asri pada saat pengambilan data.

### 2.1. Data Mining

Data mining adalah suatu proses analisis dari sekumpulan data untuk menemukan informasi, pola, atau hubungan antar data yang sebelumnya tidak diketahui dan meringkas data dengan cara yang berbeda, serta dapat mudah dipahami bagi pemilik data[8].

### 2.2. Knowledge Discovery In Database (KDD)

KDD adalah salah satu metode dalam data mining yang digunakan untuk mencari informasi dari data yang sebelumnya tidak diketahui serta menemukan pola tersembunyi yang terdapat di dalam data[9]. Informasi yang diperoleh dapat dijadikan salah satu basis pengetahuan (*knowledge base*) yang digunakan dalam mengambil suatu keputusan. Adapun Langkah-langkah KDD secara detail sebagai Berikut [10] :



Gambar 1. Tahapan Proses KDD

#### 1. Data Selection

Menyeleksi Data dari sekumpulan data perlu dilakukan terlebih dahulu sebelum tahap penggalian informasi dalam KDD dimulai. Data yang telah terseleksi disimpan dalam berkas yang terpisah.

#### 2. Pre-processing/cleaning

Proses *cleaning* artinya pembersihan data yaitu membuang duplikasi data, memeriksa data yang

tidak konsisten, serta memperbaiki kesalahan data. Proses ini penting dilakukan sebelum melanjutkan ke tahap data mining.

### 3. Transformation

*Coding* adalah proses transformasi pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses data mining. Proses coding dalam KDD merupakan proses kreatif dan sangat tergantung pada jenis atau pola informasi yang akan dicari dalam basis data.

### 4. Data Mining

Data mining adalah proses mencari pola atau informasi yang tersembunyi dari data dengan menggunakan teknik atau metode tertentu. Metode dan teknik yang digunakan akan sangat berpengaruh terhadap hasil informasi yang akan dicari dalam basis data.

### 5. Interpretation

Pola yang telah diperoleh dari proses data mining perlu ditampilkan dalam bentuk yang dapat dimengerti oleh pihak yang berkepentingan. Tahap tersebut disebut dengan *interpretation*. Tahap ini mencakup pemeriksaan apakah pola atau informasi yang didapat sesuai dengan hipotesis atau fakta yang ada sebelumnya.

## 2.3. Clustering

*Clustering* termasuk kedalam salah satu teknik *unsupervised learning* dimana di dalamnya tidak perlu dilakukan pelatihan model atau bisa juga dikatakan dengan tidak adanya fase *learning*.

*Clustering* merupakan sebuah proses yang digunakan untuk membagi data kedalam beberapa kelas atau *cluster* berdasarkan tingkat kemiripan atau karakteristik data tersebut. Data-data yang

memiliki kemiripan akan berkumpul atau dikelompokkan di satu *cluster* yang sama. Sedangkan data-data yang memiliki karakteristik yang berbeda akan berkumpul atau dikelompokkan pada *cluster* yang berbeda. Tujuan dari Klasterisasi ini adalah pengelompokan data atau objek kedalam sebuah kelompok sehingga setiap kelompok akan berisi data-data yang memiliki karakteristik yang sangat mirip[11].

### 2.4. K-Means

K-means adalah salah satu algoritma pengelompokan yang bekerja berdasarkan prinsip *partitioned clustering*. Prinsip kerja dari *partitioned clustering* yaitu mengelompokkan item secara acak karena dipengaruhi centroid. Dan disetiap iterasi dari pengelompokan *partitioned clustering* dapat memungkinkan untuk terjadinya lebih dari satu pemilihan item yang akan digabungkan. Sedangkan Prinsip kerja dari pengelompokan *hierarchical clustering* dilakukan secara bertahap. Dan disetiap iterasi dari pengelompokan *hierarchical clustering* hanya akan ada satu pemilihan penggabungan suatu item terhadap item lainnya[11].

Tahapan-tahapan dalam melakukan clustering menggunakan algoritma K means sebagai berikut [12]:

1. Tentukan Jumlah *cluster* k.
2. Inisialisasi K pusat *cluster*. Ini dapat dilakukan dengan berbagai cara. Namun cara yang paling banyak digunakan adalah dengan menentukan secara random. Pusat-pusat *cluster* diberi nilai awal dengan angka-angka random.
3. Alokasikan semua data/objek ke dalam *cluster* terdekat. Kedekatan dua data/objek dapat didapatkan berdasarkan jarak kedua objek tersebut. Untuk dapat menghitung jarak semua data terhadap titik pusat *cluster* maka dapat menggunakan teori jarak Euclidean yang dirumuskan dibawah ini :

$$D(i,j) = \sqrt{(x_{ii} - x_{ij})^2 + (x_{2i} - x_{2j})^2 + \dots + (x_{ki} - x_{kj})^2}$$

(1)

dimana:

$D(i,j)$  = Jarak data ke i ke pusat cluster j

$X_{ki}$  = Data ke i atribut data ke k

$X_{kj}$  = Titik pusat ke j pada atribut ke k

- Hitung kembali pusat *cluster* dengan keanggotaan *cluster* yang sekarang. Pusat *cluster* adalah rata-rata dari semua objek/data dalam *cluster* tertentu. Jika dikehendaki bisa juga menggunakan median dari *cluster* tersebut. Jadi rata-rata (mean) bukan satu-satunya ukuran yang bisa dipakai.

$$R_k = \frac{1}{N_k} (X_{1k} + X_{2k} + \dots + X_{nk})$$

(2)

dimana :

$R_k$  = Rata-rata baru.

$N_k$  = Jumlah training pattern pada cluster (k).

$X_{nk}$  = Pola ke (n) yang menjadi bagian cluster (k).

- Tugaskan lagi tiap objek memakai pusat cluster yang baru. Jika pusat cluster tidak berubah lagi maka proses clustering selesai. Atau, kembali ke langkah nomor 3.

Anak Asri pada saat pengumpulan data, sehingga data laporan yang diperoleh sebanyak 5 laporan data obat bulanan. Data ini dipilih karena merupakan data yang paling terbaru yang bisa diambil dari sistem Rumah Sakit Ibu dan Anak Asri. Data Laporan Obat tersebut tercampur dengan data Alat Kesehatan. Didalamnya terdapat 942 obat dan 212 alat kesehatan. Berikut salah satu data tersebut di bulan Januari :

Report ID : ListStock C  
Sistem : Inventory  
Tanggal : 26-02-2021  
Belahan : 1

**LIST STOCK OPNAME**  
Gudang/Periode : APT / Jan-21

Rumah Sakit ASRI  
Jl. Veteran Sedang No.15 Purwakarta  
Telp. (0264) 211680 - 219447

Tipe Item : ALKES

Item	Nama Material	Satuan	Harga	Qty Stock	Qty Masuk	Qty Mutasi	Qty Saldo	Stok Minimal	Stok Maksimal	Stok Saat Ini	Stok Akhir	Wilayah Akhir
31826	HYPARIX	BOX	165,000	9	10	31,52	-12,52			47	47	150,100
324E1	UNIBILICAL KLEN	BOX	3,300	125	200	312	13			71	71	859,100
33E02	URINE BAG (AMB)	PCS	12,100	67	180	231	-4			16	69	2,301,816
AR026	ANOCATH 18 (PARET)	PCS	36,258	0	50	34				110	110	1,210,000
AR018	ANOCATH 18 (PARET)	PCS	11,000	180	368	466	62			220	220	2,420,000
AR020	ANOCATH 20 (PARET)	PCS	11,000	287	310	23				332	332	3,632,000
AR024	ANOCATH 24 (PARET)	PCS	11,000	146	150	214	77			80	80	880,000
AR024	ANOCATH 24 (PARET)	PCS	11,000	76	450	580	-54			0	0	0
ACHAB	ACH CHECK LAB	PCS	984,700	0	0	0	0			0	0	0
ACHAF	ACH SAFETY LAB	PCS	303,400	0	0	0	0			0	0	0
ALH20	ALVEIN	PCS	22,000	0	0	0	0			0	0	0
ALH10	ALKOHOL 70% 100 CC	PAK	4,583	14	240	319	-65			68	68	311,666
ALH1	ALKOHOL 70% 1 LTR	LTR	33,000	10	20	22,86	7,14			14	14	462,000
ALHWA	ALKOHOL SWAB	PAK	1,100	4500	7600	3931	8169			2100	2100	2,310,000
ANICEL	ANIOSE GEL 5 1	GALL	1,157,970	0	0	0	0			0	0	0
ANIGEL	ANIOSE GEL 500ML	BTL	188,760	0	0	0	0			0	0	0
ANIO17	ANIOSE DDI	GALL	2,236,988	1	2	28,6	-25,6			0	0	0
AP020	AP0 1 SET	PCS	452,100	1	6	1				6	6	632,500
AP020C	AP020 DISPOSIBLE	PCS	5,500	127	300	689	-262			115	115	110,000
AN00	ANM SLING	PCS	22,000	5	0	0	5			5	5	121,000
AS040	ASAM AUTOCHER	PCS	151,800	0	0	0	0			0	0	0
AS030	ASEPTIC 300ML	LTR	24,200	0	0	0	0			0	0	0
AS01	ASEPTIC 5L	LTR	207,515	0	4	0	4			4	4	830,000
AS01 B	AS01 5000	PCS	14,500	80	0	0	80			30	30	499,000
BR02E	BR02 CELANA	PCS	9,200	34	0	0	34			34	34	336,600

Gambar 2. Laporan Data Obat RSUD Asri januari

### 3. HASIL DAN PEMBAHASAN

#### 3.1. Data Selection

Data yang digunakan pada penelitian ini adalah data obat dari Rumah Sakit Ibu dan Anak Asri Purwakarta yang berbentuk Laporan bulanan dalam kurun waktu 1 bulan dari bulan januari hingga Juni 2021. Namun data bulan Mei tidak dapat diakses karena terdapat kesalahan pada sistem Rumah Sakit Ibu dan

#### 3.2. Preprocessing

Dari laporan data obat yang diperoleh kemudian dilakukan perubahan format kedalam bentuk excel dan pemisahan dengan data alat kesehatan serta dilakukannya data cleaning untuk membuang atribut-atribut yang tidak relevan atau tidak konsisten. Atribut yang dibuang adalah Item, Satuan, Harga, Qty stock, Qty

Masuk, Stock Sistem, Stock Manual, Stock Akhir, Nilai Akhir. Atribut yang akan digunakan yaitu Nama Obat, Qty Keluar. Setelah dilakukan cleaning juga dilakukan penggabungan data dari laporan bulan januari hingga bulan juni dan dilakukan pengecekan apakah ada redundansi data. Setelah selesai diproses maka didapat data obat sebanyak 939 item dan 1 data obat yang telah dibuang karena muncul 2 kali dan pembuangan data obat embalase 1 dan 2 karena embalase bukan termasuk obat. Data obat yang telah di *preprocessing* dapat dilihat pada tabel berikut :

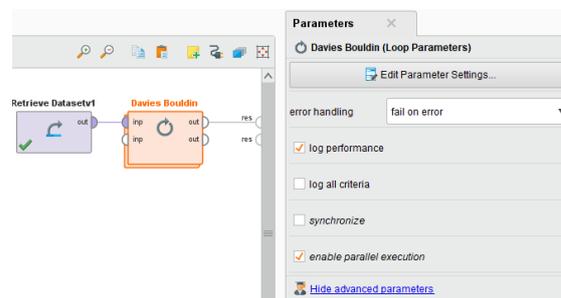
No	Nama	Jan	feb	...	Jun
1	ACARBOSE 100	7	0	...	120
2	ACETYLCYSTE IN	469	387	...	2023
3	ACRAN INJ	14	8	...	15
4	ACYCLOVIR 200	0	34	...	6
5	ACYCLOVIR 400	0	0	...	203
6	ACYCLOVIR 5%	0	2	...	7
7	ADONA AC INJ	45	30	...	15
8	AKILEN	0	2	...	3
9	ALBUMIN BEHRING 20%	0	0	...	0
10	100ML ALCO PLUS SYR	7	5	...	14
11	ALCO DROP	3	3	...	18
12	ALGANAX 1MG	7	0	...	2
13	ALGANAX 0.5MG	0	20	...	0
14	ALINAMIN F	0	0	...	0
15	ALKALIN	0	0	...	0
16	ALKAZIM	0	0	...	0
17	ALLOCLAIR GARGLE	0	0	...	0
18	ALLOPURINOL 100	0	0	...	0
19	ALPRAZOLAM 0,5 MG	47	31	...	161
20	AMLODHIPIN 10MG	406	734	...	1068
...	...	..	..	...	..
...	...	..	..	...	..
938	ZOVIRAX INJ	0	0	...	0
939	ZYRCUM SYR	72	14	...	30

Tabel 1. Data Pemakaian Obat yang sudah di *preprocessing*

### 3.3. Data Mining

Setelah data selesai di cleaning maka bisa dilanjut dengan proses data mining. Proses Data mining menggunakan teknik *clustering* dengan algoritma K-means dan perangkat lunak rapid miner.

Sebelum dilakukan *Clustering* dilakukan uji performance terhadap K-means pada data obat yang sudah di cleaning untuk mencari nilai K terbaik.



Gambar 3. Pengujian Performance Clustering K-means

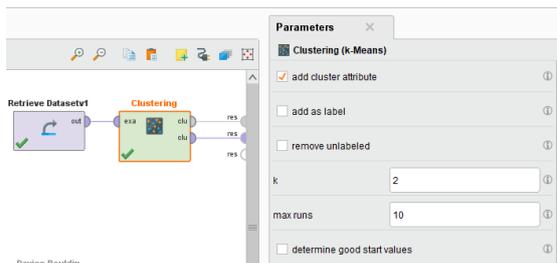
Pengujian dilakukan dengan parameter Davies Bouldin. Nilai Index Davies Bouldin (IDB) yang paling rendah maka menunjukkan jumlah cluster paling optimal. Dari hasil pengujian *performance* tersebut maka didapat hasil sebagai berikut :

Davies Bouldin (9 rows, 3 columns)

iteration	K means ...	Davies Bouldin
1	2	0.317
3	4	0.512
2	3	0.427
5	6	0.663
7	8	0.711
4	5	0.623
6	7	0.820
9	10	0.660
8	9	0.635

Gambar 4. Hasil Pengujian Performance K-Means.

Dari hasil tersebut maka dapat dipilih nilai K dengan index davies bouldin terkecil adalah dengan nilai K sebanyak 2, maka dipilihlah K dengan nilai 2. Setelah nilai K didapatkan maka dilanjutkan dengan clusterisasi menggunakan algoritma K-means dengan nilai K sebanyak 2.



Gambar 4. Proses Clustering K-Means dengan rapidminer.

Setelah proses klasterisasi selesai maka didapatkan hasil sebagai berikut :

## Cluster Model

Cluster 0: 933 items  
Cluster 1: 6 items  
Total number of items: 939

Gambar 5. Hasil dari Klasterisasi

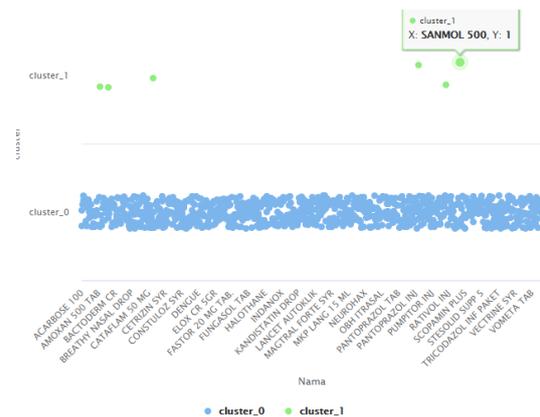
Dan nilai titik pusat cluster yang digunakan sebagai berikut :

Attribute	cluster_0	cluster_1
Jan	38.447	2807.833
Feb	47.989	2999.333
Maret	62.416	3341.667
April	51.167	3056.500
juni	67.633	3579.167

Gambar 6. Nilai Centroid

### 3.4. Interpretation

Hasil yang diperoleh dari proses klasterisasi kemudian divisualisasikan dalam bentuk diagram scatter sebagai berikut :



Gambar 7. Visualisasi data hasil klasterisasi

Dari data tersebut maka dapat dilihat bahwa Cluster 1, merupakan cluster obat yang pemakaiannya paling tinggi dan didalamnya terdapat 6 obat, kemudian Cluster 0 beranggotakan 933 obat yang merupakan cluster obat dengan pemakaian yang rendah.

No	Nama	Jan	feb	...	Jun
1	AMOXICILL IN 500 MG	2167	2064	...	2196
2	ASAM MEFENAM AT 500	4178	4554	...	4761
3	CEFADROX IL 500	2708	3271	...	2962
4	PRENAMIA	3684	3744	...	4328
5	RL SANBE PAKET	2569	2810	...	3630
6	SANMOL 500	1541	1553	...	3598

Tabel 2. Anggota Cluster 1

No	Nama	Jan	feb	...	Jun
----	------	-----	-----	-----	-----

1	ACARBOSE 100	7	0	...	120
2	ACETYLCYS TEIN	469	387	...	2023
3	ACRAN INJ	14	8	...	15
4	ACYCLOVIR 200	0	34	...	6
5	ACYCLOVIR 400	0	0	...	203
6	ACYCLOVIR 5%	0	2	...	7
7	ADONA AC INJ	45	30	...	15
8	AKILEN	0	2	...	3
9	ALBUMIN BEHRING 20%	0	0	...	0
10	100ML ALCO PLUS SYR	7	5	...	14
11	ALCO DROP	3	3	...	18
12	ALGANAX 1MG	7	0	...	2
13	ALGANAX 0.5MG	0	20	...	0
14	ALINAMIN F	0	0	...	0
15	ALKALIN	0	0	...	0
16	ALKAZIM	0	0	...	0
17	ALLOCLAIR GARGLE	0	0	...	0
18	ALLOPURIN OL 100	0	0	...	0
19	ALPRAZOLA M 0,5 MG	47	31	...	161
20	AMLODHIPI N 10MG	406	734	...	1068
...	...	..	..	...	..
...	...	..	..	...	..
932	ZOVIRAX INJ	0	0	...	0
933	ZYRCUM SYR	72	14	...	30

Tabel 3. Anggota Cluster 0

#### 4. KESIMPULAN

Data yang digunakan adalah data obat Rumah Sakit Asri Ibu dan Anak Purwakarta dengan kurun waktu bulan januari hingga bulan juni. Namun data bulan mei tidak dapat diakses saat pengumpulan data karena terjadi kesalahan pada sistem Rumah Sakit Ibu dan Anak Asri, sehingga data obat yang diperoleh sebanyak 5 bulan. Setelah data didapatkan kemudian dilakukan proses *cleaning*, lalu Dilakukan proses pemodelan atau data mining terhadap data

tersebut, dengan teknik klasterisasi dengan menggunakan Algoritma K-means dan perangkat lunak Rapidminer.

Dari hasil klasterisasi menggunakan algoritma K-means ini maka dapat disimpulkan bahwa pada *cluster* 1 terdapat 6 obat yang merupakan cluster obat dengan pemakaian yang tinggi yaitu dengan rata-rata pemakaian diatas 2.046 item, dan *cluster* 2 yang beranggotakan 933 obat dengan pemakaian rata-rata dibawah 2.046 item. Hasil klasterisasi ini dapat menjadi salah satu pertimbangan atau *knowledge base* dalam manajemen obat pada Rumah Sakit Asri untuk bulan-bulan berikutnya.

#### 5. SARAN

Kekurangan pada penelitian ini adalah data obat yang digunakan pada penelitian ini hanya 5 bulan dan atribut yang dipakai hanya nama obat dan pemakaian obat, Namun pada penelitian ini telah menetapkan nilai K terbaik dengan pengujian *performance* algoritma K-means dengan Index Davies Bouldin (IDB) terhadap data obat sehingga hasil dari klasterisasi dapat maksimal.

Data obat yang digunakan mungkin bisa lebih banyak lagi sehingga klasterisasi dapat semakin akurat, bisa juga ditambahkan atribut lain dan juga dengan metode yang lebih baik dalam menentukan nilai k pada penerapan algoritma K-means. Serta juga bisa dilakukan komparasi dengan algoritma lain sehingga nanti dapat terlihat hasil klasterisasi mana yang lebih baik antara algoritma yang dikomparasikan.

#### DAFTAR PUSTAKA

- [1] Al Syahdan S. dan Sindar A. 2018, Data Mining Penjualan Produk Dengan Metode Apriori Pada Indomaret Galang Kota. *Jurnal Nasional Komputasi dan Teknologi Informasi*, Vol. 1, No. 2, Hal 56-63.

- [2] Gustientiedina, Adiya M.H, dan Desnelita Y. 2018, Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan Pada RSUD Pekanbaru. *Jurnal Nasional teknologi dan Sistem Informasi*. Vol. 05, No. 01, hal 017-024.
- [3] Pilida Mei T. 2021. Penerapan Data Mining Dalam Analisa Data Pemakaian Obat Dengan Menerapkan Algoritma K-Means. *Jurnal Informasi dan Teknologi Ilmiah(INTI)*. Vol. 8, No. 3, hal 109-113.
- [4] A. A. Aldino dan H. Sulistiani. 2020, Decision Tree C4. 5 Algorithm For Tuition Aid Grant Program Classification (Case Study: Department Of Information System, Universitas Teknokrat Indonesia, *Edutic-Scientific J. Informatic Educ.*, No.1, Vol. 7, hal 40-50.
- [5] T. H. Sardar dan Z Ansari. 2018, An analysis of MapReduce efficiency in document clustering using parallel K-means algorithm, *Future Computing and informatics Journal*. No.2, Vol.3, hal 201-209.
- [6] X. Wu, et al. 2008, Top Algorithms in data mining, *knowl. Inf. Syst*. Vol. 14, hal 1-37.
- [7] Saxena, A, Prasad M, Gupta A, Bharill N, Patel O. P, Tiwari A, Joo Er M, Ding W. dan Lin C. 2017, A review of clustering techniques and developments. *Neorocomputing*. Vol.267, hal 664-681.
- [8] Utomo, D. P. dan Mesran.(2020). Analisis Komparasi Metode Klasifikasi Data Mining dan Reduksi Atribut Pada Data Set Penyakit Jantung. *Jurnal Media Informatika Budidarma*. Vol. 4, No. 2, Hal 437-444.
- [9] Safhi, H. M., Frikh, B., & Ouhi.(2019), Assesing realibility of Big Data Knowledge Discovery  
process.*Procedia Computer Science* Vol. 148, hal 30-36.
- [10] A.P Windarto. 2017, Penerapan Data Mining Pada Ekspot Buah-Buahan Menurut Negara Tujuan Menggunakan K-Means Clustering. *Techno.COM*, Vol. 16, No. 4, Hal. 348-357.
- [11] Y. Siyamto.2017, Pemanfaatan Data Mining Dengan Metode Clustering Untuk Evaluasi Biaya Dokumen Ekspor Di PT Winstar Batam. *Jurnal Media Informatika Budidarma*. Vol. 1, No.2, Hal
- [12] Fatmawati K., dan Windarto A., P. 2018, Data Mining : Penerapan rapidminer dengan K-Means Cluster Pada Daerah Terjangkit Demam Berdarah Dengue (DBD) Berdasarkan Provinsi. *Journal of Compter Engineering System and Science*. Vol. 3, No. 2, hal 173-178.