

Penerapan Algoritma K-Means Untuk Mengklasifikasi Data Obat

Ferdy Pangestu^[1], Nur Yasin^[2], Ronald Christover Hasugian^[3], Yunita^{[4]*}

Program Studi Informatika, Fakultas Teknologi Informasi

Universitas Nusa Mandiri, Jln. Jatiwaringin Raya No.02 RT08 RW 013 Kelurahan Cipinang Melayu Kecamatan Makassar Jakarta Timur 13620

ferdypanges357@gmail.com^[1], yasinnur782@gmail.com^[2], ronalarares816@gmail.com^[3], yunita.yut@nusamandiri.ac.id^[4]

Abstract— Classification of drug data in an agency engaged in the health sector is very important. This activity cannot be separated from supervision and monitoring every day because the processing of drug data is the core of the wheels of business in an agency or company. This research was conducted to discuss the importance of data processing in an agency or company, namely the Sawah Besar District Health Center. Currently the problem that occurs is the absence of a method in classifying drug data. Drugs have not been grouped according to their characteristics so that it becomes an obstacle when searching for data and when checking stock. With the above problem we need a method. One is by using the K-means algorithm method. K-means clustering is a non-hierarchical cluster analysis method that seeks to partition existing objects into one or more clusters based on their characteristics. These drug data are classified into 4 categories, namely over-the-counter drugs, limited over-the-counter drugs, hard drugs, and narcotics & psychotropics. The results of the research that has been done are that by using the K-Means algorithm the Sawah Besar District Health Center can classify drug data with high, moderate to low use levels based on the volume of use and intake. The data taken from the Sawah Besar District Health Center is drug stock data in December 2021. Manual calculations get an accuracy of 66.23% for cluster 1, 7.69% for cluster 2, and 23.07% for cluster 3 on over-the-counter drug data limited.

Keywords— K-Means, Clustering, Drugs

Abstrak— Pengklasifikasian data obat pada sebuah instansi yang bergerak pada bidang Kesehatan merupakan hal yang sangat penting. Kegiatan tersebut tidak lepas dari pengawasan serta monitoring setiap harinya karena pengolahan data obat termasuk inti dalam berjalannya roda bisnis pada suatu instansi atau perusahaan. Penelitian ini dilakukan untuk membahas pentingnya pengolahan data dalam suatu instansi atau perusahaan yaitu pada Puskesmas Kecamatan Sawah Besar. Saat ini permasalahan yang terjadi adalah tidak adanya suatu metode dalam pengklasifikasian data obat. Obat-obatan belum dikelompokkan sesuai dengan karakteristiknya sehingga menjadi kendala saat pencarian data serta saat pengecekan stok. Dengan masalah di atas dibutuhkan suatu metode. Salah satu dengan menggunakan metode algoritma K-means. K-means clustering merupakan salah satu metode cluster analysis non hirarki yang berusaha untuk mempartisi objek yang ada kedalam satu atau lebih cluster berdasarkan karakteristiknya data-data obat ini diklasifikasikan menjadi 4 kategori yaitu obat bebas, obat bebas terbatas, obat keras, dan narkotika & psikotropika. Hasil penelitian yang sudah dilakukan bahwa dengan menggunakan

algoritma K-Means pihak Puskesmas Kecamatan Sawah Besar dapat mengelompokkan data obat yang tingkat pemakaiannya tinggi, sedang hingga terendah berdasarkan volume pemakaian dan pemasukan. Data yang diambil dari Puskesmas Kecamatan Sawah Besar merupakan data stok obat pada bulan Desember 2021. Perhitungan manual mendapatkan akurasi yaitu 66,23% untuk cluster 1, 7,69% untuk cluster 2, dan 23,07% untuk cluster 3 pada data obat bebas terbatas.

Kata Kunci—K-Means, Clustering, Obat

I. PENDAHULUAN

Diera perkembangan teknologi yang sangat pesat ini, banyak dari instansi atau perusahaan yang menggunakannya guna membantu memecahkan permasalahan yang sedang dialami oleh instansi atau perusahaan tersebut. Salah satu kemajuan dari teknologi informasi tersebut yaitu pada bidang kesehatan. Sudah banyak dari instansi atau perusahaan yang bergerak pada bidang kesehatan yang memanfaatkan peranan teknologi informasi guna menunjang kegiatan yang dilakukan dalam instansi atau perusahaan tersebut. Salah satu dari pemanfaatan teknologi informasi tersebut yakni untuk mendiagnosa sebuah penyakit yang diderita dan pemberian obat yang sesuai dengan penyakit yang diderita oleh pasien. Pemanfaatan media informasi sangat berpengaruh pada penyampaian pesan kesehatan khususnya pada pemberian obat [1].

Pengolahan data yang kurang baik sudah pasti akan memberikan dampak yang kurang baik juga bagi instansi atau perusahaan, baik pencatatan data, penyimpanan data atau pun lainnya. Untuk itulah di perlukan sebuah langkah atau suatu proses untuk mengatasi hal tersebut yaitu salah satunya dengan menggunakan metode algoritma K-means. K-means clustering merupakan salah satu metode cluster analysis non hirarki yang berusaha untuk mempartisi objek yang ada kedalam satu atau lebih cluster berdasarkan karakteristiknya [2].

Pada tahap awal dalam melakukan penelitian ini, kami melakukan observasi dan wawancara kepada pekerja Puskesmas Kecamatan Sawah Besar guna mendapatkan informasi yang kami butuhkan. Dalam penelitian yang dilaksanakan pada Puskesmas Kecamatan Sawah Besar ini, peneliti menemukan suatu masalah yang dialami oleh instansi atau perusahaan tersebut yaitu dalam menangani pengolahan data obat yang belum maksimal dikarenakan data obat belum diklasifikasikan berdasarkan karakteristiknya. Dari hasil

observasi dan wawancara pada Puskesmas Kecamatan Sawah Besar, peneliti berusaha membantu dalam memecahkan masalah yang sedang dialami oleh Puskesmas Kecamatan Sawah Besar yang digunakan sebagai objek penelitian.

K-means clustering adalah salah satu algoritma unsupervised learning yang termasuk ke dalam analisis kluster (cluster analysis) non hirarki yang digunakan untuk mengelompokkan data berdasarkan variabel atau feature. Algoritma k-means merupakan salah satu algoritma pengelompokan. Cluster mengacu pada pengelompokan data, observasi atau kasus berdasar kemiripan objek yang diteliti. Sebuah cluster adalah suatu kumpulan data yang mirip dengan lainnya atau tidakmiripan data pada kelompok lain [3]. Data yang diambil untuk dikelompokkan dengan metode algoritma k-means berdasarkan data obat-obatan yang ada pada Puskesmas Kecamatan Sawah Besar, dikarenakan data obat pada Puskesmas Kecamatan Sawah Besar belum terkelola dengan teratur. Algoritma k-means merupakan salah satu dari algoritma yang banyak digunakan dalam pengelompokan karena kesederhanaan dan efisiensinya[3].

Berdasarkan dari uraian diatas, maka penelitian yang akan dilakukan bertujuan untuk mengetahui pengelompokan serta stok data obat dengan menggunakan algoritma k-means. Dimana pengelompokan obat berdasarkan kriteria obat bebas, obat bebas terbatas, obat keras, psikotropika dan narkotika. Penelitian dilakukan pada Puskesmas Kecamatan Sawah Besar. Output yang dihasilkan adalah kelompok data obat-obatan yang ketersediaan stoknya tidak terpantau dengan baik.

II. LANDASAN TEORI

Penelitian pada hakikatnya adalah suatu bentuk usaha untuk menguji, mengembangkan, dan menemukan kebenaran dari suatu hal dengan cara menggunakan berbagai macam metode. Secara umum, penelitian juga dapat diartikan sebagai suatu proses pengumpulan data dan analisis data yang dilakukan secara logis dan sistematis untuk mencapai suatu tujuan tertentu.

Dari hasil observasi dan penelitian yang dilakukan pada Puskesmas Kecamatan Sawah Besar ditemukan suatu permasalahan yang sedang dialami oleh instansi tersebut yaitu pengelompokan data obat-obatan yang kurang tertata dengan rapi dan pengelolaan data stok obat. Pada penelitian ini, peneliti melakukan suatu penelitian mengenai data obat pada Puskesmas Kecamatan Sawah Besar dengan menggunakan metode algoritma k-means. Untuk menyelesaikan permasalahan yang sedang dialami oleh instansi tersebut, maka peneliti berusaha untuk mencari jalan keluar agar masalah yang sedang terjadi dapat teratasi dengan baik.

A. K-Means Clustering

K-means clustering merupakan salah satu Algoritma unsupervised learning yang termasuk kedalam analisis kluster (cluster analysis) non hirarki yang digunakan untuk mengelompokkan sejumlah data berdasarkan suatu variabel (feature). Tujuan dari k-means clustering yaitu untuk mendapatkan kelompok data dengan cara memaksimalkan aspek kesamaan karakteristik data dalam suatu kluster dan memaksimalkan perbedaan antar kluster data [3].

Tahapan dalam menggunakan metode Algoritma K-means clustering yaitu:

- 1) Menentukan jumlah cluster (K).
- 2) Pilih titik acak sebanyak K.
- 3) Tandai semua data berdasarkan titik centroid terdekat.
- 4) Menentukan titik centroid baru berdasarkan cluster yang terbentuk.
- 5) Tandai ulang data berdasarkan jarak terdekat terhadap centroid baru.
- 6) Ulangi kembali langkah 4 dan 5 sampai tidak ada pergerakan lagi.

Ada beberapa keuntungan menggunakan algoritma K-Means, yaitu :

- 1) Mudah untuk diimplementasikan dan dijalankan.
- 2) Waktu yang dibutuhkan relative cepat.
- 3) Umum untuk digunakan
- 4) Mudah untuk diadaptasi

B. Clustering

Clustering adalah suatu proses pengelompokan objek serupa kedalam beberapa kelompok data yang berbeda atau lebih mudahnya partisi dari sebuah data set kedalam subset, dengan begitu data-data dalam setiap subset mempunyai arti yang bermanfaat. Sebuah cluster terbentuk dari kumpulan objek-objek yang serupa antara satu objek yang lainnya dan berbeda dengan objek yang terdapat cluster lainnya [4].

Clustering berbeda dengan klasifikasi dimana clustering tidak memiliki variabel target. Clustering tidak digunakan untuk mengklasifikasi, memprediksi, atau memperkirakan nilai suatu target. Clustering digunakan untuk membagi seluruh data menjadi kesamaan atau kelompok data yang memiliki kesamaan.

C. Data Cleaning

Data cleaning merupakan proses untuk dapat mengatasi nilai yang hilang, noise dan data yang tidak konsisten [5]. Caranya adalah dengan mendeteksi adanya error atau corrupt pada data, kemudian memperbaiki atau menghapus data jika memang diperlukan.

Terkadang, saat menggabungkan beberapa data sources sekaligus, ada kemungkinan data terduplikasi atau bahkan salah label. Situasi seperti ini juga memerlukan data cleaning agar tidak muncul masalah yang lebih rumit dalam pengolahan data kedepannya.

D. Unsupervised learning

Unsupervised learning adalah salah metode algoritma machine learning yang digunakan untuk menarik kesimpulan dari dataset. Metode ini hanya akan mempelajari suatu data berdasarkan kedekatannya saja atau yang biasa disebut dengan clustering. Metode unsupervised learning yang paling umum adalah analisis cluster, yang digunakan pada analisa data untuk mencari pola-pola tersembunyi atau pengelompokan dalam data.

Tujuan unsupervised learning adalah membangun algoritma atau model yang dapat mendeskripsikan struktur

tersembunyi pada data. *Unsupervised learning* dikenal juga sebagai model variabel tersembunyi (*latent variable models*). *Unsupervised learning* biasanya digunakan untuk *clustering*, *vector quantization*, *feature extraction*, *signal coding*, dan *data analysis*. Beberapa metode *unsupervised learning* di antaranya *k-means*, *fuzzy c-means*, *principal component analysis*, dan sebagainya. [6]

E. Obat

Obat merupakan bahan atau campuran bahan, termasuk produksi biologi yang dipakai untuk menyelidiki atau mempengaruhi sistem fisiologi dalam rangka untuk diagnosis, pemulihan, penyembuhan, pencegahan, peningkatan kesehatan, dan kontrasepsi untuk manusia [7].

BPOM (Badan Pengawas Obat dan Makanan) menyusun aturan dan klasifikasi obat berdasarkan jenisnya :

1) Obat Bebas

Obat bebas merupakan obat yang terjual secara bebas di pasar yang dapat dibeli tanpa memerlukan resep dari dokter. Jenis obat ini memiliki gambar lingkaran hijau dan batas hitam pada kemasannya.

2) Obat Bebas Terbatas

Obat jenis ini juga dapat dibeli dengan bebas tanpa memerlukan resep dari dokter, tetapi obat jenis ini mempunyai peringatan khusus dalam penggunaannya. Jenis obat ini masih merupakan obat keras, namun tersedia dalam jumlah tertentu di apotek. Obat ini diidentifikasi dengan lingkaran biru dengan batas hitam.

3) Obat Keras

Obat keras adalah obat yang tidak tersedia secara bebas tanpa resep dokter. Obat jenis ini dapat memperburuk penyakit, meracuni tubuh, dan dapat berakibat fatal jika dikonsumsi tanpa pengawasan medis. Jenis obat ini memiliki lingkaran merah dengan bingkai hitam dan ikon dengan huruf "K" di dalamnya.

4) Obat Psikotropika dan Narkotika

Obat jenis ini hanya bisa didapatkan dengan adanya resep dokter, dengan adanya tanda tangan dari dokter dan disertai nomor izin praktik dokter pada resep tersebut. Narkotika psikotropika adalah zat atau obat yang dapat menurunkan aktivitas otak dan susunan saraf pusat, menimbulkan gangguan perilaku dengan halusinasi dan gangguan berpikir, serta dapat membuat pemakainya kecanduan zat atau obat jenis tersebut.

F. Pusat Kesehatan Masyarakat (Puskesmas)

Usaha kesehatan adalah hak fundamental bagi semua individu dan semua warga negara berhak mengakses layanan kesehatan. Puskesmas merupakan organisasi fungsional yang menyelenggarakan berbagai upaya kesehatan yang bersifat menyeluruh, terpadu, merata, dan terjangkau masyarakat. Oleh karena itu, sesuai dengan kewajiban Pasal 28H Ayat 1 UUD 1945, bahwa kesehatan merupakan tanggung jawab semua pihak kepada masyarakat dan pemerintah [8].

Pusat Kesehatan Masyarakat merupakan salah satu fasilitas kesehatan masyarakat yang penting di Indonesia. Puskesmas merupakan organisasi fungsional yang menyelenggarakan kegiatan terintegrasi penuh, adil, dapat diterima, dan terjangkau oleh masyarakat sesuai dengan apa

yang diamanatkan di dalam UU No 36 Tahun 2009 Pasal 4 ayat 2 yang berbunyi “ Setiap orang yang mempunyai hak yang sama-sama memperoleh pelayanan kesehatan yang aman, bermutu dan terjangkau”.

Oleh karena itu, peran aktif masyarakat diperlukan, dan hasil pengembangan ilmu pengetahuan dan teknologi tepat guna dapat dimanfaatkan dengan biaya murah yang terjangkau oleh pemerintah dan masyarakat luas untuk mencapai derajat kesehatan yang optimal tanpa mengorbankan mutu pelayanan-pelayanan kesehatan.

Puskesmas Kecamatan Sawah Besar terletak di Kecamatan Sawah Besar kota administrasi Jakarta Pusat yang terletak antara 106^o 48' 66" Bujur Timur dan 6^o 22' 66" memiliki luas 6,61 km² terdiri atas 5 kelurahan, 49 Rukun Warga (RW) dan 598 Rukun Tetangga (RT).[9]



Gambar 1. Gambar Wilayah Jakarta Pusat
Sumber : Peta Wilayah Kecamatan Sawah Besar

Gambar 1 merupakan penjelasan gambar pemetaan wilayah Jakarta Pusat, gambar ini didapatkan sesuai kondisi geografis wilayah pada saat ini. Puskesmas Kecamatan Sawah Besar memiliki lima Kelurahan :

- 1) Kelurahan Kartini.
- 2) Kelurahan Karang Anyar.
- 3) Kelurahan Mangga Dua Selatan.
- 4) Kelurahan Gunung Sahari Utara.

III. METODE PENELITIAN

Metode penelitian merupakan langkah-langkah atau tata cara yang dilakukan oleh peneliti dalam rangka mengumpulkan data selengkap mungkin serta melakukan investigasi pada data yang telah didapatkan. Metode penelitian ini memberikan gambaran rancangan pada penelitian yang dilakukan.

A. Instrumen Penelitian

Instrumen merupakan alat yang digunakan untuk melakukan sesuatu, sedangkan penelitian memiliki arti pemeriksaan, kegiatan pengumpulan, mengolah, menyajikan dan menganalisa data-data secara sistematis secara objektif dengan tujuan memecahkan suatu persoalan atau menguji hipotesis.

B. Pengumpulan Data

Pengumpulan data merupakan salah satu langkah dalam mencapai solusi dengan mendata ulang data-data yang dibutuhkan sebagai dasar atau landasan dari suatu pemecahan masalah, maka akan memudahkan langkah berikutnya dalam

tahap penyelesaian, Adapun cara yang dilakukan peneliti dalam tahap ini yaitu wawancara, observasi, dan studi pustaka.

Pada tahap ini dalam mengumpulkan data yang kami gunakan dalam penyusunan skripsi terbagi 2 bagian yaitu :

- 1) Jenis Data, terdiri dari
 - a) Data Primer (*private*), merupakan data yang dikumpulkan langsung dari sumber data.
 - b) Data Sekunder (*public*), merupakan data
 - c) yang diperoleh dari peneliti atau pihak lain.
- 2) Teknik Pengumpulan Data
Untuk mendapatkan informasi yang di perlukan dalam pembuatan skripsi ini, ada 3 cara yang dilakukan penulis untuk mengumpulkan data yaitu:

a) Wawancara
Metode wawancara yang digunakan yaitu wawancara mendalam dengan menggunakan kisi-kisi pertanyaan. Wawancara dilakukan dengan menggali informasi, Tahapan ini mendapatkan data yang kita butuhkan yaitu data-data terkait obat-obatan, profil Puskesmas, visi dan misi, dan Struktur organisasi Puskesmas Sawah Besar.

b) Observasi
Penulis melakukan pengamatan secara langsung kegiatan yang ada di ruangan obat pada Puskesmas Sawah Besar yang berlokasi di Jl. Mangga Dua Dalam K No.13, RT.1/RW.12, Mangga Dua Sel., Kecamatan Sawah Besar, Jakarta Pusat, Daerah Khusus Ibukota Jakarta.

c) Studi Pustaka
Studi pustaka dilakukan untuk menemukan teori-teori yang berkaitan dengan permasalahan yang dihadapi. Penulis mempelajari berbagai sumber seperti jurnal, catatan selama kuliah dan buku-buku yang berkaitan dengan permasalahan yang dihadapi. Termasuk studi pustaka melalui internet seperti *digital library*, *website* dan sumber lainnya yang terdapat pada internet.

C. Pengolahan Data

Kami melakukan pengajuan untuk memperoleh data dari Puskesmas terkait obat-obatan yang ada di Puskesmas Sawah Besar, data dikirimkan melalui Email dan kami olah data-data tersebut menggunakan metode *K-means Clustering*.

D. Model yang diusulkan

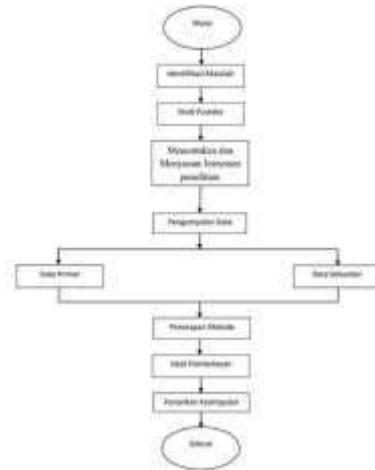
Pada penelitian ini penulis menggunakan metode Algoritma *K-Means Clustering*. *K-means* merupakan algoritma yang bersifat *unsupervised learning*. *K-means* digunakan untuk membentuk data ke dalam data *cluster*.

Dengan metode ini, penulis dapat mengelompokan data-data kedalam kelompok yang menjelaskan data-data dalam satu kelompok yang memiliki karakteristik sama dan memiliki karakteristik berbeda. Metode *K-Means Clustering* digunakan untuk megelompokan data-data kedalam *cluster* atau beberapa kelompok berdasarkan karakteristiknya. Berikut merupakan langkah-langkah dalam pengelompokan

data dengan menggunakan metode *K-Means* :

E. Tahapan Penelitian

Tahapan penelitian merupakan langkah-langkah dalam melakukan proses penelitian dari awal hingga akhir.



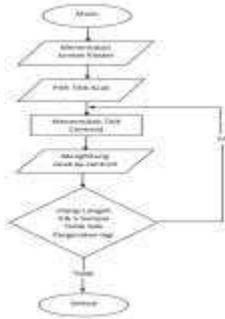
Gambar 2. Tahapan Penelitian

Tahapan penelitian merupakan langkah-langkah dalam melakukan proses penelitian dari awal hingga akhir. Tahapan penelitian sebagai berikut :

- 1) Identifikasi Masalah
Melakukan identifikasi tentang masalah apa yang akan dibahas berkaitan dengan pengelompokan data obat-obatan berdasarkan literatur dan informasi yang telah diperoleh pada saat observasi.
- 2) Studi Pustaka
Mempelajari literatur, jurnal nasional, pencarian di internet, dan sumber informasi lain yang ada kaitannya dengan topik yang akan dibahas baik berupa textbook atau paper yang akan digunakan sebagai kajian teori serta menjadi landasan dalam penelitian.
- 3) Menentukan dan menyusun instrument penelitian
Tahapan ini adalah penentuan instrument penelitian yaitu dengan cara melakukan observasi dan wawancara kepada Puskesmas Sawah Besar.
- 4) Pengumpulan Data
Dalam penelitian ini data yang dibutuhkan dibagi menjadi dua yaitu data primer dan data sekunder.
- 5) Penerapan Metode
Pengolahan data dihitung dengan menggunakan metode *K-means* yang terdiri dari pembuatan tabulasi yang nantinya akan dihitung secara manual dan menggunakan aplikasi RapidMiner untuk pengelompokan data obat berdasarkan karakteristiknya.
- 6) Hasil Pembahasan
Menganalisa hasil pengolahan data berdasarkan hasil penelitian yang telah dilakukan.
- 7) Kesimpulan
Kesimpulan diambil berdasarkan hasil penelitian dan diperiksa apakah sesuai dengan maksud dan tujuan penelitian.

IV. HASIL DAN PEMBAHASAN

Metode K-Means Clustering digunakan untuk mengelompokkan data-data kedalam cluster atau beberapa kelompok berdasarkan karakteristiknya. Berikut merupakan langkah-langkah dalam pengelompokan data dengan menggunakan metode K-Means :



Gambar 3. Skema Algoritma K-Means

A. Pengolahan Data Menggunakan Perhitungan Algoritma K-Means

Data yang akan diteliti berasal dari Laporan Pemakaian dan Lembaran Permintaan Obat (LPLPO) gudang farmasi Pusat Kesehatan Masyarakat Kecamatan Sawah Besar bulan Desember 2021.

Gambar 4. Dataset form LPLPO Desember 2021

Dataset form LPLPO Desember 2021 yang didalam berisi 18 atribut diantaranya yaitu: nama obat, satuan, stok awal, penerimaan, persediaan, apotek, TB, RB, KIA/KB, gigi, IMS, LAB, Kartini, PS. Baru, GSU, lainnya, total dan sisa stok. Terdapat ada 355 data nama obat yang ada di Puskesmas Kecamatan Sawah Besar, data tersebut akan dipisahkan kedalam 4 katagori yaitu : obat bebas, obat bebas terbatas, obat keras, psikotropika dan narkotika. Pada tahapan pembahasan penulis hanya mengambil satu sampel untuk membahasnya yaitu pada obat psikotropika dan narkotika.

B. Penerapan Metode Algoritma K-Means

Ada beberapa langkah dalam mengelola data dengan menggunakan metode k-mens data obat di gabungan menjadi satu. Berikut merupakan langkah-langkahnya :

1) Data Integration

Pada tahapan *Integration* ini dilakukan penggabungan data obat pada pengeluaran menjadi yang total pemakaian. Angka 0 sebagai obat yang tidak ada jumlah pemakaiannya

a) Obat Psikotropika dan Narkotika

TABEL I. TABEL DATA INTEGRATION OBAT PSIKOTROPIKA DAN NARKOTIKA

No	Nama Obat	Satuan	Stok Awal	Pengeluaran	Sisa Stok
1	Codeine Hcl 10 mg	Tablet	0	0	0
2	Amitipiline 25 mg	Tablet	0	0	0
3	Chlorpromazine 100mg	Tablet	2500	500	2000
4	Clobazam 10 mg	Tablet	5000	500	4500
5	Diazepam 2 mg	Tablet	0	0	0
6	Diazepam 5 mg	Tablet	300	0	300
7	Diazepam Inject 5 mg/ml	Ampul	20	10	10
8	Fluoksetin 10 mg	Tablet	0	0	0
9	Haloperidol 1.5 mg	Tablet	0	0	0
10	Haloperidol 5 mg	Tablet	3200	0	3200
11	Haloperido Inject 5mg/ml	Ampul	0	0	0
12	Phenobarbital 30 mg	Tablet	1100	0	1100
13	Proneuron (Diazepam + Metampiron)	Tablet	400	0	400
14	Risperidone 2 mg	Tablet	6500	0	6500
15	Stesolid supp 5mg/ml	Suppos	5	0	5
16	Stesolid Supp 10 mg/ml	Suppos	30	0	30
17	Trifluoperazine 5mg	Tablet	0	0	0
18	Trihexyphenidyl 2 mg	Tablet	4000	0	4000

Tabel Data *Intergration* Obat Psikotropika dan Narkotika pada tabel ini pengeluaran dijadikan satu agar lebih mudah dalam pengolahan datanya.

2) Data Clearing

Pada tahapan ini di lakukan yang dilakukan adalah membuang atribut yang tidak relevan pada Tabel I, Atribut yang dibuang adalah nomor dan satuan sedangkan nama obat menjadi kode obat agar mempermudah pemerosesan data.

TABEL II. TABEL DATA CLEARING PSIKOTROPIKA DAN NARKOTIL

Kode Obat	Stok Awal	Pengeluar	Sisa Stok
1	0	0	0
2	0	0	0
3	2500	500	2000
4	5000	500	4500
5	0	0	0
6	300	0	300
7	20	10	10
8	0	0	0
9	0	0	0
10	3200	0	3200
11	0	0	0
12	1100	0	1100
13	400	0	400
14	6500	0	6500
15	5	0	5
16	30	0	30
17	0	0	0
18	4000	0	4000

3) Inisialisasi Obat Psikotropika dan Narkotika

Dalam inialisasi pada Psiktropika dan Narkotika penulis memakai Tabel 4.2 yaitu data *Clearnig* psiktropika dan narkotika.

Tabel 4.3 merupakan pemelihan K data sebagai *centroid* awal secara acak yang di pilih dari kode obat ke-3 sebagai C1, kode obat ke-10 sebagai C2, kode Obat ke-12 sebagai C3.

TABEL III. CENTROID AWAL

Centroid	X	Y	Z
C1	2500	500	2000
C2	3200	0	3200
C3	4000	0	4000

a) Iterasi 1

Menghitung jarak setiap data ke *centroid* terdekat yang akan menjadi *cluster* yang diikuti data tersebut. Berikut perhitungan jarak setiap *centroid* pada data ke-1

Rumus *Euclidean Distance* :

$$d(X_j, C_j) = \sqrt{\sum_{j=1}^n (X_j - C_j)^2}$$

- Keterangan : d = Jarak
- : J = Banyak Data
- : c = *Centroid*
- : x = Data

Jarak *centroid* ke-1 data pada *cluster* 1 adalah :

$$d(X_j, C_j) = \sqrt{\sum_{j=1}^n (X_j - C_j)^2}$$

$$= \sqrt{(0 - 2500)^2 + (0 - 500)^2 + (0 - 2000)^2}$$

$$= 3240,37$$

Jarak *centroid* ke-1 data pada *cluster* 2 adalah :

$$d(X_j, C_j) = \sqrt{\sum_{j=1}^n (X_j - C_j)^2}$$

$$= \sqrt{(0 - 3200)^2 + (0 - 0)^2 + (0 - 3200)^2}$$

$$= 425,48$$

Jarak *centroid* ke-1 data pada *cluster* 3 adalah :

$$d(X_j, C_j) = \sqrt{\sum_{j=1}^n (X_j - C_j)^2}$$

$$= \sqrt{(0 - 4000)^2 + (0 - 0)^2 + (1425 - 4000)^2}$$

$$= 5656,85$$

Untuk jarak *centroid* selanjutnya dapat dilihat pada Tabel IV.

TABEL IV. JARAK CENTROID ITERASI 1

KODE OBAT	Centroid			Jarak Terpendek	Kelompok
	C1	C2	C3		
1	3240,37	4525,48	5656,85	3240,37	C2
2	3240,37	4525,48	5656,85	3240,37	C1
3	0,00	1476,48	2549,51	0,00	C1
4	3535,53	2275,96	1224,74	1224,74	C3
5	3240,37	4525,48	5656,85	3240,37	C1
6	2824,89	4101,22	5232,59	2824,89	C1
7	3217,23	4504,29	5635,65	3217,23	C1
8	3240,37	4525,48	5656,85	3240,37	C1
9	3240,37	4525,48	5656,85	3240,37	C2
10	1476,48	0,00	1131,37	0,00	C2
11	3240,37	4525,48	5656,85	3240,37	C1
12	1737,81	2969,85	4101,22	1737,81	C1
13	2687,01	3959,80	5091,17	2687,01	C1
14	6041,52	4666,90	3535,53	3535,53	C3
15	3233,43	4518,41	5649,78	3233,43	C1
16	3198,72	4483,06	5614,43	3198,72	C1
17	3240,37	4525,48	5656,85	3240,37	C1
18	2549,51	1131,37	0,00	0,00	C3

Selanjutnya dihitung *centroid* yang baru disetiap *cluster* berdasarkan data yang tergabung pada setiap *cluster*nya.

Tabel V. merupakan *cluster* 1 hanya terdapat 12 data yang tergabung kedalamnya.

TABEL V. DATA CLUSTER 1 ITERASI 1

KODE OBAT	x	y	z
1	0	0	0
2	0	0	0
5	0	0	0
6	300	0	300
7	20	10	10
8	0	0	0
9	0	0	0
11	0	0	0
13	400	0	400
15	5	0	5
16	30	0	30
17	0	0	0
Rata-rata	62,92	0,83	62,08

Tabel VI merupakan *cluster* 1 hanya terdapat 2 data yang tergabung kedalamnya

TABEL VI. DATA CLUSTER 2 ITERASI 1

Kode Obat	X	Y	Z
3	2500	500	2000
12	1100	0	1100
Rata-rata	1800	250	1550

Tabel VII. merupakan *cluster* 2 hanya terdapat 4 data yang tergabung kedalamnya.

TABEL VII. DATA CLUSTER 3 ITERASI 1

KODE OBAT	X	Y	Z
18	4000	0	4000
4	5000	500	4500
10	3200	0	3200
14	6500	0	6500
Rata-rata	4675	125	4550

Dari perhitungan iterasi 1, ditemukan *centroid* baru yang ditunjukkan pada hasil Tabel VIII:

TABEL VIII. CENTROID BARU ITERASI 1

Centroid	x	y	z
C1	362,92	42,50	320,42
C2	1066,67	0,00	1066,67
C3	5166,67	166,67	5000,00

b) Iterasi 2

Ulangi perhitungan sampai mendapatkan model *cluster* sehingga dapat membandingkan *cluster* iterasi 1 dan iterasi 2 apakah sama atau berbeda. Berikut perhitungan jarak kesetiap *centroid* pada ke 1 iterasi ke 2 :

Rumus *Euclidean Distance* :

$$d(X_j, C_j) = \sqrt{\sum_{j=1}^n (X_j - C_j)^2}$$

Keterangan : d = Jarak
: J = Banyak Data
: c = *Centroid*
: x = Data

Jarak *centroid* ke-1 data pada *cluster* 1 adalah:

$$d(X_j, C_j) = \sqrt{\sum_{j=1}^n (X_j - C_j)^2}$$

$$= \sqrt{(0 - 362,92)^2 + (0 - 42,50)^2 + (0 - 320,42)^2}$$

$$= 485,99$$

Jarak *centroid* ke-1 data pada *cluster* 2 adalah:

$$d(X_j, C_j) = \sqrt{\sum_{j=1}^n (X_j - C_j)^2}$$

$$= \sqrt{(0 - 1066,67)^2 + (0 - 0)^2 + (1425 - 1066,67)^2}$$

$$= 1208,49$$

Jarak *centroid* ke-1 data pada *cluster* 3 adalah :

$$d(X_j, C_j) = \sqrt{\sum_{j=1}^n (X_j - C_j)^2}$$

$$= \sqrt{(0 - 5166,67)^2 + (0 - 166,67)^2 + (0 - 5000,0)^2}$$

$$= 7191,82$$

Untuk jarak *centroid* selanjutnya dapat dilihat pada Tabel IX

TABEL IX. JARAK CENTROID ITERASI 2

KODE OBAT	Centroid			Jarak Terpendek	Kelompok
	C1	C2	C3		
1	485,99	1508,49	7191,82	485,99	C1
2	485,99	1508,49	7191,82	485,99	C1
3	2756,34	1782,01	4027,68	1782,01	C2
4	6259,45	5244,89	623,61	623,61	C3
5	485,99	1508,49	7191,82	485,99	C1
6	78,62	1084,23	6767,73	78,62	C1
7	463,69	1487,33	7170,27	463,69	C1
8	485,99	1508,49	7191,82	485,99	C1
9	485,99	1508,49	7191,82	485,99	C1
10	4042,63	3016,99	2671,25	2671,25	C3
11	485,99	1508,49	7191,82	485,99	C1
12	1073,71	47,14	5636,98	47,14	C2
13	97,54	942,81	6626,38	97,54	C1
14	8709,35	7683,89	2013,84	2013,84	C3
15	478,96	1501,42	7184,75	478,96	C1
16	443,83	1466,07	7149,41	443,83	C1
17	485,99	1508,49	7191,82	485,99	C1
18	5173,93	4148,36	1545,60	1545,60	C3

Selanjutnya dihitung *centroid* yang baru disetiap *cluster* berdasarkan data yang tergabung pada setiap *clusternya*.

Tabel X merupakan *cluster* 1 hanya terdapat 12 data yang tergabung kedalamnya.

TABEL X. DATA CLUSTER 1 ITERASI 2

KODE OBAT	x	Y	Z
1	0	0	0
2	0	0	0
5	0	0	0
6	300	0	300
7	20	10	10
8	0	0	0
9	0	0	0
11	0	0	0
13	400	0	400
15	5	0	5
16	30	0	30
17	0	0	0
Rata-rata	62,92	0,83	62,08

Tabel XI merupakan *cluster* 2 hanya terdapat 2 data yang tergabung kedalamnya.

TABEL XI. DATA CLUSTER 2 ITERASI 2

KODE OBAT	x	Y	z
3	2500	500	2000
12	1100	0	1100
Rata-rata	1800	250	1550

Tabel XII merupakan cluster 3 hanya terdapat 4 data yang tergabung kedalamnya.

TABEL XII. DATA CLUSTER 3 ITERASI 2

KODE OBAT	x	Y	z
18	4000	0	4000
4	5000	500	4500
10	3200	0	3200
14	6500	0	6500
Rata-rata	4675	125	4550

Karena ada 5 data yang berpindah cluster, yaitu data ke-1, 2, 9, 10 dan 12, maka selanjutnya dihitung centroid yang baru untuk setiap cluster berdasarkan data yang tergabung pada setiap clusternya.

TABEL XIII. CENTROID BARU ITERASI 2

Centroid	x	y	z
C1	62,92	0,83	62,08
C2	1800,00	250,00	1550,00
C3	4675,00	125,00	4550,00

c) Iterasi 3

Ulangi perhitungan sampai mendapatkan model cluster sehingga dapat membandingkan cluster iterasi 2 dan itrasi 3 apakah sama atau berbeda. Berikut perhitungan jarak kesetiap centroid itrasi ke 3 :

Rumus Euclidean Distance :

$$d(X_j, C_j) = \sqrt{\sum_{j=1}^n (X_j - C_j)^2}$$

Keterangan : d = Jarak
: J = Banyak Data
: c = Centroid
: x = Data

Jarak centroid ke-1 data pada cluster 1 adalah :

$$d(X_j, C_j) = \sqrt{\sum_{j=1}^n (X_j - C_j)^2}$$

$$= \sqrt{(0 - 62,92)^2 + (0 - 0,83)^2 + (0 - 62,08)^2}$$

$$= 88,39$$

Jarak centroid ke-1 data pada cluster 2 adalah :

$$d(X_j, C_j) = \sqrt{\sum_{j=1}^n (X_j - C_j)^2}$$

$$= \sqrt{(0 - 1800,0)^2 + (0 - 250,0)^2 + (0 - 1550,0)^2}$$

$$= 2388,51$$

Jarak centroid ke-1 data pada cluster 3 adalah :

$$d(X_j, C_j) = \sqrt{\sum_{j=1}^n (X_j - C_j)^2}$$

$$= \sqrt{(0 - 4675,0)^2 + (0 - 125,0)^2 + (1425 - 4550,0)^2}$$

$$= 6524,86$$

Untuk jarak centroid selanjutnya dapat dilihat pada Tabel XIV

TABEL XIV. JARAK CENTROID ITERASI 3

KODE OBAT	Centroid			Jarak Terpendek	Kelompok
	C1	C2	C3		
1	485,99	1508,49	7191,82	485,99	C1
2	485,99	1508,49	7191,82	485,99	C1
3	2756,34	1782,01	4027,68	1782,01	C2
4	6259,45	5244,89	623,61	623,61	C3
5	485,99	1508,49	7191,82	485,99	C1
6	78,62	1084,23	6767,73	78,62	C1
7	463,69	1487,33	7170,27	463,69	C1
8	485,99	1508,49	7191,82	485,99	C1
9	485,99	1508,49	7191,82	485,99	C1
10	4042,63	3016,99	2671,25	2671,25	C3
11	485,99	1508,49	7191,82	485,99	C1
12	1073,71	47,14	5636,98	47,14	C2
13	97,54	942,81	6626,98	97,54	C1
14	8709,35	7683,89	2013,84	2013,84	C3
15	478,96	1501,42	7184,75	478,96	C1
16	443,83	1406,07	7149,41	443,83	C1
17	485,99	1508,49	7191,82	485,99	C1
18	5173,93	4148,36	1545,60	1545,60	C3

Selanjutnya dihitung centroid yang baru disetiap cluster berdasarkan data yang tergabung pada setiap clusternya.

Tabel XV merupakan cluster 1 hanya terdapat 12 data yang tergabung kedalamnya.

TABEL XV. DATA CLUSTER 1 ITERASI 3

KODE OBAT	x	y	z
1	0	0	0
2	0	0	0
5	0	0	0
6	300	0	300
7	20	10	10
8	0	0	0
9	0	0	0
11	0	0	0

13	400	0	400
15	5	0	5
16	30	0	30
17	0	0	0
Rata-rata	62,92	0,83	62,08

Tabel XVI merupakan *cluster 2* hanya terdapat 2 data yang tergabung kedalamnya.

TABEL XVI. DATA CLUSTER 2 ITERASI 3

KODE OBAT	X	Y	Z
3	2500	500	2000
12	1100	0	1100
Rata-rata	1800	250	1550

Tabel XVII merupakan *cluster 3* hanya terdapat 4 data yang tergabung kedalamnya.

TABEL XVII. DATA CLUSTER 3 ITERASI 3

KODE OBAT	X	Y	Z
4	5000	500	4500
10	3200	0	3200
14	6500	0	6500
18	4000	0	4000
Rata-rata	4675	125	4550

Karena tidak ada data yang berpindah *cluster*, maka proses perhitungan *centroid* dihentikan dan berakhir pada iterasi ke-3

d) Pengukuran Nilai Akurasi (*AccurationRate*)

$$Accuration Rate = \frac{\text{Jumlah data yang sesuai label aslinya}}{\text{Jumlah data keseluruhan}} \times 100\%$$

Data yang ada pada *cluster 1* terdapat 12 data, maka nilai akurasinya adalah sebagai berikut :

$$Accuration Rate = \frac{12}{18} \times 100\% = 0,66 \%$$

Jadi nilai akurasi pada *cluster 1* adalah 0,66 %

Data yang ada pada *cluster 2* terdapat 2 data, maka

nilai akurasinya adalah sebagai berikut :

$$Accuration Rate = \frac{2}{18} \times 100\% = 0,11\%$$

Jadi nilai akurasi pada *cluster 2* adalah 0,11 %

Data yang ada pada *cluster 3* terdapat 4 data, maka nilai akurasinya adalah sebagai berikut :

$$Accuration Rate = \frac{4}{18} \times 100\% = 0,22\%$$

Jadi nilai akurasi pada *cluster 3* adalah 0,22 %

5. KESIMPULAN

Tingkat pemakaian obat di Puskesmas kecamatan sawah besar yg paling tinggi adalah obat bebas terbatas dan obat Psikotropika & Narkotika, untuk pemakaian sedang adalah Obat Bebas dan Obat Keras dan untuk pemakaian obat terendah tidak ada dikarenakan hasil dari perhitungan *accuration rata* memiliki dua persamaaa dalam nilainya. Dengan menggunakan algoritma *K-Means* pihak Puskesmas Kecamatan Sawah Besar dapat mengelompokkan data obat yang tingkat pemakaiannya tinggi, sedang hingga terendah berdasarkan volume pemakaian dan pemasukan. Metode algoritma *K-Means* merupakan metode yang sangat berguna untuk membantu instansi menemukan informasi yang sangat penting dari gudang data mereka yang selama ini tidak terkontrol stoknya.

REFERENCES

- [1] A. Yani, "KESEHATAN MASYARAKAT UTILIZATION OF TECHNOLOGY IN THE HEALTH OF COMMUNITY HEALTH," vol. 8, 2018.
- [2] A. Nur Khormarudin, "Teknik Data Mining: Algoritma K-Means Clustering," *J. Ilmu Komput.*, pp. 1–12, 2016.
- [3] A. Sani, "PENERAPAN METODE K-MEANS Related papers," *J. Teknol.*, vol. 1, pp. 1–7, 2014.
- [4] Simamora, "UNIVERSITAS SUMATERA UTARA Poliklinik UNIVERSITAS SUMATERA UTARA," *J. Pembang. Wil. Kota*, vol. 1, no. 3, pp. 82–91, 2018.
- [5] H. Sulastri and A. I. Gufroni, "Penerapan Data Mining Dalam Pengelompokan Penderita Thalassaemia," *J. Nas. Teknol. dan Sist. Inf.*, vol. 3, no. 2, pp. 299–305, 2017, doi: 10.25077/teknosi.v3i2.2017.299-305.
- [6] S. Anastassia, A. Kharis, A. Haqqi, and A. Zili, "Learning Analytics dan Educational Data Mining pada Data Pendidikan," vol. 6, pp. 12–20, 2022.
- [7] S. Rahayuda, "Identifikasi Jenis Obat Berdasarkan Gambar Logo Pada Kemasan Menggunakan Metode Naive Bayes," *Klik - Kumpul. J. Ilmu Komput.*, vol. 3, no. 2, p. 125, 2016, doi: 10.20527/klik.v3i2.46.
- [8] R. Ekasari, M. S. Pradana, G. Adriansyah, M. A. Prasnowo, A. F. Rodli, and K. Hidayat, "Analisis Kualitas Pelayanan Puskesmas Dengan Metode Servqual," *J. Darussalam J. Pendidikan, Komun. dan Pemikir. Huk. Islam*, vol. 9, no. 1, p. 82, 2017, doi: 10.30739/darussalam.v9i1.118.

- [9] K. S. Besar and D. Angka, "Kecamatan sawah besar dalam angka," 2021. *Bulanana Pada Dinas Kesehatan Kabupaten Lombok Timur," Infotek J. Inform. dan Teknol.*, vol. 3, no. 2, pp. 147–156, 2020, doi: 10.29408/jit.v3i2.2315.
- [10] S. Suhartini, L. Kerta Wijaya, and N. Arini Pratiwi, "Penerapan Algoritma K-Means Untuk Pendataan Obat Berdasarkan Laporan