

Penerapan Algoritma K-Means Clustering Untuk Data Obat

Tis Asy Aria, Yuliadi*, M Julkarnain, Fahri Hamdani

Fakultas Rekayasa Sistem, Program Studi Informatika, Universitas Teknologi Sumbawa, Sumbawa, Indonesia

Email: ¹tisaariaasy@gmail.com, ^{2,*}yuliadi@uts.ac.id, ³m.julkarnain@uts.ac.id, ⁴fahri.hamdani@uts.ac.id

Email Penulis Korespondensi: yuliadi@uts.ac.id

Abstrak—Dalam proses perencanaan dan pengendalian pengadaan stok obat UPT. Puskesmas Unter Iwes Sumbawa belum optimal karena permasalahan tidak akuratnya dalam menganalisa kebutuhan obat mempengaruhi jumlah kebutuhan dan permintaan obat. Dalam penelitian ini, dilakukan analisis data obat dengan menerapkan algoritma K-Means Clustering. Data obat dikelompokkan berdasarkan tingkat pemakaian C1, C2, dan C3 (rendah, sedang, tinggi) dengan atribut yaitu Nama Obat, satuan, Persediaan, Permintaan dan Pemakaian. Hasil penelitian ini dapat memberikan strategi sebagai acuan perencanaan dan kebutuhan obat kedepannya. Dari hasil pengujian data obat bulan Januari sampai bulan Desember 2022 dengan jumlah 4.642 data dan data Training berjumlah 4.622. Hasil kinerja pengelompokan obat menggunakan metode Davies Bouldin sebesar 0.513 diapatkan klaster jumlah data C1/C0 (rendah) adalah 4433 items dengan 4000 pemakaian. Klaster C2/C3 (sedang) adalah 184 items dengan 10.850 pemakaian obat dan klaster C3/C2 (tinggi) adalah 5 items dengan 12.000.

Kata Kunci: Pengolahan Data Obat; Puskesmas; Data Mining; K-Means Clustering

Abstract—In the process of planning and controlling the supply of UPT drug stock. Unter Iwes Sumbawa Community Health Center is not yet optimal because the problem of inaccuracy in analyzing drug needs affects the amount needed and demand for medicine. In this study, drug data analysis was conducted by applying the K-Means Clustering algorithm. Drug data is grouped based on usage levels C1, C2, and C3 (low, medium, high) with attributes namely Drug Name, Unit, Supply, Demand and Usage. The results of this study can provide a strategy as a reference for future drug planning and needs. From the results of drug data testing from January to December 2022 with a total of 4,642 data and training data totaling 4,622. The results of the performance of grouping drugs using the Davies Bouldin method of 0.513 clusters of C1/C0 (low) data were 4433 items with 4000 uses. C2/C3 cluster (medium) is 184 items with 10,850 drug use and C3/C2 cluster (high) is 5 items with

Keywords: Drug Data Processing; Puskesmas; Data Mining; K-Means Clustering

1. PENDAHULUAN

Pertumbuhan pembangunan suatu bangsa dapat dipengaruhi melalui banyak indikator salah satunya yaitu di bidang kesehatan. Kesehatan merupakan sebuah sumber daya yang dimiliki setiap manusia sehingga menjadi kebutuhan paling penting yang harus dimiliki. Kesehatan tidak hanya berfokus pada fisik yang bugar namun mencakup jiwa yang sehat dimana setiap individu dapat bersikap toleran serta bisa menerima perbedaan. Dalam hal ini, perkembangan teknologi juga merupakan faktor yang mempengaruhi perkembangan di dunia medis [1].

Obat adalah bagian yang tak terpisahkan dari dunia medis, dan bagian penting dalam upaya penyelenggaraan kesehatan. Oleh karena itu, diperlukan perencanaan yang tepat dalam menentukan kebutuhan obat sehingga perlu dikelola dengan efektif dan efisien. Menjamin ketersediaan dan pemerataan obat merupakan tujuan penting dari perencanaan kebutuhan obat yang cukup pada instansi pelayanan kesehatan seperti rumah sakit, puskesmas, dinas kesehatan dan sejenisnya perencanaan kebutuhan obat ini akan berdampak pada pengadaan, pendistribusian, dan penggunaan [2].

Menerapkan data mining dapat membantu menganalisis penggunaan obat untuk mendapatkan informasi yang berguna dalam perencanaan dan pengendalian stok obat. *Clustering* adalah salah satu metode pengelompokan data yang digunakan dalam data mining, dimana data dikelompokkan menjadi beberapa kelompok yang disebut cluster. Beberapa algoritma yang dapat digunakan dalam teknik ini termasuk *K-Nearest Neighbors (KNN)*, *Fuzzy C-Means*, serta *K-Means*. Algoritma *K-Means* sering digunakan dalam proses *Clustering* untuk mengelompokkan data ke dalam beberapa *Cluster* berdasarkan kemiripan antar data. Dengan menggunakan algoritma ini, data yang berbeda karakteristiknya dapat dikelompokkan dalam *Cluster* yang memiliki karakteristik yang sama [3][4][5][6].

Berdasarkan hasil observasi yang dilakukan penulis saat, dalam pengolahan data obat yang dilakukan oleh Puskesmas Unter Iwes Sumbawa masih belum optimal dalam proses perencanaan dan pengendalian pengadaan stok obat, karena terkendala terkait ketidakakuratan dalam menganalisa kebutuhan obat sehingga sulit untuk mengambil keputusan yang tepat dalam pengendalian dan perencanaan akibatnya dapat mempengaruhi jumlah kebutuhan dan permintaan obat, serta keterlambatan pengiriman obat yang mengakibatkan stok obat menjadi kosong sehingga pelayanan kepada pasien menjadi terganggu [7]. Oleh karena itu, bagian apotek Puskesmas Unter Iwes Sumbawa melakukan *Clusterisasi* data obat menggunakan teknik data mining menggunakan algoritma *K-Means* sehingga menghasilkan pengelompokan data obat. Penulis mengelompokkan data obat berdasarkan obat yang pemakaiannya rendah, sedang, dan tinggi pada tahun 2022, hal ini bertujuan sebagai acuan perencanaan dan pengendalian kebutuhan obat-obatan untuk kedepannya [3].

Penelitian yang berjudul Model Clustering Pemakaian Obat Menggunakan Algoritma K-Means (Studi Kasus Puskesmas Lempake). Dalam penelitian ini metode K-Means telah digunakan untuk melakukan analisis terhadap pemakaian obat dengan menggunakan data pemakaian obat dan permintaan obat (LPLPO) selama tiga tahun (2016-2018) dengan 225 jenis obat yang terdiri dari 80% obat generik dan 20% obat dengan resep dokter. Data kemudian kelompokkan menjadi tiga Cluster berdasarkan nilai pemakaian maksimum, rata-rata, dan minimum. Hasil dari penelitian ini menunjukkan bahwa nilai *Sum Of Squared Error* (SSE) diperoleh sebesar 77,34814 dan menandakan bahwa



pengelompokan tiga Cluster merupakan hasil terbaik. Dalam konteks monitoring persediaan obat di puskesmas, metode K-Means dapat dijadikan sebagai alternatif dalam membuat model analisis [8]. Penelitian yang berjudul Implementasi K-Means Clustering Dalam Mengelompokkan Minat Membaca Penduduk Menurut Wilayah. Penelitian ini menggunakan data pemerintah BPS (Badan Statistik) dalam upaya memberikan masukan dan informasi pada pemerintah untuk melaksanakan sosialisasi atau kegiatan yang dapat meningkatkan minat baca penduduk di setiap wilayah. Dalam penelitian ini, algoritma K-Means digunakan mengelompokkan data menjadi 2 Cluster yaitu Cluster tingkat tinggi (C1) dan rendah (C2). Hasil penelitian menunjukkan bahwa terdapat 12 Provinsi yang termasuk dalam Cluster tingkat tinggi dan 21 Provinsi lainnya termasuk dalam Cluster tingkat rendah [9]. Penelitian yang berjudul Pengelompokan Data Puskesmas Banyuwangi Dalam Pemberian Imunisasi Menggunakan Metode K-Means Clustering. Dalam penelitian ini bertujuan menurunkan angka penyakit yang dapat dicegah dengan imunisasi dengan mengelompokkan puskesmas sehingga bisa mencapai target IDL. Pengelompokan dilakukan menggunakan algoritma K-Means Clustering berdasarkan target IDL rendah, sedang, dan sangat baik dengan data yang digunakan yaitu data imunisasi pada tahun 2016-2017. Hasil dari penelitian menunjukkan acuan bagi puskesmas dalam memberikan penyuluhan kepada masyarakat dalam penanganan penyakit yang dapat dicegah dengan imunisasi berdasarkan hasil dari pengelompokan yang dilakukan menggunakan metode *K-Means Clustering* [10]. Penelitian yang berjudul *Clustering Obat Untuk Menentukan Pola Pemasaran Efektif Di Apotek Amarta Sehat*. Penelitian ini menggunakan algoritma K-Means yang digunakan untuk melakukan Clustering obat guna menentukan pola pemasaran yang efektif yang dikarenakan permintaan obat yang sangat tinggi. Menggunakan metode Elbow, *Clustering* obat dengan K-means dengan pendekatan CRISP-DM menghasilkan lima Cluster. Hasil penelitian yaitu 11 obat dengan kategori sangat laris, 76 obat dengan kategori laris, 131 obat kategori cukup laris, 399 kategori obat kurang laris, dan 326 kategori obat kurang laris. Sehingga dengan mengetahui hasil *Clustering* obat, pemilik apotek dapat menentukan pola pemasaran yang efektif dalam meningkatkan penjualan obat [11].

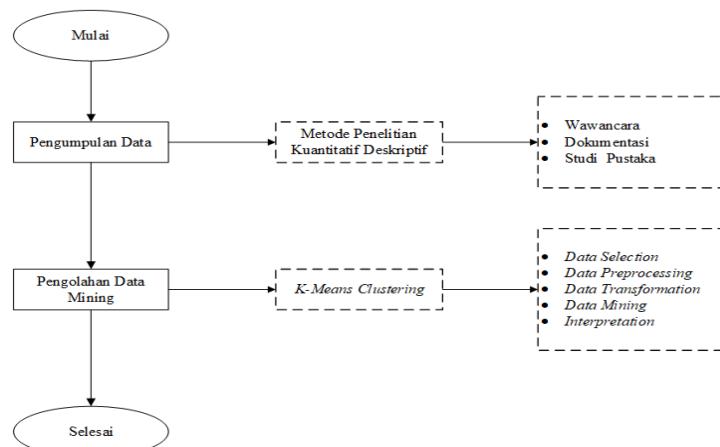
Data mining merupakan istilah yang merujuk pada proses penggalian informasi atau pengetahuan yang terdapat dalam basis data. Pertumbuhan ukuran basis data semakin pesat, bahkan terkadang menyulitkan sistem administrator untuk memahami informasi yang tersimpan atau mengaitkannya dengan pertanyaan yang muncul [12]. *Knowledge Discovery in Database* (KDD) adalah sebuah teknik untuk membentuk pola atau aturan dalam informasi. Informasi yang dihasilkan diperoleh dari data yang besar, juga dikenal sebagai tambang data dan menghasilkan data yang berpotensi bermanfaat. Iterasi dalam proses penambangan data ini disebut sebagai proses KDD [13].

K-Means adalah suatu Teknik analisis kelompok yang bertujuan untuk membagi N objek pengamatan menjadi K kelompok (*Cluster*) berdasarkan mean (rata-rata) terdekat. Metode ini mirip dengan algoritma Expectation-Maximization untuk Gaussian Mixture juga mencari pusat kelompok dalam data melalui iterasi perbaikan. K-Means adalah suatu Teknik pengelompokan data non hirarki yang bertujuan membagi data ke dalam dua atau lebih kelompok. Metode ini memisahkan data berdasarkan karakteristiknya sehingga data yang mirip dikelompokkan ke dalam satu kelompok, sedangkan data yang berbeda dikelompokkan ke dalam kelompok yang berbeda pula [14]. K-Means memiliki karakteristik sederhana dan mudah digunakan, namun pada jenis set data tertentu K-Means tidak mampu melakukan segmentasi data dengan baik. Hasil segmentasi dapat gagal dalam menunjukkan pola kelompok yang mewakili karakteristik bentuk alami data. Selain itu K-Means dapat mengalami masalah ketika mengelompokkan data yang mengandung *outlier* [15][16].

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Pada Penulisan ini penulis menggunakan dua metode, yaitu metode pengumpulan data dan pengolahan Data Mining. Jenis penulisan ini menggunakan pendekatan kuantitatif deskriptif. Dalam penelitian ini, langkah-langkah pengumpulan data dan pengolahan data mining dapat di lihat pada gambar 1 dibawah ini:



Gambar 1. Alur Metode Penulisan

Berdasarkan Gambar 1 yang merupakan tahapan-tahapan dalam penelitian yang dilakukan oleh penulis sebagai pedoman kegiatan sebanyak 2 tahap, yakni pengumpulan data dan pengolahan data mining. Tahap pertama dilakukan pengumpulan data dengan menerapkan metode penelitian kuantitatif deskriptif. Adapun proses dilakukan dalam pengumpulan data tersebut, yakni wawancara, dokumentasi dan studi Pustaka. Tahap Pengolahan data mining dengan mengadopsi siklus dalam *Knowledge Discovery in Database* (KDD) berupa *selection data*, *Preprocessing*, *Transforation*, *Data Mining* dan *Interpretation* dengan menerapkan metode *k-means clustering*.

2.2 Pengolahan Data Mining

Metode atau teknik pengumpulan data adalah metode pengumpulan data. Penulis menggunakan data sekunder dalam karya ini. Istilah "data sekunder" mengacu pada informasi yang dikumpulkan oleh orang lain selain penulis aslinya. Sebagian pintar waktu, catatan atau dokumen dari sumber tidak langsung seperti majalah, kutipan dan publikasi lainnya merupakan data sekunder [17]. Beberapa tahapan dalam metode pengumpulan data yang dilakukan oleh peneliti, yakni antara wawancara, dokumentasi dan studi Pustaka.

2.3 Metode Pengumpulan Data

Pengolahan Data Mining yang dilakukan penulisan ini yaitu, menggunakan tahapan dalam Knowledge Discovery in Database (KDD). Tujuan dari Knowledge Discovery in Database (KDD) adalah untuk menghasilkan informasi berdasarkan urutan yang telah ditentukan dan menemukan pola yang berguna dalam data [18][19][20]. KDD sendiri terdiri dari beberapa langkah yang berbeda dalam prosesnya yakni:

a. *Data Selection*

Penulis melakukan pemilihan data yang sesuai dari data nilai rapor berdasarkan atribut yang terdiri dari akidah-akhlak, al-quran-hadis, bahasa arab, fikih, bahasa inggris, bahasa indonesia, matematika, ilmu pengetahuan alam, ilmu pengetahuan sosial, pjok, ppkn, prakarya, seni budaya sejarah kebudayaan islam, bahasa lampung, total dan ekstrakurikuler.

b. *Data Preprocessing*

Penulis melakukan membersihkan data, dimana *field* yang dianggap tidak sesuai dengan kebutuhan penulisan dapat dihapus. Selain itu, format data diubah kembali untuk memastikan keseragaman format yang konsisten.

c. *Data Transformation*

Proses yang mengubah data agar dapat digunakan dan ditelusuri oleh tools *RapidMiner*. Dengan kata lain, transformasi data melibatkan pemetaan data yang rumit. Seperti mengubah data nama ekstrakurikuler menjadi bilangan angka sehingga mempermudah perhitungannya.

d. *Data Mining*

Penulis melakukan proses pengambilan pola dari data yang tersedia dengan bantuan tools *RapidMiner* menggunakan algoritma *K-Means Clustering*.

e. *Interpretation/evaluation*

Proses mengubah pola menjadi pengetahuan dengan tools *RapidMiner* sehingga dapat dimanfaatkan untuk mendukung pengambilan keputusan berupa pengelompokan data atau *cluster*.

3. HASIL DAN PEMBAHASAN

3.1 Data Selection

Data yang digunakan dalam penelitian ini merupakan data dari LPLPO tahun 2022 kemudian data digabungkan berdasarkan data dari bulan januari sampai bulan desember 2022 dengan jumlah 4.642 data. Dalam hal ini data memiliki atribut nama obat, satuan (jenis obat), stok awal, penerimaan, persediaan, pemakaian, sisa stok, kadaluarsa, permintaan, dan pemberian. Dari keseluruhan atribut yang telah disebutkan sebelumnya pada proses seleksi atribut data terdapat lima faktor atribut yang akan digunakan sebagai acuan perencanaan dan pengendalian kebutuhan obat yaitu, nama obat, satuan (jenis obat), persediaan, permintaan, dan pemakaian obat. Berikut proses hasil *selectin* sesuai tabel 1 data *selection* adalah:

Table 1. Data Selection

No	Nama Obat	Satuan	Persediaan	Permintaan	Pemakaian
1	Albendazol Tablet 400 Mg	tablet	3080	0	980
2	Attapulgite	tablet	1300	500	500
3	Alopurinol Tablet 100 Mg	tablet	1500	500	500
4	Alprazolam 0.25 Mg	Tablet	0	0	0
5	Alprazolam 0.5 Mg	Tablet	281	0	1
6	Ambroxol 30 Mg Tablet	Tablet	3800	3000	1300
7	Ambroxol Syrup 15 Mg/Ml	Botol	0	100	0
8	Aminofilin Injeksi 24 Mg/Ml - 10 Ml	Ampul	0	0	0
9	Aminofilin Tablet 200 Mg	Tablet	500	0	0
10	Amitriptilin Hcl Tablet Salut 25 Mg	Tablet	355	0	40

11	Amlodipine 10 Mg	Tablet	2000	2000	1400
12	Amlodipine 5 Mg	Tablet	3150	2000	1650
13	Amoksisilin Kaplet 500 Mg	Kaplet	7900	4000	3400
14	Amoksisilin Kapsul 250 Mg	Kapsul	1600	0	300
15	Amoksisilin Sirup Kering 125 Mg/5ml	Botol	300	100	100
16	Ampisillin Injeksi 1 Gram	Vial	70	0	13
17	Antalgin Injeksi 25 Mg/ML - 2ml	Ampul	0	0	0
18	Antasida Doen Tablet Kombinasi	Tablet	4300	4000	2300
19	Anti Bakteri Doen Salep Kombinasi	Tube	100	0	0
20	Anti Fungi Doen Kombinasi	Pot	8	0	1
...
4642	Bisacodil 5 Mg Tablet	Tablet	0	0	0

3.2 Preprocessing

Dalam langkah *Preprocessing* yang melibatkan serangkaian tindakan untuk mempersiapkan data sebelum diproses lebih lanjut. Proses *Preprocessing* melakukan pembersihan data untuk meningkatkan kualitas hasil pengelompokan data, dimana data yang dipilih diverifikasi dan diatur kembali agar format data tetap konsisten. Kemudian data dibersihkan dari *Missing Value*, *Noise* data, atau disebut data yang tidak memiliki nilai atau keterangan. Pada penelitian ini penulis menghilangkan data obat yang tidak ada keterangan sama sekali, dimana data *Missing* terdapat di atribut satuan, persediaan, pemakaian dan permintaan. Pada langkah ini, penulis berhasil mengurangi jumlah baris data dalam tabel menjadi 20 data. Berikut hasil proses *preprocessing* sesuai tabel 2 data *preprocessing* adalah:

Table 2. Data Preprocessing

No	Nama Obat	Satuan	Persediaan	Permintaan	Pemakaian
1	Albendazol Tablet 400 Mg	Tablet	3080	0	980
2	Attapulgite	Tablet	1300	500	500
3	Alopurinol Tablet 100 Mg	Tablet	1500	500	500
4	Alprazolam 0.25 Mg	Tablet	0	0	0
5	Alprazolam 0.5 Mg	Tablet	281	0	1
6	Ambroxol 30 Mg Tablet	Tablet	3800	3000	1300
7	Ambroxol Syrup 15 Mg/ML	Botol	0	100	0
8	Aminofilin Injeksi 24 Mg/ML - 10 ML	Ampul	0	0	0
9	Aminofilin Tablet 200 Mg	Tablet	500	0	0
10	Amitriptilin Hcl Tablet Salut 25 Mg	Tablet	355	0	40
11	Amlodipine 10 Mg	Tablet	2000	2000	1400
12	Amlodipine 5 Mg	Tablet	3150	2000	1650
13	Amoksisilin Kaplet 500 Mg	Kaplet	7900	4000	3400
14	Amoksisilin Kapsul 250 Mg	Kapsul	1600	0	300
15	Amoksisilin Sirup Kering 125 Mg/5ml	Botol	300	100	100
16	Ampisillin Injeksi 1 Gram	Vial	70	0	13
17	Antalgin Injeksi 25 Mg/ML - 2ml	Ampul	0	0	0
18	Antasida Doen Tablet Kombinasi	Tablet	4300	4000	2300
19	Anti Bakteri Doen Salep Kombinasi	Tube	100	0	0
20	Anti Fungi Doen Kombinasi	Pot	8	0	1
...
4642	Bisacodil 5 Mg Tablet	Tablet	0	0	0

3.3 Transformation

Dalam memungkinkan data dapat diproses dengan metode pengelompokan *K-Means*, data yang bersifat nominal seperti satuan atau jenis obat perlu diinisialisasikan terlebih dahulu dalam bentuk angka. Karena dalam proses pengelompokan (*Clustering*), data yang dianalisis harus berupa data numerik. Berikut hasil proses transformation sesuai tabel 3 hasil inisialisasi jenis obat adalah:

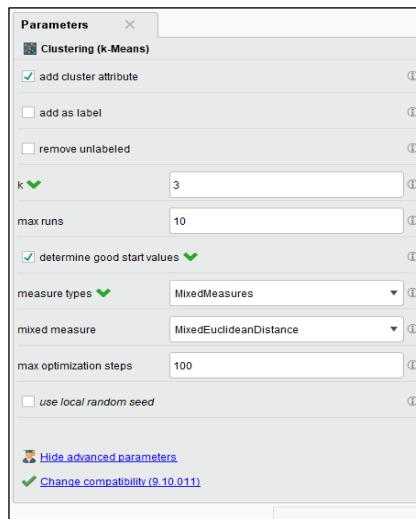
Satuan (Jenis Obat)	Ket
Tablet	1
Biji	2
Rol	3
Botol	4
Kapsul	5
Ampul	6
Kotak	7
Lembar	8
Pak	9
Paket	10
Pasang	11
Pot	12
Capsul	13
Sachet	14
Set	15
Set 2 Tube	16
Supp	17
Kaplet	18
Tube	19
Vial	20
Suppos	21

Gambar 1. Hasil Inisialisasi Jenis Obat

3.4 K-Means Clustering

Setelah mentransformasikan semua data obat pada tahun 2022 menjadi format numerik atau angka, tahap berikutnya adalah mengelompokkan data tersebut menggunakan metode *K-Means Clustering*. Dalam hal ini, penulis menggunakan Software RapidMiner sebagai alat pengolahan data dalam ini dapat mengurangi kesalahan dalam proses pengolahan data obat sehingga memiliki tingkat keakuratan yang tinggi. Untuk dapat melakukan pengelompokan data-data tersebut dalam beberapa Cluster, diperlukan beberapa langkah sesuai dengan pembahasan sebelumnya yaitu :

- Menentukan jumlah faktor (jumlah Cluster K) yang akan diklasterisasi, dalam hal ini penulis telah mengelompokkan data menjadi tiga Cluster dengan obat pemakaian rendah (C1), sedang (C2), dan tinggi (C3). Berikut gambar proses konfigurasi nilai jumlah *cluster* sesuai gambar 2 Konfigurasi Jumlah Cluster adalah:

**Gambar 2.** Konfigurasi Jumlah Cluster

- Menempatkan titik pusat awal dari tiap *Cluster*, dalam penelitian ini peneliti menempatkan titik pusat awal dari setiap *Cluster* dengan menentukan secara acak dan diperoleh titik pusat dari setiap *Cluster*. Berikut tabel titik pusat *cluster* sesuai tabel 4 Pusat *Cluster* Awal adalah:

Table 4. Pusat Cluster Awal

Cluster	Wi	Xi	Yi	Zi
C1	2	100	100	80

C2	9	3900	1000	900
C3	7	60	50	60

- c. Kemudian penting untuk menghitung jarak antara setiap data dengan setiap pusat *Cluster*. Jarak terdekat antara satu data dengan satu *Cluster* tertentu akan menentukan klaster dimana data tersebut termasuk. Untuk menghitung jarak antara semua data dengan setiap pusat klaster, dalam hal ini peneliti menggunakan konsep jarak *Euclidean* yang dirumuskan sebagai berikut :

$$D(i,j) = \sqrt{(X_{1i} - X_{1j})^2 + (X_{2i} - X_{2j})^2 + \dots + (X_{ki} - X_{kj})^2}$$

Dimana :

$D(x,y)$ = Adalah jarak antara x dan y

X_i = yaitu data ke -i pada atribut ke k

Y_i = Titik pusat ke j pada atribut ke k

Melakukan perhitungan jarak antara data pertama dengan pusat kluster pertama yaitu:

$$D11 = \sqrt{(1 - 2)^2 + (1300 - 100)^2 + (500 - 100)^2 + (500 - 80)^2} = 1335.068$$

Melakukan perhitungan jarak antara data pertama dengan pusat kluster ketiga yaitu :

$$D12 = \sqrt{(1 - 9)^2 + (1300 - 3900)^2 + (500 - 1000)^2 + (500 - 900)^2} = 82282.884$$

Melakukan perhitungan jarak antara data pertama dengan pusat kluster ketiga yaitu :

$$D13 = \sqrt{(1 - 7)^2 + (1300 - 60)^2 + (500 - 50)^2 + (500 - 60)^2} = 1391.028$$

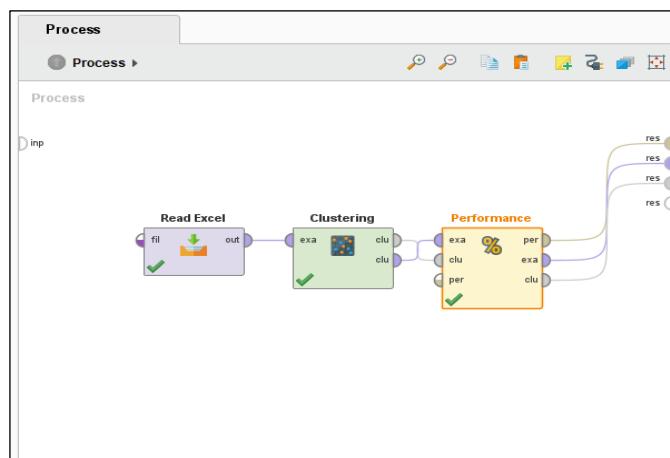
Perhitungan akan berlanjut sampai setiap baris data memiliki nilai untuk setiap klaster. Dalam hal ini, perhitungan dilakukan menggunakan aplikasi RapidMiner, sehingga penulis memperoleh hasil pengolahan data atau analisisnya.

- d. Langkah berikutnya adalah menghitung ulang pusat kluster berdasarkan keanggotaan kluster saat ini. Pusat kluster merupakan nilai rata-rata dari semua data atau objek dalam kluster tertentu. Jika diinginkan, median dari kluster tersebut juga dapat digunakan.
- e. Dan yang terakhir, setiap objek akan diberi penugasan kembali menggunakan pusat kluster yang baru. Jika tidak ada perubahan rasio lagi, maka proses pengelompokan dianggap selesai.

3.5 Interpretation/Evaluation

- a. Hasil Data Mining

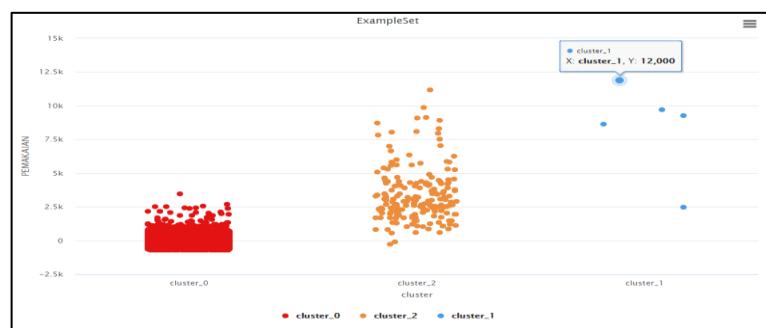
Pada perhitungan jumlah data yang sangat besar, penulis mengaplikasikan operator-operator yang tersedia di aplikasi RapidMiner dengan tujuan memodifikasi data. Data dihubungkan dengan node-node pada operator dan kemudian dapat dihubungkan ke node untuk melihat hasilnya. Dalam hal ini penulis menggunakan konfigurasi dari operator *K-Means Clustering* sebagai pengelompokan dari proses yang telah diimplementasikan. Berikut hasil *interpretation* dengan *tools* adalah:



Gambar 3. Model Penggunaan Algoritma K-Means Clustering

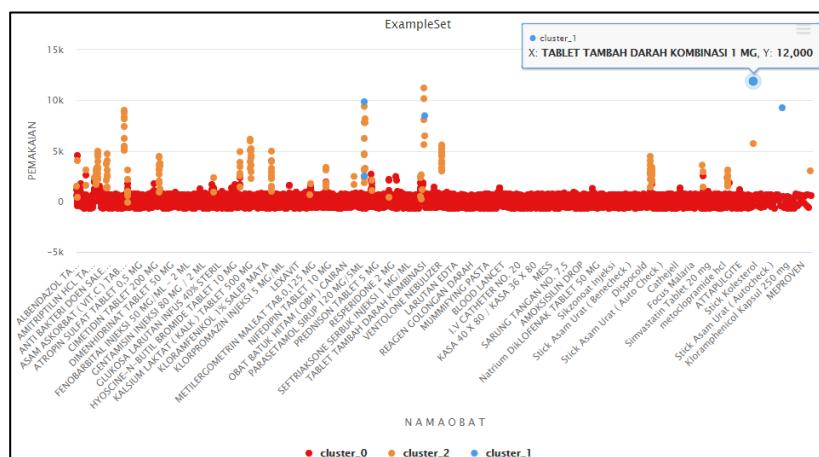
Setelah melakukan konfigurasi klaster dengan tepat, pada proses ini di diperoleh hasil nilai pengelompokan berdasarkan kluster dan jumlah entitas dalam masing-masing klaster. Adapun *Cluster* 1 dinisiasi dengan *Cluster* 0, *Cluster* 2 dengan inisiasi *Cluster* 1, dan *Cluster* 3 diinisiasi dengan *Cluster* 2. Pada *Cluster* model terdapat jumlah data obat yang di bagi menjadi tiga *Cluster* dari 4622 data obat, dimana *Cluster* C1/C0 berjumlah 4433 items, *Cluster* C2/C1 sebanyak 5 items, dan *Cluster* C3/C2 sebanyak 184 items. Dalam visualisasi ini, penulis mengamati garis yang

menunjukkan jumlah kelompok klaster berdasarkan pemakaian obat selama tahun 2022. Pada visualisasi tersebut, terdapat sumbu X di sebelah kiri dan sumbu Y di kanan, dengan warna klaster yang telah di sesuaikan berdasarkan 3 kluster. Pada hasil pengujinya ditemukan bahwa *Cluster 0* (1) memiliki pemakaian obat yang rendah dan di tandai dengan warna merah, *Cluster 1* (2) memiliki pemakaian obat yang tinggi dengan warna biru, sedangkan *Cluster 2* (3) memiliki pemakaian obat sedang yang di tandai dengan warna kuning. Berikut Grafik Visualization Hasil Persebaran *Cluster* Pemakaian Obat sesuai gambar 4 Grafik Visualization Hasil Persebaran *Cluster* Pemakaian Obat adalah:



Gambar 4. Grafik Visualization Hasil Persebaran Cluster Pemakaian Obat

Kemudian pada visualisasi ini, penulis mengamati berdasarkan nama obat dan pemakaian obat tahun 2022. Dengan visualisasi ini penulis dapat mengetahui nama obat yang tergolong ke klaster C1, C2, dan C3 berdasarkan jumlah pemakaiannya selama tahun 2022. Dimana terdapat obat paling banyak digunakan salah satunya yaitu tablet Tambah Darah Kombinasi 1 MG berjumlah 12.000 pemakaian obat dalam tahun 2022. Berikut Grafik Data Obat Berdasarkan Nama Obat dan Pemakaian dan tabel hasil Cluster Permintaan Obat sesuai gambar 5 Grafik Data Obat Berdasarkan Nama Obat dan Pemakaian adalah:



Gambar 5. Grafik Data Obat Berdasarkan Nama Obat dan Pemakaian

Pada gambar 5 diatas merupakan grafik *cluster* distribusi obat berdasarkan pemakaian yang Digambar jadi tiga warna, yakni merah, orange dan biru. Warna merah adalah cluster 0, warna orange adalah cluster 1 dan warna biru adalah cluster 2. Dari gambar tersebut jelas distribusi penggunaan obat banyak pada cluster 0, kemudian disusul cluster 2 dan yang paling sedikit adalah cluster 1.

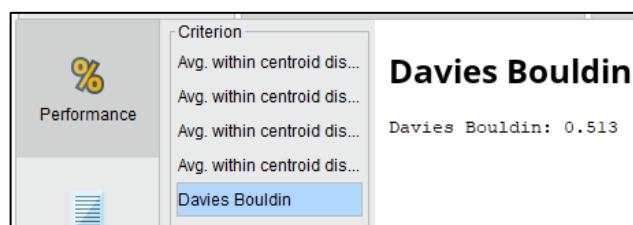
Table 6. Hasil Cluster Permintaan Obat

No.	Nama Obat	satuan	persediaan	permintaan	pemakaian	cluster
1	Albendazol Tablet 400 mg	1	3080	0	980	1
2	Attapulgite	1	1300	500	500	1
3	Alopurinol Tablet 100 mg	1	1500	500	500	1
4	Alprazolam 0.25 mg	1	0	0	0	1
5	Alprazolam 0.5 mg	1	281	0	1	1
6	Ambroxol 30 mg Tablet	1	3800	3000	1300	3
7	Ambroxol Syrup 15 mg/ml	4	0	100	0	1
8	Aminofilin Injeksi 24 mg/ml - 10 ml	6	0	0	0	1
9	Aminofilin Tablet 200 mg	1	500	0	0	1
10	Amitriptilin HCL Tablet Salut 25 mg	1	355	0	40	1
.....
4622	Bisacodil 8 mg Tablet	1	100	0	0	1

Ket:

Tanda (...) = lanjutan

Dalam pengukuran kinerja pengelompokan pemakaian obat pada tahun 2022 menggunakan Davies Bouldin, diperoleh tingkat *Performance* akurasi *Cluster* sebesar 0.513. *Davies Bouldin* memberikan skor Numerik yang menggambarkan kualitas klasterisasi, nilai tingkat akurasi yang diperoleh mendekati 0 (Nol), yang menunjukkan bahwa semakin kecil nilai *Davies Bouldin* yang diperoleh (Non-negatif ≥ 0), semakin baik klasterisasinya, nilai Davies Bouldin dapat berkisar dari 0 hingga tak terbatas, dengan nilai yang lebih rendah atau mendekati 0 menunjukkan hasil klasterisasi yang baik (Nabila et al.,2021). Oleh karena itu, berdasarkan hasil kinerja tersebut, dapat disimpulkan bahwa algoritma K-means menunjukkan *Performance* akurasi yang baik dalam pengelompokan data obat sesuai gambar 6 adalah:



Gambar 6. Performance Akurasi *Cluster*

Setelah mendapatkan hasil dari proses data mining, langkah selanjutnya adalah penyajian pola yang ditemukan sesuai dengan fakta atau hipotesa yang telah ada sebelumnya. Pola ini merupakan *Output* yang berguna dari penelitian ini. *Output* tersebut akan memberikan jawaban terhadap masalah yang telah dirumuskan sejak awal yang akan dijadikan sebagai bahan pembelajaran dalam acuan perencanaan dan pengendalian kebutuhan obat-obatan untuk kedepannya.

b. *Knowledge* (Pengetahuan) Penelitian

Berdasarkan hasil klasterisasi dan analisa tabel data obat pada UPT. Puskesmas Unter Iwes didapatkan pengetahuan sebagai berikut :

1. Sebagian besar obat masuk ke dalam kelompok klaster yang memiliki pemakaian rendah. Oleh karena itu, apotek dapat mengurangi persediaan dan stok obat untuk menghindari akumulasi obat yang pemakaianya rendah (C1/C0), salah satu obat yang pemakaian rendah yaitu Albendazol Tablet 400 MG dengan 4000 pemakaian dari hasil 4433 jenis obat rendah.
2. Pada obat-obatan dengan pemakaian sedang berada di *Cluster* (C3/C2). Dalam hal ini penting untuk menjaga persediaan dan stok obat sesuai jumlah permintaan sebelumnya, guna memastikan ketersediaan yang cukup untuk memenuhi kebutuhan pasien dan menghindari kekurangan obat yang dapat berdampak pada pengobatan pasien. Dimana pemakaian obat yang sedang salah satunya yaitu obat Tablet Tambah Darah Kombinasi dengan jumlah pemakaian yaitu 10.850 obat dari 184 jenis obat sedang.
3. Sedangkan pemakaian obat yang tinggi yang termasuk ke dalam *Cluster* (C2/C1) sebaiknya meningkatkan persediaan dan stok obat. Dengan mempertimbangkan tingkat permintaan yang tinggi dan mengantisipasi kemungkinan lonjakan permintaan, sehingga apotek dapat memastikan ketersediaan obat yang memadai untuk memenuhi kebutuhan pasien. Salah satu obat yang pemakaianya tinggi yaitu Tablet Tambah Darah Kombinasi 1 MG dengan pemakaian 12.000 dari jumlah jenis obat tinggi 5 data obat.

4. KESIMPULAN

Dari hasil analisa dalam Penerapan Algoritma K-Means Clustering Data Obat Pada UPT. Puskesmas Unter Iwes Sumbawa dapat ditarik kesimpulan yaitu, penulis melakukan penelitian di bagian apotek Puskesmas Untir Iwes Sumbawa dengan teknik data mining dan algoritma *K-Means Clustering*. Data dikelompokkan berdasarkan tingkat pemakaian obat yang rendah, sedang, tinggi. Data yang digunakan dalam penelitian ini merupakan data dari LPLPO tahun 2022 kemudian data digabungkan berdasarkan data dari bulan januari sampai bulan desember 2022 dengan jumlah 4.642 data dan data Training berjumlah 4.622 data, hasil kinerja pengelompokan obat menggunakan metode Davies Bouldin menunjukkan tingkat akurasi algoritma *K-Means Clustering* sebesar 0.513. Berdasarkan hasil pengujian dengan dapatkan hasil klaster dengan jumlah data C1/C0 (rendah) berjumlah 4433 items dengan 4000 pemakaian obat, klaster C2/C3 (sedang) dengan jumlah data obat 184 items dengan 10.850 pemakaian obat, dan klaster C3/C2 (tinggi) dengan jumlah data obat 5 items dengan 12.000 pemakaian obat, rekomendasi untuk Apotek UPT. Puskesmas Unter Iwes adalah mengurangi persediaan dan stok obat dengan permintaan rendah (C1), menjaga persediaan obat dengan pemakaian sedang (C3) sesuai jumlah permintaan sebelumnya, dan meningkatkan persediaan obat dengan pemakaian tinggi (C2). Penelitian selanjutnya dapat menentukan jumlah seberapa banyak stok obat atau pemakaian obat yang bisa di kurangi, selain hanya melakukan klasterisasi berdasarkan tingkat pemakaian obat, penelitian selanjutnya dapat menggali lebih dalam untuk menganalisis efisiensi dan efektivitas pengelolaan stok obat. Misalnya, dapat dilakukan analisis waktu pengadaan obat, estimasi persediaan yang optimal dengan menggunakan metode lain dan algoritma lain, penelitian dapat dijadikan sebagai acuan dalam perencanaan dan pengolahan data obat untuk kedepannya.

REFERENCES

- [1] D. Darmawan and Rismawati, "Hasil Pemberdayaan Masyarakat Oleh Koperasi Simpan Pinjam Sebagai Upaya Mewujudkan Kesejahteraan Ekonomi Di Koperasi Serba Usaha Mandiri Kasemen, Kota Serang," vol. 5, no. 2, pp. 205–212, 2020, [Online]. Available: <https://jurnal.untirta.ac.id/index.php/E-Plus/article/download/9260/6037>.
- [2] F. Taslim, "Penerapan algorhythma k-mean untuk clustering data obat pada puskesmas rumbai," *J. Teknol. Inf. dan Komun. Digit. Zo.*, vol. 7, no. 2, pp. 108–114, 2016, doi: 10.31849/digitalzone.v7i2.602.
- [3] B. A. Pangestu, N. A. Kristiawan, and N. Sulistiyowati, "Clustering Obat Untuk Menentukan Pola Pemasaran Efektif di Apotek Amarta Sehat," *J. Ilm. Wahana Pendidik.*, vol. 8, no. 16, pp. 115–126, 2022, [Online]. Available: <https://doi.org/10.5281/zenodo.7058995>
- [4] Z. Nabila, A. Rahman Isnain, and Z. Abidin, "Analisis Data Mining Untuk Clustering Kasus Covid-19 Di Provinsi Lampung Dengan Algoritma K-Means," *J. Teknol. dan Sist. Inf.*, vol. 2, no. 2, p. 100, 2021, [Online]. Available: <http://jim.teknokrat.ac.id/index.php/JTSI>
- [5] M. R. Nugroho, I. E. Hendrawan, and P. P. Purwantoro, "Penerapan Algoritma K-Means Untuk Klasterisasi Data Obat Pada Rumah Sakit ASRI," *Nuansa Inform.*, vol. 16, no. 1, pp. 125–133, 2022, doi: 10.25134/nuansa.v16i1.5294.
- [6] R. Haqmanullah Pembudi, B. Darma Setiawan, and Indriati, "Penerapan Algoritma C4.5 Untuk Memprediksi Nilai Kelulusan Siswa Sekolah Menengah Berdasarkan Faktor Eksternal," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 7, pp. 2637–2643, 2018, [Online]. Available: <http://j-ptiik.ub.ac.id>.
- [7] N. S. Luki Ardiantoro , Sofya Zahara, "Pemanfaatan Knowledge Data Discovery(KDD) Pada Pola Permainan Atlet Bulutangkis," *Explor. IT J. Keilmuan dan Apl. Tek. Inform.*, vol. 11, no. 1, pp. 1–6, 2019, doi: 10.35891/explorit.v11i1.1467.
- [8] H. Haviluddin, Y. Sukmono, D. Suprihanto, A. Harjanto, and O. A. Murtioso, "Drugs Consumption Clustering Model using K-Means Technique," *CESS (Journal Comput. Eng. Syst. Sci.)*, vol. 7, no. 2, p. 537, 2022, doi: 10.24114/cess.v7i2.36104.
- [9] M. Y. Rizki, S. Maysaroh, and A. P. Windarto, "Implementasi K-Means Clushtering Dalam Mengelompokkan Minat Membaca Penduduk Menurut Wilayah," *JUST IT J. Sist. Informasi, Teknol. Inf. dan Komput.*, vol. 11, no. 2, p. 41, 2021, doi: 10.24853/justit.11.2.41-49.
- [10] A. Chusyairi and P. Ramadar Noor Saputra, "Pengelompokan Data Puskesmas Banyuwangi Dalam Pemberian Imunisasi Menggunakan Metode K-Means Clustering," *Telematika*, vol. 12, no. 2, pp. 139–148, 2019, doi: 10.35671/telematika.v12i2.848.
- [11] S. Yudarwati, "Clusterisasi Pola Penjualan Obat Menggunakan Metode K-Means Clustering (Studi Kasus Di Apotek Pelangi Sei. Jering Teluk Kuantan)," *J. Perencanaan, Sains, Teknol. dan Komput.*, vol. 3, no. 1, pp. 148–162, 2020, [Online]. Available: <https://media.neliti.com/media/publications/440973-none-26a0bc65.pdf>
- [12] M. Z. Deny Jollyta, William Ramdhan, Konsep Data Mining dan Penerapan, Deepublish. Yogyakarta: Deepublish, 2020.
- [13] I. Zulfa, R. Rayuwati, and K. Koko, "Implementasi data mining untuk menentukan strategi penjualan buku bekas dengan pola pembelian konsumen menggunakan metode apriori," *Tek. J. Sains dan Teknol.*, vol. 16, no. 1, p. 69, 2020, doi: 10.36055/tjst.v16i1.7601.
- [14] M. R. Nugroho, I. E. Hendrawan, and P. P. Purwantoro, "Penerapan Algoritma K-Means Untuk Klasterisasi Data Obat Pada Rumah Sakit ASRI," *J. Nuansa Inform.*, vol. 16, no. 1, pp. 125–133, 2022, doi: 10.25134/nuansa.v16i1.5294.
- [15] A. Damuri, U. Riyanto, H. Rusdianto, and M. Aminudin, "Implementasi Data Mining dengan Algoritma Naïve Bayes Untuk Klasifikasi Kelayakan Penerima Bantuan Sembako," *JURIKOM (Jurnal Ris. Komputer)*, vol. 8, no. 6, p. 219, 2021, doi: 10.30865/jurikom.v8i6.3655.
- [16] N. Yahya and A. Jananto, "Komparasi Kinerja Algoritma C.45 Dan Naive Bayes Untuk Prediksi Kegiatan Penerimaanmahasiswa Baru (Studi Kasus : Universitas Stikubank Semarang)," in Prosiding SENDI, 2019, no. 2014, pp. 978–979. [Online]. Available: https://www.unisbank.ac.id/ojs/index.php/sendi_u/article/view/7389/2369
- [17] S. Hartati, N. A. Ramdhani, and H. A. San, "Prediksi Kelulusan Mahasiswa Dengan Naïve Bayes Dan Feature Selection Information Gain Student Graduation Prediction With Naïve Bayes and Feature Selection Information Gain," *J. Ilm. Intech Inf. Technol. J. UMUS*, vol. 4, no. 02, pp. 223–234, 2022, [Online]. Available: <http://jurnal.umus.ac.id/index.php/intech/article/view/889>
- [18] N. A. O. Saputri and M. Elvirasari, "Penerapan Metode K-Means Clustering Untuk Menentukan Jumlah Penjualan Obat Yang Banyak Terjual Pada Apotek Murbay Sekayu," *Informnika J.*, vol. 7, no. 2, pp. 44–51, 2021, [Online]. Available: <https://poltekanika.ac.id/journal/index.php/inf/article/download/253/220>
- [19] N. Safitri and C. Bella, "Penggunaan Algoritma Apriori Dalam Penerapan Data Mining Untuk Analisis Pola Pembelian Pelanggan (Studi Kasus: Toko Diengva Bandar Jaya)," *J. Portaldatala*, vol. 2, no. 1, pp. 1–8, 2022, [Online]. Available: <http://portaldatala.org/index.php/portaldatala/article/view/72/72>
- [20] I. A. Rahman, R. Y. Farifah, and F. Hamami, "Implementasi Data Mining Pada Penjualan Toko Sembako Dengan Algoritma Apriori," in eProceedings of Engineering, 2022, vol. 9, no. 2, pp. 638–643. [Online]. Available: <https://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/view/17628>