

Implementasi Algoritma K-Means Untuk Klasterisasi Data Obat Puskesmas Kotabaru

Muhamad Dicky Kurniawan¹, Bayu Priyatna², Fitria Nurapriani³

^{1,2,3}Program Studi Sistem Informasi, Universitas Buana Perjuangan Karawang, Indonesia
e-mail : si19.muhamadkurniawan@mhs.ubpkarawang.ac.id¹, bayu.priyatna@ubpkarawang.ac.id²,
fitria.apriani@ubpkarawang.ac.id³

Abstract

Drug management is one of the things needed to manage drug supplies. Proper planning of drug needs makes drug procurement efficient and effective so that drugs are available in sufficient types and quantities as needed and easily obtained when needed. The purpose of this study was to classify drug data at the Kotabaru Health Center which can be used as a reference in making decisions in planning and controlling drug needs at the Health Center. The data used in this study are the Kotabaru Health Center annual report data from 2019 to 2021. Data processing in this study uses the K-means clustering method with rapidminer tools which is a data grouping technique by dividing the existing data into one or two forms. more clusters. The results of this study divide the drug data into 4 clusters, namely the first cluster (C0) with very low usage consisting of 27 drugs, the second cluster (C2) with low usage consisting of 6 drugs, the third cluster (C3) with high usage consisting of 1 drug, and the fourth (C2) with the highest usage consisting of 1 drug.

Keywords: Data Mining, Clustering, K-Means, Drug Data, community health centre.

Abstrak

Pengelolaan obat merupakan salah satu hal yang diperlukan untuk mengelola persediaan obat. Perencanaan kebutuhan obat yang tepat membuat pengadaan obat menjadi efisien dan efektif sehingga obat tersedia dalam jenis dan jumlah yang cukup sesuai kebutuhan dan mudah diperoleh pada saat dibutuhkan. Tujuan dari penelitian ini adalah untuk mengelompokkan data obat-obatan di Puskesmas Kotabaru yang dapat digunakan sebagai referensi dalam pengambilan keputusan dalam perencanaan dan pengendalian kebutuhan obat di puskesmas tersebut. Data yang digunakan dalam penelitian ini adalah data laporan tahunan Puskesmas Kotabaru tahun 2019 sampai dengan 2021. pengolahan data dalam penelitian ini menggunakan metode K-means Clustering dengan tools rapidminer yang merupakan salah satu teknik pengelompokan data dengan cara membagi data yang ada ke dalam bentuk satu atau lebih cluster. Hasil penelitian ini membagi data obat menjadi 4 cluster yaitu cluster pertama (C0) pemakaian sangat rendah dengan beranggotakan 27 obat, cluster kedua (C2) pemakaian rendah dengan beranggotakan 6 obat, cluster ketiga (C3) dengan pemakaian tinggi yang beranggotakan 1 obat, dan cluster keempat (C2) dengan pemakaian paling tinggi yang beranggotakan 1 obat.

Kata kunci: Data Mining, Clustering, K-Means, Data Obat, Puskesmas.

1. PENDAHULUAN

Puskesmas adalah fasilitas sarana pelayanan kesehatan terdepan dan merupakan ujung tombak pelayanan kesehatan di seluruh Tanah Air, utamanya dalam era Jaminan Kesehatan Nasional (JKN)[1]. Puskesmas Kotabaru merupakan salah satu instansi kesehatan yang selalu berupaya untuk mencapai efisiensi dan efektifitas pelayanan kepada masyarakat, maka diperlukan sarana penunjang pengolahan data yang dapat memberikan informasi secara cepat dan akurat[2]. 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[3]. *Data mining* merupakan metode pengolahan data untuk menemukan pola yang tersembunyi dari data tersebut. Hasil dari pengolahan data dengan metode data mining ini dapat digunakan untuk mengambil keputusan di masa yang akan datang[4]. *Clustering* adalah suatu kegiatan yang bertujuan untuk mengelompokkan data yang memiliki tingkat kesamaan yang tinggi antara satu bagian data dengan yang lainnya, sehingga data pada satu klaster memiliki tingkat kesamaan yang tinggi dan data pada klaster lainnya memiliki tingkat kesamaan yang rendah[5].

Berdasarkan penelitian terdahulu yang dilakukan oleh Muhamad Rizki Nugroho, Iwansyah Edo Hendrawan, dan Purwantoro yang diberi judul “Penerapan Algoritma K-Means Untuk Klasterisasi Data Obat Pada Rumah Sakit ASRI” memiliki tujuan untuk mengelompokkan data obat agar 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. Hasil penelitian tersebut 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[6]. Selanjutnya penelitian yang dilakukan Gustientiedina, M.Hasmil Adiya, dan Yenny Desnelita dengan judul “Penerapan Algoritma K-Means Untuk *Clustering* Data Obat-Obatan Pada RSUD Pekanbaru” metode yang di pakai dalam penelitian ini adalah metode *clustering*. Penelitian tersebut bertujuan untuk mengelompokkan data obat-obatan di Rumah Sakit Umum Daerah Pekanbaru yang nantinya dapat digunakan sebagai referensi dalam pengambilan keputusan dalam perencanaan dan pengendalian pasokan medis di rumah sakit tersebut[7]. Penelitian selanjutnya yang dilakukan Haditsah Annur, yang berjudul “Penerapan Data Mining Menentukan Strategi Penjualan Variasi Mobil Menggunakan Metode K-Means *Clustering*” Pada penelitian ini analisa data mining dilakukan dengan Teknik *Clustering* menggunakan metode K-Means. Hasil akhir dari penelitian ini adalah mengelompokkan data produk yang terjual untuk mengetahui data yang memiliki potensi atau kecenderungan pelanggan dalam membeli barang tersebut[8]. Berdasarkan penelitian-penelitian tersebut maka dalam penelitian ini akan dilakukan penerapan salah satu teknik data mining yaitu *clustering* dengan metode K-means terhadap data obat Puskesmas Kotabaru.

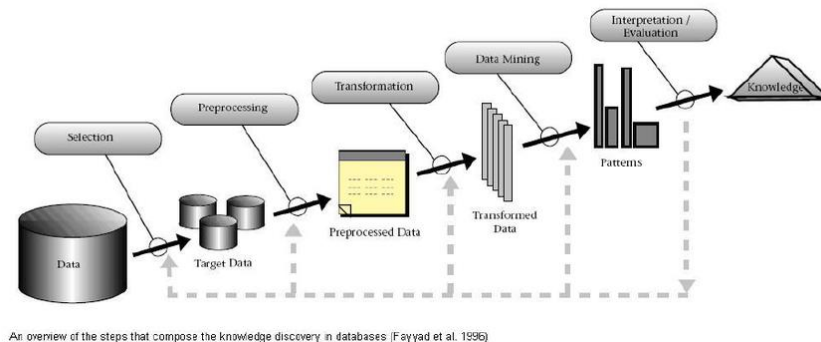
2. METODOLOGI PENELITIAN

Metode penelitian ini menerapkan Algoritma K-Means *Clustering* menggunakan tools rapidminer, Data yang digunakan berdasarkan data laporan tahunan pada Puskesmas Kotabaru tahun 2019, 2020, dan 2021.

2.1. Knowledge Discovery in Database (KDD)

Data Mining merupakan proses pencarian pola-pola yang menarik dan tersembunyi (hidden pattern) dari suatu kumpulan data yang berukuran besar yang tersimpan dalam suatu basis data, seperti data warehouse dan tempat

penyimpanan data lainnya. Data mining juga didefinisikan sebagai bagian dari proses penggalian pengetahuan dalam database yang dikenal dengan istilah Knowledge Discovery in Database (KDD)[9].



Gambar 1. Proses *Knowledge Discovery in Database (KDD)*

Knowledge Discovery in Database (KDD) memiliki segelintir proses meliputi :

2.1.1. Seleksi Data

Merupakan bagian seleksi data dari sekumpulan data operasional yang perlu dilakukan sebelum tahap penggalian informasi dalam KDD dimulai[10]. Data hasil seleksi yang akan digunakan untuk proses data mining, disimpan dalam suatu berkas, terpisah dari basis data operasional,

2.1.2. Pre-processing/Cleaning (pemilihan data)

Sebelum proses data mining dapat dilaksanakan, proses *cleaning* perlu dilakukan pada data yang menjadi fokus KDD. Proses *cleaning* mencakup memeriksa data yang inkonsisten, membuang data ganda, serta memperbaiki kesalahan pada data, seperti kesalahan cetak (tipografi)[11]. Juga dilakukan proses enrichment, yaitu proses “memperkaya” data yang sudah ada dengan data atau informasi lain yang relevan dan diperlukan untuk KDD, seperti data atau informasi eksternal.

2.1.3. Transformasi / Coding

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.

2.1.4. Data mining

Data mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik, metode, atau algoritma dalam data mining sangat bervariasi. Pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses KDD secara keseluruhan.

2.1.5. Interpretasi / Evaluasi

Pola informasi yang dihasilkan dari proses data mining perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan. Tahap ini merupakan bagian dari proses KDD yang disebut dengan interpretation. Tahap ini mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesa yang ada sebelumnya.

2.2. Rapidminer

Rapidminer adalah tools data science open-source untuk melakukan analisis data mining, text mining dan prediction analysis. Pada tools ini, RapidMiner menggunakan teknik deskriptif dan prediksi dalam memberikan wawasan kepada pengguna sehingga dapat membuat keputusan terbaik dari suatu data[12].



Gambar 2. Tools Rapidminer

3. HASIL DAN PEMBAHASAN

3.1. Seleksi Data

Data yang digunakan pada penelitian ini adalah data obat dari Puskesmas Kotabaru yang berbentuk laporan tahunan mulai dari tahun 2019 sampai dengan 2021. Data tersebut dipilih karena merupakan data yang paling terbaru yang bisa diperoleh dari Puskesmas Kotabaru. Data laporan obat tersebut tercampur dengan data alat kesehatan. Berikut salah satu laporan data obat Puskesmas Kotabaru tahun 2021:

LAPORAN TAHUNAN OBAT DAN PERBEKALAN KESEHATAN TANGGAL 31 DESEMBER TAHUN 2021 PUSKESMAS KOTABARU									
NO	NAMA OBAT	SATUAN	HARGA SATUAN	STOK AWAL TAHUN 2021	PENERIMAAN TAHUN 2021	PERSEDIAAN TAHUN 2021	PEMAKAIAN TAHUN 2021	STOK AKHIR 31 DES 2021	TOTAL (Rp)
1	2	3	4	5	6	7=5+6	8	9=7-8	10=9*4
1	ACCU CHEKK SAFE T FPRO UNO	BUAH	640	600	0	600	400	200	128.000
2	ACCU CHEKK SAFE T FPRO UNO	BUAH	650	0	0	0	0	0	0
3	ACETYL CYSTEIN 200 MG	TABLET	382	2.023	4.200	6223	2826	3.397	1.297.654
4	ACYCLOVIR 400 MG	TABLET	405	1.651	1.500	3151	491	2.660	1.077.300
5	ACYCLOVIR KRIM 5 GR	TUBE	2.710	198	0	198	119	79	214.090
6	ALBENDAZOL SYR	BOTOL	4.300	0	600	600	600	0	0
7	ALBENDAZOL TABLET	TABLET	431	0	20.500	20500	20500	0	0
8	ALOPURINOL	TABLET	105	966	4.000	4966	3461	1.505	158.025
9	AMBRDOL SYRUP	BOTOL	2.791	377	200	577	275	302	842.882
10	AMBRDOL TABLET 30 MG	TABLET	100	4.526	4.000	8526	8004	522	52.200
11	AMINOFILIN TABLET	TABLET	95	100	0	100	0	100	9.500
12	AMITRIPTILIN TABLET	TABLET	167	400	0	400	200	200	33.400
13	AMLODIPIN 5 MG	TABLET	53	794	11.700	12494	6898	5.596	296.588
14	AMOXICILIN 500 MG KAPLET	TABLET	243	24.245	10.000	34245	21455	12.790	3.184.710
15	AMOXICILIN 500 MG KAPLET	TABLET	201	0	0	0	0	0	0
16	AMOXICILIN SYRUP 250 MG/ML	BOTOL	3.438	305	200	505	505	0	0
17	AMPICILIN INJEKSI	TUBE	8.237	86	0	86	47	39	321.243
18	ANTALGIN TABLET 500 MG	TABLET	161	0	0	0	0	0	0

Gambar 3. Laporan Tahunan Dan Perbekalan Alat Kesehatan

3.2. Pre-processing/Cleaning (pemilihan data)

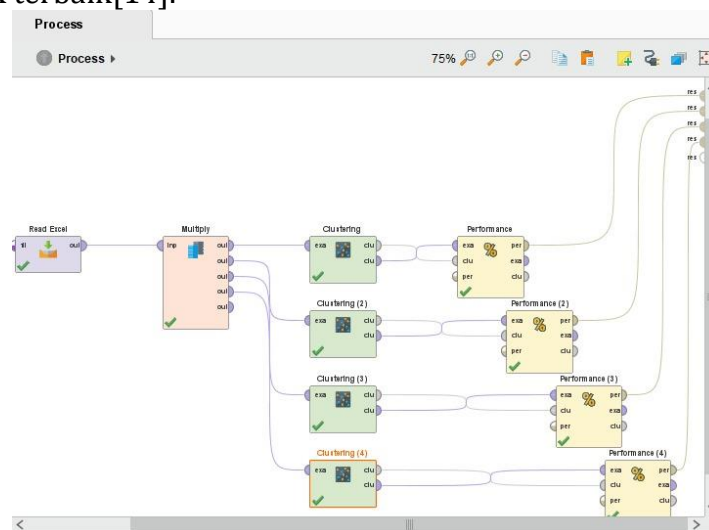
Dari laporan data obat yang diperoleh kemudian dilakukan data *cleaning* untuk membuang atribut yang tidak relevan atau tidak konsisten[13]. Atribut yang dibuang adalah item, satuan, harga satuan, stok awal, penerimaan, persediaan, stok akhir dan total. Atribut yang akan digunakan yaitu nama obat, dan pemakaian keluar. Setelah dilakukan *cleaning* kemudian diambil sampel sebanyak 35 data obat yang ada pada laporan tahunan Puskesmas Kotabaru tahun 2019, 2020, dan 2021. Data obat yang telah di *preprocessing* dapat dilihat pada gambar berikut:

NO	NAMA OBAT	PEMAKAIAN DI TAHUN		
		2021	2020	2019
1	ALBENDAZOL SYR	600	400	387
2	ALBENDAZOL TABLET	20500	24000	19045
3	ALOPURINOL	3461	3146	12679
4	AMBROXOL SYRUP	275	273	570
5	AMBROXOL TABLET 30 MG	8004	7474	4599
6	AMLODIPIN 5 MG	6898	4436	3770
7	AMOXICILIN 500 MG KAPLET	21455	32831	44026
8	ANTALGIN TABLET 500 MG	3233	278	643
9	ANTASIDA DOEN TABLET 200 MG	12921	18698	30959
10	CAPTOPRIL TABLET 25 MG	1196	30	2023
11	CIPROFLOXACIN TABLET 500 MG	962	1669	1379
...
34	TRIEKSIFENIDIL HCL TABLET 2 MG	3000	6010	2710
35	ZINC TABLET 20 MG	3290	4505	5580

Gambar 4. 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 rapidminer. Sebelum dilakukan *Clustering* dilakukan uji performance terhadap K-means pada data obat yang sudah di *cleaning* untuk mencari nilai K terbaik[14].



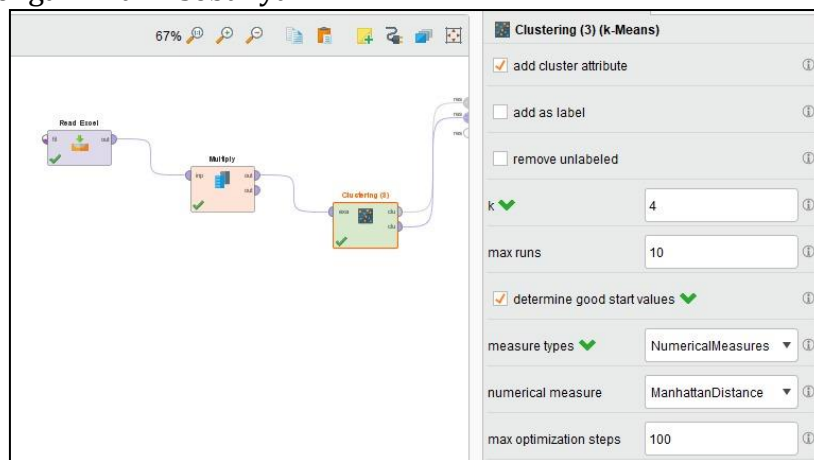
Gambar 5. Proses Pengujian K-means

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:

Tabel 1. Hasil Pengujian Performance Nilai Index Davies Bouldin

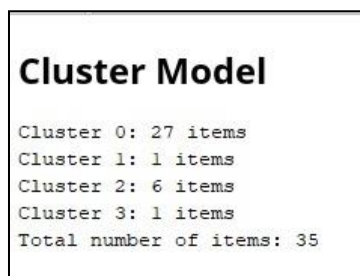
Jumlah Clustering (K)	Nilai Davies Bouldin
4	0.100
2	0.132
5	0.136
3	0.169

Dari hasil tersebut maka dapat dipilih nilai K dengan index davies bouldin terkecil adalah dengan nilai K sebanyak 4, maka dipilihlah K dengan nilai 4. Setelah nilai K didapatkan maka dilanjutkan dengan *clusterisasi* menggunakan algoritma K-means dengan nilai K sebanyak 4.



Gambar 6. Proses *Clustering* K-Means dengan rapidminer

Setelah proses klasterisasi selesai maka didapatkan hasil sebagai berikut:



Gambar 7. Hasil dari Klasterisasi

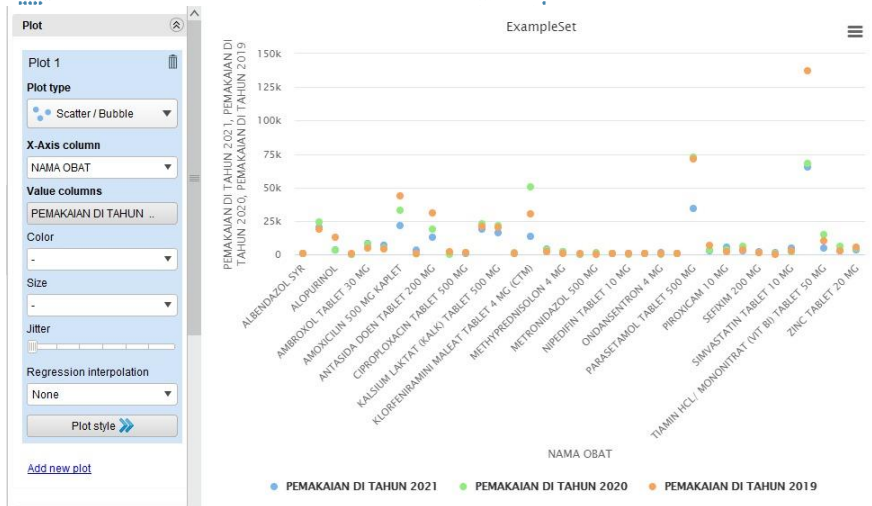
Berikut nilai titik pusat *cluster* yang digunakan sebagai berikut:

Attribute	cluster_0	cluster_1	cluster_2	cluster_3
PEMAKAIAN DI TAHUN 2021	2445.519	65410	17199.667	34169
PEMAKAIAN DI TAHUN 2020	2596.111	67844	28426.833	72454
PEMAKAIAN DI TAHUN 2019	2479.444	137170	27646.500	71085

Gambar 8. Nilai *Centroid*

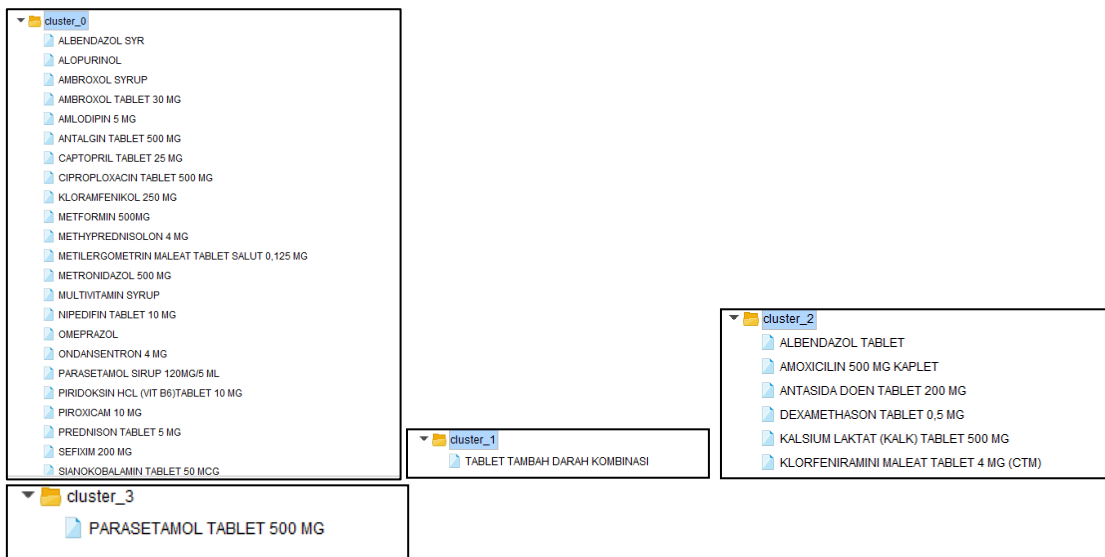
3.4. Interpretasi / Evaluasi

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



Gambar 9. Visualisasi data hasil klusterisasi

Dari data tersebut maka dapat dilihat bahwa *Cluster 0*, merupakan *cluster* obat yang pemakaiannya sangat rendah dan didalamnya terdapat 27 obat, kemudian *cluster 2* beranggotakan 6 obat yang merupakan *cluster* obat dengan pemakaian rendah, lalu *cluster 3* dengan beranggotakan 1 obat dengan pemakain tinggi, dan *cluster 1* yang beranggotakan 1 obat dengan pemakaian sangat tinggi.



Gambar 10. Hasil cluster

4. SIMPULAN

Data yang digunakan merupakan data obat di Puskesmas Kotabaru mulai dari 2019 sampai dengan 2021, sedangkan untuk data tahun 2022 belum dapat

dilihat pada saat pengambilan data karena masih dalam pendataan tahun 2022. Setelah data diperoleh, dilakukan prosedur pembersihan kemudian dilakukan prosedur data mining atau pemodelan terhadap data tersebut, menggunakan algoritma K-means dengan teknik *clustering* dan tools Rapidminer. Dari hasil *clustering* dengan algoritma K-means dapat disimpulkan bahwa *cluster* 0 dengan 27 obat merupakan kelompok pemakaian obat sangat rendah dengan rata-rata pemakaian kurang dari 15.000 obat, *cluster* 2 memiliki anggota 6 obat merupakan kelompok pemakaian obat rendah dengan rata-rata pemakaian kurang dari 29.000 obat, *cluster* 3 memiliki anggota 1 obat merupakan kelompok pemakaian obat tinggi dengan rata-rata pemakaian 58.000 obat, dan *cluster* 1 memiliki anggota 1 obat merupakan kelompok pemakaian obat Paling tinggi dengan rata-rata pemakaian di atas 60.000 obat. Hasil *clustering* ini dapat menjadi salah satu pertimbangan atau knowledge base dalam manajemen obat di Puskesmas Kotabaru pada tahun-tahun berikutnya.

DAFTAR PUSTAKA

- [1] A. Backpropagation, "Jaringan Syaraf Tiruan Memprediksi Kebutuhan Obat-Obatan Menggunakan Metode Backpropagation (Studi Kasus : UPTD Puskesmas Bahorok)," 2021.
- [2] R. Amalia and N. Huda, "Implementasi Sistem Informasi Pelayanan Kesehatan Pada Klinik Smart Medica," vol. 09, no. September, pp. 332–338, 2020.
- [3] D. A. Ramadhanty, R. Syafitri, E. Raswir, and D. Meisak, "Implementasi Data Mining Untuk Menentukan Persediaan Stok Obat Jurnal Informatika Dan Rekayasa Komputer (JAKAKOM)," *J. Inform. dan Rekayasa Komput.*, vol. 1, no. 2, pp. 155–160, 2022.
- [4] D. D. Darmansah and N. W. Wardani, "Analisis Pesebaran Penularan Virus Corona di Provinsi Jawa Tengah Menggunakan Metode K-Means *Clustering*," *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 8, no. 1, pp. 105–117, 2021, doi: 10.35957/jatisi.v8i1.590.
- [5] A. Aditya, I. Jovian, and B. N. Sari, "Implementasi K-Means *Clustering* Ujian Nasional Sekolah Menengah Pertama di Indonesia Tahun 2018/2019," *J. Media Inform. Budidarma*, vol. 4, no. 1, p. 51, 2020, doi: 10.30865/mib.v4i1.1784.
- [6] 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.
- [7] G. Gustientiedina, M. H. Adiya, and Y. Desnelita, "Penerapan Algoritma K-Means Untuk *Clustering* Data Obat-Obatan," *J. Nas. Teknol. dan Sist. Inf.*, vol. 5, no. 1, pp. 17–24, 2019, doi: 10.25077/teknosi.v5i1.2019.17-24.
- [8] H. Annur, "Penerapan Data Mining Menentukan Strategi Penjualan Variasi Mobil Menggunakan Metode K-Means *Clustering* (Studi Kasus Toko Luxor Variasi Gorontalo)," vol. 5, no. 1, 2019.
- [9] S. A. Rahmah, "Klasterisasi Pola Penjualan Pestisida Menggunakan Metode K-Means *Clustering* (Studi Kasus Di Toko Juanda Tani Kecamatan Hutabayu Raja)," vol. 1, no. 1, pp. 1–5, 2020.
- [10] R. N. Fahmi, "Implementasi Metode K-Means *Clustering* dalam Analisis Persebaran UMKM di Jawa Barat," vol. 6, no. 2, pp. 211–220, 2021, doi: 10.33633/joins.v6i2.5310.
- [11] Y. Yulia and A. D. Putri, "Prediksi Kepuasan Mahasiswa terhadap Kinerja Dosen di



- Kota Batam menggunakan Algoritma C4. 5,” *Pros. Semin. Nas. Ilmu Sos. ...*, no. September, pp. 235-240, 2019, [Online]. Available: <https://ejournal.upbatam.ac.id/index.php/prosiding/article/view/1541%0Ahttp://ejournal.upbatam.ac.id/index.php/prosiding/article/download/1541/934>
- [12] Y. R. Sari *et al.*, “Penerapan Algoritma K- Means Untuk *Clustering* Data Kemiskinan Provinsi Banten Menggunakan Rapidminer,” vol. 5, no. 2, pp. 192-198, 2020.
- [13] R. I. Borman *et al.*, “Penerapan Data Maining Dalam Klasifikasi Data Anggota Kopdit Sejahtera Bandarlampung Dengan Algoritma Naïve Bayes,” vol. 9, 2020.
- [14] M. K. Endah Patimah, Dr. Ermatita, M.Kom., Nurul Chamidah, S.Kom, *Analisis Cluster Kepuasan Pengguna Terhadap Layanan Shopee*, vol. 4221. 2021.