

K-MEANS CLUSTERING WITH COMPARISON OF ELBOW AND SILHOUETTE METHODS FOR MEDICINES CLUSTERING BASED ON USER REVIEWS

Safitri Juanita^{*1}, Raynaldi Dwi Cahyono²

^{1,2}Department of Information Systems, Universitas Budi Luhur, Indonesia
Email: ¹safitri.juanita@budiluhur.ac.id, ²raynaldidwicahyono@gmail.com

(Article received: August 25, 2023; Revision: September 19, 2023; published: February 17, 2024)

Abstract

The dissemination of medicine information allows users or customers to make massive assessments of medicines, containing positive and negative reviews—one of a site that provides medicine information online, named drugs.com, contains reviews and ratings of each medicine by variant disease. This site has extensive data collection that has not been processed to produce helpful information for medicine information seekers or the medicine industry. Therefore, research is needed to cluster medicines based on data review from drugs.com. The contribution of this study proposes the best model to cluster user reviews for medicines using K-Means by comparing 2 (two) techniques to determine the optimal number of clustering, Silhouette and Elbow. This study aims to recommend the best K-Means clustering method for processing extensive data reviews and ratings, and cluster results help medical experts, medicine information seekers, or pharmaceutical businesses determine the market share of medicines. The results show that the K-Means model performs best when clustering using the Silhouette method with a DBI value of 0.261 and producing 2 clusters. Meanwhile, the Elbow model has the best performance value of 0.460 and produces 3 clusters. This study also shows that clustering results with both methods produce three medicine cluster groups based on reviews: moderate, unpopular and famous.

Keywords: *clustering, elbow, k-means, medicine, silhouette.*

KLASTERING K-MEANS DENGAN PERBANDINGAN METODE ELBOW DAN SILHOUETTE UNTUK PENGELOMPOKAN OBAT BERDASARKAN ULASAN PENGGUNA

Abstrak

Penyebaran informasi obat memungkinkan pengguna atau pelanggan untuk membuat penilaian besar-besaran terhadap obat-obatan, yang berisi ulasan positif dan negatif-salah satu situs yang menyediakan informasi obat secara online, yaitu drugs.com, berisi ulasan dan peringkat setiap obat berdasarkan varian penyakit. Situs ini memiliki koleksi data yang luas yang belum diolah untuk menghasilkan informasi yang bermanfaat bagi pencari informasi obat atau industri obat. Oleh karena itu, diperlukan penelitian untuk mengelompokkan obat berdasarkan review data dari drugs.com. Kontribusi dari penelitian ini mengusulkan model terbaik untuk mengklaster ulasan pengguna terhadap obat-obatan menggunakan K-Means dengan membandingkan 2 (dua) teknik untuk menentukan jumlah klaster yang optimal, yaitu Silhouette dan Elbow. Penelitian ini bertujuan untuk merekomendasikan metode clustering K-Means yang terbaik untuk memproses data ulasan dan rating yang luas, dan hasil clustering membantu para ahli medis, pencari informasi obat-obatan, atau bisnis farmasi menentukan pangsa pasar obat-obatan. Hasil penelitian menunjukkan bahwa model K-Means memiliki performa terbaik ketika melakukan clustering menggunakan metode Silhouette dengan nilai DBI sebesar 0.261 dan menghasilkan 2 cluster. Sementara itu, model Elbow memiliki nilai performa terbaik sebesar 0.460 dan menghasilkan 3 cluster. Penelitian ini juga menemukan bahwa hasil clustering dengan kedua metode tersebut menghasilkan tiga kelompok klaster obat berdasarkan ulasan yaitu moderat, tidak populer dan terkenal.

Kata kunci: *clustering, elbow, k-means, obat, silhouette.*

1. PENDAHULUAN

Telemedis merupakan solusi terhadap masalah kesenjangan dalam layanan kesehatan, yang mengubah cara masyarakat berkomunikasi dalam

layanan kesehatan [1], khususnya saat pandemi Covid-19 dimana Masyarakat mulai melakukan diagnosis secara mandiri dan mencari informasi

kesehatan dan membeli obat-obatan tanpa menggunakan resep obat [2].

Tersebarinya informasi obat dan proses pemesanan secara daring, membuka peluang bagi pengguna obat untuk melakukan penilaian obat secara massif apakah obat tersebut efektif, kurang efektif dan sebagainya. Website penyedia informasi obat-obatan lengkap salah satunya ada pada situs [drugs.com](https://www.drugs.com), dimana pengguna diijinkan untuk memberikan ulasan obat berdasarkan penyakit tertentu dengan range nilai rating 0-10, sehingga terkumpul puluhan ribu data berkaitan informasi obat. Namun kumpulan data ulasan obat tersebut belum diolah menjadi informasi yang bermanfaat bagi pihak yang berkepentingan salah satunya untuk industri obat sebagai informasi penting untuk pemasaran atau penjualan produk obat.

Salah satu cara mengolah data besar yang belum memiliki label adalah dengan mengelompokkan data tersebut (klustering). Sehingga penelitian ini mengusulkan pengolahan kumpulan data ulasan obat menggunakan algoritma K-Means karena komunitas penelitian menganggap bahwa algoritma ini paling kuat dan populer diantara algoritma klustering lainnya [3], [4].

Pada penelitian ini, pemodelan klustering dengan algoritma K-Means dilakukan dengan membandingkan dua model perhitungan jarak klustering, bernama model Silhouette dan model Elbow. Perbandingan kedua metode tersebut pernah dilakukan sebelumnya untuk mengelompokkan mahasiswa *drop out* [5], dimana pada penelitian tersebut, kedua metode menghasilkan kluster terbaik yang sama yaitu 2 (dua). Begitupun penelitian untuk mengelompokkan tempat wisata berdasarkan ulasan pengunjung, dimana kedua metode tersebut menghasil nilai kluster yang sama yaitu 2 (dua) [6]. Namun berdasarkan studi literatur sebelumnya, belum menemukan penelitian sebelumnya yang melakukan perbandingan K-Means dengan kedua model tersebut untuk klustering obat.

Beberapa penelitian sebelumnya yang melakukan klustering pada kumpulan data obat hanya menggunakan salah satu metode; seperti penelitian

K-Means dengan Silhouette untuk pengelompokkan obat sehingga menjadi acuan dalam aplikasi model monitoring persediaan obat pada puskesmas, hasil klustering terbaik adalah 3 kluster [7], perencanaan persediaan obat di Klinik Citra Medika sehingga terbentuk 5 kluster terbaik [8], serta perencanaan persediaan obat di Rumah sakit ibu dan anak Asri Purwakarta yang menghasilkan 2 kluster terbaik [9]. Selain itu terdapat pula pengelompokkan obat untuk rekomendasi pengganti obat jika obat habis pada apotek RS Grestelina Makassar yang menghasilkan 2 kluster terbaik [10].

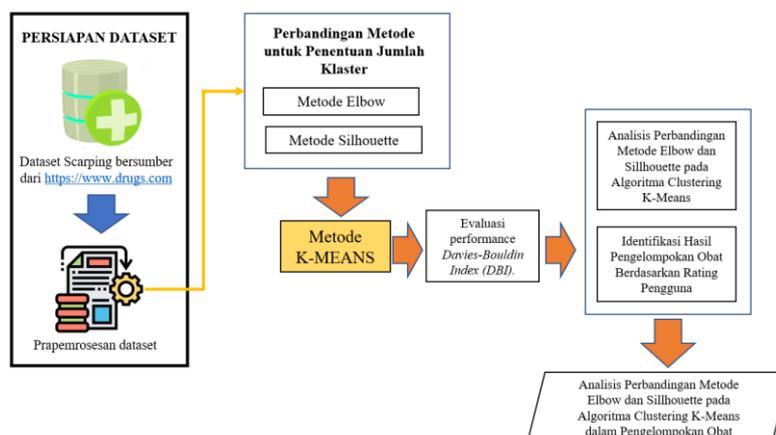
Sedangkan penelitian klustering obat menggunakan algoritma K-Means dengan metode Elbow, seperti; klustering pola pemasaran efektif penjualan obat pada Apotek Amarta Sehat dengan menghasilkan 5 kluster [11]. Pengelompokkan data obat-obatan di Rumah sakit umum daerah (RSUD) Pekan baru dan menghasilkan 3 kluster terbaik [12]. Pengelompokkan data penjualan obat pada Apotek Anugrah Farma Padalarang, yang menghasilkan 4 kelompok penjualan terbaik [13].

Sehingga berdasarkan masalah tersebut, kontribusi penelitian ini adalah menemukan model klustering K-Means terbaik dengan membandingkan performa 2 (dua) metode yaitu metode Elbow sebagai metode perhitungan jarak klustering populer [3],[10] dan metode Silhouette untuk menemukan kelompok obat berdasarkan ulasan pengguna pada [drugs.com](https://www.drugs.com).

Tujuan penelitian ini adalah mengelompokkan obat sesuai jenis penyakit menggunakan jumlah ulasan di [drugs.com](https://www.drugs.com) agar dapat membantu para ahli medis ataupun bisnis obat atau farmasi untuk mengetahui dan merencanakan tindakan pemasaran obat berdasarkan ulasan pengguna.

2. METODE PENELITIAN

Pada bagian ini menjelaskan kerangka penelitian yang berisi tahapan-tahapan yang telah dilakukan dan semua tahapan terdapat pada Gambar 1, kemudian menjelaskan setiap tahapan yang terdapat pada Gambar 1, yaitu:



Gambar 1. Tahapan penelitian untuk pengelompokkan obat berdasarkan ulasan dan rating pengguna membandingkan metode K-Means dengan metode Elbow dan Silhouette

a. Tahap Persiapan Data

Menggunakan dataset yang diambil dari [Kaggle](#) yang merupakan hasil scraping dari website [drugs](#). Dataset ini berisi kumpulan informasi obat berjumlah 2931 data, yang memiliki 17 atribut yaitu `drug_name`, `medical_condition`, `side_effects`, `generic_name`, `drug_classes`, `brand_names`, `activity`, `rx_otc`, `pregnancy_category`, `csa`, `alcohol`, `related_drugs`, `medical_condition_description`, `rating`, `no_of_reviews`, `drug_link`, `medical_condition_url`.

Pada penelitian ini, hanya akan menggunakan 3 (tiga) atribut yaitu: `drug_name` (nama obat), `medical_condition` (kondisi medis/penyakit), dan `rating` (peringkat pengguna yang berisi jawaban pengguna tentang seberapa efektif obat dengan mempertimbangkan efek positif/merugikan dan kemudahan penggunaan, dengan nilai 1 = tidak efektif, 10 = paling efektif). Contoh dataset yang akan digunakan pada penelitian ini, terdapat pada Tabel 1.

Tabel 1. Contoh Dataset Obat beserta Jenis Penyakit dan Ulasannya dari Situs Drugs

<code>drug_name</code>	<code>medical_condition</code>	<code>rating</code>	<code>no_of_reviews</code>
doxycycline	acne	6.4	1598
spironolactone	acne	6.9	668
minocycline	acne	5.9	559
singulair	asthma	6.5	204
montelukast	asthma	5.4	539
prednisone	asthma	7.5	819
benadryl	colds & flu	7.1	71
diphenhydramine	colds & flu	5.4	511
chlorpheniramine	colds & flu	8.1	48

Penelitian ini juga melakukan pembaharuan pada atribut `rating` dan `no_of_reviews`. Tujuan dilakukan pembaharuan dataset pada kedua atribut agar penelitian ini mendapatkan data terbaru dari `rating` dan jumlah pengguna yang melakukan `review`. Proses yang dilakukan pada tahap ini yaitu melakukan pengumpulan data kembali adalah menggunakan atribut `drug_link` yang terdapat pada sumber dataset asli.

Atribut `drug_link` berisi histori dari pengumpul dataset sebelumnya saat melakukan ekstraksi pada situs <https://www.drugs.com/>. Proses pengumpulan data `rating` dan jumlah pengguna yang memberikan `rating` obat pada halaman situs [drugs.com](https://www.drugs.com/) dilakukan pada tanggal 14 Maret 2023. Proses ini menggunakan bahasa pemrograman Python dan `library BeautifulSoup`.

b. Tahap Prapemrosesan Data

Berikut ini adalah tahapan prapemrosesan data menurut [15]: menghilangkan duplikasi data (`remove duplicates`), menghapus data `missing` (data yang tidak ada maka dihapus), menghilangkan inkonsistensi data, `selection` data, tahap ini menentukan atribut yang akan digunakan (3 kolom), transformasi data yang ada pada atribut terpilih ke numerik.

c. Tahap Modeling

Berikut ini adalah beberapa tahapan modeling yang dipakai pada penelitian: melakukan analisis jumlah kluster yang optimal dengan membandingkan kedua metode yaitu metode Elbow dan metode

Silhouette, melakukan proses klustering dengan K-Means.

d. Evaluasi Performa Klustering K-Means

Melakukan evaluasi performa K-Means dengan Davies-Bouldin Index (DBI), menganalisis hasil klustering K-Means dengan membandingkan proses penentuan jumlah kluster dengan 2 metode (metode Elbow dan metode Silhouette), melakukan identifikasi hasil pengelompokan obat berdasarkan `rating` pengguna dengan K-Means serta membandingkan 2 metode untuk penentuan jumlah kluster.

3. HASIL DAN PEMBAHASAN

Pada tahap ini akan dilakukan beberapa proses.

3.1. Seleksi data

Berdasarkan atribut yang terdapat pada tabel dataset, pada tahap ini melakukan seleksi atribut yang berkaitan dengan penelitian ini dan menggunakan hanya beberapa atribut yang terdapat pada dataset seperti yang terdapat pada Tabel 2.

Tabel 2. Atribut Penelitian Pengelompokan Obat Berdasarkan Rating

Nama atribut	Deskripsi
<code>drug_name</code>	Nama obat
<code>medical_condition</code>	Jenis penyakit
<code>rating</code>	kolom peringkat, pengguna ditanya seberapa efektif mereka menggunakan obat tertentu, dengan mempertimbangkan efek positif/merugikan dan kemudahan penggunaan. Nilai <code>rating</code> dari 1 hingga 10 (1 = tidak efektif, hingga 10 = paling efektif).
<code>no_of_reviews</code>	Jumlah pengguna obat yang memberikan <code>rating</code> pada obat tertentu

3.2. Pra-Pemrosesan Data

Berikut ini adalah beberapa tahapan untuk membersihkan dataset penelitian yaitu penghapusan data duplikat: Menggunakan fungsi `drop_duplicates()` dari `library pandas`. Penghapusan baris dengan data kosong: Menggunakan fungsi `dropna()` dari `library pandas`. Hapus spasi pada awal dan akhir data: Menggunakan fungsi `apply()` dari `library pandas`. Mengubah data menjadi lowercase: Menggunakan fungsi `apply()` dari `library pandas`.

Setelah melalui proses prapemrosesan, maka dataset mengalami penyusutan menjadi 1160 data. Dataset tersebut yang kemudian diolah pada tahap berikutnya.

3.3. Klustering dengan K-Means

Pada tahap ini adalah beberapa proses yang dilakukan saat melakukan klustering menggunakan K-Means untuk mengelompokkan obat berdasarkan jumlah pengguna yang memberikan `rating`. Proses eksperimen dilakukan dengan menggunakan kode program Python dengan `library SKLEARN`.

1. Penentuan Jumlah Kluster dengan Perbandingan Metode Elbow dan Silhouette.

Proses eksperimen pada tahap ini menggunakan 2 (dua) metode yang kemudian dibandingkan, yaitu metode Elbow dan Silhouette. Pada Tabel 3 adalah skenario atau model eksperimen yang dilakukan untuk membandingkan kedua metode (Elbow dan Silhouette), model ini juga membandingkan proses optimasi kluster terbaik dengan membandingkan proses saat menggunakan model scaling dan klustering serta saat tidak menggunakan keduanya.

Tabel 3. Model Eksperimen untuk Klustering K-Means untuk membandingkan Dua metode (Elbow dan Silhouette)

Percobaan	Elbow dan Silhouette	
	Cluster	Scaling
1	Ya	Ya
2	Ya	Tidak
3	Tidak	Ya
4	Tidak	Tidak

Pada Tabel 4 adalah hasil percobaan dari 4 model eksperimen yang terdapat pada Gambar 2.

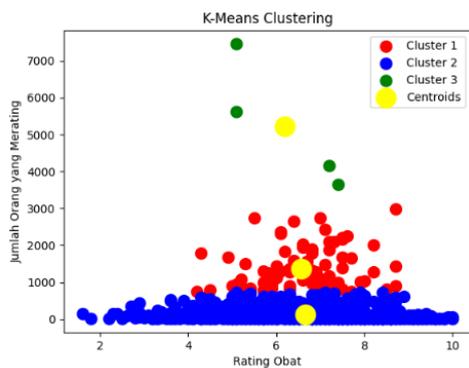
Tabel 4. Hasil Perbandingan Model untuk Penentuan Kluster Terbaik dengan Algoritma K-Means

Percobaan	Elbow		Silhouette	
	DBI	Cluster	DBI	Cluster
1	0.502	3	0.261	2
2	0.622	3	0.619	2
3	0.703	3	0.645	4
4	0.460	3	0.448	2

Berdasarkan hasil eksperimen pada Tabel 3 maka dapat diambil kesimpulan bahwa nilai DBI terbaik menggunakan klustering K-Means dengan metode Elbow terdapat pada percobaan 4, dimana nilai DBI adalah 0.460 dengan jumlah kluster sebanyak 3. Sedangkan nilai DBI terbaik dengan menggunakan klustering K-Means menggunakan metode Silhouette, terdapat pada percobaan 1, dimana nilai DBI adalah 0.261 dan jumlah kluster sebanyak 2.

2. Perbandingan Tampilan Hasil Kluster menggunakan Algoritma K-Means dengan Metode Elbow dan Silhouette

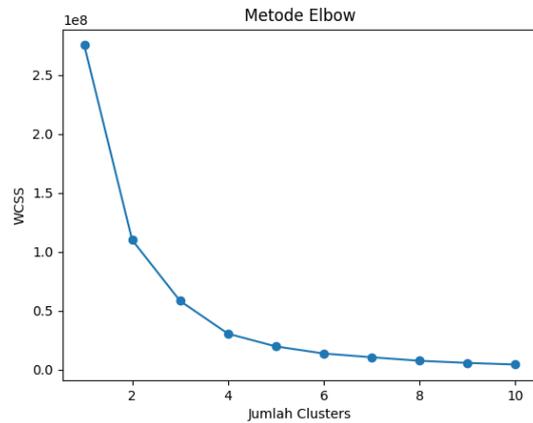
a. Algoritma K-Means dengan Metode Elbow



Gambar 2. Tampilan Hasil Terbaik Klustering K-Means Menggunakan Metode Elbow

Gambar 2 merupakan tampilan hasil terbaik klustering K-Means menggunakan metode Elbow.

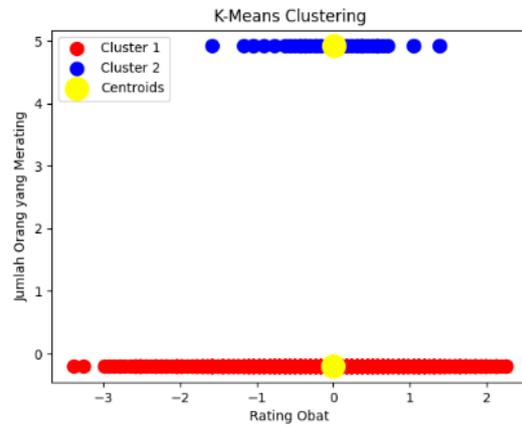
Sedangkan Gambar 3 tampilan grafik dengan performa terbaik menentukan jumlah kluster menggunakan klustering k-means dengan metode Elbow.



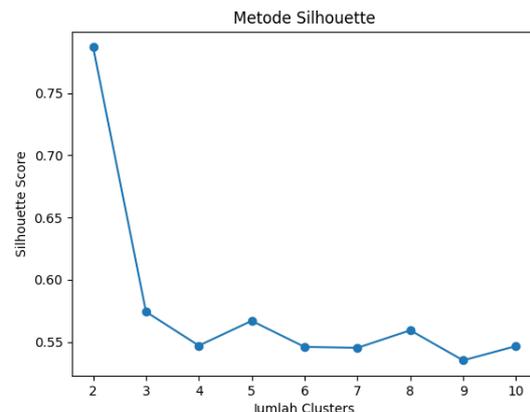
Gambar 3. Tampilan Grafik Performa Terbaik Klustering K-Means Dengan Metode Elbow untuk Menentukan Jumlah Kluster

b. K-Means dengan Metode Silhouette

Gambar 4 merupakan tampilan hasil terbaik klustering K-Means dengan metode Silhouette. Sedangkan Gambar 5 tampilan grafik dengan performa terbaik menentukan jumlah kluster menggunakan klustering K-Means dengan metode Silhouette.



Gambar 4. Tampilan hasil Terbaik Klustering K-Means dengan Metode Silhouette

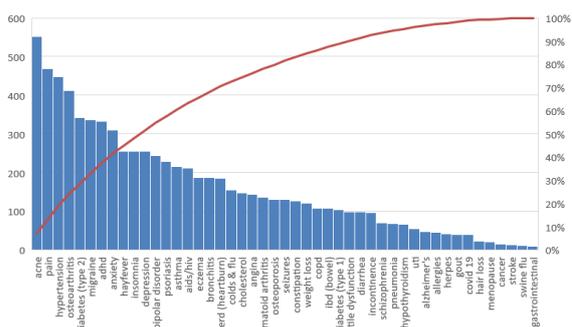


Gambar 5. Tampilan Grafik Hasil Terbaik Klustering K-Means Dengan Metode Silhouette

3. Tampilan Hasil Sebaran Obat berdasarkan Perbandingan Algoritma K-Means dengan Metode Elbow dan Silhouette.

a. Tampilan Hasil Klastering K-Means dengan menggunakan metode Elbow

Gambar 6 adalah hasil sebaran pengelompokan obat berdasarkan penyakit menggunakan klastering K-Means dengan metode Elbow, dan pada Tabel 5 merupakan detail jumlah obat berdasarkan hasil klastering terbaik, dimana pada kluster 1 terdapat 80 nama obat, kluster 2 sebanyak 1075 nama obat, dan kluster 3 sebanyak 4 nama obat.



Gambar 6 Sebaran obat berdasarkan Jenis Penyakit menggunakan K-Means dengan metode Elbow

Tabel 5. Hasil Terbaik Jumlah Klastering Menggunakan K-Means dengan Metode Elbow

	Jumlah obat	Rentang Jumlah Ulasan
Kluster 1	80	752-2981
Kluster 2	1075	8-732
Kluster 3	4	3648-7453

Pada Tabel 6 hingga Tabel 8 menampilkan contoh sebaran obat berdasarkan 3 (tiga) klaster, dimana pada Tabel 6 menampilkan contoh data yang merupakan kelompok yang berisi nama obat dan jenis penyakit pada kluster 1, pada Tabel 7 menampilkan contoh data kluster 2, begitupun pada Tabel 8 menampilkan contoh data pada kluster 3.

Tabel 6. Contoh Hasil K-Means dengan Metode Elbow untuk Kelompok Obat berdasarkan Jenis Penyakit Pada Kluster 1

Nama Obat	Jenis Penyakit	Jumlah Ulasan
victoza	diabetes (type 2)	752
amlodipine	angina	755
clindamycin	pneumonia	781
ozempic	diabetes (type 2)	785
macrobid	uti	798
diazepam	anxiety	808

Tabel 7. Contoh Hasil K-Means dengan Metode Elbow untuk Kelompok Obat berdasarkan Jenis Penyakit Pada Kluster 2

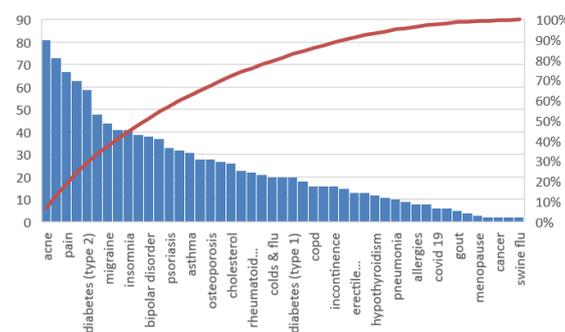
Nama Obat	Jenis Penyakit	Jumlah Ulasan
tadalafil	erectile dysfunction	732
levofloxacin	bronchitis	727
clonidine	adhd	726
lamictal	bipolar disorder	721
epiduo	acne	705
risperidone	anxiety	697
adalimumab	psoriasis	681
phentermine/ topiramate	weight loss	681

Tabel 8. Contoh Hasil K-Means dengan Metode Elbow untuk Kelompok Obat berdasarkan Jenis Penyakit Pada Kluster 3

Nama Obat	Jenis Penyakit	Jumlah Ulasan
escitalopram	anxiety	3648
sertraline	bipolar disorder	4170
Ethinyl estradiol /norgestimate	acne	5623
ethinyl estradiol/ norethindrone	acne	7453

b. Tampilan Hasil Klastering K-Means dengan menggunakan metode Silhouette.

Gambar 8 adalah hasil sebaran pengelompokan obat berdasarkan penyakit menggunakan klastering K-Means dengan metode Silhouette, dimana pada kluster 1 terdapat 1113 nama obat, kluster 2 sebanyak 46 nama obat.



Gambar 7 Sebaran obat berdasarkan Jenis Penyakit menggunakan K-Means dengan metode Silhouette

Pada Tabel 9 merupakan detail jumlah obat berdasarkan hasil DBI terbaik Implementasi klastering menggunakan K-Means dengan metode Elbow.

Tabel 9. Hasil Terbaik Jumlah Klastering menggunakan K-Means dengan metode Silhouette

	Jumlah obat	Jumlah Ulasan
Cluster 1	1113	8-1107
Cluster 2	46	1128-7453

Pada Tabel 10 dan Tabel 11 menampilkan contoh kelompok obat berdasarkan 2 klaster yang merupakan hasil klastering K-Means menggunakan metode Silhouette. Dimana pada Tabel 10 menampilkan contoh data pada kluster 1 yang berisi Kelompok obat berdasarkan jenis penyakit, pada Tabel 11 menampilkan contoh data obat berdasarkan jenis penyakit pada kluster 2.

Tabel 10. Contoh Hasil Klastering K-Means dengan Metode Silhouette yang Menampilkan Kelompok Obat berdasarkan Jenis Penyakit Pada Kluster 1

Nama Obat	Jenis Penyakit	Jumlah Ulasan
oxycodone	pain	1107
pristiq	depression	1095
amphetamine/ dextroamphetamine	adhd	1090
dulcolax	constipation	1075
trintellix	depression	1054

Tabel 11. Contoh Hasil Klastering K-Means dengan Metode Silhouette yang Menampilkan Kelompok Obat berdasarkan Jenis Penyakit Pada Klaster 2

Nama Obat	Jenis Penyakit	Jumlah Ulasan
abilify	bipolar disorder	1128
semaglutide	diabetes (type 2)	1163
prozac	depression	1172
buspirone	anxiety	1265
azithromycin	bronchitis	1269

3.4. Evaluasi Hasil Klastering K-Means

Berdasarkan hasil klastering K-Means, maka dapat disimpulkan bahwa hasil klaster merupakan kelompok obat dengan nilai rating campuran, namun klaster tersebut mengelompok berdasarkan jumlah ulasan dengan karakter sebagai berikut:

a. Kelompok obat yang moderat.

Kelompok obat ini mendapatkan ulasan yang cukup untuk memberikan gambaran tentang pengalaman pengguna obat, tetapi tidak mencapai tingkat yang sangat besar atau signifikan. Kelompok ini menunjukkan bahwa obat yang terdapat pada klaster ini bisa jadi belum terlalu menarik bagi pengguna obat.

b. Kelompok obat yang kurang dikenal.

Kelompok obat ini mendapatkan ulasan yang sedikit diantara kelompok klaster lainnya. Ada beberapa penyebab, mungkin karena obat tersebut baru atau kurang populer, kurang aktifnya perusahaan obat dalam mengenalkan produk obat, atau pengalaman negatif dari pengguna.

c. Kelompok obat yang populer.

Kelompok obat ini mendapatkan ulasan berlimpah, hal ini menunjukkan bahwa obat-obatan pada kelompok ini telah banyak dipakai atau banyak yang tertarik dengan obat-obatan pada kelompok ini. Ulasan yang banyak dari pengguna obat membantu membangun peluang dan meningkatkan kesempatan bagi pelanggan lainnya untuk yakin terhadap kualitas dan keunggulan dari obat yang terdapat pada kelompok ini.

4. DISKUSI

Hasil klastering K-Means menggunakan 2 metode yaitu Elbow dan Silhouette menghasilkan jumlah kelompok klastering yang berbeda. Menggunakan 3 jenis klastering yaitu kelompok obat moderat, obat kurang dikenal dan obat yang populer, maka dapat disimpulkan bahwa luaran dari klastering K-Means dengan metode Elbow menghasilkan 3 klaster dimana, klaster 1 adalah kelompok obat yang moderat, klaster 2 adalah kelompok obat yang kurang dikenal, dan klaster 3 adalah kelompok obat yang populer.

Sedangkan hasil klastering K-Means dengan metode Silhouette menghasilkan 2 klaster, dimana klaster 1 adalah kelompok obat yang moderat, sedangkan klaster 2 adalah kelompok obat yang populer.

5. KESIMPULAN

Penelitian ini melakukan analisis terhadap algoritma klastering K-Means untuk menemukan model terbaik yang dapat mengelompokkan obat berdasarkan jumlah ulasan. Berdasarkan hasil eksperimen dengan membandingkan kedua metode yaitu Silhouette dan Elbow, menemukan bahwa performa K-Means dengan metode Silhouette lebih unggul dengan nilai DBI 0.261 dibandingkan K-Means dengan metode Elbow mendapat nilai DBI 0.460. Hasil eksperimen pada penelitian ini juga menemukan bahwa hasil klastering dengan kedua metode menghasilkan 3 jenis kelompok klaster pada obat berdasarkan ulasan yaitu moderat, kurang dikenal dan populer. Kedua metode pengukuran jarak klaster terbaik menghasilkan nilai klaster yang berbeda, dimana K-Means dengan metode Silhouette menghasilkan 2 klaster, sedangkan metode Elbow menghasilkan 3 klaster.

UCAPAN TERIMA KASIH

Penelitian ini didanai melalui kegiatan penelitian Hibah Internal dari Universitas Budi Luhur Tahun Anggaran 2023.

DAFTAR PUSTAKA

- [1] L. M. Ganiem, "Efek Telemedicine Pada Masyarakat (Kajian Hukum Media McLuhan: Tetrad)," *Interak. J. Ilmu Komun.*, vol. 9, no. 2, pp. 87–97, 2021, doi: 10.14710/interaksi.9.2.87-97.
- [2] I. Maskanah, "Fenomena Self-Diagnosis di Era Pandemi COVID-19 dan Dampaknya terhadap Kesehatan Mental," *JoPS J. Psychol. Students*, vol. 1, no. 1, pp. 1–10, 2022, doi: 10.15575/jops.v1i1.17467.
- [3] M. Ahmed, R. Seraj, and S. M. S. Islam, "The k-means algorithm: A comprehensive survey and performance evaluation," *Electronics*, vol. 9, pp. 1–12, 2020, doi: 10.3390/electronics9081295.
- [4] Sekar Setyaningtyas, B. Indarmawan Nugroho, and Z. Arif, "Tinjauan Pustaka Sistematis: Penerapan Data Mining Teknik Clustering Algoritma K-Means," *J. Teknoif Tek. Inform. Inst. Teknol. Padang*, vol. 10, no. 2, pp. 52–61, 2022, doi: 10.21063/jtif.2022.v10.2.52-61.
- [5] M. Guntara and N. Lutfi, "Cacah Klaster pada Klasterisasi dengan Algoritma K-Means Menggunakan Silhouette Coefficient dan Elbow Method," *Juti "Jurnal Teknol. Informasi"*, vol. 2, no. 1, pp. 43–52, 2023, doi: 10.26798/juti.v2i1.944.
- [6] M. Sholeh and K. Aeni, "Perbandingan Evaluasi Metode Davies Bouldin, Elbow Dan Silhouette Pada Model Clustering Dengan

- Menggunakan Algoritma K Means,” *STRING (Satuan Tulisan Ris. dan Inov. Teknol.*, vol. 8, no. 1, pp. 56–65, 2023.
- [7] H. Haviluddin, Y. Sukmono, D. Suprihanto, A. Harjanto, and O. A. Murtioso, “Model Clustering Pemakaian Obat Menggunakan Algoritma K-Means,” *CESS (Journal Comput. Eng. Syst. Sci.*, vol. 7, no. 2, pp. 537–546, 2022, doi: 10.24114/cess.v7i2.36104.
- [8] L. Izzah and A. Jananto, “Penerapan Algoritma K-Means Clustering Untuk Perencanaan Kebutuhan Obat Di Klinik Citra Medika,” *Progresif J. Ilm. Komput.*, vol. 18, no. 1, pp. 69–76, 2022, doi: 10.35889/progresif.v18i1.769.
- [9] M. R. Nugroho, I. E. Hendrawan, and 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.
- [10] S. Paembonan and H. Abduh, “Penerapan Metode Silhouette Coefficient untuk Evaluasi Clustering Obat,” *PENA Tek. J. Ilm. Ilmu-Ilmu Tek.*, vol. 6, no. 2, pp. 48–54, 2021, doi: 10.51557/pt_jiit.v6i2.659.
- [11] 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>.
- [12] G. Gustientiedina, M. H. Adiya, and Y. Desnelita, “Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan pada RSUD Pekanbaru,” *J. Nas. Teknol. dan Sist. Inf.*, vol. 5, no. 1, pp. 17–24, 2019, doi: 10.25077/teknosi.v5i1.2019.17-24.
- [13] R. A. Febrianty, W. Witanti, and P. N. Sabrina, “Segmentasi Penjualan Obat Di Apotek Menggunakan Metode K-Means,” in *Prosiding SISFOTEK*, 2020, pp. 200–206, [Online]. Available: <http://www.seminar.iaii.or.id/index.php/SISFOTEK/article/view/213>.
- [14] M. Cui, “Introduction to the K-Means Clustering Algorithm Based on the Elbow Method,” *Accounting, Audit. Financ.*, vol. 1, pp. 5–8, 2020, doi: 10.23977/accaf.2020.010102.
- [15] C. Li, “Preprocessing Methods and Pipelines of Data Mining: An Overview,” in *Seminar Data Mining*, 2019, pp. 1–7, [Online]. Available: <http://arxiv.org/abs/1906.08510>.