

Implementasi *K-Means* Untuk Klasterisasi Kasus Penyalahgunaan Narkoba di Provinsi Sulawesi Selatan

Ikmar Mawardi^{a,1}, Dolly Indra^{b,1}, Herdianti Darwis^{c,1}, Muh. Raihan Alif Muliawan^{d,1}, Rahma Puspitasari^{e,2}

¹Universitas Muslim Indonesia, Makassar, ²Universitas Negeri Malang, Malang, Indonesia

^a13020190084@umi.ac.id; ^bdolly.indra@umi.ac.id; ^cherdianti.darwis@umi.ac.id; ^d13120210045@umi.ac.id, ^erahma.puspitasari.2505348@um.ac.id

Received: 09-08-2025 | Revised: 20-11-2025 | Accepted: 10-11-2025 | Published: 29-12-2025

Abstrak

Penyalahgunaan narkoba merupakan salah satu masalah paling mendesak dan kompleks di Indonesia, yang ditandai dengan meningkatnya jumlah pecandu narkoba, jumlah kasus kejahatan narkoba yang ditemukan, serta semakin beragamnya model dan jaringan distribusi. Penelitian ini bertujuan untuk melakukan klasterisasi penyalahgunaan narkoba di Provinsi Sulawesi Selatan dengan menerapkan metode *K-Means*. Klasterisasi dilakukan pada data pengguna dalam kasus penyalahgunaan narkoba di BNN Provinsi Sulawesi Selatan dengan mengelompokkan setiap sampel ke dalam klaster yang berbeda. Hasil penelitian ini memberikan informasi mengenai gambaran setiap variabel yang memiliki distribusi data berbeda berdasarkan jenis zat yang digunakan pada klaster tertentu. Pengujian metode menggunakan evaluasi *Silhouette Coefficient* dan *Elbow* menunjukkan terhadap 4 variabel dengan $K = 2$ memiliki nilai terbaik sebesar 0,623.

Kata kunci: Drugs, *K-Means*, Visualization, *Silhouette Coefficient*, *Elbow*.

Pendahuluan

Narkotika dan Obat Terlarang (Narkoba) merupakan zat kimia yang dapat mengubah psikologi seperti perasaan, pikiran, suasana hati serta perilaku jika masuk ke dalam tubuh manusia baik dengan cara dimakan, diminum, dihirup, suntik, intravena dan lain sebagainya [1]. Saat ini penyalahgunaan narkoba sangat merajalela, karena terlihat semakin banyaknya pengguna narkoba di masyarakat dan peredarannya yang terus meningkat. Berbagai cara telah dilakukan oleh pemerintah untuk memberantas penyalahgunaan narkoba, salah satunya dengan hukuman mati. Namun, cara tersebut tidak terlalu dipedulikan oleh para pengguna narkoba [2].

Narkoba mengacu pada zat atau obat yang berasal dari tumbuhan dan bukan dari tumbuhan. Baik sintetis maupun semi-sintetik, Obat tersebut dapat menyebabkan penurunan atau perubahan kesadaran, hilangnya rasa, hilangnya aroma, dan menyebabkan kecanduan, merupakan bahan kimia yang diharapkan dapat menyembuhkan kesehatan. Saat zat ini masuk ke dalam organ, fungsinya akan mengalami satu atau lebih perubahan. Kemudian, situasi ini akan berlanjut dengan kecanduan fisik dan psikologis. Jika tubuh tidak mengizinkan zat tersebut untuk digunakan, maka penyakit fisik dan mental akan terjadi. Hal ini menimbulkan efek mati rasa, yang dapat menimbulkan adiksi [3].

Penyalahgunaan narkoba merupakan salah satu masalah yang mendesak dan kompleks di Indonesia, yang ditandai dengan terus meningkatnya jumlah pecandu narkoba, banyaknya kasus kejahatan narkoba yang ditemukan, serta model dan jaringan pengedaran yang semakin beragam. Penyalahgunaan narkoba terjadi pada kelompok tertentu seperti kelompok umur tertentu atau kelompok ekonomi rendah [4].

Indonesia darurat narkoba tercatat di Badan Narkotika Nasional (BNN) terdapat 40.553 kasus narkoba sepanjang tahun 2018, 2.287.492 pelajar aktif menggunakan narkoba atau 3,21% dan untuk pekerja yang tercatat melakukan penyalahgunaan narkoba sebanyak 1.514.037 jiwa. *The United Nations Office on Drugs and Crime* (UNODC) juga menemukan 275 juta orang atau 5,6% penduduk dunia dalam rentang usia 15 hingga 64 tahun pernah mengonsumsi narkoba minimal sekali dalam hidupnya, Sementara di Indonesia, sebanyak 3.376.115 orang pada tahun 2017 dengan rentang usia 10-59 tahun menyalahgunakan narkoba. Penyalahgunaan narkoba dapat menimbulkan beberapa dampak psikologis bagi penggunanya, seperti lamban dan ceroboh dalam bekerja, sering tegang dan gelisah, hilangnya kepercayaan diri, apatis, pengkhayal, curiga, agitatif, ganas atau brutal, kesal, sulit berkonsentrasi, pendidikan menjadi terganggu, tidak mendapatkan penerimaan sosial, merepotkan dan menjadi beban keluarga, serta mempunyai masa depan suram [5].

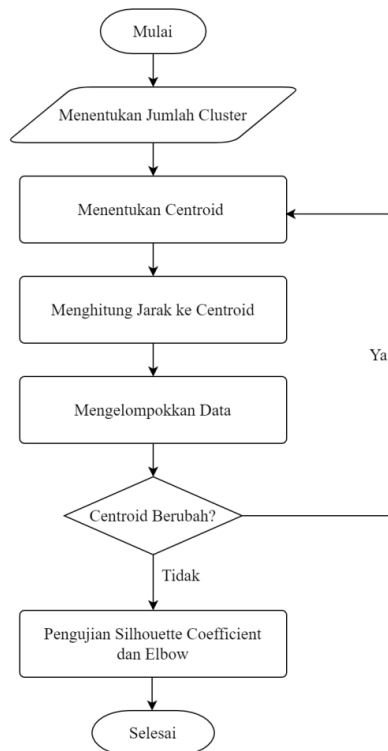
Untuk mengetahui tingkat penyalahgunaan narkoba salah satunya dengan menggunakan algoritma klasterisasi K-Means. Klasterisasi atau *clustering* merupakan metode pengelompokan data yang berarti proses partisi satu set objek data ke dalam himpunan bagian yang disebut dengan *cluster*. Objek yang di dalam *cluster* memiliki kemiripan karakteristik antar satu sama lainnya dan berbeda dengan *cluster* yang lain. Partisi tidak dilakukan secara manual melainkan dengan suatu algoritma *clustering* [6], [7]. Klasterisasi dikenal juga dengan *unsupervised learning*, yakni metode pengelompokan tanpa pengawasan yang menemukan desain dari kumpulan data tanpa mengacu pada hasil yang diketahui atau diberi nama sehingga tidak diperlukan adanya variabel target dalam pengklusteran [8].

Algoritma K-Means adalah algoritma *clustering* yang paling sederhana dibanding algoritma clustering yang lain, mudah diterapkan dan dijalankan, relatif cepat, mudah untuk diadaptasi, dan paling banyak dipraktikkan dalam tugas *Data Mining* [9]. K-Means juga merupakan metode pengelompokan paling sederhana yang mengelompokkan data ke dalam *k* kelompok berdasar centroid masing-masing kelompok. Hanya saja hasil dari K-Means sangat dipengaruhi parameter *k* dan inisialisasi centroid. Pada umumnya K-Means menginisialisasi centroid secara acak [10].

Berdasarkan latar belakang tersebut, penelitian kami berfokus pada cara menganalisis *clustering* algoritma K-Means menggunakan metode evaluasi *Silhouette Coefficient* dan *Elbow* dalam kasus penyalahgunaan narkoba di BNN Provinsi Sulawesi Selatan.

Metode

Penelitian ini dilakukan dengan menggunakan metode algoritma K-Means pada dataset kasus penyalahgunaan narkoba. Untuk metode evaluasi menggunakan *Silhouette Coefficient* dan *Elbow*. *Silhouette coefficient* merupakan gabungan dari dua metode yaitu metode *cohesion* dan *separation*. Dimana *cohesion* bertujuan mengukur seberapa dekat hubungan antara objek dalam sebuah *cluster*, dan metode *separation* bertujuan mengukur seberapa jauh sebuah *cluster* terpisah dengan *cluster* lain. Yang mana, jumlah *cluster* terbaik adalah jumlah *cluster* yang memiliki rata rata nilai *silhouette coefficient* tertinggi atau mendekati 1 [11]. *Elbow* adalah metode di mana pada suatu titik tertentu terjadi penurunan yang signifikan dalam grafik, berbentuk lengkungan yang tajam. Nilainya kemudian akan menjadi nilai *k* atau banyaknya *cluster* yang baik. Gambar 1 menunjukkan kerangka kerja metode yang digunakan [12].



Gambar 1. Algoritma K-Means Clustering

Gambar 1 merupakan tahap penerapan algoritma *K-Means* pada data-data yang telah dikumpulkan, dengan alur sebagai berikut:

- a. Menentukan jumlah *cluster* yang akan dipergunakan dalam pembagian data.
- b. Membangkitkan *centroid* awal yang diperoleh secara acak serta jumlah *centroid* sebanyak *cluster* yang akan dibuat. Pengertian *centroid* sendiri adalah titik pusat *cluster* atau awal pusat *cluster* [13]. Perhitungan jarak dilakukan dengan menggunakan persamaan *Euclidean Distance* yakni [14] :

$$D_{euclidian}(x, y) = \sqrt{\sum_{k=1}^n (x_k - y_k)^2} \tag{1}$$

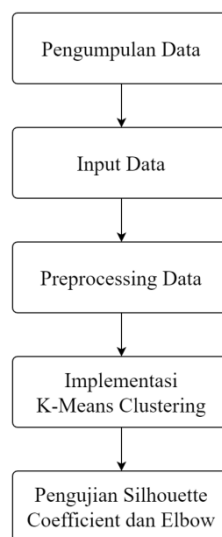
Dimana:

x : pusat *cluster*

y : data

- c. Mengelompokkan setiap data terhadap jarak pada titik pusat *centroid* terdekat.
- d. Melakukan iterasi ke-*i*, kemudian tentukan posisi *centroid* baru dengan menggunakan persamaan sebelumnya.
- e. Mengulangi langkah pada poin c jika posisi *centroid* baru tidak sama dan iterasi atau perulangan akan berhenti jika nilai rasio tidak lebih besar dari nilai rasio sebelumnya hingga hasil perngitungan di masing-masing data konvergen.
- f. Melakukan pengujian evaluasi hasil *clustering* dengan menggunakan *Silhouette Coefficient* dan *Elbow*. Pada *Elbow* hasil evaluasi hanya ditampilkan dalam bentuk grafik.

Alur sistem dalam penelitian ini berisi tahapan-tahapan yang akan dilakukan secara logis dan sistematis. Sistem penelitian yang diusulkan ditunjukkan pada Gambar 2.



Gambar 2. Alur Sistem Penelitian.

Alur sistem penelitian pada Gambar 2 dijelaskan dibawah ini.

1. Pengumpulan data, dilakukan untuk memperoleh informasi dan data yang akan dijadikan sebagai bahan penelitian. Data diperoleh dari kantor BNN Provinsi Sulawesi Selatan. Rincian jumlah data disajikan pada Tabel 1.

Tabel 1. Data jumlah pengguna narkoba

Tahun	Jumlah Pengguna Narkoba
2020	237
2021	144
2022	183

2. *Input* data, yakni tahap memasukkan data yang digunakan dalam penelitian ke dalam sistem.
3. *Preprocessing* data, mengubah data menjadi format yang lebih mudah dan efektif digunakan dalam pemrosesan selanjutnya serta menghasilkan dataset yang bersih sehingga dapat digunakan dalam pengimplementasian *K-Means* [15] dan dilakukan pelabelan data untuk mengubah label menjadi kode

biner dengan menetapkan label atau kategori tertentu pada setiap *instance* dalam dataset guna memprediksi kategori atau kelas dari suatu data [16].

4. Mengimplementasikan metode *K-Means* dan memvisualisasikan hasil *clustering* pada data kasus penyalahgunaan narkoba di Provinsi Sulawesi Selatan menggunakan pemrograman *python* di *Google Collaboratory* berupa grafik *scatter plot* untuk mengetahui pola sebaran data.
5. Pengujian hasil *clustering* pada *cluster* yang telah dihasilkan dengan menggunakan metode *Silhouette Coefficient* dan *Elbow* untuk melihat seberapa baiknya *clustering* yang sudah dilakukan [17]. Rata-rata silhouette value untuk semua *cluster* [18].

$$sil(c) = sil(k) \frac{1}{|k|} \sum_{i=1}^k sil(c_i) \tag{2}$$

Dimana:

$sil(k)$: nilai *silhouette* semua *cluster*

$|k|$: banyaknya *cluster* k

$sil(c_i)$: rata-rata nilai *silhouette*

Untuk menghitung algoritma *elbow* diterapkan dengan mempertimbangkan *sum of square error* (SSE) untuk menentukan jumlah kelompok yang terbentuk secara efektif [19].

$$SSE = \sum_{k=1}^k \sum_{x_i \in S_k} \|X_i - C_k\|^2 \tag{3}$$

Dimana:

k : banyaknya *cluster* k

X_i : jumlah data

C_k : banyaknya *clustering* pada *cluster* ke k

Hasil dan Analisis

A. Pengolahan Data

Pada tahap ini dilakukan *preprocessing* data pada data yang telah di input ke dalam sistem. Hasil *preprocessing* data ditunjukkan pada Gambar 3.

No.	Nama(Inisial)	Usia	Jenis Kelamin	Agama	Pendidikan	Pekerjaan	Alamat	Status Pernikahan	Zat Yang Digunakan	
0	1.0	W	25.0	L	ISLAM	SMA	KARYAWAN	GOWA	BELUM MENIKAH	SABU
1	2.0	AA	20.0	L	ISLAM	SMA	Tidak Bekerja	MAKASSAR	BELUM MENIKAH	SABU, SINTE
2	3.0	AMK	20.0	L	ISLAM	SMK	Tidak Bekerja	MAKASSAR	BELUM MENIKAH	SABU, SINTE
3	4.0	AJ	27.0	L	ISLAM	SMP	BURUH	MAKASSAR	MENIKAH	SABU
4	5.0	AF	25.0	L	ISLAM	SMK	SECURITY	MAKASSAR	BELUM MENIKAH	SABU
...
559	560.0	LAZ	26.0	L	ISLAM	SD	WIRASWASTA	MAKASSAR	BELUM MENIKAH	SABU
560	561.0	AG	22.0	L	ISLAM	SMP	WIRASWASTA	MAKASSAR	BELUM MENIKAH	SABU
561	562.0	PH	31.0	L	ISLAM	SD	WIRASWASTA	MAKASSAR	BELUM MENIKAH	SABU
562	563.0	IH	23.0	L	ISLAM	SD	WIRASWASTA	MAKASSAR	BELUM MENIKAH	SABU
563	564.0	KBP	33.0	L	ISLAM	SMA	KARYAWAN	MAKASSAR	MENIKAH	SABU

Gambar 3. *Preprocessing* data

Selanjutnya, dilakukan pelabelan data pada beberapa variabel yaitu usia, jenis kelamin, agama, pendidikan, pekerjaan, alamat, status pernikahan dan zat yang digunakan pada data hasil *preprocessing* yang ditunjukkan pada Gambar 4.

No.	Nama(Inisial)	Usia	Jenis Kelamin	Agama	Pendidikan	Pekerjaan	Alamat	Status Pernikahan	Zat Yang Digunakan	
0	1.0	W	2	1	1	4	4	6	1	6
1	2.0	AA	2	1	1	4	2	10	1	16
2	3.0	AMK	2	1	1	5	2	10	1	16
3	4.0	AJ	3	1	1	3	4	10	2	6
4	5.0	AF	2	1	1	5	4	10	1	6
...
559	560.0	LAZ	3	1	1	2	5	10	1	6
560	561.0	AG	2	1	1	3	5	10	1	6
561	562.0	PH	3	1	1	2	5	10	1	6
562	563.0	IH	2	1	1	2	5	10	1	6
563	564.0	KBP	3	1	1	4	4	10	2	6

Gambar 4. Pelabelan data

B. *Exploratory Data Analysis*

Tahap ini dilakukan *exploratory data analysis* dengan menghitung nilai mean, standar deviasi, minimum, maksimum, dan kuartil untuk melihat distribusi data guna membantu dalam membuat keputusan yang lebih baik terkait dengan analisis lebih lanjut atau pemodelan. Pada Gambar 5 memperlihatkan sebaran atau distribusi data untuk setiap variabel yang digunakan.

	Usia	Jenis Kelamin	Agama	Pendidikan	Pekerjaan \
count	562.000000	562.000000	562.000000	562.000000	562.000000
mean	2.617438	1.081851	1.012456	3.706406	3.046263
std	0.965306	0.274381	0.111006	1.048347	1.255858
min	1.000000	1.000000	1.000000	2.000000	1.000000
25%	2.000000	1.000000	1.000000	3.000000	2.000000
50%	2.000000	1.000000	1.000000	4.000000	4.000000
75%	3.000000	1.000000	1.000000	4.000000	4.000000
max	6.000000	2.000000	2.000000	6.000000	5.000000

	Alamat	Status Pernikahan	Zat Yang Digunakan
count	562.000000	562.000000	562.000000
mean	10.959075	1.514235	1.201068
std	4.350929	0.654596	0.534512
min	1.000000	1.000000	1.000000
25%	10.000000	1.000000	1.000000
50%	10.000000	1.000000	1.000000
75%	11.000000	2.000000	1.000000
max	20.000000	3.000000	4.000000

Gambar 5. Distribusi data

C. Implementasi Metode K-Means

1. Mencetak *Centroid*

Pada tahap ini, data setiap sampel yang telah dilakukan preprocessing akan di implementasikan dengan metode K-Means dan pada penelitian ini digunakan nilai $K = 3$ dan menggunakan nilai random state = 0. Gambar 6 menunjukkan nilai hasil centroid dari semua variabel pada data yang digunakan.

```
[[ 2.57723577  1.09756098  1.01626016  3.64769648  2.95934959 10.2899729
  1.48780488  1.19783198]
 [ 2.80769231  1.05769231  1.00961538  3.76923077  3.32692308 18.61538462
  1.59615385  1.25      ]
 [ 2.56179775  1.04494382  1.      3.87640449  3.07865169  4.78651685
  1.52808989  1.15730337]]
```

Gambar 6. Hasil *Centroid*

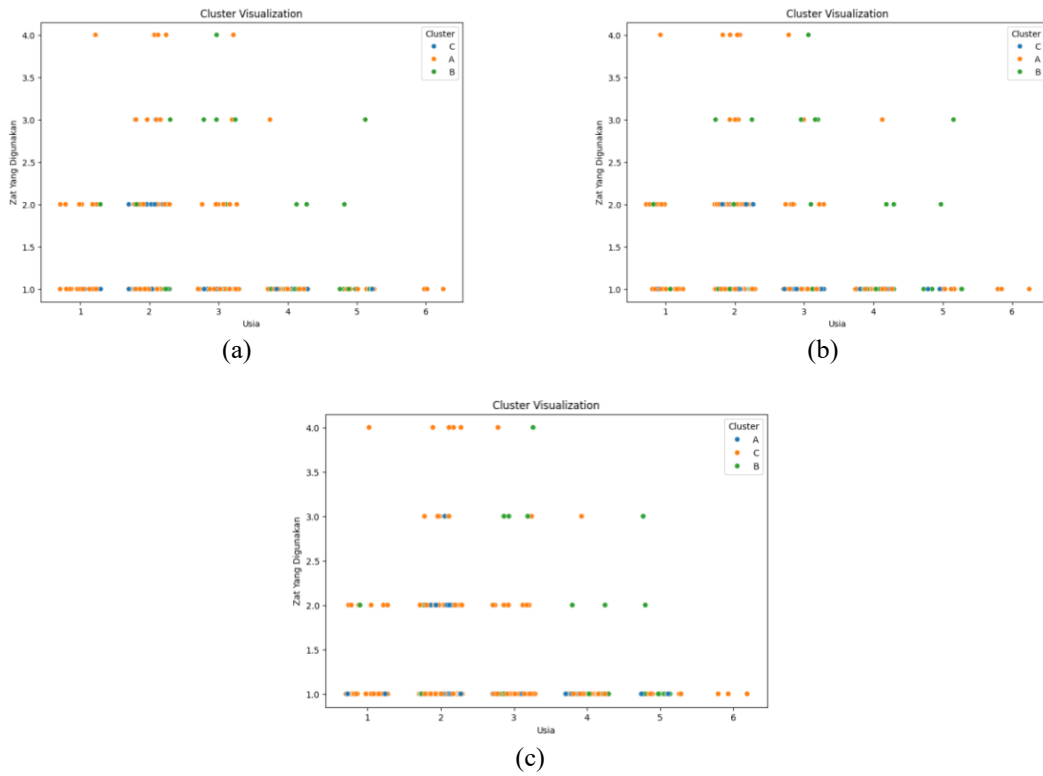
Pada Gambar 7 menunjukkan hasil pengklasteran pada seluruh data yang digunakan.

```
KMeans Labels:
[2 0 0 0 0 0 0 0 0 0 0 0 0 2 0 0 0 1 1 0 0 0 0 2 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 1 2 0 0 0 0 0 0 0 0 0 0 0 0 0 0 2 2 2 2 2 2 0 0 0 1 0 2 0 0 2 2 0
 0 1 0 2 0 1 0 0 0 0 0 0 0 0 0 0 2 0 0 0 0 0 0 0 0 0 0 0 2 2 2 0 0 0 2 2 2
 0 0 1 0 0 0 0 0 0 0 0 0 1 1 0 0 0 2 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1
 0 1 0 0 0 0 0 0 0 0 2 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 2 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 2 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 2 2 2 0 0 0 1 2 0 0 0 0 0 0 0 0 0
 0 0 0 1 0 2 2 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 2 1 1
 1 1 1 1 1 0 2 0 1 0 1 1 2 1 2 2 1 1 0 0 0 0 2 1 0 0 0 0 0 0 2 0 0 2 2 2 0
 0 2 2 0 0 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 2 2 0 0 0 2 0 0 0 2 0 0 0 1
 1 1 1 1 1 1 1 1 2 2 2 2 2 0 2 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 1 0 0 0 0 0 0 1 1 1 0 0 0 0 0 0 0 0 0 0 0 2 0 0 0 0 1 1 2 0 1 1 0 1 0
 1 0 0 1 1 1 0 2 0 0 0 0 0 0 0 0 2 0 2 0 1 1 2 0 0 0 1 1 1 1 1 1 0 0 1 0
 2 1 2 2 2 2 2 0 0 0 0 0 1 0 0 0 0 2 0 0 1 0 1 0 0 0 0 0 0 0 0 0 0 1 1 0 0
 1 1 1 1 1 1 0 0 2 0 0 0 0 2 2 2 2 0 2 2 2 2 2 2 2 2 1 0 2 0 0 1 0 0 0 0 1
 1 0 0 2 0 0 0 1 0 1 0 2 1 0 0 0 0 1 2 0 1 0 0 0 0 0 0 2 0 0 0 2 0 0 0 0 2
 2 0 0 0 0 0]
```

Gambar 7. Hasil Pengklasteran

2. Visualisasi *Cluster*

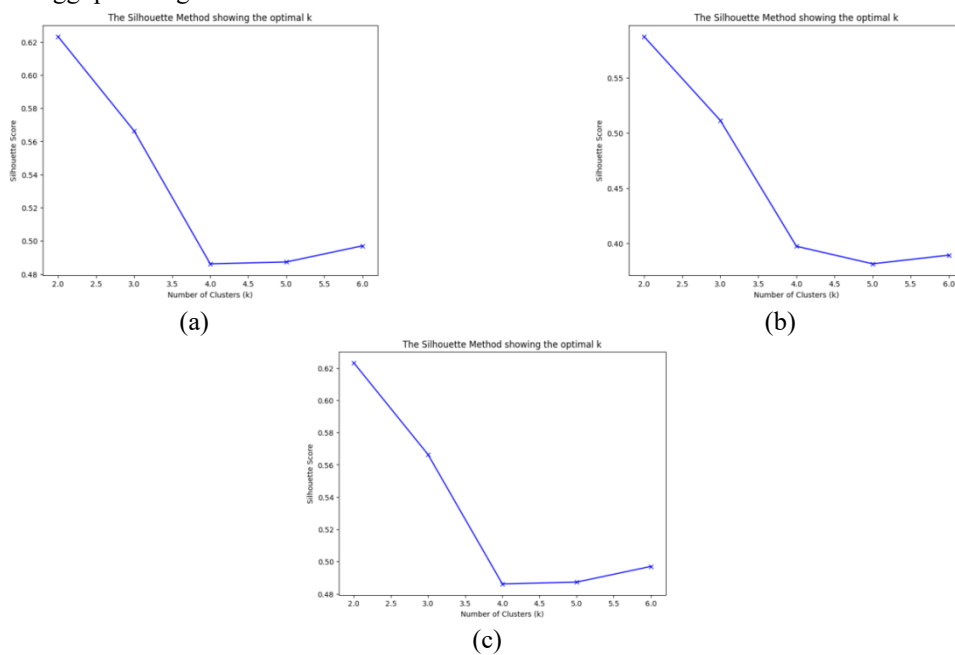
Visualisasi merupakan bentuk penyampaian berbagai informasi yang dimaksudkan untuk menguraikan berbagai hal berupa gambar hingga diagram yang dapat dilakukan berbagai pengolahan [20]. Dilakukan visualisasi terhadap beberapa *clustering* dengan menggunakan K-Means $K = 3$. Gambar 8 menampilkan hasil visualisasi *clustering* variabel usia berdasarkan zat yang digunakan yakni 8 variabel, 6 variabel, dan 4 variabel.



Gambar 8. Visualisasi *cluster* usia berdasarkan zat yang digunakan dengan $K=3$ berdasarkan (a)clustering 8 variabel, (b)clustering 6 variabel, (c)clustering 4 variabel.

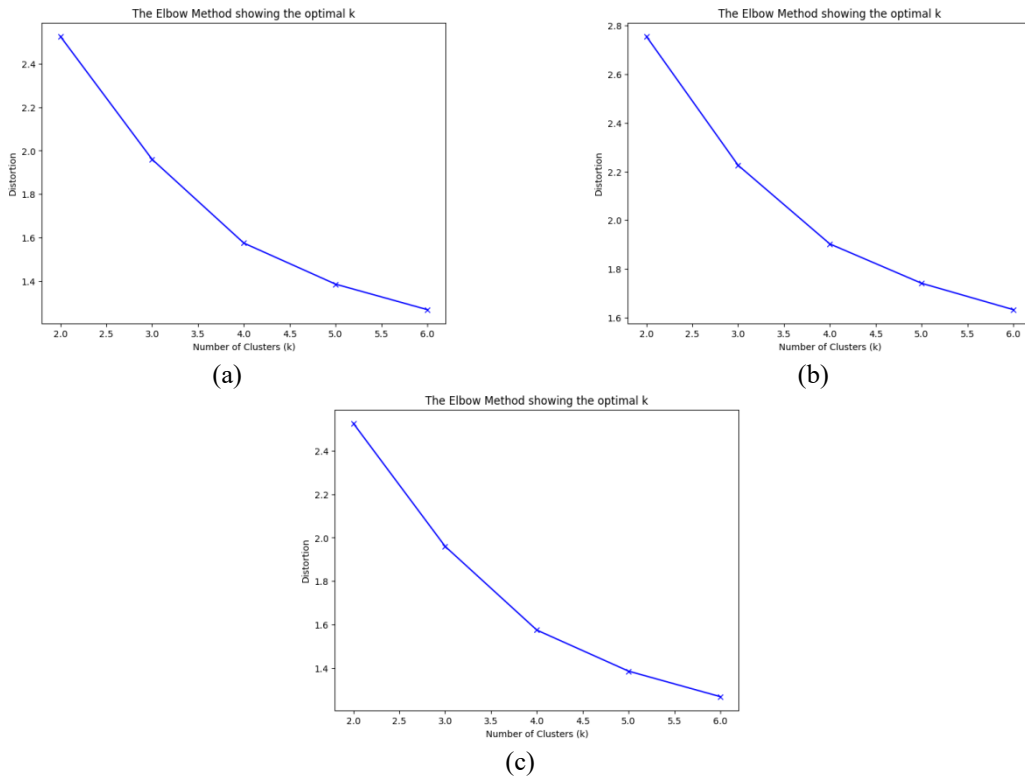
3. Evaluasi

Tahap ini ditampilkan hasil evaluasi menggunakan *Silhouette Coefficient* dan *Elbow* yang disajikan dalam bentuk grafik . Diperoleh bahwa dengan 8, 6 dan 4 variabel yang digunakan, maka jumlah kelompok disarankan *Silhouette method* sebanyak 2 *cluster* sebagaimana pada Gambar 9 dengan titik tertinggi pada angka 2.0.



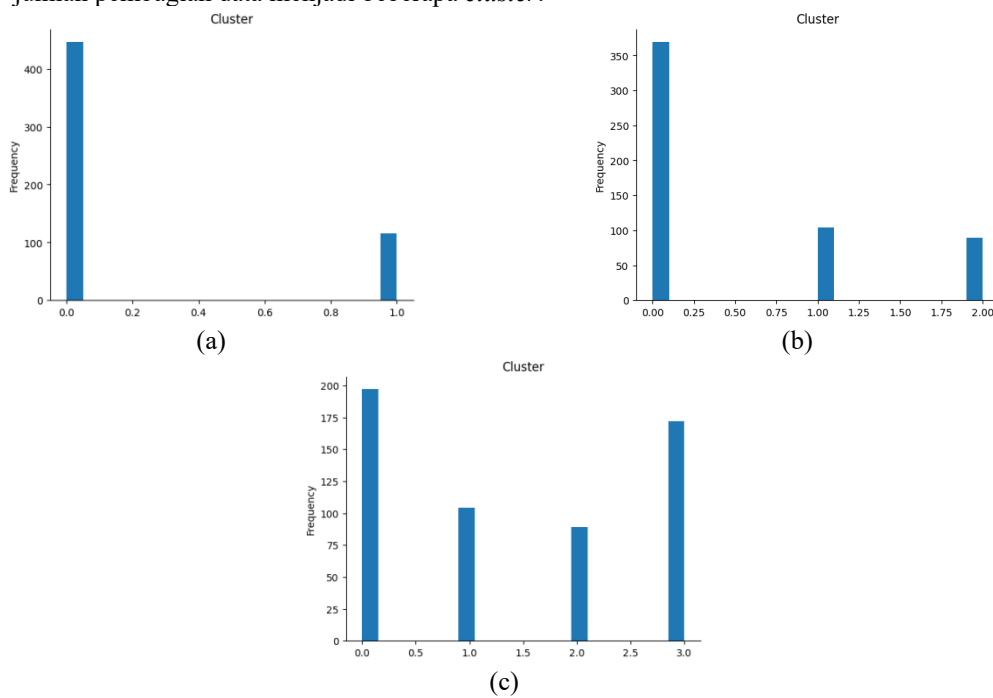
Gambar 9. Evaluasi *clustering* menggunakan *silhouette method* dengan $K = 3$ berdasarkan (a)8 variabel, (b)6 variabel, (c)4 variabel.

Sedangkan, pada *Elbow method* dihasilkan sebanyak 4 *cluster* dengan titik lengkungan yang paling tajam pada angka 4.0 sebagaimana ditunjukkan pada Gambar 10.



Gambar 10. Evaluasi *clustering* menggunakan *Elbow method* dengan $K = 3$ berdasarkan (a)8 variabel, (b)6 variabel, (c)4 variabel.

Selanjutnya, dilakukan pengujian data *clustering* dengan menggunakan nilai $K = 2$ dan $K = 4$. Lalu, didapatkan jumlah hasil pengklasteran yang berbeda tiap *clusternya*. Pada Gambar 11 ditampilkan jumlah pembagian data menjadi beberapa *cluster*.



Gambar 11. Visualisasi hasil *clustering* dengan semua variabel berdasarkan (a) $K = 2$, (b) $K=3$, (c) $K=4$.

Pada Tabel 2 berikut disajikan rincian hasil analisa *clustering* yang diperoleh.
Tabel 2. Hasil *clustering*.

K-Means	Cluster	Jumlah	Presentase
K = 2	Cluster 0	446	79,36%
	Cluster 1	116	20,64%
K = 3	Cluster 0	369	65,66%
	Cluster 1	104	18,50%
	Cluster 2	89	15,84%
K = 4	Cluster 0	197	35,05%
	Cluster 1	104	18,51%
	Cluster 2	89	15,84%
	Cluster 3	172	30,60%

D. Evaluasi *Clustering*

Dalam penelitian ini, digunakan dua metode evaluasi, yaitu *Silhouette Coefficient* dan *Elbow*. Namun, metode *Elbow* hanya membantu mengidentifikasi titik "siku" pada grafik. Maka dari itu, pada Tabel 3 disajikan hasil dari evaluasi metode *Silhouette Coefficient*.

Tabel 3. Hasil metode evaluasi

Jumlah Variabel	Cluster	Nilai Hasil
8 variabel	K = 2	0.5756520598428845
	K = 3	0.493651458125687
	K = 4	0.38078207855871304
6 variabel	K = 2	0.5874926216502723
	K = 3	0.5115597353599588
	K = 4	0.39745458980198545
4 variabel	K = 2	0.6232566842531486
	K = 3	0.5663392257970224
	K = 4	0.48614521111369613

Kesimpulan

Berdasarkan penelitian yang telah dilakukan dengan menggunakan metode *K-Means*, maka dapat disimpulkan bahwa pada pengujian *clustering* menggunakan $K = 3$ memberikan hasil grafik. Diperoleh bahwa dengan 8, 6 dan 4 variabel yang digunakan, jumlah kelompok disarankan *Silhouette method* sebanyak 2 *cluster* dan *Elbow method* sebanyak 4 *cluster*. Pengujian menggunakan metode evaluasi *Silhouette Coefficient* mendapatkan *clustering* 4 variabel dengan $K = 2$ mempunyai nilai paling baik yakni 0.6232566842531486.

Daftar Pustaka

- [1] A. N. S. Hayyun, "Pengaruh Narkoba Bagi Remaja dan Pelajar," *IIK Strada Indonesia*, pp. 1–4, 2021.
- [2] R. Tri Wahyuni, A. M. H Pardede, and T. Pasaribu, "Data Mining Pengelompokan Pecandu Narkoba Di Masyarakat Berdasarkan Faktor Penyebab Penggunaannya Menggunakan Metode K-Means Clustering," *Jurnal Informatika Kaputama (JIK)*, vol. 6, no. 3, pp. 835–844, 2022.
- [3] H. Mintawati and D. Budiman, "Bahaya Narkoba dan Strategi Penanggulangannya," vol. 1, no. 2, pp. 62–68, 2021, doi: <https://doi.org/10.52005/abdiputra.v1i2.95>.
- [4] H. Purbanto and B. Hidayat, "Systematic Literature Review: Penyalahgunaan Narkoba di Kalangan Remaja dalam Perspektif Psikologi dan Islam," *Al-Hikmah: Jurnal Agama dan Ilmu Pengetahuan*, vol. 20, no. 1, pp. 1–13, Feb. 2023, doi: [10.25299/al-hikmah:jaip.2023.vol20\(1\).11412](https://doi.org/10.25299/al-hikmah:jaip.2023.vol20(1).11412).
- [5] E. Muliati, Roswiyani, and N. Soetikno, "Pengaruh Dukungan Sosial Terhadap Self Efficacy Pada Pecandu Narkoba Di Masa Rehabilitasi," in *SENAPENMAS 2022*, SENAPENMAS 2022, Ed., Jakarta: E-Journal UNTAR, 2022, pp. 1005–1012.
- [6] Y. Mayona, R. Buaton, and M. Simanjutak, "Data Mining Clustering Tingkat Kejahatan Dengan Metode Algoritma K-Means," *Jurnal Informatika Kaputama (JIK)*, vol. 6, no. 3, pp. 345–352, 2022.
- [7] M. Mughnyanti, S. Efendi, and M. Zarlis, "Analysis of determining centroid clustering x-means algorithm with davies-bouldin index evaluation," *IOP Conf. Ser. Mater. Sci. Eng.*, no. 1, pp. 1–6, 2020, doi: [10.1088/1757-899X/725/1/012128](https://doi.org/10.1088/1757-899X/725/1/012128).

- [8] A. B. Pangestu, "Klasterisasi Tingkat Kecanduan Pemakai Napza Menggunakan Algoritma K-Medoids," UIN Sultan Syarif Kasim Riau, 2022.
- [9] B. Harahap, "Penerapan Algoritma K-Means Untuk Menentukan Bahan Bangunan Laris (Studi Kasus Pada UD. Toko Bangunan YD Indarung)," *Regional Development Industry & Health Science, Technology and Art of Life*, vol. 2, no. 1, pp. 394–403, 2019.
- [10] M. R. Muttaqin and M. Defriani, "Algoritma K-Means untuk Pengelompokan Topik Skripsi Mahasiswa," vol. 12, no. 2, pp. 121–129, 2020, doi: <https://doi.org/10.33096/ilkom.v12i2.542.121-129>.
- [11] K. P. Simanjuntak and U. Khaira, "Pengelompokan Titik Api di Provinsi Jambi dengan Algoritma Agglomerative Hierarchical Clustering," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 1, no. 1, pp. 7–16, 2021, doi: [10.57152/malcom.v1i1.6](https://doi.org/10.57152/malcom.v1i1.6).
- [12] N. Syahfitri, E. Budianita, A. Nazir, and I. Afrianty, "Pengelompokan Produk Berdasarkan Data Persediaan Barang Menggunakan Metode Elbow dan K-Medoid," *Kajian Ilmiah Informatika dan Komputer*, vol. 4, no. 3, pp. 1668–1675, 2023, doi: [10.30865/klik.v4i3.1525](https://doi.org/10.30865/klik.v4i3.1525).
- [13] B. Y. Geni, O. Kurnia, N. Hayati, M. Thoriq, and K. H. Manurung, "Analisa Algoritma K-Means Untuk Menentukan Strategi Marketing," *Jurnal Media Informatika Budidarma*, vol. 8, no. 1, pp. 211–220, 2024, doi: [10.30865/mib.v8i1.7085](https://doi.org/10.30865/mib.v8i1.7085).
- [14] P. R. Prayoga, Purnawansyah, T. Hasanuddin, and H. Darwis, "Klasifikasi Daun Herbal Menggunakan K-Nearest Neighbor dan Support Vector Machine dengan Fitur Fourier Descriptor," *Edumatic: Jurnal Pendidikan Informatika*, vol. 7, no. 1, pp. 160–168, 2023, doi: [10.29408/edumatic.v7i1.17521](https://doi.org/10.29408/edumatic.v7i1.17521).
- [15] T. Ahmad and M. N. Aziz, "Data preprocessing and feature selection for machine learning intrusion detection systems," *ICIC Express Letters*, vol. 13, no. 2, pp. 93–101, 2019, doi: [10.24507/icicel.13.02.93](https://doi.org/10.24507/icicel.13.02.93).
- [16] D. Shah, Z. Y. Xue, and T. M. Aamodt, "Label Encoding for Regression Networks," *ICLR 2022 - 10th International Conference on Learning Representations*, no. Section 2, pp. 1–35, 2022, doi: <https://doi.org/10.48550/arXiv.2212.01927>.
- [17] A. Yudhistira and R. Andika, "Pengelompokan Data Nilai Siswa Menggunakan Metode K-Means Clustering," vol. 1, no. 1, pp. 20–28, 2023, doi: <https://doi.org/10.58602/jaiti.v1i1.22>.
- [18] S. Paembonan and H. Abduh, "Penerapan Metode Silhouette Coeficient Untuk Evaluasi Clustering Obat," *PENA TEKNIK: Jurnal Ilmiah Ilmu-Ilmu Teknik*, vol. 6, no. 2, pp. 48–54, 2021, doi: [Prefix 10.51557](https://doi.org/10.51557).
- [19] N. T. Hartanti, "Jurnal Nasional Teknologi dan Sistem Informasi Metode Elbow dan K-Means Guna Mengukur Kesiapan Siswa SMK Dalam Ujian Nasional," *Jurnal Nasional Teknologi dan Sistem Informasi*, vol. 06, no. 02, pp. 82–89, 2020, doi: <https://doi.org/10.25077/TEKNOSI.v6i2.2020.82-89>.
- [20] I. Yudianto, W. Wahyudi, and A. Priyadi, "Inovasi Media Sosial dan Pengaruhnya terhadap Penerapan Fotografi Fashion," *SENIMAN: Jurnal Publikasi Desain Komunikasi Visual*, vol. 2, no. 1, pp. 273–281, 2024, doi: <https://doi.org/10.59581/seniman-widyakarya.v2i1.2578>.