



Clustering Wilayah Penyerapan Vaksinasi Covid-19 Di Kabupaten Karawang Menggunakan Algoritma K-Means Clustering

Rafif Mahfuzhin¹, Mohamad Jajuli², Rini Mayasari³

Fakultas Ilmu Komputer Universitas Singaperbangsa Karawang

Abstrak

Received: 04 Januari 2024
Revised: 10 Januari 2024
Accepted: 18 Januari 2024

Pandemi covid-19 telah menyebabkan dampak yang signifikan pada bidang kesehatan, ekonomi, dan masyarakat baik di Indonesia dan khususnya di Karawang. Pemerintah telah mengambil berbagai kebijakan dalam mencegah penyebaran covid-19 adalah program vaksinasi covid-19. Vaksin covid-19 seperti mRNA, CoronaVac, AZD1222 terbukti efektif dan aman dalam mencegah penyebaran varian covid-19. Namun, program vaksinasi covid-19 yang dilakukan di fasilitas kesehatan yang tersebar di semua kecamatan masih belum optimal dikarenakan vaksin yang seharusnya disalurkan tidak tersedia. Oleh karena itu, diperlukan suatu upaya untuk mengefisienkan program penyaluran vaksinasi di Kabupaten Karawang. Salah satunya upaya yang dapat dilakukan adalah clustering wilayah penyerapan vaksin covid-19. Algoritma yang dapat digunakan dalam clustering wilayah penyerapan vaksinasi covid-19 adalah K-Means Clustering. Pada penelitian ini data yang digunakan memiliki 10 variabel yaitu nama kecamatan, penerima dosis 1, penerima dosis 2, penerima dosis 3, penerima dosis 4, jumlah penduduk, kasus terkonfirmasi, kasus probabel, kasus suspect dan kasus kontak erat. Algoritma K-Means Clustering memberikan hasil 3 cluster dengan kategori rendah, sedang, dan tinggi dengan nilai evaluasi Silhouett Coefficient sebesar 0,77 yang termasuk strong structure.

Kata Kunci : Data Mining, K-Means Clustering, Vaksinasi, Covid-19

(*) Corresponding Author: 1910631170116@student.unsika.ac.id¹

How to Cite: Mahfuzhin, R., Jajuli, M., & Mayasari, R. (2024). Clustering Wilayah Penyerapan Vaksinasi Covid-19 Di Kabupaten Karawang Menggunakan Algoritma K-Means Clustering. <https://doi.org/10.5281/zenodo.10530803>.

PENDAHULUAN

Pandemi Covid-19 telah menyebabkan dampak yang signifikan pada bidang kesehatan, ekonomi, dan masyarakat di seluruh dunia, termasuk Indonesia. Namun, pada tanggal 30 Desember 2022, pemerintah Indonesia mengambil keputusan untuk mencabut Pemberlakuan Pembatasan Kegiatan Masyarakat (PPKM) yang merupakan salah satu langkah pencegahan penyebaran covid-19 (Instruksi Menteri Dalam Negeri Nomor 51, 2022). Keputusan diambil berdasarkan pertimbangan mengenai kasus Covid -19 yang makin terkendali. Namun, menurut berita yang diterbitkan oleh Liputan6.com pada tanggal 28 April 2023, Indonesia mengalami peningkatan kasus covid-19. Penambahan kasus pada tanggal 28 April 2023 bahkan menembus angka 2.067. Sehingga, kasus positif covid-19 meningkat sejumlah 6.771.072. Hal ini dapat menjadi masalah terutama di daerah – daerah pusat perekonomian maupun industri di Indonesia. Salah satunya adalah Kabupaten Karawang.

Salah satu strategi utama dalam menghadapi covid-19 yaitu pelaksanaan vaksinasi. Studi observasi yang dilakukan mendapatkan kesimpulan bahwa vaksin covid-19 seperti mRNA, CoronaVac dan AZD1222 sangat efektif dan aman untuk mencegah penyebaran semua varian (Alpha, Beta, Gamma, dan Delta) (Harahap,

2022). Di Indonesia, populasi yang telah divaksinasi dapat mencapai herd immunity melalui program vaksinasi (Tim Komunikasi Komite Penanganan Corona Virus 19, 2021).

Namun vaksinasi yang dilakukan masih belum mencapai target yang akan dicapai. Bahkan, menurut data Dinas Kesehatan Kabupaten Karawang yang dikutip pada tanggal 17 April 2022 melalui website republika.co.id, vaksinasi booster karawang hanya mencapai 18 persen dari sasaran 1.917.354 orang. Data terbaru diambil dari website covid19.go.id pada tanggal 1 Mei 2023 menunjukkan, Kabupaten Karawang hanya mencapai 1.83% jumlah vaksin booster dari total target provinsi Jawa Barat.

Berdasarkan permasalahan tersebut, diperlukan suatu upaya untuk mengefisienkan program vaksinasi di Kabupaten Karawang salah satunya adalah *clustering*. Salah satu algoritma clustering yang dapat digunakan adalah k-means clustering. K-Means Clustering adalah teknik pengelompokan data non-hierarkis yang mengorganisasikan data ke dalam satu atau beberapa kelompok. Data yang memiliki karakteristik serupa dikelompokkan bersama, sementara data yang memiliki karakteristik yang berbeda dikelompokkan dalam kelompok yang berbeda, sehingga menghasilkan kelompok yang memiliki variasi yang rendah di dalamnya (Supardi & Kanedi, 2020)

Pada penerapannya, algoritma k-means clustering dapat digunakan di berbagai permasalahan seperti pada perusahaan, pendidikan, dan pemerintahan. Terdapat beberapa penelitian sebelumnya yang pernah membahas mengenai penerapan algoritma K-Means Clustering mengelompokkan suatu masalah. Pada penelitian dengan judul “Perbandingan Algoritma DBSCAN dan K-Means Clustering untuk Pengelompokan Kasus Covid-19 di Dunia” yang diterbitkan tahun 2021, didapatkan hasil bahwa K-Means dengan nilai SI terbaik 0.6902 lebih unggul dibandingkan algoritma DBSCAN. Terdapat juga penelitian dengan judul “Implementasi Metode K-Means Untuk Pengelompokan Potensi Produksi Komoditas Perkebunan” yang diterbitkan tahun 2022, mendapatkan tingkat akurasi sebesar 93.75%.

Penelitian ini bertujuan untuk menerapkan algoritma k-means guna mengelompokkan wilayah penyerapan vaksinasi covid-19 di Kabupaten Karawang. Pengelompokan kecamatan di Kabupaten Karawang ke beberapa kluster tertentu sehingga dapat mengetahui daerah dengan jumlah kasus yang banyak dan yang sedikit. Hasil klusterisasi diharapkan dapat memberikan saran kepada pemerintah dalam menentukan kebijakan dalam penyerapan vaksinasi covid-19 di Kabupaten Karawang.

KAJIAN PUSTAKA

1. Data Mining

Data mining atau penambangan data merupakan suatu proses ekstraksi informasi penting dari kumpulan data yang besar. Teknik statistik, matematika dan kecerdasan buatan atau *artificial intelligence* digunakan dalam proses *data mining*. Proses dilakukan untuk mengidentifikasi pola dan informasi yang terdapat dalam kumpulan data yang besar (Husna, Syahputra, & Ginting, 2022).

2. Clustering

Clustering adalah suatu proses yang membagi himpunan objek data menjadi subset yang disebut *cluster*. Setiap objek dalam *cluster* memiliki karakteristik yang serupa dan berbeda dari *cluster* lainnya. Proses mempartisi data tidak dilakukan secara manual, melainkan menggunakan algoritma *clustering*. Oleh karena itu, *clustering* mampu mengidentifikasi grup atau kelompok yang tidak diketahui sebelumnya dalam data (Annisa, Ginting, & Syari, 2022)

3. Algoritma K-Means

K-Means Clustering adalah teknik pengelompokan data non-hierarkis yang mengorganisasikan data ke dalam satu atau beberapa kelompok. Data yang memiliki karakteristik serupa dikelompokkan bersama, sementara data yang memiliki karakteristik yang berbeda dikelompokkan dalam kelompok yang berbeda, sehingga menghasilkan kelompok yang memiliki variasi yang rendah di dalamnya (Supardi & Kanedi, 2020).

Menurut Alkhairi dan Winarto (2019), proses suatu algoritma K-Means Clustering dapat dilihat sebagai berikut :

- a. Menentukan jumlah kluster yang ingin ditetapkan untuk pusat kluster k
- b. Menggunakan jarak Euclidean, menghitung jarak antara setiap data ke pusat kluster.

$$d_{ik} = \sqrt{\sum_j^n (C_{ij} - C_{kj})^2} \quad (2.1)$$

Dengan : C_{ij} merupakan nilai atribut j dari titik data I, C_{kj} adalah nilai atribut j dari titik data I, dan n adalah banyaknya data.

- c. Mengkategorikan data ke dalam kluster dengan melihat jarak terpendek.
- d. Menghitung pusat cluster baru dengan menggunakan persamaan

$$C_{kj} = \frac{\sum_{i=1}^p x_{ij}}{p} \quad (2.2)$$

Dengan : $x_{ij} \in$ Kluster ke - k, p = banyak member kluster ke - k

- e. Langkah dua sampai empat diulangi hingga tidak ada data lagi yang berpindah ke kluster lain

4. Metode Elbow

Metode *Elbow* adalah teknik yang digunakan untuk menentukan jumlah kluster (c) optimal dalam algoritma *k-means* dengan memanfaatkan nilai *Sum of Square Error* (SSE) dari setiap kluster. SSE mengukur seberapa dekat setiap titik data dalam kluster terhadap pusat kluster (*centroid*). Pada metode *Elbow*, semakin rendah nilai SSE, semakin baik kluster tersebut dalam mengelompokkan data (Riani, Apriade, & Taufik, 2023).

5. Min-Max Normalization

Min-Max Normalization adalah salah satu teknik normalisasi yang digunakan dalam pengolahan data. Metode ini melakukan transformasi linier terhadap data asli sehingga menciptakan keseimbangan nilai perbandingan antar data sebelum dan setelah proses normalisasi. (Nasution, Khotimah, & Chamidah, 2019).

6. Silhouette Coefficient

Silhouette Coefficient adalah metode evaluasi yang digunakan untuk mengukur keoptimalan atau ketepatan kluster yang terbentuk dari proses *clustering*. Metode

ini memberikan informasi tentang kualitas visual objek dalam setiap kluster, serta memberikan indikasi sesuai dengan jumlah kluster pada dataset (Sinyor, 2021). Kriteria penilaian dari *Silhouette Coefficient* dapat dilihat pada Tabel 1.

Tabel 1. Kriteria Penilaian *Silhouette Coefficient*

Silhouette Coefficient	Kriteria Penilaian
$0.7 < \text{Silhouette} \leq 1.0$	Strong Structure
$0.5 < \text{Silhouette} \leq 0.7$	Medium Structure
$0.25 < \text{Silhouette} \leq 0.5$	Weak Structure
$\text{Silhouette} \leq 0.25$	No. Structure

7. CRISP-DM

Menurut Hasanah, Soim, dan Handayani (2021), CRISP-DM (*Cross Industry Standard Process for Data Mining*) adalah suatu proses pengembangan *data mining*. Proses ini melibatkan langkah-langkah yang terstruktur, jelas, dan efektif dalam mengolah data yang ada. Metode CRISP-DM memiliki 6 tahapan yaitu pemahaman bisnis (*Business Understanding*), pemahaman data (*Data Understanding*), pengolahan data (*Data Preparation*), pemodelan (*Modeling*), evaluasi (*Evaluation*), dan penyebaran (*Deployment*).

a. Pemahaman Bisnis (*Business Understanding*)

Business understanding merupakan tahap yang melibatkan identifikasi tujuan bisnis, pemahaman situasi dan kondisi saat penelitian, serta menetapkan tujuan penelitian untuk memecahkan masalah melalui *data mining*.

b. Pemahaman Data (*Data Understanding*)

Data understanding merupakan tahap yang melibatkan pengumpulan data dan memahami data lebih lanjut. Data yang terkumpul dieksplorasi sehingga didapatkan informasi – informasi mengenai data tersebut.

c. Pengolahan Data (*Data Preparation*)

Data preparation merupakan tahap yang dilakukan setelah pengumpulan data. Pada tahap ini, data akan melalui proses identifikasi, penyaringan data, pembersihan data, dan transformasi data.

d. Pemodelan (*Modelling*)

Modeling merupakan tahap implementasi algoritma yang digunakan untuk menemukan, mengidentifikasi, dan menghasilkan pola yang relevan dalam data penelitian.

e. Evaluasi (*Evaluation*)

Evaluation merupakan tahap yang dilakukan untuk mengukur hasil dari model yang telah dibangun pada tahap pemodelan sebelumnya. Hasil evaluasi menggambarkan proses *data mining* yang telah dilakukan dan mengukur model terbaik yang dapat digunakan.

f. Penyebaran (*Deployment*)

Deployment adalah proses pembuatan laporan atau artikel jurnal menggunakan hasil penelitian.

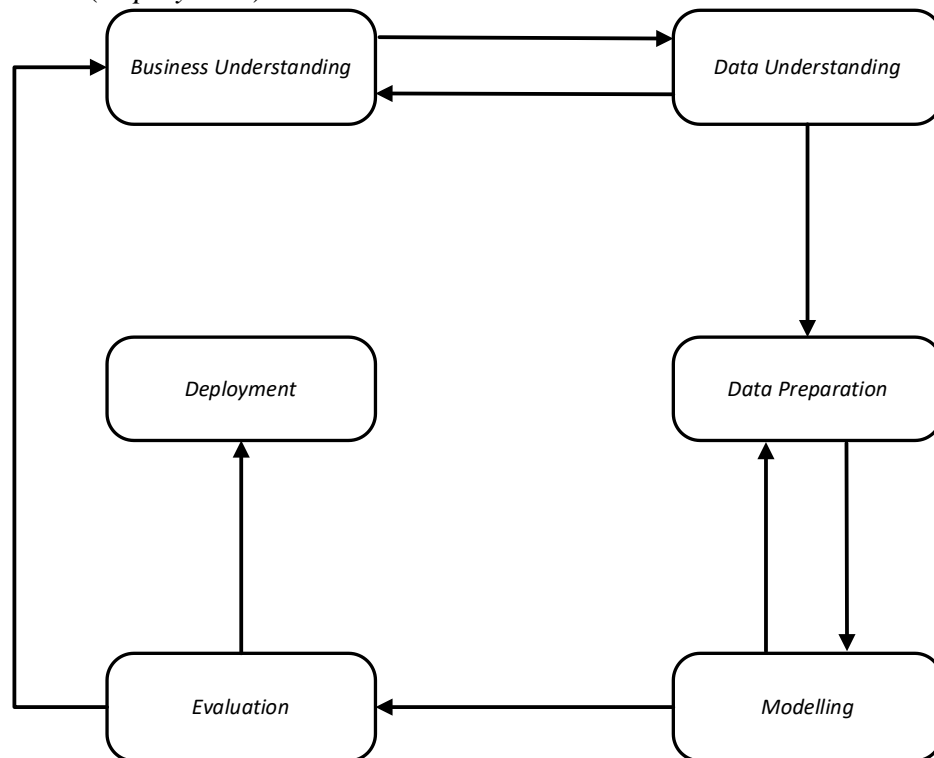
8. Python

Python adalah sebuah bahasa pemrograman tingkat tinggi yang mendukung pendekatan pemrograman berorientasi objek (OOP) dan mampu menjalankan berbagai instruksi multi-guna. Kelebihan Python sebagai bahasa pemrograman

tingkat tinggi adalah kemudahan dalam pembelajarannya berkat adanya manajemen memori otomatis yang terintegrasi (Rahmadhika & Thantawi, 2021).

METODE PENELITIAN

Penelitian ini menggunakan salah satu teknik data mining yaitu clustering menggunakan algoritma k-means clustering. Metodologi yang digunakan dalam penelitian ini yaitu metode CRISP – DM (*Cross – Industry Standard Process for Data Mining*). Metode CRISP-DM memiliki 6 tahapan yaitu pemahaman bisnis (*Business Understanding*), pemahaman data (*Data Understanding*), pengolahan data (*Data Preparation*), pemodelan (*Modeling*), evaluasi (*Evaluation*), dan penyebaran (*Deployment*).



Gambar 1. Metode CRISP-DM

1. Pemahaman Bisnis (*Business Understanding*)

Pada tahap ini dilakukan observasi ke tempat yang akan dijadikan objek penelitian yaitu Dinas Kesehatan Kabupaten Karawang serta mencari data yang akan diolah dalam penelitian. Berdasarkan informasi yang didapat, peneliti akan menilai situasi yang akan dijadikan objek penelitian tersebut. Pada proses ini akan menghasilkan suatu informasi yang dibutuhkan Tujuan dari penelitian ini adalah untuk mengetahui penyerapan vaksinasi Covid-19 di Kabupaten Karawang

2. Pemahaman Data (*Data Understanding*)

Pada tahapan ini dilakukan proses pengumpulan data yang kemudian data akan dipelajari lebih lanjut. Kemudian data dideskripsikan sehingga mendapatkan pengetahuan dari data tersebut. Setelah data dideskripsikan, data akan dieksplorasi sehingga data menentukan model analisis yang tepat.

3. Pengolahan Data (*Data Preparation*)

Pada tahapan ini melibatkan berbagai proses pengolahan dataset, seperti pemilihan data, pembersihan data, dan transformasi. Tahap ini harus dilakukan sebelum memasuki tahap pemodelan untuk menghindari keberadaan data tanpa atribut nilai atau data dengan nilai yang tidak konsisten, yang dapat mengakibatkan ketidakteroptimalan pemodelan

4. *Pemodelan (Modelling)*

Pada tahapan ini menggunakan tools Jupyter Anaconda dengan bahasa pemrograman Python. Dalam pemodelan ini diterapkan algoritma K- Means untuk mendapatkan suatu kluster. Penentuan jumlah kluster menggunakan metode elbow sehingga jumlah kluster yang terbentuk merupakan jumlah kluster optimal.

5. *Evaluasi (Evaluation)*

Pada tahapan ini dilakukan proses validasi serta pengukuran keakuratan hasil pada tahap pemodelan. Silhouette coefficient digunakan dalam proses ini untuk mengukur kualitas dari kluster yang terbentuk.

6. *Penyebaran (Deployment)*

Pada tahap ini merupakan tahap implementasi dari *data mining* berupa laporan yang dapat dijadikan informasi untuk pengambilan keputusan serta sumber penelitian yang akan datang.

HASIL DAN PEMBAHASAN

1. *Business Understanding*

Tahap pemahaman bisnis dalam proyek ini bertujuan untuk memahami kebutuhan bisnis dan merencanakan langkah-langkah awal dalam analisis data mining. Tujuannya adalah menggunakan algoritma K-Means untuk mengelompokkan wilayah penyerapan vaksinasi COVID-19 di Kabupaten Karawang dengan tujuan mengurangi penyebaran COVID-19. Permasalahan pada penelitian ini mengacu pada penyerapan vaksinasi covid-19 yang belum merata di Kabupaten Karawang. Untuk itu, diperlukan proses pengelompokan wilayah penyerapan vaksinasi covid-19 supaya dapat mengetahui daerah – daerah yang sudah terserap secara optimal atau belum di Kabupaten Karawang

2. *Data Understanding*

Pada tahap ini, data yang dikumpulkan mencakup data penerima vaksin COVID-19 dari tahun 2020 hingga Juli 2023, data kasus COVID-19 dari tahun 2020 hingga Juli 2023, serta data jumlah penduduk di Kabupaten Karawang tahun 2023. Data penerima vaksin terdiri dari 198 fasilitas kesehatan dengan atribut seperti kode fasilitas, nama fasilitas, kategori vaksin, jenis vaksin, dan jumlah penerima vaksin. Data kasus COVID-19 terdiri dari 30 kecamatan dengan atribut seperti jumlah kasus terkonfirmasi, kasus suspect, kasus probabel, dan kasus kontak erat. Data jumlah penduduk terdiri dari variabel kecamatan dan jumlah penduduk.

Analisis eksplorasi data menunjukkan bahwa terdapat banyak fasilitas kesehatan yang menyediakan vaksinasi COVID-19 di Kabupaten Karawang, dengan variasi penerimaan vaksin di berbagai kecamatan. Data jumlah penduduk menunjukkan variasi yang signifikan antara kecamatan. Kasus COVID-19 juga memiliki variasi di seluruh kecamatan, dengan tingkat terkonfirmasi, suspect, probabel, dan kontak erat yang berbeda-beda.

3. *Data Preparation*

Tahap selanjutnya dalam proyek ini adalah pengolahan data, yang melibatkan pemilihan atribut, pembersihan data, dan transformasi data untuk mempersiapkannya dengan baik sebelum dilakukan analisis lebih lanjut.

Pada tahap pemilihan data, tiga jenis data akan digunakan, yaitu data penerima vaksinasi COVID-19 berdasarkan kecamatan wilayah kerja fasilitas kesehatan, data kasus COVID-19 berdasarkan kecamatan di Kabupaten Karawang, dan data jumlah penduduk berdasarkan kecamatan di Kabupaten Karawang. Data penerima vaksinasi akan mengalami pengolahan awal dengan penambahan informasi kecamatan berdasarkan wilayah kerja fasilitas kesehatan yang ada. Data juga akan diperbarui berdasarkan kategori vaksin dan dosis yang sesuai. Validasi data akan dilakukan melalui proses wawancara dan verifikasi oleh Sub Koordinator Kelompok Sub Substansi Surveillans dan Imunisasi Dinas Kesehatan Kabupaten Karawang. Kemudian, data-data ini akan disatukan dengan atribut lain sehingga total ada 10 atribut pada dataset yang akan digunakan. Dataset yang digunakan dapat dilihat pada Tabel 2 berikut.

Tabel 2. Dataset

KECAMATAN	DOSIS 1	DOSIS 2	DOSIS 3	DOSIS 4	PEND UDUK	TER-KONFRI MASI	SUSP ECT	PROB ABEL	KONT AK ERAT
BANYUSARI	34706	28469	10593	66	57390	583	182	2	1496
BATUJAYA	50198	40579	13747	54	82230	337	49	3	340
CIAMPEL	30686	30014	13881	546	45140	1142	64	3	1658
CIBUAYA	30670	25862	8688	72	54430	568	332	1	295
CIKAMPEK	84044	74137	27085	1398	122220	4040	400	9	6086
...
TIRTAMULYA	32923	28531	8712	96	51290	391	65	1	400

Adapun, dalam dataset tersebut tidak terdapat nilai yang hilang (missing value), sehingga data sudah lengkap secara keseluruhan. Meskipun terdapat beberapa nilai yang bernilai 0, nilai ini memiliki arti bahwa kasus tidak ada dalam dataset tersebut.

Setelah itu, tahap transformasi data dilakukan dengan menggunakan teknik *min-max normalization*. *Min-max normalization* digunakan untuk menjadikan skala data pada rentang 0-1, sehingga semua atribut memiliki pengaruh yang setara dalam analisis. Teknik ini memudahkan dalam pengolahan data dan memastikan bahwa evaluasi data dapat dilakukan dengan nilai yang lebih tinggi.

```

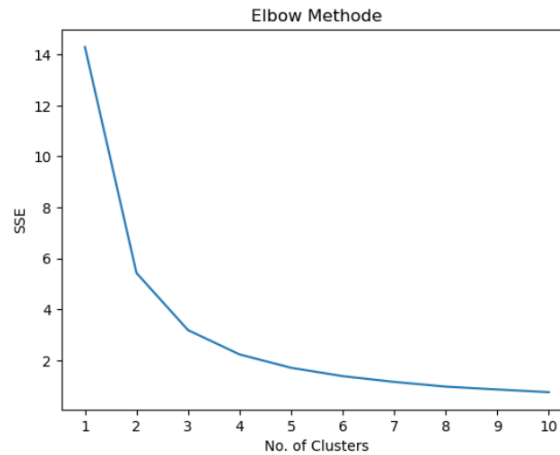
1 #Min-Max Preprocessing
2 kecamatan = df['kecamatan']
3 data = df.drop(columns=['kecamatan'])
4 from sklearn import preprocessing
5 scaler = preprocessing.MinMaxScaler()
6 data_normalized = scaler.fit_transform(data)
7 data_normalized_df = pd.DataFrame(data_normalized, columns=data.columns)
8
9 #Show Data
10 data_normalized_df
    
```

Gambar 2. Proses Transformasi Data

4. Modelling

Tahap selanjutnya dalam proyek ini adalah pemodelan, yang melibatkan penerapan teknik data mining yang telah ditentukan sebelumnya untuk mencapai tujuan yang telah ditentukan. Pada penelitian ini, teknik data mining yang digunakan adalah clustering, dan metode yang digunakan adalah algoritma K-

Means. Sebelum melakukan pemodelan, dilakukan proses menentukan jumlah cluster menggunakan metode elbow seperti pada Gambar 3 berikut.



Gambar 3. Metode Elbow

Berdasarkan hasil visualisasi metode elbow, jumlah cluster yang optimal adalah 3. Hal ini diperoleh karena pada nilai cluster 3 terbentuk sudut siku dan grafik kemudian melandai membentuk garis lurus, menunjukkan bahwa penambahan cluster tidak signifikan setelah titik tersebut.

Setelah menemukan jumlah cluster yang optimal, proses pemodelan dilanjutkan dengan membangun model K-Means clustering. Setiap data akan dibagi menjadi tiga cluster: cluster 0, cluster 1, dan cluster 2. Proses pemodelan k-means dapat dilihat di Gambar 4 berikut.

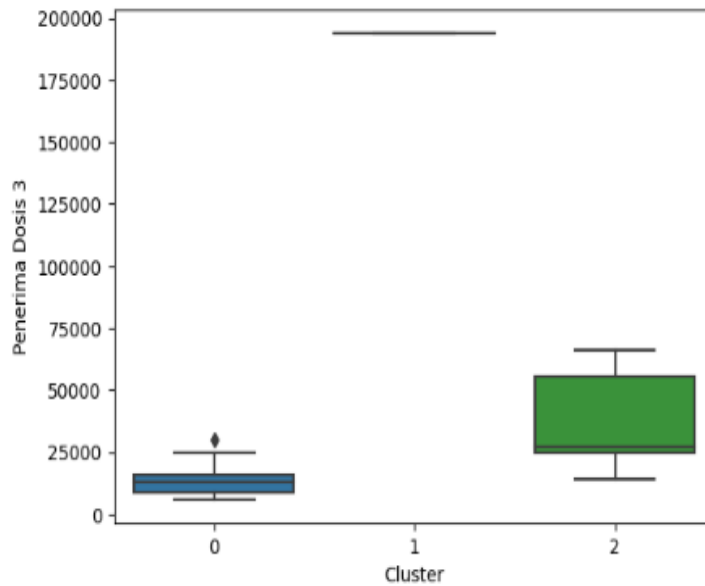
```

1 km = KMeans(n_clusters=3,init='k-means++',random_state=42)
2 km.fit(data_normalized_df)
3 centroids=km.cluster_centers_
4 centroids

array([[0.03494285, 0.04323642, 0.0389685 , 0.04509167, 0.15549336,
        0.10009116, 0.20340028, 0.15972222, 0.15479516],
       [1.          , 1.          , 1.          , 0.6786      , 0.76982827,
        0.88146718, 0.9756871 , 1.          , 1.          ],
       [0.1682314 , 0.18661613, 0.16809972, 0.34492   , 0.68381577,
        0.68236808, 0.53995772, 0.48333333, 0.8254631  ]])
    
```

Gambar 4. Pemodelan K-Means Clustering

Setelah pemodelan selesai, langkah berikutnya adalah memberikan label pada data. Label diberikan pada hasil cluster untuk memudahkan identifikasi. Pemberian label ini dilakukan berdasarkan pola persebaran data pada setiap atribut. Sebagai contoh, atribut jumlah penerima dosis 3 dapat digunakan untuk menentukan kategori hasil cluster. Kategori tersebut ditentukan dengan mengamati nilai-nilai dalam atribut dan memetakan ke dalam kategori rendah, sedang, dan tinggi. Grafik penerimaan dosis 3 berdasarkan clusternya dapat dilihat pada Gambar 5.



Gambar 5. Boxplot Penerima Dosis 3 Berdasarkan Cluster

Selanjutnya, atribut kategori dan cluster ditambahkan ke dalam data hasil pemodelan. Data hasil ini menggambarkan pengelompokan wilayah penyerapan vaksinasi COVID-19 di Kabupaten Karawang berdasarkan kategori cluster. Tabel hasil clustering dapat dilihat pada Tabel 3 berikut

Tabel 3. Hasil Clustering

Cluster	Anggota	Jumlah	Kategori
0	Banyusari, Batujaya, Ciampel, Cibuaya, Cilamaya Kulon, Cilamaya Wetan, Cilebar, Jatisari, Jayakarta, Kutawaluya, Lemah Abang, Majalaya, Pakisjaya, Pangkalan, Pedes, Purwasari, Rawamerta, Rengasdengklok, Tegalwaru, Telagasari, Telukjambe Barat, Tempuran, Tirtajaya, Tirtamulya	24	Rendah
1	Karawang Barat	1	Tinggi
2	Cikampek, Karawang Timur, Klari, Kota Baru, Telukjambe Timur	5	Sedang

5. Evaluation

Setelah melakukan tahap pemodelan, tahap selanjutnya adalah evaluasi. Evaluasi menggunakan teknik *silhouette coefficient* dengan proses pada Gambar 6.

```

1 from sklearn.metrics import silhouette_score
2 shil_avg = silhouette_score(data_normalized_df, km.labels_)
3 print('Nilai Silhouette Coefficient : '+str(shil_avg))

```

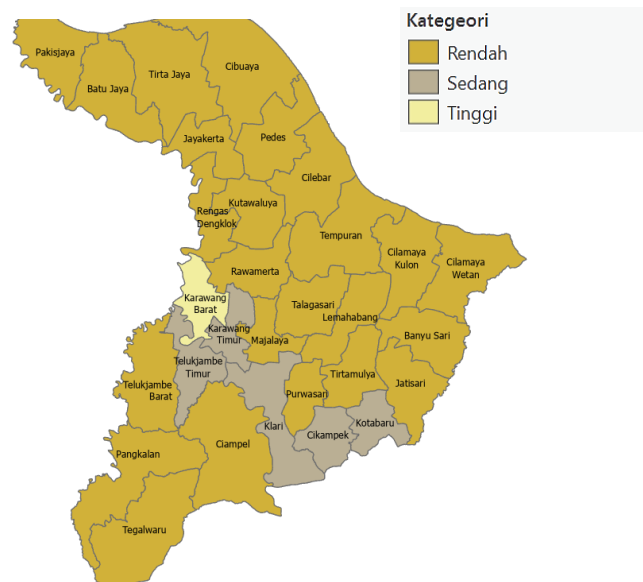
Nilai Silhouette Coefficient : 0.7770346638596172

Gambar 6. Silhouette Coefficient

Hasil evaluasi menggunakan metode silhouette coefficient diperoleh nilai 0,77. Nilai ini menunjukkan bahwa pemodelan memiliki struktur yang kuat dan anggota cluster telah ditempatkan dengan tepat dalam cluster yang sesuai.

6. Deployment

Setelah menyelesaikan semua tahapan dalam metodologi CRISP-DM untuk proses clustering wilayah penyerapan vaksinasi COVID-19 di Kabupaten Karawang menggunakan algoritma K-Means clustering, laporan akhir dapat disusun dan diserahkan. Laporan akhir ini akan berisi hasil visualisasi pemetaan menggunakan perangkat lunak Tableau. Visualisasi ini berupa pemetaan kecamatan di Kabupaten Karawang yang dapat dengan mudah dipahami oleh pihak Dinas Kesehatan Kabupaten Karawang, terutama terkait penyerapan vaksinasi COVID-19 di kecamatan yang masuk dalam kategori rendahan. Visualisasi pemetaan hasil clustering dapat dilihat pada Gambar 7.



Gambar 7. Visualisasi Pemetaan Hasil *Clustering*

KESIMPULAN

Algoritma K-Means Clustering dapat mengelompokkan wilayah penyerapan vaksinasi covid-19 di Kabupaten Karawang berdasarkan jumlah penerima vaksin, jumlah kasus covid-19, dan jumlah penduduk di Kabupaten Karawang. Algoritma K-Means Clustering memberikan hasil yaitu 3 cluster dengan kategori rendah sebanyak 24 kecamatan, kategori sedang sebanyak 5 kecamatan, dan kategori tinggi sebanyak 1 kecamatan.

Evaluasi clustering wilayah penyerapan vaksinasi covid-19 di Kabupaten Karawang menggunakan metode silhouette coefficient. Metode yang dilakukan mendapatkan nilai yaitu 0,77 yang termasuk dalam kategori strong structure sehingga data yang terhimpun dalam cluster sudah tepat

DAFTAR PUSTAKA

- Adha, R., Nurhaliza, N., Soleha, U., & Mustakim. (2021). Perbandingan Algoritma DBSCAN dan K - Means Clustering untuk Pengelompokan Kasus Covid-19 di Dunia. *Jurnal Sains, Teknologi, dan Industri*, Vol. 18 No. 2 pp. 206 - 211.

- Alkhairi, P., & Windarto, A. P. (2019). Penerapan K - Means Cluster Pada Daerah Potensi Pertanian Karet Produktif di Sumatera Utara. *Seminar Nasional Teknologi Kompter & Sains*, 762 - 767.
- Amelia, D. (2022). *Analisis Pemetaan Penyakit Diabetes Mellitus Di Kabupaten Karawang Menggunakan Algoritma K-Means*. Karawang: Universitas SIngaperbangsa Karawang.
- Annisa , K., Ginting, B. S., & Syari, M. A. (2022). Penerapan Data Mining Pengelompokan Data Pengguna Air Bersih BERdasarkan Keluhannya Menggunakan Metode Clustering Pada PDAM Langkat. *ALGORITMA : Jurnal Ilmu Komputer dan Informatika*, Vol. 06 No. 01.
- CNN Indonesia. (2023, Mei 4). *Kematian Pasien Covid Naik Lagi, Lansia Mendominasi*. Retrieved Mei 4, 2023, from <https://www.cnnindonesia.com/nasional/20230504165422-20-945340/kematian-pasien-covid-naik-lagi-lansia-mendominasi>
- covid19.go.id. (2023, Mei 1). *Status Vaksin*. Retrieved Mei 1, 2023, from <https://covid19.go.id/situasi>
- Fajrin, A. A., & Maulana, A. (2018). PENERAPAN DATA MINING UNTUK ANALISIS POLA PEMBELIAN ALGORITMA FP-GROWTH PADA DATA TRANSAKSI PENJUALAN SPARE PART MOTOR. *Kumpulan Jurnal Ilmu Komputer (KLIK)*, Vol. 05 No. 01.
- Fransiska, N. N., Anggraeni, D. S., & Enri, U. (2022). Pengelompokan Data Kemiskinan Provinsi Jawa Barat Menggunakan Algoritma K-Means dengan Silhouette Coefficient. *Jurnal Teknologi Informasi Komunikasi*, 29-35.
- goodnewsfromindonesia.id. (2020, November 25). *Karawang, Lumbung Padi Yang Jadi Kawasan Industri Terkemuka*. Retrieved April 2023, 29, from <https://www.goodnewsfromindonesia.id/2020/11/25/karawang-lumbung-padi-yang-jadi-kawasan-industri-terkemuka>
- Harahap, A. R. (2022). Efektivitas Vaksin Covid - 19 Dalam Kinerjanya Untuk Memodulasi Imunitas Tubuh Melawan Infeksi SARS-COV 2. *Indonesian Journal of Nursing and Health Sciences*, Vol 3 No. 1.
- Hasanah, M. A., Soim, S., & Handayani, A. S. (2021). Implementasi CRISP-DM Model Menggunakan Metode Decision Tree dengan Algoritma CART untuk Prediksi Curah Hujan Berpotensi Banjir. *Journal of Applied Informatics and Computing (JAIC)*, Vol. 5 No. 2.
- Husna, L., Syahputra, S., & Ginting, B. S. (2022). Penerapan Data Mining Menggunakan Metode K-Means Clustering Untuk Pengelompokan Data Perizinan Madrasah Diniyah Takmiliyah Awwaliyah (MDTA) Studi Kasus Kementerian Agama Stabat. *Jurnal Informatika Kaputama (JIK)*, Vol. 6 No. 3.
- Kementerian Kesehatan RI. (2020, Maret 30). *Info infeksi Emergen Kesehatan RI*. Retrieved Mei 1, 2023, from <https://infeksiemerging.kemkes.go.id/>.
- Kementerian Kesehatan RI. (2023, April 28). *Kemenkes Tambah Jenis Vaksin Booster Untuk Antisipasi Sub Varian Arcturus*. Retrieved from [kemkes.go.id: https://kemkes.go.id/article/view/23042900002/kemenkes-tambah-jenis-vaksin-booster-untuk-antisipasi-sub-varian-arcturus.html](https://kemkes.go.id/article/view/23042900002/kemenkes-tambah-jenis-vaksin-booster-untuk-antisipasi-sub-varian-arcturus.html)

- Kementrian Kesehatan RI. (2023, April 29). *Situasi COVID - 19*. Retrieved Mei 4, 2023, from covid19.go.id: <https://covid19.go.id/situasi>
- Kompas. (2022, Maret 26). *Selain Mal Technomart, Masyarakat Karawang Bisa Vaksinasi Booster di Tempat Ini*. Retrieved from kompas.com: <https://regional.kompas.com/read/2022/03/26/152134578/selain-mal-technomart-masyarakat-karawang-bisa-vaksinasi-booster-di-tempat>
- Kumarahadi, Y. K., Kumarahadi, B. M., & Sandradewi, K. (2022). Clustering Pelaksanaan Vaksinasi di Jawa Tengah Menggunakan Metode K - Means. *Jurnal Ilmiah Sinus (JIS)*, Vol. 20 No, 2.
- Liana, Winarno, & Darajatun, R. A. (2022). Analisis Cluster Vaksinasi Covid-19 Dengan Mempertimbangkan Jumlah Permintaan Menggunakan Metode K-Means Clustering di Kabupaten Karawang. *Jurnal Manajemen Teknologi dan Teknik Industri*, Vol. 4 No. 2 Agustus 2022, Hal 79 – 88.
- liputan6.com. (2023, April 28). *Update COVID - 19 Hari Ini 28 April 2023 : Naik Lagi, Kasus Positif Capai 2.067*. Retrieved April 29, 2023, from <https://www.liputan6.com/health/read/5272496/update-covid-19-hari-ini-28-april-2023-naik-lagi-kasus-positif-capai-2067>
- M. D. (2022). *Instruksi Menteri Dalam Negeri Nomor 51*. Jakarta: Menteri Dalam Negeri Republik Indonesia.
- Mahmud, Y., Shaeali, S., & Mutalib, S. (2021). Comparison of Machine Learning Algorithms for Sentiment Classification on Fake News Detection. *IJACSA (International Journal of Advanced Computer Science and Applications)*, 12(10), 658 - 665, www.ijacsa.thesai.org.
- Manalu, D. A., & Gunadi, G. (2022). Implementasi Metode Data Mining K-Means Clustering Terhadap Data Pembayaran Transaksi Menggunakan Bahasa Pemrograman Python Pada CV Digital Dimensi. *INFOTECH : Journal of Technology Information*, Vol. 8 No. 1.
- Marvaro, E., & Samosir, R. S. (2021). Penerapan Business Intelligence dan Visualisasi Informasi di CV. Mitra Makmur Dengan Menggunakan Dashboard Tableau. *Kalbiscentia : Jurnal Sains dan Teknologi*, Vol. 8 No. 2.
- Menteri Kesehatan RI. (2021, Januari 19). *Kemenkes Kenalkan Istilah Probable, Suspect, Kontak Erat, dan Terkonfirmasi COVID-19*. Retrieved Juli 30, 2023, from [sehatnegeriku.kemkes.go.id: https://sehatnegeriku.kemkes.go.id/baca/umum/20200714/2834469/kemenkes-kenalkan-istilah-probable-suspect-kontak-erat-dan-terkonfirmasi-covid-19/](https://sehatnegeriku.kemkes.go.id/baca/umum/20200714/2834469/kemenkes-kenalkan-istilah-probable-suspect-kontak-erat-dan-terkonfirmasi-covid-19/)
- Nasution, D. A., Khotimah, H. H., & Chamidah, N. (2019). Perbandingan Normalisasi Data Untuk Klasifikasi Wine Menggunakan Algoritma K-NN. *Journal of Computer Engineering System and Science (CESS)*, 78-82.
- Nita, D. (2023, Mei 3). *Kasus Covid - 19 Melonjak, Kemenkes Minta Warga Waspada dan Segera Vaksin Booster*. Retrieved Mei 4, 2023, from kompas.tv: <https://www.kompas.tv/article/403095/kasus-covid-19-melonjak-kemenkes-minta-warga-waspada-dan-segera-vaksin-booster>
- Pakpahan, H. S., Widians, J. A., Firmanda, H. D., & Basani, Y. (2022). Implementasi Metode K - Means Untuk Pengelompokan Potensi Produksi

- Komoditas Perkebunan. *Adopsi Teknologi dan Sistem Informasi (ATASI)*, Vol 1.
- Pradana, A. A., Casman, & Nur'aini. (2020). Pengaruh Kebijakan Social Distancing Pada Wabah Covid - 19 Terhadap Kelompok Rentan Di Indonesia. *Jurnal Kebijakan Kesehatan Indonesia : JKKI*, Vol. 09 No. 02.
- Presiden Republik Indonesia. (2021, Maret 2). Keputusan Presiden Republik Indonesia No. 14 Tahun 2021. Jakarta, Indonesia: Presiden Republik Indonesia.
- Purba, W., Siawin, W., Hardih, Nababan, M., Dharshinni, N. P., & Aisyah, S. (2018). IMPLEMENTASI DATA MINING UNTUK PENGELOMPOKKAN DAN PREDIKSI KARYAWAN YANG BERPOTENSI PHK DENGAN ALGORITMA K-MEANS CLUSTERING. *JUSIKOM PRIMA(Jurnal Sistem Informasi Ilmu Komputer Prima)*, Vol. 2 No. 1, Juli.
- Puspita, R. (2022, April 18). *Realisasi Vaksinasi Booster Di Karawang Hanya 18 Persen*. Retrieved Mei 1, 2023, from [republika.co.id: https://news.republika.co.id/berita//raihdv428/realisasi-vaksinasi-booster-di-karawang-hanya-18-persen?](https://news.republika.co.id/berita//raihdv428/realisasi-vaksinasi-booster-di-karawang-hanya-18-persen?)
- Rahmadhika, M. K., & Thantawi, A. M. (2021). Rancang Bangun Aplikasi Face Recognition Pada Pendekatan CRM Menggunakan Opencv Dan Algoritma Haarcascade. *Jurnal IKRA - ITH Informatika*, Vol. 5 No. 1.
- Ray Moynihan, c. a., Fox, M., Johansson, M., Lang, E., Duggan, A., Scott, I., & Albarqouni, L. (2021). Impact of COVID-19 pandemic on utilisation of healthcare services: a systematic review. *BMJ Journal*, 11(3).
- Riani, A. P., A. V., & T. R. (2023, Januari). Penerapan K-Means Clustering Dalam Pengelompokan Hasil Belajar Peserta Didik Dengan Metode Elbow. *Jurnal Teknologi Sistem Informasi dan Sistem Komputer TGD*, 164-172. Retrieved Juli 30, 2023, from <http://ojs.trigunadharma.ac.id/index.php/jsk/article/view/7351/1976>
- Sanubekti, M. A., & Dewi, L. P. (2022). Prinsip Klasifikasi Dan Data Mining Dengan Algoritma C4.5. *Jurnal Nuansa Informatika*, Vol. 15 No. 2.
- Sariasih, F. A. (2022). Implementasi Business Intelligence Dashboard dengan Tableau Public untuk Visualisasi Propinsi Rawan Banjir di Indonesia. *Jurnal Pendidikan Tambusai*, Vol 6, No. 2.
- Sinyor, I. J. (2021). *Pengelompokan Kasus Covid-19 Di Indonesia Menggunakan Metode Density-Based Spatial Clustering of Applications With Noise (DBSCAN)*. Semarang: Universitas Muhamadiyah Semarang.
- Sitorus, M., Antonieta.DC., C., & Larasati, C. (2022). Penerapan Algoritma K – Means Pada Clustering Vaksinasi Covid – 19 Daerah Jawa Timur. *Jurnal Teknosains Kodepena*, Vol. 03.
- Supardi, R., & Kanedi, I. (2020). IMPLEMENTASI METODE ALGORITMA K-MEANS CLUSTERING PADA TOKO EIDELWEIS. *JLTI : Jurnal Teknologi Informasi*, Vol. 4 No. 2.
- Susilo, A., Rurmende, C. M., Pitoyo, C. W., Santoso, W. D., Yulianti, M., Herikurniawan, . . . Yuniastuti, E. (2020). Coronavirus Disease 2019: Tinjauan Literatur Terkini. *Jurnal Penyakit Dalam Indonesia*, Vol. 7 No. 1.

- Tim Komunikasi Komite Penanganan Corona Virus 19. (2021, Januari 14). *Tujuan Program Vaksinasi Tercapainya Herd Immunity*. Retrieved from covid19.go.id: <https://covid19.go.id/p/vaksin/tujuan-program-vaksinasi-tercapainya-herd-immunity>
- Torence, A., Ramadhan, M., & Ginting, E. F. (2023). Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Dalam Pengelompokan Data Penerima Vaksinasi Covid-19. *Jurnal Sistem Informasi TGD*, Vol 2 No.3.
- Ulya, F. N. (2023, Januari 6). *PPKM Dicabut, Epidemiolog: Ada Kemungkinan Kasus Covid-19 Naik Lagi*. Retrieved Mei 1, 2023, from kompas.com: <https://nasional.kompas.com/read/2023/01/06/15482281/ppkm-dicabut-epidemiolog-ada-kemungkinan-kasus-covid-19-naik-lagi>
- WHO. (2020). *Pertanyaan jawaban terkait COVID-19 untuk publik*. Retrieved Mei 1, 2023, from <https://www.who.int/indonesia/news/novel-coronavirus/qa-for-public>