

Penerapan *K-Means Clustering* untuk Analisis Kondisi Lalu Lintas di Jalan Ir. H. Soekarno Surabaya

Nurul Istiqomah¹, Wika Dianita Utami^{2*}, Dian Yuliati³

^{1,2,3}Fakultas Sains dan Teknologi, UIN Sunan Ampel, Surabaya

e-mail: ¹nurulistiqomah2194@gmail.com, ²wikadianita@uinsa.ac.id, ³dian.yuliati@uinsby.ac.id

Received 23-09-2025; Reviewed 18-10-2025; Accepted 22-10-2025

Journal Homepage: <http://ktj.pktj.ac.id/index.php/ktj>

DOI: 10.46447/ktj.v12i2.725

Abstract

Growth in the number of vehicles, especially in urban areas, has a significant impact on traffic density, especially during peak hours, so an approach is needed to group traffic conditions based on the volume of all types of vehicles and the degree of saturation using the K-Means Clustering algorithm. The data used are the volume of all types of vehicles and the degree of saturation obtained from the Surabaya City Transportation Agency. The clustering results show that there are 4 clusters of different traffic characteristics, such as the volume of 2-wheeled vehicles during heavy traffic conditions of more than 4700 vehicles with a degree of saturation of more than 0.45. Evaluation using the silhouette coefficient produces a value of 0.63, which means the quality of the cluster is in a medium structure. This study shows that the clustering method is effective in understanding traffic conditions, although additional features can be done to optimize the quality of the cluster.

Keywords: Traffic, Clustering, K-Means, Silhouette Coefficient

Abstrak

Pertumbuhan jumlah kendaraan khususnya wilayah perkotaan sangat berdampak pada kepadatan lalu lintas, terutama pada jam-jam sibuk sehingga perlu adanya pendekatan untuk mengelompokkan kondisi lalu lintas berdasarkan volume semua jenis kendaraan dan derajat kejenuhan menggunakan algoritma *K-Means Clustering*. Data yang digunakan merupakan volume semua jenis kendaraan dan derajat kejenuhan yang didapat dari Dinas Perhubungan Kota Surabaya. Hasil pengelompokan menunjukkan terdapat 4 *cluster* karakteristik lalu lintas yang berbeda, seperti volume kendaraan roda 2 pada saat kondisi lalu lintas padat lebih dari 4700 kendaraan dengan derajat kejenuhan lebih dari 0.45. Evaluasi menggunakan *silhouette coefficient* menghasilkan nilai sebesar 0.63 yang berarti kualitas *cluster* dalam struktur sedang. Penelitian ini menunjukkan bahwa metode *clustering* efektif dalam memahami kondisi lalu lintas, meskipun penambahan fitur dapat dilakukan agar hasil kualitas *cluster* lebih optimal.

Kata kunci: Lalu Lintas, Clustering, K-Means, Silhouette Coefficient

PENDAHULUAN

Indonesia memiliki laju pertumbuhan penduduk yang terus meningkat setiap tahunnya, sehingga berdampak pada meningkatnya kepadatan lalu lintas, khususnya di wilayah perkotaan (Aditya et al., 2020). Kondisi ini disebabkan oleh bertambahnya kepemilikan kendaraan bermotor, baik milik pribadi maupun umum, serta kurangnya regulasi dari pemerintah terkait pembatasan kepemilikan kendaraan pribadi (Damayanti & Rachmatullah Putra, 2023). Salah satu kota besar yang turut menyumbang terhadap tingginya jumlah penduduk di Indonesia adalah Surabaya.

Surabaya turut berkontribusi terhadap kepadatan lalu lintas di Indonesia, dengan total semua jenis kendaraan mencapai 3.259.661 unit pada tahun 2020 dan terus meningkat hingga tahun 2024 mencapai 3.806.238 unit, dimana penyumbang terbesar berasal dari kendaraan sepeda motor yang mencapai 3.034.754 unit pada tahun 2024 (Statistik, 2021). Sebagai ibukota dan pusat layanan pemerintahan Jawa Timur, Surabaya dijadikan pilihan utama untuk pusat ekonomi dan pendidikan, dimana kondisi ini yang mendorong bertambahnya pendatang dari berbagai daerah sehingga ini juga berimbas pada tingginya mobilitas masyarakat (Farida et al., 2020). Hal ini memberikan dampak kemacetan lalu lintas pada titik tertentu, salah satunya di Jl. Ir. H. Soekarno.

Jl. Ir. H. Soekarno (Merr) Surabaya merupakan salah satu segmen jalan yang digunakan sebagai jalur penghubung antar kota dalam satu provinsi yang menghubungkan Kecamatan Sukolilo, Mulyorejo dan Kecamatan Rungkut dimana sepanjang koridor jalan ini, penggunaan lahannya didominasi oleh kegiatan perdagangan dan jasa serta pendidikan sehingga volume semua jenis kendaraan di jalan ini cukup tinggi hingga melebihi dari kapasitas jalan (Sirajaya et al., 2022; Tanggara et al., 2021). Jalan ini memiliki volume lalu lintas yang relatif padat pada jam-jam sibuk dan dikategorikan D pada tingkat pelayanan (*Level of Service*) dengan derajat kejenuhan mencapai 0,7674 (Agusdini et al., 2022; Tanggara et al., 2021). Derajat kejenuhan ialah perbandingan arus lalu lintas terhadap kapasitas (V/C rasio) yang dinyatakan dalam satuan smp/jam Menurut Bina Marga (1997) (Eriani Emiril, 2024). Kategori D dalam tingkat pelayanan dengan derajat kejenuhan 0,7674 berarti arus lalu lintas cenderung tidak stabil, namun kecepatan masih dapat dikontrol dengan nilai Q/C masih dalam batas wajar (Tanggara et al., 2021). Tentunya kondisi diatas perlu perhatian khusus dari pemerintah setempat.

Dinas Perhubungan Kota Surabaya dibawah naungan Pemerintah Kota telah melakukan beberapa upaya, seperti penambahan transportasi umum dan pemasangan CCTV di beberapa titik untuk mengurai kemacetan yang terjadi (Indrayana & Wahyudi, 2025). Namun, upaya ini masih belum efektif untuk mengurangi kepadatan secara sepenuhnya dikarenakan masih padatnya kondisi lalu lintas dan beberapa masyarakat masih melakukan pelanggaran lalu lintas (Kurnia Wahyu & Tukiman, 2022). Hal ini dibuktikan dengan volume lalu lintas harian di Jl. Ir. H. Soekarno sekitar 6640 ribu kendaraan (Tanggara et al., 2021). Sesuai dengan yang dilansir oleh Radar Surabaya pada 2 Desember 2023, antrian kendaraan bisa mencapai 1 km pada saat jam berangkat dan pulang kerja. Selain itu, saat terjadi insiden kecelakaan, arus lalu lintas

benar-benar terkunci dan kendaraan yang melintas hampir tidak dapat bergerak (Hurek, 2024; Sepka, 2023).

Berdasarkan kondisi diatas, perlu dilakukan analisa terkait pengelompokan kondisi lalu lintas berdasarkan volume kendaraan dan derajat kejenuhan di Jl. Ir. H. Soekarno Surabaya dengan menggunakan algoritma *K-Means Clustering*. Algoritma *K-Means Clustering* adalah satu diantara teknik data *mining* yang *simple*, mudah diterapkan, dan efektif pada jumlah data yang besar dengan waktu komputasi yang relatif cepat dan efisien (Hediyati & Suartana, 2021). Algoritma ini melakukan pengelompokan dengan cara membagi data kedalam K kelompok sesuai dengan fitur pada data tersebut (Putri Riani et al., 2023). Pada konteks pengelolaan lalu lintas, metode *clustering* memiliki peran penting karena dapat mengidentifikasi pola-pola tersembunyi pada data lalu lintas, seperti perbedaan tingkat kepadatan dan karakteristik arus kendaraan pada waktu tertentu (Rouky et al., 2024).

Beberapa penelitian terkait metode ini telah dilakukan, seperti penelitian yang dilakukan oleh Febby Arisca dkk terkait metode *K-Means Clustering* pada zonasi daerah yang menjadi korban bencana banjir di Kota Medan yang dipengaruhi oleh kondisi hilir sungai. Penelitian ini berhasil meng-*cluster* menjadi 2, yaitu daerah yang rawan banjir dan aman banjir dengan kualitas *cluster* sebesar 0.9 yang berarti kualitas *cluster* sangat kuat (Zurfani et al., 2024). Penelitian oleh Nanda dkk terkait perbandingan *K-Means* dan K-Medoids pada pengelompokan data miskin dan terbukti bahwa algoritma *K-Means* lebih baik dalam meng-*cluster* dengan nilai *Davies Bouldin Index* (DBI) sebesar 0.041 untuk 8 *cluster* (Luchia et al., 2022). Penelitian yang dilakukan oleh (Sohail et al., 2024) dengan menerapkan *K-Means Clustering* untuk mengelompokkan pola lalu lintas berdasarkan volume, kecepatan, dan jenis kendaraan. Hasil penelitian berhasil mengidentifikasi dua kondisi utama lalu lintas, yang menjadi dasar rekomendasi strategi dalam pengelolaan lalu lintas.

Berdasarkan penelitian sebelumnya, algoritma *K-Means Clustering* terbukti dapat mengelompokkan data dengan baik di beberapa kasus termasuk lalu lintas. Namun, penelitian yang dilakukan oleh (Sohail et al., 2024) belum mempertimbangkan terkait derajat kejenuhan sebagai indikator kinerja lalu lintas. Oleh karena itu, pengembangan penelitian dilakukan dengan menambahkan variabel derajat kejenuhan sehingga hasil pengelompokan kondisi lalu lintas lebih representatif. Penelitian ini bertujuan untuk mengelompokkan kondisi lalu lintas berdasarkan volume semua jenis kendaraan dan derajat kejenuhan menggunakan algoritma *K-Means Clustering*. Selain itu, penelitian ini juga untuk melihat seberapa besar hasil evaluasi dari jumlah *cluster* menggunakan *Silhouette Coefficient*. Hasil penelitian dapat dijadikan sebagai referensi untuk memahami kondisi lalu lintas yang lebih baik dan efisien.

METODE PENELITIAN

Penelitian ini menggunakan data sekunder yang diperoleh dari Dinas Perhubungan Kota Surabaya. Peneliti mengambil data terkait volume dari semua jenis kendaraan dan derajat kejenuhan di ruas Jl. Ir. H. Soekarno. Data berupa sampel volume dari semua jenis kendaraan (kendaraan bermotor dan tidak bermotor) dan derajat kejenuhan dari tahun 2022-2024 yang diambil tiap 15 menit dari jam 05.00-

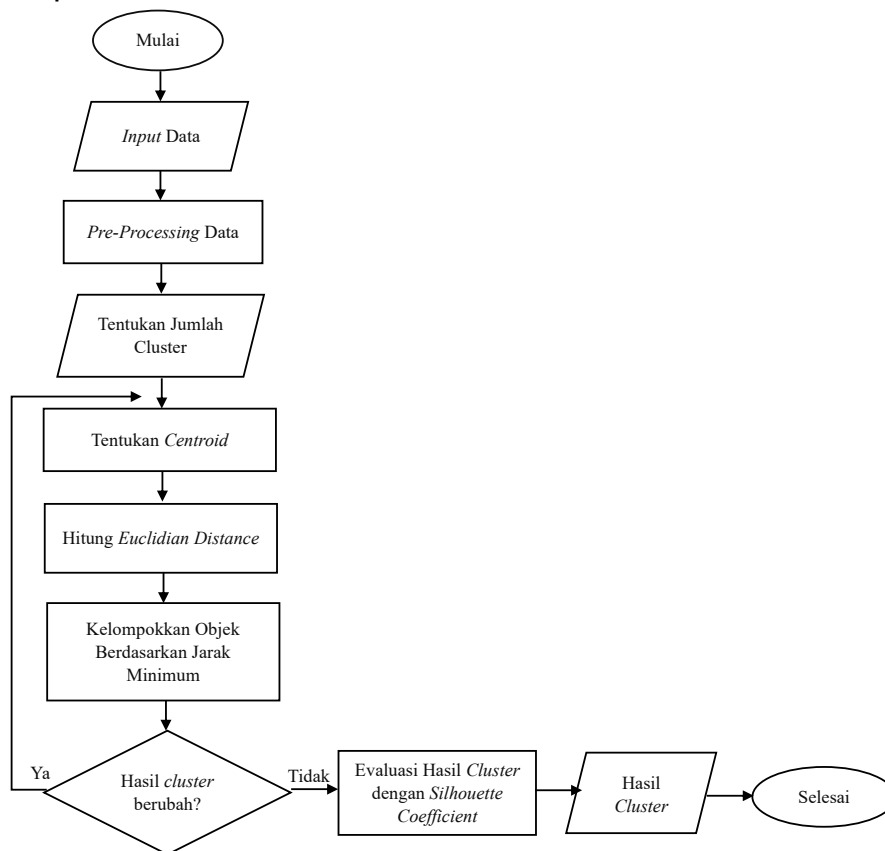
21.00 WIB. Total data yang akan diolah untuk mengelompokkan kondisi lalu lintas menggunakan volume kendaraan dan derajat kejenuhan sebanyak 503 data. Data dapat dilihat pada Tabel 1.

Tabel 1. Sampel Data Penelitian

No	Sepeda Motor	Mobil Pribadi	Angkot	Taksi	Bus	Pick Up /Box	Truck	Kendaraan Tidak Bermotor	Derajat Kejenuhan
1	1487	311	0	3	0	3	3	9	0.21
2	1526	269	5	3	0	10	6	8	0.211
3	1617	340	3	0	1	10	3	9	0.229
...
501	5994	563	17	38	0	32	13	18	0.771
502	4862	507	3	28	16	14	6	25	0.631
503	5622	614	27	5	14	38	6	39	0.736

Sumber: Dinas Perhubungan Kota Surabaya (2022-2024)

Pada Tabel 2, data terdiri dari 9 fitur dimana seluruh fitur ini akan di analisis secara kuantitatif menggunakan *K-Means Clustering* untuk mengetahui *cluster* dari data volume semua jenis kendaraan dan derajat kejenuhan. Adapun tahapan analisis data disajikan pada Gambar 1.



Gambar 1. Digram Alir (*Flowchart*) Penelitian

Berdasarkan Gambar 1, yaitu melakukan *input* data terkait volume semua jenis kendaraan dan derajat kejenuhan, kemudian dilakukan *pre-processing* data, dimana tahapan ini merupakan tahap yang sangat krusial dalam menunjang kualitas dari suatu data yang digunakan dalam penelitian. *Pre-processing* data dilakukan untuk menangani adanya data kosong, data *duplicate*, dan penanganan data *outlier* menggunakan *interquartile range*. Hasil dari *pre-processing* data akan dikelompokkan menggunakan *K-Means Clustering* yang dimulai dengan menentukan jumlah *cluster* K dengan mencoba nilai K dalam rentang 2 hingga 10, kemudian menentukan titik awal pusat *cluster* dilakukan dengan inisialisasi *centroid* secara acak. Langkah selanjutnya, menentukan setiap titik data ke *cluster* yang dibentuk berdasarkan *centroid* terdekat yang dihitung menggunakan *Euclidian Distance* pada persamaan (1) (Hakim et al., 2025).

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (1)$$

Keterangan:

$d(x, y)$ = Jarak *Euclidian* antara x dan y

n = Banyaknya data

x_i = titik data yang ingin di *cluster*

y_i = titik pusat dari *cluster*

Perhitungan nilai *centroid* terus dilakukan hingga nilai *cluster* tidak mengalami perubahan. Langkah terakhir, nilai *silhouette coefficient* dihitung menggunakan persamaan (2).

$$s_i = \frac{b(i) - a(i)}{\max \{a(i), b(i)\}} \quad (2)$$

Silhouette coefficient merupakan salah satu pengujian yang dilakukan agar mendapatkan kualitas hasil *cluster* yang baik. Nilai *Silhouette Coefficient* memiliki rentang antar -1 hingga 1 dimana sebuah hasil *cluster* dikatakan memiliki kualitas yang baik jika nilai *silhouette* mendekati 1 dan sebaliknya, nilai hasil *cluster* yang mendekati -1 berarti *cluster* memiliki kualitas yang buruk. Adapun kriteria *silhouette coefficient* dapat dilihat pada Tabel 2 (Aji et al., 2023; R et al., 2022).

Tabel 2. Kriteria Nilai *Silhouette Coefficient*

<i>Silhouette Coefficient</i>	Kriteria Penilaian
$0.70 < Silhouette \leq 1.00$	Struktur Kuat
$0.50 < Silhouette \leq 0.70$	Struktur Sedang
$0.25 < Silhouette \leq 0.50$	Struktur Lemah
$Silhouette \leq 0.25$	Tidak Ada Struktur

Sumber: (R et al., 2022)

HASIL DAN PEMBAHASAN

Pengelompokan kondisi lalu lintas berdasarkan volume kendaraan dan derajat kejenuhan menggunakan *K-Means Clustering* dimulai dengan tahap *pre-processing* data. *Pre-processing* ini dilakukan dengan menghapus data *duplicate*, menghapus data kosong (*missing value*), dan penanganan data *outlier*. Berdasarkan Tabel 3, hasil *pre-processing* data volume kendaraan dan derajat kejenuhan menunjukkan bahwa terdapat perbedaan jumlah data dibandingkan data awal dengan selisih 139 data. Hal ini disebabkan adanya proses penghapusan data yang terdeteksi *outlier*. Sementara itu, penanganan data *duplicate* dan nilai kosong (*missing value*) tidak dilakukan karena semua data tidak memenuhi kriteria tersebut sehingga data yang siap dianalisis menggunakan algoritma *K-Means Clustering* sebanyak 364 data.

Tabel 3. Hasil *Pre-Processing* Data

No	Sepeda Motor	Mobil Pribadi	Angkot	Taksi	Bus	Pick Up /Box	Truck	Kendaraan Tidak Bermotor	Derajat Kejenuhan
1	1487	311	0	3	0	3	3	9	0.21
2	1526	269	5	3	0	10	6	8	0.211
3	1617	340	3	0	1	10	3	9	0.229
...
362	677	253	10	0	9	26	3	12	0.114
363	725	265	7	17	0	6	12	9	0.12
364	672	281	3	14	16	1	3	14	0.116

Penerapan *K-Means Clustering* dimulai dengan penentuan jumlah *cluster* dengan rentang 2 hingga 10. Pada hasil dan pembahasan ini, akan dianalisa algoritma *K-Means Clustering* menggunakan 4 *cluster* untuk melihat pola distribusi data yang terbentuk pada jumlah *cluster* tersebut. Langkah pertama, menentukan pusat awal *cluster* (*centroid*) secara acak seperti pada Tabel 4, kemudian menentukan *cluster* tiap data menggunakan *Euclidian Distance*. Proses iterasi dilakukan hingga nilai *cluster* pada setiap data volume kendaraan dan derajat kejenuhan tidak mengalami perubahan.

Tabel 4. Nilai *Centroid* Iterasi Pertama

No.	<i>Centroid</i>	Sepeda Motor	Mobil Pribadi	Angkot	Taksi	Bus	Pick Up /Box	Truck	Kendaraan Tidak Bermotor	Derajat Kejenuhan
143	0	2945	402	3	14	0	27	3	12	0.394
291	1	860	468	3	22	7	42	0	11	0.163
46	2	1976	420	4	5	0	46	27	3	0.287
151	3	4184	502	3	12	6	27	3	6	0.548

Berdasarkan hasil nilai *centroid* (*C*) iterasi pertama pada Tabel 4, penentuan *centroid* 0 untuk *cluster* 0 diambil dari data ke 143, *centroid* 1 diperoleh dari data ke 291, kemudian *centroid* 2 didapat dari data ke 46, dan *centroid* 3 berdasarkan data ke 151. Langkah selanjutnya, pada Tabel 5 merupakan perhitungan jarak setiap data menggunakan *Euclidian Distance* (ED).

Tabel 5. Nilai *Centroid* Iterasi Pertama

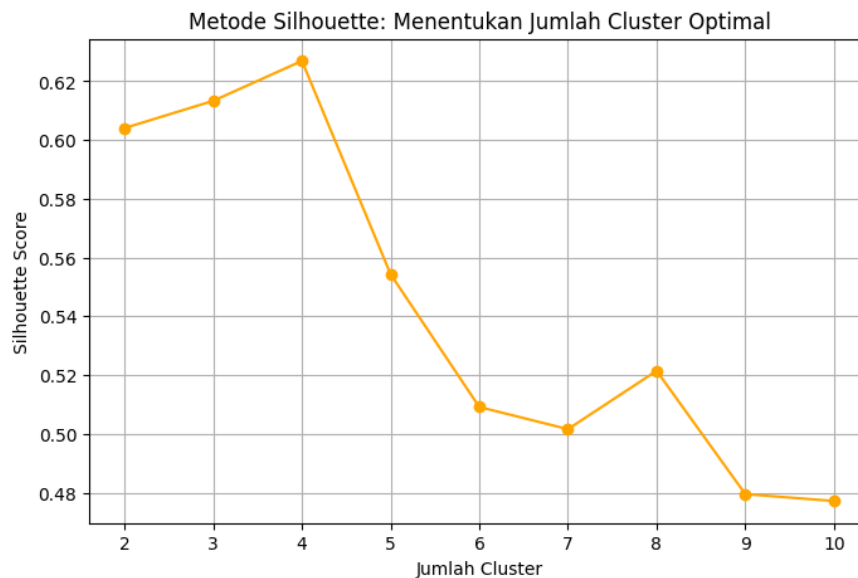
No.	ED 0	ED 1	ED 2	ED 3	Cluster
1	1461.082	647.867	503.471	2703.886	2
2	1425.373	696.161	476.517	2668.272	2
3	1329.633	768.759	370.427	2572.198	2
...
362	2272.961	283.755	1310.148	3515.863	1
363	2224.348	246.917	1261.364	3467.197	1
364	2276.424	268.621	1312.554	3519.067	1

Penentuan *cluster* setiap data dapat ditentukan dengan melihat jarak yang paling minimum diantara ke 4 jarak yang ada. Sebagai contoh, pada data pertama setelah dihitung memiliki jarak untuk setiap *cluster*, yaitu *cluster* 0 memiliki ED 0 sebesar 1461.082, *cluster* 1 memiliki ED 1 sebesar 647.867, *cluster* 2 memiliki ED 2 sebesar 503.471, dan *cluster* 3 memiliki ED 3 sebesar 2703.886, kemudian diperoleh nilai minimum pada setiap jarak yang ada pada *cluster* 2 yang berarti data pertama termasuk kedalam *cluster* tersebut. Proses penentuan *cluster* ini berlaku sama untuk seluruh data yang ada pada Tabel 6. Proses penentuan *centroid* dan perhitungan jarak antar data dilakukan secara iteratif hingga posisi *cluster* pada setiap data tidak mengalami perubahan dari iterasi sebelumnya. Tabel 6 merupakan hasil *clustering* pada iterasi terakhir untuk kondisi lalu lintas di Jl. Ir. H. Soekarno.

Tabel 6. Hasil *Clustering* Data

No	Sepeda Motor	Mobil Pribadi	...	Kendaraan Tidak Bermotor	Derajat Kejenuhan	Cluster
1	1487	311	...	9	0.21	2
2	1526	269	...	8	0.211	2
...
10	2495	660	...	5	0.371	0
11	2743	584	...	8	0.392	0
...
108	4146	541	...	14	0.547	3
109	5610	607	...	9	0.722	3
...
363	725	265	...	9	0.12	1
364	672	281	...	14	0.116	1

Tahapan selanjutnya melakukan evaluasi menggunakan *silhouette coefficient* untuk mengetahui struktur dari jumlah *cluster* yang telah ditentukan. Hasil evaluasi dapat dilihat pada Gambar 2.



Gambar 2. Hasil Nilai *Silhouette Coefficient*

Gambar 2 menunjukkan hasil evaluasi jumlah *cluster* menggunakan nilai *silhouette coefficient*, dimana *cluster* optimal berada pada *cluster* 4 sesuai dengan percobaan perhitungan *cluster* yang telah ditentukan sebelumnya dengan nilai *silhouette coefficient* sebesar 0.63. Sesuai dengan Tabel 1, nilai tersebut termasuk kedalam kategori dengan struktur sedang (*medium structure*). Nilai evaluasi sebesar 0.63 menunjukkan bahwa pemisahan antar *cluster* cukup baik, dimana sebagian data berada cukup dekat dengan hasil *cluster* sehingga metode ini masih dikatakan cukup layak untuk mengelompokkan kondisi lalu lintas di Jl. Ir. H. Soekarno.

Berdasarkan hasil pengelompokan kondisi lalu lintas di Jl. Ir. H. Soekarno Surabaya pada Tabel 7 dapat diketahui bahwa data terdiri dari 4 *cluster*. Masing-masing *cluster* memiliki karakteristik yang berbeda berdasarkan volume semua jenis kendaraan dan derajat kejenuhan. *Cluster* 1 dikategorikan sebagai kondisi lalu lintas yang tergolong sepi dengan volume sepeda motor berada pada rentang 300 hingga 1400 kendaraan dan derajat kejenuhan berada dikisaran 0.05 hingga 0.18 dan jumlah data sebanyak 134. Nilai derajat kejenuhan pada *cluster* ini menandakan arus lalu lintas masih jauh dibawah kapasitas jalan. Kondisi ini biasanya terjadi diluar jam kerja atau pada akhir pekan.

Pada *cluster* 2 dan *cluster* 0 memiliki hasil analisis *cluster* yang berbeda. *Cluster* 2 merupakan kondisi lalu lintas dikategorikan sebagai kondisi yang tergolong lancar dengan jumlah data 94. Volume sepeda motor pada *cluster* ini berada pada rentang 1400 hingga 2400 dengan derajat kejenuhan maksimal pada angka 0.35. Nilai derajat kejenuhan pada *cluster* ini menunjukkan arus kendaraan yang stabil. Sedangkan *Cluster* 0 merupakan hasil pengelompokan kondisi lalu lintas yang tergolong padat dengan jumlah data 85 dan volume sepeda motor berada pada rentang 2400 hingga 4700 dengan derajat kejenuhan maksimal pada 0.45. Nilai derajat kejenuhan pada

cluster ini menggambarkan bahwa kondisi lalu lintas mulai jenuh yang biasanya ditandai dengan menurunnya kecepatan rata-rata kendaraan. Kondisi ini sering terjadi pada jam berangkat dan pulang kerja.

Pada *cluster 3*, kondisi lalu lintas dikategorikan sebagai kondisi yang tergolong macet total (*stagnant*) dengan jumlah data 51 dan volume sepeda motor melebihi dari 4700 kendaraan dengan derajat kejenuhan berada pada kisaran 0.45 hingga 0.70. Nilai derajat kejenuhan pada *cluster* ini menunjukkan kondisi lalu lintas yang hampir atau sudah melampaui kapasitas jalan. Macetnya arus lalu lintas tidak hanya memperlambat waktu tempuh pengguna jalan, tetapi juga meningkatkan risiko kecelakaan akibat kepadatan dan potensi perilaku berkendara yang agresif. Volume kendaraan yang melebihi kapasitas ideal memperlihatkan bahwa permintaan ruang jalan jauh melampaui kemampuan ruas jalan yang tersedia, menyebabkan terjadinya antrean panjang dan waktu tunggu yang signifikan.

Hasil penelitian ini sesuai dengan yang dilakukan oleh (Sohail et al., 2024) yang menjelaskan bahwa kepadatan lalu lintas meningkat signifikan pada jam-jam kerja dan cenderung menurun pada akhir pekan. Namun, pada penelitian ini menambahkan variabel derajat kejenuhan (DS) sebagai indikator kinerja lalu lintas, sehingga hasil pengelompokan menjadi lebih representatif terhadap kondisi jalan yang sebenarnya.

SIMPULAN

Penerapan algoritma *K-Means Clustering* berhasil mengelompokkan kondisi lalu lintas di Jl. Ir. H. Soekarno menjadi 4 *cluster*. Hasil ini terdiri dari *cluster 1* sebagai kategori sepi, *cluster 2* kategori lancar, *cluster 0* kategori padat, dan *cluster 3* sebagai kategori macet total (*stagnant*). Kategori macet total (*stagnant*) pada *cluster 3* didominasi oleh kendaraan roda 2 dengan jumlah kendaraan lebih dari 4700 dan derajat kejenuhan berada pada kisaran 0.45 hingga 0.70. Kondisi ini cukup memperihatinkan karena ruas jalan telah mendekati atau bahkan melampaui kapasitas jalan. Nilai 0.63 merupakan hasil evaluasi menggunakan *silhouette coefficient* yang menunjukkan bahwa pembentukan *cluster* memiliki struktur yang sedang, dengan pemisahan antar *cluster* yang relatif jelas.

Hasil penelitian ini menunjukkan bahwa algoritma *K-Means Clustering* dapat diterapkan untuk memahami kondisi lalu lintas dengan cukup baik. Perlunya rekomendasi kebijakan dari pemerintah secara spesifik untuk mengatasi kondisi pada *cluster 3*, seperti penerapan sistem *adaptive traffic signal control*, rekayasa arus lalu lintas, pembukaan jalur alternatif, dan penertiban parkir liar. Penambahan fitur, seperti kondisi cuaca dapat dilakukan untuk membuat hasil *cluster* lebih optimal dan representatif terhadap kondisi lalu lintas yang sebenarnya.

DAFTAR PUSTAKA

- Aditya, F., Nasution, S. M., & Virgono, A. (2020). Traffic Flow Prediction using SUMO Application with K-Nearest Neighbor (KNN) Method. *International Journal of Integrated Engineering*, 12(7), 98–103.
<https://doi.org/10.30880/ijie.2020.12.07.011>
- Agusdini, T. M. C., Kusumaningrum, D., Sekartadji, R., & Agung, D. (2022). U-Turn Analysis on Road Performance by Dr. Ir. H. Soekarno Merr (Case Study: West and

- East Side U-Turn). *Journal of Civil Engineering, Planning and Design*, 1(2), 99–103. <https://doi.org/10.31284/j.jcepd.2022.v1i2.3607>
- Aji, B. G., Sondawa, D. C. A., Gifari, M. R., & Wijayanto, S. (2023). Penerapan Algoritma K-Means Untuk Clustering Harga Rumah Di Bandung. *Jurnal Ilmiah Informatika Global*, 14(2), 17–23. <https://doi.org/10.36982/jiig.v14i2.3189>
- Damayanti, D., & Rachmatullah Putra, L. (2023). Evaluasi Kebijakan Pemerintah Kota Malang tentang Rekayasa Lalu Lintas dalam Mengatasi Kemacetan (Studi Pada Dinas Perhubungan Kota Malang). *Jurnal Respon Publik*, 17(6), 17–27.
- Eriani Emiril, Y. (2024). Analisis Kinerja Ruas Jalan Akibat Parkir Di Bahu Jalan (Studi Kasus: Jalan Panjaitan Kota Probolinggo). *Jurnal Inter Tech*, 2(2), 101–107. <https://doi.org/10.54732/i.v2i2.1161>
- Farida, Y., Fanani, A., Purwanti, I., Wulandari, L., & Zaen, N. J. (2020). Pemodelan Arus Lalu Lintas dan Waktu Tunggu Total Optimal di Persimpangan Jl. Jemur Andayani Ahmad Yani sebagai Upaya Mengurai Kemacetan. *BAREKENG: Jurnal Ilmu Matematika Dan Terapan*, 14(3), 387–396. <https://doi.org/10.30598/barekengvol14iss3pp389-398>
- Hakim, B., Joanda Kaunang, F., Susanto, C., Salim, J., & Indradjaja, R. (2025). Implementasi Machine Learning dalam Pengelompokan Musik Menggunakan Algoritma K-Means Clustering. *Idealis: Indonesia Journal Information System*, 8(1), 74–83. <http://jom.fti.budiluhur.ac.id/index.php/IDEALIS/indexBhustomyHakim%7Chttp://jom.fti.budiluhur.ac.id/index.php/IDEALIS/index%7C>
- Hediyati, D., & Suartana, I. M. (2021). Penerapan Principal Component Analysis (PCA) Untuk Reduksi Dimensi Pada Proses Clustering Data Produksi Pertanian di Kabupaten Bojonegoro. *Journal Information Engineering and Educational Technology*, 05(02), 49–54.
- Hurek, L. (2024). *Macet Panjang Akibat Truk Muat Tabung Gas Terguling di Tengah Jalan MERR Surabaya*. Radar Surabaya. https://radarsurabaya.jawapos.com/surabaya/774414166/macet-panjang-akibat-truk-muat-tabung-gas-terguling-di-tengah-jalan-merr-surabaya?utm_source
- Indrayana, S. S., & Wahyudi, K. E. (2025). Strategi Dinas Perhubungan Kota Surabaya untuk Mengurangi Kepadatan Lalu Lintas Kota Surabaya. *Journal Of Social Science Research*, 5(1), 5553–5561.
- Kurnia Wahyu, A., & Tukiman. (2022). Efektivitas Program E-TLE Dalam Menangani Pelanggaran Lalu Lintas di Kota Surabaya. *Jurnal Sosial Ekonomi Dan Humaniora*, 8(3), 339–346. <https://doi.org/10.29303/jseh.v8i3.126>
- Luchia, N. T., Handayani, H., Hamdi, F. S., Erlangga, D., & Octavia, S. F. (2022). Perbandingan K-Means dan K-Medoids Pada Pengelompokan Data Miskin di Indonesia. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 2(2), 35–41. <https://doi.org/10.57152/malcom.v2i2.422>
- Putri Riani, A., Voutama, A., & Ridwan, T. (2023). Penerapan K-Means Clustering dalam Pengelompokan Hasil Belajar Peserta Didik dengan Metode Elbow. *Jurnal Teknologi Sistem Informasi Dan Sistem Komputer TGD*, 6(1), 164–172. <https://ojs.trigunadharma.ac.id/index.php/jsk/index>
- R, N. N. F., Anggraeni, D. S., & Enri, U. (2022). Pengelompokan Data Kemiskinan Provinsi Jawa Barat Menggunakan Algoritma K-Means dengan Silhouette Coefficient. *Tematik: Jurnal Teknologi Informasi Komunikasi*, 9(1), 29–35. <https://doi.org/10.38204/tematik.v9i1.901>

- Rouky, N., Bousouf, A., Benmoussa, O., & Fri, M. (2024). A spatiotemporal analysis of traffic congestion patterns using *clustering* algorithms: A case study of Casablanca. *Decision Analytics Journal*, 10(100404), 1–14. <https://doi.org/10.1016/j.dajour.2024.100404>
- Sepka, H. (2023). *Urai Kemacetan di Simpang Empat Undika MERR Saat Jam Sibuk, Dishub Kota Surabaya Sudah Siapkan Solusi*. Radar Surabaya. https://radarsurabaya.jawapos.com/surabaya/773353616/urai-kemacetan-di-simpang-empat-undika-merr-saat-jam-sibuk-dishub-kota-surabaya-sudah-siapkan-solusi?utm_source
- Sirajaya, R. D., Donny, M. S., & Rahayu, Y. E. (2022). Evaluasi Kinerja Simpang Bersinyal Jl. Dr. Ir. H. Soekarno - Jl. Mulyorejo Surabaya. *Jurnal Teknik Sipil*, 3(1), 40–48. <https://doi.org/10.31284/j.jts.2022.v3i1.3031>
- Sohail, A. M., Khattak, K. S., & Khan, Z. H. (2024). Data-driven insights: Unravelling traffic dynamics with k-means clustering and vehicle type differentiation. *Information System and Smart City*, 4(1), 1–20. <https://doi.org/10.59400/issc1737>
- Statistik, B. P. (2021). *Jumlah Kendaraan Bermotor yang Didaftarkan Menurut Kabupaten/Kota dan Jenis Kendaraan di Provinsi Jawa Timur*. Badan Pusat Statistik Provinsi Jawa Timur. <https://jatim.bps.go.id/id/statistics-table/1/MjI1MyMx/jumlah-kendaraan-bermotor-yang-didaftarkan-menurut-kabupaten-kota-dan-jenis-kendaraan-di-provinsi-jawa-timur-unit-2018-2020.html>
- Tanggara, M. A. P., Agustin, I. W., & Hariyani, S. (2021). Kinerja Jalan Di Kota Surabaya Berdasarkan Tingkat Pelayanan Jalan. *Planning for Urban Region and Environment*, 10(3), 119–128. <https://purejournal.ub.ac.id/index.php/pure/article/view/194/149>
- Zurfani, F. A., Sawaluddin, Mardiningsih, & Syahputra, M. R. (2024). Analisis Metode Clustering K-Means pada Zonasi Daerah Terdampak Banjir di Kota Medan dengan Evaluasi Silhouette Coefficient. *Algoritma : Jurnal Matematika, Ilmu Pengetahuan Alam, Kebumihan Dan Angkasa*, 2(6), 170–181. <https://doi.org/10.62383/algoritma.v2i6.270>