

SEGMENTASI MUTU JALAN NASIONAL MENGGUNAKAN METODE *CLUSTERING K-MEDOIDS* DI KABUPATEN BANDUNG BARAT DAN KOTA BANDUNG

Muhamad Sidik Firdaus^{1*}, Yulison H Chrisnanto¹ dan Edvin Ramadhan¹

¹Jurusan Informatika, Fakultas Sains dan Informatika, Universitas Jenderal Achmad Yani
Jl. Terusan Jend. Sudirman, Cibeber, Kec. Cimahi Sel., Kota Cimahi, Jawa Barat 40531.

*Email: muhamadsfirdaus17@if.unjani.ac.id

Abstrak

Jalan nasional merupakan jalan arteri yang menghubungkan antar ibu kota/provinsi, di dalam suatu jalan memiliki sebuah kriteria kondisi yang berbeda - beda. Penelitian ini tentang pengelompokan data kondisi jalan dengan menggunakan K-Medoids. Dengan dikelompokkan sebuah kelas jalan maka dapat memaknai kondisi jalan tersebut berdasarkan mayoritas dari setiap kondisinya dengan kriteria object kondisi jalan nya yaitu dengan menggunakan metode Pavement Condition Index dan International Roughness Index. Pengelompokan kondisi jalan berdasarkan 3 cluster yang memiliki kriteria kondisi pada masing – masing cluster dan membandingkan diantara cluster nya berdasarkan mayoritas dan kondisinya. Pada hasil yang diketahui bahwa mayoritas kondisi baik terdapat pada masing – masing cluster PCI, sedangkan pada IRI mayoritas masing – masing cluster yang diketahui yaitu kondisi sedang dan juga pada cluster 2 diketahui kondisi rusak ringan IRI dan kondisi rusak PCI lebih banyak dari cluster lain. Pengujian yang dilakukan menggunakan silhouette coefficient dengan pengujian sebanyak 8 kali dan k = 3 menghasilkan nilai sebesar 0,39.

Kata kunci: Data Mining, IRI, Jalan Nasional, K-Medoids, PCI.

1. PENDAHULUAN

Transportasi merupakan sektor yang sangat berperan dalam mendukung pengembangan wilayah atau pertumbuhan ekonomi suatu daerah, bahkan berfungsi sebagai urat nadi pembangunan. Sektor tersebut diperlukan karena mobilitas penumpang dan barang dari suatu tempat ketempat lain membutuhkan sarana dan prasarana transportasi(Jatmika et al., 2018). Baiknya kondisi jalan tentunya akan mengurangi tingkat kemacetan yang disebabkan oleh kondisi jalan yang tidak rata atau berlubang. Hal ini tentunya berdampak positif terhadap lancarnya distribusi barang/jasa(Wulandari Atur Rejeki et al., n.d.).

Kondisi jalan yang rusak tentu dapat menghambat aktifitas berpergian kesuatu tempat sehingga perbaikan maupun pemeliharaan harus terus dilakukan maka dari itu dalam kenyamanan berkendara akan lebih nyaman oleh pengguna jalan serta jalan yang baik dapat membuat pengguna lebih yakin untuk berkendara dengan *safety*(Yulianti & Hasanah, 2018).

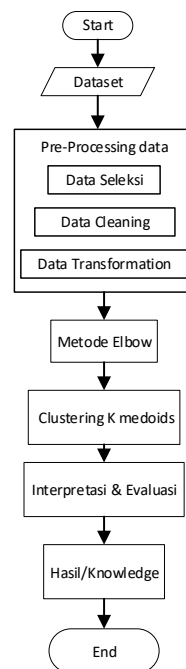
Dalam mengetahui kondisi jalan yang dianalisa di lapangan diketahui oleh pengukuran berdasarkan metode *Pavement Condition Index* dan *International Roughness Index*. Metode *PCI* merupakan metode survei visual dengan cara mengidentifikasi berbagai jenis kerusakan yang ada di lapangan dan Metode *IRI* merupakan parameter ketidakrataan suatu jalan dari naik turunnya permukaan arah profil memanjang (Tho'atin et al., 2016).

Agar tetap melaksanakan program perencanaan pemeliharaan jalan maka di lakukan metode pengelompokan dengan tujuan untuk mengetahui karakteristik antara metode *PCI dan IRI* dari sebuah *segment* jalan karena hal tersebut sama – sama tentang kondisi perkerasan jalan tetapi berbeda dalam pengukuran kondisi jalan. Hal ini dapat mengefisienkan dan efektif dalam manajemen pemeliharaan jalan karena *segment* jalan berkumpul pada masing – masing *cluster* nya sesuai dengan karakteristiknya dan kelas jalan Nasional pada pengeluaran biaya lebih banyak dikeluarkan berkali – kali lipat dari pada kelas jalan yang lain dikarenakan oleh intensitas volume lintas harian yang tinggi pada jalan nasional sehingga umur rencana pada jalan berkurang(Initiative, 2013). Penelitian ini akan berfokus pada pengelompokan kondisi jalan yang dapat mengetahui karakteristik dari metode pengukuran kondisi *IRI*

dan *PCI* pada beberapa *cluster* dengan dilihat dari nilai optimum pada k , sehingga dengan mengelompokkan dapat mengetahui berapa persentase atau mayoritas dari setiap nilai *cluster* nya dengan membandingkan hasil kondisi yang terlihat antara *PCI* dan *IRI* yang terdapat pada masing – masing *cluster* nya.

Pada penelitian tentang perbandingan antar *K-Means* dan *K-Medoids* dapat disimpulkan bahwa berdasarkan penelitian terdahulu yang telah melakukan perbandingan kedua metode *clustering partitioning* maka peneliti memutuskan untuk memakai *K-Medoids* untuk metode nya karena lebih baik dari pada dari segi eksekusi, validitas *cluster* dan tidak sensitif terhadap nilai dan juga Algoritma ini memiliki fitur yang sangat baik yang membutuhkan jarak antara setiap pasangan objek hanya sekali dan menggunakan jarak ini pada setiap setiap iterasi hingga konvergen (Arora et al., 2016; Farissa et al., 2021; Marlina et al., 2018). Kemudian pada paper lain yang membahas tentang pengelompokan kondisi permukaan jalan telah menghasilkan 4 *cluster* yaitu 4 kriteria kondisi sangat baik, sedang, rusak ringan, dan rusak berat dengan acuan dari keseluruhan ruas jalan nya dengan hasil pengujian sebesar 0,57 dalam *cluster* 4 serta memakai pengukuran dengan *Manhattan Distance* (Asmiatun, 2019) dan juga dari segi pengukuran penelitian ini menggunakan *Euclidean Distance* karena lebih baik dibandingkan dengan *Manhattan Distance* (Nahdliyah et al., 2019).

2. METODOLOGI



Gambar 1 Metodologi Penelitian

2.1. Pengumpulan Data

Data penelitian ini diperoleh dari dataset Balai Besar Pelaksanaan Jalan Nasional DKI Jakarta – Jawa barat dengan objek data kondisi jalan Nasional *IRI* dan *PCI* di Kota Bandung dan Kabupaten Bandung Barat pada tahun 2021 semester 1 yang memiliki 14 kategori jalan berdasarkan wilayah Satuan kerja PPK 4.1.

2.2. Pre Processing Data

Pada tahap ini dilakukan untuk mempersiapkan data serta memproses data mentah sebelum dilakukannya proses *data mining*. Hal ini dilakukan agar data dapat terintegrasi dengan system yang akan

digunakan pada saat proses data mining *clustering*. Pada tahap *Pre-Processing Data* ini terdapat 3 tahapan yaitu *data cleaning*, *data selection*, dan *data transformation*.

2.3. Algoritma K - Medoids

Metode *clustering K - Medoids* merupakan metode dari data mining *clustering* yang merupakan suatu proses pengelompokan record suatu, observasi, atau mengelompokkan kelas yang memiliki kesamaan objek. Perbedaan *clustering* dengan klasifikasi yaitu tidak adanya variabel target dalam melakukan suatu pengelompokan pada proses *clustering*. *Clustering* sering dilakukan sebagai langkah awal dalam proses data mining saat melakukan suatu metode analisis dan juga termasuk kedalam *unsupervised learning* yang mana menyimpulkan sebuah pola dari kumpulan data tanpa mengacu pada hasil yang diketahui atau diberi label (Pardede et al., 2016).

Algoritma *K-Medoids* adalah suatu metode partisi *clustering* atau *unsupervised learning* untuk mengelompokkan sekumpulan n objek menjadi sejumlah k *cluster* yang memiliki kedekatan atau kemiripan. Algoritma ini menggunakan objek pada kumpulan objek untuk mewakili *cluster*. *Cluster* dibangun dengan menghitung kedekatan yang dimiliki antara medoids dengan objek *non-medoids* (Irfan Rifai et al., 2015).

2.4. Metode Elbow

Tujuan dari metode elbow adalah untuk memilih nilai k yang kecil dan masih memiliki nilai *withinss* yang rendah. Penentuan optimal jumlah *cluster* diketahui dengan memperhatikan nilai perbandingan (dari perhitungan SSE untuk setiap nilai *cluster*) antara jumlah *cluster* yang akan membentuk siku pada suatu titik, sehingga semakin besar jumlah *cluster* k maka nilai SSE akan semakin kecil (Hartanti, 2020). Adapun rumus dari metode *Elbow* dibawah ini :

$$SSE = \sum_{k=1}^k \sum_{x_1 \in k} \|X_1 - C_k\|_2^2 \quad (2)$$

2.5. Silhouette Coefficient

Metode ini merupakan metode evaluasi *cluster* yang menggabungkan dua buah metode yaitu *cohesion* dan *separation*. *Cohesion* diukur dengan menghitung seluruh objek yang terdapat dalam sebuah *cluster* dan *separation* diukur dengan menghitung jarak rata-rata setiap objek dalam sebuah *cluster* dengan *cluster* terdekatnya (Farissa et al., 2021). Adapun rumus dari *Silhouette Coefficient* dibawah ini:

$$sil(c) = sil(k) \frac{1}{|k|} \sum_{i=1}^k sil \quad (3)$$

3. HASIL DAN PEMBAHASAN

3.1. Pengumpulan Data

Data yang diperoleh dan dikumpulkan berdasarkan hasil yang di dapat dari pihak Balai Besar Perencanaan Jalan Nasional Jawa Barat – DKI Jakarta, proses tahapan pengumpulan oleh pihak balai dengan menggunakan wawancara dan dataset. *Data* yang diperoleh sebanyak 587 *record* dengan total 14 buah kategori nama Jalan Nasional yang terdiri dari antara lain yaitu BTS. Kota Padalarang - BTS. Kota Bandung, Jln. Raya Cimahi (Cimahi), Jln. Raya Cibabat (Cimahi), Jln. Cibeureum (Bandung), Jln. Sudirman (Bandung), Jln. Rajawali Barat (Bandung), Jln. Nurtanio (Bandung), Jln. Abdul Rahman Saleh (Bandung), Jln. Pajajaran (Bandung), Jln. Pasir Kaliki (Bandung), Jln. Sukarno - Hatta (Bandung), Jln. Gede Bage (Bandung), Cimareme – Batujajar, Batujajar – Soreang.

3.2. Pra Processing Data

3.2.1. Data Cleaning

Pada tahapan *Data Cleaning*, *Data* akan diperbaiki untuk mencari *Data* yang yang tidak lengkap atau kosong.

Tabel 1. Data Sebelum di Cleaning

No	No Ruas	Sta From	PCI								IRI							
			L4	L3	L2	L1	R1	R2	R3	R4	L4	L3	L2	L1	R1	R2	R3	R4
1	22033	000+100				100	64							3.3	3.4			
2	22033	000+200				100	69							3.5	3.8			
3	22033	000+300				100	76							3.4	3.7			
....																		
587	22105	012+500			66	94	59	82						6.5	4.8	5.8	4.8	

Terlihat dari dataset awal yang masih terdapat data kosong atau tidak constant karna jika di tinjau di lapangan setiap lajur jalan berbeda – beda, maka dari itu dilakukan dengan mengubah 16 atribut L4, L3, L2, L1, R1, R2, R3, R4 yang merupakan lajur dari kondisi jalan tersebut menjadi 4 atribut sesuai dengan kelas *PCI* dan *IRI* dengan nama atribut yaitu Ruas Kiri *PCI*, Ruas Kanan *PCI*, Ruas Kiri *IRI*, Ruas Kanan *IRI* agar mempermudah penggunaan dan setiap 8 atribut tersebut di rata - rata kan dari setiap lajur nya antar L4-L1 (Ruas Kiri) dan R1-R4 (Ruas Kanan), kemudian menambahkan nama jalan dari setiap ruas jalan.

Tabel 2. Data Sesudah di Cleaning

No	Nama Jalan	No Ruas	StaFrom	PCI		IRI	
				Kiri	Kanan	Kiri	Kanan
1	BTS. KOTA PDL - BTS. KOTA BDG	22033	000+100	100,0	64,1	3,3	3,4
2	BTS. KOTA PDL - BTS. KOTA BDG	22033	000+200	100,0	69,4	3,5	3,8
3	BTS. KOTA PDL - BTS. KOTA BDG	22033	000+300	100,0	75,8	3,4	3,7
....
554	BTJ - SOR	22105	012+500	80,0	62,7	5,6	5,3

3.2.2. Data Selection

Data selection adalah proses dari penyeleksian *Data* yang akan di pakai untuk proses dari *clustering* serta tidak memakai atribut yang tidak menjadi tolak ukur. Dalam hal ini dilakukan penghapusan atribut No Ruas.

3.2.3. Data Transformation

Pada penelitian ini perubahan bentuk *Data* menggunakan min - max normalisasi. Normalisasi digunakan untuk mengubah atribut numerik menjadi skala yang lebih kecil yaitu dengan batas terendah 0 dan batas tertinggi. Adapun rumus pada min – max normalization ini :

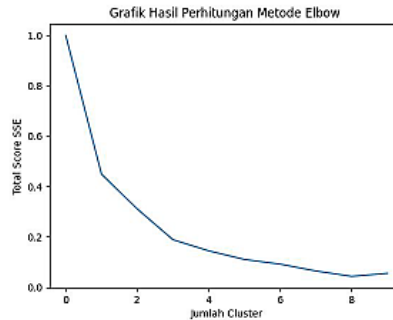
$$V'i = \frac{v_i - min_A}{max_A - min_A} (new_{maxA} - new_{minA}) + new_{minA} \tag{3}$$

Tabel 3. Data Transfomation

No	Nama Jalan	StaFrom	PCI		IRI	
			Kiri	Kanan	Kiri	Kanan
1	BTS. KOTA PDL - BTS. KOTA BDG	000+100	1,00	0,42	0,00	0,01
2	BTS. KOTA PDL - BTS. KOTA BDG	000+200	1,00	0,50	0,03	0,06
3	BTS. KOTA PDL - BTS. KOTA BDG	000+300	1,00	0,61	0,02	0,04
....
554	BTJ – SOR	012+500	0,59	0,39	0,28	0,24

3.3. Metode Elbow

Pada metode elbow terdapat penurunan yang signifikan pada jumlah selisih di setiap jumlah *cluster* nya dan Adapun hasil dari perhitungan metode *elbow* berupa grafik yang optimum pada nilai $k = 3$.



Gambar 2. Metode *Elbow*

3.4. Algoritma K-Medoids

Pada Tahapan ini merupakan hasil dari perhitungan *cluster* berdasarkan data kondisi jalan yang telah melalui proses perubahan data dari pra processing data.

Tabel 4. Hasil *Clustering*

NO	Nama Jalan	Sta From	Cluster
1	BTS. KOTA PDL - BTS. KOTA BDG	000+100	3
2	BTS. KOTA PDL - BTS. KOTA BDG	000+200	3
...
554	BTJ - SOR	012+400	2

3.5. Pengujian *Cluster*

Pada bagian ini dilakukan pengujian untuk mengukur seberapa besar keakuratan dari hasil *clustering* dengan menggunakan metode *sillhouette coefficient*. Perhitungan dilakukan dari hasil *clustering* nilai k yang dimasukan sebanyak 8. Dengan nilai $k = 3$ mendapat nilai sebanyak 0,396.

Tabel 5. Hasil Pengujian *Sillhouette*

Jumlah Cluster	Score Silhouette
2	0,62
3	0.396
4	0.444
5	0.588
6	0.446
7	0,61
8	0.582

3.6. Interpretasi

Karakteristik dari tiap *cluster* memiliki kuantitas dari setiap klasifikasi kondisi jalan nya. Data yang berkumpul pada masing – masing *cluster* dapat diketahui deskripsinya atau makna yang di dihasilkan berupa kuantitas skala mayoritas dari setiap klasifikasinya. Adapun hasil dari deskripsi yang telah disimpulkan dari *PCI* dan *IRI* pada tabel dibawah ini :

Cluster PCI memiliki 3 kondisi yaitu kondisi baik, sedang dan rusak.

1. Pada *cluster* 1 kondisi baik memiliki 87%, kondisi sedang 13% dan kondisi rusak 0,3%.
2. Pada *cluster* 2 kondisi baik memiliki 55%, kondisi sedang 35% dan kondisi rusak 7%
3. Pada *cluster* 3 kondisi baik memiliki 99%, kondisi sedang 1% dan kondisi rusak 0,4%

Cluster IRI memiliki 4 kondisi yaitu kondisi baik, sedang, rusak ringan, dan rusak berat.

1. Pada *cluster* 1 kondisi baik memiliki 35%, kondisi sedang 65%, kondisi rusak ringan 0,3% dan kondisi rusak berat 0%.
2. Pada *cluster* 2 kondisi baik memiliki 4%, kondisi sedang 88%, kondisi rusak ringan 8% dan kondisi rusak berat 0%.
3. Pada *cluster* 3 kondisi baik memiliki 26%, kondisi sedang 73%, kondisi rusak ringan 0,4% dan rusak berat 0%.

4. KESIMPULAN

Penelitian ini membuktikan bahwa *segment* jalan di bagi berdasarkan karakteristik deskripsi *cluster* nya, yang dapat mengetahui perbandingan *IRI* dan *PCI* dengan dilihat persentase dan kebanyakannya di dalam *cluster* tersebut sehingga ditetapkan pada masing – masing setiap *cluster*. Melihat dari kondisi *PCI* disetiap masing - masing *cluster* ditemukan bahwa mayoritas dan persentase kondisinya baik dan kondisi *IRI* disetiap masing - masing *cluster* mayoritas sedang diantara *cluster* lainnya, kemudian pada kondisi rusak *PCI* dan rusak ringan *IRI* ditemukan bahwa mayoritas terdapat pada *cluster* 2 diantara *cluster* lainnya dan kondisi rusak berat pada *IRI* tidak ditemukan pada setiap masing - masing dari *cluster*, artinya kondisi jalan umumnya telah berada dalam kondisi mantap. Pengujian pada penelitian kondisi jalan dengan menggunakan pengujian *Sillhouete Coefficiencie* dengan melakukan 8 kali pengujian yang mendekati nilai 1 yaitu *cluster* 2 dengan nilai sebesar = 0,62 dengan kriteria *Medium Structure* dan hasil *cluster* 3 sebesar 0.39 dengan kriteria *weak structure*, hal ini lebih baik untuk menggunakan *cluster* 2 dari kualitas *cluster* nya.

DAFTAR PUSTAKA

- Arora, P., Deepali, & Varshney, S. (2016). Analysis of K-Means and K-Medoids Algorithm for Big Data. *Physics Procedia*, 78(December 2015), 507–512. <https://doi.org/10.1016/j.procs.2016.02.095>
- Asmiatun, S. (2019). Penerapan Metode K-Medoids Untuk Pengelompokan Kondisi Jalan Di Kota Semarang. *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)*, 6(2), 171–180. <https://doi.org/10.35957/jatisi.v6i2.193>
- Farissa, R. A., Mayasari, R., & Umidah, Y. (2021). Perbandingan Algoritma K-Means dan K-Medoids Untuk Pengelompokan Data Obat dengan Silhouette Coefficient. In *Journal of Applied Informatics and Computing (JAIC)* (Vol. 5, Issue 2). <http://jurnal.polibatam.ac.id/index.php/JAIC>
- Hartanti, N. T. (2020). Metode Elbow dan K-Means Guna Mengukur Kesiapan Siswa SMK Dalam Ujian Nasional. *Jurnal Nasional Teknologi Dan Sistem Informasi*, 6(2), 82–89. <https://doi.org/10.25077/teknosi.v6i2.2020.82-89>
- Initiative, I. I. (2013). Jalan Daerah. *Jurnal Prakarsa Infrastruktur Indonesia*, 14(21), 1–47.
- Irfan Rifai Satuan Kerja Pelaksanaan Jalan Bebas Hambatan Tanjung Priok, A., Besar Pelaksanaan Jalan Nasional Jakarta, B. I., Hadiwardoyo, S. P., Gomes Correia Centre Algoritmi, A., Pereira Centre Algoritmi, P., & Cortez Centre Algoritmi, P. (2015). *IMPLEMENTASI DATA MINING UNTUK MENDUKUNG SISTEM MANAJEMEN PERKERASAN JALAN DI INDONESIA* (Vol. 1, Issue 2).
- Jatmika, B., Barkah, B., Studi Teknik Sipil, P., Nusa Putra, U., Raya Cibolong Kaler No, J., & Sukabumi, K. (2018). EVALUASI KINERJA JALAN DI JALAN RAYA BANDUNG-CIANJUR DI PT. POU YUEN INDONESIA. In *Jurnal Rekayasa Teknologi Nusa Putra* (Vol. 4, Issue 2).
- Marlina, D., Lina, N., Fernando, A., & Ramadhan, A. (2018). Implementasi Algoritma K-Medoids dan

- K-Means untuk Pengelompokan Wilayah Sebaran Cacat pada Anak. *Jurnal CoreIT: Jurnal Hasil Penelitian Ilmu Komputer Dan Teknologi Informasi*, 4(2), 64.
<https://doi.org/10.24014/coreit.v4i2.4498>
- Nahdliyah, M. A., Widiharih, T., & Prahutama, A. (2019). METODE k-MEDOIDS *CLUSTERING* DENGAN VALIDASI SILHOUETTE INDEX DAN C-INDEX (Studi Kasus Jumlah Kriminalitas Kabupaten/Kota di Jawa Tengah Tahun 2018). *Jurnal Gaussian*, 8(2), 161–170.
<https://doi.org/10.14710/j.gauss.v8i2.26640>
- Pardede, J., Gustiana, M., & Nurhasan, M. (2016). Implementasi K-Means *Clustering* Pada Aplikasi GIS (Geographic Information System) (Studi Kasus Pertanian Padi). *Seminar Nasional APTIKOM (SEMNASITIKOM)*, 274–279.
- Park, H. S., Lee, J. S., & Jun, C. H. (2006). A k-means-like algorithm for k-medoids *clustering* and its performance. *36th International Conference on Computers and Industrial Engineering, ICC and IE 2006*, 1222–1231.
- Tho'atin, U., Setyawan, A., Suprpto, M., Sipil, J., Pemeliharaan, M., Infrastruktur, R., & Tengah, J. (2016). *PENGGUNAAN METODE INTERNATIONAL ROUGHNESS INDEX (IRI), SURFACE DISTRESS INDEX (SDI) DAN PAVEMENT CONDITION INDEX (PCI) UNTUK PENILAIAN KONDISI JALAN DI KABUPATEN WONOGIRI*.
- Wulandari Atur Rejeki, P., Wulandari Atur Rejeki Pusat Kajian dan Pendidikan dan Pelatihan Aparatur Lembaga Administrasi Negara Jl KiaraPayung Km, P. I., & Barat, J. (n.d.). *EFFECTIVENESS OF RAPID RESPONSE UNIT OF ROAD PATCH IN BANDUNG*. www.bandungjuara.com,
- Yulianti, R., & Hasanah, B. (2018). Implementasi Program Pemeliharaan Jalan Di Dinas Pekerjaan Umum Dan Penataan Ruang (PUPR) Kabupaten Pandeglang.