
Penerapan Algoritma *k-Means Clustering* untuk Pengelompokan Pembangunan Jalan pada Dinas Pekerjaan Umum dan Penataan Ruang

Dede Kurniadi¹⁾, Yoga Handoko Agustin²⁾, Hari Ilham Nur Akbar³⁾, Ida Farida⁴⁾

^{1,2,3)}Jurusan Ilmu Komputer, Institut Teknologi Garut

⁴⁾Jurusan Teknik Sipil dan Perencanaan, Institut Teknologi Garut

Jl. Mayor Syamsu No.1, Jayaraga, Kec. Tarogong Kidul, Kabupaten Garut, Jawa Barat

Email: dede.kurniadi@itg.ac.id¹⁾, yoga.handoko@itg.ac.id²⁾, 1806129@itg.ac.id³⁾,
idafarida@itg.ac.id⁴⁾

Received: 08-08-2022	Riwayat artikel: Revised: 17-09-2022	Accepted: 18-09-2022
----------------------	---	----------------------

Abstract

*Determining road construction priorities at the Garut Regency Public Works and Spatial Planning (PUPR) Office could be more effective in terms of validity in achieving goals and recap time because it needs good clustering. This study uses the Cross-Industry Standard Process for Data Mining and the *k-means* algorithm to classify road construction data with type, length, and construction cost attributes. The results of this study are the selection of five clusters with the smallest Davies-Bouldin index evaluation value of 0.1617. The analysis of the characteristics of the largest cluster road construction is from Group 1, Group 2, Group 3, Group 5, and Group 4, with nominal development depending on the type of construction. From the clustering knowledge and cluster characteristics results, the Garut District PUPR office can prioritize the construction of small roads more than the construction of medium or large roads by looking at the submissions based on the cluster.*

Keywords: *k-means, algorithm, clustering, road construction, PUPR*

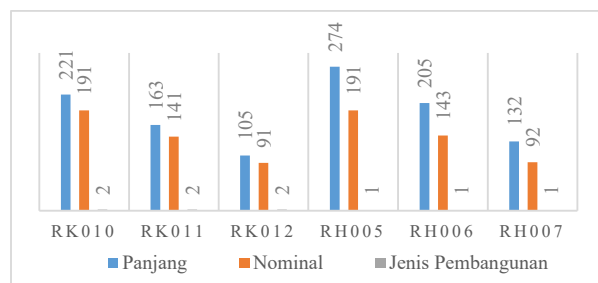
Abstrak

Proses penentuan prioritas pembangunan jalan di Dinas Pekerjaan Umum dan Penataan Ruang (PUPR) Kabupaten Garut kurang efektif, baik dalam segi keabsahan dalam mencapai tujuan maupun dalam segi waktu rekap. Hal ini dikarenakan tidak adanya pengelompokan yang baik. Penelitian ini menggunakan metode *Cross-Industry Standard Process for Data Mining* dan algoritma *k-means* dalam mengelompokkan data pembangunan jalan dengan atribut jenis, panjang, dan nominal pembangunan. Hasil penelitian ini adalah dipilihnya pengelompokan dengan lima klaster dengan nilai indeks *Davies-Bouldin* terkecil, yaitu sebesar 0,1617. Analisis karakteristik pembangunan jalan dari klaster terbesar yaitu dari Kelompok 1, Kelompok 2, Kelompok 3, Kelompok 5, dan Kelompok 4, dengan nominal pembangunan bergantung pada jenis pembangunannya. Dari hasil pengetahuan pengelompokan dan karakteristik kelompok, maka Dinas PUPR Kabupaten Garut dapat memprioritaskan pembangunan jalan kecil lebih banyak dibanding pembangunan jalan sedang ataupun besar dengan melihat pengajuan berdasarkan kelompoknya.

Kata kunci: *k-means, algoritma, klasterisasi, pembangunan jalan, PUPR*

Pendahuluan

Berdasarkan data dari Dinas Pekerjaan Umum dan Penataan Ruang (PUPR) Kabupaten Garut, setiap tahun desa mengajukan pembangunan jalan dengan panjang dan nominal yang berbeda-beda dalam musyawarah rencana pembangunan desa. Namun, penentuan prioritas pembangunan jalan masih dilakukan dengan pencocokan panjang dan nominal oleh pegawai PUPR, dengan melihat pengajuan panjang, nominal, dan jenis pembangunan. Pembangunan pendek dengan nominal yang tinggi dapat disebut menjadi pembangunan yang besar, dan pengajuan dengan ukuran yang panjang dengan nominal yang kecil disebut menjadi pembangunan yang sedang atau kecil tanpa perhitungan atau aturan yang pasti sebelumnya. Hal ini menyebabkan proses penentuan prioritas pembangunan kurang efektif baik dalam segi keabsahan, mencapai tujuan maupun dalam segi waktu rekap karena tidak memiliki pengelompokan yang baik. Berikut merupakan data pembangunan jalan yang diperoleh dari Dinas PUPR Kabupaten Garut tahun 2022 yang ditunjukkan pada Gambar 1.



Gambar 1 Data Pembangunan Jalan di Dinas PUPR

Desa di Kabupaten Garut dengan keterangan yang dibuat menjadi kode-kode pada Gambar 1 merupakan desa yang mengajukan pembangunan jalan pada Dinas PUPR Kabupaten Garut dengan panjang, nominal, dan jenis pembangunan yang berbeda. Sebagai contoh, RK010 yaitu jenis kegiatan rekonstruksi jalan dan 10 sebagai kode, dengan panjang 221 meter dan nominal biaya sekitar 191 juta. Berdasarkan permasalahan pada penentuan pengajuan jalan tersebut, dibutuhkan solusi yang dapat mengelompokkan pembangunan jalan sehingga mempermudah menentukan prioritas pembangunan jalan karena diketahui kelompok pembangunan jalan dalam pengajuan pembangunan jalan dari desa. Tujuan dari penelitian ini adalah melakukan pengelompokan pembangunan jalan pada Dinas PUPR menggunakan algoritma pengelompokan *k-means*.

Kajian Pustaka

Pada tahun 2017 penelitian oleh Puspita bertujuan untuk mengetahui rekomendasi DBMP dalam prioritas perbaikan jalan Kota Samarinda dengan mengelompokkan jalan menjadi empat kelompok jalan yang rusak dengan implementasi algoritma *fuzzy c-means* [1]. Penelitian selanjutnya oleh Asmiatun

melakukan pengelompokan jalan pada beberapa kecamatan, dengan tujuan agar diketahui kecamatan mana yang jalannya paling rawan rusak dengan algoritma *k-means* [2]. Penelitian selanjutnya oleh Mulyani melakukan pengelompokan rekomendasi bimbingan belajar berdasarkan segmentasi akademik siswa dengan hasil dua kelompok siswa, berdasarkan nilai rapor dari semester 1-5 untuk bahan pengayaan menggunakan algoritma *k-means* untuk pengelompokan dan FP-growth asosiasi [3].

Penelitian selanjutnya oleh Kurniadi melakukan analisis, prediksi, dan mengklasifikasikan mahasiswa berpotensi mendapatkan beasiswa di perguruan tinggi dengan algoritma *k-Nearest Neighbor* dengan hasil akurasi 95,83%. Penelitian ini memprediksi peluang terbesar untuk mendapatkan beasiswa Peningkatan Prestasi Akademik (PPA) yang dimiliki siswa [4]. Pada tahun 2019, penelitian oleh Listiani menerapkan *data mining* dalam rekomendasi pekerjaan dari data penduduk dengan mengelompokkan menjadi empat *klaster* yaitu C1 untuk buruh harian lepas, C2 untuk wirausaha, C3 untuk karyawan swasta, dan C4 untuk karyawan honorer menggunakan algoritma *k-means clustering* [5].

Penelitian selanjutnya pada tahun 2020 oleh Asmiatun dilakukan pengelompokan jalan dengan mengelompokkan data dari sensor *accelerometer*. Data yang digunakan berjumlah 638 data, dikelompokkan menjadi empat klaster yang menghasilkan kelompok jalan dengan keadaan baik sampai dengan rusak berat dengan metode *k-medoids* [6]. Penelitian selanjutnya oleh Vernanda membahas penerapan *clustering* dengan algoritma *k-means* untuk pemetaan daerah jalan rawan kecelakaan lalu lintas menjadi tiga klaster. Penelitian tersebut menghasilkan kelompok jalan cukup rawan, rawan, dan sangat rawan [7]. Penelitian selanjutnya oleh Nabila menerapkan algoritma *k-means* untuk menghasilkan pengelompokan pada kasus Covid-19 di Provinsi Lampung dengan hasil dari kedua perhitungan per akhir bulan menghasilkan nilai yang mendekati 0 atau klaster yang dievaluasi merupakan klaster baik [8].

Banyak algoritma yang dapat digunakan seperti *k-Nearest Neighbor* dan FP-growth, namun untuk pengelompokan data lebih sering menggunakan algoritma *k-means* dan *k-medoids*. Berdasarkan latar belakang yang diuraikan, tujuan dari penelitian ini adalah melakukan pengelompokan pembangunan jalan pada Dinas PUPR. Algoritma *k-means* menjadi alternatif pemilihan dikarenakan berdasarkan penelitian yang dilakukan Asmiatun dan Wakhidah, algoritma *k-means* cocok digunakan untuk pengelompokan yang menghasilkan klaster untuk pengelompokan kondisi permukaan jalan [2]. Begitu juga menurut Rezan, dkk., penerapan algoritma *k-means* menghasilkan nilai evaluasi *Davies-Bouldin Index* yang lebih kecil dibandingkan *k-medoids* [9]. Berdasarkan hal tersebut, maka *k-means* dijadikan pilihan karena dapat diterapkan pada kasus pengelompokan kondisi permukaan jalan dan terbukti menghasilkan evaluasi pengelompokan yang baik.

Data mining digunakan untuk mencari pola dalam menghasilkan informasi dari suatu data. *Data mining* memiliki berbagai macam metode, salah satunya adalah *clustering*. *Clustering* termasuk metode berbasis jarak yang membagi data ke dalam sejumlah kluster yang bekerja pada atribut angka atau numerik [8]. Pengelompokan ialah proses mengelompokkan satu set objek fisik atau abstrak ke dalam kelas yang mempunyai kesamaan [10].

Algoritma *k-means* merupakan algoritma untuk metode *klustering* [11] struktur non hierarki yang mempartisi data yang tersedia menjadi beberapa kelompok [12]. *k-Means* adalah teknik pengelompokan berbasis jarak yang membagi data menjadi beberapa kluster. Langkah-langkah algoritma *k-means* adalah sebagai berikut [13] [10]:

1. Menentukan banyak k kluster.
2. Menentukan nilai random awal (*centroid*).
3. Mencari jarak terdekat setiap data ke *centroid* menggunakan persamaan jarak Euclidean dengan Persamaan 1 (1).

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

dimana x adalah data, μ adalah *centroid*, dan n adalah jumlah data.

4. Mengklasifikasikan data dengan jarak terkecil kepada kelompok terdekatnya.
5. Memperbarui *centroid* dari titik tengah kluster bersangkutan menggunakan Persamaan (2).

$$\mu_k = \frac{1}{n_k} \sum_{t=1}^{n_k} x_t \quad (2)$$

dimana n_k adalah jumlah data dalam kluster k .

6. Lakukan kembali langkah ketiga sampai kelima sampai perubahan *centroid* tidak melebihi nilai toleransi error.
7. Apabila langkah enam terpenuhi, *centroid* (μ) pada iterasi terakhir digunakan sebagai parameter untuk pengklasifikasian data.

Davies-Bouldin yaitu metode evaluasi internal yang menghitung ukuran evaluasi kelompok berdasarkan nilai kohesi dan separasi. Dalam pengelompokan, kohesi merupakan jumlah kedekatan data dengan *centroid* kluster yang diikutinya, sementara separasi berdasarkan pada jarak antara kluster *centroid*. Kluster dengan jarak intra-kluster yang kecil dan jarak antar-kluster besar akan memiliki eksponen *Davies Bouldin* yang rendah. *Clustering* yang menghasilkan himpunan kluster dengan indeks *Davies-Bouldin* terkecil dianggap sebagai algoritma terbaik berdasarkan kriteria ini [9]. Langkah-langkah dalam mencari indeks *Davies-Bouldin* dijelaskan sebagai berikut:

1. Menghitung *Sum of square within cluster* (SSW) untuk mendapatkan matriks kohesi dalam sebuah kluster ke- i , sesuai Persamaan (3).

$$SSW_i = \frac{1}{m_i} \sum_{j=i}^{m_i} d(x_j, c_i) \quad (3)$$

2. Menghitung *Sum of square between cluster* (SSB) untuk mengetahui separasi antar klaster, sesuai Persamaan (4).

$$SSB_{i,j} = d(c_i, c_j) \quad (4)$$

3. Menghitung rasio ($R_{i,j}$) untuk mendapatkan nilai perbandingan antara klaster satunya dan lainnya. Klaster baik memiliki nilai kohesi paling kecil dan separasi yang paling besar. Nilai rasio dihitung menggunakan Persamaan (5).

$$R_{i,j} = \frac{SSW_i + SSW_j}{SSB_{i,j}} \quad (5)$$

4. Nilai rasio tersebut dihasilkan untuk mencari nilai indeks *Davies-Bouldin* (DBI) sesuai Persamaan (6).

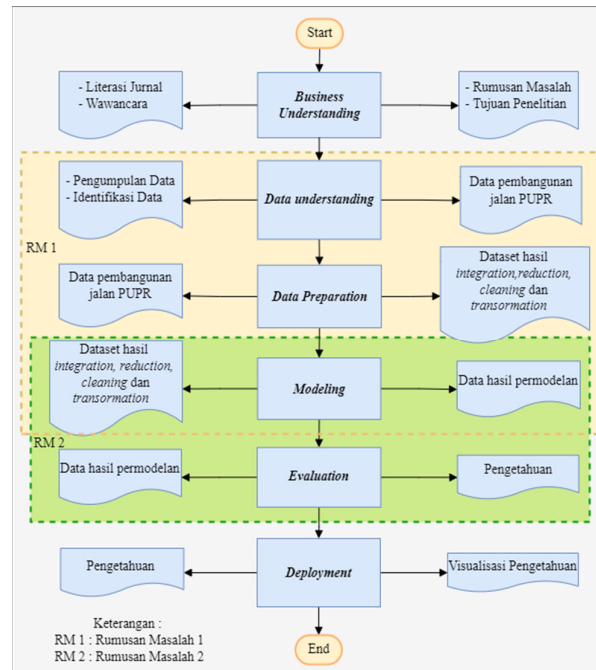
$$DBI = \frac{1}{k} \sum_{i=1}^k \max_{i \neq j} (R_{i,j}) \quad (6)$$

5. Jika nilai DBI yang diperoleh semakin kecil non-negatif atau lebih dari 0, maka klaster yang diperoleh dari pengelompokan *K-means* yang digunakan semakin baik.

Jalan merupakan sarana transportasi utama untuk mencapai suatu tujuan dari tempat satu ke tempat lain, sehingga kondisi jalan sangat berpengaruh bagi keamanan dan keselamatan [14]. Program pemerintah melalui Dinas Pekerjaan Umum dan Penataan Ruang mewujudkan perkembangan antar daerah dengan pembangunan jalan [15] dan diatur dalam PUPR No. 21/PRT/M/2019 tentang Manajemen Keselamatan Konstruksi. Seringkali rencana mutu pelaksanaan konstruksi yang direncanakan belum sepenuhnya sesuai dengan pekerjaan sehingga hasil proyeknya berbeda. Oleh karena itu diperlukan penyesuaian rencana dan standarisasi yang dikeluarkan oleh PUPR untuk mengupayakan standar sesuai aturan dari PUPR [15].

Metode Penelitian

Kerangka penelitian dalam pengelompokan pembangunan jalan ini menerapkan pendekatan *Cross-Industry Standard Process for Data Mining* (CRISP-DM) sebagaimana disajikan pada Gambar 2.



Gambar 2 Tahapan Penelitian

Kegiatan awal dimulai dengan *business understanding* dengan kegiatan memahami kebutuhan PUPR rencana dan strategi penyelesaiannya. *Data understanding* dilakukan melalui pengumpulan data dari Dinas PUPR dan dianalisis kualitas datanya. *Data preparation* dilakukan dengan membangun *dataset integration, reduction, cleaning, dan transformation*. Hal ini dilakukan dengan menggabungkan data apabila data yang didapatkan terpisah-pisah, menentukan atribut atau mengurangi data, menghapus *missing value* pada atribut yang memiliki data yang kosong atau *null*, agar data yang dihitung lebih akurat karena sudah tidak ada *missing value*, dan mentransformasi data berjenis *non nominal* menjadi bentuk angka/*numerical*. *Modeling* dilakukan dengan memilih teknik pemodelan pada dataset yang telah dibangun. Pada tahap ini *clustering* dengan *k-means* menjadi pilihan untuk pengolahan data dari Dinas PUPR. *Evaluation* dilakukan dengan menghitung DBI dan menganalisis hasil apakah sesuai dengan tujuan utama atau melihat apabila ada permasalahan yang belum terpikirkan sebelumnya. Selain itu pada tahap ini juga dipertimbangkan hasil akhir dari proses *clustering*. *Deployment* dilakukan dengan menerapkan hasil proses *data mining*, dimana pada kasus ini hasil akhirnya berupa visualisasi pengetahuan.

Pada Gambar 2 terdapat RM1 dan RM2 yang merupakan rumusan masalah yang ingin dijawab pada penelitian ini. Proses pemecahan masalah tersebut dilakukan sesuai tahapan metode CRISP-DM. Rumusan Masalah 1 (RM 1) yaitu mendapatkan model klusterisasi yang menghasilkan pengetahuan yang dapat membantu pengelompokan pembangunan jalan di Dinas PUPR. Untuk menjawab RM 1 dilakukan *data understanding* sampai dengan tahapan *modelling*. Rumusan Masalah 2 (RM 2) yaitu mengetahui karakteristik pembangunan jalan berdasarkan

klasterisasi untuk membantu pengelompokan prioritas pembangunan jalan selanjutnya di Dinas PUPR berdasarkan hasil *clustering*. Untuk menjawab RM 2 dilakukan tahap *modelling* dan tahap *evaluation*.

Hasil dan Pembahasan

Penelitian pengelompokan pembangunan jalan ini menerapkan pendekatan CRISP-DM dengan tahapan penelitian sesuai Gambar 1 pada uraian metode penelitian dengan penjelasan hasil sebagai berikut:

1. *Business Understanding*

Berdasarkan hasil literasi jurnal pada penelitian sebelumnya terdapat beberapa hasil pengelompokan dengan tujuan yang beragam dengan beberapa algoritma berbeda yang diuraikan pada Kajian Pustaka. Pada tahapan ini dilakukan wawancara dengan Dinas PUPR Kabupaten Garut pada tanggal 17 Mei 2022 dengan Ibu Riska sebagai staf Bina Marga. Dinas PUPR Kabupaten Garut setiap tahunnya menerima pengajuan pembangunan jalan dari seluruh desa di Kabupaten Garut, namun pada penentuan prioritas jalan masih memiliki kendala karena hanya dicocokkan berdasarkan jenis, panjang, dan nominal tanpa aturan yang jelas. *Data mining* dipilih untuk menyelesaikan permasalahan tersebut untuk mendapat model klasterisasi dan karakteristik pengelompokan pembangunan jalan terbaik yang diuraikan pada RM1 dan RM2 pada metode penelitian. Setelah wawancara selanjutnya dilakukan tahapan *data understanding* yaitu pengumpulan dan identifikasi data.

2. *Data Understanding*

Pada *data understanding* dilakukan pengumpulan dan identifikasi data, dimana pengumpulan data dilakukan di hari yang sama pada saat wawancara bersama Dinas PUPR Kabupaten Garut pada tanggal 17 Mei 2022 dengan Ibu Riska sebagai staf Bina Marga. *Dataset* yang diperoleh berjumlah 263 *record* data pembangunan jalan dari tahun 2015 sampai 2021. Tabel 1 menunjukkan data pembangunan jalan yang diperoleh dari Dinas PUPR Kabupaten Garut.

Tabel 1 Data Pembangunan Jalan di Dinas PUPR Kabupaten Garut

No	Kode Pembangunan	Jenis Pembangunan	Panjang	Nominal
1	rk001	2	1400	1.353.660.000
2	rk002	2	221	190.502.000
3	rk003	2	221	190.502.000
4	rk004	2	221	190.502.000
...
42	rh001	1	274	191.416.400
...
108	pb001	3	500	3.868.949.000
...
263	rh100	1	131	91.516.600

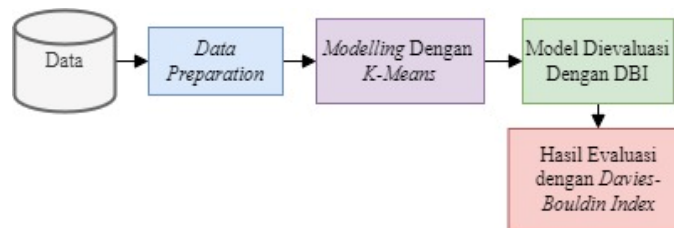
Dari Tabel 1 dapat dilihat bahwa data sejumlah 263 *record* diberi kode pb untuk pembangunan jalan, kode rk untuk rekonstruksi jalan, dan rh untuk rehabilitasi jalan. Atribut jenis pembangunan berisi angka 1, 2, dan 3, yang berarti 1 untuk rehabilitasi jalan, 2 untuk rekonstruksi jalan, dan 3 untuk pembangunan jalan dari nol. Atribut panjang diukur dalam skala meter dari 73 sampai 1400 meter dan atribut nominal diukur dalam skala rupiah dari 91 juta sampai 3,8 miliar.

3. *Data preparation*

Setelah melalui tahapan identifikasi data, pada tahapan ini data yang dikumpulkan tidak dilakukan integrasi data, karena data sudah sesuai kebutuhan yang langsung diperoleh dari Ibu Riska selaku staf Bina Marga di Dinas PUPR Kabupaten Garut. Proses *reduction* juga tidak dilakukan karena jumlah data sudah cukup baik dan atribut yang ada sudah memenuhi untuk pemodelan. *Data cleaning* juga tidak dilakukan karena data yang terkumpul tidak terdapat kekosongan atau *missing value*. Proses *data transformation* dilakukan dengan mengubah kode pembangunan menjadi ID, dimana rh dengan angka 10, rk dengan angka 20, dan pb dengan 30, yang bertujuan untuk memudahkan dalam pemodelan.

4. *Modelling*

Pada tahapan *modelling* dilakukan pemodelan dengan *k-means* untuk mengelompokkan pembangunan jalan, dengan tahapan pemodelan seperti pada Gambar 3. Tahapan *modeling* dilakukan dengan menerapkan *k-means* dan dievaluasi dengan DBI sesuai dengan tahapan-tahapan yang dijelaskan pada bagian Metode Penelitian dan telah digambarkan pada Gambar 3 pemodelan penerapan pengelompokan. Hasil *modelling* dijelaskan sebagai berikut:



Gambar 3 Pemodelan Penerapan Pengelompokan

- a. Pengelompokan pembangunan jalan di Dinas PUPR Kabupaten Garut ini dimulai dari mengelompokkan menjadi dua kelompok sampai dengan lima kelompok. Dataset pembangunan jalan di Kabupaten Garut didapat dari tahun 2015 sampai dengan 2021, berjumlah 263 *record*, dengan atribut kode yang berisikan data pembangunan jalan rehabilitasi (1), rekonstruksi (2), dan pembangunan jalan dari nol (3), panjang yang berisikan data mulai 73 meter sampai 1400 meter, dan nominal berisikan data dari 91 juta sampai 3,8 miliar.
- b. Dalam pengelompokan menjadi K2 sampai dengan K5 digunakan perhitungan titik pusat setiap klaster dengan Persamaan (1), kemudian dihitung *centroid* barunya dengan Persamaan (2). Iterasi berhenti apabila rerata perubahan

centroid di bawah nilai toleransi atau rerata menghasilkan nol atau *centroid* tidak bergerak. Pada pengelompokan ini perhitungan K2 berhenti pada iterasi ke-3, K3 berhenti pada iterasi ke-4, K4 berhenti pada iterasi ke-4, dan K5 pada iterasi ke-3, dimana hasil perhitungan rerata menghasilkan nilai nol (0) atau *centroid* tidak bergerak.

- c. Pengelompokan pembangunan jalan ini menghasilkan jumlah data yang termasuk dalam kelompok berdasarkan kemiripan dari hasil perhitungan titik tengah. Tabel 2 memperlihatkan data dalam kelompok pada pengelompokan pembangunan jalan menjadi K2 sampai K5.

Tabel 2 Data dalam Klaster Pengelompokan K2 sampai K5

Pengelompokan	Data dalam Klaster				
	C1	C2	C3	C4	C5
2	17	246			
3	11	31	221		
4	11	20	28	204	
5	11	6	22	204	20

Dapat dilihat bahwa hasil pengelompokan menghasilkan data seperti yang terdapat pada Tabel 2. Data dalam kelompoknya menghasilkan jumlah data yang relatif sama, dimana C1 berisikan 11 dan hanya berbeda pada pengelompokan K2. Setelah dilakukan *modelling*, pengelompokan dievaluasi dengan indeks *Davies-Bouldin* untuk mengetahui pengelompokan mana yang nilai DBI-nya mendekati nol (0) atau termasuk pengelompokan terbaik.

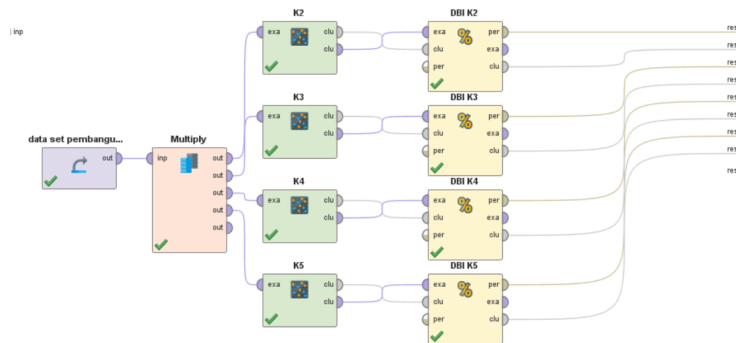
5. Evaluation

Evaluasi hasil pengelompokan pembangunan jalan dilakukan dari K2 sampai dengan K5, dengan kegiatan awal menghitung SSW menggunakan Persamaan (3). Setelah didapatkan SSW, maka selanjutnya menghitung SSB dengan Persamaan (4). Setelah SSW dan SSB dihitung, maka dihitung Rasio untuk mendapat nilai perbandingan antar klaster dengan Persamaan (5). Klaster baik memiliki nilai kohesi paling kecil dan separasi yang paling besar. Nilai R_{max} yang didapatkan merupakan total dari R maksimal pada setiap tabel matriks, kemudian dijumlahkan dan dihitung DBI dari K2 sampai K5. Jika nilai DBI yang didapatkan semakin kecil tidak negatif atau lebih dari 0, maka kelompok yang diperoleh dari pengelompokan *k-means* semakin baik. Hasil evaluasi K2 sampai dengan K5 ditampilkan pada Tabel 3.

Tabel 3 Hasil Perhitungan Evaluasi DBI

No	Pengelompokan	DBI
1.	2	0,3299
2.	3	0,2219
3.	4	0,2826
4.	5	0,1617

Hasil perhitungan evaluasi dengan DBI dari pengelompokan K2 sampai dengan K5 menggunakan *Microsoft Excel* menghasilkan perhitungan yang ditampilkan pada Tabel 3. DBI terkecil terletak pada pengelompokan K5 dengan nilai 0,1617, yang menandakan pengelompokan ini merupakan hasil kluster yang baik. Selanjutnya dilakukan *modelling* dengan *tools RapidMiner* untuk mengetahui apakah hasil perhitungan manual sudah baik dengan cara membandingkan dengan hasil *RapidMiner*. Gambar 4 merupakan operator *clustering* dengan *RapidMiner*.



Gambar 4 Pengelompokan K2 sampai K5 dengan *Tools RapidMiner*

Perbandingan dilakukan untuk membandingkan hasil manual dan *tools*, apakah perhitungan manual benar menghasilkan pengelompokan terbaik. Untuk perbandingan hasil *Microsoft Excel* dan *RapidMiner* ditampilkan pada Tabel 4.

Tabel 4 Perbandingan Data dalam Kluster

K	<i>Microsoft Excel</i>					<i>RapidMiner</i>				
	C1	C2	C3	C4	C5	C1	C2	C3	C4	C5
2	17	246				17	246			
3	11	31	221			11	31	221		
4	11	20	28	204		11	20	28	204	
5	11	6	22	204	20	11	6	22	204	20

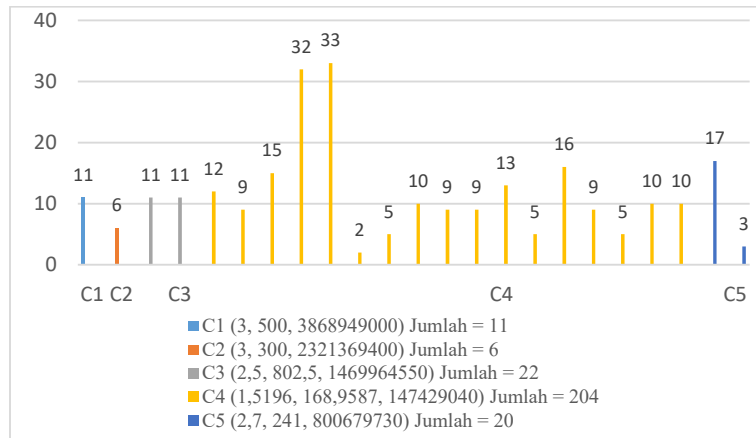
Tabel 4 merupakan perbandingan pengelompokan, dimana keduanya menghasilkan jumlah data dalam kelompok yang sama, baik pengelompokan dengan *Microsoft Excel* ataupun *RapidMiner* dari K2 sampai dengan K5. Hal ini membuktikan bahwa pengelompokan manual sudah sesuai dengan *tools*. Selanjutnya perbandingan hasil evaluasi dengan DBI diberikan pada Tabel 5. Dari Tabel 5 dapat dilihat bahwa perhitungan evaluasi DBI, keduanya menghasilkan nilai evaluasi yang serupa baik manual ataupun *RapidMiner* dan K5 merupakan pengelompokan dengan DBI terbaik.

Tabel 5 Perbandingan Hasil Evaluasi

K	<i>Microsoft Excel</i>	<i>RapidMiner</i>
2	0,3299	0,330
3	0,2219	0,222
4	0,2826	0,283
5	0,1617	0,162

6. Deployment

Pada tahap ini, pengetahuan atau informasi yang telah diperoleh diatur dan dipresentasikan dalam bentuk khusus sehingga dapat digunakan oleh Dinas PUPR. Gambar 5 merupakan hasil dari pengelompokan lima kluster.



Gambar 5 Hasil Pengelompokan Lima Kluster

Gambar 5 merupakan visualisasi grafik batang hasil pengelompokan menjadi lima kelompok, dimana sumbu *y* merupakan jumlah dalam kelompok dan sumbu *x* merupakan data pembangunan jalan yang masuk dalam kluster. Pembangunan jalan di Dinas PUPR Kabupaten Garut menghasilkan kelompok dengan karakteristik Kelompok 1 dengan jenis pembangunan yaitu pembangunan jalan dari nol, dengan panjang 500 meter, dan nominal Rp 3,8 miliar, Kelompok 2 dengan pembangunan jalan dari nol yang panjangnya 300 meter, dan nominal Rp 2,3 miliar, Kelompok 3 dengan jenis pembangunan rekonstruksi dan pembangunan jalan dari nol yang panjangnya dari 205 meter sampai 1400 meter dengan nominal Rp 1,3 miliar sampai dengan Rp 1,58 miliar, Kelompok 4 merupakan jenis pembangunan rehabilitasi dan rekonstruksi yang panjangnya dari 73 meter sampai 335 meter dengan kisaran nominal Rp 90 juta sampai dengan Rp 198 juta, dan Kelompok 5 dengan jenis pembangunan rekonstruksi dan pembangunan jalan dari nol yang panjangnya dari 100 meter sampai 1040 meter dengan kisaran nominal Rp 770 juta sampai dengan Rp 950 juta. Nilai DBI yang didapatkan cukup baik yaitu sebesar 0,1617 dan titik *centroid* akhir dari iterasi ke-3 yaitu *centroid* 1 berada pada (3, 500, 3868949000), *centroid* 2 berada pada (3, 300, 2321369400), *centroid* 3 berada pada (2,5, 802,5, 1469964550), *centroid* 4 berada pada (1,5196, 168,9587, 147429040), dan *centroid* 5 berada pada (2,7, 241, 800679730). Jika diurutkan dari kelompok terbesar sampai terkecil dalam pengelompokan K5 maka Kelompok 1 > Kelompok 2 > Kelompok 3 > Kelompok 5 > Kelompok 4.

Berdasarkan hasil penelitian diketahui pengelompokan pembangunan jalan dengan algoritma *k-means* yang terbaik adalah dengan dikelompokkan menjadi lima kelompok dengan acuan nilai DBI yang mendekati nol. Semakin dekat nilai DBI ke nol/0 maka semakin baik. Hasil ini menjawab RM 1 yaitu mendapatkan

model klasterisasi yang baik untuk proses penentuan prioritas pembangunan jalan. Hasil penelitian ini selaras dengan penelitian yang dilakukan oleh [9] dimana *k-means* mampu menghasilkan pengelompokan dengan nilai DBI yang baik. Dengan ini Dinas PUPR dapat menentukan prioritas pembangunan jalan berdasarkan kelompoknya, apakah pembangunan tahun ini lebih diprioritaskan pembangunan jalan yang mengeluarkan dana yang besar, rekonstruksi yang sedang, kecil atau lainnya. Hasil ini membantu penerapan pembangunan dapat lebih awal dilakukan dengan karakteristik pembangunan jalan sesuai dengan kelompok pembangunannya.

Berdasarkan hasil visualisasi pada Gambar 5 dapat diketahui karakteristik pengelompokan pembangunan jalan, yang mana hal ini menjawab RM 2. Jenis pembangunan bergantung pada nominal pembangunan, seperti jenis kegiatan rekonstruksi yang memiliki panjang sama dengan pembangunan jalan dari nol yaitu 100 meter, memiliki nominal pembangunan jalan yang jauh berbeda dimana rekonstruksi 100 meter bernominal 90 juta, sementara pembangunan jalan dari nol 100 meter bernominal 770 juta. Pembangunan jalan dengan nominal tertinggi yaitu pada C1 kemudian C2, C3, C5, lalu C4. Hal ini membuktikan bahwa nominal berkaitan dengan jenis pembangunan jalan dan panjang pembangunannya. Hasil tersebut dapat memberikan pemahaman kepada penerima manfaat mengenai hasil pengetahuan dari penelitian ini untuk dijadikan pertimbangan dalam menentukan prioritas pembangunan jalan di tahun selanjutnya. Hal ini dapat menjawab permasalahan yang sebelumnya terjadi mengenai penentuan prioritas pembangunan jalan bisa terminimalisir atau bahkan tidak terjadi kembali.

Simpulan

Berdasarkan hasil penelitian pengelompokan 263 data pembangunan jalan yang telah dilakukan dengan metode *k-means clustering*, maka dapat diambil kesimpulan yaitu algoritma *k-means* dapat mengelompokkan pembangunan jalan dengan baik. Hasil evaluasi DBI terbaik dalam pengelompokan lima kelompok diperoleh nilai sebesar 0,1617. Nilai ini merupakan evaluasi terbaik karena sudah mendekati nilai nol. Pembangunan jalan klaster terbesar yaitu diawali C1, C2, C3, C5, dan C4. Nominal pembangunan sangat dipengaruhi oleh jenis pembangunan jalan. Walaupun panjang pembangunan berpengaruh, namun dengan panjang yang sama, seperti panjang pembangunan 100 meter pada rekonstruksi, nominal pembangunannya jauh berbeda dengan pembangunan jalan dari nol yang panjangnya 100 meter. Dengan ini Dinas PUPR dapat memberikan kuota pembangunan jalan yang lebih banyak pada kelompok pembangunan jalan C4 atau dengan nominal yang paling kecil dibandingkan kelompok pembangunan jalan sedang atau pembangunan jalan yang besar. Untuk pengembangan penelitian yang serupa di kemudian hari, disarankan untuk menambahkan atribut seperti tinggi dan lebar atau biasa disebut dengan volume untuk mengetahui lebih banyak faktor yang mempengaruhi nominal pembangunan jalan.

Daftar Pustaka

- [1] N. Puspitasari, R. Rosmasari, and S. Stefanie, "Penentuan Prioritas Perbaikan Jalan Menggunakan Fuzzy C-Means : Studi Kasus Perbaikan Jalan Di Kota Samarinda," *J. Teknol. dan Sist. Komput.*, vol. 5, no. 1, p. 7, 2017, doi: 10.14710/jtsiskom.5.1.2017.7-14.
- [2] S. Asmiatun and N. Wakhidah, "Identifikasi Pengelompokan Kondisi Permukaan Jalan Menggunakan Algoritma *K-means*," *J. Pengemb. Rekayasa dan Teknol.*, vol. 14, no. 1, p. 17, 2018, doi: 10.26623/jprt.v14i1.1215.
- [3] D. Kurniadi, E. Abdurachman, H. L. H. S. Warnars, and W. Suparta, "A proposed framework in an intelligent recommender system for the college student," *J. Phys. Conf. Ser.*, vol. 1402, no. 6, 2019, doi: 10.1088/1742-6596/1402/6/066100.
- [4] K. A. Terbuka, "Prediksi penerima beasiswa di perguruan tinggi menggunakan algoritma k-Nearest Neighbor Prediksi penerima beasiswa di perguruan tinggi menggunakan algoritma k-Nearest Neighbor," 2018.
- [5] L. Listiani, Y. H. Agustin, and M. Z. Ramdhani, "Implementasi algoritma *k-means* klaster untuk rekomendasi pekerjaan berdasarkan pengelompokan data penduduk," *SENSITf Semin. Nas. Sist. Inf. dan Teknol. Inf.*, pp. 761–769, 2019.
- [6] S. Asmiatun, N. Wakhidah, and A. N. Putri, "Penerapan Metode K-Medoids Untuk Pengelompokan Kondisi Jalan Di Kota Semarang," *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 6, no. 2, pp. 171–180, 2020, doi: 10.35957/jatisi.v6i2.193.
- [7] A. A. Vernanda, A. Faisol, and N. Vendyansyah, "Penerapan Metode *K-means* Klastering Untuk Pemetaan Daerah Rawan Kecelakaan Lalu Lintas Di Kota Malang Berbasis Website," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 5, no. 2, pp. 836–844, 2021, doi: 10.36040/jati.v5i2.3791.
- [8] Z. Nabila, A. Rahman Isnain, and Z. Abidin, "Analisis Data Mining Untuk Klastering Kasus Covid-19 Di Provinsi Lampung Dengan Algoritma *K-means*," *J. Teknol. dan Sist. Inf.*, vol. 2, no. 2, p. 100, 2021, [Online]. Available: <http://jim.teknokrat.ac.id/index.php/JTSSI>
- [9] R. Gustrianda and D. I. Mulyana, "Penerapan Data Mining Dalam Pemilihan Produk Unggulan dengan Metode Algoritma *K-means* Dan K-Medoids," *J. Media Inform. Budidarma*, vol. 6, no. 1, p. 27, 2022, doi: 10.30865/mib.v6i1.3294.
- [10] D. J. Lubis and M. B. Tamam, "Penerapan *K-means* Untuk Pengelompokan Beasiswa Santri di Pondok Pesantren Miftahul Huda Bogor," *Teknois J. Ilm. Teknol. Inf. dan Sains*, vol. 12, no. 1, pp. 7–20, 2022, doi: 10.36350/jbs.v12i1.125.
- [11] R. Hasibuan Budiansyah, H. Hafizah, and R. Mahyuni, "Penerapan Data Mining Klastering Dengan Menggunakan Algoritma *K-means* Pada Data Nasabah Kredit Bermasalah PT. BPR Milala," *J-SISKO TECH (Jurnal Teknol. Sist. Inf. dan Sist. Komput. TGD)*, vol. 5, no. 1, p. 7, 2022, doi: 10.53513/jsk.v5i1.4767.
- [12] Y. Suhandi, I. Kurniati, and S. Norma, "Penerapan Metode Crisp-DM Dengan Algoritma *K-means* Klastering Untuk Segmentasi Mahasiswa Berdasarkan Kualitas

-
- Akademik,” *J. Teknol. Inform. dan Komput.*, vol. 6, no. 2, pp. 12–20, 2020, doi: 10.37012/jtik.v6i2.299.
- [13] L. Kartikawati, K. Kusriani, and E. T. Luthfi, “Algoritma *K-means* pada Pengelompokan Pembelajaran Tatap Muka Terbatas Sesudah Vaksinasi COVID-19,” *J. Eksplora Inform.*, vol. 11, no. 1, pp. 20–28, 2022, doi: 10.30864/eksplora.v11i1.560.
- [14] H. A. Nizar and V. E. Purba, “Evaluasi Jalan Rabat Beton baru pada STA 2.000 -STA 3.000 di Jalan Jambuara Nagori Buntu Bayu Kecamatan Hatanduhan Kabupaten Simalungun,” *Angew. Chemie Int. Ed.* 6(11), 951–952., pp. 10–27, 2018.
- [15] F. A. Jatnika, J. Akbardin, R. Natawidjana, D. Nanang, and D. Herman, “Perancangan Petunjuk Pelaksanaan (Sop) Berdasarkan Rencana Pembangunan Jalan Di Provinsi Banten,” no. 229, 2020.