

KLASTERISASI PERKARA PELANGGARAN LALU LINTAS MENGUNAKAN ALGORITMA K-MEANS DAN DAVIES- BOULDIN INDEX

Abdussalam¹, Fikria Afidatun Nisa², Ajib Susanto³, Elkaf Rahmawan Pramudya⁴, Ibnu Utomo Wahyu Mulyono⁵

^{1,2,3,4,5}*Teknik Informatika, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro Semarang*

Jl. Imam Bonjol No. 207, Semarang

E-mail : grey.salam@dsn.dinus.ac.id¹, fikrianisaa@gmail.com², ajib.susanto@dsn.dinus.ac.id³,
elkaf.rahmawan@dsn.dinus.ac.id⁴, ibnu.utomo.wm@dsn.dinus.ac.id⁵

Abstrak

Pengadilan Negeri Slawi adalah sebuah lembaga yang menyelesaikan berbagai macam perkara diantaranya yaitu perkara pelanggaran lalu lintas. Ketertiban lalu lintas di jalan sangat penting bagi tiap pengendara di jalan raya, kurangnya akan kesadaran dan perilaku buruk para pengendara akan kedisiplinan berlalu lintas berakibat banyaknya data tilang yang diterima oleh Pengadilan Negeri Slawi. Data yang banyak tersebut belum sepenuhnya dimanfaatkan untuk kepentingan terkait. Penelitian ini dilakukan untuk menganalisis dan mengelompokkan data perkara pelanggaran lalu lintas tahun 2017 dengan metode K-Means supaya mempermudah untuk mengetahui jenis pelanggaran yang sering kali dilakukan oleh pengguna kendaraan. Dimana atribut yang ditentukan sebanyak 5 atribut yaitu no register pelanggaran, form, pasal pelanggaran, jenis kendaraan, dan jenis kelamin. metode evaluasi cluster menggunakan metode Davies Bouldin Index (DBI) dengan 3 cluster dihasilkan 0,100.

Kata Kunci: *K-Means, Data Mining Pelanggaran lalu lintas, Klasterisasi, Davies-Bouldin Index*

I. PENDAHULUAN

Di Indonesia pertumbuhan jumlah kendaraan semakin pesat. Terbukti ditahun 2017 jumlah kendaraan yang dibagi sesuai jenisnya antara lain kendaraan mobil penumpang mencapai 15.423.968 unit, mobil bus 2.509.258 unit, mobil barang 7.289.910 unit, dan sepeda motor mencapai 111.988.683 unit [1]. Terutama untuk kendaraan pribadi, masyarakat lebih memilih menggunakan kendaraan pribadi karena dinilai lebih memudahkan saat menjalankan aktivitas sehari-hari. Namun, disisi lain semakin banyak permintaan tersebut masih banyak pula masyarakat yang minim edukasi dalam menggunakan kendaraan [2].

Pelanggaran banyak jenisnya, salah satunya pelanggaran lalu lintas. Dimana perbuatan atau tindakan yang tidak sesuai dengan Undang-undang mengenai lalu lintas. Maka dari itu bagi pengendara yang melanggar akan mendapat kartu tilang. Masih banyak masyarakat yang melanggar lalu lintas terutama kendaraan pribadi, salah satunya yang berada di Kabupaten Slawi dengan pelanggaran pada tahun 2017 mencapai 39395 [3].

Dengan terus bertambahnya pelanggaran lalu lintas khususnya di Kabupaten Slawi, maka jumlah perkara lalu lintas juga semakin meningkat. data pelanggaran tersebut akan diolah dengan teknik data mining untuk menggali informasi dan pengetahuan baru pada pola-pola yang terbentuk dari data yang besar. Penerapan data mining pada data pelanggaran lalu lintas untuk menentukan pola *cluster* perkara pelanggaran berlalu lintas di Kabupaten Slawi dengan memanfaatkan algoritma *K-Means*. Dari penelitian sebelumnya clustering wilayah dan pelanggaran lalu lintas di Tasikmalaya Kota[4] terbukti mampu mengelompokkan wilayah pelanggaran lalu lintas kelompok 1 dengan 14 daerah, kelompok 2 dengan 6 daerah dan kelompok 3 dengan 4 daerah. Penelitian menggunakan K-Means[5] untuk pelanggaran di

Pamekasan menghasilkan 5 kelompok dengan 1 kelompok tertinggi dengan pelanggaran tidak memiliki surat ijin mengemudi (SIM). K-Means juga diterapkan dalam analisis sintimen pada pilpres 2019[6] untuk klusterisasi data latih dengan menghasilkan nilai bobot positif dan negative. Pada penelitian ini diusulkan klusterisasi dengan menggunakan K-Means dan metode evaluasi klaster dengan menggunakan *Davies Bouldin Index (DBI)*[7]. Penemuan pola *cluster* yang dihasilkan diharapkan dapat memberikan informasi mengenai pola *cluster* pelanggaran yang sering dilakukan oleh pengendara sehingga dimanfaatkan oleh Kepolisian sebagai dasar pertimbangan dalam upaya mengurangi tingkat pelanggaran lalu lintas di Kabupaten Slawi misalnya dengan memberikan sosialisasi rutin kepada masyarakat mengenai peraturan berlalu lintas yang benar, baik dan aman serta akibat yang terjadi apabila melanggar peraturan berlalu lintas.

II. METODOLOGI PENELITIAN

1. Perkara Pelanggaran Lalu Lintas

Perkara pelanggaran lalu lintas merupakan jenis perkara yang diperiksa melalui acara cepat. pemeriksaan acara cepat diatur dalam bab XVI bagian keenam KUHAP yang terbagi menjadi dua golongan yaitu acara pemeriksaan tindak pidana ringan (pasal 205-pasal 210 KUHAP) dan acara pemeriksaan perkara pelanggaran lalu lintas (pasal 211-pasal 216 KUHAP)[8].

2. Data Mining

Data mining memiliki pengertian pengolahan data yang berbasis sistematis. Selain itu bisa juga disebut kecerdasan buatan atau *machine learning*. Dimana memiliki fungsi untuk ekstraksi data. Selain itu data mining digunakan juga untuk mengidentifikasi informasi yang terkait dari berbagai *database* [9].

3. Pre-processing

Data yang akan digunakan pada penelitian ini merupakan data register pelanggaran lalu lintas Pengadilan Negeri Slawi pada bulan Januari tahun 2017 sebanyak 500 data. Atribut dari data tersebut yaitu : No Register Tilang; Form; Nama; Alamat; Pasal Pelanggaran; Nomor Polisi; Barang Bukti; Jenis Kendaraan; Nomor Perkara atau Putusan; Denda; Biaya Perkara; dan Subsider. Data ini akan melewati tahap *preprocessing*, untuk memilih dan mendapatkan data yang baik serta berkualitas untuk kemudian masuk ke tahap *clustering*.

1. Pemilihan Data

Pada tahap ini dilakukan pemilihan atribut data yang akan digunakan pada penelitian, maka untuk atribut data yang tidak digunakan dan data yang tidak valid atau *missvalue* akan dihilangkan. Pada Gambar 1 disajikan atribut yang digunakan.

NOMOR REGISTER TILANG	FORM	JENIS KELAMIN	PASAL	JENIS KENDARAAN
2146849	Biru	L	288	Sepeda Motor
2146996	Biru	L	281	Sepeda Motor
2146999	Biru	P	281	Sepeda Motor
2146998	Biru	L	281	Sepeda Motor
2146120	Merah	P	281	Sepeda Motor
5328675	Merah	L	281	Sepeda Motor
2147106	Merah	L	281	Sepeda Motor
2147107	Merah	L	281	Sepeda Motor
2147108	Merah	P	281	Sepeda Motor
2146880	Merah	L	281	Sepeda Motor
2146878	Merah	L	281	Sepeda Motor
2147144	Merah	L	281	Sepeda Motor
2147143	Merah	P	281	Sepeda Motor
2147001	Merah	P	281	Sepeda Motor
2147142	Merah	P	281	Sepeda Motor
2147141	Merah	L	281	Sepeda Motor
2146542	Merah	P	287	Sepeda Motor
2146931	Merah	L	281	Sepeda Motor
2146932	Merah	P	281	Sepeda Motor

Gambar 1. atribut *dataset* yang digunakan

2. Transformasi Data

Transformasi data dilakukan untuk mengubah data atau dilakukan proses inialisasi dimana data diubah kedalam bentuk angka atau numerik agar data dapat diolah secara baik menggunakan metode K-Means. Berikut merupakan tabel inialisasi dari atribut data.

Tabel 1. inialisasi atribut

ATRIBUT	INISIALISASI
Nomor Register Tilang	A

Form	B
Jenis Kelamin	C
Pasal	D
Jenis Kendaraan	E

Tabel 1 di atas adalah inisialisasi atribut. Atribut yang digunakan adalah lima atribut yang akan digunakan dalam proses *clustering*.

Tabel 2. inisialisasi No register

NO REGISTER TILANG	INISIALISASI
D3479739	1
D3479823	2
D3479310	3
D3480521	4
D3480545	...
D3479603	499
D3454616	500

Pada data no register tilang data satu dengan lainnya tidak sama, maka di inisialisasikan menjadi 1-500 seperti Tabel 2.

Tabel 3. inisialisasi formulir

FORMULIR	INISIALISASI
Biru	1
Merah	2

Untuk data form terdiri dari dua jenis data yaitu form berwarna biru yang merupakan penilangan secara online, sedangkan form yang berwarna merah untuk penilangan secara langsung. diinisialisai seperti Tabel 3.

Tabel 4. inisialisasi pasal

PASAL	KETERANGAN	INISIALISAI
Pasal 104 (3)	Setiap pengguna jalan tidak mematuhi perintah yang diberikan petugas	1
Pasal 280	Tak dipasang Tanda Nomor Kendaraan	2
Pasal 281	Tidak memiliki SIM	3
Pasal 283	Kegiatan yang mengganggu konsentrasi saat berkendara	4
Pasal 284	Kendaraan bermotor tidak mengutamakan pejalan kaki atau pesepeda	5
Pasal 285	Pengemudi Sepeda Motor tidak memenuhi persyaratan teknis dan liak jalan	6
Pasal 287	Melanggar rambu lalu lintas	7
Pasal 287 (2)	Melanggar aturan perintah atau melanggar yang dinyatakan dengan Alat Pemberi Isyarat Lalu Lintas	8
Pasal 288 (1)	Tidak Memiliki STNK	9
Pasal 288 (2)	Pengendara Sepeda Motor yang memiliki SIM namun tidak dapat menunjukkannya	10
Pasal 288 (3)	Kendaraan tidak dilengkapi dengan surat keterangan uji berkala dan tanda lulus uji berkala	11
Pasal 289	Setiap pengemudi atau penumpang yang duduk disamping pengemudi mobil tidak menggunakan sabuk keselamatan	12
Pasal 291	Tidak mengenakan helm standar nasional	13
Pasal 291 (1)	Setiap pengendara atau penumpang sepeda motor tidak menggunakan helm standar nasional	14
Pasal 291 (2)	Penumpang tidak menggunakan helm standar nasional	15
Pasal 298	Kendaraan bermotor yang tidak menggunakan isyarat saat berhenti ayau parkir darurat	16
Pasal 300	Penggunaan jalur atau lajur. Tidak menggunakan lajut yang telah ditentukan	17
Pasal 303	Pengemudi mobil barang untuk mengangkut orang	18
Pasal 307	Tidak mematuhi ketentuan mengenai tata cara pemuatan,	19

daya angkut.

Data pasal yang terdiri dari 19 jenis pasal pelanggaran, diinisialisasikan seperti Tabel 4.

Tabel 5. inisialisasi jenis kendaraan

JENIS KENDARAAN	INISIALISASI
Sepeda Motor	1
Mobil Pickup	2
Mobil Penumpang Pribadi	3
Mobil Penumpang Umum	4
Bus	5
Truk	6

Data jenis kendaraan terdiri dari enam jenis, diinisialisasikan seperti Tabel 5.

Tabel 6. inisialisasi jenis kelamin

NAMA PELANGGAR	JENIS KELAMIN	INISIALISAI
Tarsono	Laki-laki	1
Sri Yulinigtyas	Perempuan	2
Titik Wijayanti	Perempuan	2
Abdul Manap	Laki-laki	1
Karyo	Laki-laki	1
Septiyan Bagas	Laki-laki	1

Data jenis kelamin terdiri dari dua jenis, diinisialisasikan seperti Tabel 6.

Setelah semua data telah melalui tahap inisialisasi, kemudian dilakukan tahap transformasi data keseluruhan sehingga lebih memudahkan dalam proses pengelompokan data, terlihat seperti Tabel 7.

Keterangan :

A = Nomor Register Tilang

B = Form

C = Jenis Kelamin

D = Pasal Pelanggaran

E = Jenis Kendaraan

Tabel 7. data setelah *preprocessing*

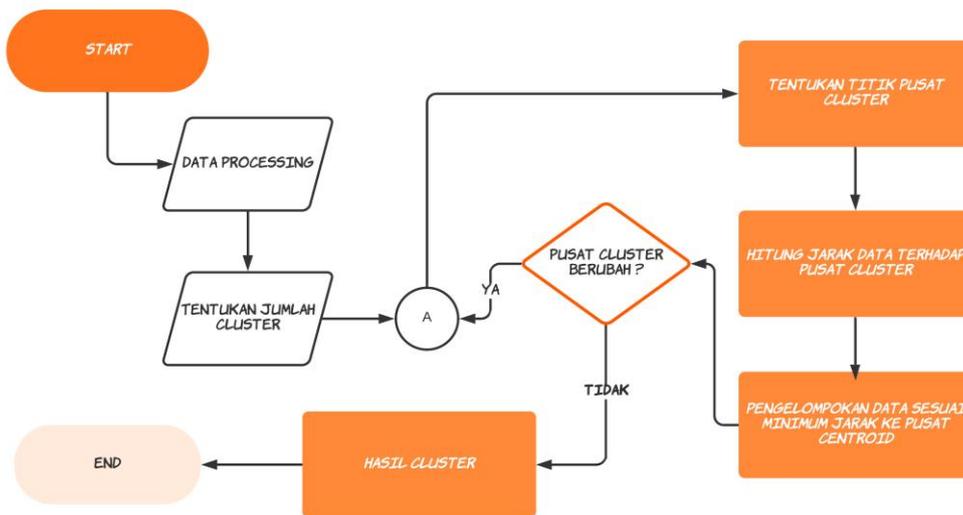
No	A	B	C	D	E
1	1	1	1	1	1
2	2	1	2	1	1
3	3	1	2	1	2
4	4	1	2	1	1
5	5	2	2	1	2
...
100	100	2	2	1	1
101	101	2	2	1	1
102	102	2	2	1	1
103	103	2	2	1	1
104	104	2	2	1	1
105	105	2	2	1	1
...
496	496	2	4	1	1
497	497	2	1	1	1
498	498	2	4	1	1
499	499	2	2	1	1
500	500	2	2	1	1

3. Metode yang Diusulkan

Pada penelitian ini metode yang digunakan adalah *clustering* dengan algoritma K-Means. *Clustering* merupakan salah satu teknik analisis dalam data mining yang melakukan pengelompokan data berdasarkan kesamaan karakteristiknya. Dengan kesamaan karakteristik [5].

Sedangkan *K-Means* merupakan salah satu algoritma atau metode *cluster analysis* non hirarki yang berusaha mempartisi objek yang ada kedalam satu atau lebih *cluster* atau kelompok objek berdasarkan karakteristiknya dan melakukan itersai dalam penentuan titik terdekatnya [4]. Berikut Gambar 2 yang

merupakan alur keseluruhan *K-Means*.



Gambar 2. Alur keseluruhan K-means

Berikut langkah-langkah proses algoritma *K-means* :

1. Menentukan *cluster* secara acak sebagai titik pusat (*centroid*).
2. Menghitung dengan rumus *Euclidian Distance* jarak yang paling terdekat dengan data terhadap titik pusat. Rumus *Euclidian* sebagai berikut :

$$d(i, j) = \sqrt{\sum_j^m (C_{ij} - C_{kj})^2} \quad (1)$$

dimana :

x_i = data kriteria

μ_j = Centroid Pada Cluster ke - j

3. Kelompokan setiap data ke dalam *cluster* dengan jarak terkecil.

$$\text{Min} \sum_{k=1}^k d_{ik} = \sqrt{\sum_j^m (C_{ij} - C_{kj})^2} \quad (2)$$

4. Memperbarui nilai titik pusat dengan melakukan rata-rata *cluster*.

$$C_{kj} = \frac{\sum_{i=1}^p x_{ij}}{p} \quad (3)$$

5. Lakukan perulangan langkah 1, 2, 3 sampai semua anggota *cluster* tidak lagi berubah.
6. Jika langkah yang kelima sudah terpenuhi yang digunakan sebagai parameter untuk menentukan akurasi data adalah nilai pusat *cluster* terakhir.

Setelah itu akan dilakukan pengujian dengan menggunakan *Davies Bouldin Index* (DBI) untuk mengetahui akurasi data yang dikelompokan dengan algoritma *K-Means*. Semakin kecil nilai DBI yang didapat (non-negatif ≥ 0), maka semakin baik kluster yang diperoleh dari pengelompokan *K-Means* yang digunakan [10].

III. HASIL DAN PEMBAHASAN

Proses *clustering* dilakukan dengan data yang telah melalui preprocessing. Data yang sudah melewati tahap *Pre-processing* selanjutnya masuk ke proses *clustering*. Pada penelitian ini menggunakan 5 atribut data yaitu, Nomor Register Tilang, Form, Jenis Kendaraan, Pasal, dan Jenis Kelamin.

1. Perhitungan *K-Means*

Rumus algoritma *K-Means* yang digunakan adalah *Euclidian Distance*, data bulan Januari 2017 sebanyak 500 data dipilih untuk melakukan perhitungan manual.pada Tabel 8 disajikan hasil transformasi 500 data dibulan Januari.

Tabel 8. Dataset Hasil *preprocessing*

No	A	B	C	D	E
1	1	1	1	1	1

2	2	1	2	1	1
3	3	1	2	1	2
...
28	28	2	4	1	2
29	29	2	2	1	1
30	30	2	2	1	1
...
108	108	2	2	1	2
109	109	2	2	1	1
110	110	2	2	1	1
...
498	498	2	4	1	1
499	499	2	2	1	1
500	500	2	2	1	1

- Langkah pertama, menentukan titik pusat kluster (centroid) yang dipilih secara acak. Pada penelitian ini ditetapkan jumlah kluster sebanyak 3 diambil dari data ke-136, 235, 389.

Tabel 9. Titik Pusat Centorid

centroid	A	B	C	D	E
C1	136	2	2	1	1
C2	235	2	4	1	1
C3	389	2	1	1	2

- Selanjutnya menghitung jarak tiap data ke *centroid* terdekat dengan menggunakan rumus *Euclidean distance*. Sehingga didapatkan hasil seperti Tabel 10.

Tabel 10. Hasil Perhitungan iterasi 1

Data ke-i	Atribut					C1	C2	C3
	A	B	C	D	E			
1	1	1	1	1	1	135.007	234.021	388.003
2	2	1	2	1	1	134.011	233.024	387.004
3	3	1	2	1	2	133.008	232.013	386.005
4	4	1	2	1	1	132.011	231.024	385.004
5	5	2	2	1	2	131.004	230.011	384.004
...
...
497	497	2	1	1	1	361.001	262.017	108.005
498	498	2	4	1	1	362.014	263.034	109.046
499	499	2	2	1	1	363.003	264.019	110.009
500	500	2	2	1	1	364.003	265.019	111.009

- Kemudian kelompokan data ke dalam kluster dengan jarak minimal

Tabel 11. Hasil Pengelompokan Jarak Minimal

Data ke-i	Atribut					C1	C2	C3	Jarak terdekat ke-c
	A	B	C	D	E				
1	1	1	1	1	1	135.007	234.021	388.003	C1
2	2	1	2	1	1	134.011	233.024	387.004	C1
3	3	1	2	1	2	133.008	232.013	386.005	C1
4	4	1	2	1	1	132.011	231.024	385.004	C1
5	5	2	2	1	2	131.004	230.011	384.004	C1
...
...
497	497	2	1	1	1	361.001	262.017	108.005	C3
498	498	2	4	1	1	362.014	263.034	109.046	C3
499	499	2	2	1	1	363.003	264.019	110.009	C3
500	500	2	2	1	1	364.003	265.019	111.009	C3

Dari Tabel 11 didapatkan hasil keseluruhan dari 500 data dibulan januari yaitu :

C1 = 185 data

C2 = 126 data

C3 = 189 data

- Langkah selanjutnya hitung kembali titik pusat *cluster* yang baru berdasarkan rata-rata anggota yang terdapat didalam klaster. Hasil perhitungan dari titik pusat *cluster* terbaru terlihat di Tabel 12.

Tabel 12. Titik Pusat *cluster* Baru

Centroid	A	B	C	D	E
C1	93	1.978	1.281	2.486	1.292
C2	248.500	2	1.302	2.746	1.230
C3	406	2	1.434	2.460	1.455

- Langkah kelima dapat mengulang Kembali perhitungan dari langkah pertama, kedua, dan ketiga. Jika pada langkah ketiga hasil perhitungan yang dihasilkan posisi data *cluster* tidak berubah diiterasi pertama maka perhitungan dihentikan. Akan tetapi jika posisi data *cluster* terjadi perubahan maka perhitungan dilanjutkan ke langkah empat. Pada perhitungan ini posisi *cluster* berhenti di iterasi-5, dapat dilihat pada Tabel 13.

Tabel 13. Hasil Akhir *cluster*

Data ke-i	Atribut					C1	C2	C3	Jarak terdekat ke-c
	A	B	C	D	E				
1	1	1	1	1	1	82.519	248.511	415.504	C1
2	2	1	2	1	1	81.508	247.505	414.502	C1
3	3	1	2	1	2	80.511	246.506	413.502	C1
4	4	1	2	1	1	79.508	245.506	412.502	C1
5	5	2	2	1	2	78.505	244.504	411.501	C1
...
...
497	497	2	1	1	1	413.503	247.509	80.515	C3
498	498	2	4	1	1	414.503	248.505	81.518	C3
499	499	2	2	1	1	415.500	249.503	82.504	C3
500	500	2	2	1	1	416.500	250.503	83.504	C3

Dari Tabel 13 didapatkan hasil perhitungan iterasi kelima menggunakan data sebanyak 500, menghasilkan yaitu:

C1 = 166 data

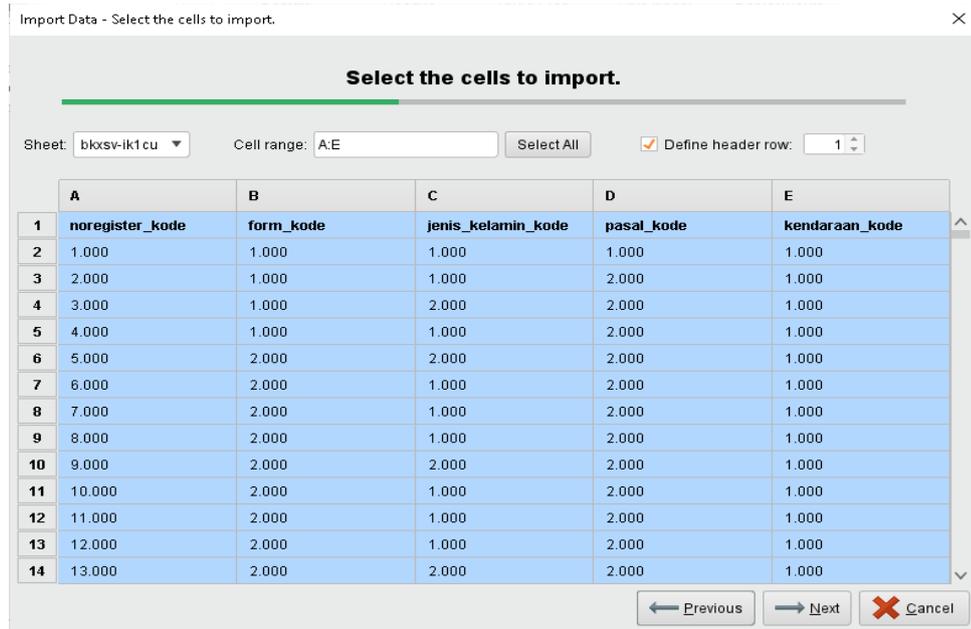
C2 = 166 data

C3 = 168 data

Perhitungan langkah ketiga pada iterasi kelima tidak mengalami perubahan pada literasi keempat sehingga perhitungan berhenti.

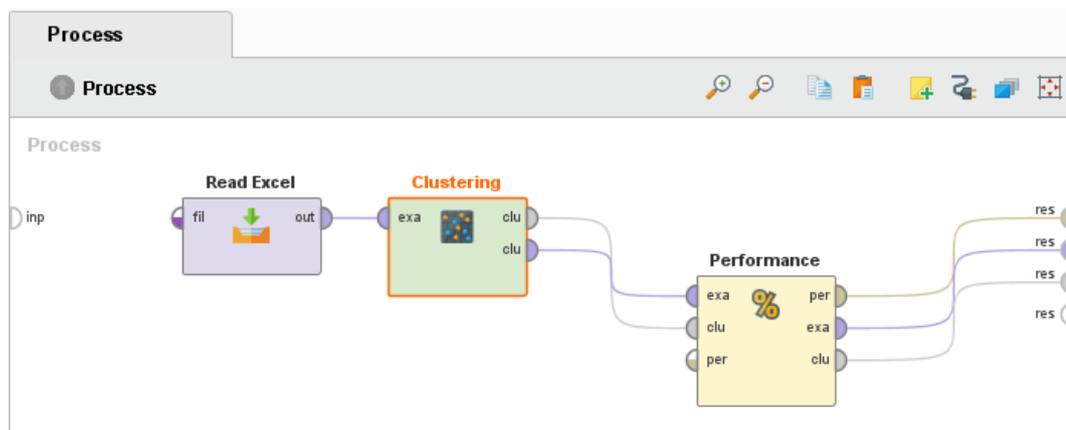
2. Pengujian

Tahap pengujian dilakukan menggunakan aplikasi *RapidMiner*. Gambar 4 berikut merupakan *import* dataset pada bulan Januari yang telah melalui tahap *Preprocessing*.



Gambar 4. Import Dataset

Berikut pada Gambar 5 merupakan desain pemodelan dataset kecelakaan menggunakan algoritma K-means yang dibentuk menjadi 3 cluster dan untuk akurasi menggunakan DBI (Davies Bouldin Index).



Gambar 5. Desain Pemodelan Data

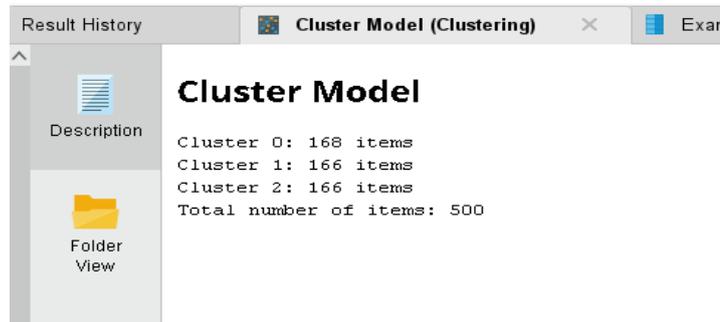
Selanjutnya pada Gambar 6 merupakan tampilan data keseluruhan hasil setelah dilakukan proses clustering.

Row No.	id	cluster	noregister_...	form_kode	jenis_kelam...	pasal_kode	kendaraan_...
1	1	cluster_0	1	1	1	1	1
2	2	cluster_0	2	1	1	2	1
3	3	cluster_0	3	1	2	2	1
4	4	cluster_0	4	1	1	2	1
5	5	cluster_0	5	2	2	2	1
6	6	cluster_0	6	2	1	2	1
7	7	cluster_0	7	2	1	2	1
8	8	cluster_0	8	2	1	2	1
9	9	cluster_0	9	2	2	2	1
10	10	cluster_0	10	2	1	2	1
11	11	cluster_0	11	2	1	2	1
12	12	cluster_0	12	2	1	2	1
13	13	cluster_0	13	2	2	2	1
14	14	cluster_0	14	2	2	2	1
15	15	cluster_0	15	2	2	2	1

ExampleSet (500 examples, 2 special attributes, 5 regular attributes)

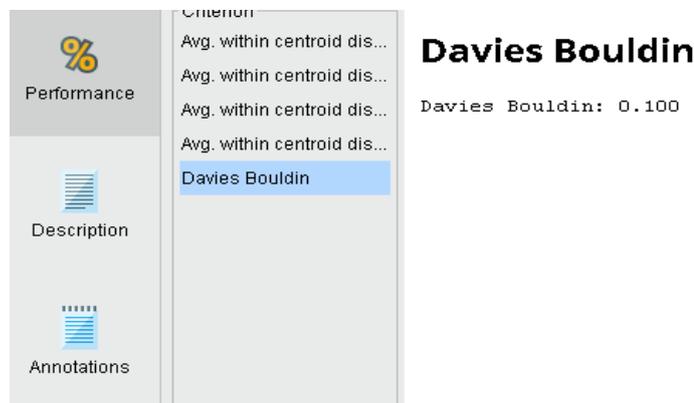
Gambar 6 Data Hasil *Clustering*

Berdasarkan pemodelan dari 3 *cluster* maka diperoleh data yang tergabung ke *cluster 0* sebanyak 168 data, *cluster1* sebanyak 166, dan *cluster2* sebanyak 166 data. Hasil akhir perhitungan manual sama dengan hasil menggunakan aplikasi *RapidMiner* seperti pada Gambar 7.



Gambar 7. *Cluster model*

Untuk menghitung seberapa tepat data yang dikelompokkan dalam perhitungan pada algoritma k-means, pengujian dilakukan dengan menggunakan rumus *DBI (Davies Bouldin Index)* di *Rapidminer* maka diperoleh 3 *cluster* yang dinilai dapat mengelompokkan data dengan baik. *PerformanceVector* hasil evaluasi *cluster* yang dibentuk sebanyak 3 *cluster* dengan nilai sebesar 0,100 yang dapat dilihat pada Gambar 8.



Gambar 8. *Performance Vector*

Pada penentuan *cluster* menggunakan *matrix Davies Bouldin Index* yaitu menggunakan data *centroid* dari iterasi akhir pada proses *clustering*.

IV. KESIMPULAN

Dari hasil penelitian yang dilakukan terhadap data register pelanggaran lalu lintas untuk pengelompokan pelanggaran yang sering terjadi di Kabupaten Slawi, maka dapat disimpulkan sebagai berikut :

1. Metode klusterisasi dengan menggunakan algoritma K-Means dapat diimplementasikan untuk mengelompokkan perkara pelanggaran lalu lintas yang ada di Pengadilan Negeri Slawi.
2. Setelah dilakukan pengelompokan data register pelanggaran lalu lintas berdasarkan no register, pasal, jenis kelamin, form, jenis kendaraan dengan 3 cluster dimana C1 terdiri dari 7 item dengan jumlah 166 pelanggar yaitu pada pasal 280 sebanyak 129, pasal 283 sebanyak 11, pasal 104(3) sebanyak 10, pasal 284 sebanyak 8, pasal 285 sebanyak 6 dan pasal 281, 287 masing-masing sebanyak 1. C2 terdiri dari 10 item dengan jumlah 166 pelanggar yaitu pada pasal 280 sebanyak 111, pasal 283 sebanyak 21, pasal 104(3) dan 284 masing-masing sebanyak 9, pasal 285 sebanyak 8, pasal 281, 287, dan 288(1) masing-masing sebanyak 2, pasal 287(2) dan 288(2) masing-masing sebanyak 1. C3 terdiri dari 8 item dengan jumlah 168 pelanggar yaitu pada pasal 280 sebanyak 123, pasal 283 sebanyak 16, pasal 104(3) sebanyak 14, pasal 284 sebanyak 8, pasal 285 sebanyak 3, pasal 281 sebanyak 2, pasal 288(2) dan 288(3) masing-masing sebanyak 1.
3. Berdasarkan validasi *Davies Bouldin* Index (DBI), dengan jumlah data sebanyak 500 terbentuk menjadi 3 *cluster* dengan nilai ketepatan sebesar 0,100.

VI. REFERENSI

- [1] B. P. Statistik, "Perkembangan Jumlah Kendaraan Bermotor Menurut Jenis, 1949-2018," 2020. .
- [2] D. Wahyuni *et al.*, "Analisa Clustering Pada Data Pelanggaran Lalulintas Di Pengadilan Negeri Dumai Dengan Menggunakan Metode K-Means," vol. 4, no. 2, pp. 1–12, 2019.
- [3] pengadilan negeri Slawi, "Statistik Perkara," 2017. .
- [4] E. D. S. Mulyani *et al.*, "Clustering Wilayah Dan Pelanggaran Berkendaraan Menggunakan Algoritma K-Means Pada Data Satlantas Polres Tasikmalaya Kota," *J. Sist. Inf. dan Teknol. Inf.*, vol. 8, no. 1, pp. 1–11, 2019.
- [5] N. Ramadhani, A. F. Rahman, D. Riskiyati, J. Raya, and T. Fax, "Analisis Cluster Data Register Perkara Lalu Lintas Menggunakan Algoritma K-Means," no. November, 2017.
- [6] I. Kurniawan and A. Susanto, "Implementasi Metode K-Means dan Naïve Bayes Classifier untuk Analisis Sentimen Pemilihan Presiden (Pilpres) 2019," *Eksplora Inform.*, vol. 9, no. 1, pp. 1–10, 2019, doi: 10.30864/eksplora.v9i1.237.
- [7] B. Jumadi Dehotman Sitompul, O. Salim Sitompul, and P. Sihombing, "Enhancement Clustering Evaluation Result of Davies-Bouldin Index with Determining Initial Centroid of K-Means Algorithm," *J. Phys. Conf. Ser.*, vol. 1235, no. 1, 2019, doi: 10.1088/1742-6596/1235/1/012015.
- [8] Undang-Undang Republik Indonesia Nomor 8 Tahun 1981 Tentang Kitab Undang Undang Hukum Acara Pidana (KUHAP), "Undang-Undang Republik Indonesia Nomor 8 Tahun 1981 Tentang Kitab Undang Undang Hukum Acara Pidana (KUHAP)," *Kpk*, vol. 1951, no. 8, 1981.
- [9] M. Riadi, "Pengertian, Fungsi, Proses dan Tahapan Data Mining," *Kajian Pustaka*, 2017. .
- [10] R. Hablum, A. Khairan, and R. Rosihan, "Clustering Hasil Tangkap Ikan Di Pelabuhan Perikanan Nusantara (Ppn) Ternate Menggunakan Algoritma K-Means," *JIKO (Jurnal Inform. dan Komputer)*, vol. 2, no. 1, pp. 26–33, 2019, doi: 10.33387/jiko.v2i1.1053.