

ANALISA CLUSTERING PADA DATA PELANGGARAN LALULINTAS DI PENGADILAN NEGERI DUMAI DENGAN MENGGUNAKAN METODE K-MEANS

Elisawati ⁽¹⁾, Deasy Wahyuni ⁽²⁾, Adi Arianto ⁽³⁾
Jurusan Sistem Informasi ^(1,3), Jurusan Teknik Informatika ⁽²⁾
^(1,2,3) Sekolah Tinggi Manajemen Informatika & Komputer (STMIK) Dumai
Jl. Utama Karya Bukit Batrem, Kota Dumai-Riau
e-mail : elisawati06@gmail.com ⁽¹⁾, deasywahyuni1@gmail.com ⁽²⁾,
adiarinato1981@gmail.com ⁽³⁾

Abstract

The order of traffic on the road is very important for motorists on the highway, the lack of awareness of motor vehicle users and the poor drivers of traffic discipline make the level of traffic violations in driving on the highway always increase so that the number of ticket data received by the Dumai District Court. This research was conducted to analyze and classify data violations using the k-means method to facilitate knowing the types of violations that are often violated by vehicle users. The attributes to be analyzed are the types of violations and types of vehicles. The test was carried out using the Rapidminer 5 application where the data tested was data from the Dumai District Court on December 2017, as many as 616 violations. Central cluster data consists of 3 clusters, namely C1 = Many, C2 = moderate and C3 = few who commit traffic violations. So the results of the data obtained where C1 produces 1 data, C2 gets as much as 4 data and C3 as many as 7 data. Where the type of violation that is often violated is the type of violation that does not use a helmet and the type of vehicle is a motorcycle. From the results of this study can be used or can be followed up with the holding of socialization to reduce the number of traffic violations.

Keywords: *Clustering Analysis, K-Means, Traffic Violations, Rapidminer*

Abstrak

Ketertiban berlalu lintas di jalan sangatlah penting bagi pengendara di jalan raya, kurangnya kesadaran pengguna kendaraan bermotor dan buruknya pengendara akan kedisiplinan berlalu lintas membuat tingkat pelanggaran berlalu lintas dalam berkendara di jalan raya selalu meningkat sehingga membuat banyaknya data tilang yang diterima oleh Pengadilan Negeri Dumai. Penelitian ini dilakukan untuk menganalisa dan mengelompokkan data pelanggaran dengan metode *k-means* guna mempermudah mengetahui Jenis pelanggaran yang sering di langgar oleh pengguna kendaraan. Adapun atribut yang akan di analisa adalah jenis pelanggaran dan jenis kendaraan. Pengujian di lakukan dengan menggunakan Aplikasi Rapidminer 5 dimana data yang di uji adalah data Pengadilan Negeri Dumai bulan Desember tahun 2017, sebanyak 616 pelanggaran. Data pusat cluster terdiri dari 3 cluster yaitu C1=Banyak, C2=sedang dan C3=sedikit yang melakukan pelanggaran lalu lintas. Sehingga hasil data yang di dapat di mana C1 menghasilkan 1 data, C2 mendapatkan sebanyak 4 data dan C3 sebanyak 7 data. Dimana jenis pelanggaran yang sering di langgar adalah jenis pelanggaran tidak menggunakan helm dan Jenis kendaraan adalah sepeda motor. Dari hasil penelitian ini dapat di pergunakan atau dapat ditindak lanjuti dengan diadakannya sosialisasi guna menekan angka pelanggaran berlalu lintas.

Kata Kunci : *Analisa Clustering, K-Means, Pelanggaran LaluLintas, Rapidminer*

1. PENDAHULUAN

Transportasi darat merupakan suatu kebutuhan yang tidak bisa di jauhan dari kalangan masyarakat baik perorangan maupun kelompok. Semakin meningkatnya masyarakat yang menggunakan kendaraan di mulai dari sepeda motor, mobil hingga truk, bertambah juga tingkat

pelanggaran lalu lintas dalam berkendara. Penyebab terjadinya pelanggaran lalu lintas sampai menyebabkan kecelakaan di karenakan pengemudi kendaraan yang buruk, yaitu masih rendahnya pengetahuan pengendara akan kedisiplinan berlalu lintas dan sedikitnya pemahaman para pemakai jalan terhadap peraturan.

Pengadilan Negeri Dumai merupakan pengadilan negeri yang memiliki jumlah pelanggaran lalu lintas yang tidak sedikit. Hal ini terlihat dari data pelanggaran lalu lintas pada tahun 2017 setiap harinya di pengadilan negeri dumai baik roda dua dan empat. Pada bulan Januari sebanyak 287 orang, bulan Februari 715 orang, bulan Maret 325 Orang, bulan April 434 Orang, bulan Mei 948 orang, bulan Juni 640, bulan Juli 588 orang, bulan Agustus 780 orang, bulan September 533 orang, bulan Oktober 657 orang, bulan November 883 orang dan bulan Desember 616 orang yang melakukan pelanggaran lalu lintas baik menggunakan sepeda motor maupun mobil. Data yang dilakukan analisa adalah data pelanggaran bulan Desember tahun 2017 sebanyak 616 orang.

Pengertian Analisis Cluster

Analisis cluster merupakan teknik multivariate yang mempunyai tujuan utama untuk mengelompokkan objek-objek berdasarkan karakteristik yang dimilikinya. Proses cluster atau pengelompokan data bisa dilakukan dengan dua metode yaitu metode hierarki dan metode non hierarki. (Sitepu & Gultom, 2011)

Pengertian Clustering

Clustering merupakan salah satu teknik analisis dalam data mining yang melakukan pengelompokan data berdasarkan kesamaan karakteristiknya. Dengan kesamaan karakteristik pada sebuah kelompok ini dapat diambil suatu pengetahuan yang memiliki arti dan berguna. (Ramadhani, Rahman, & Riskiyati, 2017). Tujuan cluster adalah agar objek-objek yang bergabung dalam sebuah cluster merupakan objek-objek yang mirip atau (berhubungan) satu sama lainnya dan berbeda atau tidak berhubungan lebih besar kemiripannya (homogenitas) dalam cluster dan lebih besar perbedaannya diantara cluster yang lainnya. (Ningrat, Maruddani, & Wuryandari, 2016)

Pengertian K-Means

Algoritma K-Means merupakan salah satu algoritma dengan partitional, karena K-Means didasarkan pada penentuan jumlah awal kelompok dengan mendefinisikan nilai centroid awalnya. Algoritma K-Means menggunakan proses secara berulang-ulang untuk mendapatkan basis data cluster. Dibutuhkan jumlah cluster awal yang diinginkan sebagai masukan dan menghasilkan jumlah cluster akhir sebagai output. Jika algoritma diperlukan untuk menghasilkan cluster K maka akan ada K awal dan K akhir. Metode K-Means akan memilih pola k sebagai titik awal centroid secara acak. Jumlah iterasi untuk mencapai cluster centroid akan dipengaruhi oleh calon cluster centroid awal secara random dimana jika posisi centroid baru tidak berubah. Nilai K yang dipilih menjadi pusat awal, akan dihitung dengan menggunakan rumus Euclidean Distance yaitu mencari jarak terdekat antara titik centroid dengan data/objek. Data yang memiliki jarak pendek atau terdekat dengan centroid akan membentuk sebuah cluster. (Putu et al., 2014)

Pengertian K-Means Cluster

K-means clustering merupakan salah satu metode data clustering non-hirarki yang mengelompokkan data dalam bentuk satu atau lebih cluster/kelompok. Data-data yang memiliki karakteristik yang sama dikelompokkan dalam satu cluster/kelompok dan data yang memiliki karakteristik yang berbeda dikelompokkan dengan cluster/kelompok yang lain sehingga data yang berada dalam satu cluster/kelompok memiliki tingkat variasi yang kecil. (Ong, 2013)

Pengertian Pelanggaran Lalu Lintas

Pelanggaran lalu lintas atau sering disebut tilang merupakan kasus hukum pidana yang diatur dalam UU No. 14 tahun 1992. Hukum pidana mengatur perbuatan yang dilarang oleh UU dan ditetapkannya hukuman bagi siapa yang melakukannya. Tujuan hukum pidana adalah untuk menakuti-nakuti orang agar tidak baik dan mendidik seseorang yang pernah melakukan perbuatan yang tidak baik menjadi baik dan dapat diterima.

Pelanggaran lalu lintas atau tilang yang sering pelanggaran terhadap pasal 54 mengenai SIM dan STNK serta pasal 59 mengenai muatan berlebihan truk angkutan kemudian pelanggaran pasal 61 memasuki jalur lalu lintas kendaraan. Persidangan kasus lalu lintas dalam proses tersebut para terdakwa pelanggaran ditempatkan di suatu ruangan. Kemudian hakim akan memanggil nama terdakwa untuk membacakan denda. Setelah denda dibacakan hakim akan menyetujui palu sebagai tanda keluarnya suatu putusan.

Bentuk-bentuk pelanggaran lalu lintas sebagai berikut :

1. Menggunakan jalan dengan cara melintangi membahayakan ketertiban keamanan dan menimbulkan kerusakan jalan.
2. Mengemudikan kendaraan bermotor yang tidak memperhatikan SIM, STNK dan Surat Tanda Uji Kendaraan (STNK) yang sah.
3. Membiarkan kendaraan bermotor dikemudikan oleh orang lain yang tidak memiliki SIM.
4. Tidak memenuhi ketentuan peraturan perundang-undangan lalu lintas.
5. Membiarkan kendaraan bermotor yang ada di jalan tanpa dilengkapi plat tanda nomor kendaraan yang sah.
6. Pelanggaran terhadap perintah yang diberikan petugas pengatur lalu lintas jalan, rambu-rambu atau tanda permukaan jalan.
7. Pelanggaran terhadap ketentuan tentang ukuran dan muatan yang diijinkan, cara menaik dan menurunkan penumpang.
8. Pelanggaran terhadap ijin trayek, jenis kendaraan yang diperbolehkan beroperasi di jalan yang ditentukan.

2. METODE PENELITIAN

Metode penelitian yang digunakan adalah metode K-Means di mana algoritma dari metode K-Means adalah sebagai berikut :

1. Menentukan k sebagai jumlah kluster yang ingin dibentuk
2. Membangkitkan nilai random untuk pusat cluster (centroid) awal sebanyak k.
3. Menghitung jarak setiap data terhadap masing-masing centroid menggunakan rumus *euclidian distance* hingga ditemukan jarak yang paling dekat dari setiap data dengan centroid. Persamaan yang digunakan untuk menghitung *euclidian distance* adalah :

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

Dimana :

x_i = data criteria

μ_j = Centroid Pada Cluster ke - j

4. Mengklasifikasi setiap data berdasarkan kedekatannya dengan centroid (jarak terkecil)

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

atau = MIN (Cij : Ckj : Clj)

5. Memperbaharui nilai centroid. Nilai centroid diperoleh dari rata-rata cluster yang bersangkutan.

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

6. Melakukan perulangan dari langkah 2 hingga langkah kelima sampai anggota cluster tidak ada yang berubah.
7. Jika langkah keenam telah terpenuhi, maka nilai pusat cluster pada iterasi terakhir akan digunakan sebagai parameter untuk menentukan klasifikasi data.

3. HASIL DAN PEMBAHASAN

Data yang akan di gunakan adalah data tilang Pengadilan Negeri Dumai bulan Desember tahun 2017 sebanyak 621 record. Untuk keperluan penelitian ini tidak semua atribut data yang akan di gunakan. Atribut data yang di pakai pada penelitian ini adalah Jenis pelanggaran (pasal) dan Jenis kendaraan.

Datasheet yang akan di cluster yaitu terkait dengan Jenis kendaraan terhadap pasal-pasal pelanggarannya. Pada tabel 1 disajikan hasil tranformasi data pada jenis kendaraan pelanggar berserta frekuensi kejadiannya.

Tabel 1. Transfomasi Jenis Kendaraan dan Frekuensi Pelanggarannya

No	Jenis Pelanggaran	Jenis Kendaraan					
		Sepeda Motor	Mobil Penumpang Pribadi	Mobil Penumpang Umum	Mobil Pickup	Truk	Lain - Lain
1	Tak dipasang Tanda Nomor Kendaraan (280)	8	2	1	0	1	2
2	Tidak memiliki SIM (281)	74	1	0	3	5	0
3	Kegiatan yang mengganggu konsentrasi saat berkendara (283)	0	1	0	0	0	0
4	Kendaraan bermotor tidak mengutamakan pejalan kaki atau pesepeda (284)	1	0	0	0	0	0
5	Pengemudi Sepeda Motor tidak memenuhi persyaratan teknis dan laik jalan (285)	57	0	0	0	0	0
6	Melanggar rambu lalu lintas (287)	37	16	0	1	23	4
7	Tidak memiliki STNK (288)	92	4	0	5	8	4
8	Setiap orang yang mengemudikan kendaraan bermotor atau Penumpang yang duduk disamping pengemudi yang tidak menggunakan sabuk keselamatan (289)	0	8	0	7	4	0
9	Tidak mengenakan helm standar nasional (291)	233	0	0	0	0	0
10	Kendaraan bermotor yang tidak menggunakan isyarat saat berhenti atau parkir	0	0	0	0	4	2

No	Jenis Pelanggaran	Jenis Kendaraan					
		Sepeda Motor	Mobil Penumpang Pribadi	Mobil Penumpang Umum	Mobil Pickup	Truk	Lain - Lain
	darurat (298)						
11	Pengemudi mobil barang untuk mengangkut orang (303)	0	0	0	1	1	0
12	Kendaraan bermotor yang mengangkut barang khusus yang tidak memenuhi persyaratan keselamatan (305)	1	0	0	0	1	4

setelah data yang di hasilkan dari tabel 1, maka langkah selanjutnya adalah melakukan analisa cluster dengan menggunakan metode k-means. Adapun langkah algoritma dari metode K-Means adalah :

- a. Langkah pertama menetapkan banyaknya jumlah cluster (C). pada penelitian ini di tetapkan jumlah cluster sebanyak 3 cluster terdiri C1 merupakan cluster pelanggaran yang terbanyak, C2 merupakan cluster pelanggaran yang sedang dan C3 merupakan Pelanggaran yang sedikit. Setelah menetapkan jumlah cluster kemudian menentukan titik pusat (*centroid*).

Tabel 2. Titik Pusat (*Centroid*)

C1	Tidak mengenakan helm standar nasional (291)	233	0	0	0	0	0
C2	Tidak memiliki SIM (281)	74	1	0	3	5	0
C3	Pengemudi mobil barang untuk mengangkut orang (303)	0	0	0	1	1	0

- b. Langkah Kedua Menghitung Jarak menggunakan Euclidian Distance.
 1. *Euclidian distance* dari semua data kesetiap titik pusat pertama :

Iterasi 1

$$1.1 = \sqrt{(8-233)^2 + (2-0)^2 + (1-0)^2 + (0-0)^2 + (1-0)^2 + (2-0)^2} = 225,0222$$

$$1.2 = \sqrt{(74-233)^2 + (1-0)^2 + (0-0)^2 + (3-0)^2 + (5-0)^2 + (5-0)^2} = 159,11$$

$$1.3 = \sqrt{(0-233)^2 + (1-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2} = 233,0021$$

$$1.4 = \sqrt{(1-233)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2} = 232$$

$$1.5 = \sqrt{(57-233)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2} = 176$$

$$1.6 = \sqrt{(37-233)^2 + (16-0)^2 + (0-0)^2 + (1-0)^2 + (23-0)^2 + (4-0)^2} = 1978,0354$$

$$1.7 = \sqrt{(92-233)^2 + (4-0)^2 + (0-0)^2 + (5-0)^2 + (8-0)^2 + (4-0)^2} = 141,4284$$

$$1.8 = \sqrt{(0-233)^2 + (8-0)^2 + (0-0)^2 + (7-0)^2 + (4-0)^2 + (0-0)^2} = 232,2767$$

$$1.9 = \sqrt{(233-233)^2 + (-0)^2 + (0-0)^2 + (0-0)^2 + (2-0)^2 + (0-0)^2} = 0$$

$$1.10 = \sqrt{(0-233)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (4-0)^2 + (2-0)^2} = 233,0429$$

$$1.11 = \sqrt{(0-233)^2 + (0-0)^2 + (0-0)^2 + (1-0)^2 + (1-0)^2 + (0-0)^2} = 232,0043$$

$$1.12 = \sqrt{(1-233)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (1-0)^2 + (4-0)^2} = 232,0366$$

2. *Euclidian distance* dari semua data kesetiap titik pusat kedua :

$$\begin{aligned}
2.1 &= \sqrt{(8-74)^2 + (2-1)^2 + (1-0)^2 + (0-3)^2 + (1-5)^2 + (2-0)^2} = 66,23443 \\
2.2 &= \sqrt{(74-74)^2 + (1-1)^2 + (0-0)^2 + (3-3)^2 + (5-5)^2 + (5-0)^2} = 0 \\
2.3 &= \sqrt{(0-74)^2 + (1-1)^2 + (0-0)^2 + (0-3)^2 + (0-5)^2 + (0-0)^2} = 74,22937 \\
2.4 &= \sqrt{(1-74)^2 + (0-1)^2 + (0-0)^2 + (0-3)^2 + (0-5)^2 + (0-0)^2} = 73,23933 \\
2.5 &= \sqrt{(57-74)^2 + (0-1)^2 + (0-0)^2 + (0-3)^2 + (0-5)^2 + (0-0)^2} = 18 \\
2.6 &= \sqrt{(37-74)^2 + (16-1)^2 + (0-0)^2 + (1-3)^2 + (23-5)^2 + (4-0)^2} = 44,02272 \\
2.7 &= \sqrt{(92-74)^2 + (4-1)^2 + (0-0)^2 + (5-3)^2 + (8-5)^2 + (4-0)^2} = 19,0263 \\
2.8 &= \sqrt{(0-74)^2 + (8-1)^2 + (0-0)^2 + (7-3)^2 + (4-5)^2 + (0-0)^2} = 73,44461 \\
2.9 &= \sqrt{(233-74)^2 + (0-1)^2 + (0-0)^2 + (0-3)^2 + (0-5)^2 + (0-0)^2} = 159,11 \\
2.10 &= \sqrt{(0-74)^2 + (0-1)^2 + (0-0)^2 + (0-3)^2 + (4-5)^2 + (2-0)^2} = 74,10128 \\
2.11 &= \sqrt{(0-74)^2 + (0-1)^2 + (0-0)^2 + (1-3)^2 + (1-5)^2 + (0-0)^2} = 74,14176 \\
2.12 &= \sqrt{(1-74)^2 + (0-1)^2 + (0-0)^2 + (0-3)^2 + (1-5)^2 + (4-0)^2} = 73,28711
\end{aligned}$$

3. Euclidian distance dari semua data kesetiap titik pusat ketiga :

$$\begin{aligned}
3.1 &= \sqrt{(8-0)^2 + (2-0)^2 + (1-0)^2 + (0-1)^2 + (1-1)^2 + (2-0)^2} = 8,602325 \\
3.2 &= \sqrt{(74-0)^2 + (1-0)^2 + (0-0)^2 + (3-1)^2 + (5-1)^2 + (5-0)^2} = 74,14176 \\
3.3 &= \sqrt{(0-0)^2 + (1-0)^2 + (0-0)^2 + (0-1)^2 + (0-1)^2 + (0-0)^2} = 1,732051 \\
3.4 &= \sqrt{(1-0)^2 + (0-0)^2 + (0-0)^2 + (0-1)^2 + (0-1)^2 + (0-0)^2} = 1,732051 \\
3.5 &= \sqrt{(57-0)^2 + (0-0)^2 + (0-0)^2 + (0-1)^2 + (0-1)^2 + (0-0)^2} = 57,01754 \\
3.6 &= \sqrt{(37-0)^2 + (16-0)^2 + (0-0)^2 + (1-1)^2 + (21-1)^2 + (4-0)^2} = 46,09772 \\
3.7 &= \sqrt{(92-0)^2 + (4-0)^2 + (0-0)^2 + (1-1)^2 + (8-1)^2 + (4-0)^2} = 92,52567 \\
3.8 &= \sqrt{(0-0)^2 + (8-0)^2 + (0-0)^2 + (7-1)^2 + (4-1)^2 + (0-0)^2} = 10,44031 \\
3.9 &= \sqrt{(233-0)^2 + (0-0)^2 + (0-0)^2 + (0-1)^2 + (0-1)^2 + (0-0)^2} = 233,0043 \\
3.10 &= \sqrt{(0-0)^2 + (0-0)^2 + (0-0)^2 + (0-1)^2 + (4-1)^2 + (2-0)^2} = 3,741657 \\
3.11 &= \sqrt{(0-0)^2 + (0-0)^2 + (0-0)^2 + (1-1)^2 + (1-1)^2 + (0-0)^2} = 0 \\
3.12 &= \sqrt{(1-0)^2 + (0-0)^2 + (0-0)^2 + (0-1)^2 + (1-1)^2 + (4-0)^2} = 4,242641
\end{aligned}$$

- c. Langkah ketiga melakukan Perhitungan pengelompokan data ke dalam cluster dengan jarak yang paling pendek. Dalam melakukan pengelompok data ke dalam cluster di cari nilai minimal dari masing-masing cluster sehingga menghasilkan jarak terpendek (JP), dapat di lihat pada tabel 3.

Tabel 3. Posisi data Cluster Pada Iterasi Pertama

No	Jenis Pelanggaran	C1	C2	C3	JP	C1	C2	C3
1	Tak dipasang Tanda Nomor Kendaraan (280)	225,0222	66,23443	8,602325	8,602325			1
2	Tidak memiliki SIM (281)	159,11	0	74,14176	0		1	
3	Kegiatan yang mengganggu konsentrasi saat berkendara (283)	233,0021	74,22937	1,732051	1,732051			1
4	Kendaraan bermotor tidak mengutamakan pejalan kaki atau pesepeda(284)	232	73,23933	1,732051	1,732051			1
5	Pengemudi Sepeda	176	18	57,01754	18		1	

No	Jenis Pelanggaran	C1	C2	C3	JP	C1	C2	C3
	Motor tidak memenuhi persyaratan teknis dan laik jalan (285)							
6	Melanggar rambu lalu lintas (287)	198,0354	44,02272	46,09772	44,02272		1	
7	Tidak memiliki STNK (288)	141,4284	19,0263	92,52567	19,0263		1	
8	Penumpang yang tidak menggunakan sabuk pengaman (289)	233,2767	74,44461	10,44031	10,44031			1
9	Tidak mengenakan helm standar nasional (291)	0	159,11	233,0043	0	1		
10	Kendaraan bermotor yang tidak menggunakan isyarat saat berhenti atau parkir darurat (298)	233,0429	74,10128	3,741657	3,741657			1
11	Pengemudi mobil barang untuk mengangkut orang (303)	233,0043	74,14176	0	0			1
12	Kendaraan bermotor yang mengangkut barang khusus yang tidak memenuhi persyaratan keselamatan (305)	232,0366	73,28711	4,242641	4,242641			1

Dari tabel 3 di atas terdapat jumlah anggota dari masing-masing cluster antara lain : pada cluster C1 berjumlah 1 item dengan data 9, C2 berjumlah 4 item dengan data 2,5,6,7 dan C3 berjumlah 7 item dengan data 1,3,4,8,10,11,12.

- d. Langkah keempat setelah selesai iterasi pertama maka langkah selanjutnya adalah menetapkan titik Pusat Cluster yang baru yang terdapat perhitungan di bawah ini:

Untuk C1 :

$$C_{kj} = \frac{\sum_{i=1}^p X_{ij}}{p} = \frac{\text{Jumlah dari cluster ke- 1 data ke- 1}}{\text{Banyak data dari cluster ke- 1}} = \frac{233}{1} = 233$$

$$C_{kj} = \frac{\sum_{i=1}^p X_{ij}}{p} = \frac{\text{Jumlah dari cluster ke- 1 data ke- 2}}{\text{Jumlah banyak data dari cluster ke- 1}} = \frac{0}{1} = 0$$

$$C_{kj} = \frac{\sum_{i=1}^p X_{ij}}{p} = \frac{\text{Jumlah dari cluster ke- 1 data ke- 3}}{\text{Banyak data dari cluster ke- 1}} = \frac{0}{1} = 0$$

$$C_{kj} = \frac{\sum_{i=1}^p X_{ij}}{p} = \frac{\text{Jumlah dari cluster ke- 1 data ke- 4}}{\text{Jumlah banyak data dari cluster ke- 1}} = \frac{0}{1} = 0$$

$$C_{kj} = \frac{\sum_{i=1}^p X_{ij}}{p} = \frac{\text{Jumlah dari cluster ke- 1 data ke- 5}}{\text{Banyak data dari cluster ke- 1}} = \frac{0}{1} = 0$$

$$C_{kj} = \frac{\sum_{i=1}^p X_{ij}}{p} = \frac{\text{Jumlah dari cluster ke- 1 data ke- 6}}{\text{Jumlah banyak data dari cluster ke- 1}} = \frac{0}{1} = 0$$

Untuk C2 :

$$C_{kj} = \frac{\sum_{i=1}^p X_{ij}}{p} = \frac{\text{Jumlah dari cluster ke-2 data ke-1}}{\text{Banyak data dari cluster ke-2}} = \frac{260}{4} = 65$$

$$C_{kj} = \frac{\sum_{i=1}^p X_{ij}}{p} = \frac{\text{Jumlah dari cluster ke-2 data ke-2}}{\text{Jumlah banyak data dari cluster ke-2}} = \frac{21}{4} = 5,25$$

$$C_{kj} = \frac{\sum_{i=1}^p X_{ij}}{p} = \frac{\text{Jumlah dari cluster ke-2 data ke-3}}{\text{Banyak data dari cluster ke-2}} = \frac{0}{4} = 0$$

$$C_{kj} = \frac{\sum_{i=1}^p X_{ij}}{p} = \frac{\text{Jumlah dari cluster ke-2 data ke-4}}{\text{Jumlah banyak data dari cluster ke-2}} = \frac{9}{4} = 2,25$$

$$C_{kj} = \frac{\sum_{i=1}^p X_{ij}}{p} = \frac{\text{Jumlah dari cluster ke-2 data ke-5}}{\text{Banyak data dari cluster ke-2}} = \frac{36}{4} = 9$$

$$C_{kj} = \frac{\sum_{i=1}^p X_{ij}}{p} = \frac{\text{Jumlah dari cluster ke-2 data ke-6}}{\text{Jumlah banyak data dari cluster ke-2}} = \frac{8}{4} = 2$$

Untuk C3 :

$$C_{kj} = \frac{\sum_{i=1}^p X_{ij}}{p} = \frac{\text{Jumlah dari cluster ke-3 data ke-1}}{\text{Banyak data dari cluster ke-3}} = \frac{11}{7} = 1,57143$$

$$C_{kj} = \frac{\sum_{i=1}^p X_{ij}}{p} = \frac{\text{Jumlah dari cluster ke-3 data ke-2}}{\text{Jumlah banyak data dari cluster ke-3}} = \frac{1}{7} = 0,14286$$

$$C_{kj} = \frac{\sum_{i=1}^p X_{ij}}{p} = \frac{\text{Jumlah dari cluster ke-3 data ke-3}}{\text{Banyak data dari cluster ke-3}} = \frac{8}{7} = 1,14286$$

$$C_{kj} = \frac{\sum_{i=1}^p X_{ij}}{p} = \frac{\text{Jumlah dari cluster ke-3 data ke-4}}{\text{Jumlah banyak data dari cluster ke-3}} = \frac{11}{7} = 1,57143$$

$$C_{kj} = \frac{\sum_{i=1}^p X_{ij}}{p} = \frac{\text{Jumlah dari cluster ke-3 data ke-5}}{\text{Banyak data dari cluster ke-3}} = \frac{8}{7} = 1,14286$$

$$C_{kj} = \frac{\sum_{i=1}^p X_{ij}}{p} = \frac{\text{Jumlah dari cluster ke-3 data ke-6}}{\text{Jumlah banyak data dari cluster ke-3}} = \frac{0}{7} = 0$$

Dari perhitungan di atas terdapat hasil perhitungan titik pusat cluster terbaru yang dapat di lihat pada tabel 4.

Tabel 4. Titik Pusat Cluster Baru

C1	233	0	0	0	0	0
C2	65	5,25	0	2,25	9	2
C3	1,57143	0,14286	1,14286	1,57143	1,14286	0

- e. Langkah kelima dapat Mengulang perhitungan pada Langkah Pertama, Kedua dan ketiga. Jika hasil dari perhitungan langkah ketiga menghasilkan posisi data cluster tidak

mengalami perubahan di iterasi pertama maka perhitungan berhenti. Tetapi jika posisi data cluster mengalami perubahan maka perhitungan di lanjutkan langkah Keempat.

Dari Hasil perhitungan langkah ketiga terdapat perhitungan akhir pada iterasi kedua karena tidak mengalami perubahan pada iterasi pertama sehingga perhitungan berhenti dan dapat dilihat di tabel 5.

Tabel 5. Posisi data Cluster Pada Iterasi Kedua

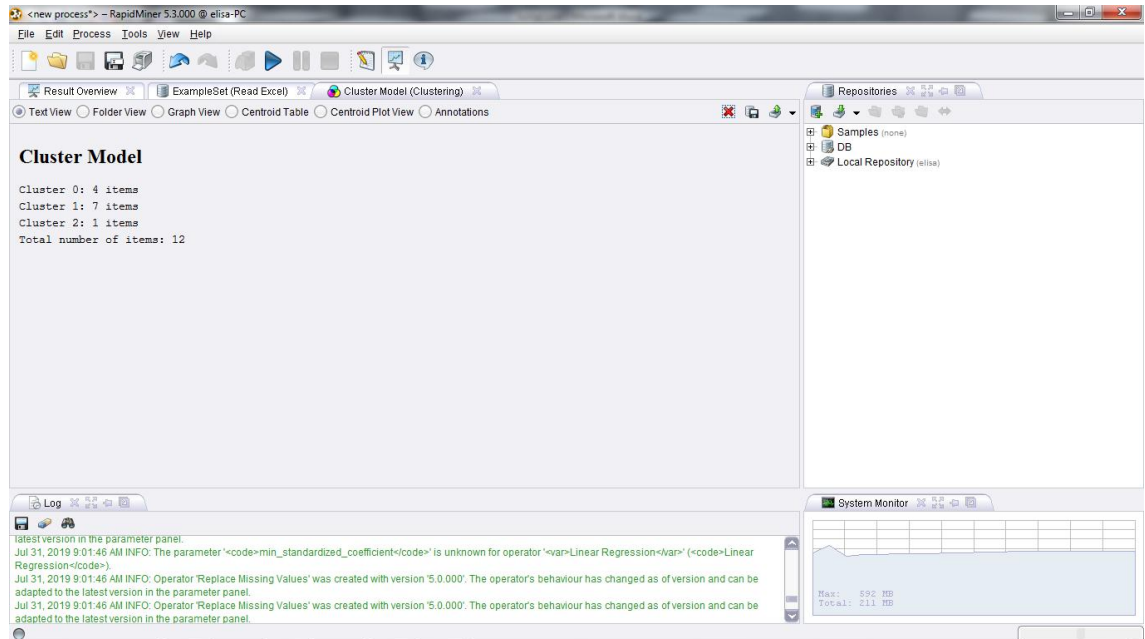
No	Jenis Pelanggaran	C1	C2	C3	JP	C1	C2	C3
1	Tak dipasang Tanda Nomor Kendaraan (280)	225,0222	57,7029	7,161404	7,161404			1
2	Tidak memiliki SIM (281)	159,11	10,93732	72,55934	10,93732		1	
3	Kegiatan yang mengganggu konsentrasi saat berkendara (283)	233,0021	65,82648	2,878492	2,878492			1
4	Kendaraan bermotor tidak mengutamakan pejalan kaki atau pesepeda(284)	232	64,91244	2,329929	2,329929			1
5	Pengemudi Sepeda Motor tidak memenuhi persyaratan teknis dan laik jalan (285)	176	13,47683	55,47458	13,47683		1	
6	Melanggar rambu lalu lintas (287)	198,0354	33,1832	44,74371	33,1832		1	
7	Tidak memiliki STNK (288)	141,4284	27,26032	90,93011	27,26032		1	
8	Penumpang yang tidak menggunakan sabuk pengaman (289)	233,2767	65,45323	10,15593	10,15593			1
9	Tidak mengenakan helm standar nasional (291)	0	168,3497	231,4396	0	1		
10	Kendaraan bermotor yang tidak menggunakan isyarat saat berhenti atau parkir darurat (298)	233,0429	65,44177	4,292851	4,292851			1
11	Pengemudi mobil barang untuk mengangkut orang (303)	233,0043	65,74287	2,035401	2,035401			1
12	Kendaraan bermotor yang mengangkut barang khusus yang tidak memenuhi persyaratan keselamatan (305)	232,0366	64,78136	4,488079	4,488079			1

Dari tabel 5 di atas terdapat jumlah anggota dari masing-masing cluster antara lain : pada cluster C1 berjumlah 1 item dengan data 9, C2 berjumlah 4 item dengan data 2,5,6,7 dan C3

berjumlah 7 item dengan data 1,3,4,8,10,11,12. Sesuai dengan posisi data cluster yang terdapat pada iterasi pertama.

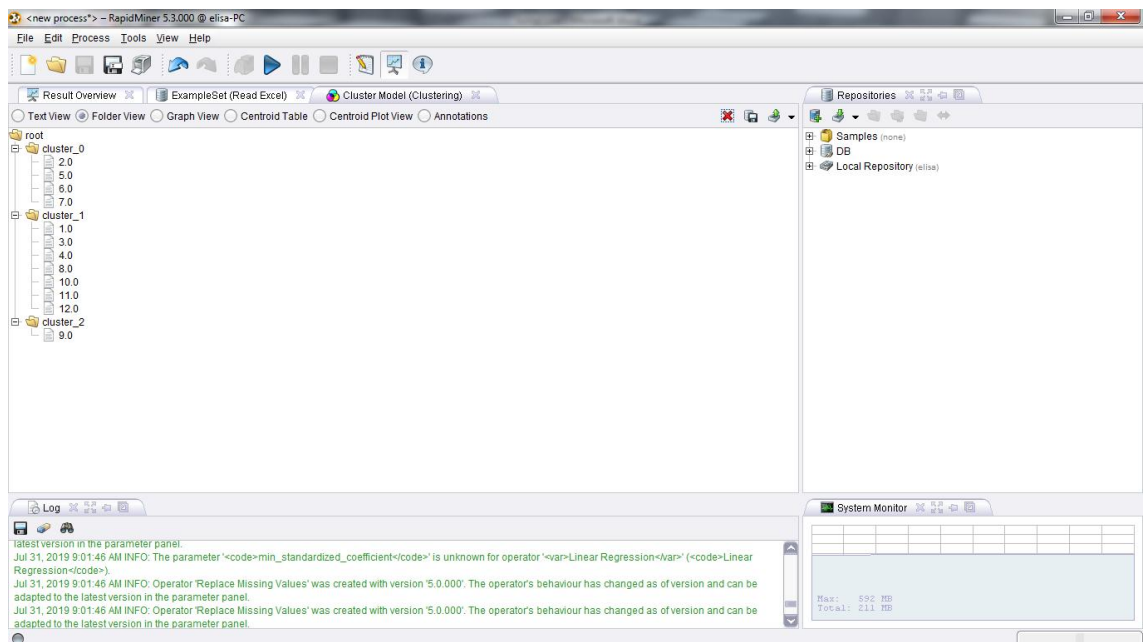
Pengujian Sistem dengan *RapidMiner*

Pengujian sistem di sini penulis menggunakan aplikasi *RapidMiner* 5.3. dimana data yang digunakan untuk pengujian adalah data bulan Desember tahun 2017.



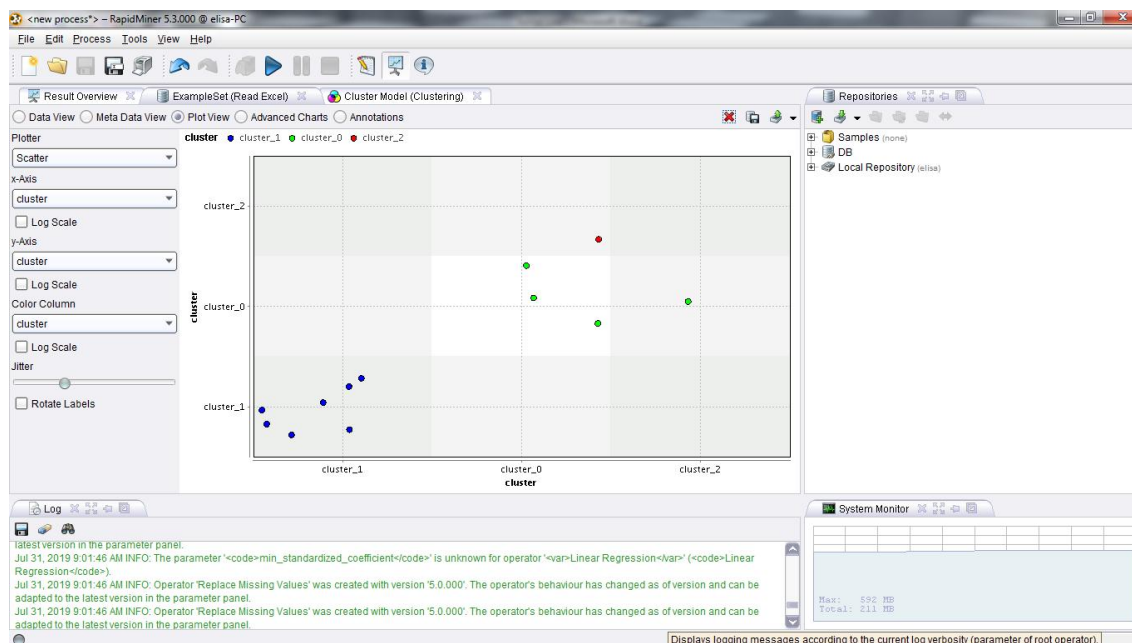
Gambar 1. Hasil Cluster dari RapidMiner

Hasil pengujian dengan menggunakan *RapidMiner* terdapat kesamaan dengan melakukan perhitungan manual dimana untuk cluster 0 terdapat 4 item, Cluster 1 terdapat 7 item dan cluster 2 terdapat 1 item.



Gambar 2. Pengelompokan Data Anggota di Tiap-Tiap cluster

Dari gambar 2 di atas memiliki kesamaan dengan perhitungan manual di mana untuk cluster 0 mempunyai 4 item dengan data 2,5,6,7, cluster 1 mempunyai 7 item dengan data 1,3,4,8,10,11,12, dan cluster 2 mempunyai 1 item dengan data 9.



Gambar 3. Pengelompokan data dengan Grafik

Pada gambar 3 adalah grafik yang dihasilkan dari pengujian dengan rapidminer di mana cluster 0 berjumlah 4 item berwarna hijau, cluster 1 berjumlah 7 item berwarna biru, dan cluster 2 berjumlah 1 item berwarna merah. Dimana cluster 2 adalah kelompok data yang jumlah pelanggaran yang paling tinggi sebanyak 233 dengan jenis pelanggaran tidak memakai helm standar nasional (291).

4. KESIMPULAN

1. Penggunaan metode clustering dengan algoritma K-Means dapat di implementasikan untuk membantu pengelompokan pelanggaran lalu lintas di pengadilan negeri Dumai.
2. Dari pengujian perhitungan manual data pelanggaran di bulan Desember tahun 2017 sebanyak 616 dengan tiga cluster dimana cluster C1 (Pelanggaran Terbanyak) dengan jumlah anggota data sebanyak 1 item yaitu pelanggaran pasal (291) tidak mengenakan helm standar nasional sebanyak 233 pelanggar, cluster 2 (Pelanggaran Sedang) dengan jumlah anggota data sebanyak 4 item yaitu Tidak memiliki SIM (281) sebanyak 83 pelanggar, Pengemudi Sepeda Motor tidak memenuhi persyaratan teknis dan laik jalan (285) sebanyak 57 pelanggar, Melanggar rambu lalu lintas (287) sebanyak 81 Pelanggar, Tidak memiliki STNK (288) sebanyak 113 pelanggar dan cluster 3 (pelanggaran sedikit) dengan jumlah anggota data sebanyak 7 item yaitu Tak dipasang Tanda Nomor Kendaraan (280) sebanyak 14 Pelanggar, Kegiatan yang mengganggu konsentrasi saat berkendara (283) sebanyak 1 pelanggar, Kendaraan bermotor tidak mengutamakan pejalan kaki atau pesepeda(284) sebanyak 1 pelanggar, Penumpang yang tidak menggunakan sabuk pengaman (289) sebanyak 19 pelanggar, Kendaraan bermotor yang tidak menggunakan isyarat saat berhenti atau parkir darurat (298) sebanyak 6 pelanggar, Pengemudi mobil barang untuk mengangkut orang (303)

sebanyak 2 pelanggar dan Kendaraan bermotor yang mengangkut barang khusus yang tidak memenuhi persyaratan keselamatan (305) sebanyak 6 pelanggar.

3. Pengujian menggunakan RapidMiner hasil pengelompokan datanya sama, hanya saja yang membedakan pada cluster di mana cluster 0 sebanyak 4 item, cluster 1 sebanyak 7 item dan cluster 2 sebanyak 1 cluster.

DAFTAR PUSTAKA

- Ningrat, D. R., Maruddani, D. A. I., & Wuryandari, T. (2016). Analisis Cluster Dengan Algoritma K-Means Dan Fuzzy C-Means Clustering Untuk Pengelompokan Data Obligasi Korporasi. *Jurnal Gaussian*, 5(4), 641–650.
- Ong, J. O. (2013). Implementasi Algoritma K-Means Clustering Untuk Menentukan Strategi Marketing, (April), 10–20.
- Putu, N., Merliana, E., Studi, P., Teknik, M., Industri, F. T., & Jaya, U. A. (2014). Analisa Penentuan Jumlah Kluster Terbaik Pada Metode K-Means Klustering. *Prosiding Seminar Nasional Multidisiplin Ilmu Dan Call For Paper Unisbank*, 978–979.
- Ramadhani, N., Rahman, A. F., & Riskiyati, D. (2017). Aplikasi Cluster Data Perkara Lalu Lintas Mingguan. *Link*, 26(2), 18–24.
- Sitepu, R., & Gultom, B. (2011). Clustering Analysis For Air Pollution Level On Industrial Sector In South Sumatera. *Jurnal Penelitian Sains*, 14(3), 11–17.
- Undang-Undang No. 14 Tahun 1992 Tentang Pelanggaran Lalu Lintas