

Clustering Wilayah Dan Pelanggaran Berkendaraan Menggunakan Algoritma K-Means Pada Data Satlantas Polres Tasikmalaya Kota

¹Evi Dewi Sri Mulyani, ²Agus Maulana Yusup, ³Alif Kurnia Tisna, ⁴Fahmi Akbar Fauzi, ⁵Irawan Bayu Seta, ⁶Rizky Khairunas, ⁷Wendi Ardiansyah

STMIK Tasikmalaya

Program Studi Teknik Informatika

Jl. RE Martadinata No. 272 A Telp/Fax (0265) 310830 Indihiang Tasikmalaya

Email: eviajadech@gmail.com¹, agusmy3@gmail.com², alifkurniatisna@gmail.com³, akbarfauzif@gmail.com⁴, irawanbayuseta@outlook.com⁵, khairunas299@gmail.com⁶, wendiardiansyah66@gmail.com⁷

Abstrak

Banyaknya pengguna jalan yang tidak mematuhi peraturan berlalu lintas dengan baik, setiap harinya dapat menambah tingkat kecelakaan dan pelanggaran tata tertib lalu lintas dalam berkendara pada wilayah Kota Tasikmalaya, sehingga masyarakat kurang dalam memahami ketertiban di jalan raya. Penelitian ini menerapkan data mining dengan menggunakan metode clustering pada data pelanggaran lalu lintas Polres Tasikmalaya Kota, algoritma yang digunakan yaitu K-Means clustering berupa proses pengelompokan sejumlah data atau objek ke dalam cluster atau group sehingga setiap dalam cluster tersebut akan berisi data yang semirip mungkin dan berbeda dengan objek dalam cluster lainnya. Data pelanggaran lalu lintas Polres Tasikmalaya kota ini diproses melalui Knowledge Data Discovery (KDD) sehingga dapat diketahui pengujian dengan rapidminer, menghasilkan cluster-cluster pelanggaran lalu lintas. Sampel yang digunakan di ambil dari tabel data pelanggaran lalu lintas yang telah ditrasformasikan. Dimana atribut yang ditentukan sebanyak 6 atribut yaitu wilayah, tidak menggunakan helm, sabuk keselamatan, melanggar rambu lintas, tidak membawa sim dan stnk dan kelebihan muatan. Dimana akan mempresentasikan cluster-cluster tiap kelompok wilayah dan jenis pelanggaran lalu lintas.

Kata Kunci : Data Pelanggaran Lalu Lintas, K-Means Clustering, Rapidminer

Abstract

The number of road users who do not comply with traffic regulations properly every day can increase the level of accidents and violations of traffic rules in driving in the Tasikmalaya City area, so that the community is lacking in understanding order on the highway. This research applies data mining by using clustering methods on data traffic violations in Tasikmalaya City Police, the algorithm used is K-Means clustering in the form of a process of grouping a number of data or objects into clusters or groups so that each cluster will contain data as closely as possible and different from objects in other clusters. The data on the Tasikmalaya Police Precinct traffic violations are processed through Knowledge Data Discovery (KDD) so that it can be known to test with rapidminer, resulting in clusters of traffic violations. The sample used is taken from the data table of traffic violations that have been transformed. Where the attributes specified are 6 attributes, namely region, do not use a helmet, safety belt, violate cross signs, do not carry sim and stnk and overload. Where will present the clusters of each regional group and type of traffic violations.

Keywords: Traffic Violation Data, K-Means Clustering, Rapidminer

1. PENDAHULUAN

Transportasi darat merupakan suatu kebutuhan yang tidak bisa dihindarkan dari kalangan masyarakat baik perorangan ataupun kelompok, untuk diketahui sejumlah kendaraan yang beredar dikalangan masyarakat kini dari tahun ke tahun semakin meningkat. Angka kemacetan dan terutama pelanggaran lalu lintas terjadi dimana-mana di daerah Kota Tasikmalaya terutama pada simpang-simpang kota. Banyaknya pengguna jalan yang tidak mematuhi peraturan berlalu lintas dengan baik, setiap harinya dapat menambah tingkat kecelakaan dan pelanggaran tata tertib lalu lintas dalam berkendara pada wilayah Kota Tasikmalaya.

Dikarenakan banyaknya tumpukan data pelanggaran lalu lintas di wilayah hukum Polres Tasikmalaya Kota, setiap tahunnya yang hanya data tersebut diproses ke dalam microsoft excell sehingga membuat data tidak beraturan. Data yang sudah masuk hanya diproses secara manual sehingga menyulitkan pihak kepolisian lalu lintas untuk mengetahui wilayah mana saja dan jenis pelanggaran berkendara yang sering melakukan pelanggaran. Maka dari itu di perlukan suatu pengelompokan pelanggaran berkendara dengan data-data pelanggaran berkendara tersebut yang akan menjadi variabel adalah wilayah dan jenis pelanggaran apa saja yang dilakukan.

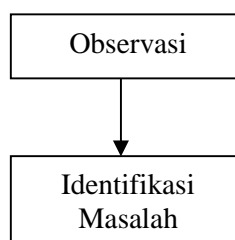
Algoritma k-means dapat mengkluster wilayah pada pelanggaran dalam berkendara berdasarkan banyak, sedang dan sedikit wilayah yang melakukan jenis pelanggaran dalam berkendara[1]. Mengambil data-data penting yang tersembunyi dan dikelola kembali menjadi informasi yang akan dikelompokkan berdasarkan kriteria atau variabel tersebut. Sehingga dapat dihasilkan seluruh pengetahuan hubungan terdekat data-data pelanggaran berkendara tersebut[2]. *Cluster* yang terbentuk menghasilkan pengetahuan terkait kasus perkara lalu lintas dan bisa digunakan sebagai kelas target pada teknik klasifikasi[3].

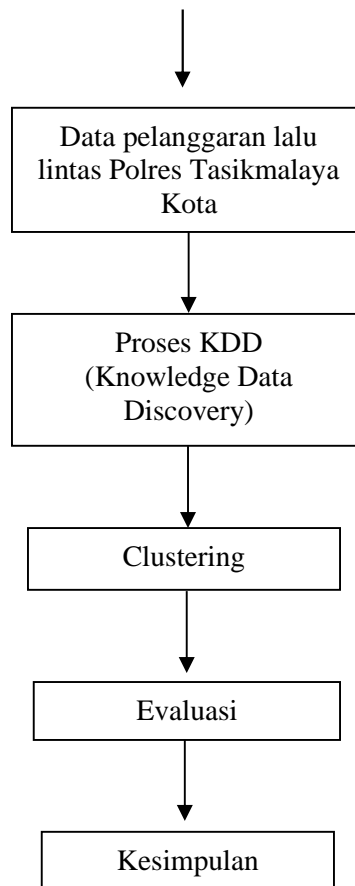
Penerapan data mining pada data pelanggaran lalu lintas untuk mengetahui pelanggaran dan daerah yang sering melanggar serta dilakukan analisa data tersebut dengan *clustering* dan algoritma k-means[4]. Analisis *cluster* berguna dalam meringkas data atau sejumlah variabel untuk menjadi lebih sedikit. Dalam melakukan proses data ini dapat dilakukan dengan mengelompokkan objek-objek berdasarkan kesamaan karakteristik tertentu diantara objek-objek yang hendak diteliti[5]. Pada penelitian sebelumnya dengan algoritma K-Means dapat menampilkan pengelompokkan kemampuan siswa pada mata pelajaran Ujian Nasional, juga dapat digunakan untuk memantau perkembangan kemampuan setelah mengikuti pengayaan materi[6].

Berdasarkan uraian diatas, maka penelitian ini akan mengolah data pelanggaran lalu lintas tahun 2018 dengan menggunakan algoritma k-means. Dengan begitu tools atau implementasi hasil data pelanggaran lalu lintas dengan Rapidminer, sehingga dapat diketahui wilayah mana saja yang sering melanggar lalu lintas dan jenis pelanggaran apa saja yang dilanggar oleh para pengemudi di daerah Kota Tasikmalaya.

2. METODE PENELITIAN

Metode yang digunakan dalam penelitian ini adalah metode deskriptif analisis dengan pendekatan kuantitatif, yaitu penelitian yang kemudian diolah dan dianalisis untuk diambil kesimpulan. Berikut ini adalah kerangka penelitian clustering wilayah dan pelanggaran berkendara menggunakan algoritma k-means pada data Satlantas Polres Tasikmalaya Kota :





Gambar 1.Kerangka penelitian

K-means clustering merupakan salah satu metode *cluster analysis* non hirarki yang berusaha untuk mempartisi objek yang ada kedalam satu atau lebih *cluster* atau kelompok objek berdasarkan karakteristiknya, sehingga objek yang mempunyai karakteristik yang sama dikelompokkan dalam satu *cluster* yang sama dan objek yang mempunyai karakteristik yang berbeda dikelompokkan kedalam cluster yang lain [7]. Jarak antara data dan pusat cluster dihitung menggunakan Euclidian Distance. Untuk menghitung jarak semua data ke setiap titik pusat cluster dapat menggunakan teori jarak Euclidean yang dirumuskan sebagai berikut:

$$D(i,j) = \sqrt{(X_{1i} - X_{1j})^2 + (X_{2i} - X_{2j})^2 + \dots + (X_{ki} - X_{kj})^2}$$

dimana:

D(i,j) = Jarak data ke i ke pusat cluster j

X_{ki} = Data ke i pada atribut data ke k

X_{kj} = Titik pusat ke j pada atribut ke k

3. HASIL DAN PEMBAHASAN

3.1 Sumber Data

Data yang digunakan dalam penelitian ini adalah dataset pelanggaran lalu lintas Tahun 2018 pada data Satlantas Polres Tasikmalaya Kota. Data yang digunakan berjumlah 343 yang akan diproses melalui tahapan KDD (*Knowledge Data Discovery*). Berikut ini adalah dataset dari 343 pelanggaran lalu lintas Tahun 2018 :

DATA PELANGGARAN LALU LINTAS POLRES TASIKMALAYA KOTA												
TAHUN 2018												
No	Tanggal	Waktu	Wilayah	Tidak Menggunakan Helm	Sabuk Keselamatan	Melanggar Rambu Lintas	Tidak Membawa SIM - STNK	Kelebihan Muatan	Total Pelanggaran	Roda 2	Roda 4/Lebih	Total Kendaraan
1	07/01/2018	Pagi	Karawang	1	1	7	1	0	10	5	7	10
2	07/01/2018	Pagi	Mida Sakik	0	2	0	14	0	16	15	1	16
3	07/01/2018	Pagi	Cimulu	0	4	0	5	0	9	7	2	9
4	07/01/2018	Sore	Cibogor Arah Suana Sonjaya	19	4	6	17	3	59	45	8	53
5	07/01/2018	Sore	Dow Selekta	0	6	4	10	0	20	16	4	20
...
343	20/07/2018	Sore	Karawang	5	22	0	16	0	43	36	8	44

Tabel 1. Dataset Pelanggaran Lalu Lintas

3.2 Pemilihan Data

Pada tahap ini dataset pelanggaran lalu lintas tahun 2018 melalui pemilihan data yang semua dataset pelanggaran lalu lintas berjumlah 343 akan dihilangkan melalui data wilayah yang sama dan total jenis pelanggaran menurut wilayah yang sama serta tidak terdapat pada data pelanggaran lalu lintas tahun 2018 yang tidak valid atau *missvalue* pada proses pemilihan data.

3.3 Transformasi Data

Transformasi data dilakukan dilakukan untuk mengubah data tujuannya adalah agar data dapat diolah dengan menggunakan metode K-means Clustering. Adapun atribut yang di gunakan pada data pelanggaran lalu lintas yaitu wilayah, tidak menggunakan helm, sabuk keselamatan, melanggar rambu lintas, tidak membawa SIM+STNK, dan kelebihan muatan. Dalam menstransformasikan data pelanggaran lalu lintas digabungkan berdasarkan wilayah yang sama, dan total jenis pelanggaran menurut wilayah yang sama, sehingga jumlah dataset berjumlah 24. Berikut ini adalah sampel 5 dataset dari 24 pelanggaran lalu lintas Tahun 2018 setelah di transformasikan :

Wilayah	Tidak Menggunakan Hel	Sabuk Keselamatan	Melanggar Rambu Lintz	Tidak Membawa SIM + STN	Kelebihan Muatan
Baypass	8	112	2	4	0
Bojong Jengkol Arah Leuwidahu	6	45	9	27	0
Bypass	0	0	1	0	0
Cibogor Arah Ciamis	0	32	4	84	0
Cibogor Arah M.Hatta	62	153	211	174	4

Tabel 2. Penggabungan Dataset Pelanggaran Lalu Lintas

3.3 Pengolahan Data

Pengolahan data pelanggaran lalu lintas dilakukan setelah proses transformasi sehingga data pelanggaran lalu lintas baru bisa diolah menggunakan metode clustering dengan algoritma k-means. Adapun langkah –langkah perhitungan manual adalah sebagai berikut :

a. Menentukan jumlah *cluster*

Diketahui jumlah *cluster* adalah 3 dikarenakan untuk mengetahui wilayah mana saja yang sering banyak sering terjadi pelanggaran lalu lintas, wilayah yang paling jarang terjadi dan jarang terjadi pelanggaran lalu lintas. Jumlah data adalah 24 dikarenakan telah dilakukan proses transformasi data dengan melalui penggabungan menurut wilayah yang sama dan total jenis pelanggaran wilayah yang sama menjadi 24 dataset. Jumlah atribut yaitu 6, sama halnya seperti jumlah data yang telah ditransformasikan.

No	Wilayah	Tidak Menggun akan Helm	Sabuk Kesemala tan	Melangg ar Rambu Lintas	Tidak Membaw a SIM + STNK	Kelebihan Muatan
1	Baypass	8	112	2	4	0
2	Bojong Jengkol Arah Leuwidahu	6	45	9	27	0
3	Bypass	0	0	1	0	0
4	Cibogor Arah Ciamis	0	32	4	84	0
5	Cibogor Arah M.Hatta	62	153	211	174	4
6	Cibogor Arah Sutisna Senjaya	39	6	66	40	3
7	Cilembang arah cikurubuk	0	20	4	35	2
8	Cimulu	6	38	8	29	6
9	Dewi Sartika	34	93	56	57	0
10	Dewi Sartika Arah Masjid Agung	0	162	29	39	0
11	Jati	10	17	21	8	3
12	Leuwidahu Arah Bojong Jengkol	14	192	4	207	1
13	Leuwidahu Arah Jati	3	135	247	128	3
14	Masjid Agung	72	87	52	107	6
15	Mitra Batik	14	204	120	104	0
16	Rancabango	11	139	85	208	1
17	Simpang Lima	39	93	18	48	4
18	Sutisna senjaya arah pancasila	0	3	6	0	0
19	Taman Alun-Alun	25	173	31	178	13
20	Taman Kota	28	16	34	65	22
21	Wasita Kusuma	58	195	71	270	1
22	Wasita Kusuma Arah Bandung	28	212	115	305	3
23	Wasita Kusuma Arah Bojong Jengkol	2	51	150	169	3
24	Wasita Kusuma Arah Letnan Harun	16	95	117	229	0

Tabel 3. Pengabungan Dataset Pelanggaran Lalu Lintas

b. Menentukan awal pusat cluster

Wilayah	Tidak Menggun akan Helm	Sabuk Kesemala tan	Melangg ar Rambu Lintas	Tidak Membaw a SIM + STNK	Kelebihan Muatan
Iterasi 1					
Bojong Jengkol Arah Leuwidahu	6	45	9	27	0
Leuwidahu Arah Bojong Jengkol	14	192	4	207	1
Wasita Kusuma	58	195	71	270	1

c. Perhitungan Jarak Pusat Cluster

Untuk mengukur jarak antara dengan pusat cluster digunakan Euclidian Distance dengan rumus :

$$D(i,j) = \sqrt{(X_{1i} - X_{1j})^2 + (X_{2i} - X_{2j})^2 + \dots + (X_{ki} - X_{kj})^2} \quad [7]$$

Sehingga didapatkan C1, C2 dan C3 :

						c1	c2	c3	
1	Baypass	8	112	2	4	0	71,21095	218,2888	291,3881
2	Bojong Jengkol Arah Leuwidahu	6	45	9	27	0	0	232,5919	296,8131
3	Bypass	0	0	1	0	0	53,42284	282,6995	345,239
4	Cibogor Arah Ciamis	0	32	4	84	0	58,98305	202,3018	262,7147
5	Cibogor Arah M.Hatta	62	153	211	174	4	277,9011	218,5681	174,9428
6	Cibogor Arah Sutisna Senjaya	39	6	66	40	3	77,69813	258,7624	298,3471
7	Cilembang arah cikurubuk	0	20	4	35	2	27,45906	243,6493	306,1111
8	Cimulu	6	38	8	29	6	9,486833	235,595	299,0451
9	Dewi Sartika	34	93	56	57	0	78,72103	188,1648	237,855
10	Dewi Sartika Arah Masjid Agung	0	162	29	39	0	119,4529	173,0491	244,0881
11	Jati	10	17	21	8	3	36,24914	265,5843	324,2468
12	Leuwidahu Arah Bojong Jengkol	14	192	4	207	1	232,5919	0	101,9951
13	Leuwidahu Arah Jati	3	135	247	128	3	273,7937	262,0382	240,3518
14	Masjid Agung	72	87	52	107	6	120,0208	163,4564	197,0152
15	Mitra Batik	14	204	120	104	0	208,7942	155,5956	178,8155
16	Rancabango	11	139	85	208	1	217,7131	96,8504	96,87621
17	Simpang Lima	39	93	18	48	4	62,69769	189,5046	250,7329
18	Sutisna senjaya arah pancasila	0	3	6	0	0	50,37857	280,6617	342,5697
19	Taman Alun-Alun	25	173	31	178	13	200,4969	46,8615	108,5403
20	Taman Kota	28	16	34	65	22	62,27359	229,5147	277,0848
21	Wasita Kusuma	58	195	71	270	1	296,8131	101,9951	0
22	Wasita Kusuma Arah Bandung	28	212	115	305	3	341,9093	150,0833	65,98485
23	Wasita Kusuma Arah Bojong Jengkol	2	51	150	169	3	200,2648	206,855	200,7934
24	Wasita Kusuma Arah Letnan Harun	16	95	117	229	0	234,6657	150,5556	124,7477

d. Pengelompokan Data

Jarak hasil perhitungan akan dilakukan perbandingan dan dipih jarak terdekat antara data dengan pusat cluster, jarak ini menunjukkan bahwa data tersebut berada dalam satu kelompok dengan pusat cluster terdekat. Berikut ini data matriks pengelompokan data, nilai 1 berarti data tersebut berada dalam kelompok data :

no	C1	C2	C3
1	1		
2	1		
3	1		
4	1		
5			1
6	1		
7	1		
8	1		
9	1		
10	1		
11	1		
12		1	
13			1
14	1		
15		1	
16		1	
17	1		
18	1		
19		1	
20	1		
21			1
22			1
23	1		
24			1
Total Anggota	15	4	5

Kelompok Data Ke-1

Selanjutnya pada iterasi 2 dan seterusnya dengan menghitung pusat *cluster* baru dan mengelompokan data seperti iterasi 1, sehingga didapatkan nilai pusat setiap *cluster* sebagai berikut :

	Nama Pusat Cluster	Tidak Menggunakan Helm	Sabuk Kesemalaman	Melanggar Rambu Lintas	Tidak Membawa SIM + STNK	Kelebihan Muatan
Iterasi 1	Bojong Jengkol Arah Leuwidahu	6	45	9	27	0
	Leuwidahu Arah Bojong Jengkol	14	192	4	207	1
	Wasita Kusuma	58	195	71	270	1
Iterasi 2	C1	16,26667	51,66667	30,66667	47,46667	3,266667
	C2	16	177	60	174,25	3,75
	C3	33,4	158	152,2	221,2	2,2
Iterasi 3	C1	17,28571	51,71429	22,14286	38,78571	3,285714
	C2	16	177	60	174,25	3,75
	C3	28,16667	140,1667	151,8333	212,5	2,333333
Iterasi 4	C1	17,28571	51,71429	22,14286	38,78571	3,285714
	C2	24,4	180,6	62,2	193,4	3,2
	C3	22,2	129,2	168	201	2,6
Iterasi 5	C1	17,28571	51,71429	22,14286	38,78571	3,285714
	C2	25	185,8333	71	212	3,166667
	C3	20,75	108,5	181,25	175	2,5

Setelah didapatkan nilai pusat *cluster* baru serta pengelompokan data, sampai data sudah tidak ada lagi yang berpindah ke cluster 1, 2 dan 3. Sehingga proses iterasi pun berhenti pada iterasi ke 5 sebagai berikut :

a. Menentukan pusat cluster baru

Wilayah	Tidak Menggunakan Helm	Sabuk Kesemalaman	Melanggar Rambu Lintas	Tidak Membawa SIM + STNK	Kelebihan Muatan
Iterasi 5					
	1	2	3	4	5
C1	17,28571	51,71429	22,14286	38,78571	3,285714
C2	25	185,8333	71	212	3,166667
C3	20,75	108,5	181,25	175	2,5

b. Perhitungan jarak pusat cluster baru

						C1	C2	C3	
1	Baypass	8	112	2	4	0	73,12433	231,8952	248,098
2	Bojong Jengkol Arah Leuwidahu	6	45	9	27	0	22,24596	241,4002	236,284
3	Bypass	0	0	1	0	0	70,25177	291,57	274,4524
4	Cibogor Arah Ciamis	0	32	4	84	0	55,42328	212,5388	214,4473
5	Cibogor Arah M.Hatta	62	153	211	174	4	257,3103	153,2701	67,6027
6	Cibogor Arah Sutisna Senjaya	39	6	66	40	3	66,97993	249,2891	205,7842
7	Cilembang arah cikurubuk	0	20	4	35	2	40,61699	252,8736	243,4761
8	Cimulu	6	38	8	29	6	24,87161	244,2984	237,7638
9	Dewi Sartika	34	93	56	57	0	58,92996	181,5463	173,3021
10	Dewi Sartika Arah Masjid Agung	0	162	29	39	0	111,891	181,3727	212,0734
11	Jati	10	17	21	8	3	46,98203	269,8995	249,1127
12	Leuwidahu Arah Bojong Jengkol	14	192	4	207	1	219,8211	68,39388	198,6495
13	Leuwidahu Arah Jati	3	135	247	128	3	256,2432	202,7315	86,88858
14	Masjid Agung	72	87	52	107	6	98,94817	152,8759	156,303
15	Mitra Batik	14	204	120	104	0	192,4616	120,5241	134,0322
16	Rancabango	11	139	85	208	1	200,6191	51,04954	106,68
17	Simpang Lima	39	93	18	48	4	47,73465	196,2644	208,2189
18	Sutisna senjaya arah pancasila	0	3	6	0	0	66,69031	288,4997	270,0086
19	Taman Alun-Alun	25	173	31	178	13	185,2651	54,93076	163,9287
20	Taman Kota	28	16	34	65	22	50,67851	228,4405	206,813
21	Wasita Kusuma	58	195	71	270	1	279,3577	67,3923	173,3555
22	Wasita Kusuma Arah Bandung	28	212	115	305	3	324,4979	106,2013	179,0367
23	Wasita Kusuma Arah Bojong Jengkol	2	51	150	169	3	183,1323	163,7042	68,34197
24	Wasita Kusuma Arah Letnan Harun	16	95	117	229	0	216,9457	103,6664	85,17702

c. Pengelompokan Data

Jarak hasil perhitungan akan dilakukan perbandingan dan dipih jarak terdekat antara data dengan pusat cluster, jarak ini menunjukkan bahwa data tersebut berada dalam satu kelompok dengan pusat cluster terdekat. Berikut ini data matriks pengelompokan data, nilai 1 berarti data tersebut berada dalam kelompok data yang tetap berwarna kuning :

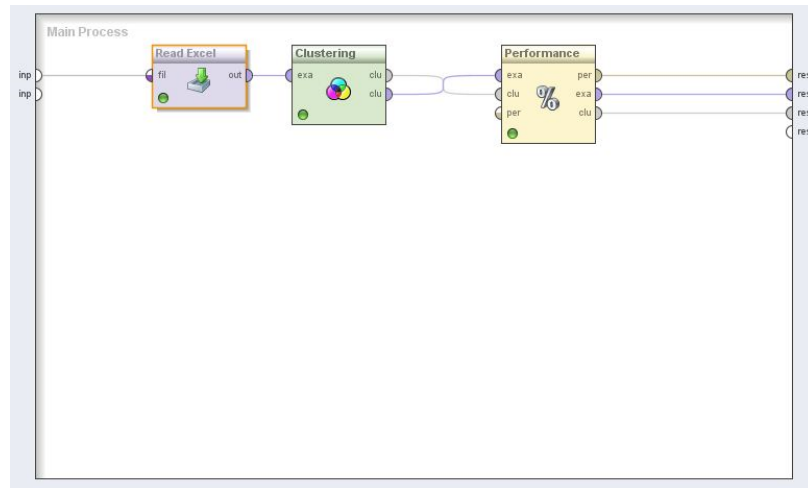
	C1	C2	C3
1	1		
2	1		
3	1		
4	1		
5			1
6	1		
7	1		
8	1		
9	1		
10	1		
11	1		
12		1	
13			1
14	1		
15		1	
16		1	
17	1		
18	1		
19		1	
20	1		
21		1	
22		1	
23			1
24			1
Total Anggota	14	6	4

Kelompok Data Ke-5

Dalam penelitian ini proses iterasi berhenti pada iterasi ke 5 karena kondisi *cluster* sudah mencapai konvergen dan proses iterasi pun berhenti. Sehingga pada iterasi ke 5, titik pusat dari setiap *cluster* sudah tidak berubah dan tidak ada lagi data yang berpindah dari satu *cluster* ke *cluster* yang lain.

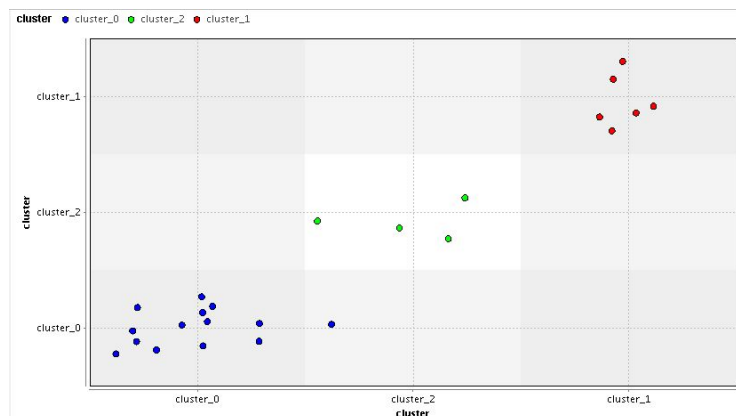
3.4 Implementasi Data Pelanggaran Lalu Lintas Ke Rapidminer

Adapun pengolahan data pelanggaran lalu lintas dengan menggunakan algoritma k-means clustering dan software Rapidminer dapat dilihat pada gambar berikut :



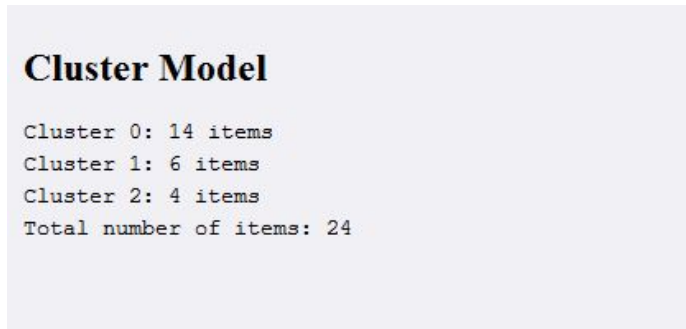
Gambar 2. Pemodelan K-means Pada Rapidminer

Dengan menggunakan pemodelan k-means clustering pada gambar 2, maka dengan jumlah data 24 dan inisialisasi jumlah cluster sebanyak 3 buah, sesuai dengan pendefinisian nilai K dengan jumlah *cluster 0* : 14 *items*, *cluster 1* : 6 *items*, dan *cluster 2*: *items*.



Gambar 3. *Exampleshet* (Read Excell) Flot view

Merupakan sheet untuk menampilkan database yang telah diolah secara keseluruhan lengkap dengan *clusternya*, urutan *clusternya* yaitu *cluster_0* terdiri dari 14 titik yang berwarna biru, dengan kelompok 1, 2, 3, 4, 6, 7, 8, 9, 10, 11, 14, 17, 18 dan 20 yang terdiri dari wilayah Bypass, Bojong Jengkol, Bypass, Cibogor Arah Ciamis, Cibogor Arah M.Hatta, Cibogor Arah Sutisna Senjaya, Cilembang Arah Cikurubuk, Cimulu, Dewi Sartika, Dewi Sartika Arah Masjid Agung, Jati, Masjid Agung, Simpang Lima, Sutisna Senjaya Arah Pancasila serta Taman Dadaha. Sedangkan *cluster_1* terdiri dari 6 titik berwarna merah dengan kelompok 12, 16, 19, 21, 22 dan 24 yang terdiri dari wilayah Leuwidahu Arah Bojong Jengkol, Rancabango, Taman Alun-Alun, Wasita Kusuma, Wasita Kusuma Arah Bandung, Wasita Kusuma Arah Letnan Harun dan terakhir *cluster_2* terdiri dari 4 titik berwarna hijau muda dengan kelompok 5,13,15 dan 23 yang terdiri dari wilayah Cibogor Arah M.Hatta, Leuwidahu Arah Jati, Mitra Batik, Wasita Kusuma Arah Bojong Jengkol.



Gambar 4. Cluster Model (Clustering)

Merupakan sheet untuk menampilkan database yang telah diolah secara keseluruhan dengan Rapidminer, algoritma k-means clustering dengan hasil yang didapatkan yaitu cluster_0 memiliki 14 items kelompok 1, 2, 3, 4, 6, 7, 8, 9, 10, 11, 14, 17, 18 dan 20 yang terdiri dari wilayah Bypass, Bojong Jengkol, Bypass, Cibogor Arah Ciamis, Cibogor Arah M.Hatta, Cibogor Arah Sutisna Senjaya, Cilembang Arah Cikurubuk, Cimulu, Dewi Sartika, Dewi Sartika Arah Masjid Agung, Jati, Masjid Agung, Simpang Lima, Sutisna Senjaya Arah Pancasila serta Taman Dadaha. Sedangkan cluster_1 memiliki 6 items kelompok 12, 16, 19, 21, 22 dan 24 yang terdiri dari wilayah Leuwidahu Arah Bojong Jengkol, Rancabango, Taman Alun-Alun, Wasita Kusuma, Wasita Kusuma Arah Bandung, Wasita Kusuma Arah Letnan Harun serta yang terakhir cluster_2 memiliki 4 items kelompok 5,13,15 dan 23 terdiri dari wilayah Cibogor Arah M.Hatta, Leuwidahu Arah Jati, Mitra Batik, Wasita Kusuma Arah Bojong Jengkol, dengan keseluruhan berjumlah 24 items.

Row No.	Wilayah	id	cluster	Tidak Meng...	Sabuk Kes...	Melanggar ...	Tidak Mem...	Kelebihan ...
1	Baypass	1	cluster_0	8	112	2	4	0
2	Bojong Jengkol Arah Leuwidahu	2	cluster_0	6	45	9	27	0
3	Bypass	3	cluster_0	0	0	1	0	0
4	Cibogor Arah Ciamis	4	cluster_0	0	32	4	84	0
5	Cibogor Arah M.Hatta	5	cluster_2	62	153	211	174	4
6	Cibogor Arah Sutisna Senjaya	6	cluster_0	39	6	66	40	3
7	Cilembang arah cikurubuk	7	cluster_0	0	20	4	35	2
8	Cimulu	8	cluster_0	6	38	8	29	6
9	Dewi Sartika	9	cluster_0	34	93	56	57	0
10	Dewi Sartika Arah Masjid Agung	10	cluster_0	0	162	29	39	0
11	Jati	11	cluster_0	10	17	21	8	3
12	Leuwidahu Arah Bojong Jengkol	12	cluster_1	14	192	4	207	1
13	Leuwidahu Arah Jati	13	cluster_2	3	135	247	128	3
14	Masjid Agung	14	cluster_0	72	87	52	107	6
15	Mitra Batik	15	cluster_2	14	204	120	104	0
16	Rancabango	16	cluster_1	11	139	85	208	1
17	Simpang Lima	17	cluster_0	39	93	18	48	4
18	Sutisna senjaya arah pancasila	18	cluster_0	0	3	6	0	0
19	Taman Alun-Alun	19	cluster_1	25	173	31	178	13
20	Taman Kota	20	cluster_0	28	16	34	65	22
21	Wasita Kusuma	21	cluster_1	58	195	71	270	1
22	Wasita Kusuma Arah Bandung	22	cluster_1	28	212	115	305	3
23	Wasita Kusuma Arah Bojong Jengkol	23	cluster_2	2	51	150	169	3
24	Wasita Kusuma Arah Letnan Harun	24	cluster_1	16	95	117	229	0

Gambar 5. Tampilan Examplset DataView

Merupakan sheet untuk menampilkan database yang telah diolah secara keseluruhan lengkap dengan cluster berjumlah 24 items. Sehingga dapat diketahui kelompok cluster_0, cluster_1 dan cluster_2.

3.5 Evaluasi

Berdasarkan hasil dari proses KDD (Knowledge Data Discovery) dapat di evaluasi bahwa pada perhitungan algoritma k-means clustering menggunakan perhitungan manual menggunakan microsoft excell dengan dataset berjumlah 24 setelah di transformasikan, bahwa

menghasilkan iterasi 1 sampai 5, dengan hasil akhir pada iterasi 5 pun berhenti sehingga diketahui jumlah cluster 3 yaitu cluster_1 berjumlah 14 anggota kelompok 1,2,3,4,6,7,8,9,10,11,14,17,18 dan 20, cluster_2 berjumlah 6 anggota kelompok 12,15,16,19,21 dan 22, cluster_3 berjumlah 4 anggota kelompok 5,13,23 dan 24.

Implementasi data pelanggaran lalu lintas ke rapidminer menghasilkan 3 cluster yaitu cluster_0 memiliki 14 *items* kelompok 1, 2, 3, 4, 6, 7, 8, 9, 10, 11, 14, 17, 18 dan 20 yang terdiri dari wilayah Bypass, Bojong Jengkol, Bypass, Cibogor Arah Ciamis, Cibogor Arah M.Hatta, Cibogor Arah Sutisna Senjaya, Cilembang Arah Cikurubuk, Cimulu, Dewi Sartika, Dewi Sartika Arah Masjid Agung, Jati, Masjid Agung, Simpang Lima, Sutisna Senjaya Arah Pancasila serta Taman Dadaha. Sedangkan cluster_1 memiliki 6 *items* kelompok 12, 16, 19, 21, 22 dan 24 yang terdiri dari wilayah Leuwidahu Arah Bojong Jengkol, Rancabango, Taman Alun-Alun, Wasita Kusuma, Wasita Kusuma Arah Bandung, Wasita Kusuma Arah Letnan Harun serta yang terakhir cluster_2 memiliki 4 *items* kelompok 5,13,15 dan 23 terdiri dari wilayah Cibogor Arah M.Hatta, Leuwidahu Arah Jati, Mitra Batik, Wasita Kusuma Arah Bojong Jengkol, dengan keseluruhan berjumlah 24 *items*. Sehingga kedua hasil tersebut sama antara perhitungan menggunakan microsoft excell dan rapidminer dengan cluster_0 berjumlah 14, cluster_1 berjumlah 6 dan cluster_2 berjumlah 4.

4. KESIMPULAN

Berdasarkan uraian yang telah dikemukakan pada bab-bab sebelumnya, maka dapat ditarik beberapa kesimpulan sebagai berikut :

1. Algoritma *k-means clustering* dapat diterapkan pada data pelanggaran lalu lintas Polres Tasikmalaya Kota, sehingga metode ini dapat membantu dalam mengelompokan data jenis pelanggaran lalu lintas.
2. Hasil perhitungan menggunakan microsoft excell dan rapidminer menghasilkan 3 cluster yaitu cluster_0 memiliki 14 *items* kelompok 1, 2, 3, 4, 6, 7, 8, 9, 10, 11, 14, 17, 18 dan 20 terdiri dari wilayah Bypass, Bojong Jengkol, Bypass, Cibogor Arah Ciamis, Cibogor Arah M.Hatta, Cibogor Arah Sutisna Senjaya, Cilembang Arah Cikurubuk, Cimulu, Dewi Sartika, Dewi Sartika Arah Masjid Agung, Jati, Masjid Agung, Simpang Lima, Sutisna Senjaya Arah Pancasila serta Taman Dadaha. Sedangkan cluster_1 memiliki 6 *items* kelompok 12, 16, 19, 21, 22 dan 24 terdiri dari wilayah Leuwidahu Arah Bojong Jengkol, Rancabango, Taman Alun-Alun, Wasita Kusuma, Wasita Kusuma Arah Bandung, Wasita Kusuma Arah Letnan Harun serta yang terakhir cluster_2 memiliki 4 *items* kelompok 5,13,15 dan 23 terdiri dari wilayah Cibogor Arah M.Hatta, Leuwidahu Arah Jati, Mitra Batik, Wasita Kusuma Arah Bojong Jengkol, dengan keseluruhan berjumlah 24 *items*.

DAFTAR PUSTAKA

- [1] Fajariyanti, W. (2017). Clustering Wilayah Pada Pelanggaran Berkendaraan Menggunakan Metode K-Means (Studi Kasus:Polres Kediri), Sikmi-Techsain, Vol 01, No.09
- [2] Ramadhani, N., Rahman, A. F., & Riskiyati, D. (2017). Analisis Cluster Data Register Perkara Lalu Lintas Menggunakan Algoritma K-Means. *SESINDO* 9, 2017.
- [3] Delita Buaton¹⁾ Fitri Nurhayati²⁾ "Clustering Pelanggaran Berkendaraan Menggunakan Algoritma K-Means Pada Polres Binjai", Jurnal STMIK Kaputama.
- [4] Lestian Cahya Ardianata¹⁾ Ahmad Zainul Fanani²⁾ "Implementasi Algoritma K-Means Pada Data Pelanggaran Lalu Lintas Di Pengadilan Negeri Purwodadi", Jurnal UDINUS, Semarang.
- [5] Handoko, K. (2016). Penerapan Data Mining dalam Meningkatkan Mutu Pembelajaran

Menggunakan Metode K-MEANS Clustering. *Jurnal Nasional Teknologi dan Sistem Informasi*, 2(3), 31-40.

- [6] Mulyani, E. D. S., Agustin, Y. H., Surgawi, N. M., & Susanto, S. (2018). Implementasi Algoritma K-Means Dan Fp-Growth Untuk Rekomendasi Bimbingan Belajar Berdasarkan Segmentasi Akademik Siswa. *IT (Informatic Technique) Journal*, 6(2), 160-173.
- [7] Agus Nur Khomarudin "*Algoritma K-Means Clustering*", Ilmu Komputer, 2016.