

ANALISIS DATA KERAWANAN KECELAKAAN LALU LINTAS MENGGUNAKAN METODE K-MEANS (STUDI KASUS POLRES BANTUL)

Eko Andriyanto Wicaksono¹⁾, Kusri²⁾, Emha Taufiq Lutfi³⁾

Mahasiswa Magister Teknik Informatika¹⁾, Dosen STMIK AMIKOM Yogyakarta^{2),3)}
Jl Ring road Utara, Condongcatur, Sleman, Yogyakarta 55281
Email : ekoaw@gmail.com¹⁾, kusrini@amikom.ac.id²⁾, emhataufiq@amikom.ac.id³⁾

Abstrak

World Health Organization (WHO) mencatat hampir 3.400 orang meninggal dunia di jalan setiap hari. Kecelakaan lalu lintas berada di urutan nomor 2 penyebab kematian setelah stroke. Kepolisian Republik Indonesia (POLRI) mencatat setiap hari ada 80 orang meninggal dunia akibat kecelakaan lalu lintas atau 3 orang per jam selama 2013.

Berbagai kebijakan dalam menekan angka kecelakaan lalu lintas telah diterapkan oleh Satlantas Polres Bantul diantaranya dengan rekayasa lalu lintas dan kegiatan pendidikan masyarakat. Kebijakan tersebut tentunya harus berasal dari data yang selama ini telah dihimpun.

Dengan bekal data kecelakaan lalu lintas, peneliti memproses data menggunakan metode K-Means untuk mencari informasi baru terkait waktu terjadinya kecelakaan lalu lintas dan menganalisisnya.

Dari analisa hasil pemrosesan menggunakan K-Means, dapat disimpulkan bahwa waktu rawan kecelakaan lalu lintas adalah malam hari yaitu dari pukul 19.30 sampai dengan 23.59.

Kata kunci: K-Means, kecelakaan, lalu lintas, data mining

1. Pendahuluan

World Health Organization (WHO) mencatat 34.153 orang meninggal dunia di jalan setiap hari [1]. Kepolisian Republik Indonesia (POLRI) mencatat setiap hari ada 80 orang meninggal dunia akibat kecelakaan lalu lintas atau 3 orang per jam selama 2013.

Satuan Lalu Lintas Polisi Resort Bantul (Satlantas Polres Bantul) pada tahun 2015 mencatat telah terjadi 1.339 kecelakaan lalu lintas. Kecelakaan lalu lintas tersebut mengakibatkan 70 orang korban meninggal dunia, 10 orang korban luka berat, 1.469 korban luka ringan dan Rp 940.920.000,- kerugian material. Angka kecelakaan lalu lintas tersebut terjadi peningkatan dari tahun sebelumnya (2014) yaitu 857 jumlah kecelakaan. Penanganan angka kecelakaan lalu lintas ini perlu mendapat penanganan yang lebih efektif, jika tidak

diperkirakan angka korban kecelakaan lalu lintas akan meningkat 2 kali lipat tiap tahunnya [2].

Berbagai kebijakan dalam menekan angka kecelakaan lalu lintas telah diterapkan oleh Satlantas Polres Bantul diantaranya dengan rekayasa lalu lintas dan kegiatan pendidikan masyarakat. Kebijakan tersebut tentunya harus berasal dari data yang tersedia. Data yang direkam pada suatu kecelakaan lalu lintas antar lain dimana, kapan dan bagaimana kecelakaan terjadi. Satlantas Polres Bantul telah memiliki sistem untuk mengarsipkan setiap kejadian kecelakaan lalu lintas dari waktu ke waktu. Berdasarkan rekaman data inilah secara rutin Satlantas Polres Bantul merekapitulasi jumlah kecelakaan, jumlah korban, dan jumlah kerugian material disuatu wilayah yang kemudian dianalisis untuk menentukan daerah yang rawan terhadap kecelakaan lalu lintas.

Penelitian ini mengangkat tema pengelompokan data kecelakaan lalu lintas. Pengelompokan tersebut ditujukan untuk menentukan waktu rawan kecelakaan lalu lintas. Metode yang digunakan untuk mencapai tujuan adalah algoritma K-Means. Algoritma K-Means dipilih karena algoritma tersebut mudah diimplementasikan dan memiliki kemampuan mengelompokkan data yang besar. Adapun objek penelitian adalah data kecelakaan lalu lintas di Polres Bantul tahun 2014 dan 2015. Diharapkan dengan penelitian ini membantu pihak kepolisian untuk menekan jumlah kecelakaan lalu lintas.

Pelitian ini memiliki batasan sebagai berikut:

- Lokasi studi adalah jalan nasional, jalan provinsi, jalan kabupaten dan jalan kecamatan yang tercatat di Satlantas Polres Bantul terhadap peristiwa kecelakaan lalu lintas.
- Data kecelakaan menggunakan data sekunder tahun 2014 dan 2015 yang diperoleh dari Satlantas Polres Bantul.
- Menggunakan data mining metode K-means clustering untuk menentukan kerawanan kecelakaan lalu lintas.

Data mining adalah suatu metode pengolahan data untuk menemukan pola yang tersembunyi dari data tersebut. Hasil dari pengolahan data dengan metode data mining

ini dapat digunakan untuk mengambil keputusan di masa depan. Data mining ini juga dikenal dengan istilah pattern recognition [3].

K-means merupakan metode pengklasteran secara partitioning yang memisahkan data ke dalam kelompok yang berbeda. Dengan partitioning secara iteratif, K-Means mampu meminimalkan rata-rata jarak setiap data ke klasternya [4].

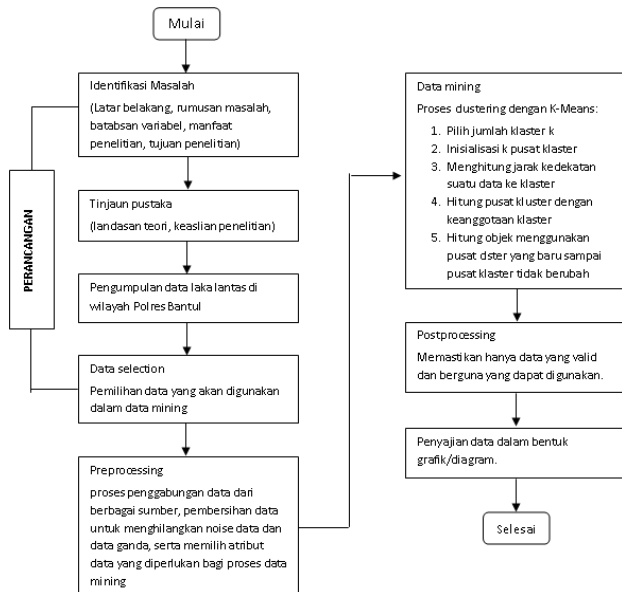
Daerah rawan kecelakaan adalah suatu lokasi dimana angka kecelakaan tinggi dengan kejadian kecelakaan berulang dalam suatu ruang dan rentang waktu yang relatif sama, diakibatkan oleh suatu penyebab tertentu. Suatu daerah dinyatakan sebagai lokasi rawan kecelakaan apabila memiliki kriteria sebagai berikut [5]:

1. Memiliki angka kecelakaan yang tinggi
2. Lokasi kejadian kecelakaan relatif menumpuk
3. Lokasi kecelakaan berupa simpangan atau segmen ruas jalan sepanjang 100-300 m untuk jalan perkotaan, ruas jalan sepanjang 1 km untuk jalan antar kota.
4. Kecelakaan terjadi dalam ruang dan rentang waktu yang relatif sama.
5. Memiliki penyebab kecelakaan dengan faktor yang spesifik.

Untuk mengidentifikasi daerah rawan kecelakaan ada dua tahapan yang harus dilewati [6], yaitu:

1. Mempelajari sejarah kecelakaan dari seluruh wilayah studi untuk kemudian memilih lokasi yang dianggap rawan kecelakaan.
2. Mempelajari secara detail lokasi terpilih untuk menemukan penanganan yang dapat dilakukan.

Penelitian ini menggunakan alur penelitian yang ditunjukkan pada gambar 1.

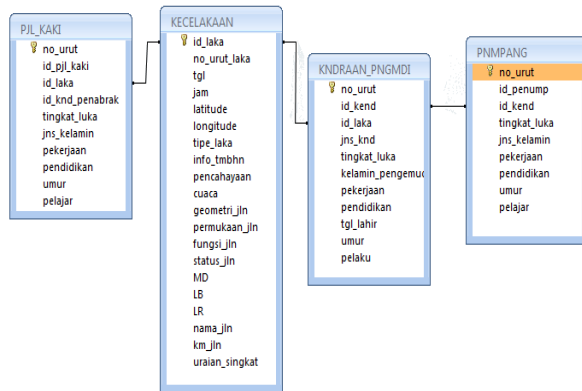


Gambar 1. Alur Penelitian

2. Pembahasan

2.1. Data Selection

Data selection adalah tahap pemilihan data yang akan diuji atau dievaluasi. Data yang dipilih dalam penelitian ini adalah data kecelakaan lalu lintas Polres Bantul tahun 2014 sampai tahun 2015 yang berasal dari aplikasi IRSMS (Integrated Road Safety Management System). Database IRSMS terdiri dari 4 buah tabel yaitu tabel KECELAKAAN, tabel KNDRAAN_PNGMDI, tabel PENUMPANG dan tabel PJJ_KAKI.



Gambar 2. Database

Sesuai kebutuhan peneliti maka data yang dipakai berada di tabel KECELAKAAN. Tabel KECELAKAAN mempunyai 20 atribut

Tabel 1. Atribut Tabel Kecelakaan

NO	NAMA	TIPE	DESKRIPSI
1	Id_laka	Int	Nomor kecelakaan, auto numbering, primary key
2	No_urut_laka	Int	Nomor urut kecelakaan, nomor lokal
3	Tgl	Date	Tanggal, bulan dan tahun kecelakaan
4	Jam	Time	Waktu terjadinya kecelakaan
5	Latitude	Text	Lokasi berdasarkan garis lintang
6	Longitude	Text	Lokasi berdasarkan garis bujur
7	Tipe_laka	Text	Tipe kecelakaan: depan-belakang, depan-depan, depan-samping dll
8	Info_tmbhn	Text	-
9	Pencahayaan	Text	Pencahayaan: terang, redup atau gelap
10	Cuaca	Text	Cuaca: cerah, mendung, hujan, kabut.
11	Geometri_jln	Text	Geometri jalan: simpang 4, simpang Y, simpang T dll.

12	Permukaan_jln	Text	Permukaan jalan: bagus atau berlubang.
13	Fungsi_jln	Text	Fungsi jalan: kolektor, arteri, lokal
14	Status_jln	Text	Status jalan: desa, kab/kota, provinsi, nasioal.
15	MD	Int	Jumlah korban meninggal dunia
16	LB	Int	Jumlah korban luka berat
17	LR	Int	Jumlah korban luka ringan
18	Nama_jln	Text	Nama jalan
19	Km_jln	Int	KM jalan (lokasi jalan)
20	Uraian_singkat	Text	Kronologi kejadian secara singkat

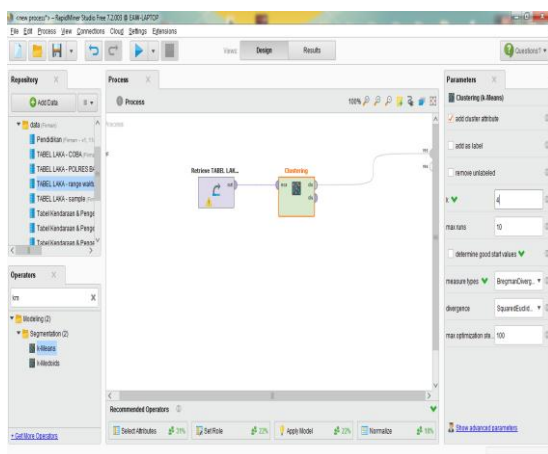
Pada tabel KECELAKAAN dari tahun 2014 sampai dengan 2015 data diperoleh sebanyak 2.182 data kecelakaan.

2.2. Preprocessing

Pada tahap ini data diteliti untuk memastikan bahwa tidak ada data kosong, data salah sebelum diproses. Selain itu pada tahap ini data dilakukan proses perubahan dari waktu (jam, menit) menjadi format menit. Hal ini dilakukan untuk mempermudah proses klustering.

2.3. Data Mining

Pada tahap ini data siap untuk diproses. Peneliti menggunakan aplikasi Rapidminer untuk membantu proses klustering menggunakan metode K-Means. Jumlah kluster yang dicoba peneliti adalah 2, 3 dan 4. Hasil clustering ditunjukkan pada tabel 1.



Gambar 3. Penggunaan Rapidminer

Row No.	id	cluster	MENIT
1	1	cluster_2	930
2	2	cluster_1	1230
3	3	cluster_3	810
4	4	cluster_2	900
5	5	cluster_0	495
6	6	cluster_0	470
7	7	cluster_2	960
8	8	cluster_1	1330
9	9	cluster_0	470
10	10	cluster_2	1080
11	11	cluster_2	990
12	12	cluster_3	570
13	13	cluster_0	420
14	14	cluster_2	1080
15	15	cluster_0	120
16	16	cluster_2	930
17	17	cluster_0	525

Gambar 4. Hasil Klustering Waktu

2.4. Postprocessing

Setelah diperoleh hasil klustering diperoleh, langkah selanjutnya adalah menggabungkan dengan data awal (menggunakan Microsoft excel). Pada tahap ini memastikan data valid dan bisa digunakan.

Tabel 1. Klustering (3 kluster)

NO_URUT	ID_LAKA	TANGGAL	MD	LB	LR	CLUSTER	MENIT	JAM
625	1122460	2014-01-05	0	0	1	cluster_2	930	15:30
626	1122467	2014-01-06	0	0	1	cluster_1	1230	20:30
627	1122471	2014-01-21	0	0	1	cluster_3	810	13:30
628	1122475	2014-01-21	0	0	2	cluster_2	900	15:00
629	1122655	2014-01-13	0	0	1	cluster_0	495	8:15
630	1122668	2014-01-13	0	0	1	cluster_0	470	7:50
631	1122690	2014-01-10	0	0	1	cluster_2	960	16:00
632	1122991	2014-01-24	0	0	1	cluster_1	1330	22:10
633	1123136	2014-01-26	0	0	2	cluster_0	470	7:50
634	1123324	2014-01-25	0	0	1	cluster_2	1080	18:00
635	1123329	2014-01-25	0	0	1	cluster_2	990	16:30

Keterangan

MD: Meninggal Dunia, LB : Luka Berat, LR: Luka Ringan

2.5. Analisa Hasil

Dari proses klustering waktu kecelakaan lalu lintas dengan jumlah kluster 2 dapat diperoleh informasi yang ditunjukkan oleh tabel 2.

Tabel 2. Hasil clustering (2 kluster)

KLUSTER	WAKTU	JML	MD	LB	LR
0	15.15 – 23.59	1390	98	14	1588
1	00.05 – 15.10	791	59	12	953

Dengan proses klustering waktu kecelakaan lalu lintas dengan jumlah kluster 3 dapat diperoleh informasi yang ditunjukkan oleh tabel 3.

Tabel 3. Hasil clustering (3 kluster)

KLUSTER	WAKTU	JML	MD	LB	LR
0	18.25 - 23.59	1020	71	12	1145
1	11.59 - 18.20	652	42	6	786
2	00.05 - 11.50	509	44	8	610

Dengan proses klustering waktu kecelakaan lalu lintas dengan jumlah kluster 4 dapat diperoleh informasi yang ditunjukkan oleh tabel 4.

Tabel 4. Hasil clustering (4 kluster)

KLUSTER	WAKTU	JML	MD	LB	LR
0	00.05- 08.55	253	32	6	279
1	19.30- 23.59	845	57	10	938
2	14.25- 19.25	633	45	5	767
3	09.00- 14.20	450	23	5	557

Dengan proses klustering waktu kecelakaan lalu lintas dengan jumlah kluster 5 dapat diperoleh informasi yang ditunjukkan oleh tabel 5.

Tabel 5. Hasil clustering (5 kluster)

KLUSTER	WAKTU	JML	MD	LB	LR
0	00:05-05:45	68	15	5	64
1	19:55-23:59	771	52	7	854
2	15:30-19:53	608	44	7	722
3	10:50-15:20	399	23	5	500
4	06:00-10:45	335	23	2	401

Dengan proses klustering waktu kecelakaan lalu lintas dengan jumlah kluster 6 dapat diperoleh informasi yang ditunjukkan oleh tabel 6.

Tabel 6. Hasil clustering (6 kluster)

KLUSTER	WAKTU	JML	MD	LB	LR
0	05:00-09:20	240	21	3	276
1	09:25-13:05	311	16	1	399
2	20:30-23:59	685	45	5	763
3	16:55-20:25	539	38	7	629
4	00:05-04:45	48	12	4	44
5	13:10-16:50	358	25	6	430

Dari percobaan pengklusteran dengan jumlah kluster 2, 3, 4, 5 dan 6 dapat dihasilkan range waktu jumlah kecelakaan lalu lintas dan jumlah korban. Semakin banyak jumlah klusternya maka semakin detail range waktu.

Dari tabel 6 ditunjukkan waktu rawan kecelakaan lalu lintas terjadi dari pukul 20:30 sampai dengan 23:59 dengan jumlah kecelakaan 685, korban meninggal dunia 45 orang, korban luka berat 5 orang dan korban luka ringan 763 orang.

3. Kesimpulan

Dari penelitian yang telah dilakukan terdapat beberapa kesimpulan yaitu :

- Jumlah kecelakaan paling tinggi terjadi pada pukul 19.55 sampai dengan 23.59 dengan jumlah kecelakaan 771 (meninggal dunia 52 orang, luka berat 7 orang dan luka ringan 854 orang).
- Jumlah kecelakaan paling rendah terjadi pada pukul 00.05 sampai dengan 05.45 dengan jumlah kecelakaan 68 (minggal dunia 15 orang, luka berat 5 orang, dan luka ringan tidak ada).
- Dari hasil kluster 4, 5 dan 6, kecelakaan tertinggi terjadi pada malam hari.
- Dari hasil kluster 1, 2, 3, 4, 5 dan 6, kecelakaan paling rendah terjadi pada pagi/dini hari.
- Semakin banyak jumlah kluster, maka semakin akan diperoleh range waktu kecelakaan yang detail.

Daftar Pustaka

- WHO, "Global status report on road safety", 2015.
- Prihartono, "Koordinasi Keselamatan Jalan (Implementasi RUNK Jalan 2011-2035), Jakarta: Bappenas, 2012.
- B. Santosa, "Data Mining: Teknik Pemanfaatan Data untuk Keperluan Bisnis", Yogyakarta: Graha Ilmu, 2007.
- J.B. MacQueen, "Some Methods for classification and Analysis of Multivariate Observations, Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability," Berkeley: University of California Press.
- Depkimpraswil, "Penanganan Lokasi Rawan Kecelakaan Lalu Lintas", Departemen Pemukiman dan Prasarana Wilayah, Jakarta.
- A.S. Wedasana, "Analisa Daerah Rawan Kecelakaan dan Penyusunan Database Berbasis Sistem Informasi Geografis (Studi Kasus Kota Denpasar)", Program Magister Teknik Sipil Universitas Udayana Denpasar.

Biodata Penulis

Eko Andriyanto Wicaksono, memperoleh gelar Sarjana Komputer (S.Kom), Jurusan Teknik Informatika STMIK AMIKOM Yogyakarta, lulus tahun 2009.

Kusrini, memperoleh gelar Doktor di Program Pasca Sarjana Ilmu Komputer UGM pada tahun 2010. Sebelumnya beliau mendapatkan gelar Magister Ilmu Komputer (M.Kom) pada Program Pasca Sarjana Ilmu Komputer UGM pada tahun 2006 dan sarjana Ilmu Komputer dari Program Studi Ilmu Komputer UGM pada tahun 2002. Kini menjadi dosen tetap Strata-1 dan Pascasarjana di STMIK AMIKOM Yogyakarta.

Emha Taufik Lutfi, memperoleh gelar Magister Ilmu Komputer (M.Kom) pada Program Pasca Sarjana Ilmu Komputer UGM dan sarjana Teknik Elektro pada Program Studi Teknik Elektro UGM. Kini menjadi dosen tetap Strata-1 di STMIK AMIKOM Yogyakarta.

