

Analisis Pola Kecelakaan Lalu Lintas Menggunakan Metode Clustering Studi Kasus Polresta Samarinda

Traffic Accident Pattern Analysis Using Clustering Method: A Case Study of Polresta Samarinda

Hardheana Eka Rahma¹

2111102441053@umkt.ac.id

Teknik Informatika, Sains dan Teknologi, Universitas Muhammadiyah Kalimantan Timur

Asslia Johar Latifah²

ajl722@umkt.ac.id

Teknik Informatika, Sains dan Teknologi, Universitas Muhammadiyah Kalimantan Timur

Korespondensi penulis : 2111102441053@umkt.ac.id

Abstract. *The increase in traffic accidents which cause major consequences, such as injuries and fatalities, can be caused by an increase in population and the need for vehicles in a certain area. It is difficult to predict the time and location of traffic accidents, which are increasing as road length and vehicle movements increase. In the Samarinda Police area, the rapid growth of urbanization and vehicle mobility is a significant challenge in reducing the number of accidents. Statistics show an alarming number of accidents, requiring more effective handling to prevent them from increasing every year. The K-Means clustering algorithm is used in this research to identify accident-prone hours and group road accident data. Six variables including accident rate, number of fatalities, serious injuries, minor injuries, type of accident, and weather were taken from data collected in 2023 by the Samarinda Police Traffic Unit. After data collection, preprocessing and K-Means Clustering stages were carried out using RapidMiner. The results of the analysis help design more effective prevention strategies.*

Keywords: *Traffic Accidents, K-Means Clustering, Polresta Samarinda, Prevention, RapidMiner.*

Abstrak. Meningkatnya kecelakaan lalu lintas yang menimbulkan akibat besar, seperti korban luka dan korban jiwa, dapat disebabkan oleh peningkatan jumlah penduduk dan kebutuhan kendaraan di suatu wilayah tertentu. Kecelakaan lalu lintas sulit diprediksi waktu dan lokasinya, yang semakin meningkat seiring pertambahan panjang jalan dan pergerakan kendaraan. Di wilayah Polresta Samarinda, pertumbuhan urbanisasi dan mobilitas kendaraan yang pesat menjadi tantangan signifikan dalam mengurangi angka kecelakaan. Statistik menunjukkan jumlah kecelakaan yang mengkhawatirkan, memerlukan penanganan lebih efektif untuk mencegah peningkatan setiap tahunnya. Algoritma K-Means clustering digunakan dalam penelitian ini untuk mengidentifikasi jam rawan kecelakaan dan mengelompokkan data kecelakaan di jalan raya. Enam variabel antara lain tingkat kecelakaan, jumlah korban meninggal, luka berat, luka ringan, jenis kecelakaan, dan cuaca diambil dari data yang dihimpun pada tahun 2023 oleh Satuan Lalu Lintas Polresta Samarinda. Setelah pengumpulan data, dilakukan tahapan preprocessing dan K-Means Clustering menggunakan RapidMiner. Hasil analisis membantu merancang strategi pencegahan lebih efektif.

Kata kunci: Kecelakaan Lalu Lintas, K-Means Clustering, Polresta Samarinda, Pencegahan, RapidMiner.

PENDAHULUAN

Kebutuhan suatu daerah akan mobil semakin meningkat seiring dengan bertambahnya jumlah penduduk. Kenaikan ini dapat menyebabkan kecelakaan mobil yang mengakibatkan korban jiwa atau cedera serius. Hoobs, F.D. (2005) mendefinisikan kecelakaan lalu lintas sebagai peristiwa yang tidak mungkin diperkirakan waktu dan tempat terjadinya secara pasti.

Semakin panjang jalan dan semakin banyak kendaraan yang melewatinya, semakin sulit untuk mengurangi kejadian ini. [1].

World Health Organization (WHO) mencatat terdapat lebih dari 34.153 orang meninggal disetiap harinya [2]. Kecelakaan lalu lintas merupakan permasalahan serius yang berdampak luas terhadap keselamatan masyarakat, infrastruktur dan perekonomian suatu wilayah. Wilayah hukum Polresta Samarinda, sebagai upaya Polri untuk menjaga ketertiban dan keselamatan lalu lintas, juga menghadapi tantangan signifikan dalam mengurangi angka kecelakaan dan memitigasi dampak negatifnya, Oleh karena itu, analisis mendalam terhadap pola kecelakaan lalu lintas di wilayah ini menjadi sangat penting guna merancang langkah-langkah pencegahan yang lebih efektif. Sebagai ibukota Provinsi Kalimantan Timur, mengalami pertumbuhan urbanisasi dan mobilitas kendaraan yang pesat. Meskipun memberikan kontribusi pada perkembangan dan kemajuan, pertumbuhan ini juga membawa risiko tinggi terkait terjadinya kecelakaan lalu lintas yang dapat membahayakan keselamatan. Statistik kecelakaan lalu lintas di wilayah Polresta Samarinda menunjukkan angka yang mengkhawatirkan dengan jenis kecelakaan dan faktor penyebab yang bervariasi. Jika penanganan angka kecelakaan tidak lebih efektif, diperkirakan jumlah korban kecelakaan lalu lintas akan meningkat setiap tahunnya[3].

Bagi pengguna jalan raya, permasalahan kecelakaan lalu lintas di Kota Samarinda sangatlah memprihatinkan. Tercatat kuantitas kecelakaan lalu lintas di Kota Samarinda menurut data anatomi kecelakaan lalu lintas Satlantas Polresta Samarinda menyatakan bahwa pada 2023 terhitung mulai Januari sampai dengan Juli jumlah laka lintas itu sementara sebanyak 1078 kasus kecelakaan lalu lintas.

METODE PENELITIAN

Data yang diperlukan terdiri dari data primer yang dikumpulkan dari Polres Samarinda Kota pada bulan Januari hingga Juni 2023, serta dokumentasi kondisi wilayah yang rawan kecelakaan lalu lintas. Enam karakteristik menjadi data yang digunakan dalam penelitian ini: jumlah kecelakaan, jumlah kematian global, jumlah korban luka berat, jumlah korban luka ringan, jenis kecelakaan, dan cuaca.

Permasalahan pengkategorian data kecelakaan lalu lintas diangkat dalam penelitian ini, dimaksudkan untuk memperkirakan waktu yang paling mungkin terjadinya kecelakaan lalu lintas. Dan algoritma k-means adalah cara untuk mencapai tujuan tersebut.

James B. MacQueen awalnya mempresentasikan teknik pengelompokan K-Means pada tahun 1976. Ini adalah teknik pengelompokan non-hierarki yang relatif mudah untuk dipahami dan diterapkan, dan biasanya digunakan untuk mengelompokkan data dalam jumlah besar. Diharapkan dengan dilakukannya penelitian ini, Polresta Samarinda mampu menurunkan frekuensi kecelakaan lalu lintas.

Pemrosesan awal data, yang mencakup transformasi dan pemilihan data, dilakukan berikutnya setelah pengumpulan data. Proses pemilihan banyak elemen dengan informasi kunci untuk digunakan sebagai atribut pengelompokan dikenal sebagai pemilihan data. Proses mengubah data dari format tertulis ke tipe data nominal dikenal sebagai transformasi data. Setelah pengolahan data, digunakan K-Means Clustering untuk mengelompokkan data dengan memanfaatkan pendekatan basic random sample untuk mengidentifikasi centroid awal. Kekuatan dan kualitas cluster akan dipastikan dengan menghitung Koefisien Silhouette setelah akuisisi data cluster kecelakaan (Anggara et al, 2016 dan Anggodo et al, 2017). Melakukan analisis terhadap cluster yang dihasilkan adalah langkah terakhir.

2.1. Data Selection

Salah satu langkah dalam memilih data yang akan diuji adalah pemilihan data. Terdapat enam atribut pada data kecelakaan lalu lintas Satlantas Polresta Samarinda tahun 2023 yang dipilih untuk penelitian ini.

Tabel 1. *Atribut Tabel Kecelakaan*

NO	NAMA	DESKRIPSI
1	Tingkat Kecelakaan	dibagi menjadi 3 kelas berdasarkan tingkat kecelakaan yang diakibatkan yaitu ringan, sedang dan berat.
2	Jumlah meninggal Dunia	Berisikan data korban meninggal dunia yang diakibatkan oleh kecelakaan tersebut
3	Jumlah Luka Berat	Berisikan data korban yang mengalami luka berat akibat kecelakaan.
4	Jumlah Luka Ringan	Berisikan jumlah data korban kecelakaan yang mengalami luka ringan
5	Tipe Kecelakaan	Berisikan data jenis kecelakaan
6	Cuaca	Berisi keadaan cuaca saat kecelakaa tersebut terjadi.

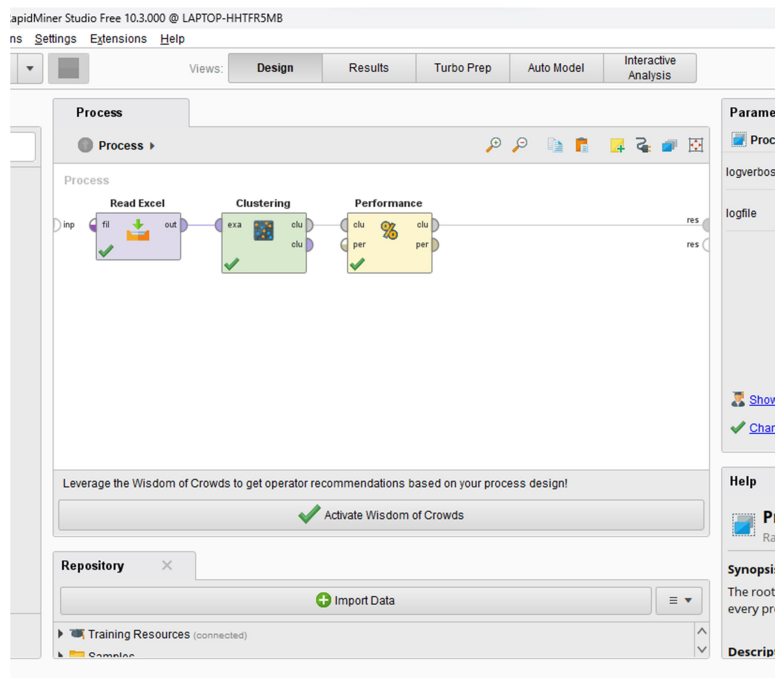
Berdasarkan atribut diatas data yang didapat diketahui sebanyak 1078 jumlah kecelakaan lalu lintas yang terjadi.

2.2. Preprocessing Data

Data diteliti untuk memastikan tidak ada kekosongan atau kesenjangan di dalamnya. Tujuannya adalah untuk menyederhanakan proses clustering.

2.3. Data Mining

Di tahap ini, data siap untuk diproses. Peneliti menggunakan pendekatan K-Means untuk mendukung proses clustering menggunakan aplikasi Rapidminer. Grafik di bawah ini menampilkan hasil clustering.



Gambar 1. Penggunaan Rapidminer

HASIL DAN PEMBAHASAN

Penulis menggunakan aplikasi Rapidminer dalam penelitian ini. Berikut tindakan dan hasil implementasinya:

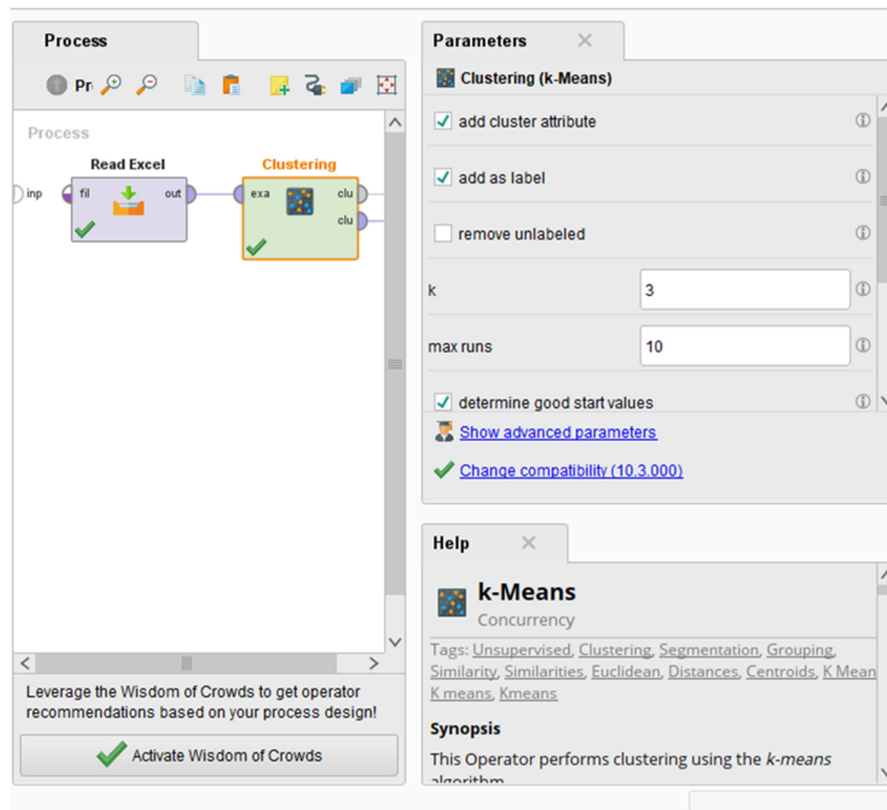
a. Pemilihan Operator

Berikut adalah operator yang digunakan pada penelitian ini adalah:

1.) Read Excel : data disini di olah menggunakan *Microsoft Excel* maka dengan itu format yang digunakan dalam data tersebut memiliki fungsi format .xlsx.

2.) Clustering (K-Means) : mengimplementasikan algoritma *K-Means*.

Peneliti melakukan langkah-langkah berikut ini: memasukkan data excel ke dalam RapidMiner, memilih operator Clustering, mengisi nilai parameter dan jumlah cluster (k), serta mengelompokkan data



Gambar 2. Pemilihan Operator

b. Hasil Olahan Data

Setelah tahap pemilihan operator, hasil dari himpunan contoh (example set) akan ditampilkan.

Cluster Model

```
Cluster 0: 305 items
Cluster 1: 773 items
Cluster 2: 0 items
Total number of items: 1078
```

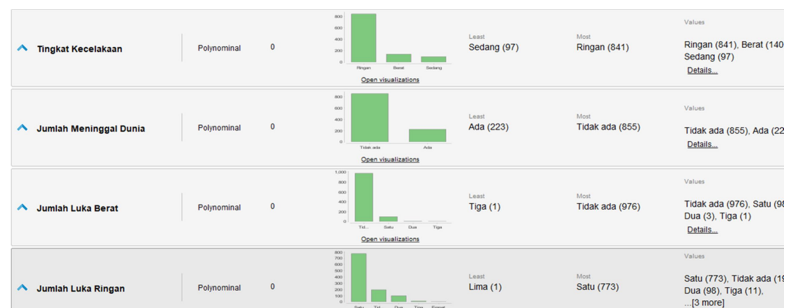
Gambar 3. Hasil Cluster Model

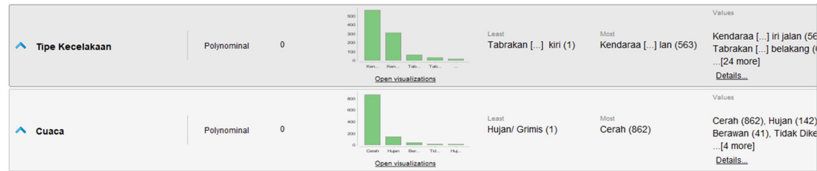
Hasil dari ExampleSet data (Clustering) akan menampilkan penjelasan setiap cluster pada semua data.

Row No.	id	label	Tingkat Kec...	Jumlah Men...	Jumlah Luk...	Jumlah Luk...	Tipe Kecela...	Cuaca
1	1	cluster_1	Ringan	Tidak ada	Tidak ada	Satu	Kendaraan O...	Cerah
2	2	cluster_1	Ringan	Tidak ada	Tidak ada	Satu	Kendaraan O...	Cerah
3	3	cluster_1	Ringan	Tidak ada	Tidak ada	Satu	Kendaraan O...	Cerah
4	4	cluster_1	Ringan	Tidak ada	Tidak ada	Satu	Kendaraan O...	Hujan
5	5	cluster_1	Ringan	Tidak ada	Tidak ada	Satu	Kendaraan O...	Cerah
6	6	cluster_1	Ringan	Tidak ada	Tidak ada	Satu	Kendaraan O...	Cerah
7	7	cluster_1	Ringan	Tidak ada	Tidak ada	Satu	Kendaraan O...	Cerah
8	8	cluster_1	Ringan	Tidak ada	Tidak ada	Satu	Kendaraan O...	Tidak Diketah...
9	9	cluster_1	Ringan	Tidak ada	Tidak ada	Satu	Kendaraan O...	Cerah
10	10	cluster_1	Ringan	Tidak ada	Tidak ada	Satu	Kendaraan O...	Cerah
11	11	cluster_1	Ringan	Tidak ada	Tidak ada	Satu	Kendaraan O...	Hujan
12	12	cluster_1	Ringan	Tidak ada	Tidak ada	Satu	Kendaraan O...	Cerah
13	13	cluster_1	Ringan	Tidak ada	Tidak ada	Satu	Kendaraan O...	Hujan
14	14	cluster_1	Ringan	Tidak ada	Tidak ada	Satu	Kendaraan O...	Cerah
15	15	cluster_0	Ringan	Tidak ada	Tidak ada	Dua	Kendaraan O...	Cerah

ExampleSet (1,078 examples, 2 special attributes, 6 regular attributes)

Gambar 4. Hasil Data Cluster





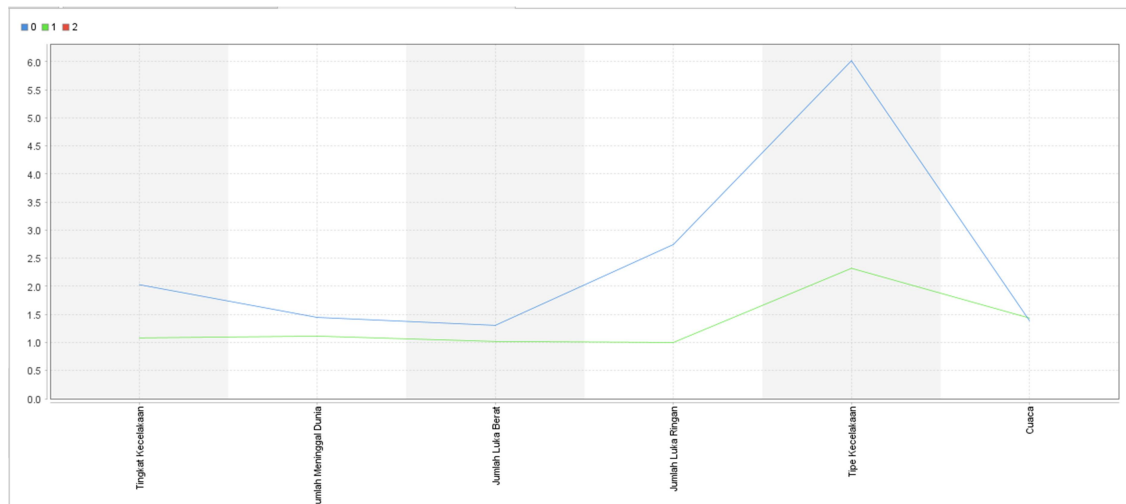
Gambar 5. Statistik Example Set

Attribute	cluster_0	cluster_1	cluster_2
Tingkat Kecelakaan	2.030	1.082	?
Jumlah Meninggal Dunia	1.446	1.113	?
Jumlah Luka Berat	1.305	1.018	?
Jumlah Luka Ringan	2.744	1	?
Tipe Kecelakaan	6.016	2.321	?
Cuaca	1.387	1.435	?

Gambar 6. Centroid Tabel

Tabel Centroid menampilkan pusat (centroid) dalam setiap *kluster*.

Plot yang menunjukkan lokasi pusat massa setiap cluster. dimana grafik clusternya ditampilkan dengan garis biru, cluster 1 diwakili dengan garis hijau, dan cluster 2 diwakili dengan garis merah.



Gambar 7. Grafik Hasil Plot

KESIMPULAN

Dengan menggunakan metode K-Means, dilakukan analisis data terhadap kejadian kecelakaan lalu lintas di Samarinda, ditemukan bahwa penyebab utama kecelakaan adalah kendaraan yang mengalami kehilangan kendali dan keluar ke kiri jalan, terutama pada cuaca cerah.

Dengan menggunakan data real-time dari Kasat Satlantas Polresta Samarinda, penelitian ini berhasil mengelola data kecelakaan lalu lintas untuk mengidentifikasi pola dan faktor penyebab utama. Diharapkan hasil ini dapat membantu pihak Polresta Samarinda merancang strategi pencegahan yang lebih efektif untuk mengurangi angka kecelakaan lalu lintas di wilayah tersebut.

DAFTAR PUSTAKA

- Yuhefizar, Santosa B., Eddy I. K. P, and Suprpto Y. K, 2013, Combination of Cluster Method for Segmentation of Web Visitors. *TELKOMNIKA*, 11(1), pp. 207-214. doi: <http://dx.doi.org/10.12928/telkomnika.v11i1.906>.
- Na'am J., Harlan J., Madenda S., and Wibowo E. P. 2016. Identification of the Proximal Caries of Dental X-Ray Image with Multiple Morphology Gradient Method. *International Journal on Advanced Science, Engineering and Information Technology (IJASEIT)*, 6(3), pp. 343-346. doi:10.18517/ijaseit.6.3.827.
- Na'am J., 2017. Edge Detection on Objects of Medical Image with Enhancement multiple Morphological Gradient (EmMG) Method. *4th Proc. EECSI*. 23-24 Sep. 2017. Yogyakarta: Indonesia. doi=10.1109/EECSI.2017.82390