

**ANALISIS MODEL CLUSTERING KECELAKAAN LALU
LINTAS DI KOTA PALEMBANG MENGGUNAKAN
PENDEKATAN *MACHINE LEARNING***

**Berdasarkan Data Sekunder dari Kepolisian Resor Kota Besar Palembang
Tahun 2020 – 2022**

SKRIPSI

**Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer**



Oleh

PUTRI SHOBIROH UTAMI

09011182025021

**JURUSAN SISTEM KOMPUTER
FAKULTAS ILMU KOMPUTER
UNIVERSITAS SRIWIJAYA**

2024

LEMBAR PENGESAHAN

**ANALISIS MODEL-CLUSTERING KECELAKAAN LALU LINTAS
DI KOTA PALEMBANG MENGGUNAKAN PENDEKATAN *MACHINE
LEARNING***

**Berdasarkan Data Sekunder dari Kepolisian Resor Kota Besar Palembang
Tahun 2020 – 2022**

SKRIPSI

**Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer**

Oleh

PUTRI SHOBIROH UTAMI

09011182025021

Pembimbing I

**Dr. Rossi Passarella, M.Eng.
NIP. 197806112010121004**

**Palembang, Juli 2024
Pembimbing II**

**Huda Ubaya, M.T.
NIP. 198106162012121003**

**Mengetahui, 11/7/24
Ketua Jurusan Sistem Komputer,**



**Dr. Ir. Sukemi, M.T.
NIP. 196612032006041001**

HALAMAN PERSETUJUAN

Pada hari kamis tanggal 20 Juni 2024 telah dilaksanakan ujian sidang skripsi oleh Jurusan Sistem Komputer Fakultas Ilmu Komputer Universitas Sriwijaya.

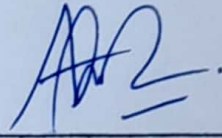
Nama : Putri Shobiroh Utami

NIM : 09011182025021

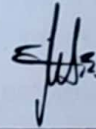
Judul : Analisis Model Clustering Kecelakaan Lalu Lintas Di Kota Palembang Menggunakan Pendekatan *Machine Learning*

Tim Penguji

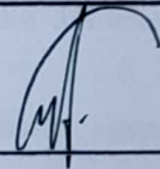
1. Ketua : Aditya Putra Perdana P, M.T.



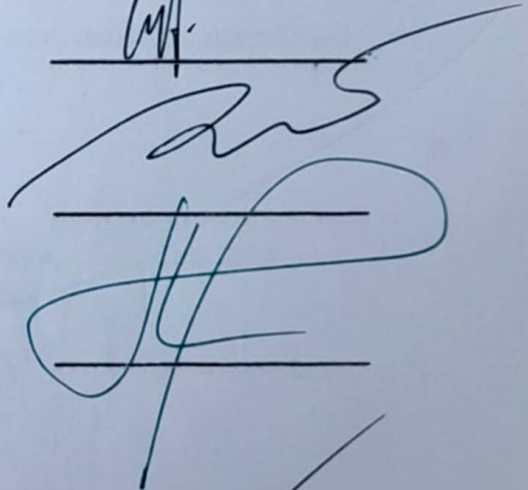
2. Sekretaris : Muhammad Ali Buchari, M.T.



3. Penguji : Dr. Ahmad Zarkasi, M.T.



4. Pembimbing I : Dr. Rossi Passarella, M.Eng.



5. Pembimbing II : Huda Ubaya, M.T.



Mengetahui, 11/7/24
Ketua Jurusan Sistem Komputer,



Dr. Ir. Sukemi, M.T.
NIP. 196612032006041001

LEMBAR PERNYATAAN

Yang bertanda tangan di bawah ini:

Nama : Putri Shobiroh Utami
NIM : 09011182025021
Program Studi : Sistem Komputer
Judul : Analisis Model Clustering Kecelakaan Lalu Lintas Di Kota Palembang Menggunakan Pendekatan *Machine Learning*

Hasil Pengecekan Software Ithenticate/Turnitin: 6%

Menyatakan bahwa laporan tesis saya merupakan hasil karya sendiri dan bukan hasil penjiplakan/plagiat. Apabila ditemukan unsur penjiplakan/plagiat dalam laporan tesis ini, maka saya bersedia menerima sanksi akademik dari Universitas Sriwijaya sesuai dengan ketentuan yang berlaku.

Demikian, pernyataan ini saya buat dengan sebenarnya dan tidak ada paksaan oleh siapapun.



Indralaya, Juli 2024
Penulis,



Putri Shobiroh Utami
NIM. 09011182025201

KATA PENGANTAR

Assalamu'alaikum Warahmatullahi Wabarakatuh

Puji syukur penulis panjatkan kehadirat Allah SWT yang telah memberikan segala rahmat, hidayah, dan karunia-Nya, sehingga penulis dapat menyelesaikan penulisan Proposal Skripsi yang berjudul “Analisis Model Clustering Kecelakaan Lalu Lintas Di Kota Palembang Menggunakan Pendekatan Machine Learning”. Proposal ini disusun sebagai salah satu syarat untuk menyelesaikan Program Studi Sistem Komputer di Universitas Sriwijaya.

Dalam laporan ini penulis menjelaskan mengenai kecelakaan yang terjadi di lalu lintas yang ada di Kota Palembang dengan menggunakan beberapa metode clustering dengan disertai data-data yang diperoleh penulis saat melakukan penelitian dan pengujian data. Penulis berharap agar tulisan ini dapat bermanfaat bagi orang banyak.

Pada kesempatan ini penulis ingin mengucapkan terima kasih kepada beberapa pihak atas ide dan saran serta bantuannya dalam menyelesaikan penulisan skripsi ini. Oleh karena itu, penulis ingin menyampaikan rasa syukur kepada Allah SWT dan terima kasih kepada:

1. Kedua orang tua penulis serta keluarga besar atas doa dan dukungan moral dan materil yang tiada henti kepada penulis. Tanpa doa restu dan kasih sayang dari kedua orang tua dan keluarga, penulis tidak akan bisa mencapai tahap ini.
2. Bapak Dr. Ir. Sukemi, M.T., selaku Ketua Jurusan Sistem Komputer Universitas Sriwijaya, atas segala arahan dan dukungannya yang sangat berarti dalam kelancaran proses akademik penulis. Kehadiran Bapak sebagai pemimpin di jurusan ini sangat memberikan inspirasi dan semangat kepada penulis.
3. Bapak Huda Ubaya, M.T., selaku Sekretaris Jurusan Sistem Komputer, serta Dosen Pembimbing Akademik dan Pembimbing 2 yang telah berkenan meluangkan waktunya guna membimbing, memberikan saran, dan motivasi serta bimbingan terbaik untuk penulis dalam menyelesaikan skripsi. Bantuan dan perhatian Bapak sangat memudahkan penulis dalam menyelesaikan setiap tahapan penelitian ini.

4. Bapak Dr. Rossi Passarella, M.Eng., selaku Dosen Pembimbing Skripsi 1 yang telah berkenan meluangkan waktunya guna membimbing, memberikan saran, dan motivasi serta bimbingan terbaik untuk penulis dalam menyelesaikan skripsi. Kesabaran dan dedikasi Bapak dalam membimbing sangat membantu penulis untuk tetap fokus dan termotivasi.
5. Pak Angga dan Mba Sari, selaku admin Jurusan Sistem Komputer yang telah membantu mengurus seluruh berkas administrasi dengan sangat cepat dan efisien. Bantuan dan keramahan Anda sangat membantu dalam kelancaran proses administratif yang penulis lalui.
6. Semua pihak yang telah membantu dalam bentuk apapun yang tidak bisa disebutkan satu persatu. Setiap bantuan, sekecil apapun, sangat berarti bagi penulis dalam menyelesaikan skripsi ini.

Penulis menyadari dalam pembuatan skripsi ini masih terdapat kekurangan serta sangat jauh dari kata sempurna. Oleh karena itu, penulis sangat membutuhkan kritik dan saran yang bersifat membangun untuk perbaikan di masa mendatang. Harapan penulis adalah dengan adanya Skripsi ini dapat bermanfaat dan berguna bagi semua pihak yang berkepentingan serta dapat menjadi referensi yang berguna bagi penelitian selanjutnya.

Wassalamu'alaikum Warahmatullahi Wabarakatuh

Palembang, Juni 2024
Penulis,

PUTRI SHOBIROH UTAMI
NIM. 09011182025021

**ANALISIS MODEL CLUSTERING KECELAKAAN LALU LINTAS DI
KOTA PALEMBANG MENGGUNAKAN PENDEKATAN *MACHINE
LEARNING***

**Berdasarkan Data Sekunder dari Kepolisian Resor Kota Besar Palembang
Tahun 2020 – 2022**

PUTRI SHOBIROH UTAMI (09011182025021)

Jurusan Sistem Komputer, Fakultas Ilmu Komputer, Universitas Sriwijaya

Email: utamiputrishobiroh@gmail.com

ABSTRAK

Antara 20 hingga 50 juta orang mengalami cedera non-fatal akibat kecelakaan lalu lintas setiap tahunnya, sementara lebih dari satu juta di antaranya menyebabkan kematian. Kecelakaan lalu lintas jalan raya merupakan ancaman yang signifikan tidak hanya bagi perekonomian tetapi juga bagi kesehatan masyarakat. Angka-angka yang diperoleh merupakan statistik obyektif dan bukan berdasarkan opini pribadi. Penelitian ini bertujuan untuk mengembangkan model pengelompokan dengan menggunakan pendekatan *machine learning* berdasarkan karakteristik kejadian dan jumlah korban kecelakaan lalu lintas di Kota Palembang, Provinsi Sumatera Selatan, Indonesia. Penelitian ini akan menganalisis pola yang mengamati hubungan yang berbeda di antara kecelakaan dengan menggunakan berbagai teknik pengelompokan seperti pengelompokan *K-means*, model campuran *Gaussian (GMM)*, berbasis kepadatan (*DBSCAN*), pengelompokan hirarkis, pengelompokan spektral, dan *OPTICS* (titik-titik yang diurutkan untuk mengidentifikasi struktur pengelompokan). Hasilnya menunjukkan bahwa algoritma yang disebut pengelompokan spektral mengungguli algoritma lainnya dengan skor *Davies-Bouldin* yang rendah (0,3221), skor *Calinski-Harabasz* yang tinggi (14789,9374), dan koefisien siluet (0,7695). Pengelompokan spektral ditemukan sebagai algoritma terbaik dari enam algoritma yang dievaluasi untuk makalah ini.

Kata Kunci: Kecelakaan Lalu Lintas, Model Clustering, *Machine Learning*, *Silhouette Coefficient*, *Davies Bouldin Index*, *Calinski-Harabasz*

**CLUSTERING MODEL ANALYSIS OF TRAFFIC ACCIDENTS IN
PALEMBANG CITY USING MACHINE LEARNING APPROACH
Based on Secondary Data from the Palembang City Police in 2020 – 2022**

PUTRI SHOBIROH UTAMI (09011182025021)

*Computer Engineering Department, Computer Science Faculty, Sriwijaya
University*

Email: utamiputrishobiroh@gmail.com

ABSTRACT

Between 20 and 50 million people suffer non-fatal injuries in road accidents every year, while more than a million of these incidents cause death. Road traffic crashes are a significant menace not only to the economy but also to public health. The numbers obtained are objective statistics not subject of personal opinion. This research aims to develop a clustering model using a machine learning approach based on characteristics of incidents and number of victims in traffic accidents in Palembang City, South Sumatra Province, Indonesia. Research will analyse the pattern which observes different relationships among accidents using various clustering techniques such as K-means clustering, Gaussian mixture models (GMM), density-based (DBSCAN), hierarchical clustering, spectral clustering, and OPTICS (ordering points to identify the clustering structure). The result shows that the algorithm called spectral clustering outperforms others with a low Davies-Bouldin score (0.3221), a high Calinski-Harabasz score (14789.9374), and a silhouette coefficient (0.7695). Spectral clustering was found to be the best algorithm out of the six algorithms evaluated for this paper.

Keywords: *Traffic Accidents, Clustering Model, Machine Learning, Silhouette Coefficient, Davies Bouldin Index, Calinski-Harabasz*

DAFTAR ISI

LEMBAR PENGESAHAN	i
HALAMAN PERSETUJUAN.....	ii
LEMBAR PERNYATAAN	iii
KATA PENGANTAR.....	iv
ABSTRAK	vi
ABSTRACT.....	vii
DAFTAR ISI	viii
DAFTAR GAMBAR	xi
DAFTAR TABEL.....	xii
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	4
1.6 Metodologi Penelitian	5
1.7 Sistematika Penulisan	5
BAB 2 TINJAUAN PUSTAKA.....	7
2.1 <i>Knowledge Discovery in Databases (KDD)</i>	7
2.2 Prapemrosesan Data	9
2.3 Clustering	10
2.4 <i>K-Means Clustering</i>	11
2.5 <i>Gaussian Mixture Models (GMM)</i>	13
2.6 <i>Density-Based Spatial Clustering of Application with Noise (DBSCAN)</i>	15
2.7 <i>Hierarchical Clustering</i>	16
2.8 <i>Spectral Clustering</i>	18
2.9 <i>Ordering Points to Identify the Clustering Structure (OPTICS)</i>	20
2.10 <i>Silhouette Coefficient</i>	21
2.11 <i>Davies-Bouldin Index</i>	23
2.12 <i>Calinski-Harabasz Index</i>	25

2.13	Penelitian Terkait	26
2.14	Kecelakaan Lalu Lintas.....	30
2.15	Faktor-faktor Penyebab Kecelakaan Lalu Lintas.....	30
2.16	Daerah Rawan Kecelakaan Lalu Lintas	32
BAB 3 METODOLOGI PENELITIAN.....		34
3.1	Data	34
3.2	Analisis Data Eksplorasi	35
3.2.1	Jumlah Tingkat Kecelakaan	38
3.2.2	Jumlah Korban Meninggal Dunia	39
3.2.3	Jumlah Korban Luka Berat	39
3.2.4	Jumlah Korban Luka Ringan	40
3.2.5	Jumlah Kecamatan	41
3.3	Kerangka Kerja Penelitian	42
3.3.1	Pengumpulan Data	43
3.3.2	Analisis Data Eksplorasi	43
3.3.3	Pengolahan Data.....	43
3.3.4	Tahap Pemodelan	44
3.3.5	Evaluasi Model.....	45
3.3.6	Pemilihan Model	45
3.3.7	Analisis Hasil Cluster.....	46
3.3.8	Kesimpulan	46
3.4	Spesifikasi Perangkat Keras dan Perangkat Lunak.....	47
3.4.1	Perangkat Keras	47
3.4.2	Perangkat Lunak.....	48
BAB 4 HASIL DAN PEMBAHASAN.....		50
4.1	Analisis Deskriptif	50
4.1.1	Perbandingan Jumlah Meninggal Dunia, Luka Berat, dan Luka Ringan	50
4.1.2	Perbandingan Antara Kategori Ringan, Sedang, dan Berat	51
4.2	Pengolahan Data	53
4.2.1	<i>Feature Selection</i>	53
4.2.2	<i>Label Encoding</i>	54
4.3	Tahap Pemodelan	56
4.4	Evaluasi Model	56

4.4.1	Visualisasi <i>Scatter Plot</i>	57
4.4.2	<i>Silhouette Coefficient</i>	58
4.4.3	<i>Davies-Bouldin Index</i>	59
4.4.4	<i>Calinski-Harabasz Index</i>	60
4.5	Pemilihan Model.....	61
4.6	Analisis Hasil <i>Clustering</i>	62
BAB 5 KESIMPULAN.....		69
5.1	Kesimpulan	69
5.2	Saran	70
DAFTAR PUSTAKA		71
LAMPIRAN.....		77

DAFTAR GAMBAR

Gambar 2.1 Tahapan KDD.....	8
Gambar 2.2 Algoritma <i>K-Means</i>	13
Gambar 3.1 Jumlah Tingkat Kecelakaan	38
Gambar 3.2 Jumlah Meninggal Dunia dalam Kecelakaan.....	39
Gambar 3.3 Jumlah Luka Berat dalam Kecelakaan	40
Gambar 3.4 Jumlah Luka Ringan dalam Kecelakaan	41
Gambar 3.5 Jumlah Kecamatan	42
Gambar 3.6 Kerangka Kerja Penelitian	42
Gambar 4.1 Perbandingan Jumlah Meninggal Dunia, Luka Berat, dan Luka Ringan.....	51
Gambar 4.2 Perbandingan Kategori Tingkat Kecelakaan	52
Gambar 4.4 Hasil visualisasi scatter plot dari beberapa algoritma clustering	58

DAFTAR TABEL

Tabel 3.1 Deskriptif Statistik Variabel Numerik	36
Tabel 3.2 Deskriptif Statistik Variabel Kategorikal	37
Tabel 3.3 Spesifikasi Perangkat Keras	47
Tabel 3.4 Spesifikasi Perangkat Lunak	48
Tabel 4.1 <i>Feature Selection</i>	53
Tabel 4.2 Hasil Score Menggunakan <i>Silhouette Coefficient</i>	59
Tabel 4.3 Hasil Score Menggunakan <i>Davies-Bouldin Score</i>	59
Tabel 4.4 Hasil Score Menggunakan <i>Calinski-Harabasz Score</i>	60
Tabel 4.5 Perbandingan Evaluasi Model dari Metode <i>Clustering</i>	61
Tabel 4.6 Jumlah Data pada Tiap Cluster dan Total dari Jumlah Korban	62
Tabel 4.7 Hasil Cluster 1	63
Tabel 4.8 Persentase Kecamatan pada Cluster 1	64
Tabel 4.9 Hasil Cluster 2	65
Tabel 4.10 Persentase Kecamatan pada Cluster 2	66
Tabel 4.11 Hasil Cluster 3	67
Tabel 4.12 Persentase Kecamatan pada Cluster 3	68

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Selain mempengaruhi berbagai aspek kehidupan bangsa dan negara, transportasi merupakan alat yang sangat penting dan strategis untuk mempercepat perekonomian dan memperkuat persatuan dan kesatuan nasional [1]. Pentingnya transportasi terlihat dari semakin tingginya permintaan akan layanan angkutan untuk mobilitas orang dan barang ke seluruh penjuru tanah air, bahkan hingga ke luar negeri. Selain itu, transportasi berfungsi sebagai penunjang, pendorong, dan penggerak bagi pertumbuhan daerah berpotensi yang belum berkembang, guna meningkatkan dan meratakan pembangunan serta hasil-hasilnya. [2].

Lalu lintas merupakan sebuah sistem yang sangat penting dalam menghubungkan dan memfasilitasi komunikasi antara masyarakat dalam suatu wilayah. Dengan kata lain, lalu lintas memainkan peran penting dalam memastikan kelancaran dan kemajuan pembangunan di suatu daerah atau negara. Dengan adanya lalu lintas, mobilitas dan konektivitas yang dibutuhkan masyarakat untuk melakukan kegiatan perekonomian menjadi lebih lancar. Akses menuju tempat kerja atau tempat melakukan aktivitas yang melibatkan penggunaan jalan raya pun menjadi lebih mudah. Namun, disadari bahwa kecelakaan lalu lintas merupakan dampak negatif dari mobilitas ini. Kecelakaan tersebut dapat mengancam keselamatan dan kesejahteraan masyarakat jika tidak ditangani dengan serius.

Kecelakaan lalu lintas adalah salah satu masalah dalam transportasi. Masalah ini biasanya timbul ketika sarana transportasi, termasuk jalan, kendaraan, dan fasilitas pendukung lainnya, belum mampu mengimbangi perkembangan di masyarakat. Pertumbuhan ekonomi dan populasi yang besar menyebabkan peningkatan aktivitas pemenuhan kebutuhan, yang pada gilirannya meningkatkan permintaan terhadap alat transportasi, baik pribadi maupun umum [3]. Belakangan ini, kecelakaan lalu lintas sering terjadi dan menyebabkan banyak kerugian. Dampak dari kecelakaan lalu lintas meliputi kerusakan pada fasilitas-fasilitas publik serta korban jiwa yang meninggal dunia [4]. Dengan bertambahnya jumlah

kendaraan bermotor, baik roda dua maupun roda empat, kondisi lalu lintas menjadi semakin rumit, yang kemudian berkontribusi terhadap peningkatan insiden kecelakaan lalu lintas [5].

Kecelakaan lalu lintas merupakan isu utama dalam ranah pemerintahan dan sosial karena dapat menimbulkan kerugian secara finansial dan mengancam keselamatan manusia. Setiap tahun di seluruh dunia, ada lebih dari 1,2 juta orang meninggal dan 50 orang terluka akibat kecelakaan [6]. Sebuah studi yang dilakukan [7] menyatakan bahwa penyebab utama kematian setelah kardiovaskular adalah kecelakaan lalu lintas jalan.

Menurut [8] kecelakaan di jalan raya menjadi penyebab utama kematian di negara-negara dengan pendapatan rendah dan menengah, juga menjadi penyebab kematian terbanyak di seluruh dunia bagi pengemudi di atas usia 15 tahun dan di bawah 29 tahun. Menurut WHO, sekitar 20 hingga 50 juta cedera setiap tahunnya disebabkan oleh kecelakaan di jalan raya. Keselamatan di jalan melibatkan pengemudi, kendaraan, dan infrastruktur jalan. Kecepatan berlebih, pengaruh alkohol atau obat-obatan, tidak mengikuti aturan keselamatan, menggunakan telepon saat mengemudi, mengemudikan kendaraan yang tidak lolos uji inspeksi, dan masalah pada infrastruktur jalan adalah beberapa penyebab utama kecelakaan.

Ada tiga kelompok faktor yang menyebabkan kecelakaan lalu lintas, salah satunya adalah faktor manusia seperti penggunaan kendaraan dengan kecepatan tinggi, perilaku ugal-ugalan, membawa muatan berlebihan, dan ketidakpatuhan terhadap aturan lalu lintas [9]. Kelompok kedua faktor adalah terkait dengan kendaraan, seperti kondisi kaca spion, lampu kendaraan, dan kelompok ketiga adalah faktor lingkungan fisik jalan, seperti keadaan jalan yang rusak, berlubang, minim penerangan, kekurangan rambu lalu lintas, dan tikungan yang tajam.

Penelitian [10] menunjukkan bahwa penggunaan prinsip komponen utama dan algoritma Canopy-Kmeans dalam mengidentifikasi titik-titik rawan kecelakaan lalu lintas dapat memberikan hasil yang efektif. Setelah dilakukan Analisis data menggunakan metode pengelompokan ensemble clustering, ditemukan beberapa titik rawan kecelakaan yang perlu mendapat perhatian lebih dalam upaya meningkatkan keselamatan lalu lintas.

Oleh karena itu, penelitian ini bertujuan untuk mendalami permasalahan terkait dan mengeksplorasi pola kejadian kecelakaan lalu lintas di Kota Palembang, baik berdasarkan jenis kecelakaan maupun jumlah korban yang terlibat. Selain itu, penelitian ini juga bertujuan untuk mengevaluasi kinerja metode analisis model clustering dalam mengelompokkan pola kecelakaan lalu lintas di wilayah tersebut.

Hasil dari penelitian ini diharapkan dapat memberikan wawasan yang lebih mendalam tentang faktor-faktor penyebab kecelakaan serta lokasi titik rawan kecelakaan di Kota Palembang, sehingga dapat menjadi dasar bagi pihak terkait dalam pengembangan kebijakan dan strategi pencegahan kecelakaan lalu lintas di masa mendatang.

1.2 Rumusan Masalah

Dari penelitian yang dilakukan terdapat rumusan masalah yang dapat diidentifikasi:

1. Bagaimana analisis dari model clustering untuk menentukan metode terbaik pada kecelakaan lalu lintas di Kota Palembang?
2. Bagaimana pengelompokan kecelakaan lalu lintas berdasarkan tipe kecelakaan dan jumlah korban yang terjadi di Kota Palembang?
3. Bagaimana mengidentifikasi lokasi titik rawan kecelakaan berdasarkan hasil dari setiap cluster yang terbentuk?
4. Faktor-faktor apa saja yang mempengaruhi kecelakaan yang terjadi di Kota Palembang berdasarkan variabel yang dipilih?

1.3 Batasan Masalah

Dalam penelitian ini, terdapat beberapa batasan yang perlu diperhatikan.

1. Data kecelakaan lalu lintas yang digunakan berasal dari Kepolisian Resor Kota Besar Palembang.
2. Data yang digunakan untuk keperluan Analisis diambil dari tanggal 1 Januari 2020 – 30 Desember 2022.
3. Metode yang digunakan adalah metode clustering berbasis partisi (*partition-based*) yaitu metode *K-Means Clustering*, *GMM*, *DBSCAN*, *Hierarchical Clustering*, *Spectral Clustering*, dan *OPTICS*

4. Bahasa pemrograman yang dipakai adalah Python, dan untuk mengeksekusi kode, digunakan perangkat lunak Visual Studio Code.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah sebagai berikut:

1. Mengembangkan model clustering yang efektif untuk menganalisis data kecelakaan lalu lintas di Kota Palembang.
2. Mengelompokkan kecelakaan lalu lintas berdasarkan tipe kecelakaan dan jumlah korban yang terjadi di Kota Palembang.
3. Mengidentifikasi lokasi titik rawan kecelakaan berdasarkan hasil dari setiap cluster yang terbentuk.
4. Mengidentifikasi dan menganalisis faktor-faktor yang mempengaruhi kecelakaan lalu lintas di Kota Palembang berdasarkan variabel-variabel yang dipilih.

1.5 Manfaat Penelitian

Adapun manfaat dari penelitian ini adalah sebagai berikut:

1. Memungkinkan identifikasi metode terbaik dalam menganalisis data kecelakaan lalu lintas, yang dapat membantu dalam perencanaan dan implementasi strategi pencegahan kecelakaan.
2. Memungkinkan pemahaman yang lebih baik tentang pola kecelakaan berdasarkan jenis dan tingkat keparahan, sehingga dapat diambil langkah-langkah pencegahan yang lebih tepat sasaran.
3. Memungkinkan penempatan sumber daya pencegahan kecelakaan secara efisien dengan fokus pada area yang memiliki risiko tertinggi, sehingga dapat meningkatkan efektivitas upaya pencegahan dan pengurangan kecelakaan di Kota Palembang.
4. Memberikan wawasan mendalam tentang faktor-faktor yang memengaruhi tingkat keparahan kecelakaan, sehingga dapat dilakukan intervensi yang lebih efektif untuk mengurangi jumlah korban.

1.6 Metodologi Penelitian

Metode yang akan digunakan dalam penelitian ini sebagai berikut:

1. Tahap Pertama (Studi Pustaka dan Literatur)

Pada tahap awal ini, dimulai dengan mengidentifikasi masalah yang relevan dan berkaitan untuk menjadi fokus penelitian. Kemudian, dilakukan pencarian sumber informasi seperti artikel, jurnal, buku, internet, dan sumber lain yang terkait dengan topik penelitian.

2. Tahap Kedua (Perancangan)

Pada tahap ini, akan dibahas tentang perancangan proses yang meliputi langkah-langkah atau pendekatan yang akan diambil dalam menerapkan metode analisis yang dipilih.

3. Tahap Ketiga (Pengujian)

Pada tahap ini, akan dijelaskan bagaimana setiap langkah penelitian yang telah direncanakan pada tahap sebelumnya dijalankan. Hal ini bertujuan untuk memperoleh hasil pengujian yang sesuai dan relevan baik secara konseptual maupun praktis.

4. Tahap Keempat (Analisis)

Pada tahap ini, dilakukan pengolahan dan analisis data yang diperoleh dari hasil pengujian berdasarkan pendekatan tertentu untuk mendapatkan data yang objektif.

5. Tahap Kelima (Kesimpulan)

Pada tahap terakhir ini, akan disusun kesimpulan dari seluruh tahapan yang telah dilakukan sebelumnya. Kesimpulan disusun secara ringkas, padat, dan jelas, serta disertai dengan saran yang dapat menjadi dasar untuk penelitian lanjutan agar lebih berkualitas.

1.7 Sistematika Penulisan

Agar memudahkan penyusunan tugas akhir dan menjelaskan isi setiap bab dan sub-bab secara lebih terperinci, berikut ini adalah sistematika penelitian yang disusun:

BAB 1 PENDAHULUAN

Bab ini menjelaskan secara sistematis tentang topik yang diamati, meliputi latar belakang penulisan tugas akhir, perumusan masalah yang akan dianalisis, batasan masalah yang dihadapi, tujuan penelitian, manfaat yang diharapkan, metodologi yang digunakan untuk menyelesaikan masalah, dan sistematika pembahasan yang akan diikuti.

BAB 2 TINJAUAN PUSTAKA

Bab ini merincikan dasar teori yang relevan dengan topik penelitian yang akan dilaksanakan.

BAB 3 METODOLOGI PENELITIAN

Bab ini menjelaskan secara sistematis proses observasi terkait pengumpulan data penelitian, serta spesifikasi perangkat keras dan lunak yang diperlukan.

BAB 4 HASIL DAN PEMBAHASAN

Bab ini menguraikan hasil dari metode yang diterapkan dan analisis yang dilakukan dalam penelitian.

BAB 5 KESIMPULAN

Bab ini berisi rangkuman dari hasil penelitian dan jawaban terhadap pertanyaan yang diajukan dalam rumusan masalah pada Bab 1. Kesimpulan ini menjadi dasar bagi penelitian berikutnya untuk mencapai hasil yang lebih baik.

DAFTAR PUSTAKA

- [1] S. Suwardi and W. Boediningsih, "KAJIAN SARANA TRANSPORTASI UMUM SEBAGAI JAMINAN KESELAMATAN PENUMPANG," *TSL: The Spirit of Law*, vol. 6, no. 01, pp. 1-15, 2019.
- [2] M. Taswin and A. A. Kadir, "Perencanaan Pembangunan Jalan Griya Abdi Negara Kota Pagar Alam," *JURNAL ILMIAH BERING'S*, vol. 3, no. 02, pp. 56-61, 2016.
- [3] F. C. Wardani and A. Kurniawan, " Kinerja Koridor Jaringan Jalan Kaliurang dan Jalan Parangtritis, Propinsi Daerah Istimewa YOGYAKARTA," *Jurnal Bumi Indonesia*, vol. 1, no. 2, 2012.
- [4] A. Surya, A. Abdurrahman and B. Buzman, "Analisis faktor Penyebab Kecelakaan pada Ruas Jalan Tanjung Serdang-Stagen Kabupaten Kotabaru," *Jurnal Kacapuri: Jurnal Keilmuan Teknik Sipil*, vol. 3, no. 2, pp. 227-232, 2020.
- [5] S. A. Heply and D. Z. Siahaan, "Analisis Penyebab Kecelakaan Lalu Lintas di Kota Tanjung Pinang Kepulauan Riau," *Aufklarung: Jurnal Pendidikan, Sosial dan Humaniora*, vol. 2, no. 3, pp. 200-214, 2020.
- [6] T. Beshah and S. Hill, " Mining road traffic accident data to improve safety: role of road-related factors on accident severity in Ethiopia," *In 2010 AAAI Spring symposium series*, 2010.
- [7] A. Bener and D. Crundall, " Road traffic accidents in the United Arab Emirates compared to Western countries," *Advances in Transportation Studies an international Journal*, vol. 6, pp. 5-12, 2005.
- [8] "Atlas Magazine," INSURANCE NEW AROUND THE WORLD, October 2021. [Online]. Available: <https://www.atlas-mag.net/en/article/road-safety-in-2017>. [Accessed 10 12 2023].
- [9] M. Azizirrahman, E. Normelani and D. Arisanty, "Faktor penyebab terjadinya kecelakaan lalu lintas pada daerah rawan kecelakaan di Kecamatan

- Banjarmasin Tengah Kota Banjarmasin," *JPG (Jurnal Pendidikan Geografi)*, vol. 2, no. 3, pp. 20-37, 2015.
- [10] S. Leonardi, N. Distefano and G. Pulvirenti, "Identification of road safety measures by elderly pedestrians based on K-means clustering and hierarchical cluster analysis," *Archives of transport*, vol. 56, no. 4, pp. 107-118, 2020.
- [11] B. R. Siburian, "Aplikasi Data Mining Untuk Menampilkan Tingkat Kelulusan Mahasiswa Dengan Algoritma Apriori," *Teknik Informatika STMIK Budi Darma Medan*, vol. 7, no. 2, 2014.
- [12] K. Tampubolon, H. Saragih, B. Reza, K. Epicentrum and A. Asosiasi, "Implementasi Data Mining Algoritma Apriori pada sistem persediaan alat-alat kesehatan," *Informasi dan Teknologi Ilmiah (INTI)*, vol. 1, no. 1, pp. 93-106, 2013.
- [13] S. Sularno and P. Anggraini, "Penerapan Algoritma C4. 5 Untuk Klasifikasi Tingkat Keganasan Hama Pada Tanaman Padi (Studi Kasus: Dinas Pertanian Kabupaten Kerinci)," *Jurnal Sains dan Informatika: Research of Science and Informatic*, vol. 3, no. 2, pp. 161-170, 2017.
- [14] A. Azhari and A. Anshori, " Pendekatan Aturan Asosiasi Untuk Analisis Pergerakan Saham," *In Seminar Nasional Informatika (SEMNASIF)*, vol. 1, no. 5, 2015.
- [15] S. D. Ali Ikhwan, "Penerapan Data Mining dengan Algoritma Fp-Growth untuk Mendukung Strategi Promosi Pendidikan (Studi Kasus Kampus STMIK Triguna Dharma)," *Jurnal Ilmiah Saindikom*, vol. 14, no. 3, pp. 211-226, 2015.
- [16] F. Wang, Z. Li, F. He, R. Wang, W. Yu and F. Nie, "Feature learning viewpoint of AdaBoost and a new algorithm," *IEEE Access*, vol. 7, pp. 149890-149899, 2019.
- [17] I. Hardiyanto, Y. Purwananto, S. Kom, M. Kom and R. Soelaiman, "Implementasi Segmentasi Citra dengan Menggunakan Metode Generalized Fuzzy C-Means Clustering Algorithm with Improved Fuzzy Partitions," *Teknik Pomits*, vol. 1, no. 1, pp. 1-5, 2012.

- [18] D. T. Larose and C. D. Larose, "Discovering knowledge in data: an introduction to data mining," vol. 4, 2014.
- [19] S. Khalid, F. Shoaib, T. Qian, Y. Rui, A. I. Bari, M. Sajjad, M. Shakeel and J. Wang, "Network constrained spatio-temporal hotspot mapping of crimes in Faisalabad," *Applied Spatial Analysis and Policy*, vol. 11, pp. 599-622, 2018.
- [20] Y. Agusta, "K-means–penerapan, permasalahan dan metode terkait," *Jurnal Sistem dan informatika*, vol. 3, no. 1, pp. 47-60, 2007.
- [21] M. A. W. K. Murti, "Penerapan metode K-means clustering untuk mengelompokan potensi produksi buah–buahan di provinsi daerah istimewa yogyakarta," *Universitas Sanata Dharma*, 2017.
- [22] B. Harahap, "Penerapan Algoritma K-Means Untuk Menentukan Bahan Bangunan Laris (Studi Kasus Pada UD. Toko Bangunan YD Indarung)," *Ready Star*, vol. 2, no. 1, pp. 394-403, 2019.
- [23] E. T. L. Kusriani, "Algoritma data mining," *Yogyakarta: Andi Offset*, 2009.
- [24] E. Prasetyo, "Prasetyo, E., 2012. Data mining konsep dan aplikasi menggunakan matlab," *Yogyakarta: Andi*, vol. 1, 2012.
- [25] G. L. McLachlan, K. E. Basford and M. Dekker, "Mixture Models: Inference and Applications to Clustering," *J Am Stat Assoc*, vol. 84, no. 405, pp. 337-338, 1989.
- [26] S. Theodoridis and K. Koutroumbas, "Pattern Recognition," *Elsevier*, 2009.
- [27] M. Afzalan and F. Jazizadeh, "An automated spectral clustering for multi-scale data," *Neurocomputing*, vol. 347, pp. 94-108, 2019.
- [28] N. Tremblay and A. Loukas, "Approximating spectral clustering via sampling: a review," *Sampling Techniques for Supervised or Unsupervised Tasks*, pp. 129-183, 2020.
- [29] C. D. Correa and P. Lindstrom, "Locally-scaled spectral clustering using empty region graphs," *In Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 1330-1338, 2012.

- [30] N. Alami, M. Meknassi, N. En-nahnahi, Y. El Adlouni and O. Ammor, "Unsupervised neural networks for automatic Arabic text summarization using document clustering and topic modeling," *Expert Systems with Applications*, vol. 172, p. 114562, 2021.
- [31] S. E. Schaeffer, "Graph clustering," *Computer science review*, vol. 1, no. 1, pp. 27-64, 2007.
- [32] L. Duan, S. Ma, C. Aggarwal and S. Sathe, "Improving spectral clustering with deep embedding, cluster estimation and metric learning," *Knowledge and Information Systems*, vol. 63, pp. 675-694, 2021.
- [33] M. Ester, H. P. Krigel, J. Sander and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," *Int. Conference of Knowledge Discovery and Data Mining (KDD'96), Portland, USA*, vol. 96, no. 34, pp. 226-231, 1996.
- [34] S. Nurlaela, A. Primajaya and T. N. Padilah, "Algoritma K-Medoids Untuk Clustering Penyakit Maag Di Kabupaten Karawang," *Jurnal Informatika, Manajemen dan Komputer*, vol. 12, no. 2, pp. 56-62, 2020.
- [35] S. and B. J. Dehotman, "Peningkatan Hasil Evaluasi Clustering Davies Bouldin Dengan Penentuan Titik Pusat Cluster Awal K-Means," *The University Institutional Repository*, 2018.
- [36] M. Herviany, S. P. Delima, T. Nurhidayah and K. , "Perbandingan Algoritma K-Means dan K-Medoids untuk Pengelompokkan Daerah Rawan Tanah Longsor Pada Provinsi Jawa Barat," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 1, no. 1, pp. 34-40, 2021.
- [37] T. K. Titus and M. Jajuli, "Clustering Data Kecelakaan Lalu Lintas di Kecamatan Cileungsi Menggunakan Metode K-Means," *Generation Journal*, vol. 6, no. 1, pp. 1-12, 2022.
- [38] M. Afdal and R. P. Disastra, "Analisis pola kecelakaan lalu lintas menggunakan algoritma K-means dan FP-growth studi kasus: Polres Solok," *Jurnal Ilmiah Rekayasa dan Manajemen Sistem Informasi*, vol. 8, no. 1, pp. 31-40, 2021.

- [39] A. A. Vernanda, A. Faisol and N. Vendyansyah, "Penerapan Metode K-Means Clustering Untuk Pemetaan Daerah Rawan Kecelakaan Lalu Lintas Di Kota Malang Berbasis Website," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 5, no. 2, pp. 836-844, 2021.
- [40] A. M. A. I. Firdaus and L. Lasmito, "ANALISA KECELAKAAN LALU LINTAS DI KABUPATEN LAMONGAN (JALAN RAYA GRESIK-BABAT)," *DEARSIP: Journal of Architecture and Civil*, vol. 2, no. 1, pp. 18-26, 2022.
- [41] L. L. Latifah, S. A. Hudjimartsu and I. Yanuarsyah, "ANALISIS DAERAH RAWAN KECELAKAAN LALU LINTAS MENGGUNAKAN CLUSTER ANALYSIS DI KOTA BOGOR BERBASIS WEBGIS," *Jurnal Ilmiah Teknologi Informasi Terapan*, vol. 8, no. 2, pp. 235-244, 2022.
- [42] Q. Xu and G. Tao, "Traffic accident hotspots identification based on clustering ensemble model," *In 2018 5th IEEE International Conference on Cyber Security and Cloud Computing (CSCloud)/2018 4th IEEE International Conference on Edge Computing and Scalable Cloud (EdgeCom)*, pp. 1-4, 2018.
- [43] H. Shirmohammadi, F. Hadidi and M. Saeedian, "Clustering analysis of drivers based on behavioral characteristics regarding road safety," *International Journal of Civil Engineering*, vol. 17, pp. 1327-1340, 2019.
- [44] S. S. Yassin and Pooja, "Road accident prediction and model interpretation using a hybrid K-means and random forest algorithm approach," *SN Applied Sciences*, vol. 2, pp. 1-13, 2020.
- [45] M. F. Labib, A. S. Rifat, M. M. Hossain, A. K. Das and F. Nawrine, "Road accident analysis and prediction of accident severity by using machine learning in Bangladesh," *In 2019 7th international conference on smart computing & communications (ICSCC)*, pp. 1-5, 2019.
- [46] C. F. Daiyah, "Beberapa Faktor Penyebab Kecelakaan di Indonesia," *Jurnal Ilmu Teknik*, vol. 2, no. 2, 2022.

- [47] R. Y. A. Pratama and H. Koesyanto, "Kejadian Kecelakaan pada Pengemudi Ojek Online," *HIGELA (Journal of Public Health Research and Development)*, vol. 4, no. 1, pp. 13-24, 2020.
- [48] N. Triuspita and A. P. Belladonna, "UPAYA KEPOLISIAN DALAM MENANGGULANGI KASUS KECELAKAAN LALU LINTAS DI KOTA CIMAHI," *Mores: Jurnal Pendidikan Hukum, Politik, dan Kewarganegaraan*, vol. 2, no. 1, pp. 33-44, 2020.
- [49] G. Harahap, "Masalah Lalu lintas dan Pengembangan Jalan (DPU)," *Bandung*, 2015.
- [50] A. Dharma, "IDENTIFIKASI KECELAKAAN LALU LINTAS (Study Kasus Jalan Dalu-Dalu sampai Pasir Pengaraian)," (*Doctoral dissertation, Universitas Pasir Pengaraian*), pp. 1-6, 2012.
- [51] D. Permukiman and W. Prasarana, "Penanganan lokasi rawan kecelakaan lalu lintas," *Pedoman Konstruksi dan Bangunan Pd T-09-2004-B*, 2004.
- [52] A. S. Wedasana, " Analisis Daerah Rawan Kecelakaan dan Penyusunan Database Berbasis Sistem Informasi Geografis (Studi Kasus Kota Denpasar)," *Universitas Udayana*, 2011.