

**KLASTERISASI DATA KECELAKAAN LALU LINTAS  
JALAN RAYA DI PROVINSI JAMBI MENGGUNAKAN  
METODE *K-MEANS***

Diajukan Sebagai Syarat Untuk Menyelesaikan  
Pendidikan Program Strata-1 Pada  
Jurusan Teknik Informatika



Oleh:

Bulan Fitri Dahlan  
NIM: 09021282025071

**Jurusan Teknik Informatika  
FAKULTAS ILMU KOMPUTER UNIVERSITAS SRIWIJAYA  
2024**

LEMBAR PENGESAHAN SKRIPSI

KLASTERISASI DATA KECELAKAAN LALU LINTAS  
JALAN RAYA DI PROVINSI JAMBI MENGGUNAKAN METODE  
*K-MEANS*

Oleh:

Bulan Fitri Dahlan  
NIM: 09021282025071

Palembang, 23 Januari 2024

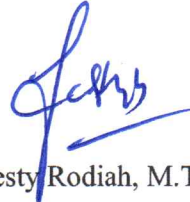
Pembimbing I



Rizki Kurniati, M.T.

NIP. 199107122019032016

Pembimbing II



Desty Rodiah, M.T.

NIP. 198912212020122011

Mengetahui,

Ketua Jurusan Teknik Informatika



Atvi Syahrini Utami, M.Kom.

NIP. 197812222006042003

## TANDA LULUS UJIAN KOMPREHENSIF SKRIPSI

Pada hari Rabu tanggal 10 Januari 2024 telah dilaksanakan ujian komprehensif skripsi oleh Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.

Nama : Bulan Fitri Dahlan  
NIM : 09021282025071  
Judul : Klasterisasi Data Kecelakaan Lalu Lintas Jalan Raya Di Provinsi Jambi Menggunakan Metode *K-Means*

Dan dinyatakan **LULUS**

1. Ketua

Osvari Arsalan, S. Kom., M.T.  
NIP. 198806282018031001



2. Penguji

Kanda Januar Miraswan, M.T.  
NIP. 199001092019031012



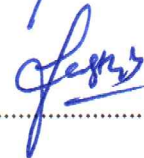
3. Pembimbing I

Rizki Kurniati, M.T.  
NIP. 199107122019032016



4. Pembimbing II

Desty Rodiah, M.T.  
NIP. 198912212020122011



Mengetahui,  
Ketua Jurusan Teknik Informatika



Alvi Syahrini Utami, M.Kom.  
NIP. 097812222006042003

## HALAMAN PERNYATAAN

Yang bertanda tangan di bawah ini:

Nama : Bulan Fitri Dahlan

NIM : 09021282025071

Program Studi : Teknik Informatika

Judul Skripsi : Klasterisasi Data Kecelakaan Lalu Lintas Jalan Raya Di  
Provinsi Jambi Menggunakan Metode *K-Means*

Hasil Pengecekan *Software iThenticate/Turnitin*: 15%

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

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



Palembang, Januari 2024



Bulan Fitri Dahlan  
NIM. 09021282025071

## **MOTTO DAN PERSEMBAHAN**

“... dan aku belum pernah kecewa dalam berdo’a kepada Engkau, Ya Rabb-Ku”

QS. Maryam:4

Kupersembahkan karya tulis ini kepada:

- Rabb-Ku, Allah SWT. yang telah memberikan kemudahan dalam segala proses penulisan.
- Nabiku, Muhammad Saw tercinta sholawat serta salam selalu terlimpah kepadanya.
- Mamak, Bapak, saudara kandung serta keluarga intiku
- Sahabat dan Teman-temanku

## ABSTRACT


*Traffic accidents are one of the main problems faced by many regions, including Jambi Province, where accidents often result in losses both in terms of casualties and material. This research focuses on applying the K-Means clustering method to analyze traffic accident data in Jambi Province. The analyzed data are from 2018 to 2022, obtained from the Traffic Directorate of Jambi Province and Muaro Jambi Police Station, and involves data collection, preparation, and preprocessing processes. Data clustering produced seven clusters for the accident data dataset with the smallest value of 0.1202 as the optimal number of clusters, while the incident location data dataset produced 6 clusters with the smallest DBI value of 0.3945 as the optimal number of clusters. This confirms the effectiveness of the K-Means method in traffic accident analysis.*

*Keywords: Davies-Bouldin Index, Jambi Province, K-Means clustering, risk analysis, traffic accidents*

Supervisor I

  
Rizki Kurniati, M.T.  
NIP. 199107122019032016

Supervisor II

  
Desty Rodiah, M.T.  
NIP. 198912212020122011

Approved,  
Head of Informatics Engineering Department

  
Alvi Syahrini Utami, M.Kom.  
NIP. 197812222006042003

## ABSTRAK

Kecelakaan lalu lintas adalah salah satu masalah utama yang dihadapi oleh banyak wilayah, termasuk Provinsi Jambi, dimana kecelakaan sering kali mengakibatkan kerugian baik dalam hal korban jiwa maupun materi. Penelitian ini berfokus pada penerapan metode *K-Means clustering* untuk menganalisis data kecelakaan lalu lintas di Provinsi Jambi. Data yang dianalisis berasal dari tahun 2018 hingga 2022, yang diperoleh dari Direktorat Lalu Lintas Provinsi Jambi dan Polres Muaro Jambi, dan melibatkan proses pengumpulan, persiapan, dan *preprocessing* data. Klasterisasi data menghasilkan tujuh klaster untuk *dataset* data kecelakaan dengan nilai terkecil yaitu 0.1202 sebagai jumlah klaster optimal, sedangkan *dataset* data lokasi kejadian menghasilkan 6 klaster dengan nilai DBI terkecil 0.3945 sebagai jumlah klaster optimal. Hal ini menegaskan efektivitas metode *K-Means* dalam analisis kecelakaan lalu lintas.

Kata Kunci: Analisis risiko, kecelakaan lalu lintas, *K-Means clustering*, Provinsi Jambi.

## KATA PENGANTAR

Puji syukur penulis panjatkan kehadirat Allah SWT, Tuhan Semesta Alam atas berkat, rahmat, rahim dan karunia-Nya yang telah diberikan kepada penulis sehingga Tugas Akhir berjudul “Klasterisasi Data Kecelakaan Lalu Lintas Jalan Raya di Provinsi Jambi Menggunakan Metode *K-Means*” dapat disusun dengan baik sebagai syarat dalam menyelesaikan studi Strata-1 program studi Teknik Informatika, Fakultas Ilmu Komputer Universitas Sriwijaya.

Pada kesempatan ini, penulis ingin mengucapkan terima kasih yang luar biasa besar kepada semua pihak yang memberikan dukungan, motivasi, dan bimbingan selama penyusunan Tugas Akhir dan penelitian. Secara khusus ucapan terima kasih ini ditujukan kepada:

1. Bapak Prof. Dr. Erwin, S.Si., M.Si. selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
2. Ibu Alvi Syahrini Utami, M.Kom. selaku Ketua Jurusan Teknik Informatika Fasilkom Universitas Sriwijaya.
3. Ibu Rizki Kurniati, M.T. dan Ibu Desty Rodiah, M.T. selaku dosen pembimbing skripsi yang telah membimbing, mengarahkan, serta memberikan motivasi kepada penyusun selama kegiatan perkuliahan dan pengerjaan skripsi.
4. Ibu Nabila Rizky Oktadini, M.T. selaku dosen pembimbing akademis yang sudah memberikan bimbingan selama proses kegiatan perkuliahan sejak awal semester.
5. Seluruh dosen Program Studi Teknik Informatika Fasilkom Unsri
6. Seluruh staff yang ada di fasilkom unsri yang telah membantu dalam kelancaran proses administrasi dan akademik selama masa perkuliahan.
7. Kepolisian Daerah Jambi yang telah mengizinkan dan membantu dalam proses pengambilan data.
8. Orang tua tersayang, Mamak dan Bapak yaitu Syafrina dan Dahlan yang telah memberikan doa, dukungan, dan semangat untuk penulis.



9. Saudara Kandung yaitu abang ijal, abang joni, dan abang iyan yang telah memberikan dukungan dan doanya untuk penulis.
10. Sahabat penulis yaitu Tamara Islahunnufus yang telah membantu dan menemani penulis dalam menyelesaikan Tugas Akhir.
11. Sahabat Queen Penulis Nida Shentia Puspita, Naifa Amanda Phatona, Selfia Herlina, Wahyuni Febrity, dan Divia Melanie Oliza Putri yang telah memberikan dukungan kepada penulis.
12. Sahabat Pentol Penulis Carolline Djuliana Aritonang, Delia Tri Gustini, Dini Putri Utami, Syahrani Puspita Wijaya, dan Yolanda Marevin yang telah menemani dan memberikan dukungan kepada penulis dalam menyelesaikan Tugas Akhir.
13. Semua pihak yang telah memberikan dukungan, semangat, dan membantu penulis dalam menyelesaikan Tugas Akhir ini.

Penulis menyadari dalam penulisan Tugas Akhir ini masih terdapat banyak kekurangan karena keterbatasan pengetahuan dan kemampuan. Oleh karena itu, kritik dan saran sangat diharapkan dalam membangun kemajuan penelitian selanjutnya.

Akhir kata, semoga Tugas Akhir ini dapat bermanfaat bagi kita semua.

Penulis, 04 Januari 2024

Bulan Fitri Dahlan

## DAFTAR ISI

LEMBAR PENGESAHAN SKRIPSI .....	ii
TANDA LULUS UJIAN KOMPREHENSIF SKRIPSI .....	iii
HALAMAN PERNYATAAN .....	iv
MOTTO DAN PERSEMBAHAN .....	v
<i>ABSTRACT</i> .....	vi
ABSTRAK .....	vii
KATA PENGANTAR.....	viii
DAFTAR ISI .....	x
DAFTAR TABEL.....	xiii
DAFTAR GAMBAR .....	xv
I. BAB I PENDAHULUAN.....	I-1
1.1 Pendahuluan .....	I-1
1.2 Latar Belakang Masalah.....	I-1
1.3 Rumusan Masalah .....	I-4
1.4 Tujuan Penelitian.....	I-4
1.5 Manfaat Penelitian.....	I-5
1.6 Batasan Masalah.....	I-5
1.7 Sistematika Penulisan.....	I-5
1.8 Kesimpulan.....	I-7
II. BAB II KAJIAN LITERATUR .....	II-1
2.1 Pendahuluan .....	II-1
2.2 Landasan Teori .....	II-1
2.2.1 <i>Data Mining</i> .....	II-1
2.2.2 <i>Unsupervised Learning</i> .....	II-1
2.2.3 Klasterisasi .....	II-2
2.2.4 <i>K-Means</i> .....	II-2
2.2.5 <i>K-Means++</i> .....	II-4
2.2.6 <i>Euclidean Distance</i> .....	II-5
2.2.7 Normalisasi: MinMaxScaler .....	II-6

2.2.8	<i>Davies-Bouldin Index (DBI)</i> .....	II-7
2.3	<i>Rational Unified Process (RUP)</i> .....	II-8
2.4	Penelitian Lain yang Relevan.....	II-10
2.4.1	Analisis Kecelakaan Berlalu Lintas Di Kota Jakarta Dengan Menggunakan Metode <i>K-Means</i> (Purwaningsih, 2019) ..	II-10
2.4.2	Algoritma <i>K-Means Clustering</i> Untuk Pengelompokkan Ayat Al Quran Pada Terjemahan Bahasa Indonesia (Robani & Widodo, 2016).....	II-10
2.4.3	Performance Measurement of <i>K-Means Clustering</i> Algorithm Using Davies Bouldin Index Method (Stiadi & Sundani, 2021) .....	II-11
2.4.4	<i>K-Means With Three Different Distance Metrics</i> (Singh et al., 2013) .....	II-11
2.5	Kesimpulan.....	II-12
III.	BAB III METODOLOGI PENELITIAN .....	III-1
3.1	Pendahuluan .....	III-1
3.2	Pengumpulan Data .....	III-1
3.3	Tahapan Penelitian .....	III-1
3.3.1	Menetapkan Kerangka Kerja Penelitian .....	III-2
3.3.2	Menetapkan Kriteria Penelitian .....	III-6
3.3.3	Menentukan Format Data Pengujian.....	III-7
3.3.4	Menentukan Alat Bantu Penelitian .....	III-7
3.3.5	Melakukan Pengujian Penelitian.....	III-8
3.3.6	Melakukan Analisis dan Menarik Kesimpulan Penelitian	III-8
3.4	Metode Pengembangan Perangkat Lunak .....	III-8
3.5	Kesimpulan.....	III-10
IV.	BAB IV PENGEMBANGAN PERANGKAT LUNAK.....	IV-1
4.1	Pendahuluan .....	IV-1
4.2	<i>Rational Unified Process (RUP)</i> .....	IV-1
4.2.1	Fase Insepsi .....	IV-1
4.2.2	Fase Elaborasi .....	IV-18

4.2.3	Fase Konstruksi.....	IV-27
4.2.4	Fase Transisi.....	IV-37
4.3	Kesimpulan.....	IV-42
V.	BAB V HASIL DAN PEMBAHASAN.....	V-1
5.1	Pendahuluan .....	V-1
5.2	Data Hasil Penelitian.....	V-1
5.2.1	Konfigurasi Percobaan.....	V-1
5.2.2	Data Hasil Konfigurasi.....	V-2
5.2.3	Hasil Pengujian Data Kecelakaan.....	V-3
5.2.4	Hasil Pengujian Data Lokasi Kejadian .....	V-3
5.3	Analisis Hasil Pengujian .....	V-4
5.4	Analisis Hasil Penelitian .....	V-5
5.5	Kesimpulan.....	V-14
VI.	BAB VI KESIMPULAN DAN SARAN .....	VI-1
6.1	Pendahuluan .....	VI-1
6.2	Kesimpulan.....	VI-1
6.3	Saran.....	VI-2
	DAFTAR PUSTAKA .....	xvii

## DAFTAR TABEL

Tabel III-1. Jumlah Kecelakaan Lalu Lintas dan Kerugian Material.....	III-3
Tabel III-2. Jumlah Kendaraan Yang Terlibat .....	III-3
Tabel III-3.Usia Korban Kecelakaan Lalu Lintas .....	III-3
Tabel III-4. Lokasi Kejadian Kecelakaan Lalu Lintas .....	III-4
Tabel III-5. Format Tabel <i>Dataset</i> Pertama.....	III-5
Tabel III-6. Deskripsi Variabel <i>Dataset</i> Pertama .....	III-5
Tabel III-7. Format Tabel <i>Dataset</i> Kedua .....	III-5
Tabel III-8. Deskripsi Variabel <i>Dataset</i> Kedua .....	III-6
Tabel III-9. Format Hasil Pengujian.....	III-7
Tabel IV-1. Kebutuhan Fungsional .....	IV-3
Tabel IV-2. Kebutuhan Non Fungsional .....	IV-3
Tabel IV-3. <i>Dataset</i> 1 setelah dilakukan preparation data .....	IV-5
Tabel IV-4. <i>Dataset</i> 1 yang telah dinormalisasikan .....	IV-6
Tabel IV-5. Probabilitas Proporsional untuk menentukan <i>centroid</i> awal kedua	IV-8
Tabel IV-6. Probabilitas Proporsional untuk menentukan <i>centroid</i> awal ketiga	IV-9
Tabel IV-7. Hasil Perhitungan Jarak antar titik dengan pusat klaster menggunakan <i>Euclidean Distance</i> .....	IV-10
Tabel IV-8. <i>Cluster</i> Hasil Iterasi Pertama .....	IV-11
Tabel IV-9. Hasil iterasi kedua (terakhir) .....	IV-13
Tabel IV-10. Definisi Aktor .....	IV-14
Tabel IV-11. Definisi <i>Use Case</i> .....	IV-15
Tabel IV-12. Skenario <i>Use Case Upload Dataset</i> .....	IV-16

Tabel IV-13. Melakukan klasterisasi menggunakan metode <i>K-Means</i> .....	IV-17
Tabel IV-14. Implementasi Kelas .....	IV-28
Tabel IV-15. Rencana Pengujian <i>Use Case Upload dataset</i> .....	IV-39
Tabel IV-16. Rencana Pengujian <i>Use Case Melakukan Klasterisasi Menggunakan Metode K-Means</i> .....	IV-39
Tabel IV-17. Pengujian <i>Use Case Upload Dataset</i> .....	IV-40
Tabel IV-18. Pengujian <i>Use Case Melakukan Klasterisasi Menggunakan Metode K-Means</i> .....	IV-41
Tabel V-1. Hasil Pengujian <i>Dataset Pertama</i> dengan DBI .....	V-3
Tabel V-2. Hasil Pengujian <i>Dataset Kedua</i> dengan DBI.....	V-4
Tabel V-3 Hasil Nilai DBI <i>dataset 1</i> dan <i>dataset 2</i> .....	V-4
Tabel V-4. Tingkat Resiko Kecelakaan Perdaerah.....	V-7
Tabel V-5. Tingkat Resiko Kecelakaan berdasarkan data lokasi kejadian.....	V-10

## DAFTAR GAMBAR

Gambar II-1. <i>Flowchart</i> Metode <i>K-Means</i> .....	II-4
Gambar II-2. Fase <i>Rational Unified Process</i> (RUP).....	II-8
Gambar III-1. Rincian Kegiatan Penelitian.....	III-1
Gambar III-2. Kerangka Kerja Penelitian .....	III-2
Gambar IV-1. <i>Use Case Diagram</i> .....	IV-14
Gambar IV-2. Rancangan Antarmuka Halaman Awal Sistem .....	IV-19
Gambar IV-3. Rancangan Antarmuka <i>Upload</i> Data Kecelakaan.....	IV-20
Gambar IV-4. Rancangan Antarmuka <i>Upload</i> Data Lokasi Kejadian .....	IV-20
Gambar IV-5. Rancangan Antarmuka ketika data telah di <i>upload</i> .....	IV-21
Gambar IV-6. <i>Activity Diagram Upload dataset</i> .....	IV-23
Gambar IV-7. <i>Activity Diagram</i> Melakukan Klasterisasi Menggunakan Metode <i>K-Means</i> .....	IV-24
Gambar IV-8. <i>Sequence Diagram upload dataset</i> .....	IV-25
Gambar IV-9. <i>Sequence Diagram</i> Melakukan Klasterisasi Menggunakan Metode <i>K-Means</i> .....	IV-26
Gambar IV-10. <i>Class Diagram</i> .....	IV-28
Gambar IV-11. Antarmuka awal sistem.....	IV-32
Gambar IV-12. Antarmuka <i>Upload File</i> .....	IV-32
Gambar IV-13. Antarmuka ketika data telah di- <i>upload</i> (a) .....	IV-33
Gambar IV-14. Antarmuka ketika data telah di- <i>upload</i> (b) .....	IV-34
Gambar IV-15. Antarmuka ketika data telah di- <i>upload</i> (c).....	IV-35
Gambar IV-16. Antarmuka ketika data telah di- <i>upload</i> (d).....	IV-36

Gambar IV-17. Antarmuka ketika data telah di-upload (e).....	IV-37
Gambar V-1. <i>Scatter plot</i> Klaster (Berdasarkan jumlah kecelakaan) .....	V-6
Gambar V-2. Histogram Jumlah Kecelakaan.....	V-6
Gambar V-3. <i>Scatter plot Klasterisasi</i> (Data Lokasi Kejadian) .....	V-9
Gambar V-4. Lokasi Kejadian Teratas berdasarkan Total Kecelakaan .....	V-9



# **BAB I**

## **PENDAHULUAN**

### **1.1 Pendahuluan**

Bab I akan membahas mengenai latar belakang masalah yang terjadi, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, dan sistematika penulisan.

### **1.2 Latar Belakang Masalah**

Peningkatan jumlah kecelakaan lalu lintas di Indonesia terjadi setiap tahunnya. Data dari Badan Pusat Statistik (BPS)<sup>1</sup> menunjukkan bahwa pada tahun 2020 jumlah kecelakaan lalu lintas mencapai 100.028 kasus dengan total korban mencapai 1447.798 jiwa dan kerugian material mencapai 198.456 juta rupiah sedangkan pada tahun 2021 jumlah kecelakaan lalu lintas mencapai 103.645 kasus kecelakaan dengan total korban mencapai 153.732 jiwa dan kerugian material mencapai 246.653 juta rupiah

Provinsi Jambi sebagai salah satu wilayah di Indonesia memiliki karakteristik geografis dan demografis yang unik, mencakup keragaman kondisi jalan, cuaca, kepadatan lalu lintas, dan perilaku pengemudi. Lalu lintas di Provinsi Jambi dapat dikatakan cukup padat karena selain ruas jalan raya yang tidak terlalu besar, di Provinsi Jambi sendiri belum memiliki jalan transportasi kereta api sehingga kendaraan penumpang maupun kendaraan muatan barang hanya dapat

---

<sup>1</sup> <https://www.bps.go.id/indicator/17/513/1/jumlah-kecelakaan-korban-mati-luka-berat-luka-ringan-dan-kerugian-materi.html> (diakses pada 8 November 2023, pukul 02.15 WIB)

melintas melewati jalan raya saja. Berdasarkan hal ini, Badan Pusat Statistik (BPS)<sup>2</sup> Indonesia menunjukkan bahwa angka kecelakaan lalu lintas di jalan raya Provinsi Jambi mengalami peningkatan. Pada tahun 2016 jumlah kecelakaan lalu lintas mencapai 1.167 kasus, pada tahun 2017 jumlah kecelakaan lalu lintas yang terjadi mencapai 1.124 kasus, dan pada tahun 2018 jumlah kecelakaan lalu lintas yang terjadi mencapai 1.252 kasus. Data ini memberikan gambaran yang jelas tentang tingkat keparahan masalah kecelakaan lalu lintas di Provinsi Jambi. Berdasarkan penjelasan diatas, diperlukan suatu upaya untuk mengidentifikasi pola dan karakteristik dari kecelakaan lalu lintas di jalan raya Provinsi Jambi guna mengurangi insiden kecelakaan dan dampak negatif yang dihasilkannya.

Penelitian klasterisasi data adalah salah satu metode yang dapat digunakan untuk mengidentifikasi pola dalam data kecelakaan lalu lintas. Klasterisasi adalah metode untuk menemukan struktur *cluster* dalam kumpulan data yang dicirikan oleh kesamaan terbesar dalam *cluster* yang sama (Sinaga & Yang, 2020). Salah satu metode yang dapat digunakan dalam klasterisasi adalah metode *K-Means*. *K-Means* saat ini banyak digunakan karena terbilang cukup mudah dipahami untuk diimplementasikan (Wahidin et al., 2021). Selain itu, algoritma *K-Means* lebih efisien untuk kumpulan data yang lebih kecil (Batra, 2011).

Penelitian sebelumnya yang berjudul *K-Means Clustering Untuk Data Kecelakaan Lalu Lintas Jalan Raya di Kecamatan Pelaihari* memiliki kesimpulan bahwa penerapan *K-Means Clustering* pada data kecelakaan lalu lintas di

---

<sup>2</sup> <https://jambi.bps.go.id/indicator/17/1089/1/kecelakaan-lalu-lintas.html> (diakses pada 2 November 2023, pukul 22.17 WIB)

Kecamatan Pelaihari berhasil dilakukan dengan perhitungan pendekatan jarak berupa *Euclidean distance* dan hasil pengujian *silhouette coefficient* pada *dataset* 1 memiliki nilai *distance measure* sebesar 0.685351106 dengan jumlah *cluster* optimal sebanyak 4 *cluster*. Dan hasil pengujian *silhouette coefficient* pada *dataset* 2 memiliki nilai *distance measure* sebesar 0.383686134 dengan jumlah *cluster* optimal sebanyak 2 *cluster* (Aprianti & Permadi, 2018). Selanjutnya, pada penelitian yang dilakukan oleh (Kurniawan & Jajuli, 2022) dengan menggunakan metode *K-Means Clustering* dengan metode pendekatan jarak *Euclidean distance* menghasilkan 3 *cluster* yaitu tidak rawan, rawan, dan sangat rawan. Pengujian penelitian ini menggunakan *silhouette coefficient* dan menghasilkan nilai evaluasi sebesar 0.35 untuk 1 *cluster* , 0.22 untuk 2 *cluster*, dan 0.38 untuk 3 *cluster*.

Penelitian yang dilakukan oleh (Supriyadi et al., 2021.) membahas tentang perbandingan algoritma *K-Means* dengan *K-Medoids* pada pengelompokan armada kendaraan truk berdasarkan produktivitas menghasilkan kesimpulan hasil *cluster* untuk algoritma *K-Means* memiliki nilai validitas DBI yang lebih rendah dibandingkan dengan nilai DBI algoritma *K-Medoids*. Selanjutnya pada penelitian yang dilakukan oleh (Arbain et al., 2023) membahas tentang perbandingan kinerja algoritma *K-Medoids* dan *K-Means* untuk klasifikasi penyakit kanker serviks menyimpulkan bahwa algoritma *K-Means* lebih unggul dibandingkan *K-Medoids* dalam memperoleh klasterisasi penyakit kanker serviks. Hasil perbandingan tersebut menggunakan evaluasi DBI dengan nilai 1,108 untuk *K-Means* dan 3.517 untuk *K-Medoids*.

Berdasarkan latar belakang yang telah diuraikan, selanjutnya akan dilakukan klasterisasi menggunakan metode *K-Means* terhadap data kecelakaan lalu lintas di Provinsi Jambi. Metode *K-Means* dipilih karena mudah untuk di implementasikan dan banyak digunakan pada penelitian sebelumnya.

### **1.3 Rumusan Masalah**

Rumusan masalah dalam penelitian ini adalah:

1. Bagaimana klasterisasi Data Kecelakaan Lalu Lintas di Provinsi Jambi menggunakan metode *K-Means*?
2. Bagaimana kinerja dari metode *K-Means* dalam mengklasterisasikan data kecelakaan lalu lintas di Provinsi Jambi?

### **1.4 Tujuan Penelitian**

Tujuan dari penelitian ini yang ingin dicapai yaitu:

1. Mengetahui cara implementasi metode *K-Means* dalam Klasterisasi Data Kecelakaan Lalu Lintas.
2. Menentukan jumlah *cluster* yang optimal untuk mengelompokkan data kecelakaan lalu lintas sehingga memudahkan pemahaman pola-pola tersebut.
3. Mengetahui hasil kinerja dari metode *K-Means* pada klasterisasi data kecelakaan lalu lintas di Provinsi Jambi

## 1.5 Manfaat Penelitian

Manfaat dari penelitian ini adalah:

1. Memahami dan mempelajari hasil kinerja dari metode *K-Means* pada Klasterisasi
2. Hasil penelitian dapat digunakan oleh pihak kepolisian dan masyarakat untuk mengetahui lokasi yang menjadi daerah rawan kecelakaan lalu lintas di Provinsi Jambi.
3. Hasil penelitian ini juga dapat digunakan sebagai dasar untuk penelitian lanjutan dalam bidang keamanan lalu lintas dan klasterisasi data.

## 1.6 Batasan Masalah

Pada penelitian ini mempunyai batasan masalah yaitu:

1. Pada penelitian ini batasan masalah menggunakan data kecelakaan lalu lintas di Provinsi Jambi dari tahun 2018 – 2022 yang diambil dari Polda Jambi.
2. Inisialisasi *centroid* awal menggunakan *K-Means++*.

## 1.7 Sistematika Penulisan

Sistematika penulisan proposal skripsi ini adalah sebagai berikut:

### **BAB I. PENDAHULUAN**

Bab ini menjelaskan prinsip-prinsip dasar yang menjadi pondasi dalam pelaksanaan penelitian ini. Hal-hal yang dibahas meliputi latar belakang

penelitian, perumusan masalah, tujuan serta kegunaan dari penelitian tersebut, pembatasan isu yang dihadapi, serta struktur dari penulisan.

## **BAB II. KAJIAN LITERATUR**

Bab ini menguraikan teori-teori fundamental yang diterapkan dalam penelitian algoritma ini dan melakukan tinjauan atas literatur penelitian sebelumnya yang berkaitan dan relevan dengan penelitian ini.

## **BAB III. METODOLOGI PENELITIAN**

Bab ini merincikan langkah-langkah yang akan dilaksanakan selama penelitian, dari pengumpulan data, detail perencanaan setiap tahap penelitian berdasarkan kerangka kerja tertentu, hingga pengembangan rancangan manajemen proyek yang akan diaplikasikan selama penelitian berlangsung.

## **BAB IV. PENGEMBANGAN PERANGKAT LUNAK**

Bab ini menguraikan tentang proses pengembangan perangkat lunak yang diterapkan dalam penelitian ini, termasuk desain perangkat lunak untuk mengklusterisasikan data kecelakaan lalu lintas di jalan raya Provinsi Jambi. Proses ini dilakukan dengan menggunakan metode *K-Means*.

## **BAB V. HASIL DAN ANALISIS**

Bab ini membahas tentang hasil yang diperoleh dari pengujian yang telah dilaksanakan dalam penelitian ini, serta melakukan analisis mendalam terhadap hasil-hasil pengujian tersebut.

## **BAB VI. KESIMPULAN DAN SARAN**

Bab ini menjelaskan kesimpulan yang ditarik dari seluruh hasil penelitian serta memberikan rekomendasi yang bisa dijadikan dasar untuk penelitian lanjutan.

### **1.8 Kesimpulan**

Penelitian ini akan melaksanakan klasterisasi menggunakan metode *K-Means*. Selanjutnya, hasil yang diperoleh akan diuji untuk menentukan klaster optimal dan mengevaluasi kinerja dari metode *K-Means*.

## DAFTAR PUSTAKA

- Anwar, A. (2014). A Review of RUP (Rational Unified Process). In *Ashraf Anwar International Journal of Software Engineering (IJSE)* (Issue 5).
- Aprianti, W., & Permadi, J. (2018). *K-MEANS CLUSTERING UNTUK DATA KECELAKAAN LALU LINTAS JALAN RAYA DI KECAMATAN PELAIHARI*. 5(5), 613–620. <https://doi.org/10.25126/jtiik2018551113>
- Arbain, D., Triloka, J., & Sriyanto. (2023). *Perbandingan Kinerja Algoritma K-Medoids Dan K-Means Untuk Klasifikasi Penyakit Kanker Serviks*. 118–131.
- Bahmani, B., Moseley, B., Vattani, A., Kumar, R., & Vassilvitskii, S. (2012). *Scalable K-Means++*. <http://arxiv.org/abs/1203.6402>
- Batra, A. (2011). *Analysis and Approach: K-Means and K-Medoids Data Mining Algorithms*.
- Han, J., Kamber, M., & Pei, J. (2011). *Data Mining. Concepts and Techniques, 3rd Edition (The Morgan Kaufmann Series in Data Management Systems)*.
- Kurniawan, T., & Jajuli, M. (2022). Clustering Data Kecelakaan Lalu Lintas di Kecamatan Cileungsi Menggunakan Metode K-Means. In *Januari 2022 Generation Journal* (Vol. 6, Issue 1).
- Larose, D. T., & Larose, C. D. (2014). *Discovering Knowledge in Data 2 nd Edition*. <https://ebookyab.ir/solution-manual-discovering-knowledge-in-data-larose-larose/>
- Lubis, D. J., & Tamam, M. B. (2022). *Penerapan K-Means Untuk Pengelompokan Beasiswa Santri di Pondok Pesantren Miftahul Huda Bogor*. 12, 7–20. <https://doi.org/10.36350/jbs.v12i1>



- Nishom, M. (2019). Perbandingan Akurasi Euclidean Distance, Minkowski Distance, dan Manhattan Distance pada Algoritma K-Means Clustering berbasis Chi-Square. *Jurnal Informatika: Jurnal Pengembangan IT*, 4(1), 20–24. <https://doi.org/10.30591/jpit.v4i1.1253>
- Permana, I., & Salisah, N. F. (2022). Pengaruh Normalisasi Data Terhadap Performa Hasil Klasifikasi Algoritma Backpropagation. *Indonesian Journal of Informatic Research and Software Engineering*, 2, 67–72.
- Purwaningsih, E. (2019). ANALISIS KECELAKAAN BERLALU LINTAS DI KOTA JAKARTA DENGAN MENGGUNAKAN METODE K-MEANS. *JURNAL ILMU PENGETAHUAN DAN TEKNOLOGI KOMPUTER*, 5. [www.bsi.ac.id](http://www.bsi.ac.id)
- Ramdani, C., & Safadila, N. (2022). LEDGER: Journal Informatic and Information Technology Analisis Data Akademis dengan Menerapkan Algoritme K-Means dan K-Means++. *OPEN ACCESS LEDGER*, 1(4). <https://doi.org/10.20895/LEDGER.V1I4.918>
- Rizaldi, P. A., Hakimah, M., & Indriyani, T. (2022). Penentuan Jurusan Siswa SMA Menggunakan Metode K-Means++.
- Robani, M., & Widodo, A. (2016). Algoritma K-Means Clustering Untuk Pengelompokan Ayat Al Quran Pada Terjemahan Bahasa Indonesia. *JURNAL SISTEM INFORMASI BISNIS*, 6(2), 164. <https://doi.org/10.21456/vol6iss2pp164-176>

- Sinaga, K. P., & Yang, M. S. (2020). Unsupervised K-means clustering algorithm. *IEEE Access*, 8, 80716–80727. <https://doi.org/10.1109/ACCESS.2020.2988796>
- Singh, A., Yadav, A., & Rana, A. (2013). K-means with Three different Distance Metrics. *International Journal of Computer Applications*, 67(10), 13–17. <https://doi.org/10.5120/11430-6785>
- Sofi Defiyanti. (2015). PENENTUAN PEMINATAN TUGAS AKHIR MAHASISWA TEKNIK INFORMATIKA UNSIKA. In *Jurnal ilmiah Solusi* (Vol. 2, Issue 5).
- Stiadi, M. J., & Sundani, D. (2021). Performance Measurement of K-Means Clustering Algorithm Using Davies Bouldin Index Method. *International Research Journal of Advanced Engineering and Science*, 6(4), 72–75. <https://www.kaggle.com>
- Sunhare, P., Chowdhary, R. R., & Chattopadhyay, M. K. (2022). Internet of things and data mining: An application oriented survey. In *Journal of King Saud University - Computer and Information Sciences* (Vol. 34, Issue 6, pp. 3569–3590). King Saud bin Abdulaziz University. <https://doi.org/10.1016/j.jksuci.2020.07.002>
- Supriyadi, A., Triayudi, A., Sholihati, I. D., Fakultas, ), Komunikasi, T., Informatika, D., Nasional, U., Sawo Manila, J., Minggu, P., Selatan, J., Khusus, D., Jakarta, I., Penulis, ) \*, Abstrak, K., & Kunci, K. (2021). *PERBANDINGAN ALGORITMA K-MEANS DENGAN K-MEDOIDS PADA*

*PENGELOMPOKAN ARMADA KENDARAAN TRUK BERDASARKAN PRODUKTIVITAS.* <https://doi.org/https://doi.org/10.29100/jipi.v6i2.2008>

Tempola, F., & Assagaf, A. F. (2018). *Clustering of Potency of Shrimp in Indonesia with K-Means Algorithm and Validation of Davies-Bouldin Index.*

Wahidin, A. J., Sensuse, D. I., & Kunci, K. (2021). Perbandingan Algoritma K-Means, X-Means Dan K-Medoids Untuk Klasterisasi Awak Kabin Lion Air. *Jurnal ICT : Information Communication & Technology*, 20(2), 298–302.

Whendasmoro, R. G., & Joseph, J. (2022). Analisis Penerapan Normalisasi Data Dengan Menggunakan Z-Score Pada Kinerja Algoritma K-NN. *JURIKOM (Jurnal Riset Komputer)*, 9(4), 872. <https://doi.org/10.30865/jurikom.v9i4.4526>

Wu, X., & Kumar, V. (2009). *The Top Ten Algorithms in Data Mining.*

Yao, Y., Liu, Y., Yu, Y., Xu, H., Lv, W., Li, Z., & Chen, X. (2013). K-SVM: An effective SVM algorithm based on K-means clustering. *Journal of Computers (Finland)*, 8(10), 2632–2639. <https://doi.org/10.4304/jcp.8.10.2632-2639>