



UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN" JAKARTA

**ANALISIS *CLUSTER* UNTUK MENENTUKAN WILAYAH BERPOTENSI
TINGKAT RESIKO KECELAKAAN PADA 13 RUAS TOL DI JABODETABEK
DAN JABAR**

SKRIPSI

DAFFA MALAEKA ADYORI

2110512129

PROGRAM STUDI S1 SISTEM INFORMASI

FAKULTAS ILMU KOMPUTER

UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN" JAKARTA

2024



UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN" JAKARTA

**ANALISIS *CLUSTER* UNTUK MENENTUKAN WILAYAH BERPOTENSI
TINGKAT RESIKO KECELAKAAN PADA 13 RUAS TOL DI JABODETABEK
DAN JABAR**

SKRIPSI

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

DAFFA MALAEKA ADYORI

2110512129

PROGRAM STUDI S1 SISTEM INFORMASI

FAKULTAS ILMU KOMPUTER

UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN" JAKARTA

2024

LEMBAR PERSETUJUAN

LEMBAR PERSETUJUAN

Yang bertanda tangan di bawah ini:

Nama : Daffa Malaeka Adyori

NIM. : 2110512129

Program Studi : SI Sistem Informasi

Judul Skripsi/TA. : Analisis Cluster Untuk Menentukan Wilayah Berpotensi Tingkat Resiko Kecelakaan Pada 13 Ruas Tol di Jabodetabek Dan Jabar

Dinyatakan telah memenuhi syarat dan menyetujui untuk mengikuti ujian sidang skripsi/tugas akhir.

Jakarta, 04 Desember 2024

Menyetujui,

Dosen Pembimbing I



Nur Hafifah Matondang S.Kom, MM.

Dosen Pembimbing II



Nindy Izavika S.SI., M.T.

Mengetahui,

Koordinator Program Studi SI Sistem Informasi,



Anita Muliawati, S. Kom., MTI.

PERNYATAAN ORISINALITAS

PERNYATAAN ORISINALITAS

Skripsi ini merupakan hasil karya sendiri serta semua sumber referensi yang dikutip maupun yang dirujuk telah saya nyatakan benar.

Nama : Daffa Malaeka Adyori

NIM : 2110512129

Tanggal : 04 Desember 2024

Bilamana di kemudian hari ditemukan ketidaksesuaian dengan pernyataan ini, maka saya bersedia dituntut dan diproses sesuai dengan ketentuan berlaku.

Jakarta, 04 Desember 2024

Yang Menyatakan,



Daffa Malaeka Adyori

PERNYATAAN PERSETUJUAN

PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIK

Sebagai civitas akademik Universitas Pembangunan Nasional Veteran Jakarta, saya yang bertanda tangan di bawah ini:

Nama : Daffa Malaeka Adyori
NIM : 2110512129
Fakultas : Ilmu Komputer
Program Studi : S1 Sistem Informasi

Demi pembangunan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Pembangunan Nasional Veteran Jakarta Hak Bebas Royalti Eksklusif (*Non-exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul:

Analisis Cluster Untuk Menentukan Wilayah Berpotensi Tingkat Resiko Kecelakaan Pada 13 Ruas Tol Di Jabodetabek Dan Jabar

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti ini Universitas Pembangunan Nasional Veteran Jakarta berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan Skripsi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sesungguhnya.

Dibuat di : Jakarta

Pada Tanggal : 04 Desember 2024

Yang Menyatakan,



Daffa Malaeka Adyori

LEMBAR PENGESAHAN

LEMBAR PENGESAHAN

Judul : Analisis *Cluster* Untuk Menentukan Wilayah Berpotensi Tingkat
Resiko Kecelakaan Pada 13 Ruas Tol di Jabodetabek dan Jabar.
Nama : Daffa Malaeka Adyori
NIM : 2110512129

Disetujui oleh :

Penguji 1:
Ika Nurlaili Isnainiyah, S.Kom., M.Sc.

Penguji 2:
Ati Zaidiah, S.Kom., MTI.

Pembimbing 1:
Nur Hafifah Matondang, S.Kom., MM., MTI.

Pembimbing 2:
Nindy Irzavika, S.SI., M.T.

Diketahui oleh:

Koordinator Program Studi:
Anita Muliawati, S.Kom., MTI.
NIP. 197005212021212002

Dekan Fakultas Ilmu Komputer:
Prof. Dr. Ir. Supriyanto, S.T., M.Sc., IPM
NIP. 197605082003121002

Tanggal Ujian Tugas Akhir :
14 Januari 2025




ANALISIS *CLUSTER* UNTUK MENENTUKAN WILAYAH BERPOTENSI TINGKAT RESIKO KECELAKAAN PADA 13 RUAS TOL DI JABODETABEK DAN JABAR

Daffa Malaeka Adyori

ABSTRAK

Kecelakaan lalu lintas di jalan tol merupakan salah satu permasalahan yang signifikan di Indonesia, terutama di wilayah Jabodetabek dan Jawa Barat. Dengan dilatar belakangi meningkatnya tingkat potensi risiko kecelakaan yang berpengaruh kepada hal-hal seperti korban meninggal, dan kendaraan rusak berat, penelitian ini mengimplementasikan metode *Clustering K-Means* untuk mengelompokkan wilayah dengan tingkat risiko kecelakaan di 13 ruas tol pada Jabodetabek dan Jabar. Data yang dimanfaatkan mencakup periode 2020 hingga 2023 dengan meliputi jumlah kecelakaan, jumlah korban meninggal, dan jumlah kendaraan rusak berat yang diperoleh dari PT Jasa Marga. Hasil analisis menampilkan nilai K optimal yaitu 3, sebagaimana didukung oleh *Elbow Method*, yang menunjukkan penurunan nilai Sum of Squared Errors (SSE) signifikan hingga jumlah *cluster* mencapai 3. Penelitian ini menyelesaikan pengelompokkan ruas tol ke dalam tiga *cluster*: Wilayah Berpotensi Risiko Tinggi dengan jumlah kecelakaan dan fatalitas sedang, Wilayah Berpotensi Risiko Sedang dengan jumlah kecelakaan sedang dan fatalitas tinggi, dan Wilayah Berpotensi Risiko Rendah dengan jumlah kecelakaan rendah dan fatalitas rendah. Evaluasi model menggunakan Davies-Bouldin Index menghasilkan skor sebesar 0.962, yang menunjukkan kualitas *clustering* yang baik. Hasil visualisasi data dilakukan menggunakan *framework* Streamlit, yang menampilkan tampilan interaktif untuk eksplorasi dan pemahaman hasil *clustering* pada berbagai wilayah tahun 2020-2023. Hasil penelitian ini memberikan kontribusi pada identifikasi wilayah rawan kecelakaan dan dapat digunakan oleh pihak pengelola jalan tol untuk merancang strategi mitigasi yang lebih tepat sasaran.

Kata Kunci: Kecelakaan lalu lintas, *Clustering K-Means*, risiko kecelakaan, visualisasi data.

CLUSTER ANALYSIS TO DETERMINE POTENTIAL ACCIDENT RISK AREAS ON 13 TOLL ROADS IN JABODETABEK AND JABAR

Daffa Malaeka Adyori

ABSTRACT

Traffic accidents on toll roads are one of the significant problems in Indonesia, especially in the Jabodetabek and West Java regions. With the background of the increasing level of potential risk of accidents that affect things such as casualties, and heavily damaged vehicles, this study implements the K-Means Clustering method to group areas with accident risk levels in 13 toll roads in Jabodetabek and West Java. The data utilized covers the period 2020 to 2023 and includes the number of accidents, the number of fatalities, and the number of heavily damaged vehicles obtained from PT Jasa Marga. The analysis results show the optimal K value of 3, as supported by the Elbow Method, which shows a significant decrease in the Sum of Squared Errors (SSE) value until the number of clusters reaches 3. This study completed the grouping of toll road sections into three clusters: High Risk Potential Areas with moderate number of crashes and fatalities, Medium Risk Potential Areas with moderate number of crashes and high fatalities, and Low Risk Potential Areas with low number of crashes and low fatalities. Model evaluation using the Davies-Bouldin Index resulted in a score of 0.962, indicating good clustering quality. The results of data visualization were carried out using the Streamlit framework, which displays an interactive display for exploration and understanding of clustering results in various regions in 2020.

Keywords: *Traffic accident, K-Means Clustering, accident risk, data visualisation.*

KATA PENGANTAR

Alhamdulillah puji dan syukur penulis panjatkan kehadirat Allah SWT atas segala rahmat dan karunia yang telah memberikan kelancaran dan kemudahan dalam penulisan skripsi ini hingga selesai. Proposal Skripsi ini merupakan salah satu syarat untuk mendapatkan gelar Sarjana Komputer Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta.

Penulis menyadari bahwa dalam penulisan penyusunan Tugas Akhir Skripsi, tidak mungkin selesai tanpa adanya dukungan, dan pengarahan dari berbagai pihak. Maka dari itu, penulis ucapkan rasa hormat dan terima kasih yang sebesar-besarnya kepada:

1. Pintu surga dan cinta tertulus penulis, Ibunda Riri Setiyowati. Mustahil penulis dapat melewati segala permasalahan yang penulis alami selama ini tanpa kehadirannya. Izinkan penulis mengucapkan terima kasih yang seluas-luasnya, terimakasih atas segala pengorbanan dan tulus kasih yang telah engkau berikan. Terimakasih telah mendidik, menemani berproses, bertumbuh, dan selalu membuat bangkit agar penulis selalu menjadi pribadi yang lebih baik dan kuat selamanya. Terimakasih telah memberikan kasih sayang, nasihat, dan dukungan yang tiada hentinya. Terimakasih atas ridho dan segala doa yang engkau panjatkan untuk penulis, menjadi pelindung yang lebih luas dari langit dan selalu mengiri kemanapun penulis berada untuk penulis selalu berteduh dibawahnya. Semoga lama hidupmu disini, melihatku berjuang sampai akhir.
2. Panutan terbaik penulis, Ayahanda Suryono. Terimakasih sudah berjuang keras untuk kehidupan penulis. Terimakasih atas segala pengorbanan, didikan, pesan-pesan dan dukungan kepada penulis yang penuh keikhlasan tak terhingga. Semoga ayah panjang umur dan sehat selalu.
3. Ibu Nur Hafifah Matondang, S.Kom., MM., MTI. Selaku Dosen Pembimbing 1, dan Ibu Nindy Irzavika, S.SI., M.T. Selaku Dosen Pembimbing 2, yang telah bersedia meluangkan waktu dan tenaga untuk membimbing dan mengarahkan penulis dalam proses pengerjaan skripsi.
4. Bapak dan Ibu Dosen Fakultas Ilmu Komputer UPN Veteran Jakarta yang telah memberikan ilmu, serta arahan selama masa perkuliahan.

5. Keluarga Besar, Terkhusus Enti, Kukung, dan Mbah Jawa yang senantiasa memberikan doa dan dukung tak terhenti.
6. Sahabat penulis, Kartika, Arya, Rahma, dan Nesa, yang telah setia memberikan dukungan, semangat, kerja sama, dan doa. Serta membantu baik dengan waktu, tenaga, ataupun moril, agar skripsi dapat selesai secara lancar dan baik.
7. Seluruh pihak yang telah berkontribusi dalam membantu penulis yang tidak bisa disebutkan satu per satu, terimakasih atas doa dan dukungan yang sangat berharga.
8. Faiza Shafiqa Putri Setiawan, terimakasih telah kebersamai dan menemani proses penulis yang tidak mudah, menjadi pendengar yang baik untuk keluhan ataupun suka dan duka. Terimakasih atas segala cinta, dukungan, semangat dan doa untuk penulis. Terimakasih atas seluruh kepercayaan yang senantiasa engkau tuturkan. Terimakasih telah selalu ada dan menjadi bagian indah dari perjalanan hidup penulis. Semoga selalu diberkahi dengan segala yang terbaik untukmu, Fiqa.

Penulis menyadari bahwa dalam penulisan penyusunan Tugas Akhir Skripsi ini masih terdapat banyak kekurangan dan kesalahan. Maka dari itu, penulis akan sangat terbuka atas saran dan kritik yang membangun untuk kemudian hari. Semoga skripsi ini dapat bermanfaat untuk penulis sendiri maupun pembaca.

Jakarta, 06 Desember 2024



Daffa Malaeka Adyori

DAFTAR ISI

LEMBAR PERSETUJUAN.....	iii
PERNYATAAN ORISINALITAS.....	iv
PERNYATAAN PERSETUJUAN	v
LEMBAR PENGESAHAN	vi
ABSTRAK	vii
<i>ABSTRACT</i>	viii
KATA PENGANTAR	ix
DAFTAR ISI	xi
DAFTAR GAMBAR.....	xiii
DAFTAR TABEL	xiv
DAFTAR SIMBOL	xv
BAB I 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 Luaran yang Diharapkan	4
1.7 Sistematika Penulisan.....	4
BAB II LANDASAN TEORI	6
2.1 Kecelakaan Lalu Lintas.....	6
2.2 Machine Learning	7
2.3 K-Means Clustering	8
2.4 Clustering	9
2.5 Data Pre-Processing	10
2.6 Evaluasi Model	11
2.7 Python	13
2.8 Metode Elbow	14
2.9 Streamlit	15
2.10 Penelitian Terdahulu	15
BAB III METODOLOGI PENELITIAN	20
3.1 Alur Penelitian.....	20
3.2 Tahapan Penelitian	20
3.2.1 Identifikasi Masalah.....	21
3.2.2 Pemahaman Data	21
3.2.3 Preprocessing Data.....	21
3.2.4 Identifikasi Masalah.....	22

3.2.5 Evaluasi Model	22
3.2.6 Analisis dan Visualisasi Hasil Clustering	22
3.3 Perangkat Penelitian	23
3.3.1 Hardware	23
3.3.2 Software	23
3.4 Jadwal Rencana Penelitian	24
BAB IV HASIL DAN PEMBAHASAN	38
4.1 Pemahaman Data	25
4.2 Pre-Processing Data	26
4.2.1 Pemilihan Atribut	26
4.2.2 Normalisasi Data	27
4.3 Penerapan K-Means Clustering	28
4.3.1 Identifikasi Masalah	28
4.3.2 Menentukan Nilai Centroid Awal	29
4.3.3 Menghitung Jarak Objek ke Centroid	30
4.3.4 Menghitung Ulang Posisi Centroid	30
4.3.5 Mengulangi Tahapan Hingga Data Stabil	30
4.3.6 Hasil Centroid Akhir	31
4.3.5 Identifikasi Masalah	30
4.4 Evaluasi Model	32
4.5 Analisis dan Visualisasi Data	33
4.5.1 Analisis Cluster	33
4.5.2 Visualisasi Data	45
4.5.3 Tampilan Halaman Streamlit	50
BAB V PENUTUP	63
5.1 Kesimpulan	54
5.2 Saran	55
DAFTAR PUSTAKA	56
LAMPIRAN	60

DAFTAR GAMBAR

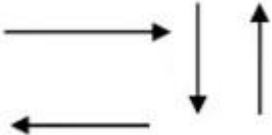
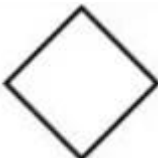
Gambar 2.1. Penerapan Aplikasi Python (Suharto, 2023)	26
Gambar 2.2. Grafik Elbow Method	27
Gambar 3.1. Diagram Alur Penelitian.....	33
Gambar 4.1. Data Kejadian Kecelakaan di 13 Ruas Tol Tahun 2020-2023	38
Gambar 4.4. Hasil Evaluasi Davies Bouldin Index	46
Gambar 4.5. Pie chart Distribusi Cluster.....	46
Gambar 4.6. Visualisasi Data Box Plot Jumlah_Kecelakaan antar Cluster.....	47
Gambar 4.7. Visualisasi Data Box Plot Jumlah_Meninggal antar Cluster	48
Gambar 4.8. Visualisasi Data Box Plot Jumlah_Meninggal antar Cluster	49
Gambar 4.9. Visualisasi Data Box Plot Jumlah_Kendaraan_Rusak_Berat antar Cluster	50
Gambar 4.9. Pemilihan Jenis Filtrasi	57
Gambar 4.10. Filter Cluster	57
Gambar 4.11. Filter Tahun.....	58
Gambar 4.12. Filter Ruas_Tol	58
Gambar 4.13. Hasil Filtrasi Tahun dan Ruas Tol	59
Gambar 4.14. Tabel Filtrasi	59
Gambar 4.15. Line Plot Jumlah_Kecelakaan per Tahun	60
Gambar 4.16. Jumlah_Kecelakaan per Ruas Tol.....	61
Gambar 4.17. Tampilan Halaman Menu Analisis Cluster 1	62
Gambar 4.18. Tampilan Halaman Menu Analisis Cluster 2	63
Gambar 4.19. Tampilan Halaman Menu Visualisasi Data 1	64
Gambar 4.20. Tampilan Halaman Menu Visualisasi Data 2	64
Gambar 4.21. Tampilan Halaman Menu Visualisasi Data 3	64

DAFTAR TABEL

Tabel 2.1 Penelitian Terdahulu	28
Tabel 3.1. Jadwal Rencana Penelitian	37
Tabel 4.1. Rincian Dataset	38
Tabel 4.2. Centroid Awal.....	42
Tabel 4.3. Hasil Centroid Akhir.....	44
Tabel 4.4. Kode Clustering K-Means.....	44
Tabel 4.5. Pendistribusian Data.....	45
Tabel 4.6. Wilayah Cluster 1.....	52
Tabel 4.7. Wilayah Cluster 2.....	54
Tabel 4.8. Wilayah Cluster 3.....	56

DAFTAR SIMBOL

1. Flowchart

Simbol	Nama	Keterangan
	<i>Flow</i>	Simbol untuk menghubungkan antar prosedur / proses
	<i>Terminal</i>	Simbol untuk permulaan atau akhir suatu program
	<i>Document</i>	Simbol untuk menyatakan bahwa masukan berasal dari dokumen fisik
	<i>Input</i>	Simbol untuk menyatakan proses masukan tanpa tergantung peralatan
	<i>Decision</i>	Simbol untuk kondisi yang akan menghasilkan beberapa kemungkinan
	<i>Process</i>	Simbol untuk menunjukkan pengolahan yang dilakukan