

**ANALISIS PENGARUH KONDISI CUACA DAN JALAN
TERHADAP KECELAKAAN LALU LINTAS
MENGUNAKAN *K-MEANS***

SKRIPSI

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



OLEH :

MUHAMMAD ROMADON

09011182025015

JURUSAN SISTEM KOMPUTER

FAKULTAS ILMU KOMPUTER

UNIVERSITAS SRIWIJAYA

2024

LEMBAR PENGESAHAN

**ANALISIS PENGARUH KONDISI CUACA DAN JALAN TERHADAP
KECELAKAAN LALU LINTAS MENGGUNAKAN *K-MEANS***

SKRIPSI



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

Oleh :


**MUHAMMAD ROMADON
09011182025015**

Indralaya, ⁹ Desember 2024

Ketua Jurusan Sistem Komputer



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

Pembimbing Tugas Akhir


Dr. Rossi Passarella, M.Eng
NIP. 197811062010121004

LEMBAR PERSETUJUAN

Telah diuji dan lulus pada :

Hari : Jum'at

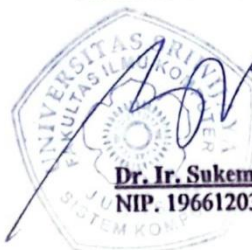
Tanggal : 6 Desember 2024

Tim Penguji :

1. Ketua : Ahmad Fali Oklilas, M.T
2. Penguji : Sutarno, M.T
3. Pembimbing : Dr.Rossi Passarella,M.Eng

Handwritten signatures and dates for the examiners. The first signature is for Ahmad Fali Oklilas, M.T. The second signature is for Sutarno, M.T, with the date '7/12' written below it. The third signature is for Dr. Rossi Passarella, M.Eng.

Mengetahui, *W/rv/m*
Ketua Jurusan Sistem Komputer



W/rv/m
Dr. Ir. Sukemi, M.T.
NIP. 196612032006041001

HALAMAN PERNYATAAN

Yang bertanda tangan di bawah ini :

Nama : Muhammad Romadon
NIM : 09011182025015
Judul : Analisis Pengaruh Kondisi Cuaca Dan Jalan Terhadap
Kecelakaan Lalu Lintas Menggunakan *K-Means*

Hasil pengecekan *Software Turnitin* : 4%

Menyatakan bahwa Laporan Skripsi saya merupakan hasil karya sendiri bukan hasil penjiplakan atau plagiat. Apabila ditemukan unsur penjiplakan atau plagiat dalam Laporan Skripsi ini, maka saya bersedia menerima sanksi akademik dari Universitas Sriwijaya. Demikian, pernyataan ini saya buat dalam keadaan sadar dan tidak dipaksakan.



Indralaya, Desember 2024



Muhammad Romadon

KATA PENGANTAR

Assalamu'alaikum Warahmatullahi Wabarakatuh

Puji dan syukur atas kehadiran Allah SWT yang telah melimpahkan rahmat, kasih sayang dan karunia-Nya, sehingga penulis dapat menyelesaikan skripsi dengan judul “Analisis Pengaruh Kondisi Cuaca Dan Jalan Terhadap Kecelakaan Lalu Lintas Menggunakan *K-Means*”.

Dalam laporan ini, penulis menjelaskan mengenai pengaruh kondisi cuaca dan kondisi jalan terhadap kecelakaan lalu lintas. Penelitian dilakukan dengan menggunakan algoritma *K-Means* sebagai metode utama, serta membandingkan hasilnya dengan metode clustering lain seperti *GMM* dan *Spherical K-Means*. Penulis berharap agar tulisan ini dapat bermanfaat bagi orang banyak.

Selama penulisan Skripsi ini, penulis banyak mendapatkan ide, bantuan, serta saran dari semua pihak, baik secara langsung maupun tak langsung. Oleh karena itu, pada kesempatan ini penulis menyampaikan ucapan terima kasih kepada:

1. Allah SWT yang telah melimpahkan berkah serta nikmat kesehatan dan kesempatan kepada penulis dalam melaksanakan Tugas Akhir.
2. Orang tua, saudara, dan keluarga besar yang telah mendoakan dan memberikan motivasi serta support.
3. Bapak Prof. Dr. Erwin, S.Si., M.Si. selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
4. Bapak Dr. Ir. Sukemi, M.T., selaku Ketua Jurusan Sistem Komputer Fakultas Ilmu Komputer Universitas Sriwijaya.
5. Bapak Dr. Rossi Passarella, M.Eng. selaku Dosen Pembimbing Tugas Akhir yang telah berkenan meluangkan waktunya guna membimbing, memberikan saran dan motivasi serta bimbingan terbaik untuk penulis dalam menyelesaikan Tugas Akhir ini.

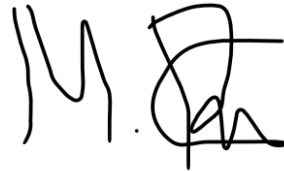
6. Bapak Dr.Ahmad Zarkasih, M.T selaku Dosen Pembimbing Akademik di Jurusan Sistem Komputer Universitas Sriwijaya.
7. Dan semua pihak yang telah membantu.

Penulis menyadari bahwa laporan ini masih sangat jauh dari kata sempurna. Untuk itu kritik dan saran yang membangun sangatlah diharapkan penulis. Akhir kata penulis berharap, semoga Skripsi ini bermanfaat dan berguna bagi khalayak.

Wassalamu'alaikum Warahmatullahi Wabarakatuh.

Indralaya, Desember 2024

Penulis,



Muhammad Romadon
NIM. 09011182025015

ANALISIS PENGARUH KONDISI CUACA DAN JALAN TERHADAP KECELAKAAN LALU LINTAS MENGGUNAKAN *K-MEANS*

MUHAMMAD ROMADON (09011182025015)

Jurusan Sistem Komputer, Fakultas Ilmu Komputer, Universitas Sriwijaya

Email : muhammadromadon734@gmail.com

ABSTRAK

Penelitian ini menganalisis pengaruh cuaca dan kondisi jalan terhadap kecelakaan lalu lintas menggunakan *K-Means*, serta membandingkannya dengan *Gaussian Mixture Model* (GMM) dan *Spherical K-Means*. Data mencakup variabel cuaca, kondisi jalan, dan karakteristik kecelakaan 2018–2023. Hasilnya, cuaca hujan dan jalan basah meningkatkan risiko kecelakaan. GMM unggul dalam evaluasi metrik, tetapi *K-Means* dipilih karena kesederhanaan, efisiensi, dan interpretabilitasnya. Hasil ini mendukung kebijakan peningkatan keselamatan melalui perbaikan infrastruktur dan pengelolaan risiko cuaca.

Kata Kunci: *K-Means*, *GMM*, *Spherical K-Means*, Kecelakaan LaluLintas,Cuaca,Kondisi jalan

ANALYSIS OF THE EFFECT OF WEATHER AND ROAD CONDITIONS ON TRAFFIC ACCIDENTS USING K-MEANS

MUHAMMAD ROMADON (09011182025015)

Department of Computer Systems, Faculty of Computer Science,
Sriwijaya University
Email : muhammadromadon734@gmail.com

ABSTRACT

This study analyzes the effects of weather and road conditions on traffic accidents using K-Means, and compares them with Gaussian Mixture Model (GMM) and Spherical K-Means. The data includes weather variables, road conditions, and accident characteristics from 2018–2023. The results show that rainy weather and wet roads increase the risk of accidents. GMM excels in metric evaluation, but K-Means is chosen because of its efficiency, efficiency, and interpretability. These results support safety improvement policies through infrastructure improvements and weather risk management.

Keywords: K-Means, GMM, Spherical K-Means, Traffic Accidents, Weather, Road Conditions

DAFTAR ISI

LEMBAR PENGESAHAN	i
LEMBAR PERSETUJUAN	ii
HALAMAN PERNYATAAN	iii
KATA PENGANTAR	iv
ABSTRAK	vi
ABSTRACT	vii
DAFTAR ISI	viii
DAFTAR GAMBAR	x
DAFTAR TABEL	xi
DAFTAR LAMPIRAN	xii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	3
1.3 Tujuan	3
1.4 Manfaat	3
1.5 Batasan Masalah	4
1.6 Metodologi Penelitian	4
1.7 Sistematika Penulisan	5
BAB II TINJAUAN PUSTAKA	7
2.1 Kajian Pustaka	7
2.2 Clustering.....	7
2.3 Kecelakaan.....	7
2.4 Lalu Lintas	8
2.5 <i>Machine Learning</i>	8
2.6 <i>K-Means</i>	9
2.7 <i>Gaussian Mixture Model (GMM)</i>	11
2.8 <i>Spherical K-Means</i>	12
BAB III METODELOGI PENELITIAN	15
3.1 Pendahuluan.....	15
3.2 Analisis Data Eksplorasi	15
3.3 Kerangka Penelitian	26
3.3.1 Pengolahan Data	27

3.3.2	Analisis Data Eksplorasi	28
3.3.3	Pengolahan Data.....	28
3.3.4	Tahap Pemodelan	29
3.3.5	Evaluasi Model.....	30
3.3.6	Pemilihan Model	30
3.3.7	Analisis Hasil Cluster.....	31
3.3.8	Kesimpulan.....	32
3.4	Menentukan Alat Penelitian.....	32
BAB IV HASIL DAN PEMBAHASAN		34
4.1	Pengumpulan Data.....	34
4.1	Pengolahan Data	36
4.1.1	Feature Selection	36
4.1.2	Encoding Data	40
4.2	Tahap Pemodelan	41
4.3	Evaluasi Model	43
4.3.2	Davies-Bouldin Index	46
4.3.3	Calinski-Harabasz Index.....	47
4.1	Pemilihan Model.....	48
4.1.1	Alasan Metode <i>K-Mean</i> Bisa Kurang Optimal.....	50
BAB V KESIMPULAN DAN SARAN		61
5.1	Kesimpulan	61
5.2	Saran	62
DAFTAR PUSTAKA.....		63

DAFTAR GAMBAR

Gambar 2. 1 <i>Flowchart Clustering</i>	11
Gambar 3. 1 <i>Distribution of accidents by Weather (2018 – 2023)</i>	20
Gambar 3. 2 <i>Distribution of accidents by Surface Condition (2018 – 2023)</i> .	21
Gambar 3. 3 <i>Distribution of accidents by ACRS Report Type (2018 – 2023)</i> .	23
Gambar 3. 4 <i>Distribution of accidents by Vehicle Year (2018 – 2023)</i>	25
Gambar 3. 5 Kerangka Kerja Penelitian.....	27
Gambar 4. 1 Korelasi Variabel Dengan <i>Heatmap</i>	37
Gambar 4. 2 Visualisasi <i>Scatter Plot K-Means , GMM, Spherical K-Means</i> ...	44

DAFTAR TABEL

Tabel 3. 1 <i>Data Deskriptif</i>	16
Tabel 4. 1 Keterangan Variabel Data.....	34
Tabel 4. 2 Hasil <i>F_Score</i> Masing – Masing Variabel.....	38
Tabel 4. 3 Encoding Data Variable <i>Weather</i> dan <i>Surface Condition</i>	40
Tabel 4. 4 Penentuan Inisial <i>Cluster</i> Menggunakan Metode <i>Elbow</i>	42
Tabel 4. 5 Perbandingan Algoritma Dengan Metode <i>Silhouette coefficient</i>	45
Tabel 4. 6 Perbandingan Algoritma Dengan Metode <i>Davies-Bouldin Index</i>	46
Tabel 4. 7 Perbandingan Algoritma Dengan Metode <i>Calinski-Harabasz Index</i>	48
Tabel 4. 8 Perbandingan Algoritma Dengan Metode <i>silhouette coefficient, Davies- Bouldin Index, Calinski-Harabasz Index</i>	49
Tabel 4. 9 Jumlah Data Pada Tiap <i>Cluster K-Means</i>	52
Tabel 4. 10 Hasil Cluster 1 <i>K-Means</i>	53
Tabel 4. 11 Hasil Cluster 2 <i>K-Means</i>	54
Tabel 4. 12 Hasil Cluster 3 <i>K-Means</i>	57
Tabel 4. 13 Hasil Cluster 4 <i>K-Means</i>	58
Tabel 4. 14 Ringkasan Hasil Mayoritas Untuk Setiap Klaster.....	59

DAFTAR LAMPIRAN

Lampiran 1 Form Perbaikan Pembimbing

Lampiran 2 Form Perbaikan Penguji

Lampiran 3 Hasil Cek Plagiarisme di Turnitin Halaman Judul

Lampiran 4 Hasil Cek Plagiarisme di Turnitin

Lampiran 5 Surat Keterangan Pengecekan Similarity

BAB I

PENDAHULUAN

1.1 Latar Belakang

Kecelakaan lalu lintas adalah suatu Observasi dimana kendaraan bermotor bertabrakan dengan benda lain dan menimbulkan kerusakan. Terkadang kecelakaan ini dapat mengakibatkan cedera atau kematian pada manusia atau hewan [1]. Banyak kecelakaan yang dapat mengakibatkan hilangnya nyawa, cedera serius, kerugian materil dan kerusakan kendaraan. Oleh karena itu, pemahaman menyeluruh tentang faktor-faktor yang mempengaruhi jumlah kecelakaan di jalan raya penting untuk merancang tindakan pencegahan yang efektif.

Salah satu faktor yang dapat berkontribusi terhadap kecelakaan lalu lintas adalah cuaca. Kondisi seperti hujan lebat, kabut tebal atau salju dapat mengurangi jarak pandang, mengubah perilaku pengemudi dan meningkatkan risiko kecelakaan. Di sisi lain, kondisi jalan seperti aspal rusak, berlubang, atau perbaikan jalan juga dapat mempengaruhi angka kecelakaan.

Oleh karena itu, Penelitian ini bertujuan untuk mengetahui pengaruh kedua faktor tersebut terhadap kecelakaan lalu lintas menggunakan *K-Means Clustering*, sebuah teknik data mining yang bertujuan untuk mencocokkan data dengan pola serupa [2]. Objek-objek dalam suatu cluster memiliki sifat yang serupa dan berbeda dengan cluster lainnya. Penyortiran tidak dilakukan secara manual melainkan dengan algoritma. Oleh karena itu, Clustering sangat berguna dan dapat menemukan kelompok yang tidak diketahui dalam statistik [3].

K-Means adalah algoritma populer yang membagi data menjadi beberapa cluster berdasarkan kemiripan pola. Namun, metode ini memiliki keterbatasan, seperti sensitivitas terhadap centroid awal dan kesulitan menangani data yang kompleks. Untuk mengatasi kelemahan ini, penelitian juga menggunakan metode pembandingan, yaitu *Gaussian Mixture Model (GMM)* dan *Spherical K-Means*.

Gaussian Mixture Model (GMM) merupakan teknik probabilistik yang lebih fleksibel dibandingkan K-Means karena mampu menangani data dengan distribusi yang tumpang tindih dan memberikan probabilitas keanggotaan untuk setiap cluster. Metode ini menggunakan prinsip campuran distribusi Gaussian untuk memodelkan data dengan distribusi yang kompleks [4][5]. Selain itu, GMM memberikan fleksibilitas tambahan dalam menangani data berdimensi tinggi dan memungkinkan analisis yang lebih mendalam terhadap pola hubungan antar variabel.

Spherical K-Means (SKM) adalah varian dari *K-Means* yang menggunakan cosine similarity sebagai ukuran kedekatan antar data, sehingga cocok untuk data berdimensi tinggi seperti teks atau data *vektor term-weight* [6][7]. Teknik ini memproyeksikan data ke ruang bola berdimensi tinggi, memastikan setiap data memiliki panjang unit. Keunggulan metode ini meliputi efisiensi dalam menangani data teks yang jarang dan pengelompokan berdasarkan orientasi dokumen dibandingkan panjang dokumen [6]. Selain itu, SKM mendukung pengembangan lebih lanjut seperti algoritma akselerasi untuk meningkatkan efisiensi [7] dan adaptasi untuk skenario pengelompokan multi-cluster [8].

Dalam "*Analysis of Traffic Flow Behavior at the Entrance of Transboundary Rivers and Oceans in Different Seasons: Case Study of the Shanghai Yangtze River*" oleh Ting Zhang, Feng Chen, Yadi Huang, Mingtao Song, dan Xiao Hu [9]. Hasil penelitian menunjukkan bahwa kondisi lalu lintas mempunyai pengaruh yang signifikan terhadap perilaku mengikuti kendaraan dan perubahan pola lalu lintas tergantung arus lalu lintas. Jika jarak antara pejalan kaki di lalu lintas padat dan lalu lintas padat kurang dari jarak aman yang disyaratkan secara hukum, hal ini menimbulkan risiko kecelakaan yang lebih besar. Studi tersebut juga mengungkapkan bahwa cuaca berdampak pada perkembangan musim, dengan kemajuan yang lebih besar terlihat pada kondisi hujan dan bersalju dibandingkan dengan cuaca cerah. Selain itu, ketinggian pintu masuk terowongan mempunyai dampak signifikan terhadap kecepatan lalu lintas, yang menyoroti pentingnya mempertimbangkan kemiringandalam desain dan pengelolaan.

Oleh karena itu, penelitian ini bertujuan untuk mengetahui pengaruh

cuaca dan kondisi jalan terhadap jumlah kecelakaan lalu lintas. data diekstraksi dari Data.Gov[10] Dengan menggunakan metode *K-Means* , kami berharap dapat menggambarkan struktur cluster dan memungkinkan kami untuk memahami sepenuhnya hubungan antara faktor- faktor ini. Kami yakin temuan penelitian ini akan memberikan informasi yang berguna bagi penyelenggara, pemerintah, dan masyarakat umum tentang bagaimana faktor-faktor tersebut berkontribusi terhadap angka kecelakaan lalu lintas

Berdasarkan uraian diatas, maka peneliti akan membahas tentang **Analisis Pengaruh Kondisi Cuaca Dan Jalan Terhadap Kecelakaan Lalu Lintas Menggunakan *K-Means***

1.2 Rumusan Masalah

Rumusan masalah dari Tugas Akhir ini dirumuskan sebagai berikut :

1. Bagaimana memetakan kecelakaan berubah sehubungan dengan variasi kondisi cuaca dan kondisi jalan raya dalam masing-masing klaster?
2. Apakah ada perbedaan signifikan dalam pengaruh variasi kondisi cuaca dan kondisi jalan raya antara klaster-klaster?

1.3 Tujuan

Adapun tujuan dari penelitian tugas akhir ini sebagai berikut:

1. Memahami kontribusi kondisi cuaca dan kondisi jalan raya terhadap kecelakaan lalu lintas berdasarkan hasil klaster dari algoritma utama *K-Means*.
2. Menggunakan algoritma utama *K-Means* serta metode pembandingan, yaitu *GMM* dan *Spherical K-Means*, untuk mengelompokkan data kondisi cuaca, kondisi jalan raya, dan kecelakaan lalu lintas.
3. Mengevaluasi kemampuan algoritma *K-Means* sebagai metode utama dalam menganalisis pola kecelakaan lalu lintas, dibandingkan dengan pendekatan *GMM* dan *Spherical K-Means* sebagai pembandingan.

1.4 Manfaat

Adapun manfaat dari penelitian tugas akhir ini sebagai berikut:

1. Memberikan gambaran mengenai kontribusi kondisi cuaca dan kondisi jalan raya terhadap kecelakaan lalu lintas dengan menggunakan algoritma utama *K-Means* dan metode pembandingan *GMM* serta *Spherical K-Means*.

2. Mendukung identifikasi wilayah dengan risiko tinggi kecelakaan lalu lintas berdasarkan pola kluster yang dihasilkan oleh ketiga metode tersebut, sehingga dapat dirumuskan tindakan preventif yang lebih tepat sasaran.
3. Menyediakan informasi dan hasil analisis yang dapat digunakan sebagai dasar untuk penelitian lebih lanjut dalam bidang keselamatan lalu lintas dan analisis risiko, khususnya dalam mengaplikasikan metode *K-Means*, *GMM*, dan *Spherical K-Means*.

1.5 Batasan Masalah

Berdasarkan rumusan masalah diatas, akan dilakukan batasan masalah agar Tugas Akhir ini tetap terarah.

1. Penelitian ini menggunakan Algoritma *K-Means* untuk melakukan klustering.
2. Penelitian ini difokuskan pada dua variabel, yaitu kondisi cuaca dan kondisi jalan raya terhadap kecelakaan lalu lintas.
3. Penelitian ini menggunakan dataset Data.Gov
4. Menggunakan algoritma pembanding yaitu *Gaussian Mixture Model (GMM)* dan *Spherical K-Means*
5. Data yang digunakan dari 2018 sampai 2023

1.6 Metodologi Penelitian

Tahapan dari metodologi penulisan yang dilakukan dalam Tugas Akhir ini adalah :

1. Rumusan Masalah

Tahapan ini akan merumuskan masalah yang dihadapi dalam melakukan pre-processing dataset, pemilihan fitur serta penyelesaiannya

2. Studi Pustaka

Tahapan ini bertujuan untuk mencari literatur yang sesuai dalam rangka menunjang penelitian yang dilakukan dengan cara mencari literatur menggunakan kata kunci spesifik dalam penelitian ini.

3. Perancangan

Tahapan ini merupakan proses perancangan dari sistem yang akan dibangun serta membahas runtutan proses dari pembuatan program untuk persiapan dataset, pemilihan fitur dan klustering.

4. Pengujian

Tahapan ini merupakan proses pengujian dari sistem yang dibangun sesuai dengan rancangan yang dibuat.

5. Validasi

Tahapan ini diperlukan untuk mengecek apakah sistem telah melakukan klustering dari dataset yang diberikan.

6. Evaluasi

Tahap ini bertujuan untuk mengetes sistem yang dibuat menggunakan data tes yang telah disediakan agar dapat mengetahui seberapa akurat sistem klustering data.

7. Analisis

Tahapan ini adalah proses menganalisis hasil dari sistem yang dibangun. Analisis ditampilkan dalam rangka mengukur seberapa baik sistem yang dibangun dalam menjalankan fungsinya.

8. Kesimpulan dan Saran

Tahapan ini adalah proses memberikan kesimpulan berdasarkan hasil dari analisis dari sistem yang dibangun dan saran bagi penelitian yang akan datang.

1.7 Sistematika Penulisan

Sistematika penulisan yang dilakukan dalam Tugas Akhir ini adalah :

BAB 1 PENDAHULUAN

Pada bab pertama ini berisi mengenai uraian singkat tentang latar belakang, tujuan, manfaat, batasan masalah, metodologi penelitian serta sistematika penulisan oleh penulis.

BAB II TINJAUAN PUSTAKA

Pada bab kedua ini menjelaskan mengenai landasan teori yang dapat mendukung pembahasan dari penelitian ini, dasar teori tersebut meliputi setiap aspek yang akan digunakan dalam penelitian ini.

BAB III METODOLOGI PENELITIAN

Pada bab ini menjelaskan mengenai proses dan berbagai rangkaian kegiatan yang dilakukan dalam penelitian. Penelitian ini diawali dengan studi literatur dan diikuti dengan konsultasi, pengujian simulasi program yang telah dirancang yaitu Algoritma *K-Means* dan juga proses analisis data pada dataset Data.Gov serta melakukan Analisa dan menyimpulkan hasil yang telah didapat.

BAB IV HASIL DAN ANALISIS

Pada bab ini akan memaparkan hasil pengujian data dari Data.Gov mengenai kepadatan lalu lintas menggunakan Algoritma *K-Means* dan menjelaskan Analisa yang dihasilkan melalui penelitian dan pengujian yang telah dilakukan.

BAB V KESIMPULAN

Pada bab ini akan dilakukannya pengumpulan seluruh hal yang telah didapat dari bab- bab sebelumnya dan menyimpulkan hal tersebut dari hasil serta analisa terhadap penelitian yang telah dilakukan.

DAFTAR PUSTAKA

- [1] A. D. Saputra, “Studi Tingkat Kecelakaan Lalu Lintas Jalan di Indonesia Berdasarkan Data KNKT (Komite Nasional Keselamatan Transportasi) dari Tahun 2007- 2016,” *Warta Penelitian Perhubungan*, vol. 29, no. 2, p. 179, Jul. 2018, doi: 10.25104/warlit v29i2.557.
- [2] L. Wulandari and B. Olga Yogantara, “Algorithm Analysis of *K-Means* and *Fuzzy C-Means* for *Clustering* Countries Based on Economy and Health,” vol. 15, no. 2, pp. 1979–276, 2022, doi: 10.30998/faktorexacta v15i2.12106.
- [3] T. Soni Madhulatha, “*An Overview On Clustering Methods*,” vol. 2, no. 4, pp. 719– 725, 2012, [Online]. Available: www.iosrjen.org
- [4] L. Zhao *et al.*, “Adaptive parameter estimation of GMM and its application in clustering,” *Future Generation Computer Systems*, vol. 106, pp. 250–259, May 2020, doi: 10.1016/j.future.2020.01.012.
- [5] Y. Fu, X. Liu, S. Sarkar, and T. Wu, “Gaussian mixture model with feature selection: An embedded approach,” *Comput Ind Eng*, vol. 152, Feb. 2021, doi: 10.1016/j.cie.2020.107000.N.
Pandiangnan and M. L. Buono, “*K-Means Clustering* Dalam Mengelompokkan Nilai Akhir Skripsi Mahasiswa”, *MJTI*, vol. 1, no. 02, pp. 42-46, Sep. 2019.
- [6] E. Schubert, A. Lang, and G. Feher, “Accelerating Spherical *k-Means*,” Jul. 2021, doi: 10.1007/978-3-030-89657-7_17.
- [7] K. Hornik, I. Feinerer, M. K. Wu, W. Wien, and C. Buchta, “*Journal of Statistical Software Spherical k-Means Clustering*,” 2012. [Online]. Available: <http://www.jstatsoft.org/>
- [8] V. Tunali *et al.*, “Article in The International Arab Journal of

Information Technology,” 2016. [Online]. Available: <https://www.researchgate.net/publication/273632084>

- [9] T. Zhang, F. Chen, Y. Huang, M. Song, and X. Hu, “Analysis of Car-Following Behaviors under Different Conditions on the Entrance Section of Cross-River and Cross-Sea Tunnels: A Case Study of Shanghai Yangtze River Tunnel,” *Int J Environ Res Public Health*, vol. 19, no. 19, Oct. 2022, doi: 10.3390/ijerph191911975.
- [10] data.montgomerycountymd.gov, “Crash Reporting - Drivers Data”, 2020, [Online]. Tersedia: <https://catalog.data.gov/dataset/crash-reporting-drivers-data>
- [11] S. Dewi, S. Defit, and Y. Yuhandri, “Akurasi Pemetaan Kelompok Belajar Siswa Menuju Prestasi Menggunakan Metode *K-Means*,” *Jurnal Sistim Informasi dan Teknologi*, pp. 28–33, Mar. 2021, doi: 10.37034/jsisfotek.v3i1.40.
- [12] Irwansyah, E., & Faisal, M. (2015). *Advanced Clustering: Teori dan Aplikasi*. Deepublish.
- [13] I. R. Hardini, “A Survey on *Machine Learning* and IoT,” *ITEJ Desember-2019*, vol. 4, no. 2.
- [14] J. Xie, F. R. Yu, T. Huang, R. Xie, J. Liu, and Y. Liu, “A Survey of *Machine Learning* Techniques Applied to Software Defined Networking (SDN): Research Issues and Challenges,” *IEEE Commun. Surv. Tutorials*, vol. PP, no. c, p. 1, 2018.
- [15] S. Jaspers, E. De Troyer, and M. Aerts, “*Machine Learning* techniques for the automation of literature reviews and systematic reviews in EFSA,” *EFSA Support. Publ.*, vol. 15, no. 6, 2018.
- [16] R. Rahmattullah, I. Indwiarti, and A. A. Rohmawati, “*Clustering* Harga Rumah: Perbandingan *Model K-Means* dan *Gaussian Mixture Model*,” *eProceedings of Engineering*, vol. 10, no. 3,

2023.

- [17] S. Rostianingsih, R. Adipranata, and F. Setiawan Wibisono, “Adaptive Background Dengan Metode Gaussian Mixture Models Untuk Real-Time Tracking.”
- [18] Agusta, Y. 2007. *K-Means* - Penerapan, Permasalahan dan Metode Terkait. Jurnal Sistem dan Informatika Vol. 3 (Februari 2007):47-60