

**ANALISIS PERBANDINGAN METODE *K-MEANS* DAN METODE
K-MEDOIDS PADA *CLUSTERING* DATA PELANGGARAN LALU LINTAS**

TUGAS AKHIR

Oleh:

NATASHA WAHYUNI

1710817320008



PROGRAM STUDI TEKNOLOGI INFORMASI

FAKULTAS TEKNIK

UNIVERSITAS LAMBUNG MANGKURAT

BANJARMASIN

2023

**ANALISIS PERBANDINGAN METODE *K-MEANS* DAN METODE
K-MEDOIDS PADA *CLUSTERING* DATA PELANGGARAN LALU LINTAS**

Tugas Akhir

Diajukan Untuk Memenuhi Salah Satu Syarat
Sarjana Strata-1 Teknologi Informasi

Oleh:

NATASHA WAHYUNI

1710817320008



**PROGRAM STUDI TEKNOLOGI INFORMASI
FAKULTAS TEKNIK
UNIVERSITAS LAMBUNG MANGKURAT
BANJARMASIN, 2023**

LEMBAR PERNYATAAN

Yang bertanda tangan di bawah ini,

Nama : Natasha Wahyuni

NIM : 1710817320008

Fakultas : Teknik

Prodi : Teknologi Informasi

Judul Tugas Akhir : Analisis Perbandingan Metode *K-Means* Dan Metode *K-Medoids* Pada Clustering Data Pelanggaran Lalu Lintas

Pembimbing Utama : Nurul Fathanah Mustamin, S.Pd., M.T

Pembimbing Pendamping : Eka Setya Wijaya, S.T., M.Kom

Dengan ini saya menyatakan bahwa dalam Tugas Akhir ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar akademik di suatu perguruan tinggi, dan sepanjang pengetahuan saya, juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar rujukan.

Banjarmasin, Agustus 2023

Penulis,

Natasha Wahyuni

LEMBAR PENGESAHAN

LEMBAR PENGESAHAN

SKRIPSI PROGRAM STUDI S-1 TEKNOLOGI INFORMASI

Analisis Perbandingan Metode *K-Means* dan Metode *K-Medoids* Pada Clustering Data
Pelanggaran Lalu Lintas

Oleh

Natasha Wahyuni (1710817320008)

Telah dipertahankan di depan Tim Penguji pada 26 Juni 2023 dan dinyatakan

L U L U S

Komite Penguji :

Ketua : Andreyan Rizky Baskara, S.Kom., M.Kom.
NIP 199307032019031011

Anggota 1 : Muti'a Maulida, S.Kom., M.T.I.
NIP 198810272019032013

Anggota 2 : Andry Fajar Zulkarnain, S.ST., M.T
NIP 199007272019031018

Pembimbing : Nurul Fathanah Mustamin, S.Pd., M.T.
Utama NIP 199110252019032018

Pembimbing : Eka Setya Wijaya, S.T., M.Kom
Pendamping NIP 198205082008011010



Banjarbaru, ... 15 AUG 2023
diketahui dan disahkan oleh:

Wakil Dekan Bidang Akademik
Fakultas Teknik ULM,



Koordinator Program Studi
S-1 Teknologi Informasi,


Dr. Ir. Yustena Sari, S.Kom., M.Kom.
NIP 198411202015042002

PERSETUJUAN TUGAS AKHIR

ANALISIS PERBANDINGAN METODE *K-MEANS* DAN METODE
K-MEDOIDS PADA CLUSTERING DATA PELANGGARAN LALU LINTAS

OLEH
NATASHA WAHYUNI
NIM.1710817320008

Telah diperiksa dan terpenuhi semua persyaratan akademik, administrasi, dan
disetujui untuk dipertahankan di hadapan dewan penguji

Banjarmasin, 1 Agustus 2023

Pembimbing Utama



Nurul Fathanah Mustamin, S.Pd., M.T
NIP. 19911025 201903 2 018

Pembimbing Pendamping



Eka Setya Wijaya, S.T., M.Kom
NIP. 19820508 200801 1 010

ABSTRAK

Berdasarkan data yang di informasikan dari halaman website resmi SIPP (Sistem Informasi Penelusuran Perkara) Pengadilan Negeri Banjarmasin. Data pelanggaran lalu lintas yang diperoleh pertahun 2020-2022 bulan Oktober-November. Pelanggaran lalu lintas terbanyak terjadi ditahun terakhir 2022. Hal ini menjadi perbandingan bahwa di kota Banjarmasin selama kurang lebih 3(tiga) tahun terakhir untuk kasus pelanggaran lalu lintas mengalami peningkatan setiap tahunnya. Penelitian ini bertujuan mengklasterisasi data pelanggaran lalu lintas di kota Banjarmasin menggunakan perbandingan metode *k-means* dan *k-medoids*. Dilakukan percobaan menggunakan *cluster* 2 sampai *cluster* 10 dengan uji validitas setiap *cluster* menggunakan parameter *davies bouldin index* pada aplikasi *rapidminer*. Hasilnya Jumlah *k* paling baik pada metode *k-means* adalah $k=10$ dengan nilai DBI sebesar 0.018, sedangkan metode *k-medoids* terletak pada $k=2$ dengan nilai DBI sebesar 0.421. Diperoleh metode *K-Means* lebih baik dibandingkan metode *K-Medoids* dalam mengelompokkan data pelanggaran lalu lintas, dengan nilai performa 0,018.

Kata Kunci : Pelanggaran Lalu Lintas, *Clustering*, *K-Means*, *K-Medoids*, *Davies Bouldin Index*, *Rapidminer*

ABSTRACT

Based on the data that was informed from the official website page of the SIPP (Sistem Informasi Penelusuran Perkara) Pengadilan Negeri Banjarmasin. Traffic violation data obtained for the 2020-2022 year October-November. The most traffic violations occurred in the last year of 2022. This is a comparison that in the city of Banjarmasin for the last 3 (three) years traffic violation cases have increased every year. This study aims to cluster data on traffic violations in the city of Banjarmasin using a comparison of the k-means and k-medoids methods. Experiments were carried out using clusters 2 to 10 with a validity test for each cluster using the davies bouldin index parameter in the rapidminer application. The result is that the best number of k in the k-means method is k=10 with a DBI value of 0.018, while the k-medoids method lies in k=2 with a DBI value of 0.421. It was found that the K-Means method was better than the K-Medoids method in classifying traffic violation data, with a performance value of 0.018.

Keywords: *Traffic Violations, Clustering, K-Means, K-Medoids, Davies Bouldin Index, Rapidminer*

HALAMAN PERSEMBAHAN

Penulis mempersembahkan Tugas Akhir ini kepada :

1. Papah, Mama, dan Adik-adik ku yang sudah membantu memberikan semangat, dukungan moral dan materi serta mendoakan saya dalam keberlangsungan penyelesaian Tugas Akhir ini.
2. Ibu Nurul Fathanah Mustamin, S.Pd., M.T selaku Dosen Pembimbing Utama yang selalu menyempatkan waktu untuk memberikan bimbingan, arahan dan dukungan kepada penulis dari awal sampai penyelesaian Tugas Akhir ini.
3. Bapak Eka Setya Wijaya, S.T., M.Kom selaku Dosen Pembimbing Pendamping yang selalu menyempatkan waktu untuk memberikan bimbingan, arahan dan dukungan kepada penulis dari awal sampai penyelesaian Tugas Akhir ini.
4. Ibu Ir. Yuslena Sari, S.Kom., M.Kom. selaku Koordinator Program Studi Teknologi Informasi dan Dosen Pembimbing Akademik yang telah memberikan banyak sekali dukungan terkait penyelenggaraan perkuliahan hingga Tugas Akhir ini dapat diselesaikan dengan baik.
5. Seluruh Dosen beserta Staf Program Studi Teknologi Informasi yang turut membantu dan mengarahkan dalam penyelesaian Tugas Akhir ini.
6. Teman-teman penulis Diza, Hafiza, dan Vicka yang telah mendengarkan keluh kesah penulis serta selalu memberikan saran dan dukungan dalam menyelesaikan Tugas Akhir ini.
7. Teman-teman penulis seluruh angkatan 17 yang telah memberikan saran dan dukungan dalam menyelesaikan Tugas Akhir ini.

KATA PENGANTAR

Puji syukur kehadiran Allah SWT yang telah melimpahkan rahmat dan hidayah-nya kepada penulis sehingga dapat menyelesaikan Tugas Akhir dengan judul “**Analisis Perbandingan Metode *K-Means* dan Metode *K-Medoids* Pada Clustering Data Pelanggaran Lalu Lintas**”. Tugas Akhir ini disusun untuk memenuhi salah satu syarat untuk memperoleh gelar Sarjana Strata-1 Teknologi Informasi di Fakultas Teknik Universitas Lambung Mangkurat, Banjarmasin. Dalam kesempatan ini penulis sangat bersyukur dan mengucapkan banyak terimakasih kepada semua pihak yang membantu, terutama kepada:

1. Rektor Universitas Lambung Mangkurat, Bapak Prof. Dr. Ahmad, S.E., M.Si. yang memimpin dan memanajemen jalannya seluruh perkuliahan yang ada di Universitas Lambung Mangkurat.
2. Dekan Fakultas Teknik, Bapak Prof. Dr. Ir. Irphan Fitrian Radam, S.T., M.T., IPU yang telah memberikan layanan terbaik dalam perkuliahan, terkhusus pada pelaksanaan Tugas Akhir di lingkungan Fakultas Teknik.
3. Ketua Program Studi Teknologi Informasi, Ibu Ir. Yuslena Sari, S.Kom., M.Kom. yang telah memberikan arahan dan solusi dalam penyelesaian Tugas Akhir.
4. Pembimbing Utama, Ibu Nurul Fathanah Mustamin, S.Pd., M.T yang selalu menyempatkan waktu untuk memberikan bimbingan, arahan dan dukungan kepada penulis dari awal sampai penyelesaian Tugas Akhir ini.
5. Pembimbing Pendamping, Bapak Eka Setya Wijaya, S.T., M.Kom. yang telah memberikan waktu, bimbingan, dan arahan dalam proses penelitian dan penulisan laporan Tugas Akhir.
6. Dosen-dosen beserta staf Program Studi Teknologi Informasi yang telah mengarahkan dan teman-teman yang membantu dalam proses penyelesaian Tugas Akhir.

Akhir kata, penulis menyampaikan terimakasih kepada semua pihak yang turut membantu dalam penyelesaian laporan Tugas Akhir ini. Harapan dari penyusunan laporan ini adalah, semoga apa yang penulis susun dapat bermanfaat bagi para pembaca dan semua pihak yang membutuhkan. Serta diharapkan dapat menjadi acuan informasi dasar dalam penelitian-penelitian terkait.

Banjarmasin, Agustus 2023

Penulis,

Natasha Wahyuni

DAFTAR ISI

LEMBAR PERNYATAAN	iii
LEMBAR PENGESAHAN.....	iv
PERSETUJUAN TUGAS AKHIR.....	v
ABSTRAK	vi
ABSTRACT	vii
HALAMAN PERSEMBAHAN	viii
KATA PENGANTAR	ix
DAFTAR ISI	xi
DAFTAR TABEL	xiv
DAFTAR GAMBAR	xvi
DAFTAR RIWAYAT HIDUP	xvii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Batasan Masalah	4
1.4 Tujuan Penelitian	5
1.5 Manfaat Penelitian	5
BAB II TINJAUAN PUSTAKA	6
2.1 Landasan Teori.....	6
2.1.1 <i>Machine Learning</i>	6
2.1.2 Metode Clustering	7
2.1.3 K-Means Clustering	9
2.1.4 K-Medoids Clustering	10
2.1.5 Pelanggaran lalu lintas.....	11
2.1.6 Rapid Miner	16

2.1.7 Davies Index Bouldin (DBI).....	16
2.2 Penelitian Terkait.....	19
2.2.1 <i>Analysis of K-Means and K-Medoids's Performance Using Big Data Technology</i>	19
2.2.2 <i>Perbandingan Algoritma K-means Dengan K-Medoids Pada Pengelompokan Armada Kendaraan Truk Berdasarkan Produktivitas</i>	20
2.2.3 <i>Klasterisasi Perkara Pelanggaran Lalu Lintas Menggunakan Algoritma K-Means dan Davies Bouldin Index</i>	21
2.2.4 <i>Clustering Wilayah dan Pelanggaran Berkendara Menggunakan Algoritma K-Means pada Data Satlantas Polres Tasikmalaya Kota</i>	22
2.2.5 <i>Analisa Clustering Pada Data Pelanggaran Lalu Lintas di Pengadilan Negeri Dumai dengan Menggunakan Metode K-Means</i>	22
2.2.6 <i>Comparison of K-Means and K-Medoids Algorithm for Grouping Landslide Prone Areas in West Java Province</i>	23
2.2.7 <i>Perbandingan Algoritma K Means dan K Medoids Untuk Clustering Kelas Siswa Tunagrahita</i>	24
BAB III METODOLOGI PENELITIAN.....	29
3.1 Alat dan Bahan Penelitian	29
3.1.1 Alat Penelitian	29
3.1.2 Bahan Penelitian.....	30
3.2 Alur Penelitian	30
3.2.1 Studi Literatur.....	30
3.2.2 Pengumpulan Data.....	30
3.2.3 Preprocessing Data	32
BAB IV HASIL DAN PEMBAHASAN	38
4.1 Preprocessing Data.....	38
4.2 Selection Data	38
4.3 Cleaning Data.....	39
4.4 Transformation Data	41
4.5 Data Mining	47
4.5.1 Implementasi Metode <i>K-Means</i> Pada Aplikasi <i>Microsoft Excel</i>	47
4.5.2 Implementasi Metode <i>K-Medoids</i> Pada Aplikasi <i>Microsoft Excel</i>	59

4.6 Pengujian Performa metode <i>K-Means</i> dan <i>K-medoids</i> Menggunakan Metode <i>Davies Bouldin Index (DBI)</i>	63
4.6.1 Metode <i>K-Means</i>	63
4.6.2 Metode <i>K-Medoids</i>	70
4.7 Penerapan Algoritma <i>K-Means</i> dan <i>K-Medoids</i> pada Aplikasi <i>Rapidminer Studio</i>	74
4.8 Analisa dan Hasil	79
BAB V KESIMPULAN DAN SARAN.....	85
5.1 Kesimpulan	85
5.2 Saran.....	85
DAFTAR PUSTAKA	86
LAMPIRAN	89

DAFTAR TABEL

Tabel 2.1 Kelebihan dan Kekurangan Metode Clustering	7
Tabel 2.2 Penelitian Terkait	25
Tabel 3.1 Atribut Pada Data Pelanggaran Lalu Lintas	31
Tabel 3.2 Atribut yang Digunakan	33
Tabel 3.3 Keterangan atribut	33
Tabel 3.4 Inisialisasi PasaI	34
Tabel 4.1 Atribut yang Digunakan	39
Tabel 4.2 Data Kosong.....	39
Tabel 4.3 Kesalahan Penulisan (typo)	40
Tabel 4.4 Asal Alamat Diluar Kota Banjarmasin.....	40
Tabel 4.5 Hasil Cleaning Data.....	40
Tabel 4.6 Hasil Konversi Alamat Menjadi Kecamatan	42
Tabel 4.7 Atribut Terbaru.....	42
Tabel 4.8 Inisialisasi Kecamatan	43
Tabel 4.9 Inisialisasi PasaI	43
Tabel 4.10 Keterangan PasaI	43
Tabel 4.11 Data setelah diinisialisasi.....	46
Tabel 4.12 Perbandingan Nilai Performa Metode Clustering	47
Tabel 4.13 Titik Pusat Centroid Awal (K-Means).....	48
Tabel 4.14 Jumlah Data Cluster	51
Tabel 4.15 Pusat Centroid Baru.....	52
Tabel 4.16 Pusat Centroid Baru (K-Means).....	54
Tabel 4.17 Kelompok Cluster Iterasi 3(Terakhir)	55
Tabel 4.18 Anggota Kelompok Cluster K-Means	56
Tabel 4.19 Titik Pusat Medoid Awal	59
Tabel 4.20 Hasil Perhitungan Iterasi 1	60
Tabel 4.21 Pusat Medoid Baru	60
Tabel 4.22 Hasil Perhitungan Iterasi 2	61
Tabel 4. 23 Anggota Kelompok Cluster K-Medoids.....	62

Tabel 4.24 Pusat Centroid Terakhir K-Means.....	63
Tabel 4.25 Hasil Perhitungan Cluster 1	64
Tabel 4.26 Hasil Perhitungan Cluster 2	65
Tabel 4.27 Hasil Perhitungan Cluster 3	65
Tabel 4.28 Hasil Perhitungan Cluster 4	65
Tabel 4.29 Hasil Perhitungan Cluster 5	66
Tabel 4.30 Hasil Perhitungan Cluster 6	66
Tabel 4.31 Hasil Perhitungan Cluster 7	66
Tabel 4.32 Hasil Perhitungan Cluster 8	67
Tabel 4. 33 Hasil Perhitungan Cluster 9	67
Tabel 4.34 Hasil Perhitungan Cluster 10	67
Tabel 4.35 Nilai SSW Cluster	68
Tabel 4.36 Pusat Medoid.....	70
Tabel 4.37 Hasil Perhitungan Cluster 1	71
Tabel 4.38 Hasil Perhitungan Cluster 2	72
Tabel 4.39 Hasil Matric SSB.....	73
Tabel 4.40 Nilai Rasio	73
Tabel 4.41 Perbandingan Nilai DBI Cluster	80
Tabel 4.42 Nilai DBI Terkecil.....	80
Tabel 4.43 Anggota Kelompok Cluster Metode K-Means	81

DAFTAR GAMBAR

Gambar 1.1 Data Pelanggaran Lalu Lintas Oktober-November 2020-2022....	2
Gambar 2.1 Clustering	7
Gambar 3.1 Alur Penelitian.....	30
Gambar 3.2 Potongan Dataset Mentah	31
Gambar 3.3 Tahapan Preprocessing Data	32
Gambar 3.4 Kesalahan Penulisan (typo).....	34
Gambar 4.1 Atribut pada Data Pelanggaran Lalu Lintas	38
Gambar 4.2 Atribut pada Data Pelanggaran Lalu Lintas	38
Gambar 4.3 Pencarian Asal Kecamatan Berdasarkan Asal Alamat	41
Gambar 4.4 Hasil Perhitungan Iterasi 1	50
Gambar 4.5 Hasil Perhitungan Iterasi 2	54
Gambar 4.6 Hasil Perhitungan Iterasi 3	55
Gambar 4.9 Nilai SSB.....	69
Gambar 4.10 Nilai Rasio	70
Gambar 4.11 Proses Import Data ke Rapidminer	74
Gambar 4.12 Penyimpanan Data Pada Rapidminer	75
Gambar 4.13 Desain Pemodelan (K-Means).....	75
Gambar 4.14 Desain Pemodelan (K-Medoids)	76
Gambar 4.15 Parameter K-Means	76
Gambar 4.16 Parameter K-Medoids	76
Gambar 4.17 Parameter Performance	77
Gambar 4.18 Cluster Model K-Means.....	77
Gambar 4.19 Cluster Model K-Medoids.....	77
Gambar 4.22 Visualisasi Cluster K-Means	79
Gambar 4.23 Visualisasi Cluster K-Medoids.....	79