

**IMPLEMENTASI DATA MINING DENGAN METODE  
*CLUSTERING* ALGORITMA *K-MEANS* UNTUK  
PENGELOMPOKAN DATA TILANG DI POLDA  
KEPRI**

**SKRIPSI**



**Oleh :**

**Reni Alda Aristawaty Manihuruk**

**190210092**

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS TEKNIK DAN KOMPUTER  
UNIVERSITAS PUTERA BATAM  
TAHUN 2023**

**IMPLEMENTASI DATA MINING DENGAN METODE  
*CLUSTERING* ALGORITMA *K-MEANS* UNTUK  
PENGELOMPOKAN DATA TILANG DI POLDA  
KEPRI**

**SKRIPSI**

**Untuk memenuhi salah satu syarat  
memperoleh gelar sarjana**



**Oleh :  
Reni Alda Aristawaty Manihuruk  
190210092**

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS TEKNIK DAN KOMPUTER  
UNIVERSITAS PUTERA BATAM  
TAHUN 2023**

## SURAT PERNYATAAN ORISINALITAS

Yang bertanda tangan di bawah ini saya:

Nama : Reni Alda Aristawaty Manihuruk  
NPM : 190210092  
Fakultas : Teknik dan Komputer  
Program Studi : Teknik Informatika

Menyatakan bahwa “skripsi” yang saya buat dengan judul:

**IMPLEMENTASI DATA MINING DENGAN METODE *CLUSTERING*  
ALGORITMA *K-MEANS* UNTUK PENGELOMPOKAN DATA TILANG DI  
POLDA KEPRI**

Adalah hasil karya sendiri dan bukan “duplikasi” dari karya orang lain. Sepengetahuan saya, didalam naskah skripsi ini tidak terdapat karya ilmiah atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis dikutip didalam naskah ini dan disebutkan dalam sumber kutipan dan daftar pustaka. Apabila ternyata di dalam naskah skripsi ini dapat dibuktikan terdapat unsur-unsur PLAGIASI, saya bersedia naskah skripsi ini digugurkan dan hasil penilaian yang saya peroleh dibatalkan, serta diproses sesuai dengan peraturan perundang-undangan yang berlaku.

Demikian pernyataan ini saya buat dengan sebenarnya tanpa ada paksaan dari siapapun.

Batam, 27 Januari 2023



**Reni Alda Aristawaty Manihuruk**  
190210092

**IMPLEMENTASI DATA MINING DENGAN METODE  
CLUSTERING ALGORITMA K-MEANS UNTUK  
PENGELOMPOKAN DATA TILANG DI POLDA  
KEPRI**

**SKRIPSI**

**Untuk memenuhi salah satu syarat  
memperoleh gelar sarjana**

**Oleh:**

**Reni Alda Aristawaty Manihuruk**

**190210092**

**Telah disetujui oleh Pembimbing pada tanggal  
seperti yang tertera di bawah ini**

**Batam, 27 Januari 2023**



**Rahmat Fauzi, S.Kom., M.Kom.**

**Pembimbing**

## ABSTRAK

Setiap komunitas mengandalkan transportasi, tetapi pengemudi harus mengikuti aturan agar aman. Ada beberapa variabel yang menyebabkan banyak individu ditilang antara lain kurangnya pengetahuan dan kesadaran masyarakat akan standar berkendara yang baik, akurat dan aman, serta kegagalan masyarakat untuk memeriksa kondisi kendaraan dan dokumen sebelum bepergian untuk menghindari tilang selama operasi khusus (penggerebekan). Penelitian ini mengkaji pelanggaran lalu lintas pengendara roda dua di Kota Batam yang menjadi permasalahan yang diteliti oleh penulis karena banyaknya pelanggaran lalu lintas yang mengganggu sistem regulasi selama ini. Kluster metode *K-Means*, penerapan Algoritma *K-Means* untuk mengelompokkan data pelanggaran lalu lintas membantu Polda Kepri menemukan pelanggaran lalu lintas terbanyak dan dinas tilang menemukan pengelompokan data pelanggaran lalu lintas di Kota Batam. Penulis akan memanfaatkan pendekatan *data mining* algoritma *K-Means* untuk melakukan segmentasi pelanggaran selama ini. Analisis data dari klaster dipisahkan berdasarkan jenis pelanggaran, tingkat pelanggaran dan besaran denda, menghasilkan temuan klaster 0: 11 item, kluster 1: 13 item, kluster 2: 11 item, dari 35 data yang diperlakukan sebagai data sampel dengan vektor kinerja pengelompokan terbaik adalah 0,499.

Kata kunci : Transportasi, Pelanggaran Lalu Lintas, Pengelompokan Data, Algoritma *K-Means*, cluster, RapidMiner.

## ABSTRACT

*Every community relies on transportation, but drivers must follow the rules to be safe. There are several variables that lead many individuals to get ticketed, including the public's lack of knowledge and awareness of excellent, accurate and safe driving standards and the community's failure to check vehicle conditions and paperwork before traveling to avoid tickets during special operations (raids). This research examined two-wheeled drivers' traffic offences in Batam City, an issue investigated by the author due to the numerous traffic offences that have disrupted the regulatory system thus far. The K-Means method clusters, implementing the K-Means Algorithm to group traffic violation data helps the Riau Island Regional Police discover the most traffic infractions and the ticketing service department locate Batam traffic violation data groupings. The author will utilize a K-Means algorithm data mining approach to segment infractions thus far. Data analysis from the cluster, segregated by mopabudget type, degree of violation and fine amount, yielded the findings of Cluster 0: 11 items, Cluster1: 13 items, Cluster 2: 11 items, from 35 data treated as sample data with performance vector best clustering is 0.499.*

*Keywords : Transport, Traffic Violation, Data Mining, K-Means Algorithm, cluster, RapidMiner.*

## **KATA PENGANTAR**

Puji dan syukur penulis panjatkan ke hadirat Tuhan yang Maha Esa, karena atas segala berkat dan rahmat-Nya sehingga tugas akhir ini dapat diselesaikan sesuai dengan waktu yang telah ditentukan. Penyusunan tugas akhir ini merupakan syarat untuk mencapai kelulusan dan meraih gelar Strata Satu (S1) Program Studi Akademik Teknik Informatika pada Universitas Putera Batam.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Karena itu, kritik dan saran akan senantiasa penulis terima dengan senang hati. Dengan segala keterbatasan penulis, penulis menyadari pula bahwa skripsi ini dapat diselesaikan dengan tepat waktu karena berbagai dukungan, dorongan bimbingan dari berbagai pihak. Untuk itu, dengan segala kerendahan hati, penulis menyampaikan ucapan terima kasih kepada:

1. Ibu Dr. Nur Elfi Husda, S.Kom., M.SI. selaku Rektor Universitas Putera Batam;
2. Bapak Welly Sugianto, S.T., M.M. selaku Dekan Fakultas Teknik dan Komputer Universitas Putera Batam;
3. Bapak Andi Maslan, S.T., M.SI. selaku Ketua Program Studi Teknik Informatika Universitas Putera Batam;
4. Bapak Rahmat Fauzi, S.Kom., M.Kom. selaku pembimbing Skripsi pada Program Studi Teknik Informatika Fakultas Teknik dan Komputer Universitas Putera Batam;
5. Ibu Alfannisa Annurrullah Fajrin, S.Kom., M.Kom. selaku pembimbing Akademik Program Studi Teknik Informatika Fakultas Teknik dan Komputer Universitas Putera Batam;
6. Dosen dan Staff Universitas Putera Batam;
7. Pihak Subdit Gakkum Ditlantasi Polda Kepulauan Riau;
8. Pihak Bagdalpers Ro SDM Polda Kepulauan Riau;
9. Kepada kedua orang tua dan keluarga tercinta atas nasihat, doa dan dukungannya;
10. Teman-teman dan sahabat yang selalu memberikan dukungan dan

semangat;

11. Pihak lain yang turut memberikan dorongan dalam dukungan penyusunan skripsi ini;

Semoga Tuhan Yang Maha Esa membalas kebaikan dan selalu mencurahkan hidayat serta taufik-Nya, Amin.

Batam, 27 Januari 2022



Reni Alda Aristawaty Manihuruk



## DAFTAR ISI

<b>HALAMAN SAMPUL</b> .....	i
<b>HALAMAN JUDUL</b> .....	ii
<b>SURAT PERNYATAAN ORISINALITAS</b> .....	iii
<b>HALAMAN PENGESAHAN</b> .....	iv
<b>ABSTRAK</b> .....	v
<b>ABSTRACT</b> .....	vi
<b>KATA PENGANTAR</b> .....	vii
<b>DAFTAR ISI</b> .....	ix
<b>DAFTAR GAMBAR</b> .....	xi
<b>DAFTAR TABEL</b> .....	xii
<b>DAFTAR RUMUS</b> .....	xiii
<b>BAB I PENDAHULUAN</b> .....	1
1.1. Latar Belakang.....	1
1.2. Identifikasi Masalah.....	4
1.3. Batasan Masalah .....	4
1.4. Rumusan Masalah.....	4
1.5. Tujuan Penelitian .....	5
1.6. Manfaat Penelitian .....	5
1.6.1. Manfaat Teoritis.....	5
1.6.2. Manfaat Praktis .....	6
<b>BAB II TINJAUAN PUSTAKA</b> .....	7
2.1. Teori Umum.....	7
2.1.1. <i>Knowledge Discovery in Database</i> (KDD) .....	7
2.1.2. Data Mining .....	9
2.1.3. <i>Clustering</i> .....	9
2.1.4. <i>K-Means</i> .....	11
2.1.5. <i>RapidMiner</i> .....	12
2.2. Teori Khusus.....	14
2.2.1. Pengertian Lalu Lintas .....	14
2.2.2. Sepeda Motor.....	16
2.2.3. Pelanggaran Lalu Lintas .....	17
2.2.4. Faktor Pelanggaran Lalu Lintas.....	17
2.2.5. Dampak Pelanggaran Lalu Lintas.....	19
2.3. Penelitian Terdahulu .....	20
2.4. Kerangka Pemikiran .....	23
<b>BAB III METODE PENELITIAN</b> .....	26
3.1. Desain Penelitian .....	26
3.2. Teknik Pengumpulan Data .....	28
3.3. Operasional Variabel .....	28
3.4. Algoritma.....	29
3.5. Lokasi dan Jadwal Penelitian.....	29
3.5.1. Lokasi .....	29
3.5.2. Jadwal Penelitian .....	30

<b>BAB IV HASIL DAN PEMBAHASAN</b> .....	31
4.1. Hasil Analisa dengan data mining Algoritma <i>clustering K-Means</i> .....	31
4.2. Melakukan Pra Proses Data Penelitian .....	31
4.3. Analisa Data.....	37
4.3.1. Pembersihan Data .....	38
4.3.2. Seleksi Data .....	38
4.3.3. Transformasi Data .....	38
4.3.4. Perhitungan Algoritma <i>Clustering K-Means</i> .....	41
4.4. Pengujian Menggunakan Software RapidMiner 5.3.....	116
<b>BAB V SIMPULAN DAN SARAN</b> .....	127
5.1. Simpulan .....	127
5.2. Saran .....	127
<b>DAFTAR PUSTAKA</b> .....	129
<b>LAMPIRAN</b>	
Lampiran 1. Dokumen Pendukung	
Lampiran 2. Daftar Riwayat Hidup	
Lampiran 3. Surat izin penelitian	
Lampiran 4. Balasan Surat Izin Penelitian	
Lampiran 5. Turnitin Skripsi	
Lampiran 6. Turnitin Jurnal	
Lampiran 7. Data Tilang	
Lampiran 8. Data Cleaning	
Lampiran 9. Transformasi Data	

## DAFTAR GAMBAR

<b>Gambar 2.1</b> <i>Knowledge Discovery in Database</i> .....	7
<b>Gambar 2.2</b> Rapidminer .....	12
<b>Gambar 2.3</b> Sistem Komponen Lalu Lintas .....	15
<b>Gambar 2.4</b> Kerangka Pemikiran .....	24
<b>Gambar 3.1</b> Desain Penelitian .....	26
<b>Gambar 3.2</b> <i>Flowchart</i> Algoritma <i>K-Means</i> .....	29
<b>Gambar 3.3</b> Lokasi Penelitian .....	30
<b>Gambar 4.1</b> Data Tilang Skripsi.xlsx .....	117
<b>Gambar 4.2</b> Tampilan Awal Rapidminer 5.3 .....	117
<b>Gambar 4.3</b> Menu <i>New Process</i> .....	118
<b>Gambar 4.4</b> Menu <i>New Process</i> .....	118
<b>Gambar 4.5</b> Mencari <i>Read Excel</i> .....	119
<b>Gambar 4.6</b> <i>Read Excel</i> .....	119
<b>Gambar 4.7</b> <i>Import Configuration Wizard</i> .....	120
<b>Gambar 4.8</b> <i>Data Import Wizard Step 1</i> .....	120
<b>Gambar 4.9</b> <i>Data Import Wizard Step 2</i> .....	121
<b>Gambar 4.10</b> <i>Data Import Wizard Step 3</i> .....	121
<b>Gambar 4.11</b> <i>Data Import Wizard Step Terakhir</i> .....	122
<b>Gambar 4.12</b> <i>Clustering K-Means</i> .....	122
<b>Gambar 4.13</b> Penentuan Titik Pusat .....	123
<b>Gambar 4.14</b> Hubungkan <i>Read Excel</i> ke <i>Clustering K-Means</i> .....	123
<b>Gambar 4.15</b> <i>Cluster Distance Performance</i> .....	124
<b>Gambar 4.16</b> Proses <i>Run</i> .....	124
<b>Gambar 4.17</b> <i>Example Set (Read Excel)</i> .....	125
<b>Gambar 4.18</b> <i>Cluster Model</i> .....	125
<b>Gambar 4.19</b> <i>Centroid Table</i> .....	126
<b>Gambar 4.20</b> <i>Graph View</i> .....	126
<b>Gambar 4.21</b> <i>Centroid Plot View</i> .....	126
<b>Gambar 4.22</b> <i>Performance Vector</i> .....	127

## DAFTAR TABEL

<b>Tabel 3.1</b> Jadwal Penelitian.....	30
<b>Tabel 4.1</b> Data Sampel .....	32
<b>Tabel 4.2</b> Transformasi Data .....	38
<b>Tabel 4.3</b> Perhitungan Iterasi Pertama.....	43
<b>Tabel 4.4</b> Perhitungan Iterasi Kedua .....	53
<b>Tabel 4.5</b> Perhitungan Iterasi Ketiga .....	63
<b>Tabel 4.6</b> Perhitungan Iterasi Keempat .....	69
<b>Tabel 4.7</b> Perhitungan Iterasi kelima.....	75
<b>Tabel 4.8</b> Perhitungan iterasi keenam .....	81
<b>Tabel 4.9</b> Perhitungan iterasi ketujuh.....	87
<b>Tabel 4.10</b> Perhitungan iterasi kedelapan.....	93
<b>Tabel 4.11</b> Perhitungan Iterasi kesembilan .....	99
<b>Tabel 4.12</b> Perhitungan Iterasi Kesepuluh.....	105
<b>Tabel 4.13</b> Hasil Akhir Perhitungan Algoritma <i>Clustering</i> K-Means.....	111

## DAFTAR RUMUS

<b>Rumus 2.1</b> Perhitungan Euclidean .....	11
--	----