ARTIKEL

CLUSTERING WILAYAH PADA PELANGGARAN BERKENDARAAN MENGGUNAKAN METODE *K-MEANS*

(Studi Kasus : Polres Kediri)

Oleh:

WULAN FAJARIYANTI

13.1.03.02.0359

Dibimbing oleh :

1. **Intan Nur Farida, M.Kom.**
2. **Danar Putra Pamungkas, M.Kom.**

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS TEKNIK

UNIVERSITAS NUSANTARA PGRI KEDIRI

2017

SURAT PERNYATAAN
ARTIKEL SKRIPSI TAHUN 2017

Yang bertanda tangan di bawah ini:

Nama Lengkap	: Wulan Fajariyanti
NPM	: 13.1.03.02.0359
Telepon/HP	: 085749457455
Alamat Surel (Email)	: wulan.fr28@gmail.com
Judul Artikel	: *CLUSTERING* WILAYAH PADA PELANGGARAN BERKENDARAAN MENGGUNAKAN METODE *K-MEANS* (Studi Kasus : Polres Kediri)
Fakultas – Program Studi	: Fakultas Teknik – Teknik Informatika
Nama Perguruan Tinggi	: Universitas Nusantara PGRI Kediri
Alamat Perguruan Tinggi	: Jl. KH. Achmad Dahlan 76 Kediri

Dengan ini menyatakan bahwa :

a. artikel yang saya tulis merupakan karya saya pribadi (bersama tim penulis) dan bebas plagiarisme;
b. artikel telah diteliti dan disetujui untuk diterbitkan oleh Dosen Pembimbing I dan II.

Demikian surat pernyataan ini saya buat dengan sesungguhnya. Apabila di kemudian hari ditemukan ketidaksesuaian data dengan pernyataan ini dan atau ada tuntutan dari pihak lain, saya bersedia bertanggungjawab dan diproses sesuai dengan ketentuan yang berlaku.

Mengetahui		Kediri, 09 Agustus 2017
Pembimbing I	Pembimbing II	Penulis,
Intan Nur Farida, M.Kom. 0704108701	Danar Putra Pamungkas, M.Kom. 0708028704	Wulan Fajariyanti 13.1.03.02.0359

CLUSTERING WILAYAH PADA PELANGGARAN BERKENDARAAN MENGGUNAKAN METODE *K-MEANS* (Studi Kasus : Polres Kediri)

Wulan Fajariyanti
13.1.03.02.0359
Fakulitas Teknik – Teknik Informatika
wulan.fr28@gmail.com
Intan Nur Farida, M. Kom. dan Danar Putra Pamungkas, M.Kom.
UNIVERSITAS NUSANTARA PGRI KEDIRI

ABSTRAK

Penelitian ini dilatar belakangi oleh kurangnya kesadaran masyarakat kediri yang sering melakukan pelanggaran dalam berkendaraan. Sehingga menyulitkan untuk para pihak kepolisian lalu lintas kediri dalam upaya memberikan himbauan – hibauan demi keselamatan para pengguna jalan. Kemudian terjadinya penumpukan data pelanggaran atau data tilang yang hanya diproses kedalam *Microsoft Excel* sehingga membuat data tersebut tidak beraturan. Data yang sudah masuk hanya di proses secara manual sehingga menyulitkan untuk mengetahui wilayah – wilayah mana saja yang sering melakukan pelanggaran.

Permasalahan dalam penelitian ini adalah Bagaimana merancang sistem yang dapat klasterisasi wilayah pada pelanggaran dalam berkendaraan menggunakan metode *K-Means* di Polres Kediri. Bagaimana menerapkan metode *K-Means* pada sistem klasterisasi wilayah yang sering melakukan pelanggaran berkendaraan di Polres Kediri. Penelitian ini menggunakan pendekatan penelitian kualitatif dikarenakan data kriteria diolah dengan menggunakan metode *K-Means* untuk mendapatkan hasil *Cluster* wilayah pada jenis pelanggaran dalam berkendaraan.

Kesimpulan hasil penelitian ini dengan menghasilkan aplikasi yang dapat klasterisasi wilayah pada pelanggaran dalam berkendaraan pada polres kediri. Dengan penerapan algoritma *K-Means* dapat mengklaster wilayah pada pelanggaran dalam berkendaraan berdasarkan banyak, sedang dan sedikit wilayah yang melakukan jenis pelanggaran dalam berkendaraan.

Kata Kunci : pelanggaran berkendaraan, wilayah, *kmeans*, klasterisasi

I. LATAR BELAKANG

Trasnportasi darat merupakan suatu kebutuhan yang tidak bisa dijauhkan dari kalangan masyarakat baik perorangan ataupun kelompok. Untuk diketahui sejumlah kendaraan yang beredar di kalangan masyarakat kini dari tahun ketahun semakin meningkat. Angka kemacetan pun terjadi di mana-mana, terutama di wilayah Kediri yang memasuki kawasan sekolah maupun pusat pembelanjaan. Banyaknya pengguna jalan yang tidak mematuhi peraturan berlalu lintas dengan baik, setiap harinya dapat menambah tingkat kecelakaan dan pelanggaran tata tertib lalu lintas dalam berkendara pada wilayah kediri.

Menurut Undang-Undang No. 22 Tahun 2009 dikatakan tertib, lancar, aman dan terpadu apabila dalam berlalu lintas berlangsung secara teratur sesuai dengan hak dan kewajiban pengguna jalan serta bebas hambatan dan kemacetan jalan. Adanya tata tertib lalu lintas diadakan untuk mewujudkan, mendukung, dan memelihara keamanan, keselamatan, ketertiban dan kelancaran lalu lintas. Berbagai tindak penertiban terus diupayakan para pihak kepolisian lalu lintas Kediri demi mewujudkan ketertiban lalu lintas dan kenyamanan berkendara, serta keselamatan para pengguna jalan raya, baik melalui razia penilangan kelengkapan berkendara, kelayakan mengemudi, serta himbauan– himbauan melalui poster dengan tujuan meningkatkan ketertiban dalam berlalu lintas. Disisi lain karena banyaknya wilayah yang sering terjadi pelanggaran dalam berkendara, mengharuskan para pihak kepolisian lebih teliti dalam memperhatikan kelengkapan berkendaraan dikalangan masyarakat khususnya pada wilayah Kediri dan sekitarnya.

Dikarenakan banyaknya penumpukan data pelanggaran berkendaraan yang hanya diproses kedalam *Microsoft Excel* sehingga membuat data tidak beraturan dan pihak kepolisian lalu lintas sulit untuk melalukan himbauan – himbauan kepada wilayah yang sering melakukan pelanggaran dalam berkendaraan. Data yang sudah masuk hanya di proses secara manual sehingga menyulitkan untuk mengetahui wilayah – wilayah mana saja yang sering melakukan pelanggaran.

Berdasarkan uraian diatas penulis akan melakukan penelitian dengan mengolah data pelanggaran dalam berkendaraan di Polres Kediri menggunakan algoritma *K-Means.* Dengan begitu aplikasi yang akan dirancang dapat mempermudah pihak kepolisian khususnya di bagian satlantas untuk mempermudah mengelompokkan wilayah yang sering

melakukan pelanggaran dalam berkendaraan.

II. METODE

K-means merupakan algoritma yang sangat sederhana dan kuat, sangat efisien, dan dapat digunakan untuk berbagai macam jenis data. Selain itu, *K-means* memiliki kelemahan mengenai optimasi yang masih memiliki beberapa teori tantangan, misalnya masalah generalisasi jarak *centroid*. Muncul dalam data dengan sifat yang rumit, seperti skala besar, tinggi dimensi, dan kelas ketidakseimbangan, juga perlu beradaptasi klasik *K-means* untuk skenario data yang berbeda[2].

Jarak dari data yang ada masing-masing *cluster* dihitung menggunakan rumus *Euclidean distance* [1] *:*

$$D(x,y) = |x-y| = \sqrt{\sum_{t-1}^{n}(x_1 - y_1)^2}$$

Dimana :

D = Jarak kuadrat Eucliden antar objek ke X_1 dengan X_2.

N = Jumlah Variabel *cluster*.

X_1 = Nilai atau data dari objek ke -1.

X_2 = Nilai atau data dari objek ke -2.

Berikut adalah perhitungan algoritma *K-Means* menggunakan sepuluh data yang berisi delapan *variable* untuk tiap datanya dan tiga *cluster*.

Tabel 1 Dataset 1

No	Wilayah	Helm	SIM	STNK	Kelengkapan
1	Blitar	9	20	29	6
2	Mojoroto	5	26	25	9
3	Pesantren	7	13	21	4
4	Kediri	3	25	12	1
5	Nganjuk	5	21	15	3
6	Trenggalek	9	16	24	2
7	Tulungagung	4	9	11	8
8	Pagu	4	18	21	5
9	Pare	2	20	23	1
10	Ngasem	6	21	10	4

Tabel 2 Dataset 2

No	Wilayah	Marka-Rambu	Melawan Arus	Safet ybelt	Kelebihan Muatan
1	Blitar	3	1	4	5
2	Mojoroto	2	3	2	2
3	Pesantren	5	1	1	2
4	Kediri	1	1	3	3
5	Nganjuk	3	2	2	1
6	Trenggalek	2	5	6	4
7	Tulungagung	4	8	2	2
8	Pagu	1	4	4	3
9	Pare	3	2	1	2
10	Ngasem	2	2	3	3

Tabel 1 belum bisa secara langsung dilakukan pemrosesan dikarenakan terdapat besaran angaka yang cukup jauh antar variabel. Perbedaan jarak atau besaran angka yang cukup jauh ini dapat menyulitkan dalam proses pengelompokan. Salah satu solusi yang digunakan untuk meperkecil besaran angka antar variabel adalah dengan melakukan normalisasi angka – angka yang ada di variabel jenis pelanggaran dalam berkendaraan. Berikut ini rumus nilai normalisasi.

$$\text{Normalisasi} = \frac{(Nilai\ Awal - Nilai\ Minimal)}{(Nilai\ Maksimal - Nilai\ Minimal)}$$

Nilai variabel jenis pelanggaran dalam berkendaraan akan dinormalisasi ke dalam rentang 0 – 1. Normalisasi angka pada tiap variabel ini sangat dibutuhkan sebelum proses perhitungan centroid oleh algoritma *K-Means* agar tidak ada parameter yang mendominasi dalam perhitungan jarak antar data. Adapaun tahapan yang dilakukan untuk proses normalisasi yaitu :

a Mencari nilai maksimum dan minimum untuk variabel jenis pelanggaran.

b Menghitung nilai normalisasi menggunakan persamaan.

Tabel 3 Data normalisasi jenis pelanggaran

No	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
1	1	0.65	1	0.63	0.5	0	0.6	1
2	0.43	1	0.79	1	0.25	0.29	0.2	0.25
3	0.71	0.24	0.58	0.38	1	0	0	0.25
4	0.14	0.94	0.11	0	0	0	0.4	0.5
5	0.43	0.71	0.26	0.25	0.5	0.14	0.2	0
6	1	0.41	0.74	0.13	0.25	0.57	1	0.75
7	0.29	0	0.05	0.88	0.75	1	0.2	0.25
8	0.29	0.53	0.58	0.5	0	0.43	0.6	0.5
9	0	0.65	0.68	0	0.5	0.14	0	0.25
10	0.57	0.71	0	0.38	0.25	0.14	0.4	0.5

Menentukkan *centroid* secara *random.*

Tabel 4 *Centroid* Awal

No	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
1	1	0.65	1	0.63	0.5	0	0.6	1
7	0.29	0	0.05	0.88	0.75	1	0.2	0.25
9	0	0.65	0.68	0	0.5	0.14	0	0.25

Setiap data akan diukur jaraknya dengan tiap *centroid* awal menggunakan rumus *Euclidean distance* kemudian diambil jarak terpendek.

Tabel 5 Iterasi 1

No	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
1	1	0.65	1	0.63	0.5	0	0.6	1
2	0.43	1	0.79	1	0.25	0.29	0.2	0.25
3	0.71	0.24	0.58	0.38	1	0	0	0.25
4	0.14	0.94	0.11	0	0	0	0.4	0.5
5	0.43	0.71	0.26	0.25	0.5	0.14	0.2	0
6	1	0.41	0.74	0.13	0.25	0.57	1	0.75
7	0.29	0	0.05	0.88	0.75	1	0.2	0.25
8	0.29	0.53	0.58	0.5	0	0.43	0.6	0.5
9	0	0.65	0.68	0	0.5	0.14	0	0.25
10	0.57	0.71	0	0.38	0.25	0.14	0.4	0.5

Tabel 6 Iterasi 1

No	C_1	C_2	C_3	Jarak
1	0	1.92	1.56	0
2	1.23	1.53	1.2	1.2
3	1.29	1.37	1.05	1.05
4	1.6	1.83	0.97	0.97
5	1.48	1.35	0.73	0.73
6	0.99	1.74	1.6	0.99
7	1.92	0	1.58	0
8	1.18	1.35	1.05	1.05
9	1.56	1.58	0	0
10	1.27	1.39	1.11	1.11

Tabel 7 Kelompok *Cluster* iterasi 1

No	C_1	C_2	C_3
1	1		
2	1		
3			1
4			1
5			1
6	1		
7		1	
8			1
9			1
10			1

Setelah diketahui anggota tiap – tiap *cluster*, kemudian centroid baru dihitung berdasarkan data anggota tiap – tiap

cluster berdasarkan rumus pusat anggota *cluster*.

Tabel 8 *Centroid* Baru

No	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
C_1	0.81	0.69	0.84	0.58	0.33	0.29	0.6	0.67
C_2	0.29	0	0.05	0.88	0.75	1	0.2	0.25
C_3	0.36	0.63	0.37	0.25	0.38	0.14	0.27	0.33

Tabel 9 Iterasi 2

No	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
1	1	0.65	1	0.63	0.5	0	0.6	1
2	0.43	1	0.79	1	0.25	0.29	0.2	0.25
3	0.71	0.24	0.58	0.38	1	0	0	0.25
4	0.14	0.94	0.11	0	0	0	0.4	0.5
5	0.43	0.71	0.26	0.25	0.5	0.14	0.2	0
6	1	0.41	0.74	0.13	0.25	0.57	1	0.75
7	0.29	0	0.05	0.88	0.75	1	0.2	0.25
8	0.29	0.53	0.58	0.5	0	0.43	0.6	0.5
9	0	0.65	0.68	0	0.5	0.14	0	0.25
10	0.57	0.71	0	0.38	0.25	0.14	0.4	0.5

Tabel 10 Iterasi 2

No	C_1	C_2	C_3	Jarak terpendek
1	0.53	1.92	1.24	0.53
2	0.87	1.53	0.96	0.87
3	1.18	1.37	0.91	0.91
4	1.29	1.83	0.69	0.69
5	1.12	1.35	0.39	0.39
6	0.77	1.74	1.23	0.77
7	1.57	0	1.33	0
8	0.73	1.35	0.69	0.69
9	1.27	1.58	0.62	0.62
10	0.95	1.39	0.51	0.51

Tabel 11 Kelompok *Cluster* Iterasi 2

No	C_1	C_2	C_3
1	1		
2	1		
3			1
4			1
5			1
6	1		
7		1	
8			1
9			1
10			1

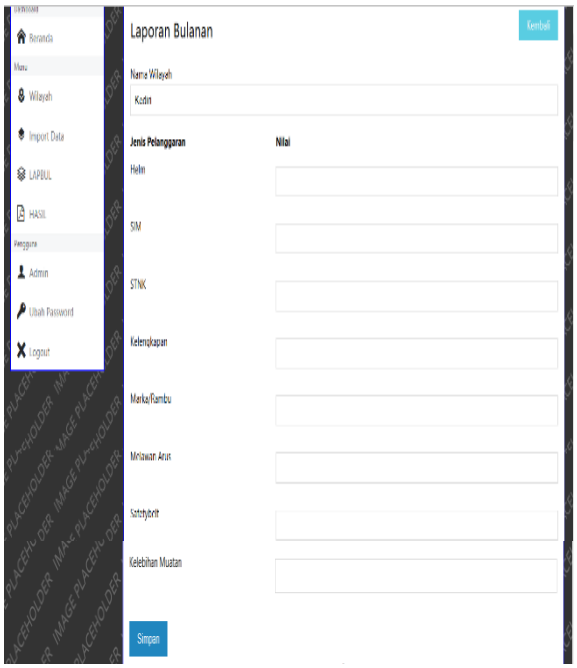
Kemudian lakukan perbandingan antara kelompok *cluster* sebelum dan setelahnya. Setiap data harus diukur jaraknya dengan tiap-tiap *centroid* baru. Jika data masih berubah atau tidak sama dengan kelompok *cluster* sebelumnya, maka dilakukan perhitungan lagi dimulai dari penentuan *centroid* baru, ulangi langkah kedua hingga data tidak berubah atau sama.

Karena tabel kelompok data kedua telah sama dengan tabel kelompok data pertama, maka hasil dari proses perhitungan dengan menggunakan metode K-Means sudah didapatkan dan proses berhenti. Ada 3 wilayah yang ada di *cluster* 1, 1 ada di *cluster* 2, dan 6 wilayah ada di *cluster* 3. Dimana *cluster* 1 berlabel Banyak, *cluster* 2 Sedang, dan *cluster* 3 Sedikit.

III. HASIL DAN KESIMPULAN

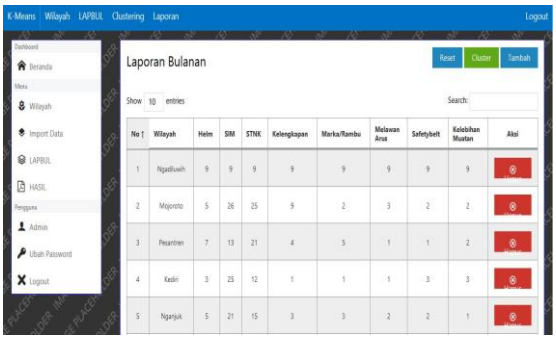
A. Hasil

Penelitian ini menghasilkan sebuah sistem aplikasi *clustering* wilayah pada pelanggaran berkendaraan menggunakan metode *K-Means* Berikut beberapa tampilan dari proses pada sistem yang telah dibangun.



Gambar 1 Halaman *Clustering*

Gambar 1 merupakan halaman yang digunakan admin untuk melakukan penginputan banyaknya jumlah pelanggar dalam berkendaraan. Admin harus mengisi semua *field* yang tersedia pada menu. Setelah *field* terisi, maka klik simpan kemudian pilih proses *clutering* untuk melakukan proses perhitungan, sehingga akan tampil hasil dari proses *clustering* dan akan disimpan ke dalam *database*. Pada halaman inilah admin juga bisa melakukan perubahan data atau mengupdate data.



Gambar 2 Halaman Laporan Bulanan

Gambar 2 merupakan halaman untuk menampilkan banyaknya jumlah pelanggar kendaraan. Pada halaman ini ada 5 proses pengolahan yang bisa dilakukan, yaitu : tambah data, hapus data, *reset* data, *cluster* dan pencarian data. Sebelum melakukan proses perhitungan *K-Means*, admin diharuskan untuk mereset data terlebih dahulu agar disaat perhitungan dijalankan tidak terjadi adanya penggandaan data. Proses tambah digunakan untuk menambahkan banyaknya jumlah pelanggar kendaraan. Kemudian admin melakukan proses *cluster* untuk mengetahui wilayah mana saja yang paling sering, sedikit dan sedang pelanggarannya.

Berikut adalah hasil pengujian dari 5 skenario yang dilakukan.

Tabel 10 Hasil Pengujian

No	Data	Centroid	Iterasi	C_1	C_2	C_3
1.	10	3	4	1	6	3
2.	20	3	8	7	4	9
3.	30	3	7	9	11	10
4.	40	3	5	18	12	10
5.	50	3	9	25	9	16

Skenario pengujian 1 - 5 dilakukan untuk mengetahui pengaruh jumlah data makssimum iterasi terhadap hasil klusterisasi. Hasil pengujian yang diperoleh, bahwa jumlah data tidak berpengaruh terhadap jumlah iterasi yang dihasilkan. Namun berpengaruh terhadap waktu yang dibutuhkan dan keakuratan hasil klasterisasi. Semakin banyak iterasi

maka semakin lama waktu yang dibutuhkan pula. Selain itu pengambilan *centroid* secara *random* sangat berpengaruh terhadap hasil klasterisasi dan hasil menjadi kurang akurat. Agar hasil bisa akurat maka *centroid* awal harus ditentukan terlebuh dahulu.

B. Kesimpulan

Berdasarkan pada analisis dan pembahasan di bab sebelumnya, maka dapat diambil kesimpulan bahwa :

1. Telah dihasilkan sistem untuk menentukan kelompok wilayah pada pelanggaran berkendaraan yang akan membantu petugas baur tilang dalam mengelompokkan suatu kelompok wilayah untuk menentukan wilayah yang sering melakukan pelanggagaran dalam berkendaraan tersebut secara efektif dan efisien.
2. Telah dihasilkan sistem untuk mengelompokkan wilayah pada pelanggaran dalam berkendaraan menggunakan metode K-Means yang dimulai dari pengambilan *centroid* secara *random*. Jumlah kelompok *cluster* ada 3. Tiap *cluster* akan dihitung jarak terdekat/ kemiripannya dengan *centroid* yang telah diambil secara *random* sesuai dengan kelompoknya. Setelah itu akan dihasilkan hasil klasterisasi kelompok pengayaan berdasarkan kelompoknya atau *cluster* yang telah ditempati.

IV. DAFTAR PUSTAKA

[1] Agusta, Y. 2007. *K-Means*-Penerapan, Permasalahan dan Metode Terkait. *Jurnal Sistem dan Informatika*, 3: 47-60.

[2] Dhani M. 2016. Definisi Data Mining Jurusan Teknik Informatika STMIK AMIKOM Yogyakarta.

[3] Edmira R. 2010. *Aplikasi K-Means Clustering* Untuk Pengelompokan Provinsi Berdasarkan Produksi Padi, Jagung, Kedelai Dan Kacang Hijau Tahun 2009. Jurnal Mat Stat Vol 10 No 2: 122-134.

[4] Krisna P. 2015. Penerapan Metode *Clustering K-Means* Untuk Mengukur Tingkat Kedisiplinan Siswa Di SMK Perikanan Nusantara Demak.

[5] Muhammad S. 2015. Analisis Kecelakaan Lalu Lintas Jalan Raya Di Kota Semarang Menggunakan Metode *K-Means Clustering*. Jurnal, Universitas Negeri Semarang.

[6] Undang-Undang Republik Indonesia Nomor 22 Tahun 2009 tentang berlalu lintas dan angkutan umum.

[7] Windha M. 2015. *Clustering* Menggunakan Metode *K-Means* Untuk Menenetukan Status Gizi Balita. *Jurnal Informatika*, Vol.15, No. 2.