



**Analisis Daerah Kerawanan Kriminalitas di Wilayah Hukum Polres Lebak
menggunakan Algoritma K – Means Clustering**

TUGAS AKHIR

An Nisaa' Muzdalifah
41518010141

UNIVERSITAS
**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2022**



**Analisis Daerah Kerawanan Kriminalitas di Wilayah Hukum Polres Lebak
menggunakan Algoritma K – Means Clustering**

Tugas Akhir

Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer

Oleh:

An Nisaa' Muzdalifah

41518010141

UNIVERSITAS
MERCU BUANA

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2022

LEMBAR PERNYATAAN ORISINALITAS

Yang bertanda tangan dibawah ini:

NIM : 41518010141

Nama : An Nisaa' Muzdalifah

Judul Tugas Akhir : Analisis Daerah Kerawanan Kriminalitas di Wilayah Hukum
Polres Lebak menggunakan Algoritma K-Means Clustering

Menyatakan bahwa Laporan Tugas Akhir saya adalah hasil karya sendiri dan bukan plagiat. Apabila ternyata ditemukan didalam laporan Tugas Akhir saya terdapat unsur plagiat, maka saya siap untuk mendapatkan sanksi akademik yang terkait dengan hal tersebut.

Jakarta, 07 Juli 2022



An Nisaa' Muzdalifah



UNIVERSITAS
MERCU BUANA

SURAT PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR

Sebagai mahasiswa Universitas Mercu Buana, saya yang bertanda tangan di bawah ini :

Nama Mahasiswa : An Nisaa' Muzdalifah
NIM : 41518010141
Judul Tugas Akhir : Analisis Daerah Kerawanan Kriminalitas di Wilayah Hukum
Polres Lebak menggunakan Algoritma K-Means Clustering

Dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Noneksklusif** (*None-exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul diatas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti/Noneksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat dan mempublikasikan tugas akhir saya.

Selain itu, demi pengembangan ilmu pengetahuan di lingkungan Universitas Mercu Buana, saya memberikan izin kepada Peneliti di Lab Riset Fakultas Ilmu Komputer, Universitas Mercu Buana untuk menggunakan dan mengembangkan hasil riset yang ada dalam tugas akhir untuk kepentingan riset dan publikasi selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 07 Juli 2022

UNIVERSITAS
MERCU BUANA



An Nisaa' Muzdalifah

SURAT PERNYATAAN LUARAN TUGAS AKHIR

Sebagai mahasiswa Universitas Mercu Buana, saya yang bertanda tangan di bawah ini :
Nama Mahasiswa : An Nisaa' Muzdalifah
NIM : 41518010141
Judul Tugas Akhir : Analisis Daerah Kerawanan Kriminalitas di Wilayah Hukum
Polres Lebak menggunakan Algoritma K-Means Clustering

Menyatakan bahwa :

1. Luaran Tugas Akhir saya adalah sebagai berikut :

No	Luaran	Jenis	Status
1	Publikasi Ilmiah	Jurnal Nasional Tidak Terakreditasi	Diajukan ✓
		Jurnal Nasional Terakreditasi	✓
		Jurnal International Tidak Bereputasi	Diterima
		Jurnal International Bereputasi	
	Disubmit/dipublikasikan di :	Nama Jurnal : ILKOM Jurnal Ilmiah ISSN : 2548-7779 Link Jurnal : http://jurnal.fikom.umi.ac.id/index.php/ILKOM/index Link File Jurnal Jika Sudah di Publish :	

2. Bersedia untuk menyelesaikan seluruh proses publikasi artikel mulai dari submit, revisi artikel sampai dengan dinyatakan dapat diterbitkan pada jurnal yang dituju.
3. Diminta untuk melampirkan scan KTP dan Surat Pernyataan (Lihat Lampiran Dokumen HKI), untuk kepentingan pendaftaran HKI apabila diperlukan

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 07 Juli 2022



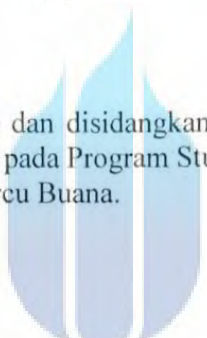
An Nisaa' Muzdalifah

LEMBAR PERSETUJUAN PENGUJI

NIM : 41518010141
Nama : An Nisaa' Muzdalifah
Judul Tugas Akhir : Analisis Daerah Kerawanan Kriminalitas di Wilayah Hukum Polres Lebak menggunakan Algoritma K – Means Clustering

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 26 Juli 2022



UNIVERSITAS
MERCU BUANA
(Ir. Emil R. Kaburuan, Ph.D., IPM)

LEMBAR PERSETUJUAN PENGUJI

NIM : 41518010141
Nama : An Nisaa' Muzdalifah
Judul Tugas Akhir : Analisis Daerah Kerawanan Kriminalitas di Wilayah Hukum Polres Lebak menggunakan Algoritma K – Means Clustering

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 26 Juli 2022



(Wawan Gunawan, S.Kom, MT)



UNIVERSITAS
MERCU BUANA

LEMBAR PERSETUJUAN PENGUJI

NIM : 41518010141

Nama : An Nisaa' Muzdalifah

Judul Tugas Akhir : Analisis Daerah Kerawanan Kriminalitas di Wilayah Hukum Polres Lebak menggunakan Algoritma K – Means Clustering

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 26 Juli 2022



(Dwi Anindyani Rochmah, ST, MTI)

UNIVERSITAS
MERCU BUANA

LEMBAR PENGESAHAN

NIM : 41518010141
Nama : An Nissa' Muzdalifah
Judul Tugas Akhir : Analisis Daerah Kerawanan Kriminalitas di Wilayah Hukum Polres Lebak menggunakan Algoritma K - Means Clustering

Tugas Akhir ini telah diperiksa dan disetujui sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 26 Juli 2022

Menyetujui,



(Elyani, Dr. Ir.)
Dosen Pembimbing

Mengetahui,



(Wawan Gusriawan, S.Kom, MT) (Ir. Emil R. Kabunan, Ph.D., IPM.)
Koord. Tugas Akhir Teknik Informatika Ka. Prodi Teknik Informatika

UNIVERSITAS
MERCU BUANA

KATA PENGANTAR

Segala puji dan syukur kehadirat Allah SWT atas berkah, rahmat dan hidayah- Nya yang senantiasa dilimpahkan kepada penulis, sehingga bisa menyelesaikan skripsi dengan judul “*Analisis Daerah Kerawanan Kriminalitas di Wilayah Hukum Polres Lebak menggunakan Algoritma K-Means Clustering*” sebagai syarat untuk menyelesaikan Program Sarjana (S1) pada Program Sarjana Fakultas Ilmu Komputer Jurusan Teknik Informatika Universitas Mercu Buana.

Penulis menyadari bahwa pembuatan Tugas Akhir ini tidak terlepas dari bantuan dan bimbingan berbagai pihak. Oleh karena itu, penulis mengucapkan terimakasih kepada:

1. Kedua orang tua yang telah memberikan support serta kasih sayang hinggasaya bisa menyelesaikan tugas akhir
2. Ibu Dr.Ir.Eliyani selaku dosen pembimbing saya yang telah membantu danmemberikan arahan kepada saya dalam menyelesaikan tugas akhir
3. Bapak Emil R. Kaburuan Ph.D selaku Kepala Program Studi Informatika yang telah memberikan arahan dan bimbingan selama perkuliahan
4. Bapak Wawan Gunawan S.Kom.,MT selaku koordinator tugas akhir Prodi Informatika yang telah memberikan arahan selama periode tugas akhir
5. Ibu Saruni Dwiasnati,ST, MM, M.Kom selaku dosen pembimbing akademik saya yang telah membimbing selama perkuliahan hingga sampaipada tugas akhir
6. Teman – teman Informatika 2018, sahabat, dan kerabat yang selalu memberikan support dan waktu untuk selalu mengingatkan dalam menyelesaikan tugas akhir

Akhir kata, penulis berharap jika ada kritik atau saran apapun yang sifatnya membangun bagi saya dengan senang hati akan saya terima. Semoga Tugas Akhir ini.

Jakarta, 07 Juli 2022
Penulis

DAFTAR ISI

HALAMAN SAMPUL	i
HALAMAN JUDUL.....	i
LEMBAR PERNYATAAN ORISINALITAS	ii
SURAT PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR	iii
SURAT PERNYATAAN LUARAN TUGAS AKHIR.....	iv
LEMBAR PERSETUJUAN.....	v
LEMBAR PERSETUJUAN PENGUJI.....	vi
LEMBAR PERSETUJUAN PENGUJI.....	vii
LEMBAR PERSETUJUAN PENGUJI.....	viii
LEMBAR PENGESAHAN.....	x
ABSTRAK.....	ix
ABSTRACT	x
KATA PENGANTAR	xi
DAFTAR ISI.....	xii
NASKAH JURNAL.....	1
KERTAS KERJA	12
BAB 1. LITERATUR REVIEW.....	13
BAB 2. ANALISIS DAN PERANCANGAN	23
BAB 3. SOURCE CODE.....	28
BAB 4. DATASET	33
BAB 5. TAHAPAN EKSPERIMEN	36
BAB 6. HASIL SEMUA EKSPERIMEN	41
DAFTAR PUSTAKA	48
LAMPIRAN DOKUMEN HAKI	50
LAMPIRAN KORESPONDENSI.....	53
LAMPIRAN SURAT IZIN PENELITIAN	54

Analisis Daerah Kerawanan Kriminalitas di Wilayah Hukum Polres Lebak menggunakan Algoritma K-Means Clustering

An Nisaa' Muzdalifah¹⁾; Eliyani²⁾

*Prodi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana
Jl. Raya Meruya Selatan, Kembangan, Jakarta, Indonesia 11650
41518010141@student.mercubuana.ac.id;*

Article history: Received Month xx, 2021; Revised Month xx, 2021; Accepted Month xx, 2021; Available online Month xx, 2021

Kriminalitas atau tindak kejahatan merupakan suatu tindakan yang melanggar hukum, undang - undang, norma, dan nilai yang berlaku dalam masyarakat, kriminalitas adalah masalah umum yang sering terjadi di kehidupan sehari – hari, termasuk di wilayah Lebak. Untuk mengetahui daerah kerawanan kriminalitas maka dalam penelitian ini menggunakan algoritma k – means yang bertujuan untuk mengclustering daerah kerawanan kriminalitas di Wilayah Hukum Polres Lebak menggunakan tools kaggle. Berdasarkan hasil penelitian algoritma k – means clustering telah menunjukkan bahwa dalam cluster 0 (Rawan), cluster 1 (Cukup Rawan), dan Cluster 2 (Sangat Rawan). Teknik data mining yang digunakan pada penelitian ini adalah metode k – means dengan optimasi penentuan jumlah cluster menggunakan metode Elbow. Dari hasil analisis dapat disimpulkan kepada masyarakat di Wilayah Lebak dimana tempat-tempat yang mempengaruhi terjadinya kriminalitas pada cluster 0 memiliki 491 Data kriminalitas, Cluster 1 memiliki 323 Data kriminalitas, dan Cluster 2 memiliki 204 Data kriminalitas. Dari hasil analisis dapat disimpulkan bahwa kerawanan kriminalitas di wilayah hukum Polres Lebak, sehingga diharapkan kepada Polres Lebak lebih meningkatkan keamanan dan juga masyarakat dapat meningkatkan kesadaran pentingnya keselamatan di wilayah Lebak.

Keywords: Kriminalitas, Kaggle, Data Mining, K-Means, Clustering

Pendahuluan

Kriminalitas atau tindak kejahatan merupakan suatu tindakan yang melanggar hukum, undang - undang, norma, dan nilai yang berlaku dalam masyarakat [1]. Kriminalitas adalah masalah umum yang sering terjadi kehidupan sehari – hari termasuk di wilayah lebak. Informasi tentang tingkat kerawanan tindak kejahatan diperlukan untuk menambah kewaspadaan terjadinya tindak kejahatan tersebut [2]. Meningkatnya jumlah tindak kejahatan setiap tahunnya karena beberapa faktor salah satunya faktor ekonomi dan faktor sosial [3].

Semakin tinggi resiko maka semakin tidak aman suatu daerah yang sangat berpengaruh bagi kesejahteraan masyarakat [4]. Penelitian ini bertujuan untuk membantu mengambil keputusan tentang daerah kerawanan kriminalitas di wilayah lebak. Dari hal ini diharapkan pihak kepolisian bisa melakukan strategi lebih baik dari sebelumnya. Untuk dapat mengatasi permasalahan yang terjadi, maka dibutuhkan suatu metode untuk mengidentifikasi dan mengelompokkan daerah rawan kriminalitas menjadi beberapa cluster atau kelompok.

Teknik data mining yang digunakan pada penelitian ini adalah metode k – means dengan optimasi penentuan jumlah cluster

menggunakan metode Elbow. Berdasarkan identifikasi masalah maka penelitian ini membatasi permasalahan hanya pada penentuan clustering jenis kasus, kecamatan, dan jumlah agar lebih akurat menggunakan metode k – means dengan optimasi jumlah cluster yaitu metode Elbow yang mendapatkan perbandingan jumlah cluster dengan cara menghitung SSE (Sum of Square Error) dari masing-masing nilai cluster [5]. Dari penelitian diketahui tingkat daerah kerawanan berdasarkan jenis kasus, kecamatan, dan jumlah. Untuk software pendukung pengolahan data dalam penelitian ini menggunakan kaggle.

Clustering adalah proses pengelompokan objek data ke dalam cluster yang terpisah-pisah sehingga data menjadi satu cluster serupa, tetapi data milik cluster yang berbeda berbeda[6]. Teknik clustering adalah suatu metode dalam analisa data untuk menyelesaikan persoalan tentang pengelompokan data. Salah satu metode clustering adalah K-Means, yang mempunyai kemampuan dalam klasifikasi data yang cukup besar jumlahnya relatif cepat dan efisien. Kekurangannya, hasil clustering dengan menerapkan algoritma K-Means bergantung pada penentuan awal pusat cluster, sehingga hasil perhitungan clustering dengan metode K-Means akan baik jika penentuan pusat cluster tepat.

K-Means merupakan suatu algoritma yang digunakan dalam pengelompokkan secara partisi yang memisahkan data ke dalam kelompok yang berbeda – berbeda. K-means telah banyak digunakan baik sendiri atau bersama-sama dengan pra-pemrosesan yang sesuai, karena kesederhanaannya dan efektivitas [7]. K-means cocok untuk mengelompokkan data yang tersebar merata di sekitar beberapa centroid [8]. Algoritma ini mampu meminimalkan jarak antara data ke clusternya. Pada dasarnya penggunaan algoritma ini dalam proses clustering tergantung pada data yang didapatkan dan konklusi yang ingin dicapai di akhir proses. Sehingga dalam penggunaan algoritma k means terdapat aturan sebagai berikut :

- a. Berapa jumlah cluster yang perlu dimasukkan.
- b. Hanya memiliki atribut bertipe numeric.

Pada dasarnya algoritma k-means hanya mengambil sebagian dari banyaknya komponen yang didapatkan untuk kemudian dijadikan pusat cluster awal, pada penentuan pusat cluster ini dipilih secara acak dari populasi data. Kemudian algoritma k-means akan menguji masing – masing dari setiap komponen dalam populasi data tersebut dan menandai komponen tersebut ke dalam salah satu pusat cluster yang telah didefinisikan sebelumnya tergantung dari jarak minimum antar komponen dengan tiap – tiap pusat cluster. Selanjutnya posisi pusat cluster akan dihitung kembali sampai semua komponen data digolongkan ke dalam tiap – tiap cluster dan terakhir akan terbentuk cluster baru [9].

Metode Elbow adalah metode yang digunakan untuk menginterpretasikan dan uji performa tingkat konsistensi jumlah cluster yang tepat dengan melihat nilai SSE. Pada titik tertentu akan terjadi grafik penurunan secara drastis dengan sebuah lekukan yang disebut dengan kriteria siku. Nilai itu kemudian menjadi nilai k atau jumlah cluster yang terbaik [10].

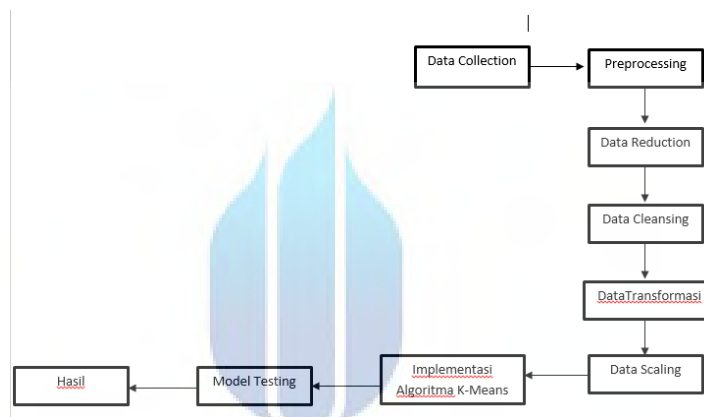
Dari penelitian sebelumnya, dilakukan oleh Suprihatin, Yustina Retno Wahyu Utami, & Didik Nugroho, 2019). Dengan studi kasus “ K – means clustering untuk pemetaan daerah rawan demam berdarah ” penelitian ini membahas tentang Sistem informasi geografis pemetaan daerah rawan DBD telah dapat dibangun dan memberikan informasi daerah rawan DBD kepada pengguna. Pada kasus pemetaan hasil dari pengujian akurasi sistem dengan data riil di lapangan menunjukkan bahwa kinerja sistem sudah baik dengan persentase akurasi sistem terhadap data riil di lapangan mencapai 92.31 %. Dengan hasil dari sistem terdapat tambahan level dari setiap kategori. daerah rawan demam berdarah ini pengkategorian endemis, sporadis dan bebas menggunakan dasar penentuan dari UPTD Puskesmas Kecamatan Nogosari. Selanjutnya diproses untuk pengelompokan Tingkat Kerawanan dari masing-masing kategori yang terbentuk. Sistem Informasi Geografis Pemetaan Daerah Rawan Demam Berdarah Menggunakan Metode K-Means Di Kecamatan Nogosari Kabupaten Boyolali ini terdapat 2 level pengguna yaitu admin yang mendapatkan hak akses sebagai administrator untuk mengelola data master, proses pemetaan dengan clustering, melihat peta dan melihat profil DBD. Dari hasil pengujian validitas cluster level zona sporadis DBD yang dilakukan peneliti dengan metode koefisien silhouette, nilai rata-rata koefisien silhouette untuk level 1 adalah 0.837, level 2 adalah 0.858, dan level 3 adalah 0.773 itu artinya objek telah berada di kelompok tepat [11].

Pada penelitian lain, dilakukan oleh Deni Triyansyah, Devi Fitriana K-Means adalah metode yang cocok digunakan dengan satu tipe data dan ditunjang dengan validitas kluster yang menggunakan DaviesBouldin Index untuk mengetahui seberapa baik proses kinerja pengklusteran. Hasil pengklusteran data untuk mendapatkan informasi reseller yang potensial, dengan menggunakan dua kluster, sedangkan untuk mendapatkan informasi tipe sepatu yang laris adalah dengan menggunakan tiga cluster. Dan pengklusteran ini menghasilkan 1 reseller yang potensial yaitu reseller Hoy, sedangkan proses mencari tipe sepatu yang laris mendapatkan 3 tipe sepatu yang terlaris yaitu Neo Adventage, Cloudfoam Swift, dan Bennasi. Hasil dari nilai Davies Bouldin Index yang diperoleh pada clustering pencarian reseller yang potensial adalah 0.271 dan 0.196 pada kluster pencarian tipe sepatu yang laris [12].

Dari permasalahan tersebut, maka dari itu penulis menggunakan metode k – means dengan optimasi penentuan jumlah cluster menggunakan metode Elbow. Berdasarkan identifikasi masalah maka penelitian ini membatasi permasalahan hanya pada penentuan clustering jenis kasus, kecamatan, dan jumlah agar lebih akurat menggunakan metode k – means dengan optimasi jumlah cluster yaitu metode Elbow yang mendapatkan perbandingan jumlah cluster dengan cara menghitung SSE (Sum of Square Error) dari masing-masing nilai cluster [5]. Dari penelitian diketahui tingkat daerah kerawanan berdasarkan jenis kasus, kecamatan, dan jumlah. Untuk software pendukung pengolahan data dalam penelitian ini menggunakan kaggle. sehingga model yang telah dibuat untuk clustering Daerah Kerawanan Kriminalitas di Wilayah Hukum Lebak menggunakan Algoritma K – Means Clustering agar dapat memperoleh hasil yang baik sehingga dapat diimplementasikan di wilayah Lebak.

Metode

Metodologi pengolahan data yang digunakan adalah metodologi data mining yaitu data cleasing, data reduction, data transformasi, dan data scaling. Dalam tahap pembersihan data hal yang dilakukan adalah pembersihan data yang merupakan outlier, tidak lengkap, dan memilih sampel data yang akan digunakan untuk penelitian ini yaitu sebanyak 1.100 jumlah data..



Gambar 1 Metode Penelitian

Data Collection

Pada data collection diambil dari Polres Lebak Banten yang berlokasi di Jl. Siliwangi No.KM.01, Cileuweung, Kec. Rangkasbitung, Kabupaten Lebak, Banten. Untuk data yang digunakan adalah data Kasus Kriminal. Data yang diambil dari tahun 2017-2022 dengan total data sebanyak 1.800 data.

B. Preprocessing

Pada tahap pre-processing dilakukan pengujian missing value dan normalisasi. Pengujian missing value berguna untuk menghindari error dengan cara menghilangkan data yang kosong atau null[9]. Adapun tahap preprocessing data yang dilakukan berikut ini:

a. Data Reduction

Pada data ini, terdapat 5 variabel yang diberikan, pada penelitian ini tidak semua variabel digunakan. Untuk menentukan variabel apa saja yang akan digunakan peneliti menyesuaikan variabel yang bisa digunakan dalam menentukan rawan, cukup rawan, dan sangat rawan. Untuk atribut JENIS_KELAMIN_PELAPOR tidak termasuk dalam menentukan jumlah cluster, untuk TAHUN, KECAMATAN, JENIS_KASUS, JUMLAH peneliti gunakan karena menentukan rawan, cukup rawan, dan sangat rawan, pada penelitian ini menggunakan 4 variabel yaitu, TAHUN, KECAMATAN, JENIS_KASUS, JUMLAH.

b. Data Cleansing

Pembersihan data adalah proses pada tahap preprocessing untuk mengisidata blank, mengulang data, mengoreksi data yang tidak sesuai dengan ketentuan atau salah ketik (seperti huruf yang tidak mencukupi dan huruf yang berlebihan),

mengubah dan memodifikasi persyaratan konfirmasi data [13]. Penelitian ini menggunakan data sebanyak 1.800 data. Pada proses cleansing data peneliti menghapus data yang jumlahnya bernilai kosong. Pada proses cleansing data dihasilkan 1.018 data.

c. Data Transformasi

Metode-metode transformasi untuk mengurangi banyaknya variabel efektif di bawah pertimbangan atau menemukan representasi invarian bagi data, dan memproyeksikan data pada ruang-ruang yang di dalamnya sebuah solusi lebih mudah ditemukan [14]. Data Transformasi dilakukan pada atribut JENIS_KASUS. Transformasi atribut JENIS_KASUS. Setelah itu, semua atribut independent bertipe data string/varchar diubah menjadi data integer. Data diubah agar terbaca pada saat proses perhitungan, jika tidak menggunakan string/varchar tidak dapat masuk ke proses perhitungan aritmatika algoritma. Perubahan data menjadi integer disajikan dalam gambar berikut ini.

d. Data Scaling

Pada metode ini untuk mengubah nilai pada dataset ke dalam skala atau range yang lebih kecil. Scaling menggunakan metode min-max scaler

C. Implementasi Algoritma K-Means

Hasil dari nilai centroid setiap Cluster menentukan tingkat Kerawanan Kriminalitas, semakin tinggi nilai centroid suatu cluster maka tingkat kerawanan semakin tinggi [15].

Adapun tahapan implementasi algoritma yang akan digunakan adalah sebagai berikut:

- a. Menentukan jumlah cluster yang ingin dibentuk
- b. Menentukan titik pusat cluster, setiap cluster memiliki centroid masing – masing
- c. Menghitung jarak data terhadap pusat cluster, hitung jarak dari masing – masing objek dengan centroid pada tiap cluster
- d. Mengelompokkan data sesuai minimum jarak ke pusat cluster, obyek akan terkelompok dalam cluster berdasarkan jarak minimum dengan centroid

D. Model Testing

Model akan diuji pada data yang telah diproses menggunakan metode elbow. Penentuan optimal jumlah cluster pada penelitian ini menggunakan salah satu metode analisis cluster yaitu metode Elbow, dengan memperhatikan nilai perbandingan (dari perhitungan SSE untuk setiap nilai cluster) antara jumlah cluster yang akan membentuk siku pada suatu titik. Sehingga semakin besar jumlah cluster k maka nilai SSE akan semakin kecil [10].

$$SSE = \sum_{K=1}^K \sum_{x_i \in S_k} \|x_i - c_k\|_2^2$$

Setelah dilihat akan ada beberapa nilai K yang mengalami penurunan paling besar dan selanjutnya hasil dari nilai K akan turun secara perlahan-lahan sampai hasil dari nilai K tersebut stabil. Misalnya nilai cluster K=2 ke K=3, kemudian dari K=3 ke K=4, terlihat penurunan drastis membentuk siku pada titik K=3 maka nilai cluster k yang ideal adalah K=3 [16].

Hasil dan Diskusi

Hasil dari penelitian yang dilakukan adalah melakukan analisis daerah kerawanan di Wilayah Hukum Polres Lebak dengan menggunakan algoritma k – means clustering. Pada bagian ini menampilkan hasil semua eksperimen yang telah dilakukan. Masing – masing memiliki jumlah anggota cluster adalah 491 data untuk data cluster 0, 323 data untuk data cluster 1, 204 data untuk data cluster 2. Informasi Cluster Optimal dapat dilihat pada tabel 1.

Tabel 1 Informasi berdasarkan Cluster Optimal

Cluster	Informasi
0(491)	Kelompok ini merupakan daerah kerawanan dengan tingkat Rendah
1(323)	Kelompok ini merupakan daerah kerawanan dengan tingkat

	Sedang
2(204)	Kelompok ini merupakan daerah kerawanan dengan tingkat Tinggi

Hasil selanjutnya dari eksperimen yang dilakukan pada dataset yang telah tersedia tersebut menemukan bahwa dataset tersebut berhasil di import dari excel menjadi data frame yang berisi no, TAHUN, KECAMATAN, JENIS_KASUS, JENIS_KELAMIN_PELAPOR, JUMLAH untuk dapat diolah menggunakan algoritma clustering. Berikut adalah dataset kriminalitas dapat dilihat pada tabel 2.

Tabel 2 Dataset Kriminalitas

no	TAHUN	KECAMATAN	JENIS_KASUS	JENIS_KELAMIN_PELAPOR	JUMLAH
1	2022	Banjarsari	Pencurian Dengan Pemberatan	Laki-Laki	1
2	2022	Banjarsari	Pencurian Dengan Pemberatan	Perempuan	1
3	2022	Banjarsari	Tipu Gelap	Laki-Laki	25
4	2022	Banjarsari	Pemerasan	Laki-Laki	1
5	2022	Banjarsari	Pengeroyokan	Laki-Laki	1
6	2022	Banjarsari	Tipu Gelap	Laki-Laki	1
7	2022	Banjarsari	Cabul	Laki-Laki	5
8	2022	Banjarsari	Miras	Laki-Laki	1
9	2022	Banjarsari	Penggelapan	Laki-Laki	3
10	2022	Cibadak	Miras	Laki-Laki	1
11	2022	Cibadak	Penggelapan	Perempuan	1
12	2022	Cibadak	Cabul	Laki-Laki	8
13	2022	Cibadak	Cabul	Perempuan	5
14	2022	Cibadak	Pencurian Dengan Pemberatan	Laki-Laki	28
15	2022	Cibadak	Pencurian Dengan Pemberatan	Perempuan	17
16	2022	Cibadak	Tipu Gelap	Laki-Laki	1
17	2022	Cibadak	Pemerasan	Laki-Laki	1
18	2022	Cibadak	Pengeroyokan	Laki-Laki	5
19	2022	Cibadak	KDRT	Laki-Laki	1
20	2022	Cibadak	KDRT	Perempuan	1
21	2022	Cibadak	UU CIPTA KERJA	Laki-Laki	1
22	2022	Cibadak	Senjata Tajam	Laki-Laki	6
23	2022	Cibeber	Korupsi	Laki-Laki	3

Kemudian hasil dari transformasi data dari sebelumnya dataset yang berhasil di import menjadi data frame kemudian dilakukan pengolahan menjadi data array untuk dapat diterapkan pada algoritma clustering dilihat pada gambar 3.


```

84      1
9       1
13      1
14      1
22      1
...
34     45
75     47
50     82
33    100
2     114
Name: JENIS_KASUS, Length: 84, dtype: int64

```

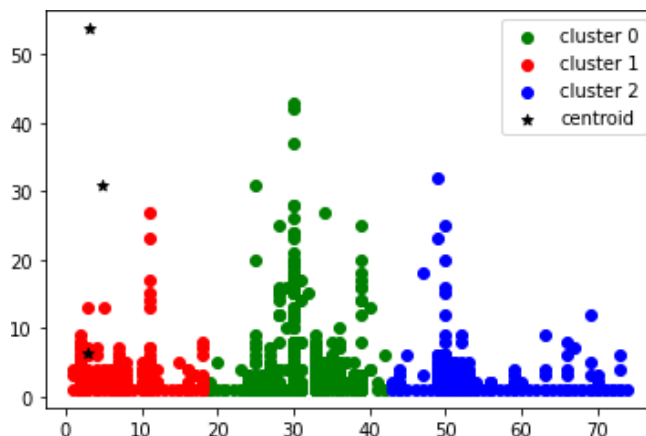
Gambar 3. Transformasi Data

Tahap selanjutnya peneliti melakukan scaling menggunakan MinMax Scaler pada atribut JUMLAH dan JENIS_KASUS dan membuat atribut baru dengan nama JUMLAH_CLUSTER dan JENIS_KASUS_CLUSTER berikut hasil data yang telah di scaling pada tabel 3.

Tabel 3. Data Hasil Scaling

no	TAHUN	KECAMATAN	JENIS_KASUS	JENIS_KELAMIN_PELAPOR	JUMLAH	Cluster	JUMLAH_CLUSTER	JENIS_KASUS_CLUSTER
1	2022	Banjarsari	30	Laki - laki	1	0	0.000000	0.397260
2	2022	Banjarsari	30	Perempuan	1	0	0.000000	0.397260
3	2022	Banjarsari	50	Laki - laki	25	2	0.571429	0.671233
4	2022	Banjarsari	25	Laki - laki	1	0	0.000000	0.328767
5	2022	Banjarsari	34	Laki - laki	1	0	0.000000	0.452055
6	2022	Banjarsari	50	Laki - laki	1	2	0.000000	0.671233
7	2022	Banjarsari	2	Laki - laki	5	1	0.095238	0.013699
8	2022	Banjarsari	16	Laki - laki	1	1	0.000000	0.205479
9	2022	Banjarsari	36	Laki - laki	3	0	0.047619	0.479452
10	2022	Cibadak	16	Laki - laki	1	1	0.000000	0.205479

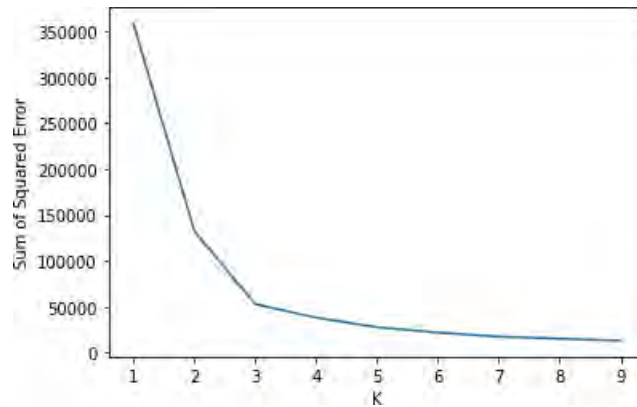
Setelah data di Scaling dari data kasus kriminalitas terbagi terbagi menjadi 3 warna cluster dan Centroid sebagai titik data (imajiner atau nyata) di pusat cluster. Urutan dari clusternya yaitu cluster 0 dengan warna hijau, cluster 1 dengan warna merah, cluster 2 dengan warna biru dapat dilihat pada Gambar 4.



Gambar 4. Hasil Klasterisasi dan Centroid

Untuk melakukan inisialisasi pusat cluster sebanyak jumlah cluster (K) menggunakan Metode Elbow. Metode Elbow digunakan untuk menentukan jumlah cluster (k), dimulai dari k=2 dan seterusnya. Berikut adalah Metode Elbow pada Algoritma K-Means

pada Gambar 5.



Gambar 5 Metode Elbow pada Algoritma K-Means

Dilihat dari hasil pengelompokan K=3 yang ditampilkan mulai dari cluster 0 hingga cluster 2, memiliki jumlah anggota masing-masing cluster adalah 491 untuk cluster 0, 323 untuk cluster 1, dan 204 untuk cluster 2.

Tabel 4 Anggota Cluster 0

No	Tahun	Kecamatan	Jenis Kasus	Jumlah
1	2020	Banjarsari	30	1
2	2020	Sobang	30	1
3	2020	Sobang	39	20
4	2020	Wanasalam	33	6
5	2020	Wanasalam	27	4
6	2020	Warunggunung	29	1
7	2020	Warunggunung	30	4
8	2019	Banjarsari	30	1
9	2019	Banjarsari	27	2
10	2019	Banjarsari	29	16
...
...
491	2021	Rangkasbitung	25	1

Tabel 5 Anggota Cluster 1

No	Tahun	Kecamatan	Jenis Kasus	Jumlah
1	2021	Cikukur	7	7
2	2019	Cimarga	11	2
3	2019	Muncang	2	5
4	2022	Cihara	11	6
5	2019	Muncang	2	1
6	2021	Cilograng	8	5
7	2021	Cikukur	7	3
8	2019	Muncang	11	2
9	2019	Malingping	2	4
10	2017	Leuwidamar	6	3
...
...
323	2021	Wanasalam	11	14

Tabel 6 Anggota Cluster 2

No	Tahun	Kecamatan	Jenis Kasus	Jumlah
1	2017	Cigemblong	50	8
2	2017	Cihara	50	4
3	2022	Cipanas	50	1
4	2017	Cihara	49	2
5	2017	Warunggunung	50	1
6	2022	Banjarsari	50	25
7	2017	Sobang	47	3
8	2017	Cigemblong	49	2
9	2022	Banjarsari	50	1
10	2017	Cibeber	45	1
...
...
204	2017	Sobang	50	2

Hasil dari penelitian cluster 0 hingga cluster 2 selanjutnya dilakukan penjabaran Informasi Jenis Kasus dengan menggunakan kode jenis kasus yang dijelaskan pada Tabel 7 sebagai berikut.

Tabel 7 Informasi Jenis Kasus

Jenis Kasus	Keterangan
7	KDRT
11	Korupsi
2	Cabul
8	Kekerasan Anak
6	Judi
50	Tipu Gelap
45	Pernikahan Tanpa Izin
49	Senjata Tajam
47	Pungutan Liar
30	Pencurian Dengan Pemberatan
39	Penipuan
33	Penganiayaan
27	Penadahan
29	Pencurian
25	Pemerasan

Dari hasil ringkasan diatas dapat dicluster optimalkan menjadi 3 cluster dengan masing-masing tingkatann yaitu cluster 0 sebanyak 491 kasus untuk tingkat kerawanan rendah, cluster 1 sebanyak 323 kasus untuk tingkat kerawanan sedang, dan cluster 2 sebanyak 204 untuk tingkat kerawanan tinggi.

Kesimpulan

Berdasarkan hasil penelitian yang dilakukan pada Polres Lebak dan setelah menganalisa permasalahan yang ada yaitu bagaimana memetakan wilayah-wilayah tindak kriminalitas di Lebak agar dapat dengan mudah mengetahui wilayah-wilayah terjadinya tindak kriminalitas dan memudahkan aparat kepolisian untuk mengetahui dimanakah tempat yang layak mendapatkan perhatian lebih untuk mengurangi tindakan kriminalitas tersebut. Dari hasil penelitian yang dilakukan menggunakan algoritma k-means untuk mengelompokkan data kasus kriminalitas yang ada di Polres Lebak. Menghasilkan 3 Cluster dari hasil pengujian menggunakan K-Means Clustering dengan Metode Elbow. Dalam 3 Cluster tersebut dibagi ke dalam 3 tingkat kerawanan yaitu Rawan (0), Cukup Rawan (1) dan Sangat Rawan(2). Pada Cluster 0 dengan 31 Kecamatan yaitu Banjarsari, Sobang, Wanasalam, Warunggunung, Banjarsari, Bayah, Bojong Manik, Cibadak, Sobang, Sajira, Muncang, Pangarangan, Rangkasbitung, Leuwidamar, Maja, Malingping, Kalanganyar, Cijaku, Cikulur, Cileles, Cimarga, Cipanas, Cihara, Curugbitung, Banjarsari, Bayah, Ciligrang, Lebak Gedong, Ciligrang, Cigemblong, dan Cibeber memiliki 491 Data, Cluster 1 dengan 30 Kecamatan yaitu Cikulur, Cimarga, Muncang, Cihara, Ciligrang, Cikulur, Maja, Leuwidamar, Cigemblong, Rangkasbitung, Cimarga, Kalanganyar, Malingping, Cibadak, Banjarsari, Bojong Manik, Bayah, Banjarsari, Cileles, Warunggunung, Wanasalam, Lebak Gedong, Sajira, Muncang, Sobang, Cirinten, Gunung Kencana, Pangarangan, Cipanas, dan Cijaku memiliki 323 Data, dan Cluster 2 dengan 20 Kecamatan yaitu Cigemblong, Cihara, Cipanas, Warung Gunung, Banjarsari, Sobang, Cibeber, Lebak Gedong, Cikulur, Leuwidamar, Malingping, Maja, Cimarga, Rangkasbitung, Cibadak, Kalanganyar, Pangarangan, Bojong Manik, Bayah, dan Cileles memiliki 204 Data dengan jenis kasus yang berbeda – beda.



Daftar Pustaka

- [1] D. P. T. Hapsari and E. Widodo, "Pengelompokan Daerah Rawan Kriminalitas di Indonesia Menggunakan Analisis K-Means Clustering," *Pros. SI MaNIs (Seminar Nas. Integr. Mat. dan Nilai-Nilai Islam.*, vol. 1, no. 1, pp. 147–153, 2017, [Online]. Available: <http://conferences.uin-malang.ac.id/index.php/SIMANIS/article/view/52>
- [2] H. Lubis, R. Salkiawati, and M. A. D. Ulhaq, "Sistem Informasi Kriminalitas Berbasis Android," *J. Algoritma.*, vol. 15, no. 2, pp. 84–94, 2019, doi: 10.33364/algoritma/v.15-2.84.
- [3] R. Sugiharto and R. Lestari, "UPAYA KEPOLISIAN DALAM PENANGGULANGAN KEJAHATAN PERAMPASAN SEPEDA MOTOR DI JALAN RAYA (Studi Kasus di Polrestabes Semarang)," *J. Pembaharuan Huk.*, vol. II, no. 2, pp. 339–347, 2015.
- [4] P. P. Nabilah, R. M. Zidni, N. L. Humairoh, and E. Widodo, "Penerapan Spatial Error Model (SEM) untuk Mengetahui FaktorFaktor yang Memengaruhi Kriminalitas," *Semin. Nas. Off. Stat.*, vol. 2021, no. 1, pp. 333–342, 2021, doi: 10.34123/semnasoffstat.v2021i1.870.
- [5] R. Yuliana Sari, H. Oktavianto, and H. Wahyu Sulisty, "Algoritma K-Means Dengan Metode Elbow Untuk Mengelompokkan Kabupaten/Kota Di Jawa Tengah Berdasarkan Komponen Pembentuk Indeks Pembangunan Manusia K-Means Algorithm With Elbow Method To Grouping District/City in Central Java Based on Components of Human D," *J. Smart Teknol.*, vol. 3, no. 2, pp. 2774–1702, 2022, [Online]. Available: <http://jurnal.unmuhjember.ac.id/index.php/JST>
- [6] S. P. Tamba, M. D. Batubara, W. Purba, M. Sihombing, V. M. Mulia Siregar, and J. Banjarnahor, "Book data grouping in libraries using the k-means clustering method," *J. Phys. Conf. Ser.*, vol. 1230, no. 1, 2019, doi: 10.1088/1742-6596/1230/1/012074.
- [7] Sudirman, A. P. Windarto, and A. Wanto, "Data mining tools | rapidminer: K-means method on clustering of rice crops by province as efforts to stabilize food crops in Indonesia," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 420, no. 1, 2018, doi: 10.1088/1757-899X/420/1/012089.
- [8] H. Hartkamp, "Versuche zur Aultrennung von Gemischen anorganischer Kationen mit Hilfe der Dünnschichtchromatographie," *Fresenius' Zeitschrift für Anal. Chemie*, vol. 188, no. 1, pp. 48–49, 1962, doi: 10.1007/BF00501033.
- [9] M. Algoritma and K. Pulau, "Klasterisasi Dampak Bencana Gempa Bumi," vol. 8, no. 1, pp. 175–179, 2022.
- [10] N. T. Hartanti, "Metode Elbow dan K-Means Guna Mengukur Kesiapan Siswa SMK Dalam Ujian Nasional," *J. Nas. Teknol. dan Sist. Inf.*, vol. 6, no. 2, pp. 82–89, 2020, doi: 10.25077/teknosi.v6i2.2020.82-89.
- [11] S. Suprihatin, Y. R. W. Utami, and D. Nugroho, "K-Means Clustering Untuk Pemetaan Daerah Rawan Demam Berdarah," *J. Teknol. Inf. dan Komun.*, vol. 7, no. 1, 2019, doi: 10.30646/tikomsin.v7i1.408.
- [12] D. Triansyah and D. Fitrihanah, "Analisis Data Mining Menggunakan Algoritma K-Means Clustering Untuk Menentukan Strategi Marketing," *J. Telekomun. dan Komput.*, vol. 8, no. 3, p. 163, 2018, doi: 10.22441/incomtech.v8i3.4174.
- [13] T. I. Hermanto *et al.*, "Analisis Sebaran Titik Rawan Bencana dengan K-Means Clustering dalam Penanganan Bencana," *J. Sains Komput. Inform. (J-SAKTI)*, vol. 5, no. 1, p. 406, 2021.
- [14] D. Firdaus, "Penggunaan Data mining dalam kegiatan pembelajaran," vol. 6, no. 2, pp. 91–97, 2017.
- [15] D. Marlina, N. Lina, A. Fernando, and A. Ramadhan, "Implementasi Algoritma K-Medoids dan K-Means untuk Pengelompokkan Wilayah Sebaran Cacat pada Anak," *J. CoreIT J. Has. Penelit. Ilmu Komput. dan Teknol. Inf.*, vol. 4, no. 2, p. 64, 2018, doi: 10.24014/coreit.v4i2.4498.
- [16] E. Muningsih and S. Kiswati, "Sistem Aplikasi Berbasis Optimasi Metode Elbow Untuk Penentuan Clustering Pelanggan," *Joutica*, vol. 3, no. 1, p. 117, 2018, doi: 10.30736/jti.v3i1.196.

KERTAS KERJA

Ringkasan

Kertas kerja ini merupakan material kelengkapan artikel jurnal dengan judul “Analisis Daerah Kerawanan Kriminalitas di Polres Lebak Menggunakan Algoritma K – Means Clustering”. Kertas kerja berisi semua material hasil penelitian Tugas Akhir yang tidak dimuat / atau disertakan di artikel jurnal. Di dalam kertas kerja ini disajikan: Literatur Review, dataset yang digunakan, source code, dan hasil eksperimen secara keseluruhan.

