

# KLASTERISASI TINDAK KRIMINALITAS DI PROVINSI JAWA BARAT DENGAN MENGGUNAKAN ALGORITMA *K-MEDOIDS*

Elina Sarastuti<sup>1\*</sup>, Deni Mahdiana<sup>2</sup>, Nidya Kusumawardhany<sup>3</sup>

<sup>123</sup>Sistem Informasi, Fakultas Teknologi Informasi, Universitas Budi Luhur, Jakarta Selatan, Indonesia

Email: <sup>1</sup>\*elinasarastuti88@gmail.com, <sup>2</sup>deni.mahdiana@budiluhur.ac.id,

<sup>3</sup>nidya.kusumawardhany@budiluhur.ac.id

(\*: *Corresponding Author*)

(Naskah masuk: 1 Maret 2024, diterima untuk diterbitkan: 30 April 2024)

## Abstrak

Pada tahun 2022, Jawa Barat mengalami peningkatan angka tindak kriminalitas. Merujuk pada data Badan Pusat Statistik (BPS), pada 2021 Jawa Barat menduduki tindak kriminalitas pada posisi kesembilan di Indonesia. Selanjutnya berdasarkan Kementerian Koordinator Bidang Politik Hukum dan Keamanan (Kemenko Polhukam) pada tahun 2022 Jawa Barat mengalami lonjakan angka kriminalitas hingga sebesar 69 persen, dan menduduki peringkat kelima di Indonesia. Tujuan pada penelitian ini untuk menggunakan pendekatan *data mining* dalam mengelompokkan Kabupaten/Kota di Jawa Barat berdasarkan angka kriminalitas sehingga dapat membantu pihak kepolisian setempat untuk dapat melakukan pemetaan agar pencegahan serta penindakan kriminalitas dapat sesuai lebih efektif serta tegas, sehingga dapat menekan angka kriminalitas di Jawa Barat. Data penelitian yang digunakan adalah data kriminalitas di Jawa Barat pada tahun 2019 sampai 2021 dengan 9 jenis tindak kriminalitas, diantaranya pencurian, penipuan, penganiayaan, pembakaran, pemerkosaan, pengedar/penyalahgunaan narkoba, perjudian, pembunuhan, dan perdagangan manusia. Metode analisis menggunakan algoritma *K-Medoids* dengan *tools Rapidminer* untuk mengelompokkan pola kriminalitas ke dalam kelompok-kelompok yang memiliki kesamaan karakteristik. Pengelompokan dilakukan berdasarkan atribut-atribut tertentu seperti 9 jenis tindak kejahatan dan kabupaten/kota terjadinya tindak kejahatan tersebut. Hasil penelitian ini dapat dinyatakan akurat karena hasil akhir penggunaan *tools* dan perhitungan manual mendapatkan hasil yang sama, dengan hasil berupa Klasterisasi status kabupaten/kota yang dibagi menjadi 2 *cluster* yaitu aman dan rawan dengan nilai DBI 0,761. Penelitian ini mendapatkan hasil yaitu, *cluster* aman terdiri dari 3 Kabupaten/Kota, diantaranya Kabupaten Kuningan, Kabupaten Pangandaran, dan Kota Banjar sedangkan *cluster* rawan terdiri dari 16 Kabupaten/Kota, diantaranya Kabupaten Bandung, Kabupaten Bandung Barat, Kabupaten Bekasi, Kabupaten Bogor, Kabupaten Ciamis, Kabupaten Cianjur, Kabupaten Cirebon, Kabupaten Garut, Kabupaten Indramayu, Kabupaten Karawang, Kabupaten Majalengka, Kabupaten Purwakarta, Kabupaten Subang, Kabupaten Sukabumi, Kabupaten Sumedang, dan Kabupaten Tasikmalaya.

**Kata kunci:** *data mining, clustering, k-medoids, kriminalitas*

## CLASSTERIZATION OF CRIMINAL ACTS IN WEST JAVA PROVINCE USING THE *K-MEDOIDS* ALGORITHM

### Abstract

In 2022, West Java will experience an increase in crime rates. Referring to data from the Central Statistics Agency (BPS), in 2021 West Java was ranked ninth in crime in Indonesia. Furthermore, according to the Coordinating Ministry for Political, Legal and Security Affairs (Kemenko Polhukam), in 2022 West Java will experience a spike in crime rates of up to 69 percent, and is ranked fifth in Indonesia. The aim of this research is to cluster regencies/cities in West Java so that it can help the local police to reduce the crime rate that occurs in West Java. The research data used is crime data in West Java from 2019 to 2021 with 9 types of crime, including theft, fraud, assault, arson, rape, drug dealing/abuse, gambling, murder, and human trafficking. The research method used is *K-Medoids Clustering*. The research results show that there are 2 optimal clusters with a DBI value of 0.761, consisting of a safe cluster and a vulnerable cluster. This research obtained results, namely, the safe cluster consists of 3 regencies/cities, including Kuningan Regency, Pangandaran Regency, and Banjar City, while the vulnerable cluster consists of 16 regencies/cities, including Bandung Regency, West Bandung Regency, Bekasi Regency, Bogor Regency, Ciamis, Cianjur Regency, Cirebon Regency, Garut Regency, Indramayu Regency, Karawang Regency, Majalengka Regency, Purwakarta Regency, Subang Regency, Sukabumi Regency, Sumedang Regency, and Tasikmalaya Regency.

**Keywords:** *data mining, clustering, k-medoids, crime*

## 1. PENDAHULUAN

Kriminalitas termasuk kejahatan yang merugikan bahkan mengancam keselamatan seseorang yang dilakukan pada individu, kelompok maupun komunitas. Tindak kejahatan tersebut melanggar hukum, undang-undang, serta norma yang berlaku di masyarakat. Menurut [1], tindak kriminalitas adalah

segala sesuatu perbuatan yang melanggar hukum dan melanggar norma-norma sosial, sehingga masyarakat menentangnya. Pelaku kejahatan ini tidak berpikir panjang karena ia tidak memikirkan konsekuensi atau dampak yang terjadi setelah ia melakukan perbuatannya, sehingga pelaku juga tidak segan – segan melukai korban bahkan membunuhnya.

Pada hal dampak yang dirasakan tidak hanya untuk dirinya sendiri tetapi juga dirasakan oleh keluarga dan orang – orang terdekatnya. Tindakan kriminalitas terjadi dikarenakan faktor ekonomi yang tidak mampu memenuhi standar hidup, faktor lingkungan yang mendorong untuk melakukan kejahatan, faktor kesenjangan sosial, faktor tekanan mental, serta faktor kebencian atau memiliki dendam kepada individu maupun kelompok. Tindakan ini bisa menyebabkan seseorang mengalami trauma dan juga bisa berpengaruh terhadap keamanan dan ketenangan masyarakat.

Kejahatan kriminalitas sedang menjadi hal yang sangat diperhatikan oleh pihak kepolisian. Kepolisian Republik Indonesia (Polri) mencatat, 276.507 kejahatan terjadi di Indonesia sepanjang 2022. Jumlah tersebut mengalami kenaikan 7,3% dibandingkan pada tahun sebelumnya yang sebanyak 257.743 perkara. Dengan demikian, ada 1 kejahatan setiap dua menit dua detik. Jika dihitung setiap jamnya, terdapat 31,6 kejahatan yang terjadi di dalam negeri[2]. Pada tahun 2022, Jawa Barat mengalami peningkatan angka tindak kriminalitas. Merujuk pada data Badan Pusat Statistik (BPS), pada 2021 Jawa Barat menduduki tindak kriminalitas pada posisi kesembilan di Indonesia[3]. Selanjutnya berdasarkan Kementerian Koordinator Bidang Politik Hukum dan Keamanan (Kemenko Polhukam) pada tahun 2022 Jawa Barat mengalami lonjakan angka kriminalitas hingga sebesar 69 persen, dan menduduki peringkat kelima di Indonesia[4].

*Data mining* bisa dikatakan sebagai suatu alat dan aplikasi menggunakan analisis statistik pada data. *Data mining* menggambarkan sebuah pengumpulan teknik-teknik dengan tujuan untuk menemukan pola-pola yang tidak diketahui pada data yang telah dikumpulkan[5]. Teknik *data mining* digunakan untuk mendapatkan informasi, seperti penelitian ini yang menggunakan *data mining* sebagai teknik analisis pada data kriminalitas di Jawa Barat.

Metodologi *Cross Industry Standard Process* for *Data mining* (CRISP-DM) yang digunakan pada penelitian ini merupakan suatu metodologi *Data mining* yang disusun oleh konsorsium perusahaan yang didirikan oleh Komisi Eropa pada tahun 1996 dan telah ditetapkan sebagai proses standar dalam *Data mining*[6]. Menurut [7], CRISP-DM memiliki enam fase, diantaranya: (1) *Business Understanding* adalah memahami situasi dan menetapkan tujuan yang ingin dicapai, (2) *Data Understanding* yaitu tahapan pengumpulan data, (3) *Data Preparation* meliputi segala kegiatan membangun dataset akhir untuk perhitungan, (4) *Modelling* fase implementasi model, (5) *Evaluation* yaitu mengevaluasi kualitas dari penelitian, dan (6) *Deployment* merupakan tahapan membuat laporan hasil penelitian.

*K-Medoids* atau algoritma *Partition Around Medoids* (PAM) dikembangkan oleh Leonard Kaufman dan Peter J. Rousseeuw pada tahun 1987. Algoritma PAM termasuk metode *Partitioning*

*Clustering* untuk mengelompokkan sekumpulan objek menjadi *cluster*. *Medoid* adalah representasi *cluster* pada PAM dari sekumpulan objek yang mewakili *cluster* [8]. Dengan langkah – langkah implementasinya sebagai berikut: (1) inialisasi pusat *cluster* sebanyak k (jumlah *cluster*), (2) alokasikan setiap data (objek) ke *cluster* terdekat menggunakan ukuran jarak *Euclidean Distance* seperti pada Persamaan 1,

$$d(x, y) = \|x - y\| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}; i = 1, 2, 3, \dots, n \quad (1)$$

(3) pilih secara acak objek pada masing-masing *cluster* sebagai kandidat *Medoid* baru, (4) hitung jarak setiap objek yang berada pada setiap masing-masing *cluster* dengan menempuh *Medoid* baru, (5) hitung total simpangan (S) dengan menghitung nilai total distance baru – total distance lama. Jika  $S > 0$ , maka ganti objek dengan data *cluster* untuk memperoleh sekelompok k objek yang baru sebagai *Medoids*, (6) ulangi tahap 3-5 hingga tidak terjadi perubahan *Medoids*, sehingga didapatkan *cluster* beserta anggota *cluster* masing-masing[9].

*Tools Rapidminer* dikhususkan untuk penggunaan *Data mining* sehingga dapat mempermudah dan mempercepat penelitian ini. Dengan menggunakan *RapidMiner*, tidak dibutuhkan kemampuan koding khusus, karena semua fasilitas sudah disediakan[10].

Beberapa penelitian terkait klasterisasi tindak kriminalitas telah dilakukan menggunakan metode *K-Means*[11][12], *Kernel Density*[13], *Association Rule FP-Growth*[14], dan *Fuzzy C-Means* Dan *K-Means*[15]. Pada penelitian [11] yang terfokus hanya pada 1 jenis kriminalitas saja, sedangkan pada penelitian ini menggunakan 9 jenis kriminalitas seperti pencurian, penipuan, penganiayaan, pembakaran, pemerkosaan, pengedar/penyalahgunaan narkoba, perjudian, pembunuhan, dan perdagangan manusia. Penelitian [12] selanjutnya berfokus di Kabupaten Karawang, Jawa Barat. Sedangkan penelitian ini berfokus pada seluruh Kabupaten/Kota di Jawa Barat. Penelitian [13] berikutnya menggunakan metode *Kernel Density* yang mengukur kerapatan berdasarkan peta dan dataset. Selain itu, penelitian tersebut berfokus di Kota Semarang, Jawa Tengah. Sedangkan penelitian ini menggunakan metode *K-Medoids* dan berfokus di Provinsi Jawa Barat. Selanjutnya pada penelitian [14] yang mengidentifikasi kriteria pelapor tindak kriminalitas di Kabupaten Tebo, Jambi. Sedangkan penelitian ini bertujuan untuk Klasterisasi Kabupaten/Kota di Provinsi Jawa Barat. Dan penelitian [15] yang membandingkan nilai uji dari kedua metode yang dipakai dan berfokus di Kota Semarang, Jawa Tengah. Sedangkan penelitian ini menggunakan dengan metode *K-Medoids* dan berfokus di Provinsi Jawa Barat. Berdasarkan 5 jurnal dari penelitian sebelumnya, maka diperlukan penelitian dengan metode lain, contohnya dengan

Algoritma *K-Medoids*. Kelebihan menggunakan *K-Medoids* pada penelitian ini adalah tidak sensitif pada *outlier* dan waktu perhitungan singkat dengan hasil yang stabil.

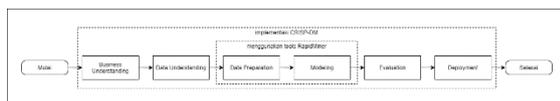
Data yang digunakan adalah data sekunder dari situs resmi Jawa Barat yaitu [opendata.jabarprov.go.id](http://opendata.jabarprov.go.id) dan penelitian ini menggunakan data kriminalitas yang terjadi di Jawa Barat pada tahun 2019 sampai 2021 sebanyak 15.937 baris. Penelitian ini menggunakan Algoritma *K-Medoids* yang akan diterapkan untuk Klasterisasi jenis kriminalitas pencurian, penipuan, penganiayaan, pembakaran, pemerkosaan, pengedar/penyalahgunaan narkoba, perjudian, pembunuhan, dan perdagangan manusia, sehingga hasil klasterisasinya lebih optimal.

Tujuan pada penelitian ini untuk menggunakan pendekatan *data mining* dalam mengelompokkan Kabupaten/Kota di Jawa Barat berdasarkan angka kriminalitas sehingga dapat membantu pihak kepolisian setempat untuk dapat melakukan pemetaan agar pencegahan serta penindakan kriminalitas dapat sesuai lebih efektif serta tegas, sehingga dapat menekan angka kriminalitas di Jawa Barat. Hasil penelitian ini adalah pengelompokan Kabupaten/Kota di Jawa Barat yang dibagi menjadi 2 *cluster*, yaitu *cluster* aman dan *cluster* rawan.

## 2. METODE PENELITIAN

### 2.1 Tahap Penelitian

Penelitian ini menggunakan metodologi *Cross Industry Standart Process Model for Data mining* (CRISP-DM) untuk membantu mencari solusi dari suatu masalah. Untuk tahapan penelitian dapat dilihat pada Gambar 1.



Gambar 1. Tahapan Penelitian

Berdasarkan Gambar 1, berikut penjelasannya: (1) *Business Understanding* merupakan tahap pertama yang dilakukan adalah memahami situasi dan menetapkan tujuan yang ingin dicapai. Pada penelitian ini memiliki tujuan untuk mengelompokkan wilayah di Jawa Barat berdasarkan tingkat kriminalitas berdasarkan angka kriminalitas di Jawa Barat yang meningkat dari tahun ke tahun. Kemudian (2) *Data Understanding* dilakukan pengumpulan data kriminalitas di Jawa Barat yang berkualitas serta memahami data. Data yang digunakan adalah data kriminalitas yang terjadi pada tahun 2019 sampai 2021 di Jawa Barat sebanyak 15.937 baris dengan membahas 9 jenis kriminalitas. Data tersebut diambil dari situs resmi Jawa Barat [opendata.jabarprov.go.id](http://opendata.jabarprov.go.id). Selanjutnya (3) *Data Preparation* melakukan *Data Preprocessing* atau mempersiapkan data untuk dimodelkan dengan bantuan *tools Rapidminer*, yang terdiri dari pemilihan data (*data selection*) menggunakan operator *Select*

*Attributes*, pembersihan data (*data cleaning*) menggunakan operator *Replace Missing Value*, transformasi terhadap data (*data transformation*) menggunakan operator *Map* dan *Rename*, agregasi data (*data aggregation*) menggunakan operator *Aggregate* dan normalisasi data (*data normalization*) menggunakan operator *Normalize*. Setelah itu tahapan (4) *Modeling* dengan penentuan metode yang dipakai dan implementasinya dengan *tools Rapidminer*. Diawali dengan mengkomparasi model dan jumlah *cluster* berdasarkan nilai DBI menggunakan operator *Cluster Distance Performance* dengan *main criterion Davies Bouldin* untuk mendapatkan *cluster* yang optimal lalu mengimplementasikan model menggunakan operator *Clustering (K-Medoids)* serta menghitung manual menggunakan persamaan *Euclidean Distance*. Kemudian tahapan (5) *Evaluation* dengan melakukan evaluasi kualitas dan apakah metode *K-Medoids Clustering* dapat memenuhi tujuan yang telah ditetapkan dengan membandingkan nilai DBI (*Davies Bouldin Index*) di beberapa jumlah *cluster*. Dan (6) *Deployment* adalah pembuatan laporan hasil penelitian yang telah dilakukan.

### 2.2 Tahap Penelitian

Selanjutnya dilakukan pengumpulan teori pendukung yang diambil dari buku, penelitian sebelumnya hingga media *online* yang berhubungan dengan topik penelitian.

## 3. HASIL DAN PEMBAHASAN

### 3.1 Business Understanding

Jawa Barat mengalami lonjakan angka kriminalitas pada tahun 2022. Oleh karena itu, dibutuhkan kerjasama seluruh pihak untuk menekan angka kriminalitas di Jawa Barat. Sebelum itu, dibutuhkan informasi yang valid mengenai daerah-daerah di Jawa Barat yang memiliki jumlah kriminalitas tinggi.

### 3.2 Data Understanding

Pada penelitian ini, penulis menggunakan data kriminalitas yang terjadi pada tahun 2019 sampai 2021 di Jawa Barat sebanyak 15.937 baris dan memiliki 23 atribut dengan membahas 9 jenis kriminalitas. Data tersebut diambil dari situs resmi Jawa Barat [opendata.jabarprov.go.id](http://opendata.jabarprov.go.id).

Data asli yang didapatkan terdapat 23 atribut diantaranya adalah *kode\_provinsi*, *nama\_provinsi*, *bps\_kode\_kabupaten\_kota*, *bps\_nama\_kabupaten\_kota*, *bps\_kode\_kecamatan*, *bps\_nama\_kecamatan*, *bps\_kode\_desa\_kelurahan*, *bps\_nama\_desa\_kelurahan*, *kemendagri\_kode\_kecamatan*, *kemendagri\_nama\_kecamatan*, *kemendagri\_kode\_desa\_kelurahan*, *kemendagri\_nama\_desa\_kelurahan*, *terdapat\_pencurian*, *terdapat\_penipuan*, *terdapat\_penganiayaan*, *terdapat\_pembakaran*, *terdapat\_perkosaan*, *terdapat\_pengedar\_penyalahgunaan\_narkoba*, *terdapat\_perjudian*, *terdapat\_pembunuhan*, *terdapat\_perdagangan\_manusia*, *tahun*, dan

id. Untuk tampilan data asli dapat dilihat pada Tabel 1.

Tabel 1. Tampilan Data Asli

kode_provinsi	nama_provinsi	bps_kode_kabupaten_kota	bps_nama_kabupaten_kota	bps_kode_kecamatan	--	id
32	JAWA BARAT	3201	KABUPATEN BOGOR	3201190	--	1
32	JAWA BARAT	3201	KABUPATEN BOGOR	3201190	--	2
32	JAWA BARAT	3201	KABUPATEN BOGOR	3201190	--	3
32	JAWA BARAT	3201	KABUPATEN BOGOR	3201190	--	4
32	JAWA BARAT	3201	KABUPATEN BOGOR	3201190	--	5
32	JAWA BARAT	3201	KABUPATEN BOGOR	3201190	--	6
32	JAWA BARAT	3201	KABUPATEN BOGOR	3201190	--	7
32	JAWA BARAT	3201	KABUPATEN BOGOR	3201190	--	8
--	--	--	--	--	--	--
32	JAWA BARAT	3279	KOTA BANJAR	3279040	--	15936

### 3.3 Data Preparation

*Data Preprocessing* yang dilakukan pada penelitian ini adalah sebagai berikut:

1. Menyeleksi atribut yang berkesinambungan dengan topik atau masalah. Data asli memiliki 23 atribut lalu di seleksi menjadi 10 atribut untuk diteliti.
2. Pembersihan data dengan menghilangkan *missing value* dan jika ditemukan *missing value* maka satu row atau baris akan dihapus.
3. Transformasi data dilakukan dengan pemberian bobot pada atribut yang menyatakan ada atau tidaknya kejadian pada jenis kriminalitas. jika

menyatakan “ADA” berbobot 1 dan “TIDAK ADA” berbobot 0.

4. Inisialisasi nama atribut menggunakan abjad.
5. Inisialisasi pada value atribut nama\_kabupaten\_kota atau id menggunakan angka.
6. Mengagregasi atribut jenis – jenis kriminalitas berdasarkan nama Kabupaten/Kota.
7. Normalisasi data menggunakan *Min Max Normalization* dengan jarak diantara 0 sampai 1. Tampilan data setelah tahapan *Preprocessing* dapat dilihat pada Tabel 2.

Tabel 2. Data setelah Tahapan *Data Preprocessing*

id	a	b	c	d	e	f	g	h	i
1	0,423	0,513	0,597	1	0,327	0,594	0,6	0,233	0,535
2	0,462	0,231	0,357	0,607	0,226	0,427	0,56	0,176	0,628
3	0,885	1	0,425	0,738	0,56	0,669	0,92	0,382	0,791
4	1	0,692	1	0,762	1	1	0,64	0,385	0,977
5	0,462	0,385	0,471	0,303	0,101	0,27	0,2	0,073	0,372
6	0,885	0,538	0,706	0,672	0,327	0,645	0,92	0,29	0,767
7	0,231	0,231	0,752	0,902	0,393	0,58	0,16	0,477	0,465
8	0,846	0,436	0,857	0,598	0,14	0,447	0,6	0,103	0,605
9	0,885	0,615	0,716	0,852	0,521	0,823	1	1	1
10	0,808	0,615	0,747	0,77	0,778	0,686	0,88	0,718	0,86
11	0,269	0,026	0,247	0,057	0,148	0,164	0,04	0,024	0,14
12	0,385	0	0,566	0,295	0,195	0,321	0,08	0,168	0,256
13	0,115	0,026	0,117	0,148	0,023	0,126	0	0,024	0,07
14	0,346	0,128	0,304	0,246	0,276	0,242	0,32	0,06	0,302
15	0,577	0,256	0,497	0,344	0,288	0,392	0,6	0,29	0,535
16	0,731	0,308	0,812	0,705	0,533	0,683	0,72	0,195	0,86
17	0,846	0,538	0,442	0,443	0,241	0,379	0,64	0,127	0,628
18	0,385	0,308	0,602	0,59	0,249	0,454	0,44	0,1	0,465
19	0	0,051	0	0	0	0	0	0	0

### 3.4 Modeling

Mengkomparasi model dilakukan untuk menemukan metode dengan jumlah *cluster* terbaik. Komparasi model dilakukan dengan membandingkan nilai DBI (*Davies Bouldin Index*) berdasarkan jumlah *cluster* pada setiap algoritma atau model. Pengukuran DBI menggunakan cara pemaksimalan jarak inter *cluster* juga meminimalkan jarak intra *cluster* dinyatakan bahwa semakin kecil nilai DBI maka hasil menunjukkan skema klaster yang paling baik[16]. Algoritma yang dibandingkan antara lain, algoritma *K-Means*, dan algoritma *K-Medoids*. Tahapan ini dilakukan menggunakan *tools Rapidminer*. Berikut merupakan hasil komparasi model dari ketiga

algoritma tersebut: Untuk perbandingan nilai DBI dapat dilihat pada Tabel 3.

Tabel 3. Perbandingan Nilai DBI

Jumlah Cluster	Algoritma	
	K-Means	K-Medoids
2 cluster	0,897	0,761
3 cluster	0,972	1,616
4 cluster	1,059	0,986

Berdasarkan Tabel 3 maka diketahui jumlah *cluster* terbaik pada algoritma *K-Medoids* berjumlah 2 *cluster*. Hal tersebut karena jumlah *cluster* yang terbaik adalah nilai yang paling mendekati nilai 0 atau nilai yang paling kecil. Nilai DBI terbaik berjumlah 2

cluster dengan nilai DBI sebesar 0,761. Inisialisasi pusat cluster sebanyak k (jumlah cluster).

sebagai Cluster 0 dan id 15 sebagai Cluster 1 seperti pada Tabel 4.

Menentukan *Centroid* awal pada tiap cluster secara acak pada dataset. Penulis mengambil id 19

**Tabel 4.** Centroid Awal

	id	a	b	c	d	e	f	g	h	i
<b>Cluster 0</b>	19	0	0,051	0	0	0	0	0	0	0
<b>Cluster 1</b>	15	0,577	0,256	0,497	0,344	0,288	0,392	0,6	0,29	0,535

1. Alokasikan setiap data (objek) ke cluster terdekat menggunakan ukuran jarak *Euclidean Distance*. Menghitung jarak terdekat menggunakan

*Euclidean Distance* seperti pada Persamaan 1. Hasil perhitungannya dapat dilihat pada Tabel 5.

**Tabel 5.** Hasil Perhitungan Iterasi 1

Id	Cluster 0	Cluster 1	Jarak Terdekat	Cluster
1	1,705	0,758	0,758	Cluster 1
2	1,309	0,361	0,361	Cluster 1
3	2,188	1,065	1,065	Cluster 1
4	2,550	1,371	1,371	Cluster 1
5	0,951	0,562	0,562	Cluster 1
6	2,003	0,740	0,740	Cluster 1
7	1,557	0,881	0,881	Cluster 1
8	1,706	0,603	0,603	Cluster 1
9	2,508	1,284	1,284	Cluster 1
10	2,287	1,059	1,059	Cluster 1
11	0,456	0,951	0,456	Cluster 0
12	0,894	0,698	0,698	Cluster 1
13	0,267	1,112	0,267	Cluster 0
14	0,782	0,570	0,570	Cluster 1
15	1,307	0	0	Cluster 1
16	1,950	0,728	0,728	Cluster 1
17	1,538	0,451	0,451	Cluster 1
18	1,270	0,428	0,428	Cluster 1
19	0	1,307	0	Cluster 0
<b>Total Cost</b>			12,281	

2. Pilih secara acak objek pada masing-masing cluster sebagai kandidat *Medoid* baru. Selanjutnya dilakukan perhitungan Iterasi-2, maka menentukan *Medoids* baru secara acak.

Penulis mengambil id 13 sebagai Cluster 0 dan id 5 sebagai Cluster 1 seperti pada Tabel 6.

**Tabel 6.** Medoids Baru

id	a	b	c	d	e	f	g	h	i	
<b>Cluster 0</b>	13	0,462	0,385	0,471	0,303	0,101	0,270	0,2	0,073	0,372
<b>Cluster 1</b>	5	0,115	0,026	0,117	0,148	0,023	0,126	0	0,024	0,07

3. Hitung jarak setiap objek yang berada pada setiap masing-masing cluster dengan menempuh *Medoids* baru.

Tahapan selanjutnya adalah mengitung jarak terdekat menggunakan *Euclidean Distance*. Hasil perhitungannya dapat dilihat pada Tabel 7.

**Tabel 7.** Hasil Perhitungan Iterasi 2

Id	Cluster 0	Cluster 2	Jarak Terdekat	Cluster
1	1,490	0,942	0,942	Cluster 1
2	1,100	0,612	0,612	Cluster 1
3	2,015	1,381	1,381	Cluster 1
4	2,338	1,695	1,695	Cluster 1
5	0,748	0	0	Cluster 1
6	1,800	1,143	1,143	Cluster 1
7	1,325	0,932	0,932	Cluster 1
8	1,491	0,796	0,796	Cluster 1
9	2,312	1,723	1,723	Cluster 1
10	2,092	1,489	1,489	Cluster 1
11	0,269	0,610	0,269	Cluster 0
12	0,652	0,459	0,459	Cluster 1
13	0	0,748	0	Cluster 0
14	0,586	0,402	0,402	Cluster 1
15	1,112	0,562	0,562	Cluster 1
16	1,730	1,112	1,112	Cluster 1
17	1,354	0,697	0,697	Cluster 1
18	0,071	0,399	0,071	Cluster 0
19	0,428	0,100	0,100	Cluster 1
<b>Total Cost</b>			14,698	

4. Hitung total simpangan ( $S$ ) dengan menghitung nilai total distance baru – total distance lama. Jika  $S > 0$ , maka ganti objek dengan data *cluster* untuk memperoleh sekelompok  $k$  objek yang baru sebagai *Medoids*.

Setelah itu menghitung total simpangan ( $S$ ) dengan menghitung perbandingan total *cost* baru – total *cost* lama sampai  $S > 0$  seperti Persamaan 2.

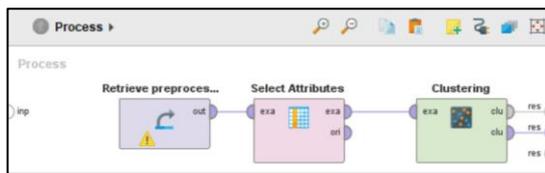
$$S = \text{total cost baru} - \text{total cost lama} \quad (2)$$

$$S = 14,698 - 12,281$$

$$S = 2,417$$

Karena nilai  $S > 0$  maka proses clustering dihentikan. Cluster 0 memiliki anggota 3 anggota dan Cluster 1 memiliki 16 anggota.

Selanjutnya dilakukan implementasi metode menggunakan *tools Rapidminer*. Menggunakan operator *Clustering K-Medoids* untuk klasterisasi seperti pada Gambar 2 dan untuk hasil *cluster* yang diperoleh dapat dilihat pada Gambar 3.



Gambar 2. Klasterisasi dengan Rapidminer

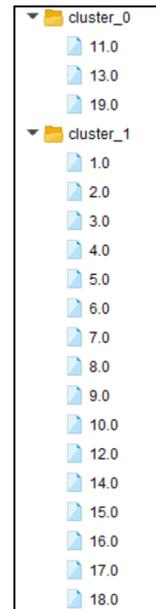
```

Cluster Model
Cluster 0: 3 items
Cluster 1: 16 items
Total number of items: 19
    
```

Gambar 3. Hasil Cluster Algoritma K-Medoids

Maka dapat diketahui total kabupaten/kota yang termasuk *cluster* 0 sebanyak 3 *items*, sedangkan *cluster* 1 sebanyak 16 *items*. Untuk mengetahui *items* dari masing-masing *cluster* dapat dilihat pada Gambar 4.

Dapat dilihat kabupaten/kota yang termasuk ke *cluster\_0* adalah Kabupaten Kuningan, Kabupaten Pangandaran, dan Kota Banjar. Sedangkan *cluster\_1* terdiri dari Kabupaten Bandung, Kabupaten Bandung Barat, Kabupaten Bekasi, Kabupaten Bogor, Kabupaten Ciamis, Kabupaten Cianjur, Kabupaten Cirebon, Kabupaten Garut, Kabupaten Indramayu, Kabupaten Karawang, Kabupaten Majalengka, Kabupaten Purwakarta, Kabupaten Subang, Kabupaten Sukabumi, Kabupaten Sumedang, dan Kabupaten Tasikmalaya.



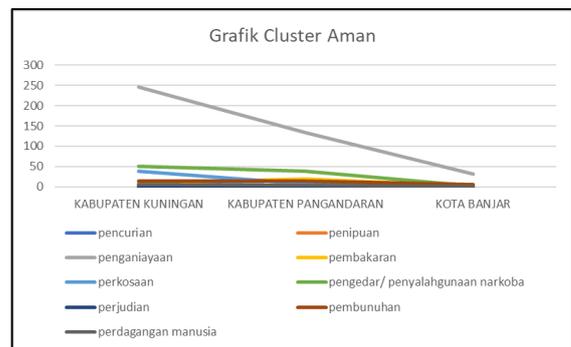
Gambar 4. Hasil Anggota tiap Cluster

Berikut merupakan nilai minimal, maksimal dan rata-rata dari masing-masing *cluster* seperti pada Gambar 4.

Tabel 8. Hasil Nilai Minimal, Maksimal, dan Rata-rata anggota cluster

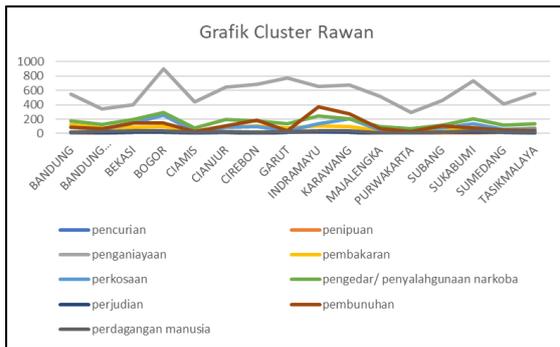
cluster	Nilai	Berdasarkan
		Jumlah Kriminalitas
cluster_0	Minimal	45
	Maksimal	374
	Rata-rata	213,3333333
cluster_1	Minimal	535
	Maksimal	1805
	Rata-rata	1092,3125

Berdasarkan Tabel 8, maka *cluster\_0* merupakan Cluster Aman dan *cluster\_1* merupakan Cluster Rawan. Untuk grafik masing-masing *cluster* dapat dilihat pada Gambar 5 dan Gambar 6.



Gambar 5. Grafik Cluster Aman

Berdasarkan grafik pada Gambar 5, pada Cluster Aman dapat diketahui bahwa tindak kriminalitas paling banyak terjadi di Kabupaten Kuningan dan jenis kriminalitas yang paling sering terjadi adalah penganiayaan.



Gambar 6. Grafik Cluster Rawan

Berdasarkan grafik pada Gambar 6, pada Cluster Rawan dapat diketahui bahwa tindak kriminalitas paling banyak terjadi di Kabupaten Bogor dan jenis kriminalitas yang paling sering terjadi adalah penganiayaan.

### 3.5 Evaluation

Tahapan *evaluation* atau pengujian pada klusterisasi dilakukan dengan membandingkan nilai DBI (*Davies Bouldin Index*) algoritma *K-Medoids* pada jumlah 2 *cluster* hingga 5 *cluster* yang dapat dilihat pada Tabel 9.

Jumlah Cluster	Nilai DBI
2 cluster	0,761
3 cluster	1,616
4 cluster	0,986
5 cluster	1,088

Nilai DBI paling kecil berada di jumlah 2 *cluster* sebesar 0,761 dan jumlah *cluster* tersebut dipakai untuk penelitian ini. Jadi, dapat disimpulkan jumlah *cluster* yang dipakai pada penelitian ini yang paling optimal.

## 4. KESIMPULAN

Berdasarkan penelitian menggunakan *K-Medoids Clustering* dengan *tools Rapidminer* pada data tindak kriminalitas di Jawa Barat, maka dapat disimpulkan penerapan algoritma *K-Medoids* dapat menentukan 2 buah *cluster* yang paling optimal dengan nilai DBI sebesar 0,761. Penggunaan *tools Rapidminer* pada penelitian memiliki hasil yang sama dengan perhitungan manual. Maka dari itu, hasil dari penggunaan *tools Rapidminer* dapat dinyatakan akurat.

## DAFTAR PUSTAKA

[1] K. Kartono, *Patologi Sosial*, 2(1). Jakarta: Rajagrafindo persada, 1999.

[2] S. Sadya, "Polri: Kejahatan di Indonesia Naik Jadi 276.507 Kasus pada 2022 Artikel ini telah tayang di DataIndonesia.id dengan judul 'Polri: Kejahatan di Indonesia Naik Jadi 276.507 Kasus pada 2022'.. Author: Sarnita Sadya. Editor: Dimas Bayu. Klik selengkapnya di s," *dataindonesia.id*, 2023. <https://dataindonesia.id/varia/detail/polri->

kejahatan-di-indonesia-naik-jadi-276507-kasus-pada-2022 (accessed Nov. 12, 2023).

[3] D. Setyowati, R. Dewi, N. Sari, and R. Fadhila Indra Putra, *Statistik Kriminal* 2022. 2022.

[4] Deputi V Bidkor Kamnas, "Polda Jabar Jadi Target Sasaran Pengendalian Tingkat Kriminalitas Tinggi, Asdep Penanganan Kejahatan Konvensional Kemenko Polhukam: Butuh Kerjasama Seluruh Pihak Ambil Langkah Strategis," *KEMENKO POLHUKAM RI*, 2023. <https://polkam.go.id/polda-jabar-jadi-target-sasaran-pengendalian-tingkat-kriminalitas-tinggi-asdep-penanganan-kejahatan-konvensional-kemenko-polhukam-butuh-kerjasama-seluruh-pihak-ambil-langkah-strategis/> (accessed Dec. 12, 2023).

[5] F. Marisa, S. S. S. Ahmad, Z. I. M. Yusoh, T. M. Akhriza, W. Purnomowati, and R. K. Pandey, "Performance Comparison of Collaborative-Filtering Approach with Implicit and Explicit Data," *Int. J. Adv. Comput. Sci. Appl.*, vol. 10, no. 10, 2019, doi: 10.14569/IJACSA.2019.0101016.

[6] R. Wirth and J. Hipp, "CRISP-DM: towards a standard process model for *data mining*. Proceedings of the Fourth International Conference on the Practical Application of Knowledge Discovery and *Data mining*, 29-39," *Proc. Fourth Int. Conf. Pract. Appl. Knowl. Discov. Data Min.*, no. 24959, pp. 29-39, 2000, [Online]. Available: [https://www.researchgate.net/publication/239585378\\_CRISP-DM\\_Towards\\_a\\_standard\\_process\\_model\\_for\\_data\\_mining](https://www.researchgate.net/publication/239585378_CRISP-DM_Towards_a_standard_process_model_for_data_mining).

[7] C. D. Larose, Daniel T.; Larose, "An Introduction to *Data mining*," in *Discovering Knowledge in Data*, WILEY Online Library, 2014, pp. 1-15.

[8] E. Febrianti, R. W. Sembiring, and D. Suhada, "Mengkluster Jumlah Kabupaten/Kota Yang Melaksanakan Kawasan Tanpa Rokok (Ktr) Di 50% Sekolah Menurut Provinsi Dengan *K-Medoids*," *KOMIK (Konferensi Nas. Teknol. Inf. dan Komputer)*, vol. 3, no. 1, 2019, doi: 10.30865/komik.v3i1.1672.

[9] N. Pulungan, S. Suhada, and D. Suhendro, "Penerapan Algoritma *K-Medoids* Untuk Mengelompokkan Penduduk 15 Tahun Keatas Menurut Lapangan Pekerjaan Utama," *KOMIK (Konferensi Nas. Teknol. Inf. dan Komputer)*, vol. 3, Nov. 2019, doi: 10.30865/komik.v3i1.1609.

[10] S. Hendrian, "Algoritma Klasifikasi *Data mining* Untuk Memprediksi Siswa Dalam Memperoleh Bantuan Dana Pendidikan," *Fakt. Exacta*, vol. 11, no. 3, pp. 266-274, 2018, doi: 10.30998/faktorexacta.v11i3.2777.

[11] U. T. Suryadi and Y. Supriatna, "Sistem Clustering Tindak Kejahatan Pencurian Di Wilayah Jawa Barat Menggunakan Algoritma *K-Means*," *J. Teknol. Inf. dan Komun.*, vol. 22, no. April, p. 111, 2019.

[12] R. N. Fahmi, M. Jajuli, and N. Sulistiyowati, "Analisis Pemetaan Tingkat Kriminalitas di Kabupaten Karawang menggunakan Algoritma *K-Means*," *INTECOMS J. Inf. Technol. Comput. Sci.*, vol. 4, no. 1, pp. 67-79, 2021, doi: 10.31539/intecomsv4i1.2413.

[13] C. A. Nanda, A. L. Nugraha, and H. S. Firdaus,

- “Analisis Tingkat Daerah Rawan Kriminalitas Menggunakan Metode Kernel Density Di Wilayah Hukum Polrestabes Kota Semarang,” *J. Geod. Undip*, vol. 8, no. 4, pp. 50–58, 2019.
- [14] D. Winarti, M. Kom, E. Revita, and M. Kom, “Penerapan *Data mining* untuk Analisa Tingkat Kriminalitas Dengan Algoritma Association Rule Metode FP-Growth,” *J. SIMTIKA*, vol. 4, no. 3, pp. 8–22, 2021.
- [15] H. S. Firdaus, A. L. Nugraha, B. Sasmito, and M. Awaluddin, “Perbandingan Metode Fuzzy C-Means Dan K-Means Untuk Pemetaan Daerah Rawan Kriminalitas Di Kota Semarang,” *Elipsoida J. Geod. dan Geomatika*, vol. 4, no. 01, pp. 58–64, 2021, doi: 10.14710/elipsoida.2021.9219.
- [16] Y. A. Wijaya, D. A. Kurniady, E. Setyanto, W. S. Tarihoran, D. Rusmana, and R. Rahim, “Davies Bouldin Index Algorithm for Optimizing Clustering Case Studies Mapping School Facilities,” *TEM J.*, vol. 10, no. 3, pp. 1099–1103, 2021, doi: 10.18421/TEM103-13.