Journal of Innovative and Creativity
Journal Homepage: https://joecy.org/index.php/joecy
e-mail: journaljoecy@gmail.com

Faculty of Education
University of Pahlawan Tuanku Tambusai
ISSN 2775-771X (Online)
ISSN 2962-570X (Printed)
6 (1) 2026, Page: 1900-1906

Klasterisasi Wilayah Rawan Tindak Kriminal Menggunakan Algoritma *K-Means* . (Studi Kasus Di Wilayah Hukum Pengadilan Negeri Painan)

Sando Eka Putra[1], Sintia[2], Diana Kemala Odang[3]
[123]Sistem Informasi, Sistem Informasi, Informatika Universitas Alifah Padang
[1]sando@alifah.ac.id, [2]sintia@alifah.ac.id, [3]odkemala@gmail.com

Published: January, 2026

ABSTRAK

Meningkatnya jumlah kasus kriminal setiap tahun di wilayah hukum Pengadilan Negeri Painan menjadi perhatian penting bagi aparat penegak hukum. Oleh karena itu, diperlukan penyediaan informasi yang bermanfaat bagi masyarakat mengenai daerah-daerah yang tergolong rawan tindak kriminal. Salah satu metode yang dapat digunakan untuk mengidentifikasi daerah rawan tersebut adalah klasterisasi dengan algoritma K-Means.

Penelitian ini bertujuan untuk mengelompokkan daerah berdasarkan tingkat kerawanan tindak kriminal di wilayah hukum Pengadilan Negeri Painan menggunakan algoritma K-Means. Metode ini termasuk dalam teknik data mining yang berfungsi untuk memisahkan atau mengelompokkan data ke dalam beberapa kelompok (cluster) berdasarkan kesamaan karakteristik tertentu.

Hasil penelitian menunjukkan bahwa penerapan algoritma K-Means dengan bantuan aplikasi RapidMiner 5.3 mampu mengelompokkan wilayah ke dalam beberapa kategori tingkat kerawanan, yaitu rawan, kurang rawan, dan aman. Informasi yang dihasilkan dari penelitian ini diharapkan dapat membantu lembaga penegak hukum dalam melakukan pemetaan daerah rawan kejahatan serta mendukung pengambilan keputusan yang lebih efektif dalam upaya pencegahan tindak kriminal.
Kata kunci: Klasterisasi, Kriminal, Algoritma K-Means, Daerah, Data Persidangan.

© The Author(s). 2021 Open Access This article is licensed under a Creative Commons Atibution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0

PENDAHULUAN

Salah satu permasalahan yang dihadapi oleh Pengadilan Negeri Painan adalah belum adanya pemetaan yang jelas terkait tingkat kerawanan daerah terhadap tindak kriminal. Kondisi ini menyebabkan pihak penegak hukum kesulitan dalam menentukan prioritas wilayah yang perlu mendapatkan perhatian lebih dalam upaya pencegahan kejahatan. Oleh karena itu, dibutuhkan suatu metode yang dapat digunakan untuk mengelompokkan daerah berdasarkan tingkat kerawanan tindak kriminal.

Data mining merupakan suatu proses penggalian atau penyaringan informasi dari kumpulan data berukuran besar melalui tahapan tertentu untuk memperoleh informasi yang bermanfaat dan memiliki nilai tambah. Dengan memanfaatkan teknik data mining, pola dan hubungan tersembunyi dalam data dapat diungkapkan, sehingga hasilnya dapat digunakan sebagai dasar pengambilan keputusan yang lebih akurat. Salah satu metode dalam data mining yang dapat digunakan untuk pengelompokan wilayah rawan kejahatan adalah algoritma K-Means Clustering, yang berfungsi untuk mengelompokkan data ke dalam beberapa kelompok berdasarkan kesamaan karakteristik tertentu. [1].

TINJAUAN LITERATUR

Penelitian di bidang klasterisasi telah banyak dikembangkan, terutama untuk tipe data numerik. Dalam penerapannya, terdapat berbagai jenis data yang dapat digunakan dalam proses data mining, di antaranya adalah:

1. Data nominal, yaitu data yang bersifat kategorikal tanpa urutan tertentu
2. Data ordinal, yaitu data yang memiliki urutan tetapi tanpa jarak yang pasti;
3. Data interval, yaitu data dengan jarak yang terukur namun tanpa titik nol absolut; dan
4. Data rasio, yaitu data yang memiliki jarak terukur serta titik nol absolut.

Metode K-Means merupakan salah satu teknik clustering non-hirarki yang bertujuan untuk membagi data ke dalam sejumlah cluster, di mana data dalam satu cluster memiliki kemiripan karakteristik, sedangkan antar-cluster memiliki perbedaan yang signifikan. Dengan demikian, metode ini memungkinkan peneliti untuk memahami distribusi dan pola data dengan lebih mudah.

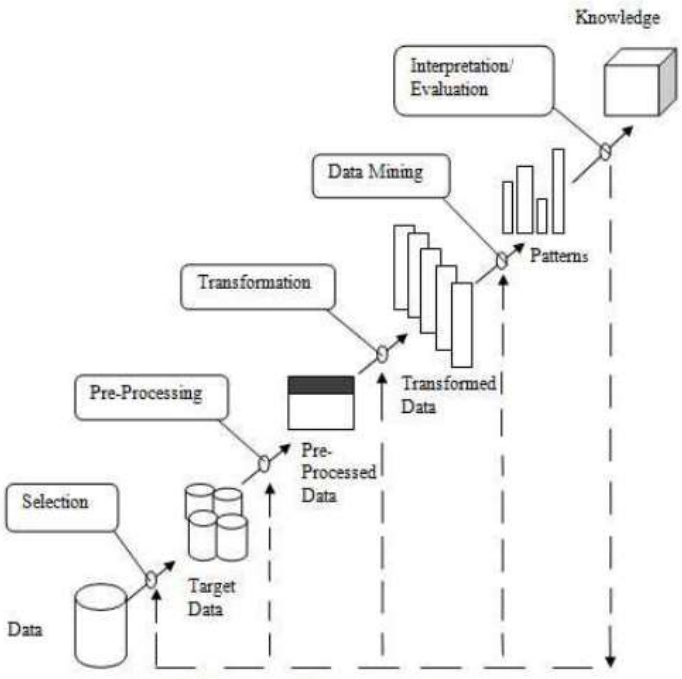
Menurut berbagai penelitian sebelumnya, metode K-Means sering digunakan dalam berbagai bidang, seperti analisis pelanggan, prediksi pola pembelian, dan pengelompokan wilayah berdasarkan indikator sosial maupun ekonomi. Dengan demikian, penerapan metode ini pada analisis wilayah rawan tindak kriminal diharapkan dapat memberikan gambaran yang lebih jelas mengenai tingkat kerawanan setiap daerah. [4].

METODOLOGI

Dalam penelitian ini bertujuan untuk mengelompokkan wilayah yang rawan tindak kriminal, Oleh karena itu dilakukan penelitian dengan mengelompokkan wilayah yang rawan tindak kriminal menggunakan algoritma *k-means clustering.*

3.1. *Knowledge discovery in Database*

Knowledge Discovery in Databases (KDD) adalah penerapan metode saintifik pada data mining. Dalam konteks ini data mining merupakan satu langkah dari proses KDD[5]. Sedangkan definisi lain dari *Knowledge discovery in Database* (KDD) didefinisikan sebagai ekstraksi informasi potensial, implisit dan tidak dikenal dari sekumpulan data. Proses *knowledge discovery* melibatkan hasil dari proses *Data Mining* (proses mengekstrak kecenderungan pola suatu data) dimana pola yang ditemukan bersifat sah, baru, dapat bermanfaat dan dapat dimengerti[6].



Gambar 1 Proses *Data Mining*

3.2. *Data Mining*

K-means is an algorithm that be used in partial classification that separate datas into different groups. This algorithm is able to minimalize the gap between data to its cluster. Basically, the application of this algorithm in the clustering process depends on the received datas and the conclusions that expected to be achieved in the end of the process. [7]. Salah satu teknik yang dibuat dalam data mining adalah adalah bagaimanana menelusuri data yang ada untuk membangun sebuah model, kemudian menggunakan model tersebut agar dapat mengenali pola data yang lain yang tidak berada dalam basis data yang tersimpan[8].

3.3. *Clusterisasi*

Clustering atau mengklasifikasikan adalah metode itu digunakan untuk membagi sirkuit data menjadi beberapa kelompok berdasarkan kesamaan yang dimiliki ditentukan sebelumnya. *Cluster* adalah sekelompok data serupa di *cluster* yang sama, dan berbeda dengan objek *cluster* lain. Obyek akan diklasifikasikan menjadi satu atau lebih kelompok, Metode ini mempartisi data ke dalam kelompok sehingga data berkarakteristik sama dimasukkan ke dalam satu kelompok yang sama dan data yang berkarakteristik berbeda dikelompokkan ke dalam kelompok yang lain. [9].
Berikut ini adalah rumus untuk menentukan jumlah *cluster*[10].

$$K \sqrt{\frac{N}{2}} \quad (1)$$

Keterangan :
K= klaster
N= jumlah data

Menghitung jarak pada ruang jarak *Euclidean* menggunakan formula :

$$D(x_2\, x_1) = \|x_{2-}x_1\|2 = \sqrt{\sum_p^e |x_{2j} - x_{1j}|\, 2} \quad (2)$$

Keterangan :
D = *euclidean distance*
x = banyaknya objek
$\sum p$ = jumlah data *record*

3.4. Algoritma *K-Means*

Algoritma K-Means merupakan salah satu teknik pengelompokan data yang sering digunakan. *K-Means* mengelompokkan objek yang mirip dalam *cluster* yang sama. Gambaran umum dari algoritma tersebut adalah sebagai berikut [11].

1. Menentukan nilai "K" atau jumlah *cluster* yang diinginkan serta menentukan *centroid* atau pusat *cluster*. *Centroid* ditentukan secara acak dari objek data yang ada sesuai dengan nilai "K", jika nilai "K" adalah dua maka *centroid* yang ditentukan juga dua.
2. Membagi setiap objek ke dalam *cluster* yang paling mirip. *Cluster* yang paling mirip adalah *cluster* dengan *centroid* terdekat. Perhitungan kedekatan antara objek dengan centroid dapat dilakukan menggunakan fungsi jarak, seperti *Euclidean Distance*.
3. Menghitung ulang "K" *cluster* dengan rata-rata dari semua objek yang diberikan untuk tiap cluster.
4. Ulangi langkah kedua dan ketiga sampai tidak ada lagi objek yang berpindah *cluster*.

Adapun formula yang digunakan untuk k-means yaitu:

$$d_{ij} = \sqrt{\sum_{k=1}^{p} \{x_{jk} - x_{jk}\} 2} \quad (3)$$

dimana:
d : distance
i : i1, i2, i3,……, ip
j : j1, ji2, ji3,……, jip
k : nilai atribut
n : dimensi data
x : objek data

Algoritma K-means mudah diimplementasikan Pada metode ini, peneliti menentukan sendiri jumlah klaster yang akan dibentuk. Peneliti mengelompokkan entitas ke dalam k-kelompok, biasanya dilakukan secara acak. Pada masing-masing kelompok dihitung rata-ratanya. Hitung jarak setiap entitas terhadap pusat masing-masing kelompok (rata-rata kelompoknya)[12].

HASIL DAN DISKUSI

Disini penulis menjelaskan langkah langkah yang digunakan dalam menyelesaikan masalah yang di hadapi yaitu menganalisa masalah menggunakan metode K-Means.
1.Data yang akan di masukandan di olah dalam penelitian ini yaitu data yang sudah disidang.
2.Kemudian menentukan jumlah kelompok atau cluster yang akan digunakan dalam penelitian ini.

3.Selajutnya dilakukan analisa dengan rincian sebagai berikut ini:
a.Menentukan titik pusat cluster awal.
b.Menentukan jarak euclidean.
c.Menentukanletakcluster
d.Menentukan pusat cluster barujika data yang di olah belum mempunyai hasiil yang sama dan apabila hasil pengolahan sudah sama maka tahapan di nyatakan selesai.

4.langkah selanjutnya yaitu hasil tersebut dibandingkan untuk melihat perbedaan dan kesamaan hasil analisa data menggunakan Software *Rapid Miner*. Data yang diambil merupakan 3 (tiga) tahun terakhir yaitu 2018, 2017, 2016 yang ditunjukkan pada tabel 1.

Tabel 1: Data Awal

No	Daerah	Jenis Kejahatan	Status Perkara
1	Sungai Lundang	1	1
2	Barung Barung Belantai	2	3
3	Kapuh utara	3	1
4	Talaok	2	2
5	Ampuan Lumpo	5	2
6	Sawah Laweh	3	4
7	Kapeh Panji	2	1
8	Gurun Panjang Selatan	4	1
9	Gurun Panjang	4	1
10	Puluik Puluik	4	1
11	Limau Gadang	4	1
12	Salido	4	1
13	Sago Salido	2	1
14	Painan	2	1
15	Painan Selatan	6	1
16	IV Koto Hilir	7	1
17	Surantih	1	1
18	Sago	1	1
19	Batang Kapas	4	1
20	Cerocok Tarusan	4	1

4.1 Proses Data

Dalam penelitian ini, pengelompokkan data berdasarkan tingkat kerawanan yang terjadi pada suatu daerah.
1. Menentukan Nilai *k* dari jumlah *cluster* yang akan di bentuk. Di mana jumlah *cluster* yang akan digunakan yaitu sebanyak 3 *cluster*.
a. C0 Merupakan Rawan
b. C1 Merupakan Kurang Rawan
a. C2 Merupakan Aman
2.Menentukan Titik Pusat (*centroid*) awal dari setiap *cluster*, seperti pada tabel 2 .

Tabel 2: *Centroid* Awal

Centroid Awal		
C0	1	1
C1	2	3

3. Setelah titik pusat *cluster* di tentukan, maka tahap selanjutnya adalah menghitung jarak terdekat atau euclidean data tingkat kerawanan pada daerah Nol ke titik pusat *cluster*.

1. Proses Hitung *cluster* (C_0)
   a. Pada *stage* (D_1), jarak dengan titik pusat *cluster* (C_0)yaitu:

$X_1 = (1 : 1)$ $C_0 = (1 : 1)$

$D_1 = \sqrt{(X1_1 - C0_1)^2 + (X1_2 - C0_2)^2}$

$= \sqrt{(1-1)^2 + (1-1)^2}$

$= 0{,}0000$

   b. Pada *stage* (D_2), jarak dengan titik pusat *cluster* (C_0) yaitu:

$X_2 = (2 : 3)$ $C_0 = (1 : 1)$

$D_2 = \sqrt{(X2_1 - C0_1)^2 + (X2_2 - C0_2)^2}$

$= \sqrt{(2-1)^2 + (3-1)^2}$

$= 2{,}2361$

Lakukan pengolahan sampai dengan data ke 20, setelah data di olah seperti diatas maka di dapat nilai *cluster* yang baru, adapun nilai tersebut dapat dilihat pada Tabel 3 di bawah ini:

Tabel 3: Hasil Cluster Iterasi I

No	C0	C1
1	0,0000	2,2361
2	2,2361	0,0000
3	2,0000	2,2361
4	1,4142	1,0000
5	4,1231	3,1623
6	3,6056	1,4142
7	1,0000	2,0000
8	3,0000	2,8284
9	3,0000	2,8284
10	3,0000	2,8284
11	3,0000	2,8284
12	3,0000	2,8284
13	1,0000	2,0000
14	1,0000	2,0000
15	5,0000	4,4721
16	6,0000	5,3852
17	0,0000	2,2361
18	0,0000	2,2361
19	3,0000	2,8284
20	3,0000	2,8284

Langkah berikutnya adalah menentukan letak *cluster* baru seperti yang di lakukan pada awal tadi yaitu dengan cara membandingkan antara tiga *cluster* yaitu *cluster* CO, *cluster* C1, nilai minimum merupakan nilai yang menjadi nilai pilihan pada tahapan ini. Jika ditemukan nilai yang paling kecil (minimum) maka dapat dimasukkan kedalam *cluster* tersebut. Untuk lebih jelas di lihat Table 4 dibawah ini:

Tabel 4: Penentuan Letak Cluster Iterasi II

No	C0	C1	C0	C1
1	0,0000	2,2361	C0	

2	2,2361	0,0000		C1
3	2,0000	2,2361	C0	
4	1,4142	1,0000		C1
5	4,1231	3,1623		C1
6	3,6056	1,4142		C1
7	1,0000	2,0000	C0	
8	3,0000	2,8284		C1
9	3,0000	2,8284		C1
10	3,0000	2,8284		C1
11	3,0000	2,8284		C1
12	3,0000	2,8284		C1
13	1,0000	2,0000	C0	
14	1,0000	2,0000	C0	
15	5,0000	4,4721		C1
16	6,0000	5,3852		C1
17	0,0000	2,2361	C0	
18	0,0000	2,2361	C0	
19	3,0000	2,8284		C1
20	3,0000	2,8284		C1

Setelah tahapan di atas selesai di kerjakan, kemudian langkah selanjutnya yaitu menentukan nilai *centroid* yang baru, nilai ini ditentukan oleh data yang masuk kedalam *cluster*, berdasarkan table diatas diperoleh nilai sebagai berikut :

1. *Cluster* 0 terdapat 7 data.
2. *Cluster* 1 terdapat 13 data.

Setelah langkah di atas selesai di lakukan maka secara keseluruhan diperoleh nilai centroid atau nilai *cluster* yang baru seperti Tabel 5 di bawah ini:

Tabel 5: *Centroid* Baru

Centroid Baru		
C0	1,714	1,000
C1	4,077	1,538

Adapun tabel hasil pengunyian dapat di lihat pada Tabel 6 :

Tabel 6: Hasil Pengujian

No	Cluster	Daerah	Tingkat Keamanan		Jumlah	Total	Ket
			Rawan	Tidak Rawan			
1	C0	Sungai Lundang	1	1	35	1,75	rawan
2		Barung Barung Belantai	2	3			
3		Kapuh utara	3	1			
4		Talaok	2	2			
5		Sawah Laweh	3	4			
6		Kapeh Panji	2	1			
7		Sago Salido	2	1			
8		Painan	2	1			

9		Surantih	1	1			
10		Sago	1	1			
11	C1	Ampuan Lumpo	5	2	57	2,85	Tidak Rawan
12		Gurun Panjang Selatan	4	1			
13		Gurun Panjang	4	1			
14		Puluik Puluik	4	1			
15		Limau Gadang	4	1			
16		Salido	4	1			
17		Painan Selatan	6	1			
18		IV Koto Hilir	7	1			
19		Batang Kapas	4	1			
20		Cerocok Tarusan	4	1			

KESIMPULAN

Berdasarkan hasil penelitian yang dilakukan, dapat disimpulkan bahwa penerapan algoritma K-Means Clustering dalam pemetaan wilayah rawan tindak kriminal di wilayah hukum Pengadilan Negeri Painan mampu menghasilkan pengelompokan daerah berdasarkan tingkat kerawanan kejahatan. Melalui proses analisis menggunakan perangkat lunak RapidMiner, wilayah dapat dikelompokkan menjadi beberapa kategori, yaitu rawan, kurang rawan, dan aman.

Hasil klasterisasi menunjukkan bahwa algoritma K-Means efektif dalam mengidentifikasi pola dan distribusi tindak kriminal berdasarkan data perkara yang telah disidangkan. Informasi ini dapat dimanfaatkan oleh aparat penegak hukum dan pemerintah daerah untuk mendukung pengambilan keputusan strategis, seperti penempatan sumber daya keamanan, peningkatan patroli, serta perencanaan kebijakan pencegahan kejahatan yang lebih tepat sasaran.

Selain itu, penelitian ini juga membuktikan bahwa penerapan metode data mining, khususnya K-Means, dapat menjadi alat analisis yang andal dalam mendukung sistem informasi geografis (SIG) untuk pemetaan kriminalitas. Ke depan, penelitian serupa dapat dikembangkan lebih lanjut dengan menambahkan faktor-faktor sosial, ekonomi, dan demografis agar hasil pengelompokan menjadi lebih komprehensif dan representatif terhadap kondisi nyata di lapangan.

REFERENSI

[1] Kusbianti, S. F. S., Yaqin, M. I. A., & Agustin, A. D. (2025). *Analysis of Mapping the Distribution Zone of Crime Cases in Indonesia: A Case Study of 2018 to 2022.* AIP Conference Proceedings, 3316(1), 080007.

[2] Rahmawati, D., & Prasetyo, E. (2024). *Crime Data Clustering Using K-Means and DBSCAN for Predictive Policing.* Journal of Data Science and Applications, 12(2), 55–64.

[3] Lestari, I., & Hidayat, R. (2024). *Implementation of K-Means Clustering in Crime Mapping: Case Study of Urban Security Zones.* Indonesian Journal of Computer Science, 9(3), 201–210.

[4] Ahmed, T., & Chatterjee, S. (2023). *Geospatial Crime Hotspot Detection Using K-Means and Heatmap Visualization.* International Journal of Information Technology, 15(4), 142–151.

[5] Suryani, M., & Ramadhan, F. (2023). *Comparative Analysis of Clustering Algorithms in Crime Data Segmentation Using RapidMiner.* Jurnal Teknologi Informasi dan Komputer (JTIK), 10(1), 87–96.

[6] Mulyono, A., & Sari, N. A. (2024). *Optimization of K-Means Algorithm for Regional Crime Analysis Using Silhouette Coefficient.* Procedia Computer Science, 239, 120–129.

[7] Wu, H., Zhang, J., & Li, X. (2025). *Machine Learning for Urban Crime Prediction: An Empirical Comparison of K-Means and Neural Networks.* Applied Intelligence, 75(2), 452–468.

[8] Yuliana, R., & Hapsari, D. (2023). *Penerapan Metode K-Means dalam Pemetaan Daerah Rawan Kriminal di Kota Bandung.* Jurnal Informatika dan Komputasi, 8(2), 135–143.

[9] Farhan, M., & Nugroho, T. (2025). *Application of K-Means and GIS for Spatial Crime Risk Clustering.* International Journal of Geographical Information Systems, 14(1), 33–42.

[10] Zhang, L., & Kim, J. (2024). *Spatial Clustering of Crime Data Using Enhanced K-Means with Adaptive Centroid Initialization.* IEEE Access, 12, 14285–14297.