

CLUSTERING DAERAH RAWAN KRIMINALITAS MENGUNAKAN ALGORITMA FUZZY C-MEANS

¹Jauharotul Inayah, ²Diva Ayu Safitri Nur Maghfiroh, ³Dian C. Rini Novitasari
^{1,2,3}Program Studi Matematika, Universitas Islam Negeri Sunan Ampel Surabaya
Jl. Ahmad Yani 117 Surabaya
¹jauharotul.inayah0@gmail.com, ³diancrini@uinsby.ac.id

Abstrak

Kriminalitas merupakan kejahatan atau tindakan ilegal yang dapat di hukum. Kejahatan seperti pembunuhan, penyerangan, pemerkosaan dan lain-lain, tumbuh secara luas saat ini. Indonesia berada dalam peringkat ke-70 sebagai salah satu negara di Dunia dengan banyaknya kriminalitas berdasarkan pada Numbeo 2022. Pengklasteran daerah rawan kejahatan ini dilakukan di Indonesia agar masyarakat dan pihak keamanan dapat lebih waspada terhadap kriminalitas, sehingga dapat menurunkan angka kriminalitas di Indonesia. Metode yang digunakan pada pengklasteran ini adalah metode Fuzzy C-Means dan menghasilkan 3 cluster, yakni daerah rawan kriminalitas tingkat tinggi, sedang, dan rendah pada tahun 2018 hingga 2021 dengan diperoleh nilai uji silhouette coefficient rata-rata sebesar 0,8322.

Kata kunci: *Clustering, Fuzzy C-Means, Kriminalitas, Silhouette.*

Abstract

A crime is a illegal act that can be punished. Crimes like murder, assault, rape etc, are growing widely nowadays. Indonesia is ranked 70th as the country with the most crime in the world based on the 2022 Numbeo. Clustering crime-prone areas will be implemented in Indonesia, so that the public and security forces can be more alert to crime, so as to reduce the crime rate in Indonesia. The method used in this clustering is the Fuzzy C-Means method and generate 3 groups or clusters, these are areas prone to high, medium, and low crime rates in 2018 to 2021 with the silhouette coefficient value obtained an average of 0,8322.

Keywords: *Clustering, Crime, Fuzzy C-Means, Silhouette.*

PENDAHULUAN

Kriminalitas merupakan kejahatan atau tindakan ilegal yang dapat di hukum. Kejahatan seperti pembunuhan, penyerangan, pemerkosaan dan lain-lain, tumbuh secara luas saat ini [1]. Setiap manusia pasti memiliki kebutuhan fundamental yaitu kebutuhan biologis seperti tidur, makan, serta minum, dan kebutuhan sosial seperti peranan sosial, rasa aman. Maka semakin tumbuh angka

kriminalitas, semakin terancam rasa aman pada diri manusia [2]. Indonesia berada dalam peringkat ke-70 sebagai salah satu negara di Dunia dengan banyaknya kriminalitas berdasarkan pada Numbeo 2022, sedangkan pada benua Asia, Indonesia menduduki peringkat ke-15. Dan menempati peringkat 5 diantara Negara di Asia-Tenggara [3]. Kejahatan tidak dapat diprediksi, namun kejahatan tidak benar-benar acak. Terdapat pola kejahatan yang dapat dianalisis.

Berdasarkan teori, kejahatan terjadi ketika ruang aktivitas korban atau sasaran bersinggungan dengan ruang aktivitas pelaku. Pola kejahatan berperan penting dalam memahami mengapa kejahatan dilakukan di daerah tertentu [4]. Tradisi klasik yang telah lama ditinggalkan dalam kriminologi didasarkan pada pilihan dan kehendak bebas, dan melihat kejahatan sebagai konsekuensi alami dari kecenderungan manusia yang tidak terkendali untuk mencari kesenangan dan menghindari rasa sakit [5][5]. Kejahatan masih menjadi ancaman terbesar bagi rasa aman manusia untuk berkehidupan dalam perubahan-perubahan yang terjadi pada nilai-nilai kemasyarakatan. Rasa aman juga merupakan salah satu hak asasi manusia, sehingga baik pemerintah maupun masyarakat itu sendiri harus saling berkerjasama untuk mengurangi jumlah kejahatan di Indonesia [6]. Penelitian ini dilakukan pengklasteran terhadap daerah rawan kriminalitas di Indonesia, menggunakan metode *fuzzy c-means*.

Pengklasteran atau clustering merupakan pengelompokan yang bertujuan untuk mengelompokkan data berdasarkan persamaannya. Fuzzy c-means merupakan teknik pengklasteran yang mana keberadaan tiap-tiap titik data dalam suatu kluster ditentukan oleh derajat keanggotaan [7]. Metode ini merupakan salah satu metode terbaik dalam melakukan clustering. Dibuktikan dengan penelitian sebelumnya oleh Aditya [8] dengan melakukan

Perbandingan K-Means dan Fuzzy C-Means untuk Pengelompokan Data User Knowledge Modeling. Hasil dari penelitian ini yaitu metode Fuzzy C-Means lebih baik dari K-Means dengan nilai validasi PCI lebih mendekati 1.

Penelitian dengan metode yang sama juga dilakukan oleh Premasundari [1] yang diperoleh hasil daerah rawan pembunuhan dan penyerangan, pemerkosaan tingkat tinggi masing - masing *Georgia, North Carolina dan Nevada*. Penelitian yang dilakukan oleh Debby Armetiyana [9] menggunakan *Fuzzy C-Means*. Hasil dari penelitian ini adalah provinsi NTT dan Papua merupakan kluster terendah indeks pembangunan manusia di Indonesia. Penelitian dengan metode yang sama juga dilakukan oleh Bhavna Saini yang menyatakan bahwa model dari metode *fuzzy c-means* dianggap optimal dalam memperoleh visualisasi data kejahatan negara bagian di India [10]. Penelitian juga dilakukan oleh Jemaictry Tamaela [11] menggunakan algoritma *fuzzy c-means* dan *k-means* untuk klusterisasi dan pemetaan lahan pertanian di Minahasa Tenggara, yang diperoleh hasil bahwa algoritma FCM dan KM dapat mengklasterkan daerah lahan pertanian sesuai komoditi berdasarkan pada jenis atribut yang digunakan. Lebih lanjut penelitian yang dilakukan oleh Salar Askari [12] dengan hasil bahwa metode *Revisi Fuzzy C-Means* (RFCM) salah satu metode terbaik untuk pengklasteran data yang bising dan data dengan *cluster* yang tidak sama. Hal tersebut membuktikan bahwa

metode *fuzzy c-means* dapat diterapkan pada penelitian ini yang memiliki tujuan untuk mengklasterkan daerah rawan kriminalitas di Indonesia.

Oleh karena itu, pada penelitian ini akan diuji lebih lanjut menggunakan 3 sampel data kejahatan yang terjadi di Indonesia dalam rentang waktu 2018-2021.

METODE PENELITIAN

Penelitian ini dilakukan dengan menggunakan metode *Fuzzy C-Means*. Metode FCM atau *Fuzzy C-Means* merupakan teknik pengelompokan data dimana setiap data dalam suatu *cluster* ditentukan oleh nilai keanggotaan. Data yang diambil sebagai objek penelitian ini adalah data kriminalitas di yang

didapatkan dari Badan Pusat Statistik (BPS). Penelitian ini menggunakan data kriminalitas terhadap nyawa yaitu kejahatan yang menyebabkan hilangnya nyawa pada seseorang seperti pembunuhan, kriminalitas terhadap fisik yaitu kejahatan yang dilakukan dengan penyerangan terhadap fisik seseorang seperti kekerasan, penculikan, dan kriminalitas terhadap asusila atau kejahatan dengan penyimpangan terhadap norma-norma atau kaidah kesopanan seperti pemerkosaan, yang terjadi pada tiap provinsi di Indonesia pada tahun 2018-2021 [13] yang dapat dilihat sampel datanya pada Tabel 1 dengan data kejahatan terhadap nyawa, sampel data kejahatan terhadap fisik yang tertera pada Tabel 2, serta sampel data terhadap asusila pada Tabel 3.

Tabel 1. Sampel Data Kejahatan terhadap Nyawa

Provinsi	Kejahatan terhadap nyawa			
	2018	2019	2020	2021
Aceh	22	26	15	124
Sumatera Utara	107	104	99	283
Sumatera Barat	14	16	20	19
...
Maluku Utara	23	3	1	23
Papua	43	52	27	70
Papua Barat	13	10	10	42

Tabel 2. Sampel Data Kejahatan terhadap Fisik

Provinsi	Kejahatan terhadap fisik			
	2018	2019	2020	2021
Aceh	1475	1256	1206	1984
Sumatera Utara	5240	4817	6207	7803
Sumatera Barat	2153	1988	1724	1470
...
Maluku Utara	210	174	255	611
Papua	1370	1413	1296	1074
Papua Barat	704	669	655	1150

Tabel 3. Sampel Data Kejahatan terhadap Asusila

Provinsi	Kejahatan terhadap asusila			
	2018	2019	2020	2021
Aceh	262	267	232	137
Sumatera Utara	321	216	774	159
Sumatera Barat	392	333	294	43
...
Maluku Utara	98	114	100	33
Papua	103	119	100	106
Papua Barat	39	33	47	64

Clustering atau klusterisasi merupakan metode dengan melakukan pengelompokan data dalam beberapa *cluster* sesuai dengan parameter yang sama [14]. Secara umum metode *clustering* dibagi menjadi dua yakni *Hierarchical Clustering* dan *Partitional Clustering*. dalam metode *Partitional Clustering* terdapat metode *Distance-Based*, *Model-Based* dan *Density-Based* [15].

Jenis *clustering* dapat dibedakan menjadi *Hard Clustering* dan *Soft Clustering* [16]. Perbedaan diantara keduanya ialah dalam *Hard Clustering* (Non Fuzzy Clustering) data dibagi dalam beberapa kelompok dan setiap data hanya bisa terdapat dalam 1 *cluster*. Sedangkan dalam *Soft Clustering* (*Fuzzy Clustering*) setiap data boleh terdapat pada lebih dari 1 *cluster* [17].

FCM merupakan perluasan dari metode *hard c-means* sehingga termasuk kedalam salah satu jenis *Soft Clustering* dimana setiap data boleh terdapat pada lebih dari 1 *cluster* dalam pengelompokannya [18]. Mengetahui kualitas dan kekuatan suatu *cluster* dapat dilakukan uji *silhouette* yaitu dengan cara membandingkan jarak data pada *cluster* yang sama dengan jarak data pada *cluster* lainnya

[19] seperti pada persamaan 1, dimana $S(i)$ adalah *silhouette coefficient* dengan a_i adalah rata-rata jarak dari objek i dengan seluruh objek yang berada pada *cluster* yang sama dan a_i nilai terkecil dari rata-rata objek i dengan objek lain pada *cluster* yang berbeda [20].

$$S(i) = \frac{b(i)-a(i)}{\max\{a(i),b(i)\}} \quad (1)$$

Langkah-langkah algoritma peng-klasteran *Fuzzy C-Means* dapat dilihat pada Gambar 1 dengan langkah yang pertama adalah menginputkan data dan diakhiri dengan pengecekan kondisi berhenti pada pembaruan matriks partisi U, dengan perincian sebagai berikut :

1. Input data berupa matriks $n \times m$ pada *cluster* x seperti pada persamaan 2, dimana n adalah sampel data dan m adalah atribut setiap data.

$$x = \begin{bmatrix} x_{1,1} & \dots & x_{1,m} \\ \vdots & \ddots & \vdots \\ x_{n,1} & \dots & x_{n,m} \end{bmatrix} \quad (2)$$

2. Menentukan parameter :
 - a. Jumlah dari Cluster (c)
 - b. Bobot pangkat (w)
 - c. Iterasi maksimum (*MaxIter*)
 - d. Kemungkinan error terkecil (ϵ)

- e. Fungsi tujuan awal (P_0)
- f. Iterasi awal ($t = 1$)
- Menentukan bilangan acak $\mu_{i,k}$ sebagai elemen matriks partisi U yang dapat dilihat pada persamaan 3 dan 4 untuk $i = 1,2,3, \dots, n$ dan $k = 1,2,3, \dots, c$. Partisi awal $\mu_{i,k}$ adalah derajat keanggotaan yang mengacu pada seberapa besar kemungkinan data tersebut termasuk dalam sebuah *cluster* dengan Q_i adalah jumlah nilai derajat keanggotaan perkolom.

$$Q_j = \sum_{k=1}^c \mu_{ik} \quad (3)$$

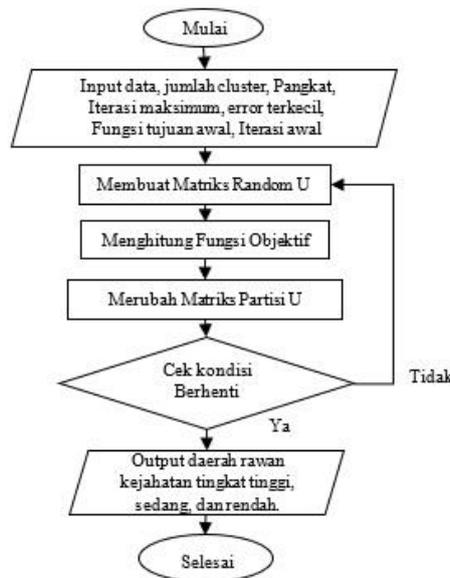
$$\mu_{i,k} = \frac{\mu_{i,k}}{Q_i} \quad (4)$$
 - Menghitung pusat cluster ke- k untuk $k = 1,2,3, \dots, c$ seperti yang terlihat pada persamaan 5, dimana V merupakan pusat *cluster* dan X_i adalah parameter ke- i .

$$V_{kj} = \frac{\sum_{i=1}^n ((\mu_{ik})^w * x_{ij})}{\sum_{i=1}^n (\mu_{ik})^w} \quad (5)$$
 - Menghitung fungsi objektif pada iterasi ke- t dengan nilai fungsi objektif dinotasikan sebagai P_t seperti pada persamaan 6.

$$P_t = \sum_{i=1}^n \sum_{k=1}^c ([\sum_{j=1}^m (x_{ik} - v_{kj})^2]) (\mu_{ik})^w \quad (6)$$
 - Menghitung pembaruan matriks partisi dengan rumusan pada persamaan 7 untuk $i = 1,2,3, \dots, n$ dan $k = 1,2,3, \dots, c$.

$$\mu_{ik} = \frac{[\sum_{j=1}^m (x_{ik} - v_{kj})^2]^{-\frac{1}{w-1}}}{\sum_{k=1}^c [\sum_{j=1}^m (x_{ik} - v_{kj})^2]^{-\frac{1}{w-1}}} \quad (7)$$
 - Cek kondisi berhenti pada pembaruan matriks. Apabila ditemukan kondisi seperti pada persamaan 8, maka pembaruan matriks berhenti. Jika ditemukan kondisi ($t = t+1$), maka tetap dilakukan pembaruan matriks atau perulangan pada langkah ke-4.

$$(|P_t - P_{t-1}|) < \xi \text{ atau } t > \text{Max Iter} \quad (8)$$



Gambar 1. Algoritma Fuzzy C-Means

Data yang diinputkan berupa matriks berukuran 34×3 dengan parameter yang sudah ditentukan, kemudian membuat matriks partisi U dengan jumlah derajat keanggotaan bernilai 1 sesuai persamaan (3). Menghitung fungsi objektif menggunakan persamaan (6) dan melakukan pembaruan sesuai persamaan (7) sehingga memenuhi minimal eror yang telah ditentukan seperti pada persamaan (8). Langkah selanjutnya melihat derajat keanggotaan matriks U yang telah diperbarui.

HASIL DAN PEMBAHASAN

Penelitian ini dilakukan untuk mengklasterkan daerah provinsi berdasarkan tingkat rawan kriminalitas yang berada di negara Indonesia dari tahun 2018 hingga 2021, dengan metode yang digunakan adalah metode *fuzzy c-means*. Nilai dari *cluster* jika di *cluster* ulang maka hasilnya tidak akan sama dikarenakan matriks partisi U yang dibuat secara random pada awal pengklasteran, akan tetapi perubahan tersebut tidak mempengaruhi

keanggotaan *cluster* karena perubahannya tidak terlalu signifikan.

Langkah pertama yang harus dilakukan dalam pengklasteran data menggunakan *fuzzy c-means* adalah menentukan komponen-komponen yang diperlukan. Penelitian ini menggunakan komponen-komponen dengan kelompok *cluster* sebanyak 3, bobot pangkat samadengan 2, iterasi maksimum sebanyak 100 iterasi, kemungkinan error terkecil senilai 1×10^{-5} dengan fungsi tujuan awal adalah 0 dan iterasi awalnya adalah 1. Langkah selanjutnya dilakukan pengklasteran daerah rawan kriminalitas pada setiap tahunnya, yaitu tahun 2018, 2019, 2020 dan 2021. Untuk menentukan *cluster* yang optimal, dilakukan proses uji validasi dalam menentukan keoptimalan suatu *cluster*. Pada penelitian ini, metode yang digunakan adalah metode *silhouette coefficient*. Masing-masing tahun di uji dan dilihat nilai indeksnya, nilai *silhouette coefficient* berada pada rentan (-1) hingga (1). Jika nilai indeks yang didapatkan mendekati 1, maka kualitas *cluster* semakin baik.

Tabel 4. Tabel Kaufman

Type	<i>Silhouette Coefficient</i>	<i>Proposed Interpretation</i>
1	0,71 – 1,00	<i>A strong structure has been found</i>
2	0,51 – 0,70	<i>A reasonable structure has been found</i>
3	0,26 – 0,50	<i>The structure is weak and could be artificial</i>
4	$\leq 0,25$	<i>No substantial structure has been found</i>

Keterangan : *Type* 1 merupakan kualitas *cluster* yang paling baik, hingga *type* 4 merupakan kualitas *cluster* yang kurang baik.

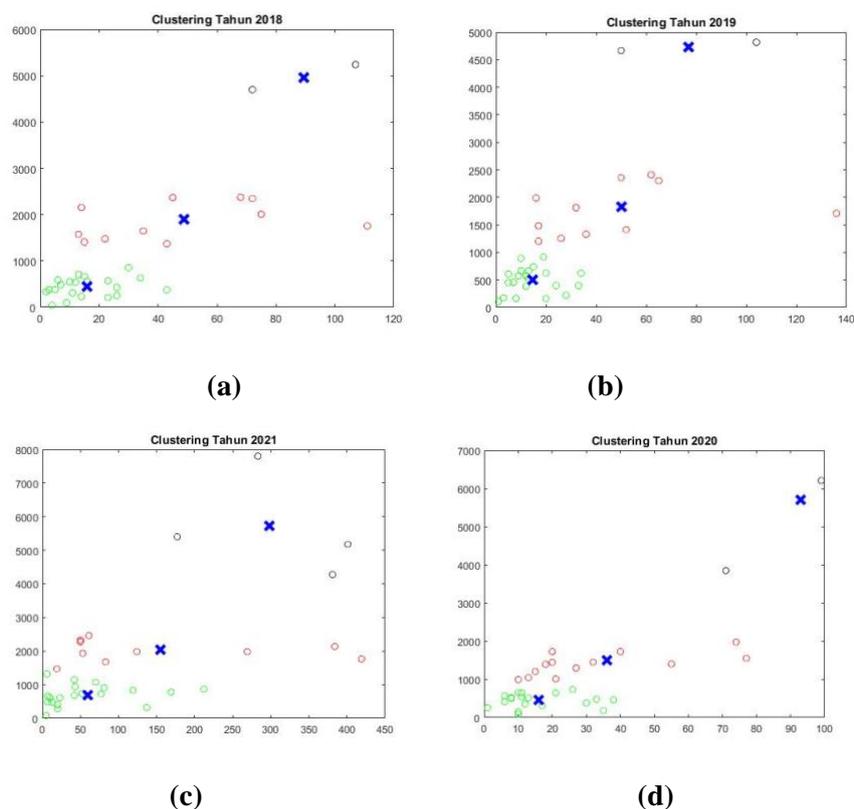
Tabel 5. Validasi *Silhouette* Data Tahun 2018, 2019, 2020 dan 2021

Jumlah <i>cluster</i>	<i>Silhouette Coefficient</i>				Rata-rata
	2018	2019	2020	2021	
2	0,7069	0,7769	0,9387	0,9058	0,8320
3	0,9058	0,8385	0,7693	0,8152	0,8322
4	0,8903	0,7861	0,8047	0,6621	0,7858

Keterangan : validasi nilai *silhouette coefficient* berdasarkan jumlah *cluster* pada tiap tahunnya. 3 *Cluster* merupakan jumlah *cluster* yang paling optimal.

Nilai rata-rata pada Tabel 5, menunjukkan bahwa jumlah *cluster* paling optimal menggunakan *silhouette coefficient* adalah 3 *cluster* dengan nilai rata-rata *silhouette coefficient* sebesar 0,8322. Berdasarkan nilai rata-rata *silhouette coefficient* pada Tabel 4, nilai *silhouette coefficient* dengan rata-rata 0,8322 masuk dalam kelompok tipe 1 yang merupakan *cluster* dengan struktur terkuat. Hasil *clustering* dapat dilihat pada Gambar 2, dimana simbol **O** berwarna hitam merupakan visualisasi dari *cluster* 1 sedangkan *cluster* 2 divisualisasikan

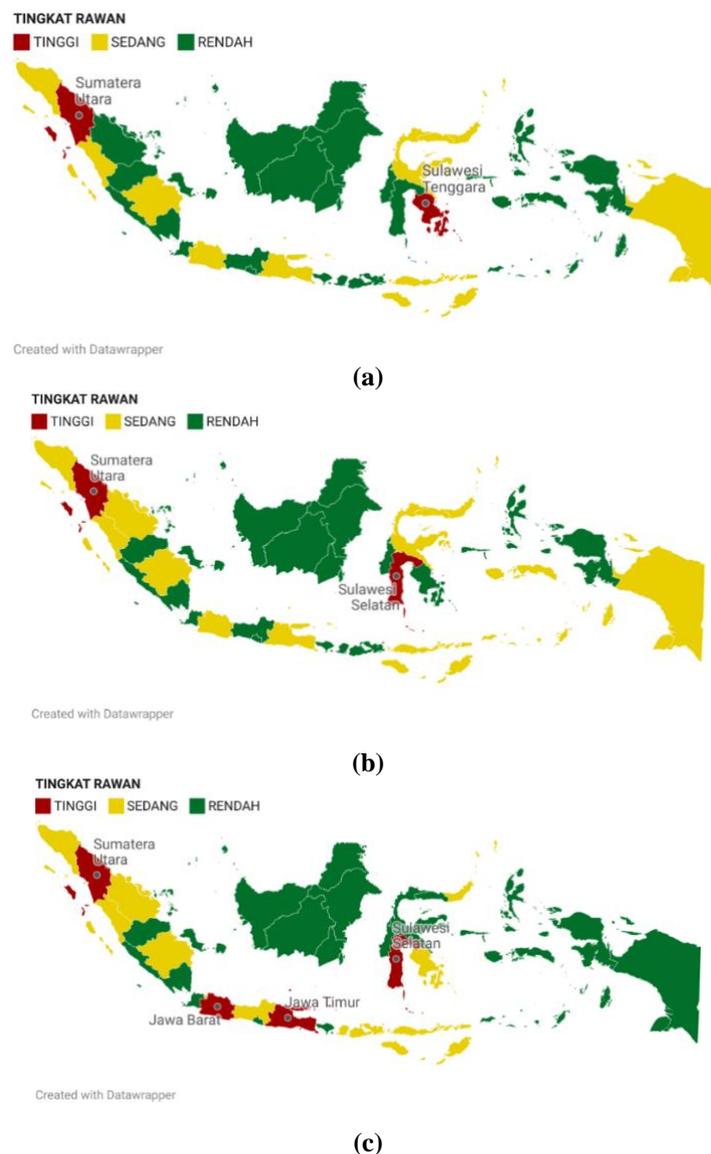
dengan simbol **O** berwarna merah, dan simbol **O** berwarna hijau merupakan kelompok *cluster* 3, serta terdapat simbol **X** berwarna biru yang menunjukkan letak pusat *cluster* pada setiap *cluster*. *Cluster* yang terletak pada barisan paling atas pada Gambar 2 merupakan pusat *cluster* terbesar, yakni merupakan daerah yang memiliki tingkat kerawanan terhadap kejahatan paling tinggi, sedangkan untuk *cluster* terletak pada barisan paling bawah merupakan *cluster* terkecil, yakni daerah yang memiliki tingkat kerawanan kejahatan paling rendah.



Gambar 2. Hasil Clustering (a) Hasil Clustering Tahun 2018, (b) Hasil Clustering Tahun 2019, (c) Hasil Clustering Tahun 2020, (d) Hasil Clustering Tahun 2021

Dari proses *clustering* ini dapat dilihat bahwa *cluster* 1 berada paling atas, sebaliknya *cluster* 3 berada paling bawah. Sehingga dapat disimpulkan bahwa *cluster* 1 merupakan daerah dengan rawan tingkat tinggi, *cluster* 2 merupakan daerah dengan tingkat kerawanan

sedang, sedangkan *cluster* 3 merupakan daerah dengan tingkat kerawanan rendah. Hasil *clustering* data pada tiap-tiap *cluster* dapat dilihat pada tabel hasil *clustering* tahun 2018-2021



Gambar 3. Peta Daerah Rawan Kriminalitas di Indonesia (a) Peta Daerah Rawan Kriminalitas di Indonesia Tahun 2018 Dan 2019 (b) Peta Daerah Rawan Kriminalitas di Indonesia Tahun 2020 (c) Peta Daerah Rawan Kriminalitas di Indonesia Tahun 2021.

Tabel 4. Hasil *Clustering* Tahun 2018 - 2021

Hasil Cluster 2018 dan 2019			Hasil Cluster 2020			Hasil Cluster 2021					
Cluster	Data Ke-	Provinsi	Cluster	Data Ke-	Provinsi	Cluster	Data Ke-	Provinsi			
1 [Tinggi]	2	Sumatera Utara	1 [Tinggi]	2	Sumatera Utara	1 [Tinggi]	2	Sumatera Utara			
	27	Sulawesi Tenggara		28	Sulawesi Selatan		12	Jawa Barat			
	1	Aceh		1	Aceh		15	Jawa Timur			
	3	Sumatera Barat		3	Sumatera Barat		28	Sulawesi Selatan			
	7	Sumatera Selatan		4	Riau		1	Aceh			
	11	DKI Jakarta		7	Sumatera Selatan		3	Sumatera Barat			
	12	Jawa Barat		11	DKI Jakarta		4	Riau			
	15	Jawa Timur		12	Jawa Barat		7	Sumatera Selatan			
	2 [Sedang]	19		Nusa Tenggara Timur	2 [Sedang]		15	Jawa Timur	2 [Sedang]	11	Dki Jakarta
		25		Sulawesi Utara			19	Nusa Tenggara Timur		13	Jawa Tengah
26		Sulawesi Tengah	25	Sulawesi Utara		18	Nusa Tenggara Barat				
30		Gorontalo	26	Sulawesi Tengah		19	Nusa Tenggara Timur				
33		Papua	30	Gorontalo		25	Sulawesi Utara				
4		Riau	31	Maluku		27	Sulawesi Tenggara				
5		Kep. Riau	33	Papua		5	Kep. Riau				
6		Jambi	5	Kep. Riau		6	Jambi				
8		Kep. Bangka Belitung	6	Jambi		8	Kep. Bangka Belitung				
9		Bengkulu	8	Kep. Bangka Belitung		9	Bengkulu				
3 [Rendah]	10	Lampung	3 [Rendah]	9	Bengkulu	3 [Rendah]	10	Lampung			
	13	Jawa Tengah		10	Lampung		14	Di Yogyakarta			
	14	DI Yogyakarta		13	Jawa Tengah		16	Banten			
	16	Banten		14	DI Yogyakarta		17	Bali			
	17	Bali		16	Banten		20	Kalimantan Barat			
	18	Nusa Tenggara Barat		17	Bali		21	Kalimantan Tengah			
	20	Kalimantan Barat		18	Nusa Tenggara Barat		22	Kalimantan Selatan			
	21	Kalimantan Tengah		20	Kalimantan Barat		23	Kalimantan Timur			
	22	Kalimantan Selatan		21	Kalimantan Tengah		24	Kalimantan Utara			
	23	Kalimantan Timur		22	Kalimantan Selatan		26	Sulawesi Tengah			
24	Kalimantan Utara	23	Kalimantan Timur	29	Sulawesi Barat						
28	Sulawesi Selatan	24	Kalimantan Utara	30	Gorontalo						
29	Sulawesi Barat	27	Sulawesi Tenggara	31	Maluku						
31	Maluku	29	Sulawesi Barat	32	Maluku Utara						
32	Maluku Utara	32	Maluku Utara	33	Papua						
34	Papua Barat	34	Papua Barat	34	Papua Barat						

Hasil penelitian ini dapat dilihat pada Tabel 4 yang menunjukkan bahwa pada tahun 2018 terhadap 2019 tidak mengalami perubahan posisi *cluster*, *cluster* dengan daerah rawan kriminalitas tertinggi berada pada provinsi Sumatera Utara dan Sulawesi Tenggara. Pada tahun 2020 dan 2021 mengalami perubahan posisi *cluster*, yakni *cluster* dengan tingkat kriminalitas tertinggi tahun 2020 berada pada provinsi Sumatera utara dan Sulawesi Selatan, sedangkan pada tahun 2021, *cluster* dengan tingkat kriminalitas tertinggi berada pada provinsi Sumatera Utara, Jawa barat, Jawa Timut, dan Sulawesi Selatan. Dapat dilihat juga pada Gambar 3, warna merah pada peta menunjukkan daerah rawan kriminalitas tingkat tinggi, warna kuning menunjukkan daerah rawan kriminalitas tingkat sedang, dan warna hijau menunjukkan daerah rawan kriminalitas tingkat rendah.

ESIMPULAN DAN SARAN

Berdasarkan hasil dan analisis algoritma *Fuzzy C-Means* untuk mengklaster daerah rawan kriminalitas dengan jumlah *cluster* sebanyak 3. Diperoleh daerah rawan kriminalitas tingkat tinggi pada tahun 2018 dan 2019 adalah Sumatera Utara dan Sulawesi Tenggara dengan nilai *silhouette coefficient* masing-masing tahun sebesar 0,9058 dan 0,8385. Pada tahun 2020, daerah rawan kriminalitas tingkat tinggi adalah Sumatera

Utara dan Sulawesi Selatan dengan nilai *silhouette coefficient* sebesar 0,7693. Pada tahun 2021, jumlah provinsi dalam *cluster* tinggi mengalami kenaikan menjadi 4 provinsi yakni provinsi Sumatera Utara, Jawa Barat, Jawa Timur dan Sulawesi Selatan dengan nilai *silhouette coefficient* sebesar 0,8152. Pada dasarnya, penelitian ini berjalan baik. Namun pada penelitian selanjutnya, hendaknya disempurnakan lagi untuk meningkatkan efektifitas dengan menambahkan variabel jumlah penduduk.

DAFTAR PUSTAKA

- [1] M. Premasundari and C. Yamini, "A violent crime analysis using fuzzy C-Means Clustering Approach," *Ictact J. Soft Comput.*, vol. 6956, no. April, p. 3, 2019, doi: 10.21917/ijsc.2019.0270.
- [2] Badan Pusat Statistik, *Statistik Kriminal 2020*. Badan Pusat Statistik, 2020.
- [3] Numbeo, "Crime Index by Country 2022," 2022. .
- [4] X. Zhao and J. Tang, "Crime in Urban Areas: A Data Mining Perspective," *ACM SIGKDD Explor. Newsl.*, vol. 20, no. 1, pp. 1–12, 2018, doi: 10.1145/3229329.3229331.
- [5] M. R. Gottfredson and T. Hirschi, *A General Theory of Crime*. 2022.
- [6] A. D. Putra, G. S. Martha, M. Fikram, and R. J. Yuhan, "Faktor-Faktor yang Memengaruhi Tingkat Kriminalitas di

- Indonesia Tahun 2018,” *Indones. J. Appl. Stat.*, vol. 3, no. 2, p. 123, 2021, doi: 10.13057/ijas.v3i2.41917.
- [7] R. Rustiyan and M. Mustakim, “Penerapan Algoritma Fuzzy C Means untuk Analisis Permasalahan Simpanan Wajib Anggota Koperasi,” *J. Teknol. Inf. dan Ilmu Komput.*, vol. 5, no. 2, p. 171, 2018, doi: 10.25126/jtiik.201852605.
- [8] A. Ramadhan, Z. Efendi, and Mustakim, “Perbandingan K-Means dan Fuzzy C-Means untuk Pengelompokan Data User Knowledge Modeling,” *Semin. Nas. Teknol. Informasi, Komun. dan Ind.* 9, pp. 219–226, 2017.
- [9] D. ARMETIYANA MARGARETTA, I. RAHMI HG, and H. YOZZA, “Pengklasteran Provinsi-Provinsi Di Indonesia Berdasarkan Indikator Yang Mempengaruhi Indeks Pembangunan Manusia Menggunakan Metode Fuzzy C-Means,” *J. Mat. UNAND*, vol. 10, no. 1, p. 79, 2021, doi: 10.25077/jmu.10.1.79-86.2021.
- [10] B. Saini, D. K. Saini, S. Srivastava, and M. Aggarwal, “A Study of Lightweight Approach to Analyze Crime Conditions in India,” *J. Appl. Secur. Res.*, pp. 1–25, 2021, doi: <https://doi.org/10.1080/19361610.2021.2006031>.
- [11] J. Tamaela, E. Sedyono, and A. Setiawan, “Cluster Analysis Menggunakan Algoritma Fuzzy C-means dan K-means Untuk Klasterisasi dan Pemetaan Lahan Pertanian di Minahasa Tenggara,” *J. Buana Inform.*, vol. 8, no. 3, pp. 151–160, 2017, doi: 10.24002/jbi.v8i3.1317.
- [12] S. Askari, “Fuzzy C-Means clustering algorithm for data with unequal cluster sizes and contaminated with noise and outliers: Review and development,” *Expert Syst. Appl.*, vol. 165, p. 113856, Mar. 2021, doi: 10.1016/j.eswa.2020.113856.
- [13] Badan Pusat Statistik, *Statistik Kriminal 2021*. Badan Pusat Statistik, 2021.
- [14] D. Abdullah, S. Susilo, A. S. Ahmar, R. Rusli, and R. Hidayat, “The application of K-means clustering for province clustering in Indonesia of the risk of the COVID-19 pandemic based on COVID-19 data,” *Qual. Quant.*, vol. 56, no. 3, pp. 1283–1291, 2022.
- [15] S. Askari, N. Montazerin, and M. H. Fazel Zarandi, “Generalized Possibilistic Fuzzy C-Means with novel cluster validity indices for clustering noisy data,” *Appl. Soft Comput.*, vol. 53, pp. 262–283, Apr. 2017, doi: 10.1016/j.asoc.2016.12.049.
- [16] S. N. Aisah, A. Nurcahyani, and D. C. Rini, “Implementasi Fuzzy C-Means Clustering (FCM) pada Pemetaan Daerah Potensi Transmigrasi di Jawa Timur,” *J. Tek. Inform. UNIKA St. Thomas*, pp. 33–40, 2022.
- [17] P. Wang, Q. Liu, G. Xu, and K. Wang,

- “A Three-Way Clustering Method Based on Ensemble Strategy and Three-Way Decision,” *Information*, vol. 10, no. 2, p. 59, Feb. 2019, doi: 10.3390/info10020059.
- [18] A. K. Dubey, U. Gupta, and S. Jain, “Comparative Study of K-means and Fuzzy C-means Algorithms on The Breast Cancer Data,” *Int. J. Adv. Sci. Eng. Inf. Technol.*, vol. 8, no. 1, p. 18, Feb. 2018, doi: 10.18517/ijaseit.8.1.3490.
- [19] A.-M. Shoolihah, M. T. Furqon, and A. W. Widodo, “Implementasi Metode Improved K-Means untuk Mengelompokkan Titik Panas Bumi,” *J. Pengemb. Teknol. Inf. dan Ilmu Komput. Univ. Brawijaya*, vol. 1, no. 11, pp. 1270–1276, 2017.
- [20] F. Novianti, Y. R. A. Yasmin, and D. C. R. Novitasari, “Penerapan Algoritma Fuzzy C-Means (FCM) dalam Pengelompokan Provinsi di Indonesia berdasarkan Indikator Penyakit Menular Manusia,” *JUMANJI (Jurnal Masy. Inform. Unjani)*, vol. 6, no. 1, pp. 23–33, 2022.