

Analisis Kasus Kriminalitas Dengan Menggunakan Metode Algoritma K-Means Pada Provinsi Sumatera Utara

Analysis of Criminal Cases Using the K-Means Algorithm Method in North Sumatra Province

**Deza Geraldin Salsabilah Saragih¹, Indah Dea Anastasia², Laily Najwa Rahmani³,
M. Safii^{4*}**

^{a,b,c,d}AMIK Tunas Bangsa, Jalan Serdang, Pematang Siantar, 21112, Indonesia

*m.safii@amiktunasbangsa.ac.id⁴

ABSTRAK

Kriminalitas adalah perilaku yang melanggar hukum dan menimbulkan kerugian bagi individu atau masyarakat luas. Variabel penelitian yang digunakan adalah jenis-jenis kriminalitas, dan obyek penelitian adalah Kepolisian Provinsi Sumatera Utara. Metode analisis kelompok yang digunakan dalam penelitian ini adalah K-Means Clustering. Dalam melakukan clustering, data yang telah diperoleh akan dihitung berdasarkan jumlah data kriminalitas pada tahun 2019 sampai 2021. Pada penelitian ini jumlah cluster yang akan digunakan adalah 3 cluster. Pada tahun 2019 memiliki cluster 1 = 21, tahun 2020 = 8 dan tahun 2021 = 8. Pada cluster 2 tahun 2019 memiliki cluster 1 = 19, tahun 2020 = 25 dan tahun 2021 = 28. Pada cluster 3 tahun 2019 = 24, tahun 2020 = 11 dan tahun 2021 = 15. Berdasarkan perhitungan yang telah dilakukan disimpulkan bahwa iterasi ke 6 memiliki jumlah cluster yang sama seperti iterasi ke 5 yaitu C1 19 data, C2 2 data dan C3 3 data. Berdasarkan hasil penelitian yang telah dilakukan dengan judul Analisis kasus kriminalitas dengan menggunakan metode Algoritma K-means pada provinsi sumatra utara, hasil penelitian memberikan wawasan yang berharga bagi pemerintah daerah, instansi terkait, dan pengambil keputusan. Pertama, identifikasi 3 kelompok utama kasus kriminalitas memungkinkan pengembangan program-program yang lebih terfokus. Untuk Kelompok C2 yang memiliki kasus kriminalitas yang sering terjadi diperlukan upaya lebih lanjut dalam menangani kasus kasus kriminalitas yang sering terjadi. Dan untuk kelompok C1 dan C3 dengan kasus kriminalitas yang jarang terjadi sebaiknya tetap diupayakan sebaik baiknya agar semua kasus kriminalitas semakin menurun.

Kata Kunci : Pengelompokan, Kriminalitas, Sumut, K-Means, Kejahatan

ABSTRACT

Crime is behavior that violates the law and causes harm to individuals or society at large. The research variables used are types of crime, and the research object is the North Sumatra Provincial Police. The group analysis method used in this research is K-Means Clustering. In conducting clustering, the data that has been obtained will be calculated based on the amount of crime data from 2019 to 2021. In research The number of clusters that will be used is 3 clusters. In 2019 it has cluster 1 = 21, in 2020 = 8 and in 2021 = 8. In cluster 2 in 2019 it has cluster 1 = 19, in 2020 = 25 and in 2021 = 28. In cluster 3 in 2019 = 24, in 2020 = 11 and 2021 = 15. so it was concluded that the 6th iteration has the same number of clusters as the 5th iteration, namely C1 19 data, C2 2 data and C3 3 data. Based on the results of research carried out entitled Analysis of crime cases using the K-means algorithm method in North Sumatra province, the research results provide valuable insight for local governments, related agencies and decision makers. First, the identification of 3 main groups of criminal cases allows the development of more focused programs. For Group C2 which has frequent criminal cases, further efforts are needed in handling frequently occurring criminal cases. And for groups C1 and C3 with rare crime cases, efforts should still be made as best as possible so that all crime cases decrease.

Keywords: Grouping, Crime, North Sumatra, K-Means, Crime

Info Artikel :

Disubmit:09 Oktober

Direview:01 November

Diterima :20 Desember

Copyright © 2023 - Journal UPU. All rights reserved.

1. PENDAHULUAN

Tindakan kejahatan merupakan tindakan yang bersifat negatif. Seringkali, tindakan ini merugikan banyak pihak dan pelaku tindakannya disebut sebagai seorang kriminal [1]. Kriminalitas adalah perilaku yang melanggar hukum dan menimbulkan kerugian bagi individu atau masyarakat luas. Ini mencakup berbagai tindakan berbeda yang dianggap ilegal, seperti penculikan, pembunuhan, pencurian ringan, pencurian dengan kekerasan, pencurian dengan pemberatan, penghinaan, pemerasan, penggelapan, penipuan, pengrusakan, penadahan, pencurian kendaraan bermotor, kejahatan narkoba, penyelundupan, korupsi, penyalahgunaan senjata api, sengketa tanah, dan ilegal loggin. Badan Pusat Statistik (BPS) menunjukkan bahwa tingkat kejahatan di Indonesia adalah 90 per 100.000 penduduk pada tahun pada tahun 2021 [2].

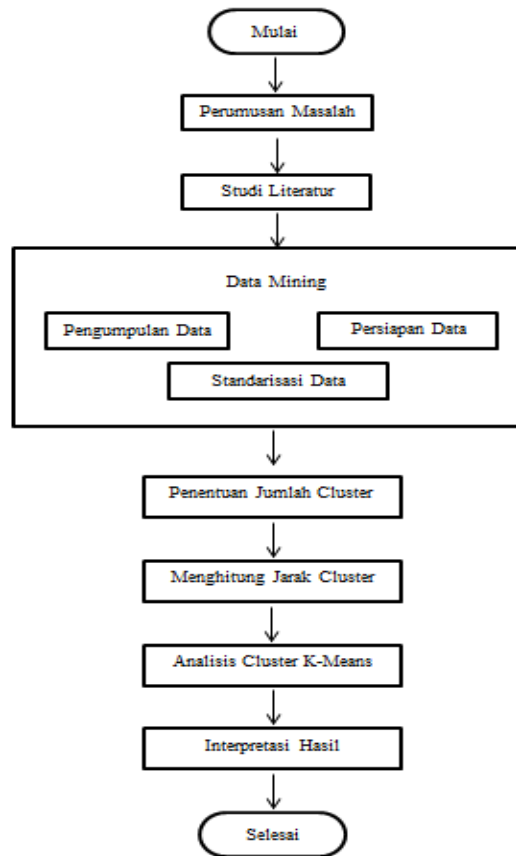
Dalam penelitian ini, kami akan melakukan pengelompokan jenis kriminalitas yang paling rawan di Sumatera Utara menggunakan metode analisis kelompok. Salah satu metode yang digunakan adalah K-Means Clustering, yang merupakan metode non-hierarki yang cocok untuk data yang besar. Metode ini juga memiliki proses pengelompokan yang lebih cepat daripada metode hierarki [3]. Variabel penelitian yang digunakan adalah jenis-jenis kriminalitas, dan obyek penelitian adalah Kepolisian Provinsi Sumatera Utara. Metode analisis kelompok yang digunakan dalam penelitian ini adalah K-Means Clustering [4].

Apalagi mengingat kondisi yang dimana kriminalitas semakin meningkat di setiap tahun khususnya yang berada pada daerah Sumatera Utara seperti pencurian kendaraan bermotor [1]. Pada penelitian ini data kriminal di Sumatera Utara akan dijadikan bahan untuk penerapan data mining. Dalam melakukan clustering, data yang telah diperoleh akan dihitung berdasarkan jumlah data kriminalitas pada tahun 2019 sampai 2021. Penulis menggunakan Data Mining untuk mengelolah data jenis kriminalitas dengan tujuan mengelompokkan data kriminalitas untuk menentukan cluster tingkat kriminalitas yang paling rawan [5].

1. METODE

Kerangka Kerja Penelitian

Dalam sebuah penelitian membutuhkan yang namanya kerangka kerja penelitian. Pada penelitian ini dilakukan sebuah proses pengumpulan data yang diperoleh dari website resmi Badan Pusat Statistika (BPS) dengan alamat <https://sumut.bps.go.id/indicator/34/493/1/banyaknya-peristiwa-kejahatan-pelanggaran-yang-dilaporkan-menurut-jenis-kejahatan-pelanggaran.html>. Data diunduh pada hari Senin, tanggal 06 November 2023. Langkah pertama dalam melakukan penelitian adalah menentukan rumusan masalah yang terjadi. Selanjutnya adalah melakukan studi pustaka atau pengumpulan referensi dari berbagai website resmi jurnal yang terdapat dari internet. Langkah selanjutnya adalah mempersiapkan data dan melakukan standarisasi data. Pada penelitian ini proses pengolahan data menggunakan algoritma K-Means untuk mengelompokkan data. Pada algoritma K-Means langkah awal yang perlu dilakukan adalah menentukan jumlah kluster yang akan digunakan, setelah itu menghitung jarak kluster dan menganalisis kluster yang telah didapatkan. Setelah mendapat iterasi 1 dalam algoritma K-Means dan melanjutkan sampai dengan iterasi 6 berdasarkan pengelompokkan kluster yang telah didapatkan dan melakukan langkah tersebut berulang kali sehingga mendapat persamaan dari kluster dan mendapatkan interpretasi hasil. Dengan adanya kerangka kerja, peneliti dapat secara sistematis mengarahkan penelitian dari rumusan masalah hingga analisis hasil dengan menggunakan algoritma K-Means dalam mengelompokkan data kriminalitas yang diperoleh dari Badan Pusat Statistik Sumut.



Gambar 1 Kerangka Kerja Penelitian

Data Mining

Data Mining sering digunakan sebagai proses analisis dan statistik data dan proses untuk memperoleh informasi berguna dari penyimpanan database besar, yang juga dapat dipahami sebagai ekstraksi informasi [6]. Berita terkini dari Big Data membantu pengambilan keputusan (penemuan pengetahuan) [7]. Dalam data mining diperlukan KDD (Knowledge Discovery in Database Process). Proses dalam KDD meliputi langkah pembersihan data, integrasi data, pemilihan data, transformasi data, evaluasi model, dan representasi pengetahuan. Penggunaan data mining melibatkan data relasional, data transaksional, multimedia, web, dan teks [8].

Algoritma K-Means

K-Means merupakan metode clustering yang paling sederhana dikarenakan dapat mengelompokkan data dalam jumlah yang cukup besar dengan waktu komputasi yang relatif cepat dan efisien [9]. Pengolahan KMeans adalah pengolahan berdasarkan perhitungan jarak terdekat antara data ke centroid [10]. Menurut [11] Data yang digunakan adalah data yang sama digunakan untuk pengolahan FCM, yaitu jumlah kejadian dan informasi kasus yang sudah dilakukan perkalian sesuai dengan bobotnya masing-masing, kemudian diolah berdasarkan langkah-langkah sebagai berikut :

1. Tentukan jumlah K cluster

2. Inisialisasikan K pusat cluster. Inisialisasi ini dilakukan dengan cara random. Namun pada penelitian ini, untuk menentukan centroid awal menggunakan software SPSS.
3. Alokasikan masing-masing data ke centroid/ratarata terdekat. Kedekatan suatu data ke cluster tertentu ditentukan jarak antara data dengan pusat cluster. Dalam tahap ini dihitung jarak tiap data ke tiap pusat cluster. Jarak paling dekat antara satu data dengan data satu cluster tertentu akan menentukan suatu data masuk dalam cluster mana. Untuk menghitung jarak digunakan rumus Euclidean Distance Space yaitu:

$$D_{12}(x_2, x_1) = \|x_2 - x_1\| = \sqrt{\sum_{j=1}^p (x_{2j} - x_{1j})^2}$$

Keterangan :

x_1 = Objek Data
 x_2 = Centroid
 P = Dimensi Data
 D = Jarak

4. Menentukan kembali pusat *cluster* yang baru dengan cara menghitung rata-rata dari keanggotaan *cluster* yang sekarang.
5. Pengalokasian objek ke *cluster* terdekat dengan menggunakan *centroid* yang baru. Perhitungan selesai jika tidak ada anggota yang berpindah *cluster*, jika ada ulangi lagi langkah ke 4.

Clustering

Pada hakikatnya clustering adalah suatu metode untuk mencari dan mengelompokkan data yang mempunyai ciri-ciri yang serupa (serupa) antara satu data dengan data lainnya [12]. Clustering adalah metode data mining tanpa pengawasan, artinya diterapkan tanpa pelatihan, tanpa guru, dan tanpa memerlukan target keluaran [13]. Tujuan dari clustering data dapat dibedakan menjadi dua, yaitu:

1. Untuk memahami clustering harus menangkap struktur alami dari data yang dirancang hanya sebagai proses awal dan kemudian dilanjutkan dengan latar belakang sebagai ringkasan (mean standar deviasi), memberikan Label Kelas masing-masing kelompok digunakan sebagai data pelatihan klasifikasi.
2. Pengelompokan yang digunakan bertujuan untuk menemukan prototipe kelompok data yang paling representatif, memberikan abstraksi setiap objek data dalam kelompok tempat data berada.

2. HASIL DAN PEMBAHASAN

Tindakan kejahatan merupakan tindakan yang bersifat negatif. Seringkali, tindakan ini merugikan banyak pihak dan pelaku tindakannya disebut sebagai seorang kriminal. Kriminalitas adalah perilaku yang melanggar hukum dan menimbulkan kerugian bagi individu atau masyarakat luas.

Tabel 1. Data Mentah Peristiwa Kejahatan/Pelanggaran yang Dilaporkan menurut Jenis Kejahatan/Pelanggaran

No	Jenis Kejahatan/Pelanggaran	2019	2020	2021
1	Pembakaran	111	79	107
2	Kebakaran	158	155	56
3	Kejahatan Mata Uang	21	8	8
4	Kejahatan Materai dan Merk	7	5	16
5	Perkosaan	211	216	203
6	Perjudian	1069	581	514
7	Penculikan	19	25	28
8	Pembunuhan	107	106	96
9	Pencurian Ringan	2015	1618	28
10	Pencurian dengan Kekerasan	716	681	532
11	Pencurian dengan Pemberatan	4679	3908	4738
12	Penghinaan	415	278	457
13	Pemerasan	697	590	715
14	Penggelapan	2667	2555	2531
15	Penipuan	2450	2379	2736
16	Pengrusakan	776	608	723
17	Penadahan	55	13	4
18	Pencurian Kendaraan Bermotor	2982	2634	2620
19	Kejahatan Narkotik	6376	6218	5950
20	Penyelundupan	11	4	4
21	Korupsi	29	27	12
22	Penyalahgunaan Senjata Api	24	11	15
23	Sengketa Tanah	201	186	375
24	Illegal Logging	25	29	18

Implementasi

Berikut ini adalah beberapa contoh metode perhitungan manual yang digunakan dalam menggunakan algoritma k-means clustering, kami akan melakukan pengelompokan data Kriminalitas di provinsi sumatera utara. Pada penelitian ini menggunakan metode k-means untuk melakukan clustering berdasarkan jenis kriminalitas di provinsi sumut berdasarkan indikator tingkat kriminalitas dengan menggunakan langkah-langkah berikut :

1. Pada penelitian ini data diperoleh dari tahun 2019-2021 dan dicari nilai rata-ratanya. Langkah awal yang dapat dilakukan setelah mendapatkan rata-rata maka melakukan penentuan jumlah cluster, pada penelitian ini jumlah cluster yang akan digunakan adalah 3 cluster. Pada tahun 2019 memiliki cluster 1 = 21, tahun 2020 = 8 dan tahun 2021 = 8. Pada cluster 2 tahun 2019 memiliki cluster 1 = 19, tahun 2020 = 25 dan tahun 2021 = 28. Pada cluster 3 tahun 2019 = 24, tahun 2020 = 11 dan tahun 2021 = 15. Ketiga cluster tersebut diambil dari data pengangguran secara acak.

Tabel 2. Menentukan *Centroid Cluster* 1 Secara Acak

Centroid	2019	2020	2021
C1	21	8	8
C2	19	25	28
C3	24	11	15

2. Melakukan pencarian nilai C1, C2 dan C3 antara setiap pusat massa objek dan titik yang sesuai dengan menggunakan Jarak Euclidean.

Tabel 3. Menghitung jumlah cluster dan jarak

Jenis	2019	2020	2021	C1	C2	C3
Kejahatan Mata Uang	21	8	8	0	26,32489	8,185353
Kejahatan Materai dan Merk	7	5	16	16,40122	26,22975	18,05547
Penyelundupan	11	4	4	11,48913	32,87856	18,41195
Pembakaran	111	79	107	151,4662	132,7441	143,7254
Kebakaran	158	155	56	206,5962	192,3668	200,9303
Perkosaan	211	216	203	342,6208	322,4438	335,1686
Perjudian	1069	581	514	1297,177	1283,679	1290,708
Penculikan	19	25	28	26,32489	0	19,74842
Pembunuhan	107	106	96	157,3023	137,5827	149,9166
Pencurian Ringan	2015	1618	28	2562,916	2553,755	2558,652
Pencurian dengan Kekerasan	716	681	532	1100,241	1081,74	1093,185
Pencurian dengan Pemberatan	4679	3908	4738	7699,342	7679,674	7691,707
Penghinaan	415	278	457	655,5433	636,2908	647,7144
Pemerasan	697	590	715	1138,222	1118,427	1130,562
Penggelapan	2667	2555	2531	4455,789	4435,968	4448,33
Penipuan	2450	2379	2736	4354,729	4334,091	4347,038
Pengrusakan	776	608	723	1200,521	1181,509	1192,97
Pencurian Kendaraan Bermotor	2982	2634	2620	4741,934	4722,787	4734,545
Kejahatan Narkotik	6376	6218	5950	10689,13	10669,34	10681,71
Sengketa Tanah	201	186	375	445,8397	423,6201	437,6688
Illegal Logging	25	29	18	23,60085	12,32883	18,27567
Penadahan	55	13	4	34,59769	44,89989	32,95451
Korupsi	29	27	12	21	18,97367	17,02939
Penyalahgunaan Senjata Api	24	11	15	8,185353	19,74842	0

3. Langkah berikutnya melakukan perhitungan jarak terdekat dan mencari nilai terkecil dari cluster 1, cluster 2 dan cluster 3. Apabila nilai terkecil cluster 1 maka cluster 1 mendapatkan nilai 1, cluster 2 mendapatkan nilai 2 dan cluster 3 mendapatkan nilai 3. Setelah itu dilakukan pengelompokan cluster C1, C2 dan C3.

Tabel 4. Hasil Iterasi 1

C1	C2	C3	JARAK TERDEKAT	CLUSTER
0	26,32489	8,185353	0	1
16,40122	26,22975	18,05547	16,40121947	1
11,48913	32,87856	18,41195	11,48912529	1
	132,7441	143,7254	132,7441147	1
206,5962	192,3668	200,9303	192,3668371	2
342,6208	322,4438	335,1686	322,4437936	2
1297,177	1283,679	1290,708	1283,679088	2
26,32489	0	19,74842	0	2
157,3023	137,5827	149,9166	137,5827024	2
2562,916	2553,755	2558,652	2553,755078	2
1100,241	1081,74	1093,185	1081,739802	2
7699,342	7679,674	7691,707	7679,673756	2
655,5433	636,2908	647,7144	636,290814	2
1138,222	1118,427	1130,562	1118,426573	2
4455,789	4435,968	4448,33	4435,968102	2
4354,729	4334,091	4347,038	4334,090562	2
1200,521	1181,509	1192,97	1181,508781	2
4741,934	4722,787	4734,545	4722,786677	2
10689,13	10669,34	10681,71	10669,3384	2
445,8397	423,6201	437,6688	423,6201128	2
23,60085	12,32883	18,27567	12,32882801	2
34,59769	44,89989	32,95451	32,95451411	3
21	18,97367	17,02939	17,02938637	3
8,185353	19,74842	0	0	3

Dapat dilihat pada tabel 3 disimpulkan bahwa pada iterasi 1, C1 memiliki 3 data yaitu (3, 4, 20) dan C2 berjumlah 18 data yaitu (1, 2, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 18, 19, 23, 24) lalu C3 memiliki 3 data yaitu (17, 21, 22). sebelum mencari iterasi selanjutnya, penulis menentukan centroid terbaru. Langkah selanjutnya penulis mencari iterasi 2 sampai dengan iterasi 6 untuk mendapatkan hasil yang sama. Dalam melakukan pencarian nilai cluster 1, 2 dan 3 pada iterasi 5, maka dilakukan perhitungan dibawah ini :

Tabel 5. Hasil Iterasi 5

Jenis Kejahatan/Pelanggaran	JARAK				
	C1	C2	C3	TERDEKAT	CLUSTER
Pembakaran	51,22557	6230,954	1126,762	51,22557	1
Kebakaran	122,2623	6190,948	1068,979	122,2623	1
Kejahatan Mata Uang	102,5786	6381,455	1269,59	102,5786	1
Kejahatan Materai dan					
Merk	107,8488	6386,845	1278,516	107,8488	1
Perkosaan	240,4463	6039,71	942,5687	240,4463	1
Penculikan	82,86401	6361,649	1253,944	82,86401	1
Pembunuhan	55,87264	6224,786	1118,369	55,87264	1
Penghinaan	554,9858	5737,166	697,2979	554,9858	1
Penadahan	89,20494	6360,775	1243,678	89,20494	1
Penyelundupan	112,8003	6391,964	1280,691	112,8003	1
Korupsi	85,7742	6363,854	1251,515	85,7742	1
Penyalahgunaan					
Senjata Api	94,92255	6373,939	1263,029	94,92255	1
Sengketa Tanah	348,737	5964,253	929,4993	348,737	1
Ilegal Logging	83,06798	6361,652	1251,043	83,06798	1
Pencurian dengan					
Pemberatan	7597,279	1379,232	6484,593	1379,232	1
Penggelapan	4353,631	1929,805	3232,055	1929,805	1
Penipuan	4252,227	2051,874	3164,055	2051,874	1
Pencurian Kendaraan					
Bermotor	4640,158	1654,735	3506,273	1654,735	1
Kejahatan Narkotik	10587,05	4318,724	9454,316	4318,724	1
Perjudian	1202,967	5159,702	189,8246	189,8246	2
Pencurian Ringan	2484,282	4536,581	1466,999	1466,999	2
Pencurian dengan					
Kekerasan	999,6598	5291,549	239,2287	239,2287	3
Pemerasan	1036,074	5245,581	360,514	360,514	3
Pengrusakan	1098,833	5183,961	309,1408	309,1408	3

Berdasarkan perhitungan yang telah dilakukan, bahwa iterasi 5 mendapatkan cluster 1 sebanyak 19 data, cluster 2 sebanyak 2 data dan cluster 3 mendapatkan 3 data, sehingga disimpulkan jumlah cluster

1, 2 dan cluster 3 tidak sama dan harus melanjutkan pada iterasi ke 6. Dalam mencari nilai cluster 1, cluster 2 dan cluster 3 pada iterasi ke 6, maka dilakukan perhitungan berikut ini :

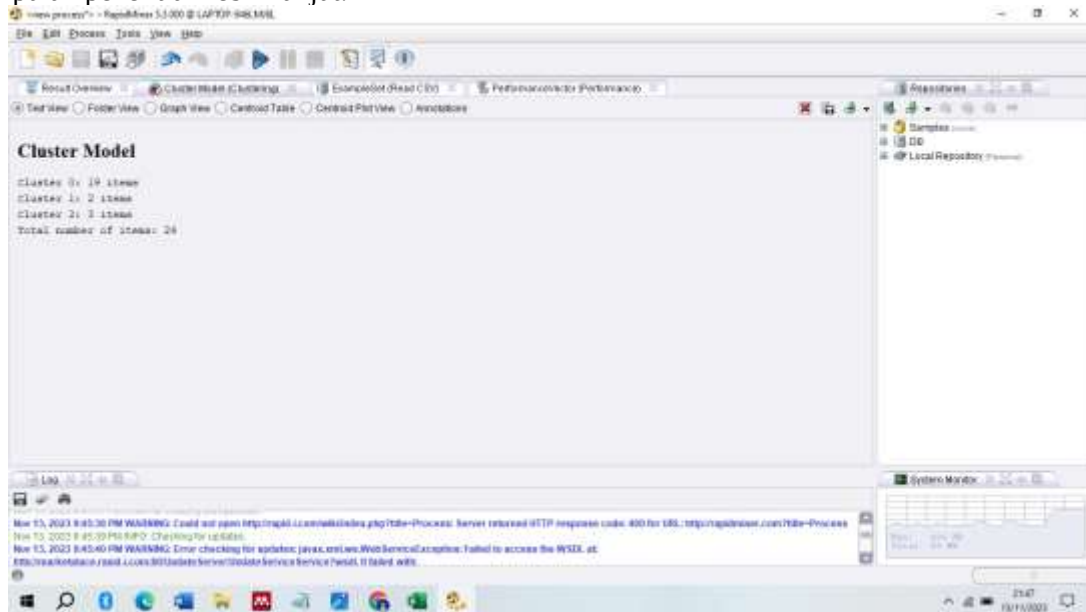
Tabel 6. Hasil Iterasi 5

Jenis Kejahatan/Pelanggaran	C1	C2	C3	JARAK TERDEKAT	CLUSTER
Pembakaran	13,68319	6230,954	1260,675	13,68319	1
Kebakaran	103,6116	6190,948	1199,815	103,6116	1
Kejahatan Mata Uang	141,5524	6381,455	1401,777	141,5524	1
Kejahatan Materai dan Merk	146,5492	6386,845	1411,072	146,5492	1
Perkosaan	202,7587	6039,71	1077,414	202,7587	1
Penculikan	121,9254	6361,649	1386,568	121,9254	1
Pembunuhan	25,83355	6224,786	1251,654	25,83355	1
Penadahan	126,0593	6360,775	1375,416	126,0593	1
Penyelundupan	151,8701	6391,964	1412,869	151,8701	1
Korupsi	125,2589	6363,854	1383,56	125,2589	1
Penyalahgunaan Senjata Api	133,8067	6373,939	1395,377	133,8067	1
Sengketa Tanah	311,2192	5964,253	1068,298	311,2192	1
Ilegal Logging	122,6252	6361,652	1383,278	122,6252	1
Penghinaan	515,3439	5737,166	836,7575	515,3439	1
Pencurian dengan Pemberatan	7558,07	1379,232	6374,913	1379,232	1
Penggelapan	4315,049	1929,805	3120,988	1929,805	1
Penipuan	4213,261	2051,874	3062,732	2051,874	1
Pencurian Kendaraan Bermotor	4601,526	1654,735	3392,002	1654,735	1
Kejahatan Narkotik	10548,59	4318,724	9337	4318,724	1
Pencurian Ringan	2456,553	4536,581	1338,383	1338,383	2
Perjudian	1166,48	5159,702	235,3276	235,3276	2
Pencurian dengan Kekerasan	962,2809	5291,549	365,5725	365,5725	3
Pemerasan	996,8618	5245,581	473,2567	473,2567	3
Pengrusakan	1059,67	5183,961	411,5581	411,5581	3

Berdasarkan perhitungan yang telah dilakukan disimpulkan bahwa iterasi ke 6 memiliki jumlah cluster yang sama seperti iterasi ke 5 yaitu C1 19 data, C2 2 data dan C3 3 data. Apabila nilai cluster sudah memiliki jumlah cluster yang sama pada objek yang sama perhitungan iterasi dapat dihentikan.

4. Langkah berikutnya melakukan pengujian *software Rapidminer* versi 5.3. Dimana data yang digunakan dalam pengujian adalah tahun 2019 sampai 2021. Berdasarkan hasil pengujian dengan menggunakan perangkat lunak RapidMiner 5.3 dan perhitungan manual menggunakan metode K-means pada aplikasi Microsoft Excel 2010, dapat disimpulkan bahwa hasil perhitungan manual dan hasil yang

dihasilkan oleh RapidMiner 5.3 sama. Hal ini menunjukkan bahwa perhitungan manual yang dilakukan dalam penelitian ini telah diverifikasi dan divalidasi dengan baik, sehingga dapat diandalkan sebagai representasi yang tepat dari data. Hasil ini memberikan kepercayaan bahwa data yang digunakan dalam penelitian memiliki keakuratan yang baik, dan bahwa algoritma K-means yang digunakan di kedua platform memberikan hasil yang konsisten. Dengan demikian, penelitian ini memberikan bukti yang kuat bahwa perhitungan manual yang dilakukan dalam analisis data menggunakan Microsoft Excel 2010 sesuai dengan metode K-means pada RapidMiner 5.3, dan hasilnya dapat dijadikan sebagai dasar untuk kesimpulan penelitian lebih lanjut.



Gambar 2. Hasil cluster

3. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan dengan judul analisis kasus kriminalitas dengan menggunakan metode algoritma k-means pada provinsi Sumatra utara, Penelitian ini mengidentifikasi tiga kelompok kasus kriminalitas di Provinsi Sumatera Utara melalui penggunaan algoritma K-Means. Cluster pertama (C1) ada pada tahun 2020, Cluster kedua (C2) tahun 2021, dan Cluster ketiga (C3) tahun 2022. Dalam hal ini, hasil penelitian memberikan wawasan yang berharga bagi pemerintah daerah, instansi terkait, dan pengambil keputusan. Pertama, identifikasi 3 kelompok utama kasus kriminalitas memungkinkan pengembangan program-program yang lebih terfokus. Untuk Kelompok C2 yang memiliki kasus kriminalitas yang sering terjadi diperlukan upaya lebih lanjut dalam menangani kasus kasus kriminalitas yang sering terjadi. Dan untuk kelompok C1 dan C3 dengan kasus kriminalitas yang jarang terjadi sebaiknya tetap diupayakan sebaik baiknya agar semua kasus kriminalitas semakin menurun. Dengan demikian, kesimpulan penelitian ini adalah pemahaman kasus kriminalitas yang sering terjadi di polresta sumatera utara untuk mengurangi kasus kriminalitas yang sering terjadi. Algoritma K-means juga membuktikan diri sebagai alat analisis yang bermanfaat dalam konteks ini. Diharapkan bahwa temuan ini akan membantu mengurangi kasus kriminalitas yang sering terjadi di wilayah tersebut.

UCAPAN TERIMA KASIH

Penulis ucapkan terima kasih untuk para rekan teman-teman yang telah memberi bantuan berupa kritik dan saran sehingga penulis dapat menyelesaikan penelitian ini.

REFERENSI

- [1] Y. Mayona, R. Buaton, and M. Simanjutak, "Data Mining Clustering Tingkat Kejahatan Dengan Metode Algoritma K-Means (Studi Kasus: Kejaksaan Negeri Binjai)," *J. Inform. Kaputama*, vol. 6, no. 3, pp. 2548–9739, 2022.
- [2] P. S. Rosiana, A. A. Mohsa, M. A. Fadila, and J. Haerul, "DATA KARAWANG DENGAN MENGGUNAKAN ALGORITMA CLUSTERING K-MEANS," vol. 11, no. 3, 2023.
- [3] M. Komariah, R. Pratama, and E. Adriansyah, "Dampak Pandemi Corona Virus Disease (Covid-19) Terhadap Tingkat Kriminalitas Di Kabupaten Ciamis," *J. Suara Huk.*, vol. 4, no. 1, pp. 22–39, 2023, doi: 10.26740/jsh.v4n1.p22-39.
- [4] U. T. Suryadi and Y. Supriatna, "Sistem Clustering Tindak Kejahatan Pencurian Di Wilayah Jawa Barat Menggunakan Algoritma K-Means," *J. Teknol. dan Komun. STMIK Subang*, vol. 12, no. 1, pp. 15–27, 2019, doi: 10.47561/a.v12i1.147.
- [5] A. Aini, A. Zakir, and D. Irwan, "Penerapan Algoritma C4.5 Dalam Pengelompokan Penanganan Kasus Kejahatan Yang Sering Terjadi Di Polsek Percut Sei Tuan," *Djtechno J. Teknol. Inf.*, vol. 3, no. 2, pp. 244–250, 2022, doi: 10.46576/djtechno.v3i2.2706.
- [6] E. D. Arifonang, H. S. Tambunan, and ..., "Penerapan Data Mining Dalam Mengelompokkan Provinsi Rawan Kejahatan Menggunakan Algoritma K-Means," *KOMIK (Konferensi ...)*, vol. 4, pp. 35–42, 2020, doi: 10.30865/komik.v4i1.2576.
- [7] Z. Nabila, A. Rahman Isnain, and Z. Abidin, "Analisis Data Mining Untuk Clustering Kasus Covid-19 Di Provinsi Lampung Dengan Algoritma K-Means," *J. Teknol. dan Sist. Inf.*, vol. 2, no. 2, p. 100, 2021, [Online]. Available: <http://jim.teknokrat.ac.id/index.php/JTSI>
- [8] P. Dwi Lestari and M. Mulyawan, "Datamining Pada Penjualan Air Bersih Di Spam Akidah Menggunakan Algoritma K-Means Clustering Menggunakan Rapidminer," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 7, no. 1, pp. 412–416, 2023, doi: 10.36040/jati.v7i1.6315.
- [9] H. D. Tampubolon, D. Gultom, L. Y. Hutabarat, F. I. R.H Zer, and D. Hartama, "Penerapan Algoritma K-Means Untuk Mengetahui Tingkat Tindak Kejahatan Daerah Pematangsiantar," *J. Teknol. Inf.*, vol. 4, no. 1, pp. 146–151, 2020, doi: 10.36294/jurti.v4i1.1263.
- [10] 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.
- [11] M. Soleh, S. R. Andani, and ..., "Algoritma C4. 5 Dalam Mengklasifikasi Tindak Kejahatan Yang Sering Terjadi Di Kecamatan Siantar Barat Pematangsiantar," *TIN Terap. Inform. ...*, vol. 2, no. 5, pp. 335–340, 2021, [Online]. Available: <http://ejurnal.seminar-id.com/index.php/tin/article/view/916%0Ahttp://ejurnal.seminar-id.com/index.php/tin/article/download/916/628>
- [12] N. Septiani, K. Erwansyah, and ..., "Analisis Data Mining Pengelompokkan Kasus Tindak Kejahatan Yang Terjadi Di Kecamatan Medan Polonia Dengan Menggunakan Metode K-Means Clustering," *J. Cyber Tech*, vol. 3, no. 2, pp. 1–13, 2020, [Online]. Available: <https://ojs.trigunadharma.ac.id/index.php/jct/article/view/3106%0Ahttps://ojs.trigunadharma.ac.id/index.php/jct/article/download/3106/1560>
- [13] R. Hidayat, "Clustering Menggunakan Algoritma K-Means untuk Mengelompokan Wilayah Rawan Kejahatan di Wilayah Kabupaten Solok," *J. Ilm. Mhs. Manajemen, Bisnis dan Akunt.*, vol. 4, no. 5, pp. 646–654, 2022, doi: 10.32639/jimmba.v4i5.169.