

## KLASTERISASI DAERAH PELAKU PENCURIAN MENGUNAKAN ALGORITMA K-MEANS

Pany Nurul Akbar  
Teknologi Informasi  
\*) panynurulakbarr@gmail.com

### Abstrak

Pencurian merupakan perilaku yang menyebabkan kerugian bagi korban yang menjadi sasaran dan bisa menyebabkan korban. Tingkat perilaku pencurian semakin meningkat di setiap daerah karena semakin banyaknya tingkat pengangguran dan sifat malas bekerja yang membuat seseorang melakukan pencurian untuk memenuhi kebutuhan hidupnya. Tujuan penelitian adalah untuk menganalisa dengan menggunakan teknik datamining pada daerah pelaku kejahatan pencurian berdasarkan provinsi. Teknik yang digunakan adalah klustering dengan metode K-means. Data bersumber dari Badan Pusat Statistik Indonesia dengan alamat url: <https://www.bps.go.id/>. Hasil dari penelitian dengan menggunakan teknik ini adalah berupa cluster terdapat daerah di Indonesia yang memiliki tingkat kejahatan pencurian tertinggi. Dari hasil penelitian menggunakan teknik K-means, bahwa terdapat 17 provinsi dari 34 provinsi yang daerah pelaku kejahatan pencurian tertinggi (C1) yakni: Aceh, Sumatera Utara, Sumatera Barat, Riau, Jambi, Sumatera Selatan, Lampung, DKI Jakarta, Jawa Barat, Jawa Tengah, Jawa Timur, Banten, Nusa Tenggara Barat, Nusa Tenggara Timur, Kalimantan Selatan, Sulawesi Selatan, Papua. Hasil penelitian diharapkan menjadi informasi bagi pemerintah dalam melakukan kebijakan mengurangi tingkat kejahatan pencurian di Indonesia yang sangat tinggi (> 50%).

**Kata Kunci:** Data Mining dan K-Means.

---

### PENDAHULUAN

Pencurian adalah perilaku yang melanggar hukum yang dapat menimbulkan korban jiwa, pencurian sering terjadi di pusat kota atau pusat perbelanjaan seperti pasar karena pusat tersebut merupakan pemicu untuk melakukan tindakan pencurian, pelakunya beroperasi secara individu atau Era yang semakin maju ini semakin meningkat sehingga sebagian besar tidak bekerja untuk memenuhi keinginan dan kebutuhannya, mereka tertekan secara ekonomi dan malas bekerja sehingga melakukan tindakan yang melanggar hukum. aktor biasanya beroperasi sendiri atau di tempat-tempat tertentu dengan tugas masing-masing untuk memfasilitasi operasi mereka (Sukawirasa et al., 2008), (Hafidz, 2021), (Celarier, n.d.). Di setiap daerah sering terjadi pencurian pencurian, setiap hari pencurian semakin banyak, pencurian semakin merajalela (Cindiyasari, 2017). Banyak faktor yang menyebabkan pencurian remaja meningkat, krisis ekonomi, lingkungan dan pergaulan yang tidak baik, usia tidak menjadi masalah bagi pelaku untuk melakukan tindakan buruk karena pola asuh yang buruk, tindakan, atau orang dewasa yang sudah tua dapat melakukan kejahatan yang merugikan tersebut (CS, 2019), (Aditomo Mahardika Putra, 2021), (Savestra et al., 2021).

Dari sekian banyak kasus pencurian di tingkat provinsi bervariasi, pengolahan data dilakukan dengan metode data mining, data mining merupakan metode yang digunakan untuk pengolahan data untuk menemukan citra tersembunyi dari data yang diolah. Data

yang diolah dengan metode data mining kemudian menghasilkan pengetahuan baru yang berasal dari data lama, hasil pengolahan data dan digunakan dalam menentukan keputusan di masa yang akan datang (BRONDONG, n.d.), (NASIONAL, n.d.). Dari kasus tersebut peneliti menerapkan algoritma k-means untuk mengelompokkan data kasus pencurian di setiap provinsi agar prosesnya lebih cepat dan efisien, algoritma k-means memberikan nilai cluster (K) secara acak, untuk sementara nilai-nilai tersebut menjadi pusat dari cluster atau disebut centroid, kemudian menghitung jarak setiap data yang ada ke setiap centroid menggunakan rumus Euclidian untuk mencari jarak terdekat dari setiap data ke centroid (Amin, 2020), (SETIYANTO, 2016), (Marlyna, 2017).

Penelitian sebelumnya yang menggunakan data mining dengan algoritma k-means seperti pada penerapan datamining pada populasi ayam pedaging di Indonesia menurut provinsi menggunakan k-means clustering (Heaverly & EWK, 2020), (Isnain et al., 2021). Dan menerapkan metode k-means dalam mengelompokkan penjualan produk di supermarket Fadhila. Dimana data yang diambil pada Supermarket Fadhilla Bengkulu menghasilkan dua jenis data kelompok yaitu data penjualan rendah dan data penjualan tinggi. Berdasarkan peneliti, ia mengangkat topik pengelompokan wilayah kejahatan pencurian menurut provinsi menggunakan metode K-means. Semoga hasil penelitian ini dapat menjadi masukan bagi pemerintah dalam kebijakan mengenai tindak pidana pencurian yang selalu meresahkan (V. A. Safitri et al., 2019), (Pinem, 2018), (Endang Woro Kasih, 2018).

## **KAJIAN PUSTAKA**

### **Data Mining**

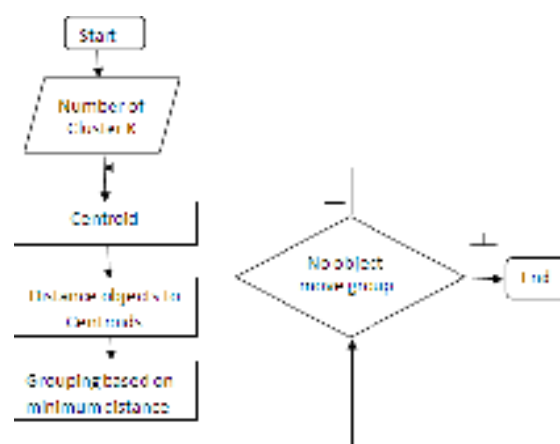
Data mining adalah proses pengumpulan dan pengolahan data yang bertujuan untuk mengekstrak informasi penting pada data. Proses pengumpulan dan ekstraksi informasi tersebut dapat dilakukan menggunakan perangkat lunak dengan bantuan perhitungan statistika, matematika, ataupun teknologi Artificial Intelligence (AI) (Hendrastuty, 2021), (Styawati et al., 2021), (Dharma et al., 2020). Data mining sering disebut juga Knowledge Discovery in Database (KDD). Data mining adalah metode dalam ilmu komputer yang biasa digunakan dalam proses pencarian knowledge. Tahapan di dalamnya berguna untuk mencari pola-pola tertentu dari data yang ada pada database (V. A. D. Safitri & Anggara, 2019), (V. A. Safitri et al., 2020), (Supriadi & Oswari, 2020). Biasanya, metode ini banyak ditemukan pada bidang machine learning dan statistika. Pada awalnya, metode penambangan data dikembangkan karena kompleksitas kerja komputer yang semakin meningkat (Putri et al., 2021), (Rossi et al., 2021). Namun, disinilah keuntungan adanya data mining adalah proses pengumpulan dan seleksi data yang lebih praktis. Data mining adalah sebuah proses menemukan sesuatu bermakna dengan memilah data melalui repository dengan bantuan teknologi sosialisasi pola, statistik, serta matematika (Susanto et al., 2021), (Pramita et al., n.d.), (Bertarina & Arianto, 2021). Data mining adalah sebuah aktivitas analisa data yang memiliki jumlah besar demi menemukan pola (pattern) dan aturan (rule) yang berguna (Agustina & Bertarina, 2022), (Sanjaya et al., 2014). Data mining adalah proses analisa yang dilakukan secara otomatis pada data yang kompleks dan berjumlah besar untuk memperoleh sebuah pola atau kecenderungan yang umumnya tidak disadari (Songati, 2018). Secara teknis, metode ini memang tidak lepas dari ilmu komputer dan kecerdasan buatan. Tetapi ternyata manfaatnya bisa dirasakan dalam berbagai bidang lainnya, termasuk bisnis dan pemasaran. Berikut sejumlah manfaat data mining. Mengetahui tren, metode untuk memprediksi keputusan bisnis di masa depan, mengetahui

produk yang dibeli bersamaan, mengamati perilaku konsumendan model sebagai sarana menyusun strategi peningkatan penjualan (Hasan, 2018), (Kurniawan, 2020), (Mathar et al., 2021).

## K-Means

K-means merupakan salah satu algoritma yang bersifat unsupervised learning. K-Means memiliki fungsi untuk mengelompokkan data kedalam data cluster. Algoritma ini dapat menerima data tanpa ada label kategori (Damayanti et al., 2021), (An'ars, 2022). K-Means Clustering Algoritma juga merupakan metode non-hierarchy. Metode Clustering Algoritma adalah mengelompokkan beberapa data ke dalam kelompok yang menjelaskan data dalam satu kelompok memiliki karakteristik yang sama dan memiliki karakteristik yang berbeda dengan data yang ada di kelompok lain (Anars et al., 2018), (Saputra, 2020b), (Suwarni et al., 2022). Cluster Sampling adalah teknik pengambilan sampel di mana unit-unit populasi dipilih secara acak dari kelompok yang sudah ada yang disebut 'cluster, nah Clustering atau klasterisasi adalah salah satu masalah yang menggunakan teknik unsupervised learning (Handayani et al., 2022), (Saputra, 2020a). K-Means Clustering adalah suatu metode penganalisaan data atau metode Data Mining yang melakukan proses pemodelan unssupervised learning dan menggunakan metode yang mengelompokan data berbagai partisi. K Means Clustering memiliki objective yaitu meminimalisasi object function yang telah di atur pada proses clasterisasi (AS & Baihaqi, 2020), (Akbar, 2019), (Bonar Siregar, 2021). Dengan cara minimalisasi variasi antar 1 cluster dengan maksimalisasi variasi dengan data di cluster lainnya (Budiman & Sidiq, n.d.). K means clustering merupakan metode algoritma dasar,yang diterapkan sebagai berikut, menentukan jumlah cluster, secara acak mendistribusikan data cluster, menghitung rata rata dari data yang ada di cluster, menggunakan langkah baris 3 kembali sesuai nilai threshold, menghitung jarak antara data dan nilai centroid(K means clustering), distance space dapat diimplementasikan untuk menghitung jarak data dan centroid. Contoh penghitungan jarak yang sering digunakan adalah manhattan/city blok distance. Clustering Algoritma (K-Means) memiliki tujuan untuk meminimalisasikan fungsi objective yang telah di set dalam proses clustering (PUSPITASARI, n.d.), (PRASETYAWAN, n.d.), (an Environmenta, n.d.). Tujuan tersebut dilakukan dengan cara meminimalikan variasi data yang ada didalam cluster dan memaksimalikan variasi data yang ada di cluster lainnya (Yuninda, 2020), (Kustinah & Indriawati, 2017).

## METODE



Gambar 1 Flowchart Metode K-Means

Berdasarkan gambar 1 dapat dijelaskan bahwa:

Tentukan jumlah cluster yang ingin dibentuk dan tetapkan pusat cluster k. Menggunakan jarak euclidean kemudian hitung setiap data ke pusat cluster.

$$d_{ik} = \sqrt{\sum_j^m (C_{ij} - C_{kj})^2}$$

Kelompokkan data ke dalam cluster dengan jarak yang paling pendek dengan persamaan

$$\text{Min} \sum_k^k = d_{ik} \sqrt{\sum_j^m (C_{ij} - C_{kj})^2}$$

Hitung pusat cluster yang baru menggunakan persamaan

$$C_{kj} = \frac{\sum_{i=1}^{x_{ij}} x_{ij}}{p}$$

Dengan  $x_{ij} \in$  cluster ke – k dan  $p$  = banyaknya anggota cluster ke – k. Ulangi langkah dua sampai dengan empat sehingga sudah tidak ada lagi data yang berpindah ke kluster yang lain.

## HASIL DAN PEMBAHASAN

Data diperoleh dari badan pusat statistik (<https://www.bps.go.id/>) untuk kategori kriminal yang kemudian di kelola dari data disetiap provinsi. Data yang digunakan merupakan data tahun 2016, 2017, dan 2018 yang terdiri dari 34 provinsi.

Tabel 1. Daerah Pelaku Kejahatan Pencurian Berdasarkan Provinsi

No	Provinsi	Tahun		
		2016	2017	2018
1	Aceh	140	63	47
2	Sumatera Utara	186	93	141
3	Sumatera Barat	57	55	71
4	Riau	149	79	101
5	Jambi	66	57	76
6	Sumatera Selatan	279	221	430
7	Bengkulu	38	33	40
8	Lampung	200	186	255
9	Kep. Bangka Belitung	24	16	15
10	Kep. Riau	21	13	17
11	DKI Jakarta	28	35	73
12	Jawa Barat	297	287	344
13	Jawa Tengah	132	146	176
14	Di Yogyakarta	7	28	20
15	Jawa Timur	269	290	419
16	Banten	78	54	49
17	Bali	8	13	23
18	Nusa Tenggara Barat	96	107	146
19	Nusa Tenggara Timur	61	54	67
20	Kalimantan Barat	39	33	37
21	Kalimantan Tengah	54	41	27
22	Kalimantan Selatan	117	85	73
23	Kalimantan Timur	55	25	37
24	Kalimantan Utara	0	0	9
25	Sulawesi Utara	22	27	21
26	Sulawesi Tengah	21	24	23

27	Sulawesi Selatan	69	54	47
28	Sulawesi Tenggara	6	23	16
29	Gorontalo	3	4	3
30	Sulawesi Barat	23	7	7
31	Maluku	0	9	20
32	Maluku Utara	5	2	7
33	Papua Barat	3	13	13
34	Papua	64	153	113

Pada penjelasan tabel 1, data akan diolah dengan mengambil nilai rata rata dari 3 tahun terakhir pada daerah pelaku kejahatan pencurian berdasarkan provinsi, sehingga diperoleh hasil seperti tabel berikut:

Tabel 2. Data yang akah digunakan

No	Provinsi	Nilai Rata - Rata
1	Aceh	83.33333333
2	Sumatera Utara	140
3	Sumatera Barat	61
4	Riau	109.6666667
5	Jambi	66.33333333
6	Sumatera Selatan	310
7	Bengkulu	37
8	Lampung	213.6666667
9	Kep. Bangka Belitung	18.33333333
10	Kep. Riau	17
11	DKI Jakarta	45.33333333
12	Jawa Barat	309.3333333
13	Jawa Tengah	151.3333333
14	Di Yogyakarta	18.33333333
15	Jawa Timur	326
16	Banten	60.33333333
17	Bali	14.66666667
18	Nusa Tenggara Barat	116.3333333
19	Nusa Tenggara Timur	60.66666667
20	Kalimantan Barat	36.33333333
21	Kalimantan Tengah	40.66666667
22	Kalimantan Selatan	91.66666667
23	Kalimantan Timur	39
24	Kalimantan Utara	9
25	Sulawesi Utara	23.33333333
26	Sulawesi Tengah	22.66666667
27	Sulawesi Selatan	56.66666667
28	Sulawesi Tenggara	15
29	Gorontalo	3.333333333
30	Sulawesi Barat	12.33333333
31	Maluku	14.5
32	Maluku Utara	4.666666667
33	Papua Barat	9.666666667
34	Papua	110

Penentuan titik cluster dilakukan dengan mengambil nilai terbesar (maksimum) untuk cluster daerah pelaku kejahatan pencurian yang tingkat tinggi (C1), nilai rata-rata (average) untuk cluster daerah pelaku kejahatan pencurian tingkat sedang (C2) dan nilai terkecil (minimum) untuk cluster daerah pelaku kejahatan pencurian tingkat rendah (C3). Nilai titik tersebut seperti yang ditunjukkan pada tabel berikut:

Tabel 3. *Centroid* awal

<i>cluster</i> tinggi	<i>cluster</i> sedang	<i>cluster</i> rendah
289.75	85.2381	18.44792

Proses cluster dengan mengambil jarak terdekat dari setiap data yang diolah. Dari data daerah pelaku kejahatan pencurian pada tahun 2016, 2017, 2018 yang diambil nilai rata-ratanya seperti yang ditunjukkan pada tabel 2, didapatkan pengelompokan pada iterasi 1 untuk 3 cluster tersebut. Cluster (C1) yakni 17 Provinsi : Aceh, Sumatera Utara, Sumatera Barat, Riau, Jambi, Sumatera Selatan, Lampung, DKI Jakarta, Jawa Barat, Jawa Tengah, Jawa Timur, Banten, Nusa Tenggara Barat, Nusa Tenggara Timur, Kalimantan Selatan, Sulawesi Selatan, Papua. cluster (C2) yakni 2 provinsi : Gorontalo, Maluku Utara dan cluster (C3) yakni 15 Provinsi : Bengkulu, Kep. Bangka Belitung, Kep. Riau, DI Yogyakarta, Bali, Kalimantan Barat, Kalimantan Tengah, Kalimantan Timur, Kalimantan Utara, Sulawesi Utara, Sulawesi Tengah, Sulawesi Tenggara, Sulawesi Barat, Maluku, Papua Barat. Proses pencarian jarak terpendek, pengelompokan data pada iterasi 1 dan Clustering data dapat digambarkan pada tabel dan gambar berikut:

Tabel 4. Perhitungan Jarak Pusat Cluster Iterasi 1

1	Aceh	83.33	242.67	5.465686272	80	5.465686272
2	Sumatera Utara	140	186	62.13235294	136.6666667	62.13235294
3	Sumatera Barat	61	265	16.86764706	57.66666667	16.86764706
4	Riau	109.67	216.33	31.79901964	106.3333334	31.79901964
5	Jambi	66.33	259.67	11.53431373	63	11.53431373
6	Sumatera Selatan	310	16	232.1323529	306.6666667	16
7	Bengkulu	37	289	40.86764706	33.66666667	33.66666667
8	Lampung	213.67	112.33	135.7990196	210.3333334	112.3333333
10	Kep. Riau	17	309	60.86764706	13.66666667	13.66666667
11	DKI Jakarta	45.33	280.67	32.53431373	42	32.53431373
12	Jawa Barat	309.33	16.67	231.4656862	306	16.6666667
13	Jawa Tengah	151.33	174.67	73.46568624	148	73.46568624
14	DI Yogyakarta	18.33	307.67	59.53431373	15	15
15	Jawa Timur	326	0	248.1323529	322.6666667	0
16	Banten	60.33	265.67	17.53431373	57	17.53431373
17	Bali	14.67	311.33	63.20098039	11.33333334	11.33333334
20	Kalimantan Barat	36.33	289.67	41.53431373	33	33
24	Kalimantan Utara	9	317	68.86764706	5.666666667	5.666666667
25	Sulawesi Utara	23.33	302.67	54.53431373	20	20
26	Sulawesi Tengah	22.67	303.33	55.20098039	19.33333334	19.33333334
27	Sulawesi Selatan	56.67	269.33	21.20098039	53.33333334	21.20098039
29	Gorontalo	3.33	322.67	74.53431372	0	0
30	Sulawesi Barat	12.33	313.67	65.53431373	8.999999997	8.999999997
31	Maluku	14.5	311.5	63.36764706	11.16666667	11.16666667
32	Maluku Utara	4.67	321.33	73.20098039	1.333333334	1.333333334
33	Papua Barat	9.67	316.33	68.20098039	6.333333334	6.333333334
34	Papua	110	216	32.13235294	106.6666667	32.13235294

Tabel 5. *Pengelompokan Data Iterasi 1*

No	Nama	Kelompok Data 1		
		C1	C2	C3
1	Aceh		1	
2	Sumatera Utara		1	
3	Sumatera Barat		1	
4	Riau		1	
5	Jambi		1	
6	Sumatera Selatan	1		
7	Bengkulu			1
8	Lampung	1		
9	Kep, Bangka Belitung			1
10	Kep, Riau			1
11	DKI Jakarta		1	
12	Jawa Barat	1		
13	Jawa Tengah		1	
14	Di Yogyakarta			1
15	Jawa Timur	1		
16	Banten		1	
17	Bali			1
18	Nusa Tenggara Barat		1	
19	Nusa Tenggara Timur		1	
20	Kalimantan Barat			1
21	Kalimantan Tengah		1	
22	Kalimantan Selatan		1	
23	Kalimantan Timur			1
24	Kalimantan Utara			1
25	Sulawesi Utara			1
26	Sulawesi Tengah			1
27	Sulawesi Selatan		1	
28	Sulawesi Tenggara			1
29	Gorontalo			1
30	Sulawesi Barat			1
31	Maluku			1
32	Maluku Utara			1
33	Papua Barat			1
34	Papua		1	

C1 : {6,8,12,15}

C2 : {1,2,3,4,5,11,13,16,18,19,21,22,27,34}

C3 : {7,9,10,14,17,20,23,24,25,26,28,29,30,31,32,33}

Proses iterasi akan terus berlangsung sampai hasil cluster terakhir sama dengan cluster sebelumnya. Pada penelitian ini cluster berhenti di iterasi 8 dengan hasil akhir seperti yang ditunjukkan pada tabel berikut:

Tabel 6. Perhitungan Jarak Pusat Cluster Iterasi 8

1	Aceh	83.33	9.67233333	80.28769333	71.96403333	9.67233333
2	Sumatera Utara	140	66.3391	136.9544	128.6307	66.3391
3	Sumatera Barat	61	12.6609	57.9544	49.6307	12.6609
4	Riau	109.67	36.0057667	106.6210667	98.2973667	36.0057667
5	Jambi	66.33	7.32756667	63.28773333	54.96403333	7.32756667
6	Sumatera Selatan	310	236.3391	306.9544	298.6307	236.3391
7	Bengkulu	37	36.6609	33.9544	25.6307	25.6307
8	Lampung	213.67	140.0057667	210.6210667	202.2973667	140.0057667
10	Kep. Riau	17	56.6609	13.9544	5.6307	5.6307
11	DKI Jakarta	45.33	28.32756667	42.28773333	33.96403333	28.32756667
12	Jawa Barat	309.33	235.6724333	306.2877333	297.9640333	235.6724333
13	Jawa Tengah	151.33	77.6724333	148.2877333	139.9640333	77.6724333
14	Di Yogyakarta	18.33	55.32756667	15.28773333	6.96403333	6.96403333
15	Jawa Timur	326	252.3391	322.9544	314.6307	252.3391
16	Banten	60.33	13.32756667	57.28773333	48.96403333	13.32756667
17	Bali	14.67	58.99423333	11.62106667	3.29736667	3.29736667
20	Kalimantan Barat	36.33	37.32756667	33.28773333	24.96403333	24.96403333
24	Kalimantan Utara	9	64.6609	5.9544	2.3693	2.3693
25	Sulawesi Utara	23.33	50.32756667	20.28773333	11.96403333	11.96403333
26	Sulawesi Tengah	22.67	50.99423333	19.62106667	11.29736667	11.29736667
27	Sulawesi Selatan	56.67	16.99423333	53.62106667	45.29736667	16.99423333
29	Gorontalo	3.33	70.32756667	0.28773333	8.03596667	0.28773333
30	Sulawesi Barat	12.33	61.32756667	9.28773333	0.96403333	0.96403333
31	Maluku	14.5	59.1609	11.4544	3.1307	3.1307
32	Maluku Utara	4.67	68.99423333	1.62106667	6.70263333	1.62106667
33	Papua Barat	9.67	63.99423333	6.62106667	1.70263333	1.70263333
34	Papua	110	36.3391	106.9544	98.6307	36.3391

Tabel 7. Pengelompokan Data Iterasi 8

No	Nama	Kelompok Data 1		
		C1	C2	C3
1	Aceh	1		
2	Sumatera Utara	1		
3	Sumatera Barat	1		
4	Riau	1		
5	Jambi	1		
6	Sumatera Selatan	1		
7	Bengkulu			1
8	Lampung	1		
9	Kep. Bangka Belitung			1
10	Kep. Riau			1
11	DKI Jakarta	1		
12	Jawa Barat	1		
13	Jawa Tengah	1		
14	Di Yogyakarta			1
15	Jawa Timur	1		
16	Banten	1		
17	Bali			1
18	Nusa Tenggara Barat	1		
19	Nusa Tenggara Timur	1		
20	Kalimantan Barat			1

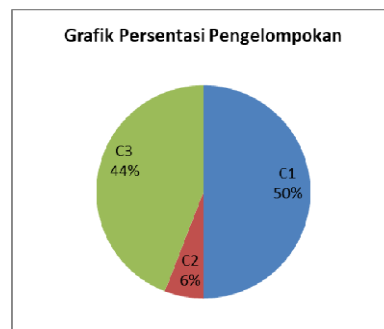


21	Kalimantan Tengah			1
22	Kalimantan Selatan	1		
23	Kalimantan Timur			1
24	Kalimantan Utara			1
25	Sulawesi Utara			1
26	Sulawesi Tengah			1
27	Sulawesi Selatan	1		
28	Sulawesi Tenggara			1
29	Gorontalo		1	
30	Sulawesi Barat			1
31	Maluku			1
32	Maluku Utara		1	
33	Papua Barat			1
34	Papua	1		

C1 : {1,2,3,4,5,6,8,11,12,13,15,16,18,19,22,27,34}

C2 : {29,32}

C3 : {7,9,10,14,17,20,21,23,24,25,26,28,31,31,33}



Gambar 2 Grafik Persentasi Pengelompokan

Berdasarkan hal tersebut bahwa teknik k-means pada kasus daerah pelaku kejahatan pencurian berdasarkan provinsi dengan menggunakan 3 cluster dengan pengelompokan sebagai berikut:

Cluster Tinggi (C1): Aceh, Sumatera Utara, Sumatera Barat, Riau, Jambi, Sumatera Selatan, Lampung, DKI Jakarta, Jawa Barat, Jawa Tengah, Jawa Timur, Banten, Nusa Tenggara Barat, Nusa Tenggara Timur, Kalimantan Selatan, Sulawesi Selatan, Papua, Cluster Normal (C2): Gorontalo, Maluku Utara, Cluster Rendah (C3): Bengkulu, Kep. Riau, Kep. Bangka Belitung, DI Yogyakarta, Bali, Kalimantan Barat, Kalimantan Tengah, Kalimantan Timur, Kalimantan Utara, Sulawesi Utara, Sulawesi Tengah, Sulawesi Tenggara, Sulawesi Barat, Maluku dan Papua Barat.

## SIMPULAN

Berdasarkan hasil penelitian dapat dikatakan bahwa teknik datamining K-Means dapat diterapkan pada kasus tindak pidana pencurian berdasarkan provinsi yang menghasilkan 3 cluster yaitu high cluster (C1) : 17 provinsi, normal cluster (C2) : 2 provinsi dan klaster rendah (C3). : 15 provinsi. Hasil kajian tersebut dapat menjadi sumber informasi bagi pemerintah khususnya yang berkepentingan bahwa angka pencurian yang ditemukan oleh provinsi masih sangat tinggi yaitu 50%.

## REFERENSI

- Aditomo Mahardika Putra, R. (2021). Underground Support System Determination: A Literature Review. *International Journal of Research Publications*, 83(1), 55–68. <https://doi.org/10.47119/ijrp100831820212185>
- Agustina, A., & Bertarina, B. (2022). ANALISIS KARAKTERISTIK ALIRAN SUNGAI PADA SUNGAI CIMADUR, PROVINSI BANTEN DENGAN MENGGUNAKAN HEC-RAS. *JICE (Journal of Infrastructural in Civil Engineering)*, 3(01), 31–41.
- Akbar, A. A. (2019). *Analisa Aplikasi OVO Menggunakan Model Delone & McLean Di Kalangan Mahasiswa Universitas Airlangga*. UNIVERSITAS AIRLANGGA.
- Amin, R. (2020). *IMPLEMENTASI RESTFULL API MENGGUNAKAN ARSITEKTUR MICROSERVICE UNTUK MANAJEMEN TUGAS KULIAH (STUDI KASUS: MAHASISWA STMIK AKAKOM)*. STMIK AKAKOM Yogyakarta.
- An'ars, M. G. (2022). Sistem Informasi Manajemen Berbasis Key Performance Indicator (KPI) dalam Mengukur Kinerja Guru. *Jurnal Data Mining Dan Sistem Informasi*, 3(1), 8–18.
- an Environmenta, C. E. (n.d.). *Pr idin*.
- Anars, M. G., Munaris, M., & Nazaruddin, K. (2018). Kritik Sosial dalam Kumcer Yang Bertahan dan Binasa Perlahan dan Rancangan Pembelajarannya. *Jurnal Kata (Bahasa, Sastra, Dan Pembelajarannya)*, 6(3 Jul).
- AS, N. R., & Baihaqi, I. (2020). Studi Inspeksi Kelayakan Instalasi Dan Instrumen Tenaga Listrik. *SINUSOIDA*, 22(2), 21–33.
- Bertarina, B., & Arianto, W. (2021). ANALISIS KEBUTUHAN RUANG PARKIR (STUDI KASUS: AREA PARKIR ICT UNIVERSITAS TEKNOKRAT INDONESIA). *Jurnal Teknik Sipil*, 2(02), 67–77.
- Bonar Siregar, B. (2021). *Pengembangan Sistem Perencanaan & Bantuan KRS*. Universitas Multimedia Nusantara.
- BRONDONG, L. (n.d.). *IDENTIFIKASI DAN PREVALENSI CACING PADA SALURAN PENCERNAAN IKAN KEMBUNG (Rastrelliger brachysoma) DI PELABUHAN PERIKANAN NUSANTARA*.
- Budiman, F., & Sidiq, M. (n.d.). *RANCANG BANGUN APLIKASI SISTEM INFORMASI APLIKASI DATA PETAMBAK*.
- Celarier, M. (n.d.). *RSS New York Times–Dealbook*.
- Cindiyasari, S. A. (2017). *Analisis Pengaruh Corporate Social Responsibility, Intellectual Capital, Dan Rasio Likuiditas Terhadap Kinerja Keuangan Perusahaan (Studi Kasus Perusahaan Perbankan yang Terdaftar di Bursa Efek Indonesia (BEI) tahun 2013-2015)*.
- CS, S. A. (2019). *Analisis Pengaruh Intellectual Capital Terhadap Kinerja Keuangan Perusahaan (Studi Kasus Perusahaan Sektor Keuangan Yang Terdaftar Di Bursa Efek Indonesia (BEI) Pada Tahun 2008-2017)*. Universitas Gadjah Mada.
- Damayanti, D., Yudiantara, R., & An'ars, M. G. (2021). SISTEM PENILAIAN RAPOR PESERTA DIDIK BERBASIS WEB SECARA MULTIUSER. *Jurnal Informatika Dan Rekayasa Perangkat Lunak*, 2(4), 447–453.

- Dharma, F., Shabrina, S., Noviana, A., Tahir, M., Hendrastuty, N., & Wahyono, W. (2020). Prediction of Indonesian inflation rate using regression model based on genetic algorithms. *Jurnal Online Informatika*, 5(1), 45–52.
- Endang Woro Kasih, E. (2018). Formulating Western Fiction in Garrett Touch of Texas. *Arab World English Journal For Translation and Literary Studies*, 2(2), 142–155. <https://doi.org/10.24093/awejtls/vol2no2.10>
- Hafidz, D. A. (2021). *Pengembangan Sistem Informasi Edukasi dan Pemasaran Hasil Pertanian di Tulang Bawang*.
- Handayani, M. A., Suwarni, E., Fernando, Y., Fitri, F., Saputra, F. E., & Candra, A. (2022). PENGELOLAAN KEUANGAN BISNIS DAN UMKM DI DESA BALAIREJO. *Suluh Abdi*, 4(1), 1–7.
- Hasan, A. F. (2018). *400 Kebiasaan Keliru dalam Hidup Muslim*. Elex Media Komputindo.
- Heaverly, A., & EWK, E. N. (2020). Jane Austen’s View on the Industrial Revolution in *Pride and Prejudice*. *Linguistics and Literature Journal*, 1(1), 1–6. <https://doi.org/10.33365/llj.v1i1.216>
- Hendrastuty, N. (2021). Rancang Bangun Aplikasi Monitoring Santri Berbasis Android (Studi Kasus: Pesantren Nurul Ikhwan Maros). *Jurnal Data Mining Dan Sistem Informasi*, 2(2), 21–34.
- Isnain, A. R., Hendrastuty, N., Andraini, L., Studi, P., Informasi, S., Indonesia, U. T., Informatika, P. S., Indonesia, U. T., Studi, P., Komputer, T., Indonesia, U. T., & Lampung, K. B. (2021). *Comparison of Support Vector Machine and Naïve Bayes on Twitter Data Sentiment Analysis*. 6(1), 56–60.
- Kurniawan, A. H. (2020). Konsep Altmetrics dalam Mengukur Faktor Dampak Artikel Melalui Academic Social Media dan Non-academic Social Media. *UNILIB: Jurnal Perpustakaan*, 11(1), 43–49.
- Kustinah, S., & Indriawati, W. (2017). Pengaruh Perputaran Persediaan dan Perputaran Piutang Terhadap Profitabilitas Pada Unit Usaha Toserba Koperasi PT LEN Bandung. *Journal Study & Accounting Research*, 14(1), 27–35.
- Marlyna, D. (2017). Pengaruh Peran Auditor Intern Terhadap Kinerja Perusahaan Angkutan Sungai, Danau Dan Penyeberangan. *Jurnal Ilmiah GEMA EKONOMI*, 3(2 Agustus), 321–332.
- Mathar, T., Hijrana, H., Haruddin, H., Akbar, A. K., Irawati, I., & Satriani, S. (2021). The Role of UIN Alauddin Makassar Library in Supporting MBKM Program. *Proceedings of the International Conference on Social and Islamic Studies (SIS) 2021*.
- NASIONAL, P. P. (n.d.). *KEMENTERIAN PENDIDIKAN DAN KEBUDAYAAN*.
- Pinem, Y. A. (2018). Encouraging healthy literacy: The interconnection between reading toward writing in social media. *Language in the Online and Offline World 6: The Fortitude*, 360–366.
- Pramita, G., Lestari, F., & Bertarina, B. (n.d.). Study on the Performance of Signaled Intersections in the City of Bandar Lampung (Case Study of JL. Sultan Agung-Kimaja Intersection durig Covid-19. *Jurnal Teknik Sipil*, 20(2).
- PRASETYAWAN, D. W. I. G. (n.d.). *LAPORAN INDIVIDU PRAKTIK PENGALAMAN LAPANGAN (PPL) DI SD NEGERI TLOGOADI PERIODE 10 AGUSTUS–12*

SEPTEMBER 2015.

- PUSPITASARI, R. D. (n.d.). *LAPORAN KEGIATAN PRAKTIK PENGALAMAN LAPANGAN (PPL) DI SD NEGERI TLOGOADI PERIODE 10 AGUSTUS–12 SEPTEMBER 2015*.
- Putri, N. U., Rossi, F., Jayadi, A., Sembiring, J. P., & Maulana, H. (2021). Analysis of Frequency Stability with SCES's type of Virtual Inertia Control for The IEEE 9 Bus System. *2021 International Conference on Computer Science, Information Technology, and Electrical Engineering (ICOMITEE)*, 191–196.
- Rossi, F., Sembiring, J. P., Jayadi, A., Putri, N. U., & Nugroho, P. (2021). Implementation of Fuzzy Logic in PLC for Three-Story Elevator Control System. *2021 International Conference on Computer Science, Information Technology, and Electrical Engineering (ICOMITEE)*, 179–185.
- Safitri, V. A. D., & Anggara, B. (2019). FACTORS THAT AFFECT THE COMPANY INNOVATION. II. In *TradersUluslararası Ticaret Kongresi Kongre Kitabı The Second InTraders International Conference on International Trade Conference Book*, 230.
- Safitri, V. A., Sari, L., & Gamayuni, R. R. (2019). Research and Development, Environmental Investments, to Eco-Efficiency, and Firm Value. *The Indonesian Journal of Accounting Research*, 22(03), 377–396. <https://doi.org/10.33312/ijar.446>
- Safitri, V. A., Sari, L., & Gamayuni, R. R. (2020). Research and Development (R&D), Environmental Investments, to Eco-Efficiency, and Firm Value. *The Indonesian Journal of Accounting Research*, 22(3).
- Sanjaya, R., Nurweni, A., & Hasan, H. (2014). The Implementation of Asian-parliamentary Debate in Teaching Speaking at Senior High School. *U-JET*, 3(8).
- Saputra, F. E. (2020a). Analisis faktor-faktor yang mempengaruhi Kinerja Keuangan Bank Umum Syariah yang terdaftar di Bursa Efek Indonesia (BEI) Periode 2016-2018. *TECHNOBIZ: International Journal of Business*, 3(1), 45–50.
- Saputra, F. E. (2020b). *ANALISIS PENGARUH FDR, BOPO, DAN NPF TERHADAP KINERJA BANK UMUM SYARIAH DI INDONESIA PERIODE TAHUN JANUARI 2015 S/D JULI 2020*. Universitas Teknokrat Indonesia.
- Savestra, F., Hermuningsih, S., & Wiyono, G. (2021). Peran Struktur Modal Sebagai Moderasi Penguatan Kinerja Keuangan Perusahaan. *Jurnal Ekonika: Jurnal Ekonomi Universitas Kadiri*, 6(1), 121–129.
- SETIYANTO, A. (2016). *PENATAAN KELEMBAGAAN PRODUKSI UNTUK PENINGKATAN NILAI TAMBAH STUDI KASUS PADA ASOSIASI PRIMA SEMBADA*. Universitas Gadjah Mada.
- Songati, N. C. (2018). *An assessment of pedagogical strategies of teaching English at ordinary secondary level: a case of Kasulu district in Tanzania*. The University of Dodoma.
- Styawati, S., Hendrastuty, N., & Isnain, A. R. (2021). Analisis Sentimen Masyarakat Terhadap Program Kartu Prakerja Pada Twitter Dengan Metode Support Vector Machine. *Jurnal Informatika: Jurnal Pengembangan IT*, 6(3), 150–155.

- Sukawirasa, I. K. A., Udayana, I. G. A., Mahendra, I. M. Y., Saputra, G. D. D., & Mahendra, I. B. M. (2008). Implementasi Data Warehouse Dan Penerapannya Pada PHI-Minimart Dengan Menggunakan Tools Pentaho dan Power BI. *Jurnal Elektronik Ilmu Komputer Udayana P-ISSN, 2301*, 5373.
- Supriadi, A., & Oswari, T. (2020). Analysis of Geographical Information System (GIS) design application in the Fire Department of Depok City. *Technium Soc. Sci. J.*, 8, 1.
- Susanto, T., Setiawan, M. B., Jayadi, A., Rossi, F., Hamdhi, A., & Sembiring, J. P. (2021). Application of Unmanned Aircraft PID Control System for Roll, Pitch and Yaw Stability on Fixed Wings. *2021 International Conference on Computer Science, Information Technology, and Electrical Engineering (ICOMITEE)*, 186–190.
- Suwarni, E., Handayani, M. A., Fernando, Y., Saputra, F. E., & Candra, A. (2022). Penerapan Sistem Pemasaran berbasis E-Commerce pada Produk Batik Tulis di Desa Balairejo. *Jurnal Pengabdian Masyarakat Indonesia, 2*(2), 187–192.
- Yuninda, P. (2020). *The Use of Macromedia Flash as a Media in Learning Vocabulary at Third Grade of SDN Pademawu Barat IV Pamekasan*. INSTITUT AGAMA ISLAM NEGERI MADURA.