

**PENERAPAN METODE *CLUSTERING* DENGAN ALGORITMA *K-MEANS*
TINDAK KEJAHATAN PENCURIAN DI KABUPATEN ASAHAN****Nur Afni Syahpitri Damanik¹, Irianto^{2*}, Dahriansah³**¹Mahasiswa Prodi Sistem Informasi, STMIK Royal²Prodi Sistem Informasi, STMIK Royal³Prodi Manajemen Informasi, STMIK Royal**email: irianto2121212@gmail.com*

Abstract:Theft is the illegal taking of property or belongings of another person without the permission of the owner. The most common crime problem in Asahan District is theft, so that the POLRES is still having trouble determining which areas are often the crime of theft. With this problem, we need to do a grouping for areas where theft often occurs, so the process used is the data mining process. Data mining is one of the processes of Knowledge Discovery from Databases (KDD). KDD is an activity that includes collecting, using historical data to find regularities, patterns or relationships in large data sets. One of the techniques known in data mining is clustering technique. The K-Means method is a method for clustering techniques, K- Means is a method that partitions data into groups so that data with the same characteristics are entered into the same set of groups and data with different characteristics are grouped into other groups. The attributes used in grouping this data are annual data, namely 2015, 2016, 2017, 2018, 2019. A case study of 9 POLSEK in the Asahan.

Keywords: Data Mining, Clustering, K-Means Algorithm, Theft Crimes Grouping.

Abstrak: Pencurian merupakan pengambilan properti atau barang milik orang lain secara tidak sah tanpa ijin dari pemilik. Masalah tindak kejahatan yang paling banyak terjadi di Kabupaten Asahan adalah tindak kejahatan pencurian sehingga pihak POLRES masih kesulitan untuk menentukan daerah mana saja yang sering terjadi tindak kejahatan pencurian. Dengan adanya masalah ini kita perlu melakukan pengelompokan untuk daerah mana saja yang sering terjadi tindak pencurian maka proses yang digunakan adalah proses data mining. Data mining adalah salah satu proses dari Knowledge Discovery from Databases (KDD). KDD adalah kegiatan yang meliputi pengumpulan, pemakaian data, historis untuk menemukan keteraturan, pola atau hubungan dalam set data besar. Salah satu teknik yang di kenal dalam data mining adalah teknik clustering. Metode K-Means merupakan metode untuk teknik clustering, K-Means adalah metode yang mempartisi data kedalam kelompok sehingga data berkarakteristik sama dimasukkan kedalam set kelompok yang sama dan data yang berkarakteristik berbeda dikelompokkan ke dalam kelompok yang lain. Atribut yang di gunakan dalam pengelompokan data ini adalah data pertahun yaitu tahun 2015, 2016, 2017, 2018, 2019. Studi kasus pada 9 POLSEK yang ada di daerah kabupaten Asahan.

Kata kunci: Data Mining, Clustering, Algoritma K-Means, Pengelompokan Tindak Kejahatan Pencurian.

PENDAHULUAN

Masalah tindak kejahatan yang paling banyak terjadi di Kabupaten Asahan adalah tindak kejahatan pencurian sehingga pihak Polres Asahan masih kesulitan untuk menentukan daerah mana saja yang sering terjadi tindak kejahatan pencurian dikarenakan belum tersediannya sistem yang dapat memudahkan dan proses perhitungan data yang di dapat masih secara manual yaitu setiap Polsek harus datang setiap satu bulan sekali untuk memberikan hasil rekapan dari Polsek ke Polres Asahan.

Hal ini yang sering membuat para petugas bingung untuk melakukan patroli di daerah mana saja yang di perlukan untuk melakukan kegiatan patroli ataupun sosialisasi agar masyarakat sekitar ikut berpartisipasi membantu kegiatan patroli tersebut, sehingga masyarakat sekitar merasa aman untuk melakukan segala aktivitas kesehariannya tanpa adanya rasa takut atau tidak aman. Dari penjelasan masalah diatas penulis memiliki tujuan dari penelitian ini adalah : Menerapkan proses *data mining* dengan metode *K-Means Clustering* untuk pengelompokan daerah yang sering terjadi pencurian di Kabupaten Asahan [1]., Menguji penerapan metode *K-Means Clustering* menggunakan *PHP* [2]., Membangun sistem baru untuk pembagian daerah yang sering terjadi tindak kejahatan pencurian berdasarkan data tindak kejahatan pencurian yang ada di Polres Asahan [3].

METODE

Clustering atau klasifikasi adalah metode yang digunakan membagi rangkaian data menjadi beberapa group berdasarkan kesamaan-kesamaan yang telah ditentukan sebelumnya [4]. Pada algoritma *K-Means Cluster Analysis* terdapat beberapa langkah yang harus dilakukan yaitu sebagai berikut: Menentukan jumlah *cluster*, jumlah *cluster* merupakan jumlah kelompok yang akan diselesaikan, Menentukan *centroid* awal, *centroid* awal diperoleh secara acak [5]. *Centroid* awal merupakan titik pusat *cluster* pertama. Rumus menentukan *centroid* baru sebagai berikut:

$$v_{ij} = \frac{1}{N_i} \sum_{k=0}^{N_i} x_{kj} \quad (1)$$

Menghitung jarak antara titik *centroid* dengan titik objek. digunakan rumus sebagai berikut:

$$D_e = \sqrt{(x_i - S_i)^2 + (y_i - t_i)^2} \quad (2)$$

Pengelompokan objek untuk menentukan anggota *cluster*, adalah dengan memperhitungkan jarak minimum objek. Kembali ke tahap 2, lakukan perulangan hingga nilai *centroid* yang dihasilkan tetap dan anggota *cluster* tidak berpinda ke *cluster* lain.

Struktur basis *data mining clustering* tindak kejahatan pencurian di Kabupaten Asahan di uraikan sebagai berikut :

Struktur data penentuan *centroid*

Cluster di kelompokkan menjadi 3 kelompok berdasarkan data yang tersedia, sebagai berikut:

Tabel 1. Data Penentuan *Centroid*

NO .	SATUAN/ POLSEK	TAHUN 2015	TAHUN 2016	TAHUN 2017	TAHUN 2018	TAHUN 2019	KRIT ERIA
		JTP	JTP	JTP	JTP	JTP	
1.	Polsek Air Jorman	25	28	89	53	40	C2
2.	Polsek simpang Emat	63	59	44	46	33	
3.	Polsek Sei Kepayang	67	63	68	52	42	
4.	Polsek Bandar Pulau	49	53	73	43	35	
5.	Polsek Pulau Raja	65	93	71	53	33	
6.	Polsek Air Batu	69	87	89	43	58	C1
7.	Polsek Perapat Janji	53	63	88	55	24	
8.	Polsek Bp.Mandoge	31	48	79	35	33	
9.	Polsek Kota Kisaran	0	138	65	21	19	C3
Jumlah		1016	1028	935	705	608	

Menentukan pusat awal *cluster "Centroid"* Untuk penentuan awal diasumsikan: (a) Diambil data ke 6 sebagai pusat *cluster* ke 1., (b) Diambil data ke 1 sebagai pusat *cluster* ke 2., (c) Diambil data ke 9 sebagai pusat *cluster* ke 3.

Tabel 2. *cluster "Centroid"*

Centroid 1	(69,87,89,43,58)
Centroid 2	(25,28,89,53,40)
Centroid 3	(0,138,65,21,19)

Struktur data hasil iterasi pertama

Perhitungan jarak dari data ke 1 sampai data ke 9 terhadap pusat *cluster*, Sehingga hasil perhitungan jarak selengkapya adalah sebagai berikut:

Tabel 3. Hasil Iterasi Pertama

Data Ke	C1	C2	C3
1	76.43	0.00	121.52
2	58.98	67.29	107.11
3	36.85	58.61	107.76
4	48.38	39.77	102.13
5	33.18	78.73	86.64
6	0.00	76.43	99.71
7	46.18	47.64	100.72
8	61.27	30.15	98.23
9	99.71	121.52	0.00

Struktur pengelompokan data

Tabel 4. Pengelompokan Data

C1	C2	C3
1	2	0
2	1	0
2	1	0
1	2	0
2	1	0
2	1	0
2	1	0
1	2	0
0	1	2

Struktur data penentuan pusat *cluster* baru

Tabel 5. Centroid Baru

Data Ke	2015	2016	2017	2018	2019
C1	63.4	73	72	49.8	38
C2	35	43	80.33	43.67	36
C3	0	138	65	21	19

Struktur data penentuan pusat *cluster* baru

Tabel 6. Hasil Iterasi Ke Dua

Data Ke	C1	C2	C3
1	61.67	22.43	121.52
2	31.93	48.73	107.11
3	12.24	41.01	107.76
4	25.76	18.74	102.13
5	20.95	59.86	86.64
6	31.03	60.43	99.71
7	26.21	32.48	100.72
8	44.36	11.26	98.23
9	97.39	106.24	0.00

Struktur pengelompokan data

Tabel 7. Pengelompokan Data Ke Dua

C1	C2	C3
1	2	0
2	1	0
2	1	
1	2	0
2	1	0
2	1	0
2	1	0
1	2	0
1	0	2

Karena C1,C2,C3 sama dimana anggota yang sama, maka tidak perlu dilakukan itersi lagi, sampai disini hasil *Clustering* sudah mencapai stabil dan kovergen [6]. Jadi dari data di atas di dapat di dapat hasil bahwa : Untuk *cluster* 5 yaitu ada 1 daerah sebagai berikut : (a) Data Ke-2 : Polsek Simpang Empat., (b) Data Ke-3 : Polsek Sei Kepayang 3.Data Ke-5 : Polsek Pulau Raja., (c) Data Ke-6 : Polsek Air Batu., (d) Data Ke-7 : Polsek Perapat Janji. Untuk *cluster* 2 yaitu ada 8 daerah sebagai berikut : (a) Data ke-1 : Polsek Air Jorman., (b) Data Ke-8 : Polsek BP. Mandoge. Untuk *cluster* 3 yaitu ada 1 daerah sebagai berikut : Data Ke-9 : Polsek Kota Kisaran.

HASIL DAN PEMBAHASAN

Tampilan Halaman Login

Untuk menguji halaman *login*, dibutuhkan *username* dan *password* agar dapat masuk ke halaman admin panelnya. Pengguna yang dapat masuk ke dalam sistem ini adalah admin, petugas polres dan petugas polsek. Masukkan *username* dan *password* pada *field* nya masing – masing seperti pada gambar 1.



Gambar 1. Tampilan Halaman *Login*

Tampilan Halaman Data Pencurian

Tampilan halaman data pencurian adalah tampilan yang digunakan untuk menampilkan data pencurian. Untuk menguji tampilan ini, klik pada menu *Trend Data Pencurian* maka akan muncul halaman data pencurian seperti gambar 2.

#	Polsek	Jumlah	Tahun	Aksi
1	POLSEK AIR JOMAN	25	2015	Edit Hapus
2	POLSEK AIR JOMAN	28	2016	Edit Hapus
3	POLSEK AIR JOMAN	99	2017	Edit Hapus
4	POLSEK AIR JOMAN	93	2018	Edit Hapus
5	POLSEK AIR JOMAN	40	2019	Edit Hapus
6	POLSEK SIMPANG EMPAT	82	2015	Edit Hapus
7	POLSEK SIMPANG EMPAT	98	2016	Edit Hapus

Gambar 2. Tampilan Halaman Data Pencurian

Tampilan Halaman Hitung K-Means

Tampilan hitung *K-means* adalah Tampilan yang digunakan untuk melihat data perhitungan. Untuk menguji tampilan ini, klik pada menu Hitung *K-means* maka akan muncul halaman hitung *K-means* seperti gambar 3.

#	Polsek	2015	2016	2017	2018	2019
1	POLSEK AIR JOHAN	25	28	89	53	40
2	POLSEK SIMPANG EMRAT	63	59	44	46	33
3	POLSEK SEI KEPERANG	67	63	68	52	42
4	POLSEK BANDAR PULAU	49	53	73	43	35
5	POLSEK PULAU RAJA	65	92	71	53	33
6	POLSEK AIR BATU	69	87	88	43	58
7	POLSEK PERAKAT JALUI	53	63	88	55	24
8	POLSEK BP MANDAGE	31	46	79	36	33
9	POLSEK KOTA KISARAN	0	138	65	21	18

Gambar 3. Tampilan Halaman Hitung *K-Means*



Gambar 4. Cetak Hasil Hitung *K-Means*

SIMPULAN

Dari penelitian ini dapat disimpulkan beberapa yang menjadi kesimpulan mendasar pada penelitian ini diantaranya adalah sebagai berikut: (1) *K-Means Clustering* dapat memberikan hasil pengelompokan data tindak kejahatan pencurian yang dilakukan oleh *admin/petugas* Polres Asahan., (2) Dalam proses pengelompokan data pencurian, sistem yang digunakan dapat melakukan pengelompokan secara otomatis dengan menggunakan metode *K-Means Clustering*., (3) Dengan adanya penerapan teknik *Data Mining* dengan Metode *K-Means Clustering* dapat disimpulkan bahwa teknik ini dapat di terapkan pada kasus kejahatan pencurian berdasarkan polsek-polsek yang tergabung yang menghasilkan 3 *Cluster*, Yakni C1 Tinggi, C2 Menengah, C3 Rendah., (4) Penerapan program yang dibuat dengan menggunakan aplikasi pemrograman berbasis *web*, dan hasilnya dengan menggunakan program yang dibuat dapat lebih memudahkan pengguna untuk menentukan daerah mana saja yang tingkat kejahatan pencuriannya tinggi di Kabupaten Asahan.

DAFTAR PUSTAKA

- [1] Y. Heriyanto, “Perancangan Sistem Informasi Rental Mobil Berbasis Web Pada PT.APM Rent Car,” *J. Intra-Tech*, vol. 2, no. 2, pp. 64–77, 2018.
- [2] Y. Di and J. Dipasaran, “1 , 1 , 2,” vol. VI, no. 2, pp. 1–6, 2013, doi: 10.2478/amsc-2014-0077.
- [3] K. Kawano, Y. Umemura, and Y. Kano, “ Field Assessment and Inheritance of Cassava Resistance to Superelongation Disease 1 ,” *Crop Sci.*, vol. 23, no. 2, pp. 201–205, 1983, doi: 10.2135/cropsci1983.0011183x002300020002x.
- [4] M. Mardalius, “Pemanfaatan Rapid Miner Studio 8.2 Untuk Pengelompokan Data Penjualan Aksesoris Menggunakan Algoritma K-Means,” *Jurteksi*, vol. 4, no. 2, pp. 123–132, 2018, doi: 10.33330/jurteksi.v4i2.36.
- [5] M. L. Sibuea and A. Safta, “Pemetaan Siswa Berprestasi Menggunakan Metode K-Means Clustring,” *Jurteksi*, vol. 4, no. 1, pp. 85–92, 2017, doi: 10.33330/jurteksi.v4i1.28.
- [6] A. Afrisawati and S. Sahren, “Analisis Perbandingan Menggunakan Metode Moora Dan Waspas Pemilihan Bibit Sapi Potong Terbaik,” *JURTEKSI (Jurnal Teknol. dan Sist. Informasi)*, vol. 6, no. 3, pp. 269–276, 2020, doi: 10.33330/jurteksi.v6i3.827.