

PENERAPAN *MACHINE LEARNING* DALAM PENGELOMPOKKAN WILAYAH RAWAN TINDAK KEJAHATAN DI KOTA BATAM MENGGUNAKAN ALGORITMA *K-MEANS*

Mega Christy Aritonang¹,
Alfannisa Annurrullah Fajrin.²

¹Mahasiswa Program Studi Teknik Informatika, Universitas Putera Batam

²Dosen Program Studi Teknik Informatika, Universitas Putera Batam

email: pb190210073@upbatam.ac.id

ABSTRACT

Issues of public security are becoming more complex in this day and age of rapid technical advancement due to the rise of societal problems including deviant behavior and social behavior as well as the collapse in moral nationalism. This has contributed to a troubling rise in crime in a number of places, including Batam City. Although Batam City is growing quickly, crime is also rising, particularly in areas that are seen to be vulnerable. Crimes including motor vehicle theft, violent theft, and aggravated theft are on the rise, according to police data. Purpose of this research is to help law enforcement Batam City take proactive measures to lower crime rates by applying machine learning techniques utilizing the k-means algorithm.

Keywords: algorithm, clustering, machine learning, k-means

PENDAHULUAN

Wilayah Kota Batam salah satu kota yang mempunyai pertumbuhan pesat dalam sebagian tahun terakhir. Tetapi pertumbuhan tersebut pula diiringi dengan meningkatnya tindak kejahatan, paling utama diwilayah- wilayah tertentu yang dikira rawan. Bagi informasi Kepolisian tindak kejahatan yang cenderung bertambah. Sebagian tindak pidana berupa pencurian dengan pemberatan (curat), pencurian dengan kekerasan (curas), serta pencurian kendaraan bermotor (curanmor). Pasti saja berat untuk pihak Kepolisian, bahwasannya bukan cuma pelakon tindak kejahatan. Hingga periset mengenali sebagian kasus ialah analisa

kenaikan keamanan, kenyamanan, serta kedisiplinan warga yang terjalin di daerah rawan terbentuknya tindak kejahatan masih belum optimal, serta belum terdapatnya wujud analisa informasi tindak kejahatan di daerah rawan terbentuknya atau terjadinya tindak pidana pencurian dengan pemberatan (curat), pencurian dengan kekerasan (curas), serta pencurian kendaraan bermotor (curanmor), sehingga dengan terdapatnya analisis dalam pengelompokkan kemampuan tingkatan kerawanan di daerah Kota Batam bisa tergambar serta teranalisa dengan hasil dalam wujud grafik. Bisa diformulasikan permasalahan dalam riset ini ialah bagaimana metode analisis data mining memakai algoritma K-Means

menganalisis pengelompokan daerah rawan tindak kejahatan di Kota Batam?, gimana hasil pengujian analisis dengan algoritma *K-Means* dengan aplikasi RapidMiner?. Tujuan dari riset ini merupakan buat menganalisis pengelompokan daerah rawan tindak kejahatan di Kota Batam dengan algoritma *K-Means*, serta buat menguji hasil analisis pengelompokan daerah rawan tindak kejahatan dengan aplikasi RapidMiner.

KAJIAN TEORI

2. 1 Knowledge Discovery in Database (KDD)

Knowledge Discovery in Database (KDD) ialah proses yang terdiri dari pengumpulan, pemaknaan informasi, historis buat menciptakan keteraturan, pola ataupun ikatan dalam set informasi berdimensi besar Knowledge Discovery in Database (KDD) didefinisikan selaku ekstraksi data potensial, implisit serta tidak diketahui dari sekumpulan informasi. Proses Knowledge Discovery in Database mengaitkan hasil proses informasi mining (proses pengekstrak kecenderungan sesuatu pola informasi setelah itu mengganti hasilnya secara akurat jadi data yang gampang dimengerti (P. P. Putra & Camp; Chan, 2018).

2. 2 Informasi Mining

Bagi (Nabila et.al, 2021) informasi data mining ialah metode bantu pengambilan keputusan pada penulis mencari pola dalam informasi. Ini bisa dicoba secara manual oleh pengguna (lewat kueri, misalnya) ataupun secara otomatis oleh fitur lunak yang menanyakan basis informasi buat mencari model informasi. Temuan mengacu pada proses mencari data baru. Pemilihan tugas informasi mining, pemilihan goal

dari proses KDD misalnya klasifikasi, regresi, clustering. Pemilihan algoritma informasi mining buat pengelompokan (clustering). Bagi (Aditya et., al, 2020) clustering informasi mining jadi nama lain buat 10 tata cara segmentation. Tujuan dari segmentasi pada tata cara informasi mining merupakan mengelompokkan sesuatu class ke dalam sebagian segmen bersumber pada atribut yang ditetapkan. Penentuan atribut wajib cocok kesamaan yang dipunyai sebagian class tadi. *K-means* ialah algoritma clustering, salah satu "unsupervised machine learning algorithms" sangat terkenal.

2. 3 Clustering

Bagi (Aditya et., al, 2020) clustering jadi nama lain buat 10 tata cara segmentation. Tujuan dari segmentasi pada tata cara data mining ialah mengelompokkan sesuatu class ke dalam sebagian segmen bersumber pada atribut yang ditetapkan. Penentuan atribut wajib cocok kesamaan yang dipunyai sebagian class tadi. *K-means* ialah algoritma clustering. *K-means Clustering* ialah salah satu "unsupervised machine learning algorithms" yang sangat simpel serta terkenal. *K-Means clustering* jadi sesuatu tata cara penganalisaan informasi ataupun tata cara informasi mining yang melaksanakan proses pemodelan tanpa supervise (unsupervised) serta ialah salah satu tata cara yang melaksanakan pengelompokan informasi dengan sistemn partisi.

2. 4 Rapidminer

Bagi (Zaki Muhammad et., al, 2018) RapidMiner ialah sesuatu platforms fitur lunak informasi ilmu pengetahuan yang dibesarkan oleh Industri dengan nama yang sama, yang sediakan area terpadu buat pendidikan mesin (machine learning), pendidikan mendalam (deep

learning, penambangan bacaan(text mining) serta analisis prediktif(predictive analytics).

2. 5 Tindak Kejahatan

Kejahatan ialah perbuatan serta aksi jahat semacam yang umum orang tahu ataupun dengarkan. Kejahatan dimaksudkan selaku sesuatu sikap yang berlawanan dengan nilai serta norma yang sudah disahkan oleh hukum tertulis. Kejahatan dilihat dari sudut pandang pendekatan sah dimaksud selaku sesuatu perbuatan yang melanggar hukum pidana ataupun UndangUndang yang berlaku di Warga. Pada hakikatnya sesuatu perbuatan melanggar hukum pidana ataupun Undang- Undang yang berlaku dalam sesuatu warga merupakan sesuatu perbuatan yang sangat merugikan yang bersangkutan.

METODE PENELITIAN

Desain penelitian secara teknis menggambarkan apa yang peneliti lakukan. Metodologi penelitian ini memiliki beberapa pedoman yang harus diikuti dan urutan peroman tersebut merupakan contoh prosedur harus diikuti untuk memastikan keberhasilan penelitian ini.

3.1 Desain Penelitian



Gambar 1. Desain Penelitian (Sumber: Data Penelitian, 2024)

3. 2 Pengumpulan Data

Pengumpulan informasi ialah salah satu aktivitas yang sangat berarti pada riset ini, buat itu hingga penulis melaksanakan Metode pengumpulan informasi berikut:

1. Observasi

Metode ini ialah metode pengumpulan informasi bersumber pada observasi area. Pada titik ini, penulis melaksanakan observasi terhadap informasi daerah terbentuknya tindak kejahatan di Kota Batm ialah objek riset buat memperoleh data bersumber pada kenyataan di Ditreskrim Polda Kepulauan Riau.

2. Wawancara

Metode ini ialah metode langsung dengan menghubungi pihak yang berkaitan dengan objek serta kasus yang diteliti setelah itu hendak melaksanakan wawancara kepada pihak Ditreskrim Polda Kepulauan Riau ialah Ayah AKBP Arthur Sitindaon, S. H., Meter. H., Bunda AKBP Adolfien C. Tuerah, S. H., Meter. H. serta Ayah Debora Leonardo, S. E.

3. Metode Riset Literatur

Metode ini ialah perdebatan dalam bidang riset tertentu. Penulis pula memperoleh cerminan pendek tentang apa yang dapat dipelajari, penalaran serta pemecahan permasalahan lewat riset permasalahan. Oleh sebab itu dalam riset ini penulis mencari bermacam data pendukung dari harian. Penulis mengkaji serta menarik kesimpulan yang bisa menolong dalam riset.

3. 3 Operasional Variabel

Ada pula operasional variable yang hendak digunakan dalam riset ini merupakan:

1. Jenis Tindak Pidana

Berdasarkan data yang diperoleh dari Ditreskrim Polda Kepulauan Riau tindak kejahatan yang cenderung meningkat seperti tindak pidana pencurian dengan pemberatan (curat), pencurian dengan kekerasan (curas), dan pencurian kendaraan bermotor (curanmor) terjadi di wilayah Kota Batam.

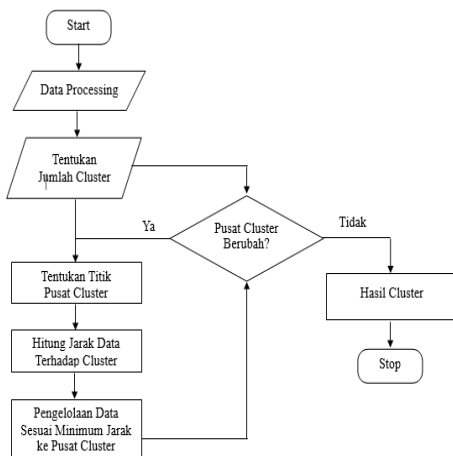
2. Tingkat Kesulitan Penanganan Tindak Pidana

Tingkat kesulitan penanganan tindak pidana dapat dipengaruhi oleh berbagai faktor yang berkaitan dengan karakteristik kasus dan kondisi di lapangan. Adapun penjelasan keterangan pada tabel yaitu “lengkap” bahwa penanganan perkara telah selesai dan yang “berjalan” dinyatakan penanganan perkara masih dalam proses penyelidikan maupun penyidikan.

3. Wilayah terjadinya tindak pidana

Wilayah terjadinya tindak pidana pencurian dapat sangat bervariasi dan seringkali mempengaruhi tingkat kesulitan dalam penanganannya. Setiap jenis wilayah memiliki karakteristik khusus yang dapat mempengaruhi cara pencurian dilakukan, jenis barang yang dicuri, serta metode yang digunakan oleh pelaku.

3.4 Flowchart Algoritma K-means



Gambar 2. Flowchart Algoritma K-means (Sumber: Data Penelitian, 2024)

3.5 Pengkodean Data

Pengkodean dalam clustering penting untuk mengoptimalkan representasi data agar algoritma clustering dapat mengenali pola atau kelompok lebih baik. Berikut pengkodean data yang digunakan dalam penelitian:

Tabel 1 Kategori Tindak Pidana

TINDAK PIDANA	KODE
CURAT	CT
CURAS	CS
CURANMOR	CR

Sumber: (Data Penelitian, 2024)

Pada tabel 3.4 pada pengkodean kategori tindak pidana maka curat menjadi CT, curas menjadi CS dan curanmor menjadi CR.

Tabel 2 Kategori Penanganan Perkara

TINDAK PIDANA	KODE
SEDANG	SD
SULIT	ST
SANGAT SULIT	SS

Sumber: (Data Penelitian, 2024)

Pada tabel 3.5 pada pengkodean kategori penanganan perkara maka sedang menjadi SD, sulit menjadi ST dan Sangat Sulit menjadi SS.

Tabel 3 Kategori Keterangan Berkas Perkara

TINDAK PIDANA	KODE
Lengkap	LK
Berjalan	BJ

Sumber: (Data Penelitian, 2024)

Pada tabel 3.6 pada pengkodean kategori keterangan berkas perkara maka lengkap menjadi LK dan berjalan menjadi BJ.

HASIL DAN PEMBAHASAN

4.1 Perhitungan Algoritma Clustering K-means

1. Menentukan centroid

Pusat awal cluster atau centroid ditentukan random atau acak. Peneliti menentukan nilai diambil dari data tindak pidana. Berikut adalah nilai centroid awal pada segmentasi pengelompokan data wilayah berdasarkan tindak pidana:

a. Kategori tindak pidana

Tabel 4. Tabel Centroid Tindak Pidana

TKP	CT	CS	CR	CLUSTER
BENKONG	5	11	22	C0
SAGULUNG	1	6	6	C1
NONGSA	10	7	5	C2

Sumber: (Data Penelitian, 2024)

Jarak centroid :

$$\text{cluster 0} = \sqrt{(10-5)^2 + (7-11)^2 + (5-22)^2} = 18$$

$$\text{cluster 1} = \sqrt{(10-1)^2 + (7-6)^2 + (5-22)^2} = 9$$

$$\text{cluster 2} = \sqrt{(10-10)^2 + (7-7)^2 + (5-5)^2} = 0$$

b. Penanganan Perkara

Tabel 5. Tabel Centroid Penanganan Perkara

TKP	SS	ST	SD	CLUSTER
BENKONG	1	22	15	C0
SAGULUNG	0	5	8	C1
NONGSA	0	19	2	C2

Sumber: (Data Penelitian, 2024)

Jarak centroid:

$$\text{cluster 0} = \sqrt{(0-1)^2 + (19-22)^2 + (2-15)^2} = 13$$

$$\text{cluster 1} = \sqrt{(0-0)^2 + (19-5)^2 + (2-8)^2} = 15$$

$$\text{cluster 2} = \sqrt{(0-0)^2 + (19-19)^2 + (2-2)^2} = 0$$

c. Penanganan Perkara

Tabel 6. Tabel centroid tindak pidana

TKP	LK	BJ	CLUSTER
BENKONG	10	5	C0
SAGULUNG	10	3	C1
NONGSA	15	1	C2

Sumber: (Data Penelitian, 2024)

Jarak centroid:

$$\text{cluster 0} = \sqrt{(15-28)^2 + (9-16)^2} = 15$$

$$\text{cluster 1} = \sqrt{(15-8)^2 + (9-5)^2} = 17$$

$$\text{cluster 2} = \sqrt{(15-15)^2 + (9-9)^2} = 0$$

4.2 Hasil Perhitungan Manual dan Rapidminer

Perbandingan	Perhitungan Manual			Perhitungan Rapidminer		
Jumlah Iterasi	3			-		
Jumlah Cluster	C0	C1	C2	C0	C1	C2
	11	5	4	2	2	4

Tabel 7 Perbandingan Hasil

Sumber: (Data Penelitian, 2024)

Pada tabel 4.21 menjelaskan perbandingan hasil antara perhitungan

manual, perhitungan menggunakan *tools Rapidminer*, pada perhitungan manual terdapat jumlah iterasi sebanyak 3 kali iterasi, pada perhitungan *Rapidminer* tidak ditampilkan berapa jumlah iterasi, jumlah perhitungan manual pada cluster adalah C0 terdapat 11 data, C1 terdapat 5 data, dan C2 terdapat 4 data. Jumlah cluster perhitungan *Rapidminer* adalah C0 terdapat 2 data, C1 terdapat 2 data dan C2 terdapat 4 data.

SIMPULAN

Mengelompokkan informasi dengan algoritma k-means dicoba dengan metode memastikan jumlah cluster, hitung jarak terdekat dengan pusat cluster. Informasi dengan jarak terdekat melaporkan daerah terbentuknya tindak kejahatan dari cluster tersebut, dicoba perhitungan kembali hingga informasi tidak berpindah pada cluster lain, buat meminimalkan guna objektif. Bersumber pada hasil pengujian informasi memakai *Rapidminer* pada riset informasi daerah terbentuknya tindak kejahatan di Kota Batam pada tahun 2022-2023, menciptakan pengelompokkan jumlah perhitungan manual pada cluster merupakan C0 ada 11 informasi, C1 ada 5 informasi, serta C2 ada 4 informasi. Jumlah cluster perhitungan *Rapidminer* merupakan C0 ada 2 informasi, C1 ada 2 informasi serta C2 ada 4 informasi.

DAFTAR PUSTAKA

- Aditya, A., Jovian, I., & Sari, B. N. (2020). Implementasi K-Means Clustering Ujian Nasional Sekolah Menengah Pertama di Indonesia Tahun 2018/2019. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 4(1), 51–58.
- Putra, P. P., & Chan, A. S. (2018). Pengembangan Aplikasi Perhitungan Prediksi Stock Motor Menggunakan Algoritma C 4.5 Sebagai Bagian dari Sistem Pengambilan Keputusan (Studi Kasus di Saudara Motor). *JURNAL INOVTEK POLBENG -SERI INFORMATIKA*, 3(1), 24–33.
- Putra, R. R., & Wadisman, C. (2018). Implementasi Data Mining Pemilihan Pelanggan Potensial Menggunakan Algoritma K Means. *INTECOMS: Journal of Information Technology and Computer Science*, 1(1), 72–77.
- Nabila, Z., Rahman Isnain, A., Permata, & Abidin, Z. (2021). ANALISIS DATA MINING UNTUK CLUSTERING KASUS COVID-19 DI PROVINSI LAMPUNG DENGAN ALGORITMA K-MEANS. *Jurnal Teknologi Dan Sistem Informasi (JTSI)*, 2(2), 100.
- Zaki Muhammad, Rifqi Rahmadhani, Hafid Rizqifaluthi, & Muhammad Ainul Yaqin. (2018). *Process Mining Akademik Sekolah Menggunakan Rapid Miner*. *MATICS : Jurnal Ilmu Komputer Dan Teknologi Informasi*, 10(2), 47–51.
- Indraputra, R. A., Fitriana, R., & others. (2020). K-Means Clustering Data COVID-19. *Jurnal Teknik Industri*, 10(3), 275–282.

	<p>Biodata Penulis pertama, Mega Christy Aritonang merupakan mahasiswa Prodi Teknik Informatika Universitas Putera Batam.</p>
	<p>Biodata Penulis kedua, Alfannisa Annurrullah Fajrin merupakan dosen program studi Teknik Informatika, Universitas Putera Batam, penulis banyak berkecimpung di bidang Teknik Informatika.</p>