

Algoritma K-Means dalam Mengelompokkan Daerah Rawan Kriminalitas Pencurian Sepeda Motor Di Kabupaten Asahan

Mustopa Husein Lubis¹, Nopi Purnomo², Fery Irawan³

^{1,2,3}Institut Teknologi dan Sains Paluta
mustopahusein1302@gmail.com
nopipurnomo2020@gmail.com
Fery.irawan1922@gmail.com

Abstrak

Kejahatan merupakan tindakan dan perbuatan yang merugikan secara ekonomis dan psikologis yang melanggar hukum yang berlaku dalam Negara Indonesia serta norma-norma sosial dan agama. Adapun yang menjadi lokasi penelitian adalah Polres Asahan yang merupakan instansi yang dapat memberikan keamanan dan perlindungan bagi masyarakat khususnya yang berada di Kabupaten Asahan. Masalah yang terjadi sulitnya Polres Asahan mengelompokkan daerah yang rawan tindak kriminalitas dalam kategori paling rawan, cukup rawan dan tidak rawan. Berdasarkan latar belakang masalah di atas, peneliti tertarik untuk mengambil penelitian dalam membuat sistem untuk meramalkan tingkat kriminalitas di Kabupaten Asahan guna mengantisipasi lonjakan kriminalitas yang akan datang. Metode yang digunakan adalah Algoritma *K-Means Clustering* sebagai salah satu metode data *clustering* non-hirarki mempartisi data yang ada ke dalam bentuk satu atau lebih *cluster* atau kelompok.. Data yang digunakan data kasus tindak kriminalitas pencurian kendaraan bermotor sebanyak 5 tahun terakhir pada tahun 2016 sampai 2020. Hasil dari pengujian didapat bahwa Kecamatan Kisaran Timur yang menjadi daerah rawan tindak kriminalitas pencurian kendaraan bermotor.

Kata Kunci: *clustering*, *data mining*, kriminalitas, Pencurian Kendaraan bermotor, *k-means*

© 2021 JSisfotek

1. Pendahuluan

Kejahatan adalah suatu tindakan yang dilarang oleh hukum publik untuk melindungi masyarakat dan diberikan hukuman oleh negara (Atmaja,dkk.2018). Masalah yang terjadi sulitnya Polres Asahan mengelompokkan daerah yang rawan tindak kriminalitas dalam kategori paling rawan, cukup rawan dan tidak rawan. Apalagi mengingat kondisi Asahan terdapat empat pulau besar terdiri dari 10 kecamatan, di mana kriminalitas semakin meningkat di setiap tahun khususnya yang berada pada daerah Kabupaten Asahan seperti pencurian kendaraan bermotor. Berdasarkan permasalahan tersebut maka digunakan data mining dalam mengelompokkan daerah rawan tindak kriminalitas di daerah Kabupaten Asahan

Knowledge Discovery in Database (KDD) atau sering disebut sebagai data mining dimana data ini berhubungan dengan metode pengintegrasian dan penemuan secara ilmiah, penjelasan yang memuat data dan untuk menggambarkan pola dari sejumlah kumpulan data. Data mining sendiri adalah bagian dari tahapan proses *Knowledge Discovery in Database* (KDD) (Kurniawan,dkk.2020).

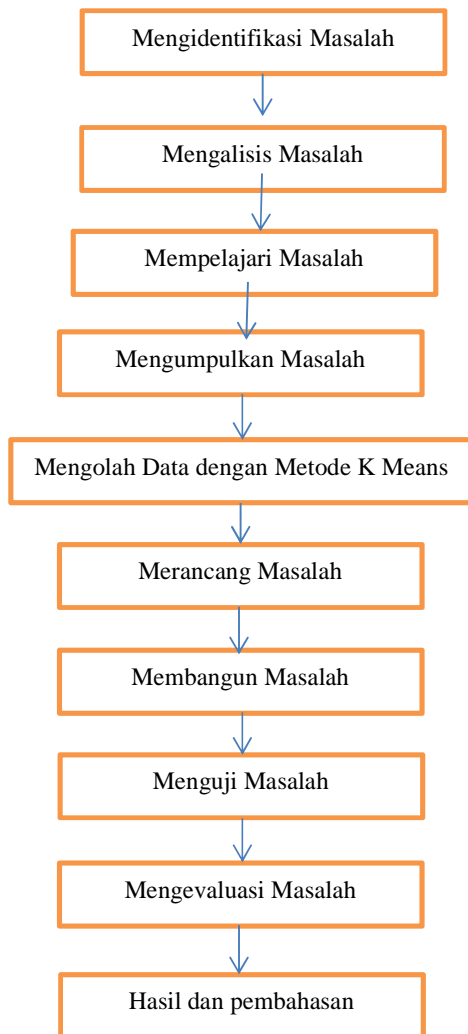
Algoritma *K-Means Clustering* sebagai salah satu metode data *clustering* non-hirarki mempartisi data yang ada ke dalam bentuk satu atau lebih *cluster* atau kelompok. *Clustering* merupakan salah satu teknik data mining yang digunakan untuk mendapatkan kelompok-kelompok dari obyek-obyek yang mempunyai karakteristik yang umum didata yang cukup besar (Suryadi & Supriatna, 2019). *Clustering* juga merupakan salah satu teknik klasifikasi data mining tanpa pengawasan yang paling umum untuk memecah objek menjadi satu set kelompok yang bermakna (Alibuhntoa & Mahatb, 2020). Tujuannya menemukan *cluster* yang berkualitas dalam waktu yang layak. *Clustering* dalam *data mining* berguna untuk menemukan pola distribusi di dalam sebuah data set yang berguna untuk proses analisa data. Analisis *clustering* dapat diterapkan secara luas pada gambar pemrosesan, pengenalan model, pengambilan dokumen, diagnosis medis, analisis web (Kumar et al, 2019).

Pada penelitian sebelumnya dengan menggunakan metode Algoritma *K-Means Clustering*. Singh et al (2019) penelitian di bidang kesehatan di mana sistem yang dibuat dengan metode ini dapat membantu dalam mengusulkan pendekatan hibrida cerdas untuk diagnosis penyakit hepatitis. Kim et al

(2021) penelitian di bidang bisnis di mana sistem yang dibuat dengan metode ini dapat memprediksi dan mengevaluasi efektivitas biaya pengenalan dan pengoperasian BIM. Febriyati et al (2020) penelitian di bidang perekonomian di mana sistem yang dibuat dapat membantu dan mempermudah dalam mengelompokkan laju pertumbuhan menurut bidang usaha di Kota Surabaya, sehingga dapat diketahui sektor mana saja yang mengalami pertumbuhan tinggi atau rendah. Westari & Halim (2021) penelitian di bidang kesehatan di mana sistem yang dapat mengklasifikasikan pasien diabetes dan non diabetes dengan menghitung jarak antara data latih dan data uji.

2. Metode Penelitian

Penelitian ini menggunakan Kerangka Kerja sebagai berikut :



Gambar 1. Kerangka Kerja Penelitian

1. Mengidentifikasi Masalah
Tahapan ini menentukan rumusan masalah yang terjadi pada Polres Asahan Sumatera Utara dalam menentukan tindak kriminalitas di Kabupaten Asahan dimasa yang akan datang.
2. Menganalisa Masalah
Pada analisa masalah ini digambarkan proses untuk pengelompokkan daerah tindak kriminalitas di Kabupaten Asahan berdasarkan data tindak kriminalitas sebelumnya..
3. Mempelajari Literatur
Literatur diambil dari berbagai sumber yaitu berupa artikel, jurnal ilmiah tentang metode *K-Means Clustering*, serta bahan bacaan lain yang mendukung.
4. Mengumpulkan Data
Data yang dibutuhkan dalam penelitian ini adalah data tindak kriminalitas pencurian sepeda motor di Kabupaten Asahan. Data yang diperlukan adalah data dalam 5 tahun terakhir dari tahun 2016 sampai dengan tahun 2020.
5. Mengolah Data Dengan Metode *K-Means Clustering*
Dengan melakukan pengolahan terhadap data tindak kriminalitas pencurian sepeda motor 5 tahun terakhir dari tahun 2016 sampai dengan tahun 2020 yang diperoleh dari Polres Kabupaten Asahan, selanjutnya ke tahapan proses analisa Metode *Algoritma K-Means*.
6. Merancang Sistem
Tahap ini dimulai dari dari pembuatan basis data, desain antar muka masukan, pembuatan algoritma, dan desain antar muka keluaran.
7. Membangun sistem
Pada tahap ini untuk melakukan pembangunan sistem tindak kriminalitas pencurian sepeda motor di Kabupaten Asahan yang digunakan bahasa pemrograman *PHP* dan *database Mysql*.
8. Menguji Sistem
Pada tahap ini dilakukan pengujian dan implementasi terhadap data tindak kriminalitas pencurian sepeda motor di Kabupaten Asahan yang telah diolah dengan bahasa pemrograman *PHP*.
9. Mengevaluasi Hasil
Pada tahap ini akan diuraikan hasil dari pengolahan dan pengujian data yang telah dilakukan dengan menggunakan metode *K-Means Clustering*. Hasil dari metode ini akan dilihat tindak keakuratannya.

10. Hasil dan Pembahasan

Pada tahapan ini akan dilakukan analisa data yang dihasilkan dari metode *K-Means Clustering* sehingga akan didapatkan hasil tindak kriminalitas untuk pengelompokan daerah tindak kriminalitas dengan cepat dan tepat.

Dalam melakukan penelitian, peneliti harus memperhatikan dan mengetahui metode apa yang akan digunakan dalam penelitiannya. Dalam penelitian ini metode yang digunakan penulis adalah metode kualitatif.

Penelitian kualitatif biasanya mengandalkan teknik observasi terlibat dan tak terkendali, juga wawancara bebas dan mendalam. Jenis data yang dikumpulkan bersifat monografis atau berwujud kasus-kasus. Data-data tersebut dianalisis secara kualitatif.

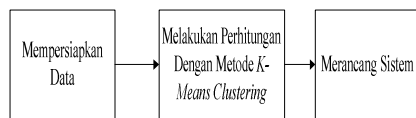
Didalam proses penelitian ini, yang dimaksud dari metode kualitatif yaitu pada masalah yang sesuai dengan fenomena dan menggunakan data. Data tersebut diambil langsung dari Kepolisian Resor (Polres) Asahan.

Adapun teknik pengumpulan data dalam penelitian ini adalah :

1. Observasi
Mengadakan penelitian langsung di Kepolisian Resor (Polres) Asahan sehingga didapat gambaran permasalahan yang ada pada Kepolisian Resor (Polres) tersebut.
2. Interview
Pengumpulan data yang dilakukan dengan bertanya langsung kepada petugas untuk mendapatkan data dan informasi yang diperlukan agar data yang didapat benar-benar akurat.
3. Studi Pustaka
Mempelajari dan mengumpulkan materi-materi yang berhubungan dengan sistem pengelolaan data daerah tindak kriminalitas pencurian sepeda motordi Kepolisian Resor (Polres) Asahan.

3. Hasil dan Pembahasan

Berdasarkan pembahasan yang telah dibuat pada bab-bab sebelumnya dapat dibentuk sebuah bagan alir dari sistem yang akan dirancang seperti Gambar 2.



Gambar 2. Bagan Alir Analisa dan Perancangan

Sebelum melakukan perancangan terhadap sistem, perlu dilakukan analisa terhadap data yang akan digunakan pada penelitian ini dengan menggunakan metode *K-Means Clustering*. Setelah selesai melakukan analisa selanjutnya akan dirancang sebuah sistem yang terkomputerisasi dengan menggunakan

metode yang sama. Analisa dan perancangan sistem tersebut selanjutnya akan dibahas secara terperinci pada sub bab di bawah ini.

3.1 Mempersiapkan Data

Dalam penelitian ini, data utama yang digunakan adalah data tahun 2016, 2017, 2018, 2019, dan 2020. Data tindak kriminalitas pencurian kendaraan motor dari tahun 2016-2020 dapat dilihat pada Tabel 1.

Tabel 1 Data Kasus Tindak Kriminalitas Pencurian Kendaraan Bermotor Tahun 2016-2020

No	Kecamatan	2016	2017	2018	2019	2020
1.	Kecamatan Kisaran Timur	6	11	19	10	7
2.	Kecamatan Kisaran Barat	3	7	15	10	5
3.	Kecamatan Meranti	5	9	6	7	5
4.	Kecamatan Rawang	1	2	5	5	3
5.	Kecamatan Panca Arga Aek Songsongan	1	2	7	5	4
6.	Kecamatan Simpang Empat	2	2	2	3	2
7.	Kecamatan Air Batu	4	8	11	15	12
8.	Kecamatan Air Joman	1	1	3	7	6
9.	Kecamatan Setia Janji	1	3	7	5	3
10.	Kecamatan BP. Mandoge	2	8	5	8	20

3.2 Melakukan Perhitungan Dengan Metode *K-Means*

Tahapan untuk menggunakan Clustering dengan metode *K-Means* sebagai berikut:

1. Memilih secara acak k buah data sebagai pusat pengelompokan (*cluster*)
Untuk menghitung jarak semua data ke setiap titik pusat *cluster* penelitian ini menggunakan teori jarak *Euclidean* berupa rumus berikut :

$$D_{(i,j)} = \sqrt{(X_{1i} - X_{1j})^2 + (X_{2i} - X_{2j})^2 + \dots + (X_{xi} - X_{xj})^2} \quad (!)$$

bahwa $D_{(i,j)}$ merupakan jarak data ke "I" ke pusat *cluster* "j", X_{ki} adalah data ke "I" pada atribut data ke "k", X_{kj} adalah titik pusat ke "j" pada atribut ke "k"

2. Penempatan data dalam *cluster* yang terdekat, dengan cara dihitung dari tengah *cluster*.

3. Pusat dari *cluster* baru ditentukan jika semua data sudah ditetapkan dalam *cluster* terdekat. Rumus penghitungan titik pusat *cluster* baru dengan cara berikut :

$$v = \sum_{i=1}^n X_i \quad ; i = 1,2,3, \dots, n \quad (2)$$

bahwa “v” merupakan *centroid* pada *cluster*, X_i adalah objek ke “i”, N merupakan banyaknya objek ataupun jumlah objek yang menjadi anggota dari *cluster*.

4. Teknik dalam menentukan *center* (pusat) *cluster* serta untuk menempatkan data dalam sebuah *cluster* dilakukan secara berulang kali sampai nilai *centroid* tidak lagi mengalami perubahan

Berdasarkan langkah-langkah dari metode *K-Means* di atas, selanjutnya dibentuklah *pseudocode* seperti yang terlihat pada Tabel 2.

Tabel 2 Pseudocode Algoritma K-Means

Algoritma 1. K-Means Clustering

Input : X,k,n,s

Output : C

Procedure

C ← 0;

For i ← 1 to n

Si ← RandomSample(x,s)

(ai,ci) ← kmeans(Si,k)

C ← AppendRows(C,ci)

End for

Return Cfinal ← kmeans(C,k)

End procedure

Disini akan dilakukan perhitungan dengan Metode *K-Means Clustering* sesuai dengan data kasus tindak kriminalitas pencurian kendaraan bermotor yang telah di buat dalam bentuk tabel yaitu Tabel 4.1 selanjutnya akan dilakukan proses perhitungan menggunakan Metode *K-Means Clustering*.

1) Menentukan Nilai *Cluster* Dari Tiap Data

Adapun pusat *Cluster* sebagai berikut :

Diambil data ke-1 sebagai pusat *Cluster*.

C1 =	6	11	19	10	7
------	---	----	----	----	---

Cluster 1 untuk *Cluster* 1 (Paling Rawan)

Diambil data ke-2 sebagai pusat *Cluster*.

C2 =	3	7	15	10	5
------	---	---	----	----	---

Cluster 2 untuk *Cluster* 2 (Cukup Rawan)

Diambil data ke-3 sebagai pusat *Cluster*.

C3 =	5	9	6	7	5
------	---	---	---	---	---

Cluster 3 untuk *Cluster* 3 (Tidak Rawan)

Iterasi-1

1. Menghitung jarak dari data ke “1” terhadap *center* (pusat) *Cluster* sebagai berikut :

$$C1 = \sqrt{(6-6)^2 + (11-11)^2 + (19-19)^2 + (10-10)^2 + (7-7)^2} = 0$$

$$C2 = \sqrt{(6-3)^2 + (11-7)^2 + (19-15)^2 + (10-10)^2 + (7-5)^2} = 5$$

$$C3 = \sqrt{(6-5)^2 + (11-9)^2 + (19-6)^2 + (10-7)^2 + (7-5)^2} = 13,67$$

2. Menghitung jarak dari data ke “2” terhadap *center* (pusat) *Cluster* sebagai berikut :

$$C1 = \sqrt{(3-6)^2 + (7-11)^2 + (15-19)^2 + (10-10)^2 + (5-7)^2} = 6,71$$

$$C2 = \sqrt{(3-3)^2 + (7-7)^2 + (15-15)^2 + (10-10)^2 + (5-5)^2} = 0$$

$$C3 = \sqrt{(3-5)^2 + (7-9)^2 + (15-6)^2 + (10-7)^2 + (5-5)^2} = 9,90$$

3. Menghitung jarak dari data ke “3” terhadap *center* (pusat) *Cluster* sebagai berikut :

$$C1 = \sqrt{(5-6)^2 + (9-11)^2 + (6-19)^2 + (7-10)^2 + (5-7)^2} = 13,67$$

$$C2 = \sqrt{(5-3)^2 + (9-7)^2 + (6-15)^2 + (7-10)^2 + (5-5)^2} = 9,90$$

$$C3 = \sqrt{(5-5)^2 + (9-9)^2 + (6-6)^2 + (7-7)^2 + (5-5)^2} = 0$$

4. Menghitung jarak dari data ke-4 terhadap pusat *Cluster*

$$C1 = \sqrt{(1-6)^2 + (2-11)^2 + (5-19)^2 + (5-10)^2 + (3-7)^2} = 18,52$$

$$C2 = \sqrt{(1-3)^2 + (2-7)^2 + (5-15)^2 + (5-10)^2 + (3-5)^2} = 12,57$$

$$C3 = \sqrt{(1-5)^2 + (2-9)^2 + (5-6)^2 + (5-7)^2 + (3-5)^2} = 8,60$$

5. Menghitung jarak dari data ke “5” terhadap *center* (pusat) *Cluster* sebagai berikut :

$$C1 = \sqrt{(1-6)^2 + (2-11)^2 + (7-19)^2 + (5-10)^2 + (4-7)^2} = 16,85$$

$$C2 = \sqrt{(1-3)^2 + (2-7)^2 + (7-15)^2 + (5-10)^2 + (4-5)^2} = 10,91$$

$$C3 = \sqrt{(1-5)^2 + (2-9)^2 + (7-6)^2 + (5-7)^2 + (4-5)^2} = 8,43$$

6. Menghitung jarak dari data ke “6” terhadap *center* (pusat) *Cluster* sebagai berikut :

$$C1 = \sqrt{(2-6)^2 + (2-11)^2 + (2-19)^2 + (3-10)^2 + (2-7)^2}$$

$$= 21,45$$

$$C2 = \frac{\sqrt{(2-3)^2 + (2-7)^2 + (2-15)^2 + (3-10)^2 + (2-5)^2}}{5} = 15,91$$

$$C3 = \frac{\sqrt{(2-5)^2 + (2-9)^2 + (2-6)^2 + (3-7)^2 + (2-5)^2}}{5} = 9,95$$

7. Menghitung jarak dari data ke “7” terhadap *center* (pusat) *Cluster* sebagai berikut :

$$C1 = \frac{\sqrt{(4-6)^2 + (8-11)^2 + (11-19)^2 + (15-10)^2 + (12-7)^2}}{5} = 11,27$$

$$C2 = \frac{\sqrt{(4-3)^2 + (8-7)^2 + (11-15)^2 + (15-10)^2 + (12-5)^2}}{5} = 9,59$$

$$C3 = \frac{\sqrt{(4-5)^2 + (8-9)^2 + (11-6)^2 + (15-7)^2 + (12-5)^2}}{5} = 11,83$$

8. Menghitung jarak dari data ke “8” terhadap *center* (pusat) *Cluster* sebagai berikut :

$$C1 = \frac{\sqrt{(1-6)^2 + (1-11)^2 + (3-19)^2 + (7-10)^2 + (6-7)^2}}{5} = 19,77$$

$$C2 = \frac{\sqrt{(1-3)^2 + (1-7)^2 + (3-15)^2 + (7-10)^2 + (6-5)^2}}{5} = 13,93$$

$$C3 = \frac{\sqrt{(1-5)^2 + (1-9)^2 + (3-6)^2 + (7-7)^2 + (6-5)^2}}{5} = 9,49$$

9. Menghitung jarak dari data ke “9” terhadap *center* (pusat) *Cluster* sebagai berikut :

$$C1 = \frac{\sqrt{(1-6)^2 + (3-11)^2 + (7-19)^2 + (5-10)^2 + (3-7)^2}}{5} = 16,55$$

$$C2 = \frac{\sqrt{(1-3)^2 + (3-7)^2 + (7-15)^2 + (5-10)^2 + (3-5)^2}}{5} = 10,63$$

$$C3 = \frac{\sqrt{(1-5)^2 + (3-9)^2 + (7-6)^2 + (5-7)^2 + (3-5)^2}}{5} = 7,81$$

10. Menghitung jarak dari data ke “10” terhadap *center* (pusat) *Cluster* sebagai berikut :

$$C1 = \frac{\sqrt{(2-6)^2 + (8-11)^2 + (5-19)^2 + (8-10)^2 + (20-7)^2}}{5} = 19,47$$

$$C2 = \frac{\sqrt{(2-3)^2 + (8-7)^2 + (5-15)^2 + (8-10)^2 + (20-5)^2}}{5} = 18,28$$

$$C3 = \frac{\sqrt{(2-5)^2 + (8-9)^2 + (5-6)^2 + (8-7)^2 + (20-5)^2}}{5} = 15,10$$

Menentukan Centroid Baru :

$$C1 = \frac{6}{1}, \frac{11}{1}, \frac{19}{1}, \frac{10}{1}, \frac{7}{1} = 6, 11, 19, 10, 7$$

$$C2 = \frac{3+4}{2}, \frac{7+8}{2}, \frac{15+11}{2}, \frac{10+15}{2}, \frac{5+12}{2} = 3,5, 7,5, 13, 12,5, 18,5$$

$$\frac{5+1+1+2+1+1+2}{7}, \frac{9+2+2+2+1+3+8}{7}$$

$$C3 = \frac{6+5+7+2+3+7+5}{7}, \frac{7+5+5+3+7+5+8}{7}, \frac{5+3+4+2+6+2+20}{7} = 1,86, 3,86, 5, 5,71, 6,14$$

Iterasi-2

Proses iterasi-2 ini akan dilakukan perhitungan dengan *centroid* (Pusat *Cluster*) nya yang tidak lagi berdasarkan sampel sebelumnya namun dari *centroid* baru.

Tabel 2 Pusat *Cluster* Baru

C1	6	11	19	10	7
C2	3,50	7,50	13	12,50	8,50
C3	1,86	3,86	5	5,71	6,14

1. Menghitung jarak dari data ke “1” terhadap *center* (pusat) *Cluster* sebagai berikut :

$$C1 = \frac{\sqrt{(6-6)^2 + (11-11)^2 + (19-19)^2 + (10-10)^2 + (7-7)^2}}{5} = 0$$

$$C2 = \frac{\sqrt{(6-3,5)^2 + (11-7,5)^2 + (19-13)^2 + (10-12,5)^2 + (7-8,5)^2}}{5} = 7,94$$

$$C3 = \frac{\sqrt{(6-1,86)^2 + (11-3,86)^2 + (19-3)^2 + (10-5,71)^2 + (7-6,14)^2}}{5} = 16,83$$

2. Menghitung jarak dari data ke “2” terhadap *center* (pusat) *Cluster* sebagai berikut :

$$C1 = \frac{\sqrt{(3-6)^2 + (7-11)^2 + (15-19)^2 + (10-10)^2 + (5-7)^2}}{5} = 6,71$$

$$C2 = \frac{\sqrt{(3-3,5)^2 + (7-7,5)^2 + (15-13)^2 + (10-12,5)^2 + (5-8,5)^2}}{5} = 4,8$$

$$C3 = \frac{\sqrt{(3-1,86)^2 + (7-3,86)^2 + (15-5)^2 + (10-5,71)^2 + (5-6,14)^2}}{5} = 11,4$$

3. Menghitung jarak dari data ke “3” terhadap *center* (pusat) *Cluster* sebagai berikut :

$$C1 = \frac{\sqrt{(5-6)^2 + (9-11)^2 + (6-19)^2 + (7-10)^2 + (5-7)^2}}{5} = 13,67$$

$$C2 = \frac{\sqrt{(5-3,5)^2 + (9-7,5)^2 + (6-13)^2 + (7-12,5)^2 + (5-8,5)^2}}{5} = 9,80$$

$$C3 = \frac{\sqrt{(5-1,86)^2 + (9-3,86)^2 + (6-5)^2 + (7-5,71)^2 + (5-6,14)^2}}{5} = 6,35$$

4. Menghitung jarak dari data ke “4” terhadap *center* (pusat) *Cluster* sebagai berikut :

$$C1 = \frac{\sqrt{(1-6)^2 + (2-11)^2 + (5-19)^2 + (5-10)^2 + (3-7)^2}}{5} = 18,52$$

$$C2 = \frac{\sqrt{(1-3,5)^2 + (2-7,5)^2 + (5-13)^2 + (5-12,5)^2 + (3-8,5)^2}}{5}$$

$$C3 = \frac{13,67}{\sqrt{(1-1,86)^2 + (2-3,86)^2 + (5-5)^2 + (5-5,71)^2 + (3-6,14)^2}} = 3,82$$

5. Menghitung jarak dari data ke “5” terhadap *center* (pusat) *Cluster* sebagai berikut :

$$C1 = \frac{16,85}{\sqrt{(1-6)^2 + (2-11)^2 + (7-19)^2 + (5-10)^2 + (4-7)^2}} = 16,85$$

$$C2 = \frac{12,21}{\sqrt{(1-3,5)^2 + (2-7,5)^2 + (7-13)^2 + (5-12,5)^2 + (4-8,5)^2}} = 12,21$$

$$C3 = \frac{3,64}{\sqrt{(1-1,86)^2 + (2-3,86)^2 + (7-5)^2 + (5-3,71)^2 + (4-6,14)^2}} = 3,64$$

6. Menghitung jarak dari data ke “6” terhadap *center* (pusat) *Cluster* sebagai berikut :

$$C1 = \frac{21,45}{\sqrt{(2-6)^2 + (2-11)^2 + (2-19)^2 + (3-10)^2 + (2-7)^2}} = 21,45$$

$$C2 = \frac{16,91}{\sqrt{(2-3,5)^2 + (2-7,5)^2 + (2-13)^2 + (3-12,5)^2 + (2-8,5)^2}} = 16,91$$

$$C3 = \frac{6,08}{\sqrt{(2-1,86)^2 + (2-3,86)^2 + (2-5)^2 + (3-3,71)^2 + 6,14^2}} = 6,08$$

7. Menghitung jarak dari data ke “7” terhadap *center* (pusat) *Cluster* sebagai berikut :

$$C1 = \frac{11,27}{\sqrt{(4-6)^2 + (8-11)^2 + (11-19)^2 + (15-10)^2 + (12-7)^2}} = 11,27$$

$$C2 = \frac{4,80}{\sqrt{(4-3,5)^2 + (8-7,5)^2 + (11-13)^2 + (15-12,5)^2 + (12-8,5)^2}} = 4,80$$

$$C3 = \frac{13,35}{\sqrt{(4-1,86)^2 + (8-3,86)^2 + (11-5)^2 + (15-3,71)^2 + (12-6,14)^2}} = 13,35$$

8. Menghitung jarak dari data ke “8” terhadap *center* (pusat) *Cluster* sebagai berikut :

$$C1 = \frac{19,77}{\sqrt{(1-6)^2 + (1-11)^2 + (3-19)^2 + (7-10)^2 + (6-7)^2}} = 19,77$$

$$C2 = \frac{13,6}{\sqrt{(1-3,5)^2 + (1-7,5)^2 + (3-13)^2 + (7-12,5)^2 + (6-8,5)^2}} = 13,6$$

$$C3 = \frac{3,82}{\sqrt{(1-1,86)^2 + (1-3,86)^2 + (3-5)^2 + (7-3,71)^2 + (6-6,14)^2}} = 3,82$$

9. Menghitung jarak dari data ke “9” terhadap *center* (pusat) *Cluster* sebagai berikut :

$$C1 = \frac{16,55}{\sqrt{(1-6)^2 + (3-11)^2 + (7-19)^2 + (5-10)^2 + (3-7)^2}} = 16,55$$

$$C2 = \frac{12,21}{\sqrt{(1-3,5)^2 + (3-7,5)^2 + (7-13)^2 + (5-12,5)^2 + (3-8,5)^2}} = 12,21$$

$$C3 = \frac{3,98}{\sqrt{(1-1,86)^2 + (3-3,86)^2 + (7-5)^2 + (5-3,71)^2 + (3-6,14)^2}} = 3,98$$

10. Menghitung jarak dari data ke “10” terhadap *center* (pusat) *Cluster* sebagai berikut :

$$C1 = \frac{19,85}{\sqrt{(2-6)^2 + (8-11)^2 + (5-19)^2 + (8-10)^2 + (20-7)^2}} = 19,85$$

$$C2 = \frac{14,80}{\sqrt{(2-3,5)^2 + (8-7,5)^2 + (5-13)^2 + (8-12,5)^2 + (20-8,5)^2}} = 14,80$$

$$C3 = \frac{14,64}{\sqrt{(2-1,86)^2 + (8-3,86)^2 + (5-5)^2 + (8-3,71)^2 + (20-6,14)^2}} = 14,64$$

Pada pengelompokan iterasi ke-2 (G2) ini, data tidak ada yang berubah dan perhitungan diberhentikan sampai iterasi ke-2.

Berdasarkan hasil iterasi 1 dan 2 diatas, maka dapat ditemukan hasil sebagai berikut :

1. Untuk *Cluster* 1 (C1) /Paling Rawan: adalah 1 *Cluster* 1: Kecamatan Kisaran Timur
2. Untuk *Cluster* 2 (C2) /Cukup Rawan adalah 2,7 *Cluster* 2: Kecamatan Kisaran Barat dan Kecamatan Air Baru
3. Untuk *Cluster* 3 (C3)/Tidak Rawan: adalah 3,4,5,6,8,9,10 *Cluster* 3: Kecamatan Meranti, Kecamatan Rawang panca Arga, Kecamatan Aek Songosongan, Kecamatan Simpang Empat, Kecamatan Air Joman, Kecamatan BP Mandoge dan Kecamatan Setia Janiji

Kelompok data yang termasuk ke dalam *Cluster* 1, *Cluster* 2, dan *Cluster* 3 dapat dilihat pada Tabel 3.

Tabel 3 Hasil Pengujian

No	Kecamatan	Cluster	Keterangan
1.	Kecamatan Kisaran Timur	1	Paling Rawan
2.	Kecamatan Kisaran Barat	2	Cukup Rawan
3.	Kecamatan Meranti	3	Tidak Rawan
4.	Kecamatan Rawang Panca Arga	3	Tidak Rawan
5.	Kecamatan Aek Songosongan	3	Tidak Rawan
6.	Kecamatan Simpang Empat	3	Tidak Rawan
7.	Kecamatan Air Batu	2	Cukup Rawan
8.	Kecamatan Air Joman	3	Tidak Rawan
9.	Kecamatan Setia Janji	3	Tidak Rawan
10.	Kecamatan BP. Mandoge	3	Tidak Rawan

4. Kesimpulan

Setelah dilakukan penelitian *data mining* dengan metode *K-Means Clustering* untuk menentukan daerah rawan tindak kriminalitas di Asahan, maka dapat diambil kesimpulan bahwa :

1. Dari proses *K-Means Clustering* yang telah dilakukan, pengelompokan daerah rawan tindak kriminalitas pencurian sepeda motor di Asahan terdiri dari *cluster* 1 (paling rawan) yaitu 1 desa, *cluster* 2 (cukup rawan) yaitu 2 desa, dan *cluster* 3 (tidak rawan) yaitu 10 Kecamatan

2. Penerapan *data mining* dengan menggunakan metode *K-Means Clustering* dapat membantu dalam menentukan daerah rawan tindak kriminalitas pencurian sepeda mootrdi Asahan.

Sebuah penelitian tidak terlepas dari yang namanya kekurangan dan kelemahan. Setelah melakukan penelitian, penulis memberikan beberapa saran antara lain:

1. Penulis menyarankan adanya penelitian *data mining* dalam menentukan daerah rawan tindak kriminalitas di Asahan ini dengan membandingkan secara langsung beberapa metode *clustering*, guna menguji metode yang paling tepat digunakan.
2. Penulis mengharapkan sistem *data mining* dengan metode *K-Means Clustering* ini dapat diimplementasikan berupa aplikasi agar dapat dimanfaatkan langsung oleh Polres Kabupaten Asahan.

Gustientiedina, G., Adiya, M. H., & Desnelita, Y. (2019). Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan. *Jurnal Nasional Teknologi Dan Sistem Informasi*, 5(1), 17–24. <https://doi.org/10.25077/tek`nosi.v5i1.2019.17-24>

Haris Kurniawan, Sarjon Defit, & Sumijan. (2020). Data Mining Menggunakan Metode K-Means Clustering Untuk Menentukan Besaran Uang Kuliah Tunggal. *Journal of Applied Computer Science and Technology*, 1(2), 80–89. <https://doi.org/10.52158/jacost.v1i2.102>

Kim, H. S., Kim, S. K., & Kang, L. S. (2021). BIM performance assessment system using a K-means clustering algorithm. *Journal of Asian Architecture and Building Engineering*, 20(1), 78–87. <https://doi.org/10.1080/13467581.2020.1800471>

Kumar, S., Kumar-Solanki, V., Choudhary, S. K., Selamat, A., & Gonzalez-Crespo, R. (2020). Comparative Study on Ant Colony Optimization (ACO) and K-Means Clustering Approaches for Jobs Scheduling and Energy Optimization Model in Internet of Things (IoT). *International Journal of Interactive Multimedia and Artificial Intelligence*, 6(1), 107. <https://doi.org/10.9781/ijimai.2020.01.003>

Mahmudan, A. (2020). Clustering of District or City in Central Java Based COVID-19 Case Using K-Means Clustering. *Jurnal Matematika, Statistika Dan Komputasi*, 17(1), 1–13. <https://doi.org/10.20956/jmsk.v17i1.10727>

Purba, W., Tamba, S., & Saragih, J. (2018). The effect of mining data k-means clustering toward students profile model drop out potential. *Journal of Physics: Conference Series*, 1007(1). <https://doi.org/10.1088/1742-6596/1007/1/012049>

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. <https://doi.org/10.31539/intecom.v1i1.141>

Singh, A., Mehta, J. C., Anand, D., Nath, P., Pandey, B., & Khamparia, A. (2021). An intelligent hybrid approach for hepatitis disease diagnosis: Combining enhanced k-means clustering and improved ensemble learning. *Expert Systems*, 38(1). <https://doi.org/10.1111/exsy.12526>

Suriani, L. (2020). Pengelompokan Data Kriminal Pada

REFERENSI

Alibuhitto, M. C., & Mahat, N. I. (2020). Distance based k-means clustering algorithm for determining number of clusters for high dimensional data. *Decision Science Letters*, 9(1), 51–58. <https://doi.org/10.5267/j.dsl.2019.8.002>

Andri Nofiar Am, Sarjon Defit, S. (2019). Penentuan Mutu Kelapa Sawit Menggunakan Metode K-Means Clustering. *KomTekInfo*, 5(3), 1–9.

Ardiada, D., Ariawan, P. A., & Sudarma, M. (2018). Evaluation of Supporting Work Quality Using K-Means Algorithm. *IJEET International Journal of Engineering and Emerging Technology*, 3(1), 3–6.

Atmaja, E. H. S. (2019). Implementation of k-Medoids Clustering Algorithm to Cluster Crime Patterns in Yogyakarta. *International Journal of Applied Sciences and Smart Technologies*, 1(1), 33–44. <https://doi.org/10.24071/ijasst.v1i1.1859>

Fatmawati, K., & Windarto, A. P. (2018). Data Mining: Penerapan Rapidminer Dengan K-Means Cluster Pada Daerah Terjangkit Demam Berdarah Dengue (Dbd) Berdasarkan Provinsi. *Computer Engineering, Science and System Journal*, 3(2), 173. <https://doi.org/10.24114/cess.v3i2.9661>

Febriyati, N. A., Gs, A. D., & Wanto, A. (2020). GRDP Growth Rate Clustering in Surabaya City uses the K- Means Algorithm. *International Journal of Information System & Technology*, 3(2), 276–283.

- Poldasu Menentukan Pola Daerah Rawan Tindak Kriminal Menggunakan Data Mining Algoritma K-Means Clustering. *Jurnal Sistem Komputer Dan Informatika (JSON)*, 1(2), 151. <https://doi.org/10.30865/json.v1i2.1955>
- Suryadi, U. T., & Supriatna, Y. (2019). Sistem Clustering Tindak Kejahatan Pencurian Di Wilayah Jawa Barat Menggunakan Algoritma K-Means. *Jurnal Teknologi Dan Komunikasi STMIK Subang*, 12(1), 15–27. <https://doi.org/10.47561/a.v12i1.147>
- Westari, D. (2021). Performa Comparison of the K-Means Method for Classification in Diabetes Patients Using Two Normalization Methods. *International Journal of Multidisciplinary Research and Analysis*, 04(01), 18–23. <https://doi.org/10.47191/ijmra/v4-i1-03>