



Algoritma K-Means Clustering dalam Mengklasifikasi Data Daerah Rawan Tindak Kriminalitas (Polres Kepulauan Mentawai)

Yoni Aswan^{1✉}, Sarjon Defit², Gnadi Widi Nurcahyo³

¹Dinas Pendidikan Sumatera Barat

^{2,3}Universitas Putra Indonesia YPTK Padang

yoniaswan@gmail.com

Abstract

Crime is all kinds of actions and actions that are economically and psychologically harmful that violate the laws in force in the State of Indonesia as well as social and religious norms. Ordinary criminal acts affect the security of the community and threaten their inner and outer peace. The research location is the Mentawai Islands Police, which is an agency that can provide security and protection for the community, especially those in the Mentawai Islands Regency. The problem is that it is difficult for the Mentawai Islands Police to classify areas that are prone to crime in the most vulnerable, moderately vulnerable and not vulnerable categories. Especially considering the condition of the Mentawai, there are four large islands consisting of 10 sub-districts, where crime is increasing every year, especially those in the Mentawai Islands Regency area such as motor vehicle theft. Based on the background of the problem above, the researcher is interested in taking research in creating a system to predict the crime rate in the Mentawai Islands Regency in order to anticipate the surge in crime that will come. The method used is the K-Means Clustering Algorithm as a non-hierarchical data clustering method to partition existing data into one or more clusters or groups. This method partitions data into clusters so that data with the same characteristics are grouped into the same cluster and data with different characteristics are grouped into other clusters. Clustering is one of the data mining techniques used to get groups of objects that have common characteristics in large enough data. The data used is data on cases of criminal theft of motor vehicles for the last 5 years from 2016 to 2020. The results of the test show that South Sipora District is an area prone to the crime of motor vehicle theft.

Keywords: Clustering, Data Mining, Crime, Motor Vehicle Theft, K-Means.

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 Kepulauan Mentawai yang merupakan instansi yang dapat memberikan keamanan dan perlindungan bagi masyarakat khususnya yang berada di Kabupaten Kepulauan Mentawai. Masalah yang terjadi sulitnya Polres Kepulauan Mentawai mengelompokkan daerah yang rawan tindak kriminalitas dalam kategori paling rawan, cukup rawan dan tidak rawan. Apalagi mengingat kondisi mentawai terdapat empat pulau besar terdiri dari 10 kecamatan, di mana kriminalitas semakin meningkat di setiap tahun khususnya yang berada pada daerah Kabupaten Kepulauan Mentawai seperti pencurian kendaraan bermotor. Berdasarkan latar belakang masalah di atas, peneliti tertarik untuk mengambil penelitian dalam membuat sistem untuk meramalkan tingkat kriminalitas di Kabupaten Kepulauan Mentawai 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 Sipora Selatan 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 [1]. Masalah yang terjadi sulitnya Polres Kepulauan Mentawai mengelompokkan daerah yang rawan tindak kriminalitas dalam kategori paling rawan, cukup rawan dan tidak rawan. Apalagi mengingat kondisi mentawai terdapat empat pulau besar terdiri dari 10 kecamatan, di mana kriminalitas semakin meningkat di setiap tahun khususnya yang berada pada daerah Kabupaten Kepulauan Mentawai seperti pencurian kendaraan

bermotor. Berdasarkan permasalahan tersebut maka digunakan data mining dalam pengelompokkan daerah rawan tindak kriminalitas di daerah Kabupaten Kepulauan Mentawai.

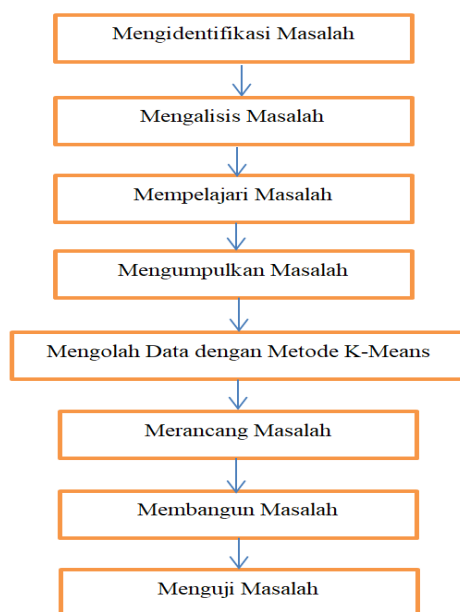
Knowledge Discovery in Database (KDD) atau sering disebut sebagai data mining dimana data ini berhubungan dengan metode pengintegrasian dan penemuan secara ilmiah, penjelasan yang memuat data dan untuk menggambarkan pola dari sejumlah kumpulan data. Data mining sendiri adalah bagian dari tahapan proses *Knowledge Discovery in Database* (KDD) [2].

Algoritma *K-Means Clustering* merupakan salah satu cara pengelompokan data yang mirip satu sama lain ataupun pengelompokan objek data yang berbeda dengan data lain dikelompokkan dalam satu kelompok data. [3][4]. Tujuan melakukan *Clustering* yaitu agar ditemukan objek-objek data yang mirip dalam satu kelompok data yang berkualitas, sehingga semakin besar perbedaan maupun kemiripan antar kelompok data dalam suatu cluster maka semakin baik dan berkualitas cluster tersebut [5].

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 [6]. 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 [7]. 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[8]. 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 [9]. Pengelompokan dengan menggunakan *K-Means*. Beberapa penelitian terdahulu dengan menerapkan *K-Means* dengan hasil yang baik dan tepat [10] - [17].

2. Metodologi Penelitian

Penelitian ini menggunakan kerangka kerja disajikan pada Gambar 1.



Gambar 1. Kerangka Kerja Penelitian

a. Mengidentifikasi Masalah

Tahapan ini menentukan rumusan masalah yang terjadi pada Polres Kepulauan Mentawai Sumatera Utara dalam menentukan tindak kriminalitas di Kabupaten Kepulauan Mentawai dimasa yang akan datang.

b. Menganalisa Masalah

Pada analisa masalah ini digambarkan proses untuk pengelompokan daerah tindak kriminalitas di Kabupaten Kepulauan Mentawai berdasarkan data tindak kriminalitas sebelumnya..

c. Mempelajari Literatur

Literatur diambil dari berbagai sumber yaitu berupa artikel, jurnal ilmiah tentang metode *K-Means Clustering*, serta bahan bacaan lain yang mendukung.

d. Mengumpulkan Data

Data yang dibutuhkan dalam penelitian ini adalah data tindak kriminalitas pencurian sepeda motor di Kabupaten Kepulauan Mentawai. Data yang diperlukan adalah data dalam 5 tahun terakhir dari tahun 2016 sampai dengan tahun 2020.

e. 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 Mentawai, selanjutnya ke tahapan proses analisa Metode Algoritma *K-Means*.

f. Merancang Sistem

Tahap ini dimulai dari dari pembuatan basis data, desain antar muka masukan, pembuatan algoritma, dan desain antar muka keluaran.

g. Membangun sistem

Pada tahap ini untuk melakukan pembangunan sistem tindak kriminalitas pencurian sepeda motor di Kabupaten Kepulauan Mentawai yang digunakan bahasa pemrograman *PHP* dan database *Mysql*.

h. Menguji Sistem

Pada tahap ini dilakukan pengujian dan implementasi terhadap data tindak kriminalitas pencurian sepeda motor di Kabupaten Kepulauan Mentawai yang telah diolah dengan bahasa pemrograman *PHP*.

i. 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.

3. Hasil dan Pembahasan

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 (Kasus)

No	Kecamatan	2016	2017	2018	2019	2020
1.	Sipora Selatan	6	11	19	10	7
2.	Siberut Utara	3	7	15	10	5
3.	Siberut Selatan	5	9	6	7	5
4.	Siberut Barat	1	2	5	5	3
5.	Siberut Tengah	1	2	7	5	4
6.	Siberut Barat Daya	2	2	2	3	2
7.	Sikakap	4	8	11	15	12
8.	Pagai Selatan	1	1	3	7	6
9.	Pagai Utara	1	3	7	5	3
10.	Sipora Utara	2	8	5	8	20

3.2 Melakukan Perhitungan Dengan Metode *K-Means*

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

3.2.1. 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*.

$$D_{(i,j)} = \sqrt{(X_{i1}-X_{j1})^2 + (X_{i2}-X_{j2})^2 + \dots + (X_{in}-X_{jn})^2} \quad (1)$$

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"

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

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

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

Dimana "v" merupakan *centroid* pada *cluster*, X_i adalah objek ke "I", N merupakan banyaknya objek ataupun jumlah objek yang menjadi anggota dari *cluster*.

3.2.4. 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 Algoritma 1.

Algoritma 1. *K-Means Clustering*

```

Input :  $X, k, n, s$ 
Output :  $C$ 
Procedure
 $C \leftarrow 0$ ;
For  $i \leftarrow 1$  to  $n$ 
 $S_i \leftarrow \text{RandomSample}(X, s)$ 
 $(a_i, c_i) \leftarrow k\text{means}(S_i, k)$ 
 $C \leftarrow \text{AppendRows}(C, c_i)$ 
End for
Return  $C_{\text{final}} \leftarrow k\text{means}(C, k)$ 
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*.

3.1.1. 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

a. Menghitung jarak dari data ke "1" terhadap *center* (pusat) *Cluster*

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

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

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

b. Menghitung jarak dari data ke "2" terhadap *center* (pusat) *Cluster*

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

$$= 11,83$$

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

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

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

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

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

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

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

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

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

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

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

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

Menentukan Centroid Baru:

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

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

$$C3 = \frac{5+1+1+2+1+1+2}{7}, \frac{9+2+2+2+1+3+8}{7}, \frac{6+5+7+2+3+7+5}{7}$$

$$= \frac{7+5+5+3+7+5+8}{7}, \frac{5+3+4+2+6+2+20}{7}$$

$$= 1,86, 3,86, 5, 5,71, 6,14$$

Iterasi-2

Proses iterasi-2 ini akan dilakukan perhitungan dengan *centroid* (pusat *Cluster*) nya yang tidak lagi

berdasarkan sampel sebelumnya namun dari *centroid* baru. Nilai pusat *Cluster* disajikan pada Tabel 2.

Tabel 2. Pusat *Cluster* Baru

Cluster	Nilai				
C1	6,00	11,00	19	10,00	7,00
C2	3,50	7,50	13	12,50	8,50
C3	1,86	3,86	5	5,71	6,14

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

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

$$= 0$$

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

$$= 7,94$$

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

$$= 16,83$$

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

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

$$= 6,71$$

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

$$= 4,8$$

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

$$= 11,4$$

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

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

$$= 13,67$$

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

$$= 9,80$$

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

$$= 6,35$$

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

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

$$= 18,52$$

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

$$= 13,67$$

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

$$= 3,82$$

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

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

$$= 16,85$$

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

$$= 12,21$$

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

$$= 3,64$$

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

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

$$= 21,45$$

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

$$= 16,91$$

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

$$= 6,08$$

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

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

$$= 11,27$$

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

$$= 4,80$$

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

$$= 13,35$$

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

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

$$= 19,77$$

$$C2 = \sqrt{(1-3,5)^2 + (1-7,5)^2 + (3-13)^2 + (7-12,5)^2 + (6-8,55)^2}$$

$$= 13,6$$

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

$$= 3,82$$

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

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

$$= 16,55$$

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

$$= 12,21$$

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

$$= 3,98$$

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

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

$$= 19,85$$

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

$$= 14,80$$

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

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 :

a. Untuk *Cluster* 1 (C1) /Paling Rawan: adalah 1

Cluster 1: Kecamatan Sipora Selatan

b. Untuk *Cluster* 2 (C2) /Cukup Rawan adalah 2,7

Cluster 2: Kecamatan Siberut Utara dan Kecamatan Sikakap

c. Untuk *Cluster* 3 (C3)/Tidak Rawan:adalah 3,4,5,6,8,9,10

Cluster 3: Kecamatan Siberut Barat, Kecamatan Siberut Tengah, Kecamatan Barat Daya, Kecamatan Pagai Utara, Kecamatan Pagai Selatan, dan Kecamatan Sipora Utara

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 Sipora Selatan	1	Paling Rawan
2.	Kecamatan Siberut Utara	2	Cukup Rawan
3.	Kecamatan Siberut Selatan	3	Tidak Rawan
4.	Kecamatan Siberut Barat	3	Tidak Rawan
5.	Kecamatan Siberut Tengah	3	Tidak Rawan
6.	Kecamatan Siberut Barat Daya	3	Tidak Rawan
7.	Kecamatan Sikakap	2	Cukup Rawan
8.	Kecamatan Pagai Selatan	3	Tidak Rawan
9.	Kecamatan Pagai Utara	3	Tidak Rawan
10.	Kecamatan Sipora Utara	3	Tidak Rawan

4 Kesimpulan

Dari hasil pengujian yang dilakukan dapat dijadikan pedoman bagi Polres Kepulauan Mentawai dalam mengetahui daerah yang paling rawan tindak kriminalitas pencurian kendaraan bermotor.

Daftar Rujukan

- [1] 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>
- [2] 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>
- [3] 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>
- [4] 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>
- [5] 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>
- [6] 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>
- [7] 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>
- [8] 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.
- [9] 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>
- [10] Andri Nofiar Am, Sarjon Defit, S. (2019). "Penentuan Mutu Kelapa Sawit Menggunakan Metode K-Means Clustering". *KomTekInfo*, 5(3), 1–9.
- [11] 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.
- [12] 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>
- [13] 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>
- [14] 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>
- [15] 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>
- [16] 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>
- [17] Suriani, L. (2020). Pengelompokan Data Kriminal Pada 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>