

Penerapan Metode K-Means ++ Untuk Pengelompokan Wilayah Rawan Kekerasan Anak dan Perempuan Di Kabupaten Nagekeo

Marshella Angela Merici Poa¹, Ahmad Fahrudi Setiawan², Joseph Dedy Irawan³

Teknik Informatika, Fakultas Teknologi Industri, Institut Teknologi Nasional Malang, Indonesia

E-mail: ¹2218054@scholar.itn.ac.id, ²fahrudi.itn@gmail.com, ³joseph@lecturer.itn.ac.id

(* : corresponding author)

Abstrak

Kekerasan terhadap perempuan dan anak di Kabupaten Nagekeo merupakan masalah sosial krusial yang memerlukan penanganan tepat sasaran. Dinas PMD-P3A menghadapi kendala dalam analisis kerawanan wilayah yang selama ini masih bersifat manual dan subjektif. Penelitian ini bertujuan membangun sistem pengelompokan wilayah rawan berbasis website menggunakan metode K-Means++ Clustering. Metode ini dipilih karena keunggulannya dalam mengoptimalkan penentuan centroid awal melalui perhitungan probabilitas jarak, sehingga menghasilkan klusterisasi yang lebih stabil dan akurat dibandingkan algoritma K-Means standar. Sistem dikembangkan menggunakan framework Laravel dan database MySQL dengan memanfaatkan data historis tahun 2020-2025. Proses pengelompokan dilakukan berdasarkan parameter Jenis Kekerasan dan Tempat Kejadian untuk memetakan wilayah ke dalam tiga tingkat kerawanan: Sangat Rawan, Rawan, dan Tidak Rawan. Hasil penelitian menunjukkan performa sistem yang sangat baik dengan nilai Silhouette Score sebesar 0.6633 dan Davies-Bouldin Indeks (DBI) sebesar 0.4520, yang mengindikasikan struktur kluster yang solid dan terpisah secara optimal. Selain data statistik, sistem juga menyediakan visualisasi pemetaan digital interaktif. Implementasi ini diharapkan dapat menjadi instrumen pendukung keputusan bagi pemerintah daerah dalam merumuskan kebijakan perlindungan sosial yang lebih efektif dan efisien di Kabupaten Nagekeo.

Kata Kunci: Clustering, Davies-Bouldin Indeks (DBI), Silhouette Score, Kekerasan Anak dan Perempuan

Abstract

Violence against women and children in Nagekeo Regency is a crucial social issue requiring targeted intervention. The Department of PMD-P3A faces challenges in analyzing regional vulnerability, which has historically been manual and subjective. This research aims to develop a web-based vulnerability grouping system implementing the K-Means++ Clustering method. This method was strategically selected for its ability to optimize initial centroid selection through distance probability calculations, resulting in more stable and accurate clustering compared to the standard K-Means algorithm. The system was developed using the Laravel framework and MySQL database, utilizing historical data from 2020 to 2025. The clustering process is based on two key parameters: Type of Violence and Place of Occurrence, mapping regions into three levels: Highly Vulnerable, Vulnerable, and Non-Vulnerable. The results demonstrate excellent system performance with a Silhouette Score of 0.6633 and a Davies-Bouldin Index (DBI) of 0.4520, indicating a solid and optimally separated cluster structure. Beyond statistical data, the system provides interactive digital mapping visualizations. This implementation is expected to serve as a decision-support tool for the local government in formulating more effective and efficient social protection policies in Nagekeo Regency.

Keywords: Clustering, Davies-Bouldin Index (DBI), Silhouette Score, Violence against Children and Women

1. PENDAHULUAN

Tindakan kekerasan adalah perampasan hak seseorang yang menimbulkan dampak negatif terhadap fisik atau psikis baik disengaja atau tidak dalam bentuk perbuatan lainnya [1]. Kekerasan terjadi karena adanya ketimpangan dalam deferensiasi sosial yang akhirnya menyebabkan terjadi tindak kekerasan. Kekerasan dalam arti luas merujuk pada tindakan fisik atau tindakan psikis yang dilakukan oleh seseorang atau sekelompok individu, baik yang dilakukan secara sengaja, langsung maupun tidak langsung, personal atau struktural [2].

Kasus kekerasan terhadap anak dan perempuan merupakan masalah yang sangat mengkhawatirkan bagi masyarakat, dengan meningkatnya kasus kekerasan yang menimpa

anak dan perempuan serta sedikitnya korban yang melaporkan kasus yang dialaminya. Ditambah lagi tipisnya jaminan perlindungan terhadap korban, yang dalam hal ini kaum wanita, yang akhirnya korban lebih banyak bersikap diam, karena mana yang benar dan mana yang salah tidak lagi ada pemahamannya yang serius [3].

Selama ini, analisis kerawanan suatu wilayah terhadap kekerasan anak dan perempuan seringkali dilakukan secara manual dan subjektif, sehingga kurang efektif dan cenderung lambat dalam merespons dinamika persebaran kasus [4]. Untuk itu diperlukan sebuah pendekatan yang lebih objektif dan berbasis data (data-driven) untuk dapat mengelompokkan wilayah berdasarkan tingkat dan karakteristik yang diukur. Dengan demikian, intervensi dan langkah pencegahan oleh pemerintah, dinas terkait, maupun lembaga masyarakat lebih tepat sasaran, efisien dan terfokus pada wilayah yang paling membutuhkan.

Kabupaten Nagekeo sebagai wilayah yang terus berkembang, tidak terlepas dari permasalahan sosial ini. Dinas PMD - P3A melalui Bidang Pemberdayaan Perempuan dan Perlindungan Anak setempat mencatat cukup banyak laporan dan kasus kekerasan anak dan dewasa yang terjadi di berbagai kecamatan. Data historis yang terkumpul memuat informasi penting seperti Jenis Kekerasan, Lokasi Kecamatan, Tempat Kejadian, Hubungan dengan Korban, Usia, dan Pendidikan Korban. Data ini merupakan aset berharga yang belum sepenuhnya dimanfaatkan untuk analisis lebih lanjut guna mendukung upaya pencegahan dan penanganan yang lebih efektif.

Metode K-Means++ Clustering sebagai salah satu yang mampu mengelompokkan data ke dalam beberapa cluster berdasarkan kesamaan karakteristiknya [5]. Dalam konteks ini, kecamatan-kecamatan yang dibagi lagi menjadi desa-desa dapat dikelompokkan berdasarkan tempat kejadian dan jenis kekerasan yang terjadi, sehingga dapat diidentifikasi kecamatan dan desa mana yang termasuk dalam kategori rawan, cukup rawan, atau tidak rawan untuk setiap jenis kekerasan tertentu.

Penerapan metode clustering untuk analisis kerawanan wilayah telah banyak dilakukan pada penelitian terdahulu, namun sebagian besar masih menggunakan algoritma K-Means standar. Sebagai contoh, penelitian oleh Febri Laksana, dkk. berhasil mengelompokkan jumlah kekerasan terhadap anak di Kabupaten Banyumas menggunakan K-Means menjadi tiga klaster (rendah, sedang, tinggi) [6]. Penelitian lain oleh Cahya Diningrat, dkk. juga menerapkan K-Means untuk mengklusterisasi kasus di Jawa Barat berdasarkan lokasi kejadian, yang juga menghasilkan tiga tingkatan cluster [7]. Meskipun kajian-kajian tersebut menunjukkan bahwa pengelompokan berbasis jarak sangat relevan untuk mengidentifikasi pola kerawanan, penggunaan K-Means standar memiliki kelemahan pada sensitivitas terhadap inisialisasi pusat klaster (*centroid*) yang acak, yang seringkali menyebabkan hasil yang tidak konsisten.

Oleh karena itu, diperlukan sebuah sistem yang mampu mengelompokkan wilayah secara objektif. Penelitian ini bertujuan untuk membangun sebuah sistem pengelompokan wilayah rawan berbasis website yang mengimplementasikan Metode K-Means++. Penggunaan K-Means++ merupakan langkah strategis untuk mengatasi kelemahan utama K-Means standar yang sangat bergantung pada inisialisasi centroid awal secara acak. Dengan algoritma ini, pemilihan pusat klaster awal dilakukan melalui perhitungan probabilitas yang memastikan jarak antar centroid saling berjauhan. Hal ini secara signifikan mengurangi risiko terjebaknya hasil pada local optima, sehingga menghasilkan klusterisasi yang lebih stabil, konsisten, dan akurat. Sistem ini dapat digunakan secara otomatis untuk mengelompokkan wilayah kecamatan dan desa berdasarkan jenis kekerasan dan tempat kejadian menjadi kategori sangat rawan, rawan, atau tidak rawan. Penerapan metode ini diharapkan dapat mendukung pemerintah daerah dalam mengambil kebijakan yang lebih tepat sasaran, berbasis data, dan mampu meminimalisir bias dalam penentuan skala prioritas penanganan kasus.

2. METODE PENELITIAN

2.1. Dataset

Data yang digunakan adalah data sekunder kasus kekerasan dari Dinas PMD-P3A Kabupaten Nagekeo periode tahun 2020-2025. Atribut yang digunakan untuk clustering meliputi Kecamatan, Desa, Jenis Kekerasan (Fisik, Psikis, Seksual, Penelantaran, TPPO), dan Tempat Kejadian.

Tabel 1. Defenisi Atribut Data

Atribut	Defenisi
Fisik	Tindakan yang bertujuan melukai tubuh secara langsung seperti pemukulan atau penganiayaan.
Psikis	Perbuatan yang mengakibatkan ketakutan, hilangnya rasa percaya diri, atau penderitaan psikologis.
Seksual	Setiap perbuatan yang berupa pemaksaan hubungan seksual atau pelecehan seksual.
Penelantaran	Tidak memberikan perawatan atau kehidupan yang layak bagi orang di lingkup rumah tangga.
TPPO	Tindak Pidana Perdagangan Orang untuk tujuan eksploitasi.
Tempat Kejadian	Lokasi spesifik terjadinya kekerasan (Rumah, Sekolah, Kebun/Sawah)

Penentuan tingkatan kerawanan dibagi menjadi tiga kelas (Sangat Rawan, Rawan, Tidak Rawan) berdasarkan parameter frekuensi dan bobot jenis kekerasan. Kelas 'Sangat Rawan' didefinisikan sebagai wilayah dengan frekuensi laporan tahunan di atas rata-rata kabupaten (>7) atau memiliki skor intensitas kekerasan seksual lebih dari 5 kasus per bulan. Kelas 'Rawan' mencakup wilayah dengan akumulasi kekerasan fisik dan psikis dalam rentang nilai median (3–6 kasus), yang menunjukkan adanya pola kekerasan yang berulang secara signifikan secara statistik. Sementara itu, wilayah 'Tidak Rawan' dikategorikan berdasarkan frekuensi laporan yang berada di bawah standar deviasi minimum (<2 kasus per bulan) atau hanya mencakup laporan kekerasan ringan tanpa adanya catatan kekerasan berat.

2.2. Algoritma K-Means++

K-Means merupakan algoritma *clustering* yang dilakukan secara berulang. Algoritma *K-Means* dimulai dengan pemilihan K yang dipilih secara acak, K merupakan jumlah *cluster* yang ingin dibuat dan nilainya ditentukan secara acak yang kemudian nilai tersebut dijadikan sebagai pusat dari *cluster* atau *centroid* [8]. Langkah selanjutnya ialah menghitung jarak dari setiap data terhadap masing-masing *centroid* sampai data tersebut ditemukan jarak yang paling dekat dengan *centroid* menggunakan rumus *Euclidean Distance*. Setiap data diklasifikasikan berdasarkan jarak kedekatan dengan *centroid*. Langkah- langkah tersebut diulang sampai nilai *centroid* yang diperoleh tidak ada perubahan lagi [9].

Sedangkan *K-Means++* merupakan pengembangan dari algoritma *K-Means*. Perbedaannya ada pada penentuan nilai awal. *K-Means++* digunakan untuk meminimalisir dampak buruk dari algoritma *K-Means* yang bergantung pada nilai awal [10].

Adapun algoritma *K-Means++* adalah sebagai berikut :

- Menentukan jumlah *cluster* yang akan dibuat.
- Memilih satu titik data secara acak sebagai pusat kluster pertama.
- Menghitung jarak semua data dengan *centroid* 1 dengan persamaan berikut [11].

$$D^2(x, y) = \left(\sum_{i=1}^n (x_i - y_i)^2 \right) \quad (1)$$

Penjelasan :

$D(x, y)$ = Nilai *euclidean distance* antara titik x dan y

x = Titik data objek

y = Titik data *centroid*

i = Jumlah atribut data

n = Banyaknya observasi

Penelitian ini menggunakan K-Means++ untuk memperbaiki kelemahan K-Means konvensional dalam pemilihan pusat kluster (centroid) awal. Algoritma K-Means++ menggunakan proses inisialisasi nilai pusat cluster (centroids) awal, hal ini mengurangi ketidakstabilan yang terjadi pada K-Means sehingga dapat memberikan hasil klasterisasi yang baik dan stabil [13], sehingga mempercepat waktu komputasi dan meningkatkan akurasi pengelompokan.

- a. Menentukan *centroid* berikutnya dengan memilih nilai maksimum dari jarak terdekat pada *centroid* sebelumnya dengan persamaan berikut [11].

$$K = \frac{D(x)^2}{\sum_{x \in X} D(x)^2} \quad (2)$$

Penjelasan :

$D(x)^2$ = Kuadrat jarak terdekat dari titik x ke pusat *cluster* terdekat

$\sum_{x \in X} D(x)^2$ = Total dari semua kuadrat jarak terdekat untuk setiap titik x

- b. Setelah semua nilai *centroid* ditentukan, lanjutkan proses iterasi dengan algoritma K-Means++. Hitung kembali jarak setiap data ke-sebanyak jumlah *centroid* dengan persamaan (1).
- c. Mengelompokkan data berdasarkan jarak terdekat antara data dengan *centroid*.
- d. Memperbarui nilai *centroid*. Nilai *centroid* baru diperoleh dari rata-rata *cluster* yang bersangkutan dengan menggunakan persamaan berikut [12].

$$C_k = \left(\frac{1}{n_k}\right) \sum d_i \quad (3)$$

Penjelasan :

C_k = *Centroid cluster k*

n_k = Jumlah data dalam *cluster k*

d_i = Jumlah dari nilai jarak yang masuk ke dalam masing-masing *cluster*

Ulangi langkah nomor 5-7, apabila posisi *centroid* baru tidak sama. Namun jika posisi *centroid* sudah sama maka proses iterasi telah selesai.

2.3. Silhouette Score

Silhouette score merupakan metode yang menekankan pada validasi dan penafsiran set data untuk mengukur kualitas sebuah cluster [14]. Penggabungan konsep *cohesion* dan *separation* digunakan untuk mengukur hasil *clustering*. Rumus *silhouette score* yaitu :

$$S(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (4)$$

Penjelasan :

S_i = Nilai *silhouette score*

a_i = Rata-rata jarak dari titik data i ke semua titik lain dalam klusternya (*cohesion*)

b_i = Rata-rata jarak dari titik data i ke semua titik dalam kluster terdekat yang berbeda (*separation*)

Nilai rata-rata pada metode *silhouette score* dari masing-masing data objek dalam suatu cluster menunjukkan seberapa layak data tersebut dimasukkan ke dalam cluster. *Silhouette Score* berkisar dari -1 hingga 1, yang berarti bahwa titik data lebih dekat ke centroid cluster mereka sendiri dibandingkan dengan ke centroid cluster lain [12].

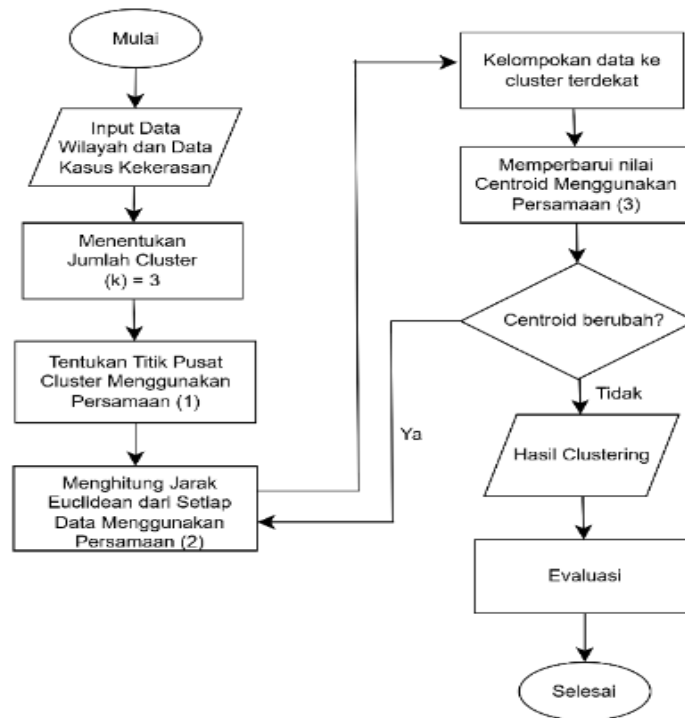
2.4. Davies-Bouldin Index (DBI)

Davies-Bouldin yaitu metode evaluasi internal yang menghitung ukuran evaluasi kelompok berdasarkan nilai kohesi dan separasi. Dalam pengelompokan, kohesi merupakan jumlah kedekatan data dengan centroid kluster yang diikutinya, sementara separasi berdasarkan pada jarak antara kluster centroid. Kluster dengan jarak intra-kluster yang kecil dan jarak antar-kluster

besar akan memiliki eksponen Davies Bouldin yang rendah. Clustering yang menghasilkan himpunan klaster dengan indeks Davies-Bouldin terkecil dianggap sebagai algoritma terbaik[15].

2.5. Flowchart Metode K-Means ++

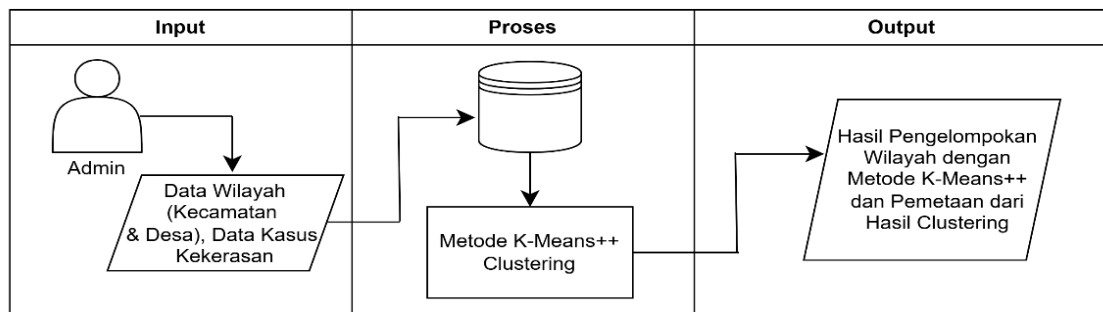
Alur K-Means++ diawali dengan input data kasus kekerasan dan penetapan jumlah k sebanyak 3 cluster. Sistem menentukan centroid awal, kemudian memulai iterasi dengan menghitung jarak Euclidean setiap data terhadap seluruh centroid untuk menentukan afiliasi klaster terdekat. Posisi centroid lalu diperbarui berdasarkan rata-rata anggota cluster baru, dan proses ini berulang hingga posisi centroid stabil atau tidak berubah lagi. Hasil akhirnya berupa pengelompokan wilayah yang disajikan sebagai luaran sistem sebagaimana ditunjukkan pada Gambar 2.



Gambar 1. Flowchart Metode K-Means ++

2.6. Diagram Blok Sistem

Berdasarkan diagram blok pada Gambar 3, sistem ini terbagi menjadi tiga tahapan utama: Input, Proses, dan Output. Tahap Input dimulai dengan administrator memasukkan data wilayah dan data kasus kekerasan. Agar dapat diolah secara matematis, data tersebut direpresentasikan dalam bentuk vektor numerik (vektor ruang), di mana setiap wilayah (desa/kecamatan) menjadi satu titik data (X_i) dengan dimensi yang ditentukan oleh frekuensi jenis kekerasan dan lokasi kejadian. Selanjutnya pada tahap Proses, data yang telah terukur tersebut disimpan di dalam database lalu diolah menggunakan algoritma K-Means++. Algoritma ini pertama-tama menentukan titik pusat awal (initial centroids) melalui pemilihan probabilitas berdasarkan jarak terjauh antar data untuk memastikan persebaran yang optimal. Sistem kemudian melakukan iterasi dengan menghitung jarak Euclidean antara vektor data wilayah terhadap centroid. Data wilayah akan berafiliasi dengan klaster yang memiliki jarak terkecil, hingga tercapai kondisi konvergensi di mana titik pusat tidak lagi bergeser. Tahap Output kemudian menyajikan hasil akhir berupa pengelompokan wilayah ke dalam kategori sangat rawan, rawan, atau tidak rawan serta visualisasi pemetaan geografis untuk mempermudah pemahaman distribusi data secara spasial.



Gambar 2. Diagram Blok Sistem

3. HASIL DAN PEMBAHASAN

3.1 Tahapan Metode K-Means++

a. Inisialisasi Centroid Awal (Iterasi ke-1)

1) Menentukan C1

Pada data perhitungan ini jumlah cluster ada 3 yakni Sangat Rawan, Rawan dan Tidak Rawan. Setelah menentukan jumlah cluster yang ingin dibentuk tahap selanjutnya yaitu menentukan centroid awal terlebih dahulu. Dalam penentuan centroid awal ini dilakukan secara random untuk centroid pertama (C1). Seperti yang ditampilkan pada Tabel 2.

Tabel 2. Centroid Pertama

Desa	Jenis Kekerasan					Tempat Kejadian		
	Fisik	Psikis	Seksual	Penelantaran	TPPO	Rumah	Sekolah	Kebun/Sawah
Danga	0	1	1	0	0	1	1	0
Aeramo	0	0	2	0	0	1	0	1
Ngededhawe	0	0	2	0	0	0	1	1
Tonggurambang	1	1	1	1	0	2	1	1
Mbay 1	0	1	2	0	0	0	1	2
Nggolonio	0	0	2	0	0	0	0	2
Dhawe	1	0	1	0	0	1	1	0
Marapokot	0	1	1	2	0	2	1	1
Nangadhero	1	1	1	0	0	1	1	1
Olaia	0	0	0	1	0	1	0	0
Kelimado	1	0	0	0	0	1	0	0
Maukeli	0	0	1	0	0	0	0	1
Nagerawe	0	0	1	0	0	0	0	1
Natanage	0	0	0	0	0	0	0	0
Langedhawe	1	0	0	0	0	1	0	0
Wuliwawo	0	0	0	0	0	0	0	0
Woewolo	0	0	2	0	0	1	0	1
Kotakeo 2	0	0	0	0	0	0	0	0
Ulupulu	1	0	0	0	1	2	0	0
Nataute	1	0	1	0	0	0	2	0
Witurombaua	0	1	0	0	0	0	0	1

Setelah memilih centroid pertama (C1) yang diambil dari data wilayah Desa Danga, hitunglah jarak semua data dengan C1 menggunakan persamaan jarak euclidean. Adapun untuk menghitung jarak adalah sebagai berikut :

$$D^2(x_1, SR) = \frac{(0-0)^2 + (1-1)^2 + (1-1)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (1-1)^2 + (1-1)^2 + (0-0)^2 + (0-0)^2}{10} = 0$$

$$D^2(x_2, SR) = \frac{(0-0)^2 + (0-1)^2 + (2-1)^2 + (0-0)^2 + (0-0)^2 + (1-1)^2 + (0-1)^2 + (1-0)^2 + (0-0)^2}{10}$$

$$= 4$$

$$D^2(x_3, SR) = \frac{(0-0)^2 + (0-1)^2 + (2-1)^2 + (0-0)^2 + (0-0)^2 + (0-1)^2 + (1-1)^2 + (1-0)^2 + (0-0)^2}{4} = 4$$

Perhitungan pada centroid tersebut dilakukan sebanyak jumlah data yang ada pada tabel Data Kekerasan Kabupaten Nagekeo. Untuk data yang dihitung diambil dari data tahun 2024.

2) Menentukan C2

Inisialisasi centroid kedua C2 dilakukan dengan memilih titik data x yang memiliki Jarak Kuadrat Euclidean Terbesar dari centroid pertama C1. Titik data ini kemudian ditetapkan sebagai C2 yaitu diambil dari data wilayah Nggolonio. Selanjutnya, untuk menentukan centroid berikutnya Ck, di mana $k > 2$, perhitungan dilakukan pada semua centroid yang telah ditetapkan sebelumnya. Centroid Ck dipilih sebagai titik data yang memiliki Jarak Kuadrat Euclidean Terdekat yang bernilai maksimum dari semua centroid sebelumnya yang dapat dilihat pada Tabel 3 berikut.

Tabel 3. Hasil Perhitungan Jarak data dengan C1

Desa	Jenis Kekerasan					Tempat Kejadian			Jarak C1
	Fisik	Psikis	Seksual	Penelantaran	TPPO	Rumah	Sekolah	Kebun/Sawah	
Danga	0	1	1	0	0	1	1	0	0
Aeramo	0	0	2	0	0	1	0	1	4
Ngedghawe	0	0	2	0	0	0	1	1	4
Tongguramban	1	1	1	1	0	2	1	1	4
Mbay 1	0	1	2	0	0	0	1	2	6
Nggolonio	0	0	2	0	0	0	0	2	8
Dhawe	1	0	1	0	0	1	1	0	2
Marapokot	0	1	1	2	0	2	1	1	6
Nangadhero	1	1	1	0	0	1	1	1	2
Olaia	0	0	0	1	0	1	0	0	4
Kelimado	1	0	0	0	0	1	0	0	4
Maukeli	0	0	1	0	0	0	0	1	4
Nagerawe	0	0	1	0	0	0	0	1	4
Natanage	0	0	0	0	0	0	0	0	4
Langedhawe	1	0	0	0	0	1	0	0	4
Wuliwawo	0	0	0	0	0	0	0	0	4
Woewolo	0	0	2	0	0	1	0	1	4
Kotakeo 2	0	0	0	0	0	0	0	0	4
Ulupulu	1	0	0	0	1	2	0	0	6
Nataute	1	0	1	0	0	0	2	0	4
Witurombaua	0	1	0	0	0	0	0	1	4

3) Menentukan C3

Inisialisasi centroid ketiga C3 dilakukan dengan memilih titik data x yang memiliki Jarak Kuadrat Euclidean Terbesar dari centroid kedua C2. Titik data ini kemudian ditetapkan sebagai C3 yang diambil dari data wilayah Desa Ulupulu seperti pada Tabel 4 berikut.

Tabel 4. Hasil Perhitungan Jarak C2

Desa	Jenis Kekerasan					Tempat Kejadian			Jarak C1	Jarak C2
	Fisik	Psikis	Seksual	Penelantaran	TPPO	Rumah	Sekolah	Kebun/Sawah		
Danga	0	1	1	0	0	1	1	0	0	8
Aeramo	0	0	2	0	0	1	0	1	4	2
Ngegedhawe	0	0	2	0	0	0	1	1	4	2
Tonggurambang	1	1	1	1	0	2	1	1	4	10
Mbay 1	0	1	2	0	0	0	1	2	6	2
Nggolonio	0	0	2	0	0	0	0	2	8	0
Dhawe	1	0	1	0	0	1	1	0	2	8
Marapokot	0	1	1	2	0	2	1	1	6	12
Nangadhero	1	1	1	0	0	1	1	1	2	6
Olaia	0	0	0	1	0	1	0	0	4	10
Kelimado	1	0	0	0	0	1	0	0	4	10
Maukeli	0	0	1	0	0	0	0	1	4	2
Nagerawe	0	0	1	0	0	0	0	1	4	2
Natanage	0	0	0	0	0	0	0	0	4	8
Langedhawe	1	0	0	0	0	1	0	0	4	10
Wuliwawo	0	0	0	0	0	0	0	0	4	8
Woewolo	0	0	2	0	0	1	0	1	4	2
Kotakeo 2	0	0	0	0	0	0	0	0	4	8
Ulupulu	1	0	0	0	1	2	0	0	6	14
Nataute	1	0	1	0	0	0	2	0	4	10
Witurombaua	0	1	0	0	0	0	0	1	4	6

4) Centroid Awal

Dari perhitungan jarak untuk menentukan C2 dan C3 sebelumnya didapat hasil akhir untuk centroid iterasi pertama yaitu, data ke-1 (C1) dengan Desa Danga, data ke-14 (C2) dengan Desa Dhawe, dan data ke-5 (C3) dengan Desa Ulupulu yang dapat dilihat pada Tabel 5 berikut.

Tabel 5. Centroid Iterasi 1

Desa	Jenis Kekerasan					Tempat Kejadian		
	Fisik	Psikis	Seksual	Penelantaran	TPPO	Rumah	Sekolah	Kebun/Sawah
Danga	0	1	1	0	0	1	1	0
Aeramo	0	0	2	0	0	1	0	1
Ngegedhawe	0	0	2	0	0	0	1	1
Tonggurambang	1	1	1	1	0	2	1	1
Mbay 1	0	1	2	0	0	0	1	2
Nggolonio	0	0	2	0	0	0	0	2
Dhawe	1	0	1	0	0	1	1	0
Marapokot	0	1	1	2	0	2	1	1
Nangadhero	1	1	1	0	0	1	1	1
Olaia	0	0	0	1	0	1	0	0
Kelimado	1	0	0	0	0	1	0	0
Maukeli	0	0	1	0	0	0	0	1
Nagerawe	0	0	1	0	0	0	0	1

Desa	Jenis Kekerasan					Tempat Kejadian		
	Fisik	Psikis	Seksual	Penelantaran	TPPO	Rumah	Sekolah	Kebun/Sawah
Natanage	0	0	0	0	0	0	0	0
Langedhawe	1	0	0	0	0	1	0	0
Wuliwawo	0	0	0	0	0	0	0	0
Woewolo	0	0	2	0	0	1	0	1
Kotakeo 2	0	0	0	0	0	0	0	0
Ulupulu	1	0	0	0	1	2	0	0
Nataute	1	0	1	0	0	0	2	0
Witurombaua	0	1	0	0	0	0	0	1

b. Menentukan Jarak Pada Iterasi Ke-1

Setelah masing-masing centroid sudah ditetapkan dan jarak C1 dan C2 sudah didapatkan, langkah selanjutnya yaitu menghitung jarak C3 dengan menggunakan persamaan Jarak Kuadrat Euclidean seperti pada perhitungan jarak C1 sebelumnya dan kemudian mengelompokkannya ke dalam kluster 1, 2, dan 3 berdasarkan jarak terdekat (nilai jarak terkecil) dari pusat cluster. Adapun hasil perhitungan jarak C3 dapat dilihat pada Tabel 6 berikut.

Tabel 6. Hasil Iterasi ke-1

ITERASI 1			
C1	C2	C3	Cluster
0	8	6	3
4	2	8	3
4	2	12	3
.....			
4	6	8	3

c. Menentukan Centroid Baru Iterasi 2

Setelah data dikelompokkan pada iterasi pertama, langkah selanjutnya dalam algoritma K-Means++ adalah memperbarui posisi Centroid (pusat kluster) untuk Iterasi berikutnya. Proses ini dimulai dengan mengidentifikasi semua titik data yang termasuk dalam cluster yang sama berdasarkan jarak terdekat yang baru ditetapkan. Kemudian, untuk setiap cluster (Cluster 1, Cluster 2, dan Cluster 3), Centroid baru dihitung sebagai rata-rata aritmatika dari nilai asli semua titik data yang menjadi anggotanya. Centroid baru ini kemudian menggantikan Centroid lama dan digunakan untuk menghitung ulang jarak dari setiap titik data dalam dataset untuk memulai Iterasi berikutnya. Proses ini diulang hingga posisi Centroid tidak lagi berubah secara signifikan (konvergen). Berdasarkan perhitungan di atas centroid baru untuk proses iterasi ke- 2 dapat dilihat pada Tabel 7 berikut.

Tabel 7. Centroid Iterasi ke-2

Iterasi 2	1	2	3	4	5	6	7	8	9
c1	0	0	0	1	0	1	0	0	0
c2	1	0	0	0	0,33	1,33	0	0	0
c3	0,23	0,35	1,05	0,17	0	0,52	0,52	0,76	0

d. Menentukan Jarak pada Iterasi ke-2

Setelah nilai centroid baru pada iterasi ke 2 sudah terbentuk, langkah selanjutnya menghitung jarak antara masing-masing data dengan 3 centroid baru tersebut, dengan menggunakan persamaan Kuadrat Euclidean Distance seperti sebelumnya pada iterasi 1. Kemudian mengelompokkannya ke dalam kluster 1, 2, dan 3 berdasarkan jarak terdekat (nilai

jarak terkecil) dari pusat cluster. Hasil dari perhitungan jarak pada iterasi 2 terdapat pada Tabel 8.

Tabel 8. Hasil Iterasi ke 2

ITERASI 2			
C1	C2	C3	Cluster
4	4,222222	1,536332	2
6	6,222222	1,653979	2
8	8,888889	1,653979	2
.....			
4	4,888889	2,242215	2

e. Menentukan Centroid Baru Iterasi ke-3

Hasil dari perhitungan rata rata setiap data, dengan menjumlahkan cluster yang sama dan dibagi dengan jumlah cluster pada iterasi 2 sebelumnya, membentuk centroid baru kembali pada iterasi ke-3. Centroid baru pada iterasi 3 ini selanjutnya digunakan untuk menghitung jarak kembali dengan persamaan Kuadrat Euclidean Distance. Adapun hasil dari centroid baru pada iterasi ke-3 dapat dilihat pada Tabel 9 berikut.

Tabel 9. Centroid Iterasi ke-3

iteras 3	1	2	3	4	5	6	7	8	9
c1	0	0	0	1	0	1	0	0	0
c2	0,33	0,41	1,16	0,08	0	0,5	0,66	0,83	0
c3	0,42	0,14	0,14	0,42	0,14	1	0,14	0,14	0

f. Menentukan Jarak Pada Iterasi ke-3

Setelah menemukan centroid untuk iterasi yang ke 3, langkah selanjutnya yaitu menghitung jarak antara setiap data dengan 3 centroid baru tersebut menggunakan persamaan Kuadrat Euclidean Distance. Kemudian kelompokkan hasil ke dalam klaster 1, 2, dan 3 berdasarkan jarak terdekat (nilai jarak terkecil) dari pusat cluster. Adapun hasil dari perhitungan jarak dari masing-masing data dengan 3 centroid baru pada iterasi yang ke -3 dapat dilihat pada Tabel 10 berikut.

Tabel 10. Hasil Iterasi ke-3

ITERASI 3			
C1	C2	C3	Cluster
4	1,541667	2,612245	3
6	1,708333	4,612245	3
8	1,375	6,326531	3
.....			
4	2,541667	2,897959	3

g. Menentukan Centroid Baru Iterasi ke-4

Hasil perhitungan rata rata setiap data, dengan menjumlahkan cluster yang sama dan dibagi dengan jumlah cluster pada iterasi 3 sebelumnya, kemudian akan digunakan untuk membentuk centroid baru kembali pada iterasi ke-4. Centroid baru pada iterasi 4 ini selanjutnya digunakan untuk menghitung jarak kembali dengan persamaan Kuadrat Euclidean Distance. Adapun hasil dari centroid baru pada iterasi yang -3 dapat dilihat pada Tabel 11 berikut.

Tabel 11. Centroid Iterasi ke-4

iteras 4	1	2	3	4	5	6	7	8	9
c1	0	0	0	1	0	1	0	0	0
c2	0,5	0	0	0	0,16	0,66	0	0	0
c3	0,27	0,36	1,18	0,18	0	0,54	0,63	0,81	0

h. Menentukan Jarak Pada Iterasi ke-4

Setelah menemukan centroid untuk iterasi yang ke 4, langkah selanjutnya yaitu menghitung jarak antara setiap data dengan 3 centroid baru tersebut menggunakan persamaan Kuadrat Euclidean Distance. Kemudian kelompokkan hasil ke dalam klaster 1, 2, dan 3 berdasarkan jarak terdekat (nilai jarak terkecil) dari pusat cluster. Adapun hasil dari perhitungan jarak dari masing-masing data dengan 3 centroid baru pada iterasi yang ke 4 dapat dilihat pada Tabel 12 berikut.

Tabel 12. Hasil Iterasi ke-4

ITERASI 4			
C1	C2	C3	Cluster
4	3,388889	1,553719	2
6	5,388889	1,553719	2
8	6,722222	1,371901	2
.....			...
4	2,722222	2,644628	1

i. Menentukan Centroid Baru Iterasi ke-5

Hasil dari perhitungan rata rata setiap data, dengan menjumlahkan cluster yang sama dan dibagi dengan jumlah cluster pada iterasi 4 sebelumnya, akan digunakan untuk membentuk centroid baru kembali pada iterasi ke-5 Bisa dilihat bahwa centroid baru pada iterasi yang ke -4 hasilnya sama dengan centroid baru pada iterasi yang ke-5 ini. Centroid baru pada iterasi 5 ini selanjutnya digunakan untuk menghitung jarak kembali dengan persamaan Euclidean Distance. Adapun hasil dari centroid baru pada iterasi ke-5 dapat dilihat pada Tabel 13 berikut.

Tabel 13. Centroid Iterasi-5

Mean Data Nyata	Fisik	Psikis	Seksual	Penelantaran	TPPO	Rumah	Sekolah	Kebun/Sawah
C1	0.375	0.375	0.125	0.125	0.125	0.625	0.500	0.125
C2	0.428	0	0.142	0.285	0	0.571	0.142	0
C3	0.375	0.125	1,375	0	0	0.500	0.625	0.500

Dari hasil perhitungan jarak menggunakan persamaan Eucliden Distance yang dilakukan pada iterasi yang ke-5. Langkah selanjutnya yaitu mengelompokkan daripada data-data tersebut, dengan mencari nilai yang paling kecil atau terdekat dengan pusat cluster Dapat dilihat bahwa hasil pengelompokkan data pada iterasi ke - 5 nilainya tetap atau tidak berubah, yakni dengan hasil pengelompokkan data pada iterasi yang ke 4. Artinya perhitungan selesai dan hanya sampai pada iterasi ke-5.

j. Menentukan Rata-Rata Akhir Centroid Tiap Cluster

Setelah algoritma K-Means++ mencapai konvergensi yang ditandai dengan stabilnya alokasi data dilakukan penetapan centroid akhir dengan menghitung nilai rata-rata (*mean*) dari seluruh anggota klaster. Centroid ini berfungsi sebagai representasi numerik pusat klaster yang menjadi acuan utama dalam menganalisis karakteristik tiap kelompok. Rincian nilai centroid akhir tersebut disajikan pada Tabel 14.

Tabel 14. Hasil Rata-Rata Centroid Akhir

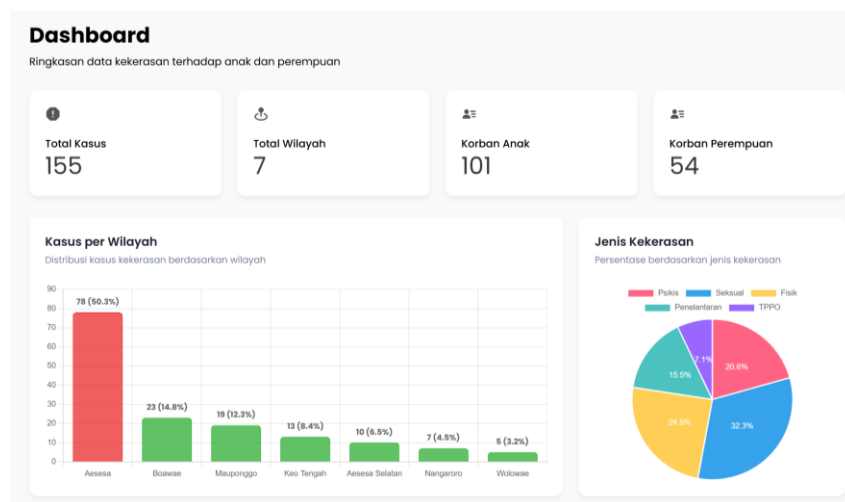
Mean Data Nyata	Fisik	Psikis	Seksual	Penelantaran	TPPO	Rumah	Sekolah	Kebun/Sawah
C1	0.375	0.375	0.125	0.125	0.125	0.625	0.500	0.125
C2	0.428	0	0.142	0.285	0	0.571	0.142	0
C3	0.375	0.125	1,375	0	0	0.500	0.625	0.500

Berdasarkan perhitungan di atas, Cluster 1 dikategorikan sebagai Tingkat Kerawanan Sedang (Rawan) karena menunjukkan karakteristik kekerasan yang beragam; meskipun nilai Seksual rendah (0,125), cluster ini memiliki rata-rata kasus Psikis (0,375) dan TPPO (0,125) tertinggi dibandingkan kelompok lain, serta didukung nilai Fisik yang cukup signifikan (0,375). Sebaliknya, Cluster 2 masuk dalam kategori Tingkat Kerawanan Rendah (Tidak Rawan). Walaupun variabel Fisik (0,428) dan Penelantaran (0,285) dominan di cluster ini, absennya kasus Psikis dan TPPO (0,000) serta minimnya kasus Seksual dan Sekolah (0,142) menjadikan intensitas kerawanan keseluruhan paling minim. Terakhir, Cluster 3 menempati posisi Tingkat Kerawanan Tinggi (Sangat Rawan) dengan indikator paling ekstrim, didominasi oleh rata-rata Kekerasan Seksual (1,375) yang jauh melampaui cluster lain, serta tingginya kejadian di Sekolah (0,625) dan Kebun/Sawah (0,500).

Hasil pengelompokan ini memiliki tingkat akurasi dan stabilitas yang lebih tinggi karena penggunaan metode K-Means++. Jika dibandingkan dengan K-Means biasa, algoritma standar seringkali terjebak dalam local optima akibat inisialisasi centroid yang murni acak, sehingga ada kemungkinan Cluster 3 (Sangat Rawan) tidak teridentifikasi secara terpisah atau bercampur dengan data dari klaster lain pada percobaan yang berbeda. Dengan K-Means++, pemilihan pusat klaster awal dihitung berdasarkan probabilitas jarak terjauh, memastikan bahwa titik pusat untuk kategori "Sangat Rawan" benar-benar mewakili pencilan (outlier) data kekerasan seksual yang ekstrem. Hal ini memberikan hasil yang konsisten dan meminimalisir bias subjektivitas dibandingkan K-Means standar, sehingga pemetaan wilayah rawan di Kabupaten Nagekeo menjadi lebih reliabel untuk pengambilan kebijakan.

3.2 Tampilan Halaman Dashboard

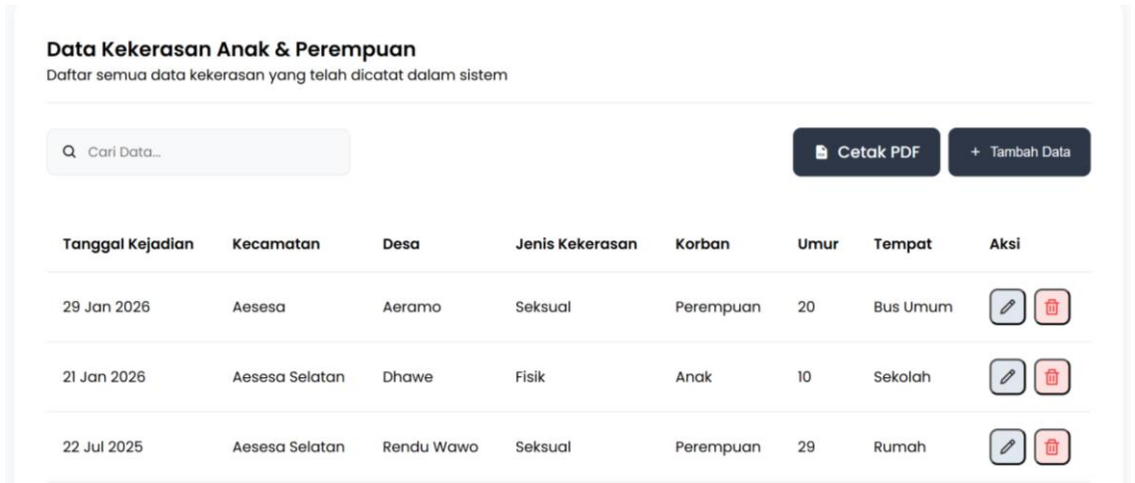
Halaman Dashboard merupakan halaman utama yang menampilkan ringkasan data secara visual. Pada halaman ini, telah berhasil diimplementasikan kartu ringkasan (*summary cards*) untuk data total kasus, korban anak, dan korban perempuan. Selain itu, terdapat dua grafik yaitu diagram batang untuk menampilkan jumlah kasus per wilayah dan diagram lingkaran (pie chart) untuk persentase jenis kekerasan. Seperti yang ditampilkan pada Gambar 11.



Gambar 3. Halaman Dashboard

3.3 Tampilan Halaman Data Kasus Kekerasan

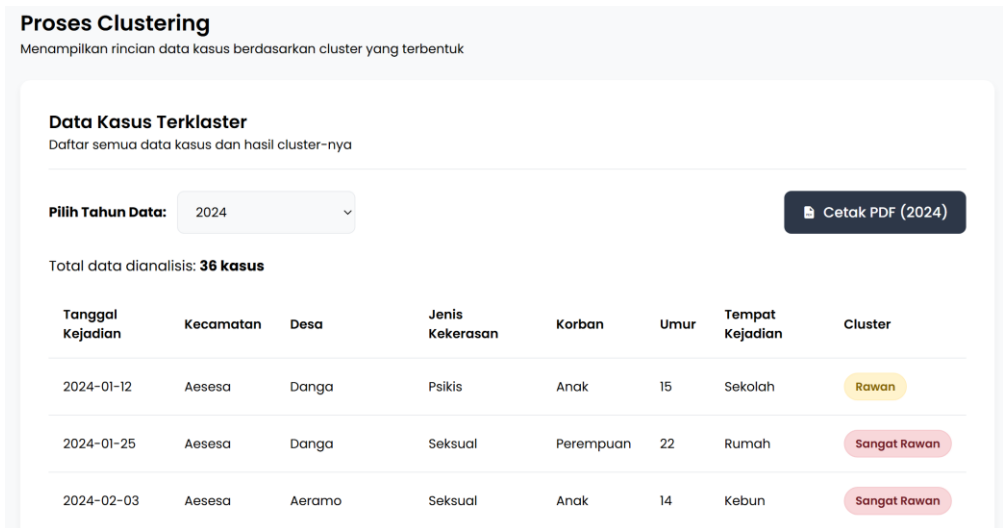
Halaman ini berfungsi untuk mengelola data utama yaitu data kasus kekerasan. Fungsionalitas untuk menambah, mengubah, dan menghapus data kasus telah berhasil diimplementasikan. Setiap data kasus mencakup informasi detail seperti tanggal kejadian, lokasi (kecamatan dan desa), jenis kekerasan, dan data korban. Fitur untuk mencetak data ke dalam format PDF juga telah ditambahkan serta dilengkapi dengan fungsi pencarian untuk mempermudah pengelolaan data. Seperti yang ditampilkan pada Gambar 12.



Gambar 4. Halaman Data Kasus Kekerasan

3.4 Tampilan Halaman Proses Clustering

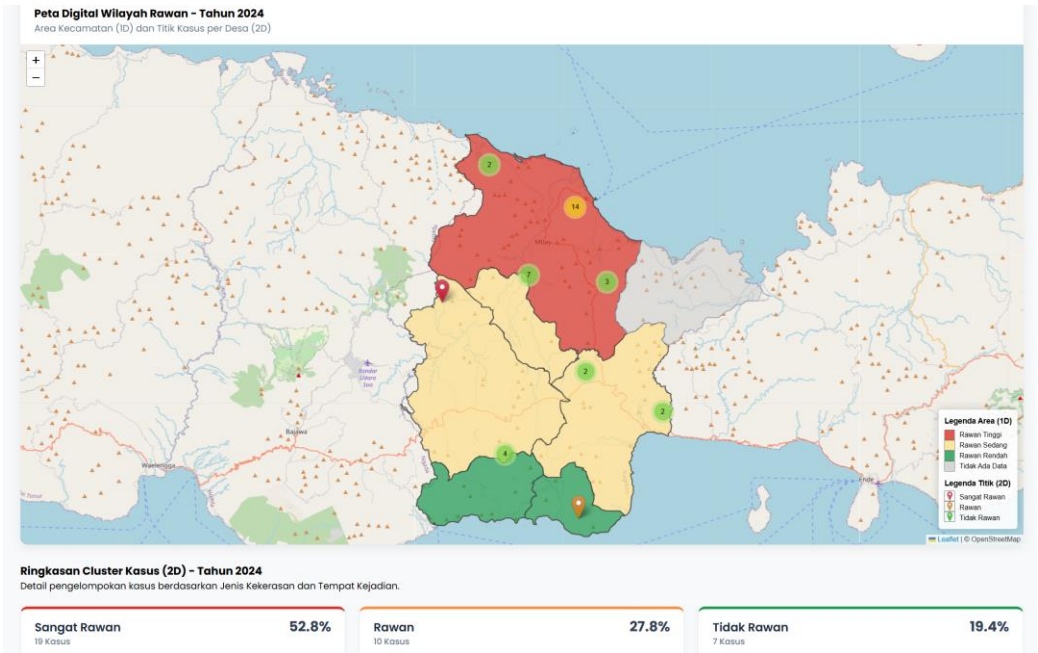
Halaman ini untuk fitur proses *clustering* dengan menggunakan metode *K-Means ++*. Halaman ini menyediakan pilihan dropdown dimana pengguna dapat memilih tahun berapa yang akan melakukan proses *clustering*. Setelah pengguna memilih tahun berapa, hasil *clustering* dari tahun tersebut akan ditampilkan pada halaman ini. Seperti yang ditampilkan pada Gambar 13.



Gambar 5. Halaman Proses Clustering

3.5 Tampilan Halaman Pemetaan

Pada halaman Pemetaan fungsi utama pada halaman ini adalah menampilkan pemetaan wilayah rawan kekerasan berdasarkan hasil pengelompokan. Terdapat sebuah fitur filter berupa dropdown "Pilih Tahun", yang memungkinkan pengguna untuk memilih dan menampilkan data pemetaan dari tahun tertentu. Seperti yang ditampilkan pada Gambar 14.



Gambar 6. Halaman Pemetaan

3.6 Hasil Clustering

Berdasarkan hasil perhitungan *clustering* sistem menggunakan metode *K-Means ++* dihasilkan jumlah anggota untuk tahun 2024 sebanyak 36 anggota yaitu : 19 anggota untuk *Cluster* Sangat Rawan, 10 anggota untuk *Cluster* Rawan dan 7 anggota untuk *Cluster* Tidak Rawan. Seperti yang ditunjukkan pada Tabel 7.

Tabel 15. Hasil Clustering

Tanggal Kejadian	Kecamatan	Desa	Jenis Kekerasan	Korban	Umur	Tempat Kejadian	Cluster
25 Jan 2024	Aesesa	Danga	Seksual	Perempuan	22	Rumah	Sangat Rawan
03 Feb 2024	Aesesa	Aeramo	Seksual	Anak	14	Kebun	Sangat Rawan
18 Feb 2024	Aesesa	Aeramo	Seksual	Anak	17	Sawah	Sangat Rawan
.....					...		
06 Desember 2024	Nangaroro	Nataute	Fisik	Perempuan	29	Rumah	Tidak Rawan

3.7 Pengujian Metode K-Means++

a. Silhouette Score

Untuk menguji hasil dari *clustering* dapat dilakukan dengan menggunakan metode *Sillhouette Score*. Pada pengujian ini peneliti menggunakan 36 data kasus kekerasan pada tahun 2024. Berdasarkan hasil dari pengujian *Sillhouette Score* didapatkan nilai 0,6633 yang mengindikasikan bahwa kualitas *clustering* berada pada kategori baik, meskipun masih ada potensi untuk perbaikan dalam pemisahan antar *cluster*.

b. Davies-Bouldin Index (DBI)

DBI adalah metrik untuk mengevaluasi model klusterisasi dengan menghitung perbandingan antara jarak di dalam kluster (intra-cluster distance) dan jarak antar kluster (inter-cluster distance). Prinsip utamanya adalah semakin kecil nilai DBI, semakin baik kualitas klusternya (kluster padat di dalam dan terpisah jauh antar kelompok). Untuk menguji hasil dari clustering dapat dilakukan dengan menggunakan metode Davies-Bouldin Index (DBI). Pada pengujian ini peneliti menggunakan 36 data kasus kekerasan pada tahun 2024. Berdasarkan hasil dari pengujian Davies-Bouldin Index (DBI) didapatkan nilai 0.4520 yang mengindikasikan bahwa kualitas clustering berada pada kategori lebih baik, meskipun masih ada potensi untuk perbaikan dalam pemisahan antar cluster.

4. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian dan implementasi metode K-Means++ untuk pengelompokan wilayah rawan kekerasan anak dan perempuan di Kabupaten Nagekeo, dapat disimpulkan bahwa penerapan algoritma ini mampu memberikan landasan objektif dalam memetakan tingkat kerawanan wilayah. Penggunaan inisialisasi pusat kluster (*centroid*) pada K-Means++ terbukti menghasilkan pengelompokan yang lebih stabil dan optimal dibandingkan dengan metode K-Means standar. Hal ini divalidasi melalui evaluasi internal dengan perolehan nilai *Silhouette Score* sebesar 0.6633 yang menunjukkan struktur kluster yang baik (*good structure*), serta didukung oleh nilai *Davies-Bouldin Index* (DBI) yang lebih rendah dimana mengindikasikan kepadatan dan keterpisahan kluster yang lebih akurat. Penerapan parameter Jenis Kekerasan dan Tempat Kejadian terbukti menghasilkan pemetaan kluster (Sangat Rawan, Rawan, dan Tidak Rawan) yang lebih spesifik dan mendalam dibandingkan dengan metode sebelumnya yang hanya menghitung total kasus secara umum. Temuan ini memberikan wawasan baru bagi Dinas PMD-P3A dalam mengidentifikasi pola kerawanan di Kabupaten Nagekeo secara lebih presisi. Untuk pengembangan selanjutnya, peneliti menyarankan dilakukannya *fine-tuning* pada parameter model dan peningkatan teknik *preprocessing* data. Selain itu, pengembangan sistem ke platform *mobile* (Android) dapat dipertimbangkan untuk meningkatkan aksesibilitas dan pelaporan data lapangan secara langsung.

DAFTAR PUSTAKA

- [1] R. Fauziah and A. I. Purnamasari, "Implementasi Algoritma K-Means pada Kasus Kekerasan Anak dan Perempuan Berdasarkan Usia," *Hello World J. Ilmu Komput.*, vol. 2, no. 1, pp. 34–41, 2023, doi: 10.56211/helloworld.v2i1.232.
- [2] N. Adawiyah, N. Sulistiyowat, and O. Jajuli, "Klusterisasi Kasus Kekerasan Terhadap Anak dan Perempuan Berdasarkan Algoritma K-Means," *Gener. J.*, vol. 5, no. 2, pp. 69–80, 2021.
- [3] Z. Hasan, A. Firly, A. P. Utami, and D. E. Sari, "Perlindungan Hukum Terhadap Perempuan Korban Kekerasan Dalam Rumah Tangga," *J. Hukum, Polit. Dan Ilmu Sos.*, vol. 2, no. 2, pp. 103–113, 2023, doi: 10.55606/jhps.v2i2.1607.
- [4] W. Agus Lestari, K. Paranita Kartika, and S. Nur Budiman, "Klusterisasi Siswa Berdasarkan Hasil Belajar Menggunakan K-Means Berbasis Web (Studi Kasus : Tk. Prima Insan Sholeh Talun)," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 6, no. 1, pp. 9–16, 2021, doi: 10.36040/jati.v6i1.4261.
- [5] A. N. Alifia, A. F. Setiawan, and D. Rudhistiar, "Penerapan Algoritma K-Means Clustering dalam Peringatan Dini Resiko Kegagalan Siswa Pada Mata Pelajaran Bahasa Indonesia," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 8, no. 2, pp. 1174–1181, 2024, doi: 10.36040/jati.v8i2.9075.
- [6] R. Hidayat, F. Laksana, and Y. Dewi, "Pengelompokan Jumlah Kekerasan Terhadap Anak Menurut Kecamatan Di Kabupaten Banyumas Menggunakan Penerapan Algoritma K-Means Clustering," *Indexia*, vol. 5, no. 02, pp. 123-135, 2023, doi: 10.30587/indexia.v5i02.6404.

- [7] C. Dinatingrat, B. Priyatna, E. Novalia, and S. S. Hilabi, “Klasterisasi Kasus Kekerasan Berdasarkan Jenis Lokasi Kejadian di Jawa Barat Menggunakan Algoritma K-Means,” *J. Minfo Polgan*, vol. 14, no. 1, pp. 518–528, 2025, doi: 10.33395/jmp.v14i1.14760.
- [8] N. M. Fithryani, *et al.*, “Algoritma K-Means untuk Meningkatkan Segmentasi Pola Kekerasan,” *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 13, no. 1, pp. 997–1003, 2025, doi: 10.23960/jitet.v13i1.5795.
- [9] M. Orisa, “Optimasi Cluster pada Algoritma K-Means,” *Pros. SENIATI*, vol. 6, no. 2, pp. 430–437, Jul. 2022, doi: 10.36040/seniati.v6i2.5034.
- [10] U. F. Laili, C. Umatin, and M. U. Ridwanulloh, “Analisis Potensial Drop Out Mahasiswa Dengan K-Means ++ Clustering dalam Upaya Peningkatan Kualitan IAIN Kediri Infor Artikel Abstrak,” *Paedagoria: Jurnal Kajian, Penelitian, dan Pengembangan Kependidikan*, vol. 14, no. 2, pp. 145–153, 2023, doi: 10.31764/paedagoria.v14i2.14077.
- [11] N. Nugroho and F. D. Adhinata, “Penggunaan Metode K-Means dan K-Means++ Sebagai Clustering Data Covid-19 di Pulau Jawa,” *Teknika*, vol. 11, no. 3, pp. 170–179, 2022, doi: 10.34148/teknika.v11i3.502.
- [12] M. Ferdiansyah and U. Chotijah, “Implementasi Algoritme K-Means ++ Untuk Clustering Penjualan Bahan Bangunan,” *J. Ilm. Tek. Inform. dan Komun.*, vol. 4, no. 1, pp. 181–193, 2024.
- [13] C. Ramdani and N. Safadila, “Analisis Data Akademis dengan Menerapkan Algoritme K-Means dan K-Means++,” vol. 1, no. 4, pp. 155–160, 2022, doi: 10.20895/ledger.v1i4.918.
- [14] Y. Hasan, “Pengukuran Silhouette Score Dan Davies-Bouldin Index Pada Hasil Cluster K-Means dan Dbscan,” *J. Inform. dan Tek. Elektro Terap.*, vol. 12, no. 3S1, 2024, doi: 10.23960/jitet.v12i3s1.5001.
- [15] M. A. K. D. K-Medoids, “Penerapan Data Mining Dalam Pemilihan Produk Unggulan dengan Metode Algoritma K-Means dan K-Medoids,” *Jurnal Media Informatika Budidarma*, vol. 6, no. 1, pp. 27–34, 2022, doi: 10.30865/mib.v6i1.3294.