KOMPARASI ALGORITMA *K-MEANS* DAN *K-MEDOIDS* DALAM ANALISIS PENGELOMPOKAN DAERAH RAWAN KRIMINALITAS DI INDONESIA

Anis Hoerunnisa [1], Gifthera Dwilestari [2], Fatihanursari Dikananda [3], Heliyanti Sunana [4], Denni Pratama [5]
[1,4] Teknik Informatika,[2,3] Sistem Informasi, [5] Komputerisasi Akuntansi, STMIK IKMI Cirebon
Jalan Perjuangan No.10 B, Kecamatan Kesambi, Kota Cirebon
anishoerunnisa29@gmail.com

ABSTRAK

Kriminalitas sebagai fenomena yang kompleks, memiliki dampak serius pada stabilitas dan ketertiban sosial, sertamemengaruhi aspek sosial dan ekonomi suatu wilayah. Pemahaman mendalam terhadap area-area rentan terhadap tindak kriminal menjadi esensial dalam upaya mencegah dan menangani kejahatan. Penelitian ini bertujuan eksploratif untuk membandingkan efektivitas dua metode clustering, yaitu *K-Means* dan *K-Medoids*, dalam menganalisis daerah-daerah rentan kriminal di Indonesia.Metode penelitian melibatkan langkah-langkah seperti pengumpulan data kriminalitas dari sumber terpercaya, penyortiran, dan pengolahan data untuk membentuk dataset yang akurat. Dengan menggunakan pendekatan *Knowledge Discovery in Databases* (KDD), penelitian ini menerapkan algoritma *K-Means*, pengelompokan dalam analisis data yang bertujuan mengelompokkan data menjadi beberapa klaster, dan *K-Medoids*, variasi *K-Means* yang menggunakan medoid sebagai representasi klaster. Hasil analisis menunjukkan bahwa kedua algoritma dapat digunakan untuk mengelompokkan daerah rawan kriminalitas di Indonesia, dengan K-Means membentuk 5 cluster dan *K-Medoids* membentuk 2 cluster. Evaluasi menggunakan *Davies-Bouldin Index* (DBI) menunjukkan bahwa *K-Means* lebih optimal (DBI=0,463) dibandingkan *K-Medoids* (DBI=1,089). Sebagai saran untuk pengembangan selanjutnya, disarankan untuk mempertimbangkan penggunaan algoritma clustering yang berbeda, melakukan perbandingan dengan metode lain, dan menambahkan kriteria atau variabel yang relevan untuk meningkatkan kualitas analisis clustering. Diharapkan hasil ini dapat memberikan kontribusi signifikan dalam upaya pencegahan dan penanggulangan kriminalitas di Indonesia.

Kata kunci : *Data Mining,Clustering, K-Means,K-Medoids, Kriminalitas*

1. PENDAHULUAN

Kriminalitas, sebagai pelanggaran hukum dengan potensi dampak negatif, sering menimbulkan kerugian luas dan menempatkan pelakunya sebagai pelaku kejahatan. Di Indonesia, seperti negara lain, berbagai jenis kriminalitas, termasuk pembunuhan, penganiayaan, pemerkosaan, pencurian, peredaran narkoba, penipuan, dan korupsi, sering terjadi [1].

Dalam studi Tampubolon tahun 2021, terkait tingginya tingkat kejahatan di Indonesia, perlu dilakukan pengelompokan wilayah potensial rawan kejahatan berdasarkan provinsi. Ini merupakan langkah penting untuk mendukung kepolisian dalam menentukan apakah pengawasan tambahan diperlukan. Oleh karena itu, diperlukan pengolahan data menggunakan metode *data mining*, khususnya analisis klaster, dengan fokus pada algoritma *K-Means* dan *K-Medoids Clustering* [2].

Penelitian Anggoro dan tim pada tahun 2019 menggunakan *K-Means Clustering*, *K-Medoids Clustering*, dan *Kernel Density* untuk klasifikasi wilayah rawan curanmor di Kota Semarang [3]. Hasil uji menunjukkan perbedaan *antara K-Means* dan *K-Medoids*, sementara *Kernel Density* memiliki nilai verifikasi tinggi, khususnya pada kerawanan tinggi dan sedang, mencapai 82,35%. Penelitian lain seperti yang dilakukan oleh [4] di bidang ekonomi menggunakan *K-Means* untuk mengelompokkan pertumbuhan berdasarkan sektor usaha di Kota Surabaya.

Penelitian ini bertujuan membandingkan efisiensi dan akurasi algoritma *clustering K-Means* dan *K-Medoids* dalam mengidentifikasi daerah rawan kriminalitas di Indonesia. Hasilnya diharapkan memberikan wawasan bagi pihak berwenang untuk merancang strategi pencegahan dan penanggulangan kriminalitas serta memperkaya literatur analisis keamanan dan kriminalitas.

Analisis klaster merupakan teknik pengelompokan objek berdasarkan karakteristiknya. Salah satu metode umumnya adalah *non-hierarki partitioning*, yang membagi objek data ke dalam k klaster, dengan dua metode populer: *k-means* dan *k-medoids*. Algoritma *K-Means* memiliki sensitivitas terhadap pencilan karena menggunakan nilai rata-rata sebagai pusat kelompok, sedangkan *K-Medoids* menggunakan medoid untuk mengatasi potensi masalah ini. *Medoid* merujuk pada objek paling dekat dengan pusat klaster [5].

Penelitian ini menggunakan data dari website Open Data Jabar yaitu *https://opendata.jabarprov.go.id/id* dengan data yang bersumber dari Badan Pusat Statistik (BPS) Jawa Barat ntuk menganalisis tindak kriminalitas berdasarkan daerah dengan algoritma *K-Means* dan *K-Medoids*. Tujuannya adalah mengetahui algoritma pengelompokan wilayah yang paling efektif guna membantu kepolisian dalam pengambilan keputusan dan perhatian khusus terhadap daerah tertentu.

2. TINJAUAN PUSTAKA

2.1. Penelitian Terdahulu

Dalam jurnal "Penerapan Algoritma *K-Means* dan *K-Medoids Clustering* untuk Mengelompokkan Tindak Kriminalitas Berdasarkan Provinsi" [2], dibahas penggunaan algoritma *clustering K-Means dan K-Medoids* untuk mengelompokkan tindak kriminal berdasarkan provinsi di Indonesia, memberikan penilaian apakah suatu daerah memerlukan pengawasan tambahan. Hasil dari kedua algoritma ini tergantung pada data yang sedang diproses, dan jurnal ini menunjukkan konsistensi hasil pengelompokan menggunakan algoritma *K-Medoids* dengan perhitungan manual menggunakan algoritma *K-Means dan K-Medoids.*

Mengevaluasi preferensi rasa *Thai tea* menggunakan algoritma *K-means* dan *K-medoids* melalui penelitian ini, yang melibatkan langkah-langkah penambangan data seperti pembersihan, transformasi, dan integrasi data. Kedua algoritma menghasilkan klaster yang identik, merekomendasikan *Thai tea* asli dengan susu dan greentea dengan susu sebagai pilihan favorit. Data penelitian ini diperoleh melalui survei pelanggan di Cirebon. Selain itu, paper ini mencantumkan sumber-sumber terkait penambangan data, pengeringan beku, dan aplikasi penambangan data di berbagai bidang [6].

Penerapan *data mining* dengan menggunakan metode K-Medoids dapat diimplementasikan untuk mengklasifikasikan bidang pekerjaan utama dari populasi yang berusia 15 tahun ke atas. Hasil pengelompokan menunjukkan bahwa ada 14 bidang pekerjaan utama dalam kelompok tinggi dan 3 bidang pekerjaan utama dalam kelompok rendah. Hal ini dapat membantu meningkatkan kinerja pemerintah dalam mengelola bidang pekerjaan utama penduduk secara akurat dan efektif [7].

Penelitian Dewi dan tim pada tahun 2019 mengevaluasi penerapan metode K-Means untuk mengklasifikasikan tingkat kejahatan berdasarkan provinsi dengan menggunakan dataset tahun 2012 dari 31 provinsi. Metode ini mengelompokkan data kejahatan menjadi dua klaster, yaitu tinggi dan rendah, memberikan hasil yang dapat menjadi panduan strategis bagi pemerintah dalam memprioritaskan provinsi dengan tingkat kejahatan tinggi. Proses data mining mencakup pengumpulan, pengolahan, pengelompokan, dan analisis data, dengan iterasi K-Means untuk mencapai pengelompokan data yang konsisten. Hasil akhir menunjukkan klaster tingkat tinggi (C1) terdiri dari 8 provinsi, sedangkan klaster tingkat rendah (C2) terdiri dari 23 provinsi [8].

2.2. *Data Mining*

Data Mining adalah metode analisis data, statistik, dan proses untuk memperoleh informasi berharga dari gudang basis data besar. Artinya, ini mengekstrak informasi baru dari data yang melimpah, memberikan kontribusi pada pengambilan keputusan (*Knowledge Discovery*). Pendekatan dasarnya adalah merangkum data dan mengekstrak informasi yang sebelumnya tidak diketahui namun masuk akal. *Data mining* terbagi menjadi dua kategori utama: *Descriptive Mining* untuk mengidentifikasi karakteristik penting dari data, dan *Predictive Mining* untuk memahami pola muncul dari dataset tertentu. Dalam proses *Knowledge Discovery in Database* (KDD), data mining memegang peran sentral, membantu menganalisis data elektronik dan melakukan pencarian otomatis menggunakan komputer [9].

2.3. *Clustering*

Prinsip dasar *clustering* melibatkan penerapan teknik untuk mengidentifikasi dan mengelompokkan data berdasarkan kesamaan karakteristiknya, seperti tingkat kemiripan. *Clustering* merupakan metode *unsupervised* dalam data *mining*, di mana tanpa pelatihan dan panduan eksternal, serta tanpa memerlukan output yang ditargetkan. Dalam konteks data *mining*, terdapat dua teknik *clustering* yang umumnya digunakan, yaitu *hierarchical clustering* dan *nonhierarchical clustering*[10].

2.4. *K-Means Clustering*

K-Means adalah salah satu algoritma *Clustering* yang termasuk dalam kategori *Unsupervised Learning* yang berfungsi untuk mengelompokkan data menjadi beberapa kelompok dengan prinsip partisi. Setiap kelompok memiliki titik pusat (*centroid*) yang mewakili kelompok tersebut [9].

Prinsip fundamental dari algoritma *K-Means* adalah mengorganisir k partisi, yang dikenal sebagai centroid atau mean, dari suatu himpunan data yang diberikan.

2.5. *K-Medoids Clustering*

Algoritma *K-Medoids* merupakan suatu teknik clustering yang serupa dengan algoritma *K-Means*. Perbedaan utama terletak pada pendekatan penentuan pusat klaster. Algoritma *K-Medoids* menggunakan medoid (objek individu) sebagai pusat klaster, sedangkan *K-Means* menggunakan nilai rata-rata objek dalam klaster. Langkah-langkah algoritma *K-Medoids* telah dijelaskan oleh [7].

1. Algoritma dimulai dengan menginisialisasi k pusat klaster, sesuai dengan jumlah klaster yang diinginkan.
2. Objek data dialokasikan ke klaster terdekat menggunakan *Euclidean Distance* (dijelaskan dalam persamaan 1).
3. Objek-objek klaster dipilih acak sebagai kandidat medoids baru.
4. Jarak antara objek dalam klaster dengan kandidat *medoids* dihitung.
5. Total simpangan (S) dihitung sebagai selisih total jarak baru dengan total jarak sebelumnya. Jika $S < 0$, dilakukan pertukaran objek untuk membentuk *k medoids* baru.

6. Langkah 3-5 diulang hingga medoids tidak berubah, menghasilkan klaster dan anggota-anggotanya.

2.6. *Davies Bouldin Index (DBI)*

Indeks Davies Bouldin (DBI) digunakan untuk mengevaluasi hasil klaster dengan cara mengukur validitas internal melalui penilaian jumlah kriteria dari dataset. Tujuannya adalah memaksimalkan jarak antar objek dalam klaster dan meminimalkan jarak antar klaster. Proses pengelompokan memilih algoritma *K-Means* atau *K-Medoids* dengan nilai DBI minimum sebagai hasil optimal dan jumlah klaster terbaik [11].

2.7. KDD

Menurut [12], istilah data mining *dan Knowledge Discovery in Database* (KDD) sering digunakan secara bergantian untuk merujuk pada proses menggali informasi tersembunyi dalam basis data luas. Tahapan integral dalam KDD mencakup:

a. Data *Selection*: Pemilihan data dari basis data operasional sebelum proses KDD dimulai.
b. *Pre-processing/Cleaning*: Pembersihan data, termasuk penghapusan duplikat dan perbaikan kesalahan.
c. *Transformation*: Transformasi data yang dipilih untuk sesuai dengan proses data mining.
d. Data *Mining*: Pencarian pola atau informasi menarik dalam data menggunakan metode tertentu.
e. *Interpretation/Evaluation*: Penyajian dan pemeriksaan pola informasi untuk diinterpretasikan dengan mudah.

2.8. Kriminalitas

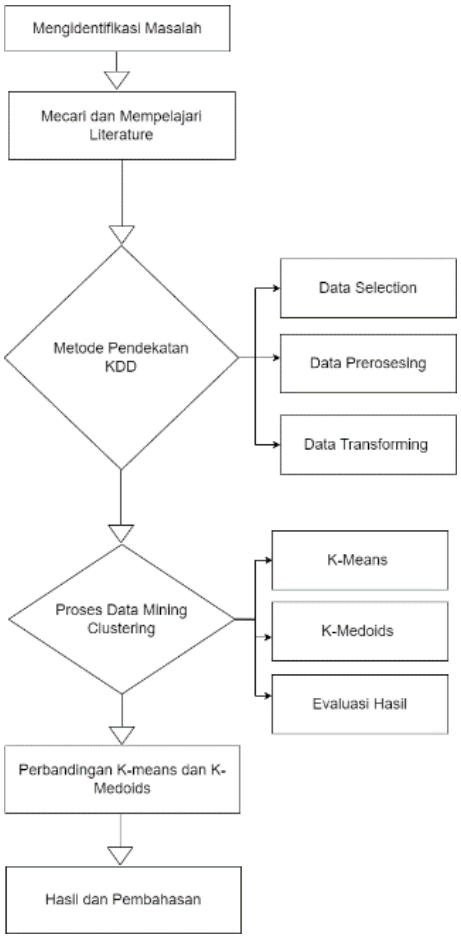
Kriminalitas dapat didefinisikan sebagai tindakan yang menyalahi norma hukum, agama, dan sosial, serta menimbulkan kerugian bagi individu atau kelompok lainnya. Fenomena ini merupakan permasalahan yang sering terjadi dalam lingkungan masyarakat, memerlukan perhatian karena dapat merugikan berbagai kepentingan dan menimbulkan dampak negatif, seperti ketidakamanan, kecemasan, kepanikan, dan ketakutan [8].

2.9. *Rapid Miner*

Rapid Miner merupakan perangkat lunak terbuka yang dapat diakses oleh semua orang, *RapidMiner* menyediakan lebih dari 500 operator data mining, termasuk fungsi input, output, preprocessing data, dan visualisasi data [13]. Dalam penelitian ini, *RapidMiner* 10.2 digunakan untuk menguji dan memvalidasi hasil perhitungan manual algoritma *K-Means* dan *K-Medoids*.

3. METODE PENELITIAN

Pada bagian ini akan dijelaskan metode penelitian yang digunakan dalam penelitian ini, metode dapat dilihat di Gambar 1



Gambar 1. Metode penelitian

a. Identifikasi Masalah: Menetapkan fokus pada permasalahan kriminalitas di Indonesia.
b. Studi Literatur: Pencarian dan kajian literatur dari artikel dan jurnal ilmiah tentang metode KDD, *Data Mining*, *K-Means*, *K-Medoids Clustering*, serta sumber pendukung lainnya.
c. Metode Pendekatan KDD: Pra-pemrosesan yang mencakup *Data Selection*, *Data Praocessing*, dan *Data Transformation*.
d. Pengolahan Data dengan Metode *Clustering*: Pengolahan data tindak kriminalitas di Indonesia tahun 2021 dari Badan Pusat Statistik dan Open Data Jabar.
e. Analisis menggunakan Metode Algoritma *K-Means dan K-Medoids*.
f. Evaluasi Hasil: Evaluasi *clustering* dengan menggunakan DBI untuk memvalidasi perbandingan antara algoritma *K-Means dan K-Medoids*.
g. Hasil dan Pembahasan: Analisis data hasil Metode *K-Means Clustering* digunakan untuk mendapatkan informasi tindak kriminalitas dengan cepat dan akurat dalam pengelompokkan daerah.

4. HASIL DAN PEMBAHASAN

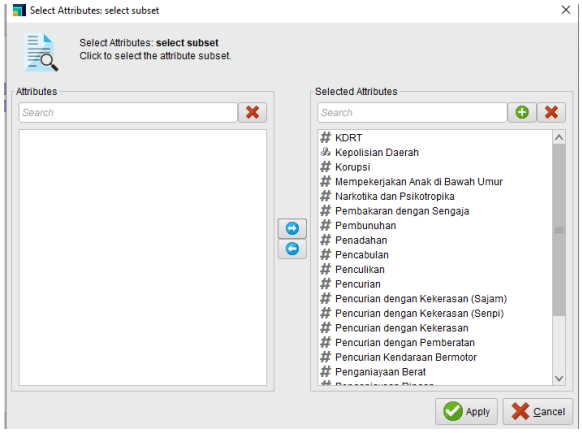
Dalam bagian ini, akan dibahas langkah-langkah yang dihasilkan selama proses pengklasteran kejadian kriminal di Indonesia pada tahun 2021.

4.1. *Data Set*

Data penelitian diperoleh dari situs web Open Data Jabar yaitu *https://opendata.jabarprov.go.id/id* yang mengambil sumber dari Badan Pusat Statistik dan mengidentifikasi data sekunder yang relevan dengan topik penelitian data yang digunakan merupakan data tindak kriminal di Indonesia pada tahun 2021.

4.2. *Data Selection*

Pemilihan Data (*Data Selection*): Atribut yang relevan untuk penelitian ini dipilih, seperti Pembunuhan, Pengeroyokan Berat, Penganiayaan Ringan, KDRT, Perkosaan, Pencabulan, Penculikan, Mempekerjakan Anak Dibawah Umur, Pencurian Dengan Kekerasan, Pencurian Dengan Kekerasan (Senpi), Pencurian Dengan Kekerasan (Sajam), Pencurian, Pencurian Dengan Pemberatan, Pencurian Kendaraan Bermotor, Penadahan, Pengrusakan/Pengancuran Barang, Pembakaran Dengan Sengaja, Narkotika Dan Psikotropika, Penipuan/Perbuatan Curang, Penggelapan, Korupsi, Terhadap Ketertiban Umum.

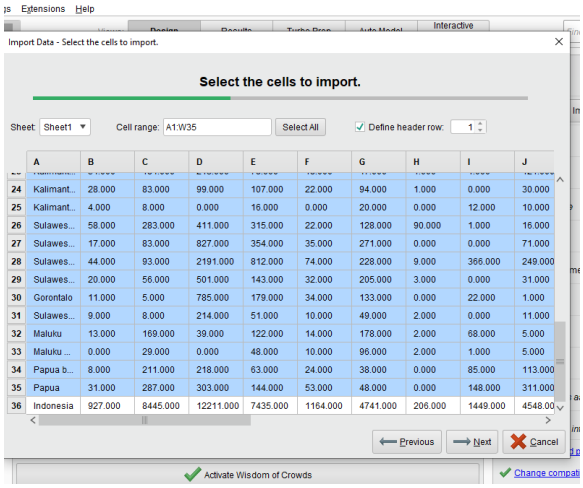


Gambar 2. *Data selection*

Gambar 2 menunjukkan tahap seleksi data, di mana data yang diperlukan dipilih dan disesuaikan dengan kebutuhan penelitian. *Dataset* yang digunakan terdiri dari 23 atribut dengan total 34 data pada setiap atribut, disusun untuk pemilihan atribut dalam pengolahan data.

4.3. *Data Preprocessing*

Pada tahap *Preprocessing*, di mana data yang tidak lengkap atau kosong dibersihkan sebelum melanjutkan ke langkah berikutnya.

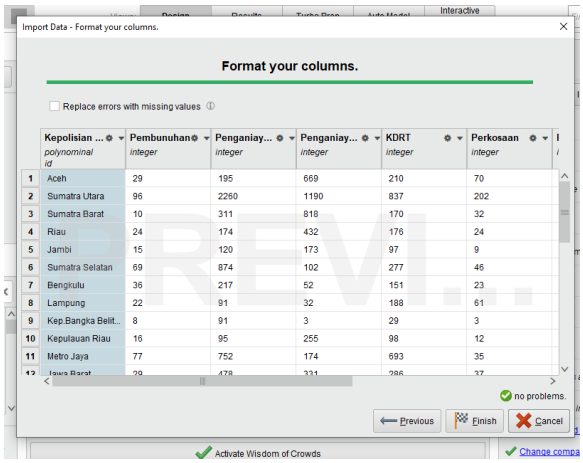


Gambar 3. *Data preprocessing*

Gambar 3 menunjukkan proses pembersihan data melibatkan penghapusan data untuk Indonesia, karena fokus hanya pada Nama Keposilian Daerah atau jumlah provinsi di Indonesia.

4.4. *Data Transformation*

Dalam tahap ini, dilakukan langkah transformasi data ke dalam format yang sesuai untuk diolah oleh teknik data mining. Hal ini dilakukan dengan tujuan untuk menyelaraskan data yang akan diproses oleh algoritma dan alat bantu yang digunakan dalam penelitian, yaitu *RapidMiner.*



Gambar 4. *Data transformation*

Gambar 4 menunjukan proses transformasi data di *Rapid Miner*.

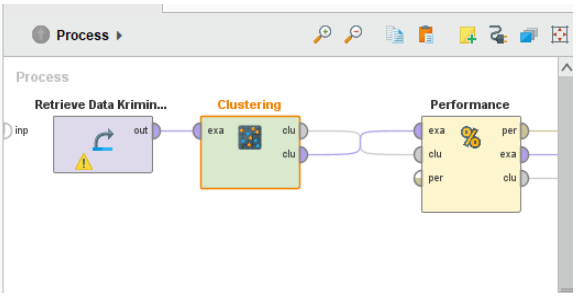
4.5. Melakukan *Data Mining* dengan aplikasi Rapid Miner ver.10.2

Melakukan proses *data mining* dengan algoritma *K-Means* atau *K-Medoids* menggunakan *RapidMiner* adalah langkah penting dalam penelitian. Konfigurasi algoritma melibatkan parameter seperti jumlah *cluster* (K), metode inisialisasi pusat *cluster*, dan pengaturan lainnya. Proses *clustering* pada 34 data tindak kriminal di Indonesia menghasilkan kelompok data daerah kriminal, memungkinkan identifikasi tingkat tindak

kriminalitas rendah, sedang, dan tinggi. Implementasi algoritma *K-Means* dan *K-Medoids* menggunakan RapidMiner dilakukan dalam proses *clustering* tersebut.

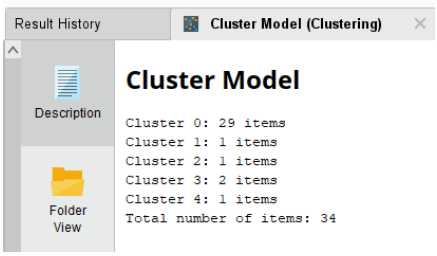
4.6. Proses *K-Means Clustering*

Tahap ini adalah tahap pemrosesan data dan hasil praprosesing menggunakan algoritma *K-Means*



Gambar 5. Proses *k-means clustering*

Pada Gambar 5, tampak proses *k-means clustering* yang menggunakan operator *Clustering* dan operator *Performance* untuk mengukur persentase nilai *Davies-Bouldin Inde*x (DBI). Proses ini memberikan pemahaman mendalam tentang efisiensi dan validitas hasil pengelompokan oleh algoritma *k-means*.



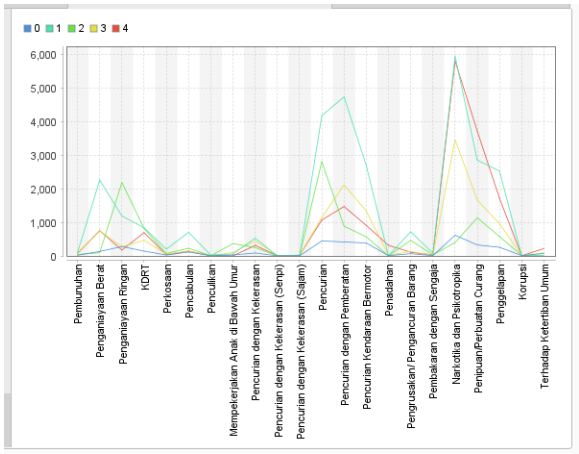
Gambar 6. Hasil *k-means* untuk 5 *cluster*

Gambar 6 menggambarkan hasil dari proses k-means, yang mencakup 29 daerah dalam *Cluster*_0, 1 daerah dalam *Cluster*_1, 1 daerah dalam *Cluste*r_2, 2 daerah dalam *Cluster*_3, dan 1 daerah dalam *Cluster*_4.

Tabel 1. Nilai *Centroid*

Attribut	*Cluster*_ 0	*Cluster*_ 1	…	*Cluster*_ 4
Pembunuhan	20.276	96	…	77
KDRT	143.621	837	…	693
Pemerkosaan	25.059	202	…	35
Pencabulan	116.138	702	…	143
Penculikan	5.138	28	…	10
…	…	…	…	…
Pencurian	446.585	4183	…	1070
Penggelapan	257.138	2531	…	1717
Penadahan	8.690	9	…	316
Korupsi	10.690	12	…	6

Tabel 1 merupakan gambar nilai *centroid* dari hasil *k-means clustering*.

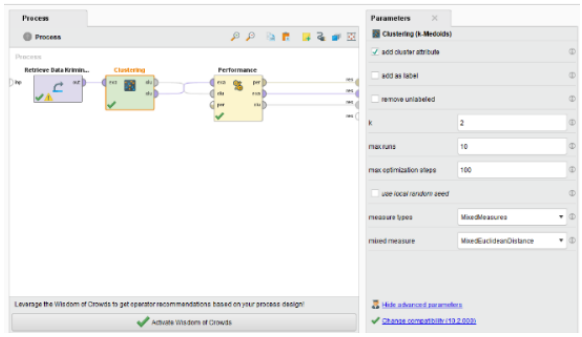


Gambar 7. *Plot* hasil *k-means*

Gambar 7 menunjukan plot hasil proses *K-Means Clustering* menggunakan *RapidMiner*.

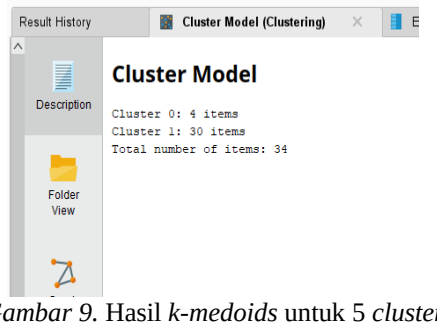
4.7. Proses *K-Medoids Clustering*

Tahap ini adalah tahap pemrosesan data dan hasil praprosesing menggunakan algoritma *K-Medoids*



Gambar 8. Proses *k-medoids clustering*

Gambar 8 menggambarkan langkah-langkah *k-medoids clustering* menggunakan aplikasi *RapidMiner*. *Data set* yang telah diproses melalui tahap *preprocessing* dimasukkan ke dalam aplikasi tersebut. Proses selanjutnya melibatkan penggunaan algoritma *K-medoids* untuk mengolah data, sementara fitur *Performance* ditambahkan untuk mengevaluasi hasil dari proses tersebut.



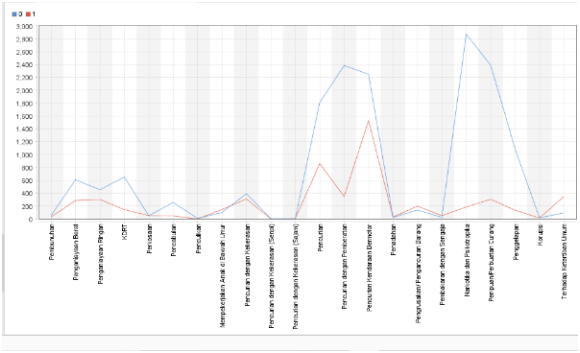
Gambar 9. Hasil *k-medoids* untuk 5 *cluster*

Gambar 4.9 Menunjukan jumlah anggota *cluster*_0 adalah 4 Daerah dan anggota *Cluster*_1 berjumlah 30 daerah.

Tabel 2. Nilai *centroid k-medoids*

Attribut	*Cluster*_0	*Cluster*_1
Pembunuhan	53	31
KDRT	651	144
Pemerkosaan	51	53
Pencabulan	257	48
Penculikan	9	0
…	…	…
Pencurian	1811	861
Penggelapan	1088	135
Penadahan	24	30
Korupsi	18	16

Tabel 2 menujukan nilai *centroid* yang dipeoleh dari hasil proses *k-medoids clustering*



Gambar 10. *Plot* hasil proses *k-medoids*

Gambar 4.14 menunjukan *plot* hasil proses *k-medoids clustering* menggunakan *RapidMiner*.

4.8. Evaluasi Algoritma *K-Means* dan *K-Medoids Clustering*

Berikut merupakah hasil evaluasi proses *K-Means* dan *K-Medoids* menggunakan DBI

Tabel 3. DBI hasil evaluasi *k-means*

Nilai DBI	
k-2	0.668
k-3	0.549
k-4	0.499
k-5	0.463
k-6	0.668

Pada Tabel 3. Terlihat bahwa nilai BDI dengan 5 klaster lebih kecil disbanding klaster lainnya.

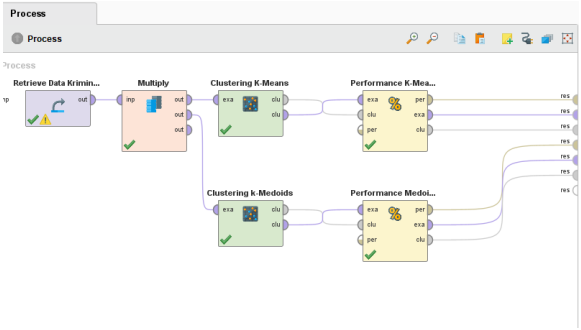
Tabel 4. DBI hasil evaluasi *k-medoids*

Nilai DBI	
k-2	1.089
k-3	1.747
k-4	1.378
k-5	1.448
k-6	1.259

Pada Tabel 4. Terlihat bahwa nilai BDI dengan 2 klaster lebih kecil disbanding klaster lainnya. Perbandingan Algoritma *K-Means* dan *K-Medoids*

4.9. *Clustering*

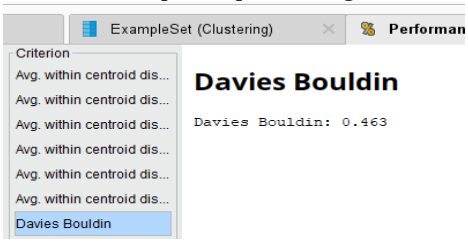
Pada tahap ini dilakukan perbandingan kedua algoritma menggunakan DBI pada aplikasi *RapidMiner*.



Gambar 11. Proses perbandingan di rapidminer

Gambar 11 menunjukan proses perbandingan dua algoritma pada aplikasi *RapidMiner* menggunakan operator *Multiply*, *Clustering K-Means*, *Clustering K-Medoids* dan *Performance.*

Berikut ini hasil proses perbandingan *K-Means*



Gambar 12. Hasil perbandingan *k-means*

Gambar 12. Menunjukan nilai DBI hasil perbandingan algoritma *K-Means* sebesar 0.463. Selanjutnya merupakan hasil proses perbandingan *K-Medoids*

Gambar 13. Hasil perbandingan *k-medoids*

Gambar 13. Menunjukan nilai DBI hasil perbandingan algoritma *K-Medoids* sebesar 1.089

4.10. Pembahasan

Pada proses *Clustering K-Means* diperoleh hasil pengelompokan yakni terdapat 5 cluster dengan anggota tiap clusternya sebagai berikut :

Tabel 5. Anggota *cluster k-means*

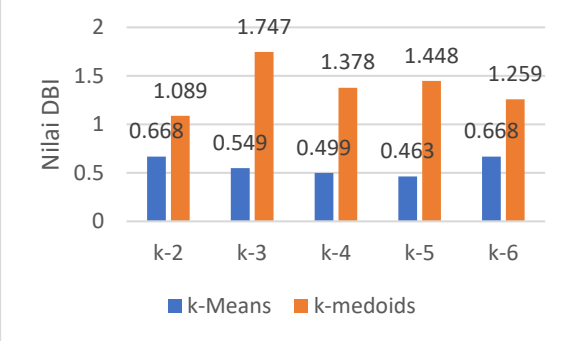
Cluster	**Anggota *Cluster***
*Cluster*_0	Aceh, Sumatra Barat, Riau, Jambi, Bengkulu, Lampung, Kep. Bangka Belitung, Kepulauan Riau, Jawa Barat, Jawa Tengah, DI Yogyakarta, Banten, Bali, Nusa Tenggara Barat, Nusa Tengggara Timur, Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Timur, Kalimantan Utara, Sulawesi Utara, Sulawesi Tengah, Sulawesi Tenggara, Gorontalo, Sulawesi Barat, Maluku, Maluku Utara, Papua Barat, Papua
*Cluster*_1	Sumatra Utara
*Cluster*_2	Sulawesi Selatan
*Cluster*_3	Sumatra Selatan, Jawa Timur
*Cluster*_4	Metro Jaya

Kemudian pada tahap *Clustering K-Medoids* diperoleh hasil 2 *cluster* dengan anggota *cluster* dapat dilihat di tabel berikut,

Tabel 6. Anggota *cluster k-medoids*

Cluster	**Anggota *Cluster***
*Cluster*_0	Sumatra Utara, Sumatra Selatan, Metro Jaya, Jawa Timur
*Cluster*_1	Aceh, Sumatra Barat, Riau, Jambi, Bengkulu, Lampung, Kep. Bangka Belitung, Kepulauan Riau, Jawa Barat, Jawa Tengah, DI Yogyakarta, Banten, Bali, Nusa Tenggara Barat, Nusa Tengggara Timur, Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Timur, Kalimantan Utara, Sulawesi Utara, Sulawesi Tengah, Sulawesi Selatan, Sulawesi Tenggara, Gorontalo, Sulawesi Barat, Maluku, Maluku Utara, Papua Barat, Papua

Selanjutnya merupakan hasil perbandingan *K-Means* dan *K-Medoids*



Gambar 14. *Barchat* hasil perbandingan *k-means dan k-medoids*

Hasil pengelompokan Provinsi dengan algoritma k-means melibatkan penentuan *Davies Bouldin Index* (DBI) untuk setiap klaster. Rekapitulasi DBI menunjukkan DBI k-2 = 0,668, k-3 = 0,549, k-4 = 0,499, k-5 = 0,463, dan k-6 = 0,668. Oleh karena itu, k yang dipilih pada percobaan ketiga (k-5) dengan DBI terkecil, yaitu 0,463, sehingga digunakan 5 klaster.

Pada pengelompokan Provinsi dengan algoritma k-medoids, hasilnya melibatkan penentuan DBI untuk setiap klaster. Rekapitulasi DBI menunjukkan DBI k-2 = 1,089, k-3 = 1,747, k-4 = 1,378, k-5 = 1,448, dan k-6 = 1,259. Kembali dipilih k pada percobaan ketiga (k-2) dengan DBI terkecil, yaitu 1,089, sehingga digunakan 2 klaster.

Berdasarkan perbandingan diatas, bahwa pada eksperimen dengan dataset kecil, *K-Means* lebih efektif daripada *K-Medoids*. Pada dataset tindak kriminal di Indonesia tahun 2021, K-Means memiliki DBI sebesar 0.463, sedangkan *K-Medoids* memiliki nilai evaluasi sebesar 1.089. Analisis ini menunjukkan bahwa hasil pengelompokan *K-Means* lebih sesuai dalam kasus tersebut. Selain itu, pengujian DBI menunjukkan bahwa *K-Means* memiliki nilai lebih rendah, mengindikasikan kualitas klaster yang lebih baik. Oleh karena itu, *K-Means* lebih superior dalam menangani data kecil, di mana nilai DBI yang lebih rendah menandakan kualitas klaster yang lebih baik.

5. KESIMPULAN DAN SARAN

Berdasarkan penelitian dengan tujuan, analisis, dan hasilnya, dapat disimpulkan bahwa algoritma *K-Means* dan *K-Medoids Clustering* dapat mengelompokkan daerah rawan tindak kriminal di Indonesia. Proses *K-Means* menghasilkan 5 Cluster, sementara *K-Medoids* menghasilkan 2 Cluster. Hasil perbandingan menggunakan DBI menunjukkan bahwa *K-Means* lebih optimal (DBI 0,463) dibandingkan *K-Medoids* (DBI 1,089). Untuk pengembangan selanjutnya, disarankan mempertimbangkan penggunaan algoritma *clustering* yang berbeda dan melakukan perbandingan dengan metode lain. Demi meningkatkan analisis *clustering*, disarankan menambahkan kriteria atau variabel yang relevan.

DAFTAR PUSTAKA

[1] S. B. Faradilla, "Komparasi Analisis K-Medoids Clustering Dan Hierarchical Clustering," *(Studi Kasus Data Kriminalitas di Indones. Tahun 2020)*, 2022.

[2] H. D. Tampubolon, S. Suhada, M. Safii, S. Solikhun, and D. Suhendro, "Penerapan Algoritma K-Means dan K-Medoids Clustering untuk Mengelompokkan Tindak Kriminalitas Berdasarkan Provinsi," *J. Ilmu Komput. dan Teknol.*, vol. 2, no. 2, pp. 6–12, 2021, doi: 10.35960/ikomti.v2i2.703.

[3] J. W. Anggoro, M. Awaluddin, and A. L. Nugraha, "Zonasi Daerah pencrian kendaraan bermotor (curanmor)di kota semarang degan menggunakan metode cluster," *J. Geod. Undip*,

vol. 8, no. 4, pp. 225–234, 2019.

[4] N. A. Febriyati, A. Daengs Gs, and A. Wanto, "GRDP Growth Rate Clustering in Surabaya City uses the K-Means Algorithm," *Int. J. Inf. Syst. Technol. Akreditasi*, vol. 3, no. 36, pp. 276–283, 2020.

[5] M. A. Nahdliyah, T. Widiharih, and A. Prahutama, "Metode K-Medoids Clustering Dengan Validasi Silhouette Index Dan C-Index (Studi Kasus Jumlah Kriminalitas Kabupaten/Kota di Jawa Tengah Tahun 2018)," *J. Gaussian*, vol. 8, no. 2, pp. 161–170, 2019, doi: 10.14710/j.gauss.v8i2.26640.

[6] A. Amalia, D. R., Narasatu, R., Faqih, "Perbandingan Hasil Klasifikasi Rasa Minuman Thai Tea yang Paling Digemari Menggunakan K-means dan K-medoids," *Pros. Semin. Nas. Unimus*, vol. 2, pp. 401–407, 2019.

[7] N. Pulungan, S. Suhada, and D. Suhendro, "Penerapan Algoritma K-Medoids Untuk Mengelompokkan Penduduk 15 Tahun Keatas Menurut Lapangan Pekerjaan Utama," *KOMIK (Konferensi Nas. Teknol. Inf. dan Komputer)*, vol. 3, no. 1, pp. 329–334, 2019, doi: 10.30865/komik.v3i1.1609.

[8] S. M. Dewi, A. P. Windarto, I. S. Damanik, and H. Satria, "Analisa Metode K-Means pada Pengelompokan Kriminalitas Menurut Wilayah," *Semin. Nas. Sains Teknol. Inf.*, pp. 620–625, 2019.

[9] H. D. Tampubolon, D. Gultom, L. Y. Hutabarat, F. I. R.H Zer, and D. Hartama, "Penerapan Algoritma K-Means Untuk Mengetahui Tingkat Tindak Kejahatan Daerah Pematangsiantar," *J. Teknol. Inf.*, vol. 4, no. 1, pp. 146–151, 2020, doi: 10.36294/jurti.v4i1.1263.

[10] J. Nasir, "Penerapan Data Mining Clustering Dalam Mengelompokan Buku Dengan Metode K-Means," *Simetris J. Tek. Mesin, Elektro dan Ilmu Komput.*, vol. 11, no. 2, pp. 690–703, 2021, doi: 10.24176/simet.v11i2.5482.

[11] I. . Supriyadi, A., Triayudi, A., Sholihati, "Perbandingan algoritma k-means dengan k-medoids pada pengelompokan armada kendaraan truk berdasarkan produktivitas," vol. 06, pp. 229–240, 2021.

[12] D. Marlina, N. Lina, A. Fernando, and A. Ramadhan, "Implementasi Algoritma K-Medoids dan K-Means untuk Pengelompokkan Wilayah Sebaran Cacat pada Anak," *J. CoreIT J. Has. Penelit. Ilmu Komput. dan Teknol. Inf.*, vol. 4, no. 2, p. 64, 2018, doi: 10.24014/coreit.v4i2.4498.

[13] E. Elisa and A. Annurullah, "Data Mining dalam Menganalisis Faktor Alasan Pemilihan Perumahan," no. September, pp. 43–48, 2020.