

Pemetaan Tingkat Kriminalitas di Indonesia: Analisis Spasial dengan Pendekatan SIG pada Tingkat Provinsi

Ronal Watrianthos*, Sudi Suryadi, Kusmanto, Samsi

Fakultas Ilmu Komputer, Program Studi Teknik Informatika, Universitas Al Washliyah, Rantauprapat, Indonesia

Email: ¹ronal.watrianthos@email.com, ²sudisuryadi28@gmail.com, ³kusnabara03@gmail.com, ⁴samsirst11@gmail.com

(*: koresponden author: ronal.watrianthos@email.com)

Abstrak- Indonesia memiliki Indeks Pembangunan Manusia yang rendah, yang menunjukkan bahwa masih ada pekerjaan yang harus dilakukan untuk meningkatkan kualitas hidup dan kesehatan masyarakatnya. Selain itu, Indonesia menghadapi banyak masalah sosioekonomi, termasuk populasi yang berlebihan, kemiskinan, tingkat pengangguran yang tinggi, dan sistem pendidikan yang buruk. Masalah-masalah ini dapat memengaruhi masyarakat, termasuk meningkatkan kejahatan. Banyak indikator yang umum digunakan dalam bidang statistik kriminal untuk mengukur kejahatan dari perspektif yang luas dan untuk menilai tingkat keparahannya. Tujuan dari penelitian ini adalah untuk menggambarkan distribusi tingkat kejahatan secara keseluruhan di antara provinsi-provinsi Indonesia, dengan penekanan khusus pada Sumatra dan Jawa. Studi ini menggunakan data dari Badan Pusat Statistik dari tahun 2010 hingga 2020 tentang jumlah kejahatan yang dilaporkan oleh petugas polisi regional. Objek yang diamati di masing-masing provinsi dikelompokkan ke dalam kelompok yang saling terkait menggunakan teknik pembelajaran tidak terbimbing dengan algoritma klasifikasi K-Means. Hasil menunjukkan bahwa antara tahun 2010 dan 2020, provinsi Bengkulu, Kepulauan Bangka Belitung, dan Banten memiliki tingkat kejahatan terendah dibandingkan provinsi lain di Sumatra dan Jawa. Hasil ini menunjukkan bahwa ketiga provinsi ini mungkin memiliki kemampuan yang lebih baik untuk mengatasi masalah sosioekonomi yang diketahui berkontribusi pada kejahatan.

Kata Kunci: Pemetaan; Spasial; K-Means; Kriminalitas; Indonesia

Abstract Indonesia's Human Development Index (HDI) underscores the imperative of sustained efforts to improve the quality of life and healthcare services of the nation. The HDI's modest ranking accentuates the substantial developmental strides required to address prevailing disparities. In addition, Indonesia is grappling with a complex matrix of socioeconomic challenges, including issues such as overpopulation, deep poverty, rising unemployment, and systemic educational deficiencies. The amalgamation of these intricate challenges engenders the potential for adverse sociocultural outcomes, with an increased vulnerability to escalating crime rates being a major concern. Within the realm of criminological discourse, a diverse spectrum of metrics is used to comprehensively assess the spectrum and gravity of criminal propensities. The present study aims to illuminate the multifaceted landscape of aggregate criminal prevalence in Indonesia's diverse provinces, with a specific lens on the geographies of Sumatra and Java. Drawing on official data archived by the Central Bureau of Statistics during the period 2010 to 2020, the research scrutinizes reported instances of criminal conduct as documented by regional law enforcement entities. The analytical pursuit adopts an unsupervised learning framework, with the K-means classification algorithm as the methodological scaffold. This approach systematically clusters observed entities within each province, unified by shared parameters. Evidently, the results of this analytical construct delineate the provinces of Bengkulu, the Bangka Belitung Islands and Banten as enclaves exhibiting notably reduced frequencies of criminal occurrences. These findings corroborate the contention that these provinces showed relatively subdued levels of criminal involvement in relation to their counterparts in the domains of Sumatra and Java during the stipulated epoch from 2010 to 2020.

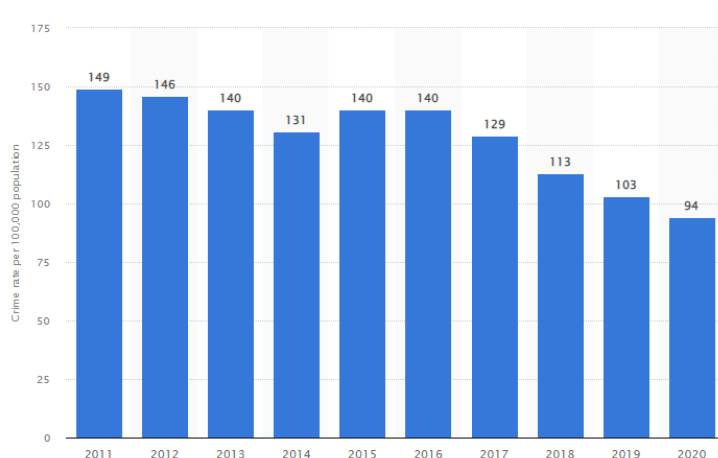
Keywords: Mapping; Spatial; K-Means; Crime; Indonesia

1. PENDAHULUAN

Indonesia memiliki Indeks Pembangunan Manusia yang rendah, yang menunjukkan bahwa masih ada pekerjaan yang harus dilakukan untuk meningkatkan kualitas hidup dan kesehatan masyarakatnya [1]. Selain itu, Indonesia menghadapi banyak masalah sosioekonomi, termasuk populasi yang berlebihan [2], kemiskinan [3], tingkat pengangguran yang tinggi [4], dan sistem pendidikan yang buruk. Masalah-masalah ini dapat memengaruhi masyarakat, termasuk meningkatkan kejahatan [5]. Pemerintah Indonesia hanya mengandalkan data deskriptif untuk proses identifikasi masalah sosial dan kejahatan, tetapi tidak memasukkan masalah sosial yang telah diakui dapat mempengaruhi kejahatan. Faktor-faktor seperti pendidikan, populasi, dan ekonomi dapat dikaitkan dengan frekuensi kegiatan kriminal di suatu tempat. Setiap daerah di Indonesia memiliki karakteristik unik, sehingga kecenderungan kejahatan tidak sama di setiap daerah. Pelanggaran kriminal terjadi di Indonesia setiap 1 menit dan 32 detik, menurut perhitungan yang dilakukan oleh Badan Pusat Statistik. Sehingga dengan 100.000 penduduk yang tinggal di Indonesia, 140 rentan menjadi korban pelanggaran hukum [6].

Bidang statistik kriminal sering menggunakan berbagai indikator untuk mengukur kejahatan dari sudut pandang yang lebih luas dan untuk mengevaluasi tingkat keparahannya [7]. Dalam konteks yang lebih luas, indikator tersebut mencakup jumlah total kejahatan yang dilakukan (jumlah kejahatan total), jumlah kejahatan yang dilakukan per 100.000 orang dalam populasi (tingkat kejahatan), dan jumlah jam yang dibutuhkan sebelum kejahatan terjadi. Informasi terbaru menunjukkan tren yang menurun dalam proporsi populasi yang menjadi korban kejahatan pada tahun 2019 dan 2020. Korban kejahatan berjumlah 0,78% dari populasi pada tahun 2020, turun dari 1,01% pada tahun 2019. Sementara itu, jumlah laporan tahunan ke polisi masih cukup rendah. Persentase penduduk Indonesia yang mengalami kejadian kriminal dan kemudian melaporkannya kepada polisi tidak lebih dari 25% pada tahun 2019 dan 2020. Ini naik sedikit dari 22,19% pada tahun 2019. Gambar 1 menunjukkan tingkat kejahatan di Indonesia tercatat sebanyak 94 insiden per seratus ribu penduduk pada tahun 2020. Ini turun dari puncak 149 insiden per seratus ribu penduduk pada tahun 2011, yang merupakan titik tertinggi

dalam tujuh tahun. Sejak itu, tingkat kejahatan telah menurun dengan konsisten, menempatkan Indonesia dalam kategori keamanan yang sedang [8].



Gambar 1. Tingkat kejahatan di Indonesia dari tahun 2011 hingga 2020 (per 100.000 penduduk) [9]

Tujuan dari penelitian ini adalah untuk mengategorikan distribusi kejadian kriminal secara sistematis di berbagai provinsi Indonesia, dengan penekanan khusus pada wilayah Sumatra dan Jawa. Data yang digunakan mencakup informasi statistik dari Badan Pusat Statistik dari 2010 hingga 2020, yang mencakup jumlah insiden kejahatan yang dilaporkan oleh otoritas penegak hukum regional masing-masing. Dengan menggunakan algoritma K-Means sebagai alat utama untuk klasifikasi dalam pendekatan pembelajaran tanpa pengawasan, penelitian ini memfasilitasi pengelompokan entitas yang diamati di setiap provinsi ke dalam kluster-kluster saling terhubung yang memiliki hubungan satu sama lain.

Prinsip-prinsip teknik pembelajaran tanpa pengawasan dan algoritma kluster K-Means berdiri di dasar metodologi ini. Karena kemampuannya untuk mengekstrak pengetahuan dan menemukan pola, metode ini telah mendapatkan perhatian akademis yang signifikan saat ini. Algoritma K-Means adalah contoh teknik klusterisasi partisi; ini adalah paradigma klusterisasi yang menggabungkan data menjadi komponen konstituen yang berbeda [10], [11]. Metode K-Means telah muncul sebagai pilihan utama untuk mengklusterisasi kumpulan data besar, menggabungkan banyak data, dan menemukan pencilaan. Ini dikenal karena kemudahan operasionalnya dan kemampuan untuk memproses data dengan cepat [12]–[14].

Para ilmuwan di Bangladesh telah menyelidiki penyebaran kegiatan kriminal secara spasial meliputi representasi kartografi insiden kejahatan dan analisis menyeluruh dari sebaran spasialnya. Pemetaan kejadian kriminal di Bangladesh saat ini masih sangat baru meskipun menggunakan alat kuat seperti Sistem Informasi Geografis (GIS). Studi ini juga melibatkan analisis menyeluruh terhadap tren kejahatan selama tiga tahun, terutama dari tahun 2016 hingga 2018. Tujuan utama dari studi ini adalah untuk mengurai konfigurasi-nuansa yang menggambarkan distribusi insiden kejahatan di seluruh batasan administratif yang ada. Selain itu, analisis proporsional perilaku kriminal di berbagai wilayah Bangladesh adalah bagian penting dari penelitian ini [15].

Teknik penambahan data yang digunakan secara strategis memfasilitasi proses kartografi yang kompleks, yang mendukung upaya penelitian ini. Untuk ketatnya analisis ini, set data yang digunakan terdiri dari 34 catatan individual. Setiap catatan menggambarkan informasi diskrit tentang aktivitas kriminal kuantitatif yang dilaporkan kepada otoritas penegak hukum regional dari tahun 2017 hingga 2019. Studi ini menunjukkan bahwa enam provinsi—Sumatera Utara, Sumatera Selatan, Metro Jaya, Jawa Barat, Jawa Timur, dan Sulawesi Selatan—menunjukkan lokasi geografis dengan tingkat insiden kejahatan yang paling mencolok [16].

RapidMiner digunakan untuk menerapkan proses pengelompokan. Selanjutnya, untuk mendapatkan representasi spasial yang lebih baik, sebuah gambar kartografi menyeluruh dari semua provinsi di Indonesia dibuat dengan bantuan Quantum GIS. Data yang diperoleh selama proses pengelompokan sebelumnya digunakan untuk membuat visualisasi ini [17]. Tujuan utama dari penelitian ini adalah untuk meningkatkan pemahaman visual tentang frekuensi kejadian kriminal di seluruh Indonesia, dengan penekanan khusus pada kejadian di provinsi di Pulau Jawa dan Sumatera.

2. METODOLOGI PENELITIAN

Metode K-Means digunakan oleh banyak peneliti dan praktisi di berbagai domain. Dalam bidang akademik, berbagai pendekatan telah digunakan untuk meningkatkan kerangka dasar K-Means. Algoritma K-Means dan varian-varianannya, yang berasal dari paradigma pembelajaran tanpa pengawasan, menjadi alat penting untuk melaksanakan upaya pengelompokan data. Ini juga berfungsi dengan baik dalam bidang pengenalan pola dan pembelajaran mesin. [14], [18]. Meskipun demikian, satu hal penting yang perlu diperhatikan adalah bahwa jumlah kluster yang digunakan untuk inisialisasi iterasi-algoritma tetap konstan, tidak peduli apakah proses pelatihan dilakukan dalam pengaturan yang terpantau atau tanpa pengawasan. Dengan demikian, jelas bahwa algoritma K-Means pada dasarnya tidak berfungsi

sebagai solusi otomatis untuk klasifikasi data tanpa intervensi manusia; ini menunjukkan betapa pentingnya intervensi manusia dalam proses klasifikasi [19].

2.1 Algoritma K-Means

Eksposisi tentang algoritma K-Means memberikan gambaran tentang teknik pengelompokan iteratif dasar yang ditandai dengan pendekatan yang dapat diakses. Perhitungan jarak rata-rata adalah langkah penting yang memudahkan derivasi sentroid awal dalam urutan operasional algoritma ini. Perhitungan ini menggunakan metrik jarak intrinsik, dan kategori K dari dataset dianggap sebagai parameter input. Hasilnya adalah formulasi sentroid primer, yang memberikan dasar untuk tahap berikutnya. Pembatasan setiap kategori yang muncul selanjutnya secara khusus didorong oleh sentroid dasar ini [20]–[22].

Untuk mengukur indeks kesamaan di antara set data X yang diberkati dengan n titik data multidimensi dan kategorisasi target yang disebut K , jarak geometris digunakan. Keputusan metodologis ini dibuat setelah berpikir tentang banyak pilihan. Modulasi yang dicurahkan dalam tujuan pengelompokan, yang secara sistematis disesuaikan untuk membatasi kumpulan kombinasi yang mungkin dalam tingkat yang praktis dapat dicapai, menyatu dalam dinamika algoritma. Dalam waktu bersamaan, upaya ini berusaha untuk mengurangi penjumlahan kuadrat untuk masing-masing kategori diskrit. Tujuannya adalah untuk mengurangi jumlah penjumlahan ini secara absolut [23]–[25].

$$d = \sum_{k=1}^k \sum_{i=1}^n \| (x_i - u_k) \|^2 \quad (1)$$

Dalam rumus yang diberikan, variabel " k " mencakup gagasan tentang jumlah total pusat kluster, " uk " menunjukkan indeks yang diberikan kepada pusat kluster k , dan " xi " menunjukkan indeks yang diberikan kepada titik data i th dalam kumpulan data. Selain itu, sesuai dengan penunjukan yang ditetapkan, " k " dapat memperoleh nilai bilangan bulat positif. Dalam situasi ini, ada metode yang dapat digunakan untuk menyelesaikan masalah khusus mengenai bagaimana " uk " diidentifikasi sebagai pusat.

Persamaan ini dapat direpresentasikan sebagai berikut:

$$\begin{aligned} \frac{\partial}{\partial u_k} &= \frac{\partial}{\partial u_k} \sum_{k=1}^k \sum_{i=1}^n (x_i - u_k)^2 \\ &= \sum_{k=1}^k \sum_{i=1}^n \frac{\partial}{\partial u_k} (x_i - u_k)^2 \\ &= \sum_{i=1}^n 2(x_i - u_k) \end{aligned} \quad (2)$$

Dengan menetapkan titik " k " sebagai pusat fokus dari kelompok awal, gagasan di balik penggunaan algoritma bergantung pada pemilihan acak titik dari kumpulan titik sampel. Dimulai dengan kategorisasi awal setiap titik sampel ke dalam kelompok yang pusatnya paling dekat dengan titik tengah yang dianalisis dalam prosedur ini. Selanjutnya, diminta untuk menghitung koordinat spasial yang menunjukkan inti kelompok tertentu dengan menggunakan rata-rata aritmetika dari seluruh kohor titik sampel. Proses ini diulangi, melacak akarnya kembali ke segmen sebelumnya, sampai titik pusat kelompok menstabilkan, selaras dengan kehabisan jumlah iterasi maksimal yang telah ditetapkan, mana pun yang terjadi lebih awal.

Penting untuk diingat bahwa kesimpulan yang dihasilkan oleh algoritma ini tetap rentan terhadap variabilitas yang bergantung pada seleksi pusat yang tepat; faktor ini secara intrinsik merangkul hasil-hasil algoritma ini dengan stokastisitas yang ada saat proses pengelompokan dimulai. Akibatnya, kepercayaan pada hasil temuan menurun. Dalam algoritma ini, nilai " k " menunjukkan penentu sentral, yang bertanggung jawab untuk menentukan koordinat yang tepat dari pusat nexus. Baik keberadaannya yang mendesak maupun efek yang dihasilkannya sangat memengaruhi hasil pengelompokan. Hal ini berdampak pada elemen penting seperti menggambarkan kluster sebagai entitas yang terlokalisasi atau termasuk dalam konfigurasi global yang lebih luas [26].

2.2 RapidMiner

Sebagai salah satu platform yang paling banyak digunakan dan digunakan di seluruh dunia, RapidMiner adalah kerangka kerja penambangan data open-source yang luar biasa. Proyek ini mulai direncanakan di Universitas Dortmund pada tahun 2001, tetapi sejak tahun 2007 telah dipercepat oleh Rapid-I GmbH [17]. Pengaruh RapidMiner melampaui dunia komersial dan bahkan masuk ke dunia akademik, mendapat dukungan dari berbagai lembaga pendidikan dan sarjana. Posisinya diperkuat oleh garis keturunan yang mendalam dari pendidikannya, yang menjadikannya alat yang tidak dapat diabaikan bagi para sarjana dari berbagai bidang penelitian akademis. RapidMiner dianggap memenuhi kebutuhan ilmuwan komputer, ahli statistik, dan matematikawan. Alat yang komprehensif dan serbaguna ini cocok untuk mengeksplorasi kompleksitas pembelajaran mesin, metodologi statistik, dan penambangan data. Secara bersamaan,

utilitasnya menyebar ke wilayah lain. RapidMiner tampak menjadi pengfasilitasi yang tangguh, mempermudah penerapan jalur penelitian baru dan analisis komparatif dengan menyediakan antarmuka yang mudah dipahami untuk menciptakan dan menerapkan paradigma dan metodologi penelitian baru [27].

Sebaliknya, RapidMiner banyak digunakan dalam berbagai bidang, seperti fisika, teknik mesin, kedokteran, kimia, linguistik, dan ilmu sosial. Banyak sub domain penelitian saat ini didorong oleh banyaknya data, yang menekankan pentingnya alat analitis yang serbaguna. Dengan menyediakan kumpulan metodologi yang beragam dan berguna, RapidMiner muncul sebagai instrumen penting dalam konteks ini. Metodologi ini mencakup penilaian statistik dasar seperti analisis korelasi hingga proses yang lebih kompleks seperti analisis regresi, klasifikasi, pengelompokan, dan bidang optimisasi parameter dan reduksi dimensi. Kemampuan RapidMiner melampaui batas prosedur ini. Aplikasi teknologi ini dapat bermanfaat dalam banyak bidang, termasuk analisis teks, gambar, audio, dan deret waktu. Terutama, seluruh cakupan upaya analitis ini dapat diotomatisasi secara menyeluruh, yang akan menghasilkan efisiensi dan akurasi. Ada banyak pilihan presentasi yang dapat digunakan untuk menyampaikan kesimpulan analisis ini [28].

2.3 Quantum GIS

Sistem Informasi Geografis (GIS) adalah rangka kerja lengkap yang memungkinkan penyimpanan, pengambilan, pengolahan, dan visualisasi data yang terkait dengan lokasi geografis tertentu. Istilah "geografis" secara inheren terkait dengan sifat data, bukan hanya menggambarkan lokasi geografis, artinya data diberi georeferensi yang tepat. Berbagai operasi yang kompleks, seperti administrasi, pemrosesan, dan pemanfaatan data digital dengan georeferensi yang tepat, termasuk dalam kategori ini [29], [30]. Bidang Sistem Informasi Geografis (GIS) telah berkembang pesat dalam sepuluh tahun terakhir. Disebabkan oleh perkembangan ini, ada variasi yang lebih besar dan berbagai aplikasi di berbagai bidang, seperti bidang bisnis dan pemerintahan, adopsi yang merata dari konsumen, dan integrasi yang semakin meningkat dalam sistem informasi, serta institusi akademik yang berfokus pada pendidikan bisnis. Konseptualisasi dan arsitektur Sistem Informasi Geografis (GIS) menggabungkan elemen penelitian kualitatif yang dipertahankan melalui metodologi penelitian deskriptif [31], [32]. Dalam konteks penelitian ini, Quantum GIS, atau QGIS, muncul sebagai alat bantu yang penting untuk pemrosesan dan pengelompokan data terkait dengan jumlah pelanggaran hukum yang terjadi di provinsi Sumatra dan Jawa [33].

2.4 Dataset

Tugas pengumpulan dan analisis data terletak di bawah yurisdiksi Badan Pusat Statistik Republik Indonesia, yang sering disingkat sebagai BPS. Dalam penelitian ini, telaah yang cermat terhadap data yang berkaitan dengan kejahatan dalam rentang waktu dari tahun 2010 hingga 2020 di dalam provinsi Sumatra dan Jawa telah dijalankan [34]. Sebagai aspek integral dari protokol visualisasi datanya, Rapid Miner secara strategis memanfaatkan metodologi pengelompokan. Serangkaian temuan yang komprehensif ini terangkum dalam Tabel 1, yang meliputi baik data mentah yang belum diolah maupun hasil-hasil setelah melalui prosedur pengolahan yang ditetapkan:

Tabel 1. Jumlah kejahatan menurut kepolisian daerah di setiap provinsi pada tahun 2010-2020

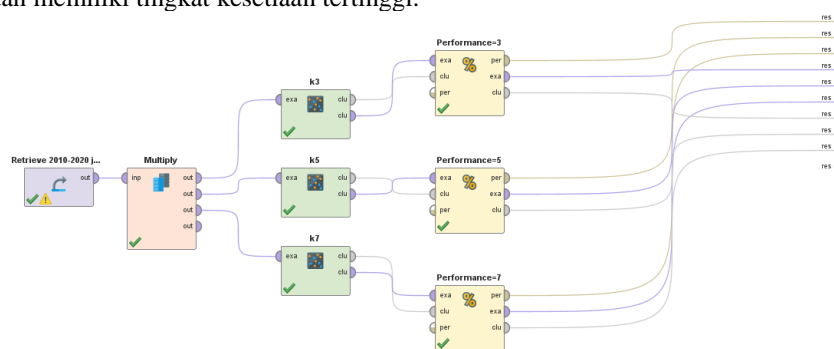
Provinsi	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Aceh	9244	9114	9200	9150	7569	8048	9646	8885	8758	7483	7745
Sumatera Utara	33227	37610	33250	40709	35728	35248	37102	39867	32922	30831	32990
Sumatera Barat	10819	11695	13468	14324	14955	16277	14921	13205	12953	11064	7992
Riau	10129	8323	12533	9399	9644	9595	8520	6869	7246	6570	8194
Jambi	3586	4450	6099	6510	7643	10564	9424	9531	6313	6848	4709
Sumatera Selatan	18288	19353	21498	22882	22708	20575	20368	15728	13558	12861	12189
Bengkulu	2717	3498	3943	4550	3847	4463	5904	4867	3389	3453	3333
Lampung	4813	6052	4383	4812	7755	9218	10485	11089	8963	8534	7594
Bangka Belitung	2642	2732	5197	2515	1796	1875	2094	1931	2048	1953	1931
Kep. Riau	4141	3643	3626	4278	4633	4892	4885	3673	3409	3159	2843
Jakarta	60989	53324	52642	49498	44298	44461	43842	34767	34655	31934	26585
Jawa Barat	16869	29296	27247	24843	27058	27805	29351	25183	16209	13145	11256
Jawa Tengah	15479	15205	11079	14859	15993	15958	14353	12033	9127	10317	10712
Yogyakarta	17622	6326	8987	6727	7135	9692	8348	7251	6731	6650	7721
Jawa Timur	16948	28392	22774	16913	14102	35437	28902	34598	26295	26985	17642
Banten	3832	3205	3804	4259	5741	5002	4570	3692	3623	3287	4250

3. HASIL DAN PEMBAHASAN

Dalam penelitian ini, metodologi pengelompokan K-Means digunakan untuk memeriksa data yang tercantum dalam Tabel 1 secara menyeluruh. Dalam analisis ini, kelompok ketiga, kelima, dan ketujuh diberikan latihan pengelompokan secara menyeluruh, menggunakan label pemetaan secara strategis. Gambar 2 menunjukkan representasi visual dari proses pemetaan yang tepat yang dilakukan oleh Rapid Miner. Secara khusus, Gambar 2 menunjukkan bagaimana metode pembacaan Excel dapat digunakan secara optimal untuk memasukkan data sesuai dengan informasi penting yang ditemukan dalam Tabel 1.

Salah satu tugas utama yang diemban oleh model K-means adalah pemetaan spasial terhadap insiden aktivitas kriminal di wilayah Sumatra dan Jawa. Selain itu, pemeriksaan yang lebih mendalam terhadap data yang disertakan dalam Tabel 2 menunjukkan proposal untuk tiga kelompok yang berbeda, yang direncanakan untuk meningkatkan efektivitas metrik

performa yang didasarkan pada Davies Bouldin Index (DBI). Tujuan akhir dari pelatihan ini adalah untuk memberikan gambaran yang akurat tentang kompetensi teknologi informasi di masing-masing provinsi dengan memilih kelompok yang paling cocok dan memiliki tingkat kesetiaan tertinggi.



Gambar 2. Tiga klaster yang berbeda dipetakan dengan pemodelan RapidMiner

Tabel 2 menampilkan hasil evaluasi Indeks Davies-Bouldin untuk masing-masing tabel. Indeks Davies-Bouldin (DBI) adalah alat yang digunakan untuk mengevaluasi efisiensi pendekatan pengelompokan. Angka DBI harus positif dan tidak negatif [35], [36]. Kelompok (k) yang ditemukan oleh teknik yang digunakan lebih kuat jika nilai ini lebih kecil. Salah satu metode pengorganisasian data yang paling efektif adalah membagi 16 provinsi di Sumatra dan Jawa menjadi tujuh kelompok, dengan skor DBI sebesar -0,432. Quantum GIS menciptakan tujuh kelompok data yang jelas dan berbeda dengan memeriksa dengan cermat korpus data yang berasal dari Open Map yang mencakup provinsi-provinsi Indonesia. Upaya yang cermat ini menghasilkan konstitusi dataset yang lengkap serta dasar untuk analisis lanjutan. Gambar 3 menunjukkan hasil ekspedisi prosedural ini dengan memungkinkan visualisasi QGIS untuk setiap kelompok.

Provinsi Bengkulu, Kepulauan Bangka Belitung, dan Banten muncul sebagai kelompok dengan skor paling rendah di antara tujuh kelompok, menurut data yang diuraikan dalam Tabel 2. Provinsi-provinsi ini menunjukkan tingkat aktivitas kriminal yang relatif rendah selama tahun 2010–2020, terutama dibandingkan dengan provinsi pesaing mereka. Namun, perbedaan yang signifikan terlihat ketika fokus beralih ke provinsi seperti Aceh, Riau, Jambi, Lampung, dan Yogyakarta. Provinsi-provinsi ini dianggap sebagai pusat pelanggaran kekerasan karena memiliki tingkat kejahatan tertinggi pada waktu yang sama.

Tabel 2. Kinerja ketiga menurut Indeks Davies-Bouldin

Klaster	DBI
Kinerja K=3	-0.584
Kinerja K=5	-0.554
Kinerja K=7	-0.432



Gambar 3. Distribusi kejahatan berdasarkan klaster di Jawa dan Sumatera, 2010-2020

Tabel 2. Distribusi warna di setiap klaster setiap provinsi

Klaster	Provinsi	Warna
Klaster 0	Bengkulu	
	Kep. Bangka Belitung	
	Banten	
Klaster 1	Sumatra Selatan	

	Jawa Barat	
Klaster 2	Jakarta	
Klaster 3	Sumatra Barat	
	Jawa Tengah	
Klaster 4	Jawa Timur	
Klaster 5	Sumatra Utara	
	Aceh	
	Riau	
Klaster 6	Jambi	
	Lampung	
	DI Yogyakarta	

Dengan memasukkan beberapa provinsi, seperti Jakarta, Jawa Timur, dan Sumatera Selatan, ke dalam klaster yang berbeda, terungkap hubungan yang kompleks. Fenomena ini menunjukkan bahwa provinsi-provinsi ini bertanggung jawab atas banyak kejadian kriminal. Selama rentang waktu yang panjang dari tahun 2010 hingga 2020, banyak kegiatan kriminal yang digabungkan ini tampaknya berperan sebagai penggerak utama dalam isolasi mereka..

4. KESIMPULAN

Berdasarkan temuan penelitian ini, metodologi K-means tampaknya dapat digunakan untuk menggambarkan secara kartografis berbagai peristiwa kriminal yang terjadi di Indonesia, dengan penekanan khusus pada wilayah demarkasi Sumatra dan Jawa. Patut ditekankan bahwa, dalam upaya untuk meningkatkan kinerja Indeks Davies Bouldin (DBI), pengklasifikasian yang mencakup tiga klaster yang berbeda telah diperkenalkan. Manuver strategis ini direncanakan dengan tujuan utama untuk memastikan konfigurasi yang terukur dan ideal untuk pemetaan berbasis provinsi. Dalam trichotomy klaster ini, jelas bahwa klaster ketujuh menonjol dengan skor DBI yang lebih rendah sebesar 0,432 poin dibandingkan dengan dua rekannya. Akibatnya, klaster ketujuh muncul sebagai pilihan utama di antara triad tersebut. Jaringan QGIS telah digunakan untuk membentuk kontur ketujuh klaster ini dengan cermat. Ini memungkinkan pemahaman menyeluruh tentang setiap klaster dalam panorama. Metode kreatif ini sepenuhnya sesuai dengan prinsip data peta terbuka yang dapat diakses untuk setiap provinsi di seluruh kain Indonesia. Provinsi Bengkulu, Kepulauan Bangka Belitung, dan Banten beralih menuju peringkat yang lebih rendah dalam hierarki klaster sebagai konsekuensi logis. Ini secara kolektif mendorong informasi bahwa selama periode dari tahun 2010 hingga 2020, ketiga provinsi ini secara umum menunjukkan tingkat kejahatan yang paling rendah dibandingkan dengan rekan provinsi mereka yang melintasi Sumatra dan Jawa.

REFERENCES

- [1] Badan Pusat Statistik, "Indeks Pembangunan Manusia (IPM) Indonesia Tahun 2020," *Bps*, no. 97, 2020.
- [2] Muttaqin, M. Zuhri, and A. Irfan, "Perancangan sistem pemetaan dan pendataan populasi penduduk miskin di Kota Banda Aceh menggunakan aplikasi Quantum GIS," *Informatics and Computer Science*, vol. 5, no. 1, 2019.
- [3] I. Lubis and D. Eva Santi Silalahi, "Faktor-Faktor Yang Mempengaruhi Kemiskinan di Provinsi Sumatera UTara," *JURNAL PENELITIAN PENDIDIKAN SOSIAL HUMANIORA*, vol. 7, no. 1, 2022.
- [4] N. P. Syabrina, Hardiani, and C. Mustika, "Pengaruh Pertumbuhan Ekonomi Dan Tingkat Pengangguran Terhadap Tingkat Kemiskinan Di Provinsi Sumatera Utara," *e-Jurnal Perspektif Ekonomi dan Pembangunan Daerah*, vol. 10, no. 1, 2021.
- [5] R. A. Pratama, T. Shafira, F. Ardiansyah, and R. F. Hakim, "Characteristics and segmentation of social problems with kohonen self-organizing maps," *Bulletin of Social Informatics Theory and Application*, vol. 1, no. 1, 2017, doi: 10.31763/businta.v1i1.19.
- [6] L. Ghiffari, N. Gusriani, and K. Parmikanti, "Pemetaan Jenis Tindak Kriminal di Indonesia Berdasarkan Karakteristik Wilayah Menggunakan Canonical Correspondence Analysis (CCA)," *Jurnal Statistika dan Aplikasinya*, vol. 5, no. 2, 2021, doi: 10.21009/jsa.05202.
- [7] A. Chairani P, I. K. Dwi Astuti, N. Panjaitan, S. R. Y. Sinurat, and R. J. Yuhan, "Analisis Jalur Pada Kejadian Kriminalitas di Indonesia Tahun 2018," *Jurnal Sains Matematika dan Statistika*, vol. 6, no. 2, 2020, doi: 10.24014/jsms.v6i2.10555.
- [8] Margo Yuwono, *STATISTIK KRIMINAL 2021*. Badan Pusat Statistik, 2021.
- [9] Statista Research Department, "Crime rate in Indonesia from 2011 to 2020(per 100,000 population)," *statista.com*, Feb. 2022.

- [10] K. P. Sinaga and M. S. Yang, "Unsupervised K-means clustering algorithm," *IEEE Access*, vol. 8, 2020, doi: 10.1109/ACCESS.2020.2988796.
- [11] S. S. Yu, S. W. Chu, C. M. Wang, Y. K. Chan, and T. C. Chang, "Two improved k-means algorithms," *Applied Soft Computing Journal*, vol. 68, 2018, doi: 10.1016/j.asoc.2017.08.032.
- [12] D. Andra and A. B. baizal, "E-commerce Recommender System Using PCA and K-Means Clustering," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 6, no. 1, pp. 57–63, 2022, doi: 10.29207/resti.v6i1.3782.
- [13] M. Jumarlis and Mirfan, "Detecting Diseases on Clove Leaves Using GLCM and Clustering K-Means," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 6, no. 4, pp. 624–631, 2022, doi: 10.29207/resti.v6i4.4033.
- [14] I. Shaik, S. S. Nittela, T. Hiwarkar, and S. Nalla, "K-means clustering algorithm based on E-commerce big data," *International Journal of Innovative Technology and Exploring Engineering*, vol. 8, no. 11, 2019, doi: 10.35940/ijitee.K2121.0981119.
- [15] Md. B. Khan and Md. I. Talukder, "Spatial Distribution of Crime in Bangladesh: An Analysis," *Journal of Penal Law and Criminology / Ceza Hukuku ve Kriminoloji Dergisi*, vol. 0, no. 0, 2021, doi: 10.26650/jplc2021-931625.
- [16] N. Azis *et al.*, "Mapping study using the unsupervised learning clustering approach," *IOP Conf Ser Mater Sci Eng*, vol. 1088, no. 1, p. 012005, Feb. 2021, doi: 10.1088/1757-899X/1088/1/012005.
- [17] S. Land and S. Fischer, *RapidMiner 5: RapidMiner in academic use*. Rapid-I GmbH, 2012.
- [18] C. Yuan and H. Yang, "Research on K-Value Selection Method of K-Means Clustering Algorithm," *J Multidisciplinary Scientific Journal*, vol. 2, no. 2, pp. 226–235, Jun. 2019, doi: 10.3390/j2020016.
- [19] R. Watrionthos, R. Handayani, A. F. P. Akhir, A. Ambiyar, and U. Verawardina, "Penerapan Algoritma K-Means Pada Pemetaan Kemampuan Penggunaan Teknologi Informasi Remaja dan Dewasadi Indonesia," *Journal of Computer System and Informatics (JoSYC)*, vol. 4, no. 1, Nov. 2022.
- [20] F. P. Dewi, P. S. Aryni, and Y. Umaidah, "Implementasi Algoritma K-Means Clustering Seleksi Siswa Berprestasi Berdasarkan Keaktifan dalam Proses Pembelajaran," *JISKA (Jurnal Informatika Sunan Kalijaga)*, vol. 7, no. 2, pp. 111–121, Jul. 2022, doi: 10.14421/jiska.2022.7.2.111-121.
- [21] F. Nur, M. Zarlis, and B. B. Nasution, "Penerapan Algoritma K-Means Pada Siswa Baru Sekolah Menengah Kejuruan Untuk Clustering Jurusan," *InfoTekJar: Jurnal Nasional ...*, 2017, [Online]. Available: <https://jurnal.uisu.ac.id/index.php/infotekjar/article/view/70>
- [22] I. Kurniawan and A. Susanto, "Implementasi Metode K-Means dan Naïve Bayes Classifier untuk Analisis Sentimen Pemilihan Presiden (Pilpres) 2019," *Eksplora Informatika*, vol. 9, no. 1, pp. 1–10, Sep. 2019, doi: 10.30864/eksplora.v9i1.237.
- [23] Qomariyah and M. U. Siregar, "Comparative Study of K-Means Clustering Algorithm and K-Medoids Clustering in Student Data Clustering," *JISKA (Jurnal Informatika Sunan Kalijaga)*, vol. 7, no. 2, pp. 91–99, Jul. 2022, doi: 10.14421/jiska.2022.7.2.91-99.
- [24] S. R. Munthe, S. Suryadi, and ..., "Algoritma K-Means Untuk Analisis Presepsi Masyarakat Labuhanbatu Dalam Promosi Produk Berbasis Digital Pasca COVID-19," *U-NET Jurnal ...*, 2021, [Online]. Available: <https://ejurnal.univalabuhanbatu.ac.id/index.php/u-net/article/view/346>
- [25] E. Kartikadarma, S. Jumini, N. B. H. Ismail, and ..., "Educational Data Mining to Improve Decision Support on the Ratio of students and Study Groups in Elementary Schools in Indonesia using K-Means Method.," *Ilkogretim ...*, 2021, [Online]. Available: <http://search.ebscohost.com/login.aspx?direct=true&profile=ehost&scope=site&authtype=crawler&jrnl=13053515&AN=150153727&h=KZBWkxTSzOwdMGGIH4O%2FJVnCGAi5ddgp5e7DFjomzmWCXdSVS4BVgJKzdCS0donDlo7%2BMKK12EevHXK9XuTg%3D%3D&crl=c>
- [26] F. Rahman, I. I. Ridho, M. Muflih, S. Pratama, M. R. Raharjo, and A. P. Windarto, "Application of Data Mining Technique using K-Medoids in the case of Export of Crude Petroleum Materials to the Destination Country," *IOP Conf Ser Mater Sci Eng*, vol. 835, no. 1, p. 012058, Apr. 2020, doi: 10.1088/1757-899X/835/1/012058.
- [27] Ainurrohman, "Akurasi Algoritma Klasifikasi pada Software Rapidminer dan Weka," *Prosiding Seminar Nasional Matematika*, vol. 4, 2021.
- [28] S. Kurniawan, W. Gata, D. A. Puspitawati, I. K. S. Parthama, H. Setiawan, and S. Hartini, "Text Mining Pre-Processing Using Gata Framework and RapidMiner for Indonesian Sentiment Analysis," *IOP Conf Ser Mater Sci Eng*, vol. 835, p. 012057, May 2020, doi: 10.1088/1757-899X/835/1/012057.
- [29] J. Karman, H. Mulyono, and A. Taqwa Martadinata, "Sistem Informasi Geografis Berbasis Android Studi Kasus Aplikasi SIG Pariwisata," *Deepublish*. 2019.
- [30] J. Karman and H. Mulyono, "Perancangan Sistem Informasi Geografis Lokasi Objek Wisata Di Kota Lubuklinggau Berbasis Android," *J-SAKTI (Jurnal Sains Komputer dan Informatika)*, vol. 3, no. 2, 2019, doi: 10.30645/j-sakti.v3i2.160.

- [31] A. Riansyah, D. Kurniadi, I. Widiastuti, and ..., “Sistem Informasi Geografis Untuk Pemetaan Lokasi Dan Bidang Kompetensi Sekolah Menengah Kejuruan (SMK) Di Kota Tegal,” *TRANSISTOR Elektro* ..., 2018, [Online]. Available: <http://jurnal.unissula.ac.id/index.php/EI/article/view/3025>
- [32] K. Hamdi, “Analisis Data Sistem Informasi Geografis Rumah Tidak Layak Huni (RTLH) Menggunakan Metode Fuzzy Logic,” *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 4, no. 6, pp. 1179 – 1189–1179 – 1189, 2020, doi: 10.29207/resti.v4i6.2658.
- [33] R. E. P. Putera, F. Ramdani, and ..., “Evaluasi Tampilan Antarmuka QGIS dan MapWindow dengan Menggunakan Pendekatan Heuristic Evaluation (Studi Kasus: Fungsi Geoprocessing Tools),” ... *Teknologi Informasi dan* ..., vol. 3, no. 11, 2020, [Online]. Available: <http://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/6681>
- [34] Badan Pusat Statistik, “Statistik kriminal 2022,” 2022.
- [35] A. Idrus, N. Tarihoran, U. Supriatna, A. Tohir, S. Suwarni, and R. Rahim, “Distance Analysis Measuring for Clustering using K-Means and Davies Bouldin Index Algorithm,” *TEM Journal*, vol. 11, no. 4, 2022, doi: 10.18421/TEM114-55.
- [36] Y. A. Wijaya, D. A. Kurniady, E. Setyanto, W. S. Tarihoran, and ..., *Davies Bouldin Index Algorithm for Optimizing Clustering Case Studies Mapping School Facilities*. ceeol.com, 2021. [Online]. Available: <https://www.ceeol.com/search/article-detail?id=977772>