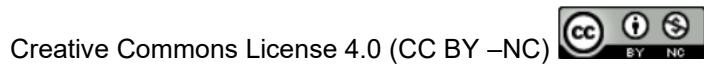


Density-Based Spatial Clustering for Assessing Public Service Accessibility in Jombang Regency

DOI: <http://dx.doi.org/10.35889/progresif.v22i1.3248>



Indiana Lazulfa^{1*}, Reza Augusta Jannatul Firdaus², Anita Andriani³, Muhammad Husein Hanafiyah⁴, Lilis Karimatun Nisa⁵

Teknologi Informasi, Universitas Hasyim Asy'ari Tebuireng, Jombang, Indonesia

*e-mail Corresponding Author: indanalazulfa@unhasy.ac.id

Abstract

The problem of unequal accessibility to public services between subdistricts in Jombang can hinder medium, long-term equitable development. Good accessibility indicates justice and improvements in quality of life in the area. In some subdistricts, school and health facilities are easy to reach, whereas in other areas, mountainous regions or border areas, accessibility is very low. This study aims to map accessibility clusters, analyze the data, and provide recommendations for priority intervention area using DBSCAN. The variables include population density, ratios of health facility, education and administrative facilities per area, geospatial public datasets such as road network density, elevation/slope and rivers. The results show three accessibility clusters (high, medium, and low) with variations in geographical constraints such as hills and major rivers that affect the distribution of services. The visualization shows that low-accessibility areas are generally located in peripheral and mountainous regions, whereas high accessibility is concentrated in the regency center.

Keywords: DBSCAN; Spatial clustering; Public service accessibility; Geospatial

Abstrak

Permasalahan ketimpangan aksesibilitas layanan publik antar kecamatan di Jombang dapat menghambat pemerataan pembangunan jangka menengah dan jangka panjang. Aksesibilitas yang baik mengindikasikan keadilan dan peningkatan kualitas hidup di wilayah tersebut. Di beberapa kecamatan, fasilitas sekolah dan kesehatan mudah dijangkau, sedangkan di daerah lain, pegunungan atau perbatasan, aksesibilitas sangat rendah. Kondisi ini menunjukkan adanya ketimpangan spasial yang perlu dianalisis lebih lanjut. Penelitian ini bertujuan untuk memetakan cluster aksesibilitas, menganalisis data hasil dan memberi rekomendasi lokasi prioritas intervensi. Clustering ini menggunakan metode DBSCAN. Variabel mencakup kepadatan penduduk, rasio fasilitas kesehatan, pendidikan, administrasi per wilayah, geospasial public dataset berupa kepadatan jaringan jalan, hambatan seperti elevasi/lereng dan sungai besar. Hasilnya terdapat tiga klaster aksesibilitas, yaitu tinggi, menengah, dan rendah, dengan variasi hambatan geografis seperti perbukitan dan sungai besar yang memengaruhi distribusi layanan. Visualisasi memperlihatkan bahwa daerah aksesibilitas rendah umumnya berada di wilayah pinggiran dan pegunungan, sedangkan aksesibilitas tinggi terkonsentrasi di pusat kabupaten.

Kata kunci: DBSCAN; Spasial clustering; Aksesibilitas layanan public; Geospasial

1. Pendahuluan

Aksesibilitas terhadap layanan publik seperti sekolah, fasilitas kesehatan, dan pusat-pusat keagamaan merupakan faktor kunci dalam mewujudkan keadilan sosial sekaligus pembangunan wilayah yang berkelanjutan. Distribusi yang merata atas fasilitas-fasilitas tersebut menjamin bahwa masyarakat, tanpa memandang lokasi geografinya, memiliki kesempatan yang adil untuk mengakses layanan esensial. Aksesibilitas yang baik juga meningkatkan kualitas hidup masyarakat karena memperpendek jarak fisik, menurunkan biaya transportasi, serta

mempercepat proses interaksi antara warga dan institusi pelayanan. Sebaliknya, aksesibilitas yang rendah seringkali mengakibatkan ketidakadilan, memperburuk kesenjangan sosial-ekonomi, dan memperlambat pembangunan wilayah. Aksesibilitas hakikatnya merupakan konsep sentral dalam kajian perencanaan kota dan wilayah. Menurut Geurs dan van Wee (2004), aksesibilitas mencakup empat dimensi yang saling berkaitan, yaitu: (1) dimensi tata guna lahan, yang mencerminkan distribusi spasial aktivitas dan layanan; (2) dimensi transportasi, yang berkaitan dengan kecepatan dan biaya perjalanan; (3) dimensi temporal, yang merujuk pada keterbatasan waktu dan jam operasional layanan; serta (4) dimensi individu, yang memperhatikan kondisi sosial-ekonomi dan kapasitas personal [1]. Kerangka ini menegaskan bahwa aksesibilitas tidak hanya soal ketersediaan fasilitas, tetapi juga keterhubungan antara penduduk, infrastruktur, dan karakteristik ruang. Dalam konteks infrastruktur publik, analisis aksesibilitas sering dipakai untuk menilai pemerataan distribusi layanan sekaligus menjadi dasar bagi kebijakan yang bertujuan mengurangi kesenjangan spasial [2], [3].

Namun demikian, saat ini di banyak wilayah terutama di negara berkembang aksesibilitas rendah menghadapi berbagai tantangan. Ketidakmerataan distribusi penduduk, keterbatasan infrastruktur, serta hambatan geografis seperti sungai besar, perbukitan, dan jalur transportasi utama menjadi faktor penghambat pemerataan layanan publik. Di Indonesia, isu pemerataan layanan publik mendapat perhatian serius dalam berbagai kebijakan nasional, terutama dalam Rencana Pembangunan Jangka Menengah Nasional (RPJMN) dan agenda Tujuan Pembangunan Berkelanjutan (SDGs). Pemerintah menekankan pentingnya pembangunan infrastruktur dasar, seperti pendidikan dan kesehatan, untuk memperkuat ketahanan sosial-ekonomi masyarakat. Akan tetapi, implementasi kebijakan ini masih menghadapi disparitas. Fenomena ini juga dialami oleh sejumlah daerah di Indonesia, termasuk Kabupaten Jombang, Jawa Timur yang memiliki 21 kecamatan, 306 kelurahan dan populasi 1,36 juta jiwa. Nilai Indeks Pembangunan Manusia (IPM) 75,67 pada tahun 2024 dengan kategori tinggi, dan laju pertumbuhan ekonomi 5,15% pada tahun 2024-2025 [2]. Namun permasalahannya persentase penduduk miskin masih berada di level menengah yakni 8,60% [3]. Hal ini menegaskan bahwa pemerataan pembangunan dipengaruhi oleh faktor geografis. Kondisi ini menunjukkan adanya ketimpangan spasial yang perlu dianalisis lebih lanjut [4], [5], [6], [7].

Dalam studi terdahulu, analisis aksesibilitas sering menggunakan metode berbasis jarak Euclidean atau model berbasis jaringan transportasi sederhana menggunakan *k-means*, *k-medoids* dan algoritma djikstra. Pendekatan ini relatif mudah diaplikasikan, namun memiliki keterbatasan dalam menangkap kompleksitas spasial, terutama ketika berhadapan dengan wilayah yang memiliki hambatan geografis nyata. Oleh karena itu, penggunaan metode *Density-Based Spatial Clustering of Applications with Noise* (DBSCAN) menjadi relevan [8], [9], [10]. DBSCAN mampu mengidentifikasi klaster dengan bentuk yang tidak beraturan, sekaligus memisahkan wilayah dengan kepadatan rendah sebagai *noise*. Keunggulan ini membuat DBSCAN lebih sesuai untuk memetakan distribusi layanan publik di daerah dengan karakteristik geografis yang beragam seperti Jombang.

Penelitian ini bertujuan untuk menganalisis pola klaster aksesibilitas layanan publik di Kabupaten Jombang dengan metode spatial clustering yakni DBSCAN, mengidentifikasi faktor-faktor geografis dan demografis yang memengaruhi pembentukan klaster tersebut. Sedangkan manfaat dari penelitian ini antara lain memberikan rekomendasi kebijakan berbasis bukti (*evidence-based policy*) guna meningkatkan pemerataan akses layanan publik, perencanaan berbasis bukti, pengurangan ketimpangan spasial, mitigasi hambatan geografis, serta mempercepat akses warga ke layanan dasar sehingga menurunkan biaya dan waktu tempuh. Hasil penelitian ini diharapkan tidak hanya memperkaya literatur akademik tentang implementasi clustering, tetapi juga menjadi kontribusi praktis bagi pembangunan wilayah yang inklusif dan berkeadilan.

2. Tinjauan Pustaka

Penelitian-penelitian sebelumnya di Indonesia telah menerapkan metode klasterisasi untuk menganalisis aksesibilitas layanan publik, namun sebagian besar berfokus pada pendekatan *k-means* atau *k-medoids* [11], [12]. Meskipun metode tersebut efektif dalam menemukan pusat klaster, keduanya cenderung kurang adaptif terhadap variasi kepadatan spasial [13]. Hal ini menimbulkan *research gap* dalam konteks pemerataan layanan publik, khususnya pada wilayah dengan pola pemukiman dan hambatan geografis yang kompleks. Berarti penelitian tersebut tidak memperhatikan faktor topografi atau *obstacle* real.

Metode klasterisasi telah lama digunakan dalam pemetaan aksesibilitas. Mayoritas studi mengandalkan metode *k-means* karena kesederhanaannya. Misalnya, penelitian Yusoff et al menggunakan *k-means clustering* untuk menganalisis distribusi fasilitas kesehatan di Malaysia [14]. Hasilnya cukup efektif, namun kelemahannya adalah asumsi jumlah klaster yang harus ditentukan sejak awal, serta keterbatasan dalam menangkap variasi kepadatan spasial. Beberapa penelitian lain mencoba mengkombinasikan pendekatan klasterisasi dengan GIS dan Dijkstra [15]. Telah dilakukan pula beberapa Upaya peningkatan algoritma DBSCAN. Misalnya, penelitian oleh Ahmed dan Razak telah menganalisis dan menyajikan berbagai peningkatan signifikan dari algoritma DBSCAN untuk bahan evaluasi yakni parameter epsilon dan minimum point [16]. Penelitian tersebut memperlihatkan bahwa faktor parameter optimal memainkan peran signifikan dalam ketimpangan layanan public. Oleh karena itu, di penelitian kali ini juga akan memasukkan unsur peningkatan algoritma sedemikian sehingga parameter yang diinputkan akan semakin baik.

DBSCAN sebagai metode berbasis kepadatan mulai digunakan dalam berbagai bidang analisis spasial. Ahmed dan Razak telah menganalisis DBSCAN dalam menemukan klaster arbitrer serta mengidentifikasi noise [17]. Chen et al. menggunakan DBSCAN untuk memetakan distribusi fasilitas kesehatan di kawasan perkotaan Tiongkok, dan menemukan hasil yang lebih akurat dibanding metode partisi tradisional [18].

Dalam konteks di Indonesia, studi terkait DBSCAN untuk layanan publik masih sangat terbatas. Sebagian besar penggunaan DBSCAN berfokus pada analisis geospasial umum, seperti deteksi pola kriminalitas atau pemetaan titik bencana. Belum ada penelitian yang secara khusus menggunakan DBSCAN untuk menganalisis ketimpangan aksesibilitas pendidikan dan kesehatan di tingkat kabupaten/kota, dan itu masih jarang ditemukan. Pada tahun 2024, Syifa Miftahurrahmi et al menggunakan metode DBSCAN untuk mengcluster provinsi di Indonesia berdasarkan kasus criminal. Meskipun fokusnya bukan layanan publik secara langsung, clustering kasus kriminalitas antar provinsi bisa menunjukkan kawasan yang membutuhkan perhatian layanan keamanan publik. Tetapi bukan layanan seperti sekolah / rumah sakit / fasilitas publik umum [19].

Berdasarkan celah dan permasalahan terbuka dari penelitian sebelumnya tersebut, terdapat beberapa *state of art* dan kebaruan yang difokuskan yaitu permasalahan aksesibilitas pada Tingkat kabupaten, kecamatan/kelurahan, dengan fokus layanan public esensial (sekolah, kesehatan, dan layanan publik) sebagai *Point of Interest* (POI). Kemudian optimalisasi parameter DBSCAN, epsilon dan MinPts, sesuai dengan karakteristik data spasial lokal untuk menangani variasi kepadatan antar wilayah, karena terkait sensitivitas parameter untuk daerah *density* sangat tinggi. Hal lainnya yaitu penggabungan faktor geografis nyata yakni kondisi topografi, hambatan fisik (sungai, pegunungan/bukit) dan infrastruktur jalan (kualitas, konektivitas), bukan hanya berdasarkan jarak Euclidean.

3. Metodologi

Penelitian ini berfokus pada analisa data desa/kelurahan dan kecamatan di Kabupaten Jombang, Jawa Timur. Kurang lebih terdapat 21 kecamatan dan 306 desa/kelurahan untuk uji *robustness*. Kemudian public dataset geospasial meliputi batas-batas wilayah, desa, kecamatan, sungai, pegunungan, dan *point of interest* (POI).

3.1. Objek dan Sumber Data

Variabel yang dibangun per unit spasial antara lain: demografi dan sosial (kepadatan penduduk, rasio usia sekolah 7-18 tahun terhadap total penduduk, rasio lansia lebih dari 60 tahun terhadap total penduduk), fasilitas layanan publik bidang pendidikan (jumlah/kapasitas sekolah TK, SD, hingga SMA), fasilitas Kesehatan (puskesmas, puskesmas pembantu, klinik, poliklinik, rumah sakit umum, rumah sakit khusus, apotek), fasilitas publik lainnya (kantor pos, pasar, tempat ibadah). Variabel selanjutnya yaitu infrastruktur dan jaringan meliputi kepadatan jaringan jalan (km/km^2), keterhubungan, akses ke ruas primer/sekunder (jarak km ke jalan arteri), peta jalan Jombang dan POI. Variabel terakhir adalah geografi dan hambatan (topografi, kemiringan lereng/elevasi, ketinggian, *obstacle*). Data-data tersebut bersumber dari BPS, Dinas Pendidikan dan Kesehatan, *openstreet map*, dan data geospasial Indonesia [16].

Tabel 1. Data sebagian variabel [2]

Kecamatan-	Luas	persentase thdp	Populasi	Tinggi	Rumah	RS	Polikli	Puskes
	wilayah	kabupaten	(2020)	wilayah	sakit			
	km ²	%	jiwa	m.a.s.l.	bangun	bangun	bangun	bangun
Bandar Kedung Mulyo	32,5	2,8	49959	35	0	0	1	1
Perak	29,05	2,51	55975	36	0	0	3	1
Gudo	34,39	2,97	55997	36	0	0	3	2
Diwek	47,7	4,11	110495	35	2	0	7	2
Ngoro	49,86	4,3	77814	41	0	0	1	2
Mojowarno	78,62	6,78	95660	41	1	0	2	2
Bareng	94,27	8,13	56539	63	0	0	3	1
Wonosalam	121,63	10,49	33520	459	0	0	1	1
Mojoagung	60,18	5,19	81051	62	1	0	5	2
Sumobito	47,64	4,11	86385	38	0	0	3	2
Jogoroto	28,28	2,44	70760	43	1	0	1	2
Peterongan	29,47	2,54	67822	42	0	0	1	2
Jombang	36,40	3,14	143803	41	6	1	12	4
Megaluh	28,41	2,45	39725	38	0	0	1	1
Tembelang	32,94	2,84	54891	41	1	0	2	2
Kesamben	51,72	4,46	66861	25	0	0	4	2
Kudu	77,75	6,71	31513	39	0	0	1	1
Ngusikan	34,98	3,02	22793	42	0	0	1	1
Ploso	25,96	2,24	41777	38	1	0	3	1
Kabuh	97,35	8,4	41311	66	0	0	4	1
Plandaan	120,4	10,38	37383	58	0	0	0	1

3.2. Pre-processing

Pada tahap ini, data yang telah terkumpul melalui proses integrasi dibersihkan dari inkonsistensi, duplikasi, maupun nilai hilang (*missing values*). Selain itu, dilakukan normalisasi agar variabel-variabel memiliki skala yang sebanding dan tidak menimbulkan bias dalam proses pengelompokan. Data spasial juga ditransformasikan ke dalam sistem koordinat yang sesuai, sehingga siap untuk diolah lebih lanjut dalam tahap clustering. Pada data numerik dilakukan normalisasi z-score agar skala sebanding.

3.3. Skema Klasterisasi dengan DBSCAN

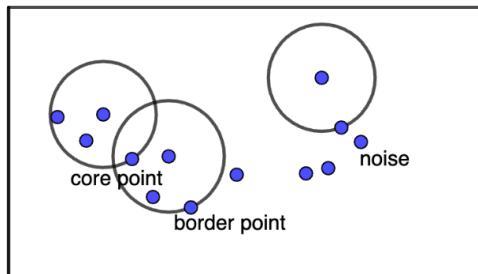
Penelitian ini mengimplementasikan metode DBSCAN untuk melakukan *clustering* dengan bantuan software Tableau, Python dan ArcGIS. Algoritma DBSCAN (*Density-based Spatial Clustering of Application with Noise*) pertama kali diperkenalkan oleh Martin Ester, Hans-Peter Kriegel, Jörg Sander, dan Xiaowei Xu pada tahun 1996. Inti dari metode ini adalah mendefinisikan suatu klaster sebagai sekumpulan titik dalam ruang data yang memiliki kepadatan tinggi, sedangkan wilayah dengan kepadatan rendah dianggap sebagai *noise* atau titik yang tidak termasuk dalam klaster manapun. Dalam DBSCAN, terdapat dua parameter penting yang menentukan proses pembentukan klaster: $\text{epsilon } \varepsilon$ sebagai radius maksimum di sekitar sebuah titik untuk mencari tetangga, dan minPts sebagai jumlah minimum titik yang harus berada dalam radius ε agar titik tersebut dianggap sebagai core point. Penentuan parameter ini merupakan hal yang terus berkembang dari masa ke masa. Nilai minPts yang akan digunakan Adalah dua kali jumlah variabel penelitian. Menurut Ester, tahapan pengelompokan dalam metode DBSCAN sebagai berikut [17], [19]:

- 1) Memilih sebuah titik acak dari dataset. Misalkan objek terpilih, x_{jl} .
- 2) Menghitung jumlah tetangga dalam radius epsilon. Tahap ini bertujuan untuk menentukan objek yang *density-reachable* terhadap objek x_{jl} menggunakan rumus jarak apapun, misalnya *Euclidean distance*

$$d(j, i) = \sqrt{\sum_{l=1}^p (x_{jl} - x_{il})^2} \quad (1)$$

x_{jl} objek ke- j pada variabel ke- l , x_{il} objek ke- i pada variabel ke- l , p jumlah variabel, $d(j, i)$ jarak antara objek j dan objek i .

- 3) Kluster kemudian diperluas dengan menambahkan semua titik yang dapat dijangkau langsung dari *core point* tersebut. *Core point* adalah titik yang memiliki jumlah tetangga lebih dari minPts dalam jarak ε . Titik ini menjadi pusat pembentukan klaster. Selain *core point*, ada juga *border point* dan *noise point*. *Border point* adalah titik yang tidak memenuhi syarat sebagai *core point*, tetapi berada dalam lingkaran ε dari *core point*. Sedangkan *noise point* adalah titik yang bukan *core* maupun *border*, sehingga dianggap sebagai *outlier*.



Gambar 1. Jenis titik dalam DBSCAN method

- 4) Proses ini berulang sampai tidak ada lagi titik yang dapat ditambahkan ke klaster.
- 5) Titik yang tidak dapat dimasukkan ke klaster manapun diklasifikasikan sebagai *noise*.

3.4. Output

Output/hasil dari *clustering* ini adalah daerah-daerah klaster dengan tingkat kesamaan tinggi pada intracluster dan kesamaan rendah pada intercluster. Kemudian pelabelan masing-masing *cluster* dengan ciri/karakteristik masing-masing *cluster* (aksesibilitas: tinggi – menengah – rendah) dari hasil DBSCAN, lalu dipetakan dan dideskripsikan.

3.5. Validasi internal (koefisien Silhouette)

Menurut Kaufman dan Rousseeuw dalam Bagirov et al, validasi hasil *cluster* dari metode DBSCAN dapat menggunakan nilai koefisien silhouette [20]. Nilai silhouette coefficient dapat menggunakan persamaan (2) berikut

$$SC = \frac{1}{n} \sum_{j=1}^n s(x_j) \quad (2)$$

Tabel 2. Kategori nilai silhouette coefficient

Silhouette coefficient	Kategori
$\leq 0,25$	Struktur <i>cluster</i> buruk
0,26 – 0,50	Struktur <i>cluster</i> lemah
0,51 – 0,70	Struktur <i>cluster</i> baik
0,71 – 1,00	Struktur <i>cluster</i> kuat

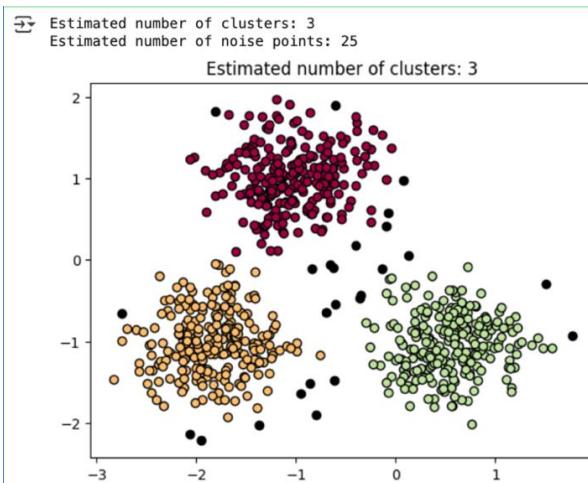
4. Hasil dan Pembahasan

4.1. Hasil Klasterisasi

Setelah melalui proses DBSCAN, didapatkan hasil 3 *cluster* utama dan beberapa titik *noise*. Dapat dilihat pada Gambar 2, *noise* disimbolkan dengan warna hitam. Pada Gambar 2

dapat dilihat bahwa estimasi jumlah *cluster* ada sebanyak 3 cluster. *Cluster* pertama ditandai dengan warna hijau, dan terkonsentrasi di daerah dengan aksesibilitas tinggi atau mudah mengakses layanan publik. *Cluster* kedua ditandai dengan warna kuning dengan aksesibilitas menengah/sedang atau cukup mudah mengakses layanan publik. Terakhir adalah *cluster* ketiga dengan warna merah adalah objek-objek di daerah pinggiran yaitu utara dan selatan Kabupaten Jombang.

Hasil clustering terbentuk dengan bentuk bebas/freeform (tidak selalu bulat/linear seperti *k-means*). *Noise* dapat diinterpretasikan sebagai wilayah yang terisolasi atau memiliki karakteristik unik. Hasil divisualisasikan dalam peta/diagram untuk interpretasi spasial. Penyajian peta tematik hasil *cluster* ada pada Gambar 2.



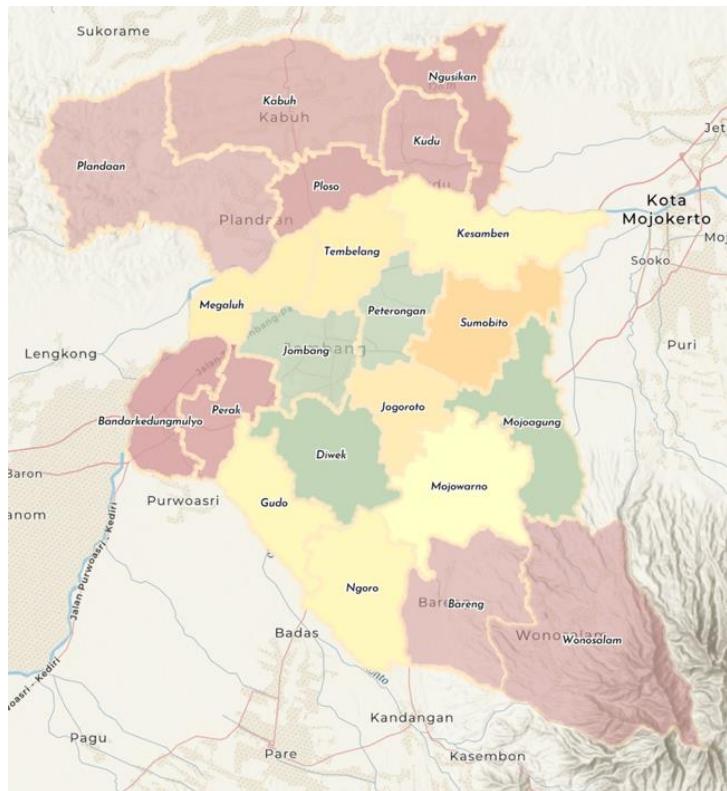
Gambar 2. Plotting titik koordinat *clustering*

Di samping itu, Gambar 3 merupakan penyajian hasil *cluster* ke dalam visualisasi bentuk peta tematik agar mudah dipahami dan sesuai kondisi nyata. Peta hasil *clustering* menggambarkan distribusi spasial setiap *cluster* beserta titik-titik yang termasuk *noise*. Visualisasi ini diinterpretasikan untuk memahami pola aksesibilitas, hambatan (*obstacles*), serta perbedaan antarwilayah. Selain visualisasi dengan peta, juga dilakukan analisis data secara naratif yang menjelaskan karakteristik masing-masing *cluster* serta rekomendasi awal pengembangan wilayah berdasarkan hasil penelitian.

Peta pada Gambar 3 memperlihatkan *cluster* 1 (warna hijau) terkonsentrasi di Kecamatan Jombang sebagai pusat pemerintahan. *Cluster* 2 (warna kuning) tersebar di daerah tengah dan utara mengelilingi sekitaran *cluster* 1. *Cluster* 3 (warna merah) di daerah selatan dengan *obstacle* pegunungan dan daerah utara dengan *obstacle* sungai besar. Berdasarkan Gambar 3, data hasil cluster algoritma DBSCAN menghasilkan 3 cluster utama dan beberapa titik *noise*. Titik berwarna merah merupakan titik koordinat fasilitas kesehatan meliputi rumah sakit dan puskesmas. Titik berwarna biru merupakan titik koordinat fasilitas pendidikan meliputi sekolah TK/RA, sekolah SD/MI, sekolah SMP/MTs/sederajat, sekolah SMA/MA/SMK/sederajat. Pada Gambar 2 juga menggambarkan jalan yang ada di Jombang. Aksesibilitas tinggi juga bisa dilihat dari kondisi kepadatan jalan. *Cluster* 1 merupakan daerah dengan kategori aksesibilitas tinggi yakni kecamatan Jombang, Diwek, Mojoagung dan Peterongan. Karakteristik dari cluster ini adalah lokasi dekat pusat kota, jumlah fasilitas publik (RS, sekolah, kantor layanan) tinggi, dan kepadatan penduduk tinggi. Interpretasi wilayah ini merupakan pusat pertumbuhan dan layanan publik pusat di Kabupaten Jombang.

Cluster 2 terdiri dari objek-objek dengan kategori aksesibilitas menengah yakni Jogoroto, Sumobito, Kesamben, Tembelang, Megaluh, Gudo, Ngoro, dan Mojowarno. Karakteristik dari cluster ini adalah memiliki fasilitas publik cukup, namun distribusinya belum merata atau lebih jarang jumlahnya dibandingkan *Cluster* 1. Kemudian akses jalan relatif baik, tetapi beberapa wilayah masih terkendala sungai besar. Interpretasi dari *cluster* ini adalah wilayah peralihan, dengan kebutuhan peningkatan pemerataan fasilitas kesehatan dan pendidikan serta kantor layanan. Terakhir, *Cluster* 3 yaitu kategori aksesibilitas rendah dengan anggota *cluster*

Kecamatan Ploso, Kudu, Kabuh, Ngusikan, Perak, Bandar kedungmulyo, Bareng dan Wonosalam. Karakteristik dari *cluster* ini adalah terletak di dataran tinggi/gunung atau daerah pinggiran dengan *obstacle* sungai sehingga membuat jarak memutar jalan lebih jauh. Kemudian fasilitas publik sangat terbatas, akses jalan sulit, dan hambatan geografis signifikan. Interpretasi untuk daerah ini adalah memerlukan prioritas pengembangan infrastruktur dasar agar aksesibilitas meningkat sesuai dengan kebutuhan tiap kecamatan.

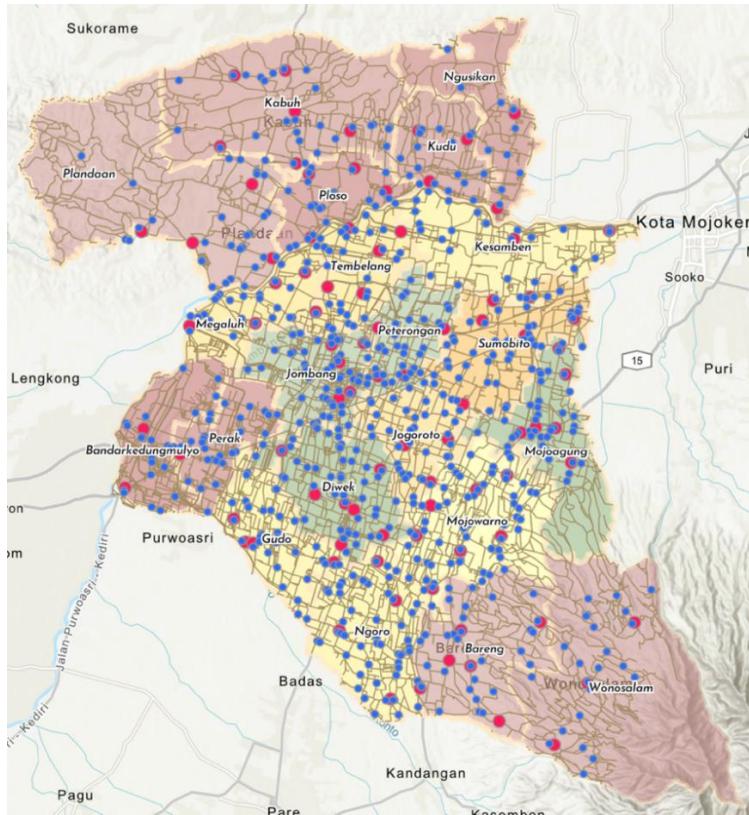


Gambar 3. Visualisasi hasil clustering kecamatan di Jombang berdasarkan aksesibilitas layanan publik

Berdasarkan Gambar 4, data hasil *cluster* algoritma DBSCAN menghasilkan 3 cluster utama dan beberapa titik *noise*. Titik berwarna merah merupakan titik koordinat fasilitas kesehatan meliputi rumah sakit dan puskesmas. Titik berwarna biru merupakan titik koordinat fasilitas pendidikan meliputi sekolah TK/RA, sekolah SD/MI, sekolah SMP/MTs/sederajat, sekolah SMA/MA/SMK/sederajat. Pada Gambar 2 juga menggambarkan jalan yang ada di Jombang. Aksesibilitas tinggi juga bisa dilihat dari kondisi kepadatan (*density*) jaringan jalan. *Cluster* 1 memiliki objek-objek dengan kategori aksesibilitas tinggi yakni kecamatan Jombang, Diwek, Mojoagung dan Peterongan. Karakteristik dari *cluster* ini adalah lokasi dekat pusat kota, jumlah fasilitas publik (RS, sekolah, kantor layanan) tinggi, dan kepadatan penduduk tinggi. Interpretasi wilayah ini merupakan pusat pertumbuhan dan layanan publik pusat di Kabupaten Jombang.

Kemudian *cluster* 2 dengan objek-objek kategori aksesibilitas menengah yakni Jogoroto, Sumobito, Kesamben, Tembelang, Megaluh, Gudo, Ngoro, dan Majowarno. Karakteristik dari cluster ini adalah memeliki fasilitas publik cukup, namun distribusinya belum merata atau lebih jarang jumlahnya dibandingkan Cluster 1. Kemudian akses jalan relatif baik, tetapi beberapa wilayah masih terkendala sungai besar. Interpretasi dari cluster ini adalah wilayah peralihan, dengan kebutuhan peningkatan pemerataan fasilitas kesehatan dan pendidikan serta kantor layanan. Terakhir, *Cluster* 3 yaitu kategori aksesibilitas rendah dengan anggota *cluster* Kecamatan Ploso, Kudu, Kabuh, Ngusikan, Perak, Bandar Kedungmulyo, Bareng dan Wonosalam. Karakteristik dari *cluster* ini adalah terletak di dataran tinggi/gunung atau daerah pinggiran dengan *obstacle* sungai sehingga membuat jarak memutar jalan lebih jauh. Kemudian

fasilitas publik sangat terbatas, akses jalan sulit, dan hambatan geografis signifikan. Interpretasi untuk daerah ini adalah memerlukan prioritas pengembangan infrastruktur dasar agar aksesibilitas meningkat sesuai dengan kebutuhan tiap kecamatan.



Gambar 4. Point of Interest (POI) pendidikan, kesehatan, ekonomi dan jaringan jalan Kabupaten Jombang tahun 2023-2024

4.2. Validasi Performa Algoritma

Setelah melakukan clusterisasi dan terbentuk cluster menurut DBSCAN, kemudian dilanjutkan dengan evaluasi untuk menguji ketepatan sebuah cluster yang telah terbentuk menggunakan *Silhouette Coefficient* (SC). Inti dari SC ini adalah tahapan perhitungan SC sebagai berikut: untuk setiap objek ke- i , yang merupakan anggota cluster C_g , dan jarak Euclidean persamaan (1) objek i dan j , D_{ij} , Perhitungan SW_i dan SC sebagai berikut:

$$SW_i = \frac{b_i - a_i}{\max \{a_i, b_i\}} ; \quad a_i = \frac{\sum_{j \in C_g (j \neq i)} D_{ij}}{n_g - 1} ; \quad b_i = \min_{h \neq g} \left(\frac{\sum_{j \in C_h (j \neq i)} D_{ij}}{n_h} \right) \quad (3)$$

n_g, n_h adalah jumlah objek pada cluster ke- g, h . Jika objek ke- i sendiri/tunggal dalam cluster tertentu, maka nilai a_i tidak bisa dihitung dan nilai $SW_i = 0$. Nilai SC didapatkan dari rata-rata nilai SW_i berikut:

$$SC = \frac{\sum_{i=1}^n SW_i}{n} \quad (4)$$

Nilai SC diantara interval $(-1,1)$. Semakin tinggi nilai berarti semakin baik hasil cluster. Jumlah cluster ditentukan Hasil pengujian akurasi menggunakan validasi internal koefisien Silhouette menurut referensi dari Bagirov, et al [20].

Tabel 3. Nilai yang diinputkan dan kategori struktur *cluster*

Silhouette coefficient	Kategori
$\leq 0,25$	Struktur cluster buruk
$0,26 - 0,50$	Struktur cluster lemah
$0,51 - 0,70$	Struktur cluster baik
$0,71 - 1,00$	Struktur cluster kuat

Dengan mengambil beberapa sampel data dari total 21 kecamatan di Kabupaten Jombang, akan dilakukan perhitungan koefisien Silhouette. Pada kasus ini, cluster pertama (aksesibilitas tinggi) objek kecamatan Jombang, kecamatan Peterongan dan Diwek selanjutnya disebut obyek 1-3. Lalu *cluster* kedua (aksesibilitas rendah) yaitu kecamatan Kabuh, selanjutnya disebut obyek 4.

Tabel 4. Nilai input dan jarak Euclidean untuk empat objek

Object	Nilai input		Jarak euclidean			
	Variabel fasilitas pendidikan	Variabel fasilitas kesehatan	Object 1	Object 2	Object 3	Object 4
1	20	28	0	16.279	18.788	20.615
2	23	12	16.279	0	5.099	8.944
3	28	11	18.788	5.099	0	13.342
4	15	8	20.615	8.944	13.342	0
:	:	:	:	:	:	:
21						

Sumber: perhitungan sendiri

Dengan menggunakan Python versi 3.11, kalkulasi koefisien silhouette untuk tiga *cluster* ditampilkan oleh Gambar 5 dan berdasarkan persamaan (3), diperoleh:

$$a_1 = \frac{D_{12} + D_{13}}{3 - 1} = \frac{16.279 + 18.788}{2} = 17.534, \quad b_1 = D_{14} = 20.615,$$

$$SW_1 = \frac{20.615 - 17.534}{20.615} = 0.149$$

Lakukan langkah serupa, sehingga diperoleh $SW_1, SW_2, SW_3, SW_4, \dots, SW_{21}$. Kemudian hitung silhouette coefficient, SC, yang merupakan rerata dari nilai $SW_1, SW_2, SW_3, SW_4, \dots, SW_{21}$.

$$SC = \frac{SW_1 + SW_2 + SW_3 + SW_4 + \dots + SW_{21}}{21} = \frac{0.149 + 0.78 + 0.676 + \dots + 0.695}{21} = 0.808 \approx 0.81$$

```
from sklearn.cluster import DBSCAN
from matplotlib import pyplot as plt
import numpy as np
from sklearn.datasets import make_blobs

X, _ = make_blobs(n_samples=100, centers=3, cluster_std=0.5, random_state=42) # Decreased cluster_std

X

Tampilkan output tersembunyi

# Resetting epsilon and min_points and performing DBSCAN on the new data
epsilon = 0.5
min_points = 5

dbSCAN = DBSCAN(eps=epsilon, min_samples=min_points)
clusters = dbSCAN.fit_predict(X)

# Print the cluster labels for each point
print("Cluster labels:", clusters)

Tampilkan output tersembunyi

from sklearn.metrics import silhouette_score
silhouette_score(X, clusters)

np.float64(0.8084301812478875)
```

Gambar 5. Perhitungan koefisien silhouette untuk tiga cluster pada Python 3.11

Validasi internal dengan persamaan (3) dan (4) menghasilkan nilai koefisien Silhouette 0,81 yang termasuk dalam bentuk *cluster* kuat, berdasarkan Tabel 3. Validasi dilakukan menggunakan tools Python versi 3.11 yang ditunjukkan oleh Gambar 5. Indeks yang peka terhadap kepadatan (*density-aware*) menunjukkan struktur *cluster* yang berarti DBCV positif.

Indeks globular yang digunakan adalah koefisien silhouette yang cenderung moderat sesuai ekspektasi untuk cluster non-spheric, namun tetap konsisten. Kemudian dari hasil inti *cluster* relative stabil pada variasi epsilon ε dan minPts, dengan kata lain pergeseran terjadi pada titik perbatasan (*border point*) bukan pada inti (*core point*). Deteksi *noise* yang dihasilkan oleh DBSCAN bermanfaat untuk mengidentifikasi daerah dengan hambatan topografi atau jaringan jalan sehingga membantu memetakan masalah akses.

4.3. Pembahasan

Hasil *clustering* menggunakan DBSCAN berhasil menemukan tiga *cluster* dengan bentuk tak beraturan dan *noise*, selaras dengan pola geografis dan topografi wilayah pegunungan dan koridor sungai besar masuk ke dalam cluster rendah. Sedangkan pusat kabupaten dengan POI terpadat masuk dalam *cluster* tinggi. Berdasarkan hasil interpretasi klasterisasi menggunakan algoritma DBSCAN, wilayah Kabupaten Jombang dapat dikelompokkan menjadi tiga zona aksesibilitas layanan publik, yaitu klaster aksesibilitas tinggi, sedang, dan rendah. Penemuan ini menunjukkan adanya disparitas spasial yang signifikan antar kecamatan, terutama antara kawasan utara–selatan serta barat–timur, yang dipengaruhi oleh persebaran fasilitas pendidikan serta kesehatan, kondisi geografis, dan jaringan jalan.

Untuk mengatasi ketimpangan tersebut, beberapa rekomendasi konstruktif dapat diajukan sebagai berikut: (1) optimalisasi perencanaan spasial berbasis data yaitu pemerintah daerah perlu memanfaatkan hasil klasterisasi sebagai dasar untuk alokasi anggaran berbasis wilayah prioritas. Penggunaan model klasterisasi spasial terbukti cukup efektif mendukung evidence-based policy dan perencanaan infrastruktur publik, sejalan dengan riset Schubert et al (2017) dan Han et al (2023) yang menegaskan efektivitas analisis spasial dalam mendukung perencanaan pembangunan basis data [21], [22] ; (2) pengembangan infrastruktur pendidikan, ekonomi dan kesehatan di wilayah akses rendah dan hambatan geografis (lereng, sungai besar), pemerintah dapat menerapkan pendekatan mobile service atau layanan terpadu bergerak (misalnya puskesmas keliling, SIM keliling, sekolah satelit, dll) untuk meningkatkan indeks aksesibilitas; (3) pemanfaatan teknologi geospasial (GIS) yang terintegrasi dengan *machine learning* untuk memperbarui hasil klasterisasi secara periodik, selaras dengan hasil Ramadhan et al [9]; dan (4) kolaborasi lintas sektor untuk menutup gap layanan publik.

Hasil temuan dalam penelitian ini berkontribusi signifikan terhadap pengembangan ilmu *clustering* dalam topik data mining, khususnya dalam penerapan algoritma DBSCAN pada data sosial ekonomi dan spasial. Penelitian ini memperluas bidang ilmu dengan menunjukkan bahwa DBSCAN mampu mengelompokkan wilayah berdasarkan pola kepadatan data dan faktor geospasial secara adaptif, tanpa harus menentukan jumlah klaster sejak awal seperti pada algoritma *k-means* atau *hierarchical clustering*. Dengan demikian, penelitian ini memberikan kontribusi empiris terhadap penguatan pemahaman bahwa pendekatan berbasis kepadatan (*density-based*) lebih representatif dalam mengatasi heterogenitas data wilayah dan variasi distribusi spasial, yang sering kali dihadapi dalam analisis sosial ekonomi dan pelayanan publik.

Selain itu, penelitian ini memberikan kontribusi metodologis melalui optimalisasi parameter epsilon dan *minimum points*, MinPts untuk menyesuaikan performa DBSCAN terhadap karakteristik data real. Pengujian dilakukan dengan berbagai kombinasi parameter, yang menghasilkan koefisien Silhouette sebesar 0,81 (telah dijelaskan pada sub bab 4.2), menunjukkan kualitas pemisahan klaster yang kuat serta tingkat homogenitas internal yang tinggi. Nilai tersebut memperkuat posisi DBSCAN sebagai algoritma yang tangguh dalam mengenali pola data non-linear dan menangani outlier secara efisien. Dengan hasil tersebut, penelitian ini memperluas pemahaman praktis tentang bagaimana *tuning* parameter dapat meningkatkan akurasi model klasterisasi dan sebagai solusi peningkatan sensitivitas terhadap variasi kepadatan data, sebagaimana direkomendasikan oleh Schubert, et al [21].

Dari sisi pengembangan aplikasi keilmuan, hasil penelitian ini berpotensi menjadi model *benchmarked* untuk studi-studi klasterisasi berbasis spasial di bidang kebijakan publik, tata kelola wilayah, dan perencanaan pembangunan. Implementasi DBSCAN pada konteks data sosial ekonomi dan spasial menghasilkan visualisasi pola keterjangkauan layanan publik yang lebih akurat dan berbasis data, sehingga dapat digunakan sebagai dasar ilmiah dalam menentukan prioritas pembangunan di wilayah dengan aksesibilitas rendah. Dengan demikian, kontribusi penelitian ini tidak hanya memperkaya literatur akademik dalam bidang data mining, tetapi juga memberikan model terapan yang dapat direplikasi dalam konteks lain dengan karakteristik data

serupa. Secara konseptual, penelitian ini juga memperkuat posisi algoritma DBSCAN sebagai metode analisis adaptif untuk data kompleks dan tidak beraturan, sekaligus menegaskan pentingnya integrasi antara analisis data kuantitatif dan konteks spasial dalam data-driven decision making. Hasil temuan ini diharapkan membuka peluang penelitian lanjutan untuk mengembangkan *hybrid clustering* yang menggabungkan DBSCAN dengan algoritma lain untuk meningkatkan sensitivitas terhadap variasi kepadatan dan kompleksitas spasial.

5. Simpulan

Hasil penelitian ini menunjukkan bahwa penerapan metode *clustering* terhadap indikator aksesibilitas layanan publik di Kabupaten Jombang berhasil mengidentifikasi perbedaan karakteristik wilayah yang signifikan. Wilayah dengan aksesibilitas rendah (*cluster* merah) umumnya berada di daerah perbatasan dan pegunungan dengan hambatan geografis serta keterbatasan infrastruktur, sementara wilayah dengan aksesibilitas tinggi (*cluster* hijau) terkonsentrasi di sekitar pusat kota yang memiliki kepadatan penduduk, fasilitas pendidikan, kesehatan, dan ekonomi yang relatif lengkap. Adapun wilayah dengan kategori sedang (kuning dan orange) berfungsi sebagai zona transisi yang berpotensi menjadi simpul penghubung antarwilayah. Hasil ini memperkuat pemahaman bahwa ketimpangan aksesibilitas tidak hanya dipengaruhi oleh faktor jumlah penduduk atau luas wilayah, tetapi juga kondisi topografi, distribusi sarana prasarana, serta posisi strategis terhadap jalur transportasi utama. Dengan demikian, penelitian ini berkontribusi pada pengembangan konsep *evidence-based spatial planning* yang menempatkan hasil analisis spasial sebagai dasar penentuan prioritas pembangunan.

Prospek pengembangan ke depan terletak pada penerapan model ini secara lebih luas, baik dengan menambahkan variabel baru seperti akses digital, tingkat ekonomi rumah tangga, maupun risiko bencana, sehingga peta klasterisasi yang dihasilkan dapat menjadi instrumen strategis dalam perumusan kebijakan daerah yang lebih inklusif dan berkeadilan. Pemerataan layanan publik di wilayah dengan aksesibilitas rendah, penguatan zona transisi, serta pengembangan pusat layanan spesialis di wilayah akses tinggi menjadi langkah yang saling melengkapi untuk mewujudkan tata kelola pembangunan daerah yang berorientasi pada pemerataan kesejahteraan masyarakat.

Ucapan Terima Kasih

Penulis mengucapkan terima kasih kepada Direktorat Penelitian dan Pengabdian kepada Masyarakat, Direktorat Jenderal Riset dan Pengembangan, Kementerian Pendidikan Tinggi, Sains dan Teknologi Republik Indonesia yang telah memberikan dukungan pendanaan.

Daftar Referensi

- [1] N. M. Mahfuz, M. Yusoff, and Z. Idrus, "Clustering Heterogeneous Categorical Data Using Enhanced Mini Batch K-Means With Entropy Distance Measure," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 13, no. 1, pp. 1048–1059, Feb. 2023, doi: <http://doi.org/10.11591/ijece.v13i1.pp1048-1059>.
- [2] Badan Pusat Statistik Kab Jombang, "Kabupaten-Jombang-dalam-Angka-2025," vol. 45, pp. 10–672, 2025.
- [3] Badan Pusat Statistik Kabupaten Jombang, "Profil Kemiskinan di Kabupaten Jombang Maret 2024," Dec. 2024.
- [4] Y. Lv *et al.*, "An Efficient and Scalable Density-based Clustering Algorithm for Datasets with Complex Structures," *Journal of Neurocomputing*, vol. 171, no. 1, pp. 9–22, May 2015, doi: [dx.doi.org/10.1016/j.neucom.2015.05.109](https://doi.org/10.1016/j.neucom.2015.05.109).
- [5] A. Arranz-López, J. A. Soria-Lara, F. Witlox, and A. Páez, "Measuring relative non-motorized accessibility to retail activities," *Int J Sustain Transp*, vol. 13, no. 9, pp. 639–651, Oct. 2019, doi: [10.1080/15568318.2018.1498563](https://doi.org/10.1080/15568318.2018.1498563).
- [6] K. Cao, C. Zhou, R. Church, X. Li, and W. Li, "Revisiting spatial optimization in the era of geospatial big data and GeoAI," *International Journal of Applied Earth Observation and Geoinformation*, vol. 129, p. 103832, 2024, doi: <https://doi.org/10.1016/j.jag.2024.103832>.
- [7] X. Tu, C. Fu, A. Huang, H. Chen, and X. Ding, "DBSCAN Spatial Clustering Analysis of Urban 'Production–Living–Ecological' Space Based on POI Data: A Case Study of Central Urban Wuhan, China," *Int J Environ Res Public Health*, vol. 19, no. 9, pp. 5153–5174, Apr. 2022, doi: doi.org/10.3390/ijerph19095153.

- [8] G. Tóth and Á. Kincses, "Accessibility models based on the gravity analogy: In theory and practice," *Regional Statistics*, vol. 5, no. 1, pp. 137–158, 2015, doi: 10.15196/RS05108.
- [9] R. Ramadhan, A. W. Wijayanto, and S. Pramana, "A Geospatial Big Data Approaches to Estimate Granular Level Poverty Distribution in East Java, Indonesia using Machine Learning and Deep Learning Regressions," in *Proceeding of 2023 International Conference On Data Science and Official Statistics (ICDSOS)*, Jakarta: Politeknik Statistika, Dec. 2023, p. 186200. doi: doi.org/10.34123/icdsos.v2023i1.359.
- [10] X. Du, Y. Du, Y. Zhang, Y. Zhu, and Y. Yang, "Urban And Rural Disparities in General Hospital Accessibility Within a Chinese Metropolis," *Sci Rep*, vol. 14, no. 1, p. 23359, Dec. 2024, doi: 10.1038/s41598-024-74816-4.
- [11] W. Li, H. Li, and Y. Luo, "Dynamic and Static Enhanced BIRCH for Functional Data Clustering," *IEEE Access*, vol. 11, pp. 111448–111465, Oct. 2023, doi: 10.1109/ACCESS.2023.3322929.
- [12] D. Syaputri, P. Herwina Noprita, and S. Romelah, "Implementation of K-Means Algorithm for Economic Distribution Clustering Base on Demographics of Population," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 1, no. 1, pp. 1–6, Apr. 2021.
- [13] A. Z. Hardi and A. K. A. Murad, "Spatial Analysis of Accessibility for Public Transportation, A Case Study in Jakarta, Bus Rapid Transit System (Transjakarta), Indonesia," *Journal of Computer Science*, vol. 19, no. 10, pp. 1190–1202, 2023, doi: 10.3844/jcssp.2023.1190.1202.
- [14] A. Comber *et al.*, "A Route Map for Successful Applications of Geographically Weighted Regression," *Geogr Anal*, vol. 55, no. 1, pp. 155–178, Jan. 2023, doi: 10.1111/gean.12316.
- [15] H. A. Salman, L. F. Ibrahim, and Z. Fayed, "Enhancing clustering technique with knowledge-based system to plan the social infrastructure services," in *ICAART 2013 - Proceedings of the 5th International Conference on Agents and Artificial Intelligence*, 2013, pp. 401–408. doi: 10.5220/0004391504010408.
- [16] K. N. Ahmed and T. A. Razak, "An Overview of Various Improvements of DBSCAN Algorithm in Clustering Spatial Databases An Overview of Various Improvements of DBSCAN Algorithm in Clustering Spatial Databases," *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 5, no. 2, 2016.
- [17] Y. Zhang, "Large Data Oriented to Image Information Fusion Spark and Improved Fruit Fly Optimization Based on the Density Clustering Algorithm," *Advances in Multimedia*, vol. 2023, pp. 1–14, Mar. 2023, doi: 10.1155/2023/5596605.
- [18] D. A. Keatley *et al.*, "Behaviour Tracking: using Geospatial and Behaviour Sequence Analysis to Map Crime," *Security Journal*, vol. 34, no. 1, pp. 184–201, Mar. 2021, doi: <https://doi.org/10.1057/s41284-019-00216-3>.
- [19] S. Miftahurrahmi, Zilrahmi, N. Amalita, and T. O. Mukhti, "DBSCAN Method in Clustering Provinces in indonesia Based on Crime Cases in 2022," *UNP Journal of Statistics and Data Science*, vol. 2, no. 3, pp. 330–337, Aug. 2024.
- [20] A. M. Bagirov, R. M. Aligulyev, and N. Sultanova, "Finding compact and well-separated clusters: Clustering using silhouette coefficients," *Journal of Pattern Recognition*, vol. 135, no. 1, pp. 109144–109158, Nov. 2023, doi: 10.1016/j.patcog.2022.109144.
- [21] E. Schubert, J. Sander, M. Ester, H. P. Kriegel, and X. Xu, "DBSCAN revisited, revisited: Why and how you should (still) use DBSCAN," *Journal of ACM Transactions on Database Systems*, vol. 42, no. 3, pp. 1–21, Jul. 2017, doi: <https://doi.org/10.1145/3068335>.
- [22] J. Han, J. Pei, and H. Tong, *Data Mining: Concepts and Techniques*, 4th ed. Cambridge, US: Morgan Kaufmann Publishers, Elsevier Inc, 2023. Accessed: Oct. 15, 2025. [Online]. Available: www.elsevier.com/books-and-journals