

Analisis Pemilihan Parameter pada Algoritma DBSCAN untuk Pengelompokan Titik Api di Indonesia

Bernadeta Avrilya Hayung Driyandita¹, I Putu Eka Nila Kencana^{*2}, I Gusti Ngurah Lanang Wijayakusuma³

^{1,2,3}Matematika, Fakultas MIPA, Universitas Udayana, Indonesia

Email: ¹bernadetaavrilya@gmail.com, ²i.putu.enk@unud.ac.id, ³lanang_wijaya@unud.ac.id

Abstrak

Kebakaran hutan, gambut, wilayah pertanian, maupun wilayah urban dan industri menjadi salah satu ancaman terbesar bagi lingkungan, menyebabkan kerusakan ekosistem dan kerugian ekonomi yang signifikan. Penelitian ini mengusulkan optimalisasi pengelompokan titik api di Indonesia sebagai salah upaya pencegahan kebakaran dengan menggunakan algoritma *Density-Based Spatial Clustering of Applications with Noise* (DBSCAN). Metodologi yang digunakan dalam penelitian mencakup pemrosesan data, pemilihan parameter tuning algoritma, dan evaluasi hasil kluster menggunakan metrik validasi internal yaitu *Silhouette Coefficient* (SC). Data penelitian diambil dari web FIRMS NASA berjumlah 12.708, dengan variabel yang digunakan yaitu *longitude*, *latitude*, *bright_t31*, *brightness*, *fire radiative power*, *scan*, dan *track*. Pada tahap pemrosesan dilakukan pembersihan dan standarisasi data memakai algoritma *z-score*. Selanjutnya percobaan klasterisasi dijalankan dengan mengubah-ubah nilai pada parameter epsilon (eps) sebesar 0.2, 2, 3, 4, 5, dan 10, yang dikombinasikan dengan nilai minimal point (MinPts) dari 2 sampai dengan 6. Hasil pengelompokan optimal yang ditemukan adalah pada saat percobaan dengan nilai eps 4 dan MinPts 4. Kondisi optimal tersebut menghasilkan 2 kluster dan 3 noise, dengan nilai SC sebesar 0.8022. Penelitian ini memberikan wawasan baru dalam pemetaan risiko kebakaran serta dapat digunakan untuk sistem pemantauan berbasis kecerdasan buatan.

Kata kunci: DBSCAN, klasterisasi, *silhouette coefficient*, titik panas

Analysis of Hotspot Clustering Optimization Using DBSCAN Algorithm

Abstract

Forest, peat, agricultural, and urban and industrial fires are some of the biggest environmental threats, causing significant ecosystem damage and economic losses. This study proposes the optimization of hotspot clustering in Indonesia as one of the fire prevention efforts using the *Density-Based Spatial Clustering of Applications with Noise* (DBSCAN) algorithm. The methodology used in the study includes data processing, selection of algorithm tuning parameters, and evaluation of cluster results using internal validation metrics, namely the *Silhouette Coefficient* (SC). The research data was taken from the NASA FIRMS website totaling 12.708, with the variables used being *longitude*, *latitude*, *bright_t31*, *brightness*, *fire radiative power*, *scan*, and *track*. At the processing stage, data cleaning and standardization were carried out using the *z-score* algorithm. Furthermore, the clustering experiment was run by changing the value of the epsilon parameter (Eps) by 0.2, 2, 3, 4, 5, and 10, which was combined with the minimum point value (MinPts) from 2 to 6. The optimal clustering results found were during the experiment with Eps values of 4 and MinPts 4. These optimal conditions produced 2 clusters and 3 noises, with an SC value of 0.8022. This study provides new insights into fire risk mapping and can be used for artificial intelligence-based monitoring systems.

Keywords: clustering, DBSCAN, hotspot, *silhouette coefficient*

1. PENDAHULUAN

Kebakaran merupakan peristiwa yang melibatkan pembakaran bahan yang mudah terbakar dan berlangsung tidak terkendali. Menurut NFPA (National Fire Protection Association), kebakaran merupakan suatu peristiwa oksidasi yang melibatkan tiga unsur penting yaitu bahan bakar yang mudah terbakar, oksigen, dan sumber panas. Kebakaran terjadi karena adanya pancaran api sejak awal kebakaran hingga penjalaran api, asap, dan gas yang ditimbulkan [1]. Kebakaran yang tidak terkendali dapat mengancam bahkan merusak lingkungan disekitarnya. Dampak kebakaran pada kondisi lingkungan hidup yaitu vegetasi lahan yang berkurang dan penurunan produksi,

serta polusi udara dan sumber air yang berkurang [2]. Kualitas udara yang buruk dan kurangnya sumber air ini dapat mempengaruhi dan mengganggu kesehatan masyarakat.

Kebakaran hutan tidak hanya menjadi masalah lokal, tetapi juga menjadi isu global yang signifikan seperti peristiwa kebakaran yang terjadi di Los Angeles, California yang bermula terjadi pada 2 Januari 2025 dan mulai meluas secara signifikan hingga pada 7 Januari 2025. Diketahui terjadi tiga kebakaran utama yaitu kebakaran Palisades, Eaton, dan Hurst. Menurut Cal Fire, lebih dari 38.000 hektar lahan terbakar dan lebih dari 12.000 bangunan hancur akibat kebakaran hutan [3]. Mark Gold, Direktur Solusi Kelangkaan Air di Natural Resource Defense Council menyatakan penyebab kebakaran besar di Los Angeles dikarenakan kombinasi dari fenomena Angin Santa Ana yang kencang, musim kemarau yang ekstrem sehingga vegetasi menjadi sangat kering, suhu yang tinggi, serta rendahnya kelembapan sehingga menciptakan kondisi yang rawan kebakaran.

Peristiwa kebakaran yang terjadi di California, dimana setiap tahunnya mengalami peristiwa kebakaran akibat kombinasi cuaca ekstrem dan perubahan iklim, memiliki beberapa kesamaan dengan karakteristik kebakaran yang terjadi di Indonesia. Kedua wilayah ini menghadapi tantangan serupa dalam hal kerusakan ekosistem, peningkatan emisi karbon, serta dampak buruk terhadap kualitas udara dan kesehatan masyarakat. Meskipun penyebab kebakaran di Indonesia juga dikarenakan kombinasi faktor yang berbeda dengan di California, tetap diperlukan upaya mitigasi dan pengelolaan kebakaran yang lebih baik dengan berbasis data dan teknologi. Badan Nasional Penanggulangan Bencana (BNPB) menyatakan tercatat ada 629 kasus kebakaran hutan dan lahan (karhutla) di sejumlah daerah di Indonesia sepanjang tahun 2024 [4]. Berdasarkan data Kementerian Lingkungan Hidup dan Kehutanan (KLHL) RI yakni melalui web [5] total luas kebakaran yang terjadi di Indonesia tahun 2024 yakni 376.805,05 hektar.

Upaya mitigasi dengan memantau dan mendeteksi titik api atau *hotspot* dapat menjadi salah satu upaya penanggulangan kebakaran. Kebakaran aktif dapat dipantau dan dideteksi melalui satelit jarak jauh [6]. Dalam hal ini, data penginderaan jauh yang disediakan oleh Fire Information for Resource Management System (FIRMS) dari NASA dapat menjadi alat penting untuk memantau titik panas. FIRMS menyediakan data yang mencakup lokasi, intensitas, dan cakupan area yang sangat berguna dalam analisis kebakaran. Namun, data perlu dilakukan pemrosesan lebih lanjut untuk dapat mengidentifikasi pola dan kelompok titik panas yang relevan sehingga dapat mempermudah dalam mengamati area kebakaran dan karakteristiknya.

Salah satu pendekatan untuk menganalisis dan mengantisipasi kebakaran menggunakan pendekatan matematis yakni dengan mengintegrasikan teknologi analisis kluster. Analisis metode kluster yang populer digunakan yaitu *density-based spatial clustering of application with noise* (DBSCAN) dan metode ini berfokus pada kepadatan. Pengelompokan data *hotspot* dapat dengan baik dianalisis menggunakan metode berbasis kepadatan, karena fleksibilitas dalam mengidentifikasi kelompok dengan berbagai bentuk dan ukuran [6].

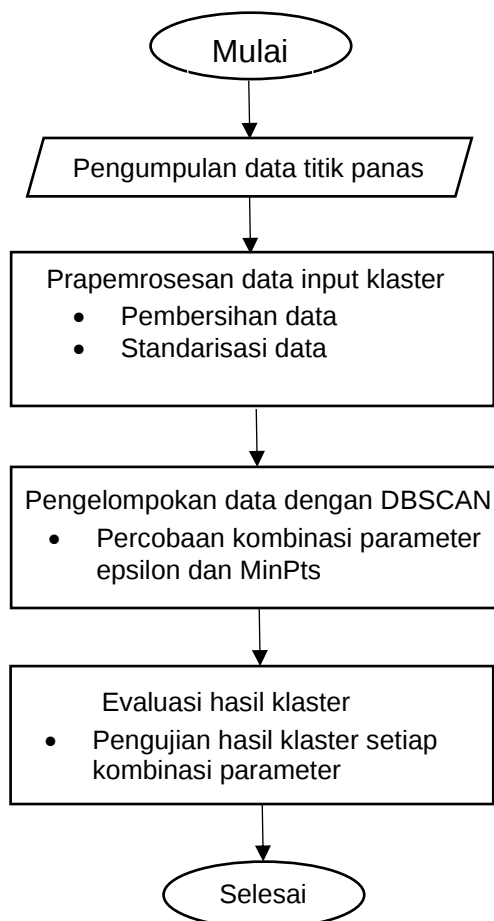
Pada beberapa penelitian sebelumnya telah dibuktikan bahwa algoritma DBSCAN mampu melakukan pengelompokan data dengan baik terutama dalam konteks data spasial maupun non-spasial. Seperti penelitian Utami et. al (2023) yang memanfaatkan DBSCAN untuk menganalisis pola pengelompokan titik panas, dengan hasil pengujian validasi kluster mendapatkan nilai *silhouette coefficient* terbaik yaitu 0,876 [7]. Penerapan algoritma DBSCAN juga dilakukan oleh Fauzan et al (2022) dalam mengelompokkan secara spasial hotel di Bali, dengan hasil evaluasi diperoleh nilai rata – rata lebar *silhouette* sebesar 0,709 atau termasuk dalam kategori struktur kuat dalam klusterisasi spasial [8]. Sementara Nurhaliza dan Mustakim (2021) mengelompokkan 197 negara di dunia berdasarkan data kasus Covid-19 dengan algoritma DBSCAN, dengan melakukan percobaan sebanyak 22 kali dan memiliki nilai SI yang paling mendekati 1 yaitu sebesar 0,326 sehingga dapat dikatakan sebagai kluster paling optimal [9]. Perbandingan dalam penerapan algoritma DBSCAN dengan algoritma kluster lainnya dilakukan oleh Sigit Susanto Putro et al. (2023) yakni untuk pengelompokan kualitas garam. Nilai kluster yang tertinggi diperoleh metode DBSCAN sebesar 0,5 sehingga struktur *silhouette coefficient* lebih kuat dibanding dengan K-means dan AHC dengan nilai berturut – turut 0,345 dan 0,32 [10]. Penelitian oleh Niko Purnomo dan Windu Gata (2024) juga membandingkan DBSCAN dengan algoritma kluster lain untuk mengelompokkan analisis sentimen komentar *Youtube* terhadap pengambilalihan jalan rusak di Lampung. Hasil penelitian menunjukkan skor *silhouette* untuk K-means sebesar -0,348, DBSCAN sebesar 0,836, dan HDBSCAN sebesar 0,106 [11]. Artinya pada penelitian tersebut DBSCAN terbukti lebih unggul.

Keberhasilan DBSCAN dalam menghasilkan kluster yang berkualitas sering kali ditunjukkan melalui nilai *silhouette coefficient* sebagai metrik evaluasi. Berdasarkan hal ini, penelitian ini akan menggunakan algoritma DBSCAN untuk mengelompokkan data titik panas di Indonesia tahun 2023 dengan 7 atribut yaitu *longitude*, *latitude*, *bright_t31*, *brightness*, *fire radiative power*, *confidence*, *scan*, dan *track*. Pendekatan ini diharapkan dapat menghasilkan kluster yang sesuai dan informatif dalam mengidentifikasi karakteristik setiap kluster.

Tujuan penelitian ini adalah untuk mengeksplorasi efektivitas algoritma DBSCAN dalam mengelompokkan titik panas di Indonesia berdasarkan parameter tertentu, serta menentukan kombinasi parameter yang menghasilkan klusterisasi paling optimal.

2. METODE PENELITIAN

Pada penelitian ini terdapat beberapa tahapan yang perlu dilakukan untuk mencapai tujuan dalam menyelesaikan permasalahan. Adapun proses pengelompokan data titik panas digambarkan melalui diagram alur penelitian yang dapat dilihat pada Gambar 1.



Gambar 1. Diagram Alir Penelitian

2.1. Pengumpulan Data Titik Panas

Data yang digunakan pada penelitian ini merupakan data sekunder yang diambil dari Fire Information of Resource Management System (FIRMS) pada web National Aeronautics and Space Administration (NASA). NASA FIRMS menyajikan informasi hasil pendeteksian titik api aktif dan *thermal anomalies* menggunakan satelit MODIS dan VIIRS secara *near real-time*. Pengambilan data dilakukan dengan memasukkan koordinat *longitude* dan *latitude* wilayah Indonesia pada pilihan *custom region* pada bagian fitur *fire alerts*. Terdapat 15 variabel pada web FIRMS NASA titik api, tetapi pada penelitian ini hanya digunakan tujuh variabel yaitu *longitude*, *latitude*, *brightness*, *confidence*, *fire radiative power (frp)*, *bright_t31*, *scan*, dan *track* [7][12][13]. Pengelompokan data titik panas menggunakan algoritma DBSCAN dilakukan dengan bantuan *google colab*.

2.2. Prapemrosesan Data Input Kluster

Pada tahapan ini dilakukan pembersihan data dengan menangani data yang hilang dan menggantikan nilai kosong serta melakukan penghapusan duplikat data untuk menghasilkan pengelompokan dataset yang akurat, sehingga analisis tidak dipengaruhi oleh informasi yang berulang.

Tahapan pemrosesan berikutnya adalah melakukan transformasi data melalui standarisasi. Hal ini dilakukan untuk memastikan setiap variabel memberikan kontribusi untuk hasil pengelompokan secara seimbang. Standarisasi data dilakukan dengan metode *z-score* [14] yang dirumuskan pada persamaan 1.

$$z = \frac{x_i - \text{center}(x)}{\text{scale}(x)} \quad (1)$$

Persamaan 1 digunakan untuk mengubah nilai rata – rata setiap variabel sama dengan 0 dan standar deviasi sama dengan satu. Variabel Z merupakan nilai setiap variabel setelah dilakukannya standarisasi dengan proses perhitungan yaitu nilai asli pada variabel (x_j) dikurangi dengan nilai rata – rata dari seluruh nilai pada variabel (μ_j) lalu dibagi dengan nilai standar deviasi pada variabel (σ_j).

2.3. Pengelompokan Data Titik Panas

Density-based spatial clustering of application with noise (DBSCAN) merupakan salah satu algoritma pengelompokan atau kluster yang populer. DBSCAN diperkenalkan oleh Ester dan tim pada tahun 1996 dengan membawa pendekatan baru yaitu pengelompokan yang berfokus pada kepadatan data. Konsep kepadatan yang dimaksud dalam DBSCAN adalah jumlah minimal data dalam radius Epsilon (ϵ) dan data tersebut termasuk dalam kategori kepadatan yang diinginkan [15].

Adapun tahapan pengelompokan berbasis kepadatan dengan algoritma DBSCAN dijelaskan sebagai berikut.

- 1) Memberikan inisialisasi pada faktor input yakni nilai Epsilon (ϵ) yang mengatur jarak maksimum antara dua titik untuk dianggap berada dalam satu kluster dan nilai *minimal points* ($MinPts$) yang mengatur jumlah minimum titik yang diperlukan untuk membentuk kluster dengan syarat nilai ϵ dan $MinPts > 0$
- 2) DBSCAN memilih titik awal secara acak yaitu $\vec{p} \in D$ kemudian menggunakan radius ϵ untuk menentukan tetangga dari titik $\vec{p}$ atau secara matematis mencari ($N_\epsilon(\vec{p})$). Titik $\vec{p} \in D$ diperiksa untuk menentukan apakah titik tersebut dapat memulai pembuatan kluster sebagai *core point*. Proses ini dilakukan dengan menghitung jarak densitas terjangkau langsung dengan titik lain terhadap $\vec{p}$ menggunakan rumus ukuran jarak yang populer digunakan yaitu *euclidean distance* [16] pada persamaan 2.

$$d(i, j) = \sqrt{(x_{i1} - x_{j1})^2 + (x_{i2} - x_{j2})^2 + \dots + (x_{ik} - x_{jk})^2} \quad (2)$$

Persamaan diatas d_{pq} adalah jarak antara objek titik ke- p dan objek titik ke- q , dengan n adalah banyaknya objek titik yang diamati, j adalah banyaknya variabel yang diamati, v_{pi} adalah nilai pada amatan objek titik ke- p untuk variabel ke- i , v_{qi} adalah nilai pada amatan objek titik ke- q untuk variabel ke- i .

- 3) Jika $N_\epsilon(\vec{p}) < minPts$, maka titik tersebut tidak dapat menjadi titik inti dan sementara dianggap sebagai noise. Namun, titik ini masih bisa diintegrasikan ke dalam kluster lain, jika dalam proses iterasi titik yang lain, titik $\vec{p}$ dapat dijangkau.
- 4) Jika titik $\vec{p}$ memenuhi syarat $|N_\epsilon(\vec{p})| \geq minPts$, maka titik $\vec{p}$ dapat dianggap sebagai *core point* dan terbentuk kluster awal C . Titik – titik yang berada dalam lingkungan radius ϵ titik $\vec{p}$, secara notasi dilambangkan sebagai $N_\epsilon(\vec{p})$ di iterasi dan dilakukan pemeriksaan dan menggunakan konsep hubungan kedekatan untuk mencari titik tetangga lain dari sekumpulan titik pada $N_\epsilon(\vec{p})$.
- 5) Untuk setiap titik $\vec{q} \in N_\epsilon(\vec{p})$, jika $|N_\epsilon(\vec{q})| \geq minPts$, maka tambahkan q ke kluster C sebagai *core point* dan perluas lagi kluster dengan mencari titik tetangga lain dari sekumpulan titik dalam $N_\epsilon(\vec{q})$. Proses ini berlanjut hingga tidak ada lagi *core point* baru dalam radius ϵ yang dapat diperluas dari titik yang dipilih.
- 6) Jika $|N_\epsilon(\vec{q})| < minPts$ namun $q \in N_\epsilon(\vec{p})$ maka tetapkan q sebagai *border point*. *Border point* tetap dimasukkan dalam kluster tetapi tidak berperan dalam perluasan kluster lebih lanjut.
- 7) Jika sudah tidak ada titik densitas terjangkau langsung (*directly density reachable*) dari titik p , artinya satu kluster sudah selesai terbentuk, maka DBSCAN mencari titik lain dalam dataset yang belum pernah diproses. Titik berikutnya dipilih secara acak dari sisa titik yang belum dikunjungi dan dilakukan perulangan proses dari langkah b-g hingga semua titik diproses.
- 8) Pada akhir proses, DBSCAN menghasilkan kumpulan kluster dengan titik yang saling terkait dalam radius ϵ dan memenuhi kriteria $minPts$. Jika titik q tidak memenuhi syarat $|N_\epsilon(\vec{q})| \geq minPts$ dan $q \notin C$, maka tetapkan q sebagai *noise*.

2.4. Evaluasi Hasil Kluster

Pengujian hasil kluster digunakan untuk memastikan bahwa hasil pengelompokan yang diperoleh mencerminkan kejadian yang sebenarnya. Evaluasi hasil kluster dilakukan dengan metode *silhouette coefficient* [14] yang dirumuskan pada Persamaan 3.

$$s(p) = \frac{b(p) - a(p)}{\max \{a(p), b(p)\}} \quad (3)$$

Variabel (SC) merupakan nilai SC masing – masing titik dengan (a) merupakan nilai rata – rata jarak antara suatu titik ke-p dengan titik lain di klaster yang sama dan (b) merupakan nilai rata – rata jarak antara suatu titik ke-p dengan titik lain di klaster yang berbeda. Nilai SC berkisar dari rentang -1 hingga 1, dengan keterangan semakin besar nilai silhouette coefficient maka titik tersebut sudah berada pada klaster yang tepat. Nilai 0 menunjukkan bahwa titik data berada pada atau sangat dekat dengan batas keputusan antara dua klaster yang berdekatan [17].

Dalam penelitian ini digunakan kombinasi parameter epsilon dan minpts seperti yang dilakukan pada penelitian-penelitian sebelumnya. Dari penelitian Nurhaliza et al (2021) digunakan rentang nilai Epsilon 0,1 dan 0,2 pada nilai MinPts 3 dan 4. Sedangkan penelitian oleh Utami et al (2023) setelah dilakukan beberapa percobaan diputuskan menggunakan nilai Epsilon 10 dan MinPts 5.

3. HASIL DAN PEMBAHASAN

3.1. Hasil-hasil Penelitian

Pada tahap “Pengumpulan Data Titik Panas” melalui web FIRM NASA, diperoleh informasi bahwa data titik panas bersifat kuantitatif. Rentang data yang diambil adalah data pada bulan Januari sampai Desember 2023 dengan total sebanyak 56614 titik panas. Akan tetapi, dalam penelitian ini hanya digunakan data titik panas dengan fitur *confidence* diatas 80% mengikuti penelitian Utami dkk. (2023), sehingga ditemukan data yang memenuhi sebanyak 12.708 titik panas. Variabel pada penelitian ini yaitu *longitude* (X1), *latitude* (X2), *confidence* (X3) *bright_t31* (X4), *brightness* (X5), *frp* (X6), *scan* (X7), dan *track* (X8). Semua variabel digunakan dalam pengelompokan kecuali *confidence*, seperti terlihat pada Tabel 1.

Tabel 1. Data Titik Panas

longitude (X1)	Latitude (X2)	confidence (X3)	bright_t31 (X4)	brightness (X5)	frp (X6)	scan (X7)	track (X8)
121,7155	-1,3757	84	287,2	337,1	35,6	1,1	1
120,3451	-2,5051	83	287,1	326,8	20,7	1	1
121,7294	-2,7547	93	296,3	344,8	50,1	1,1	1
121,7032	-2,9427	83	295,8	328,8	23,1	1	1
127,634	1,5005	100	289,5	325,7	24,7	1	1
127,6299	1,494	83	292,6	333,9	36,3	1,1	1
127,6395	1,4926	85	292,3	338	43,6	1,1	1
112,925	-8,1182	94	269,3	317	25,1	1,2	1,1
...	...	...	...	...	...	...	...
138,373	-7,7894	91	291,5	315,5	32,6	2	1,4
120,7027	-10,07	87	291,1	313,6	12,1	1,1	1,1
122,4212	-3,9057	92	294	315,8	11,7	1	1
122,4007	-3,8714	80	294,8	310,5	7,4	1	1

Setelah data diperoleh, tahap selanjutnya dilakukan pemeriksaan untuk mendeteksi data yang hilang dan melakukan penghapusan pada data duplikat. Hasil setelah dilakukan pemeriksaan yakni data tetap berjumlah 12.708 titik panas atau tidak ada data yang hilang maupun duplikat. Kemudian dilakukan transformasi data berupa standarisasi dengan metode *z-score* pada semua variabel yang digunakan untuk pengelompokan, seperti terlihat hasilnya di Tabel 2. Variabel Z1 sampai Z8 secara berturut-turut kecuali Z3, masing-masing adalah variabel untuk menyimpan hasil transformasi data *longitude*, *latitude*, *confidence*, *bright_t31*, *brightness*, *frp*, *scan*, dan *track* dari data di Tabel 1 yang sesuai.

Tabel 2. Data Titik Panas Setelah Standarisasi

Z1	Z2	Z4	Z5	Z6	Z7	Z8
0,726263	0,59915	0,154542	-0,26104	-0,46113	-0,71751	0,726263
0,584959	0,235094	-0,48479	-0,48065	-0,6469	-0,71751	0,584959
0,727696	0,154636	0,632491	-0,04732	-0,46113	-0,71751	0,727696
0,724995	0,094035	-0,36065	-0,44528	-0,6469	-0,71751	0,724995
1,336528	1,526278	-0,55307	-0,42169	-0,6469	-0,71751	1,336528
1,336105	1,524183	-0,04409	-0,25072	-0,46113	-0,71751	1,336105
1,337095	1,523732	0,210406	-0,14312	-0,46113	-0,71751	1,337095
-0,18014	-1,57426	-1,09309	-0,4158	-0,27535	-0,16102	-0,18014
...	...	...	...	...	...	...
2,443842	-1,46827	-1,1862	-0,30525	1,210823	1,508465	2,443842
0,621832	-2,20341	-1,30413	-0,60741	-0,46113	-0,16102	0,621832
0,799029	-0,21638	-1,16758	-0,6133	-0,6469	-0,71751	0,799029
0,796915	-0,20533	-1,49656	-0,67668	-0,6469	-0,71751	0,796915

Setelah data siap diproses, maka tahap berikutnya adalah pengelompokan data titik panas dengan melakukan percobaan pada kombinasi nilai Eps dan MinPts guna menemukan kluster yang optimal. Berdasarkan penelitian oleh Nurhaliza[7], kluster paling optimal yang ditentukan melalui nilai SC yaitu kombinasi nilai Eps 0,2 dan MinPts 3 dan penelitian oleh Utami, dkk.[5] menggunakan nilai Eps 10 dan Minpts 5, karena kluster yang terbentuk memiliki nilai noise yang sedikit. Maka pada penelitian ini dilakukan percobaan pengelompokan dengan kombinasi nilai Eps dan MinPts yang sama yaitu kombinasi nilai Eps 0,2 dan MinPts 3, kemudian kombinasi Eps 10 dan Minpts 5, serta kombinasi Eps pada rentang 2 hingga 5 dengan MinPts pada rentang 2 hingga 6, seperti terlihat hasilnya pada Tabel 3.

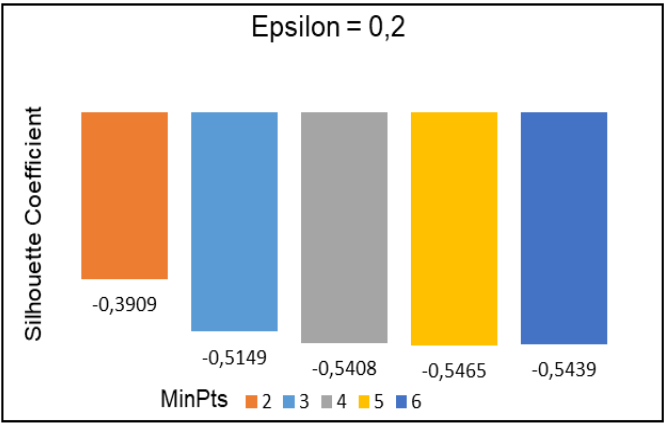
Tabel 3. Hasil Pengelompokan Data Titik Panas

Eps	MinPts	Jumlah Kluster	Jumlah Noise
0.2	2	1267	6213
0.2	3	439	7869
0.2	4	219	8906
0.2	5	142	9566
0.2	6	80	10164
2	2	3	20
2	3	2	22
2	4	2	26
2	5	2	26
2	6	2	32
3	2	3	5
3	3	2	7
3	4	3	7
3	5	2	15
3	6	1	20

4	2	1	3
4	3	1	3
4	4	2	3
4	5	2	5
4	6	2	6
5	2	1	2
5	3	1	2
5	4	1	2
5	5	1	3
5	6	1	3
10	4	1	0
10	5	1	0

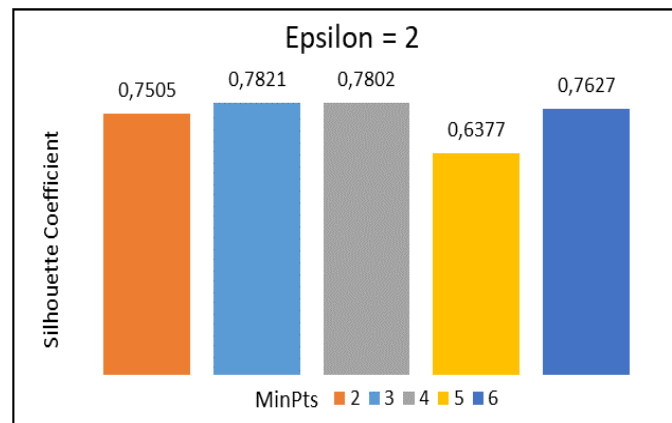
3.2. Diskusi

Uji evaluasi hasil kluster dilakukan setelah tahap pengelompokkan selesai. Pengujian ini dilakukan dengan menghitung nilai *silhouette coefficient* (SC) pada masing – masing percobaan kombinasi nilai Eps dan MinPts untuk menentukan kluster yang optimal. Hasil evaluasi kluster ditampilkan pada Gambar 2 hingga Gambar 6.



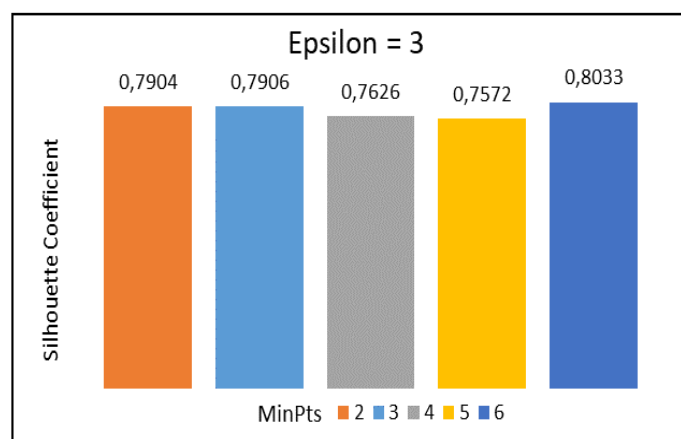
Gambar 2. Hasil Evaluasi Kluster Pada Eps 0,2

Pada Gambar 2 dapat terlihat bahwa semua nilai SC pada Eps 0,2 dan MinPts dari 2 hingga 6 menunjukan nilai yang berada dibawah 0 dan cenderung mendekati angka -1. Kluster yang terbentuk pada kombinasi ini semua bernilai negatif yang artinya titik belum dikelompokkan pada kluster yang tepat.



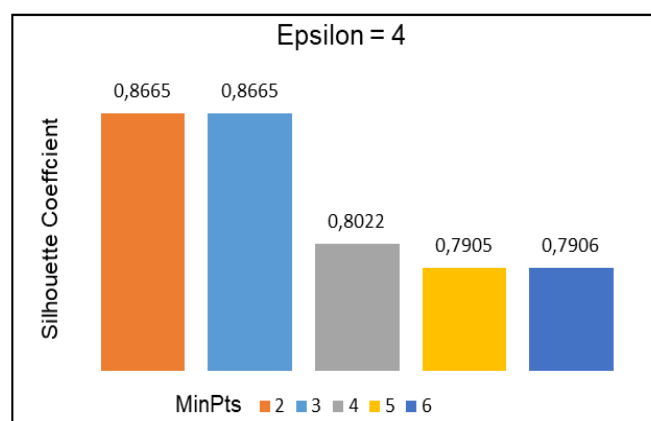
Gambar 3. Hasil Evaluasi Kluster Pada Eps 2

Pada Gambar 3 dapat terlihat bahwa nilai SC pada pengelompokan dengan Eps 2 mengalami kenaikan hingga pada MinPts 4, lalu mengalami penurunan pada MinPts 5, dan kembali naik pada MinPts 6. Nilai SC tertinggi terletak pada kombinasi Eps 2 dan MinPts 3 dengan jumlah kluster sebanyak 2 dan noise sebanyak 22.



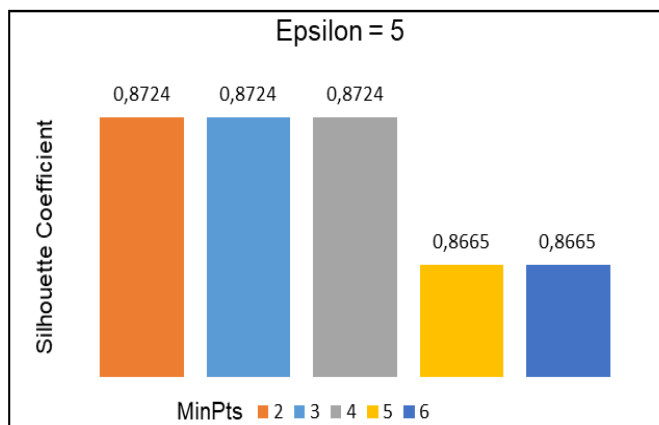
Gambar 4. Hasil Evaluasi Kluster Pada Eps 3

Pada Gambar 4 dapat terlihat pola yang belum stabil pada nilai SC dengan Eps 3. Nilai SC mengalami kenaikan hingga pada MinPts 3, kemudian terjadi penurunan dari MinPts 4 hingga MinPts 5, dan kembali naik pada MinPts 6. Nilai SC tertinggi terjadi pada kombinasi Eps 3 dan MinPts 6 dengan jumlah kluster sebanyak 1 dan noise sebanyak 20.



Gambar 5. Hasil Evaluasi Kluster Pada Eps 4

Pada Gambar 5 dapat terlihat bahwa nilai SC pada Eps 4 stabil pada MinPts 2 hingga MinPts 3 kemudian mengalami penurunan secara signifikan hingga pada MinPts 5 dan mengalami kenaikan pada MinPts 6. Nilai SC tertinggi terletak pada MinPts 2 dan 3 dengan jumlah kluster yang sama yaitu 1 dan noise 3.



Gambar 6. Hasil Evaluasi Kluster Pada Eps 5

Pada Gambar 6 dapat terlihat bahwa nilai SC stabil pada MinPts 2 hingga 4 dengan nilai sebesar 0,8724 dan mengalami penurunan yang stabil hingga pada MinPts 6 dengan nilai sebesar 0,8665. Nilai SC tertinggi memiliki jumlah kluster yang sama yaitu 1 dan noise sebanyak 2.

Dari pencermatan nilai SC pada Gambar 3 sampai Gambar 6, semua percobaan yang dilakukan menemukan hasil pengelompokan yang berada pada struktur kuat yakni dengan nilai SC terkecil adalah 0,6377 pada Eps 2 dan MinPts 5 dan nilai SC tertinggi adalah 0,8724 pada Eps 5 dan MinPts 2 hingga 4. Hal ini sejalan dengan penelitian - penelitian sebelumnya, bahwa percobaan beberapa kombinasi parameter Eps dan MinPts mampu digunakan sebagai penentu kluster yang optimal berdasarkan nilai SC. Namun dengan memperhatikan jumlah kluster yang terbentuk pada penelitian ini akan menjadi sangat padat jika hanya menjadi 1 kluster besar saja. Peneliti mencoba mencermati hasil pengelompokan dengan nilai SC 0,8022 pada parameter Eps 4 dan MinPts 4 yang membentuk 2 kluster serta mengandung noise sebanyak 3 titik. Diperoleh hasil pengelompokan seperti di Tabel 4.

Tabel 4. Data Anggota Setiap Kluster Hasil Pengelompokan

longitude	latitude	bright_t31	brightness	Frp	scan	track	kluster
121,7155	-1,3757	287,2	337,1	35,6	1,1	1	0
120,3451	-2,5051	287,1	326,8	20,7	1	1	0
121,7294	-2,7547	296,3	344,8	50,1	1,1	1	0
121,7032	-2,9427	295,8	328,8	23,1	1	1	0
127,634	1,5005	289,5	325,7	24,7	1	1	0
127,6299	1,494	292,6	333,9	36,3	1,1	1	0
127,6395	1,4926	292,3	338	43,6	1,1	1	0
112,925	-8,1182	269,3	317	25,1	1,2	1,1	0
...	...	...	...	...	...	...	...
138,373	-7,7894	291,5	315,5	32,6	2	1,4	0
120,7027	-10,07	291,1	313,6	12,1	1,1	1,1	0
122,4212	-3,9057	294	315,8	11,7	1	1	0
122,4007	-3,8714	294,8	310,5	7,4	1	1	0
102,6471	-0,5322	312,6	457,2	1156,6	1,3	1,1	1
110,1333	-1,7514	313,7	469,3	991	1	1	1
105,1805	-3,872	314,6	435,3	894	1,4	1,2	1
112,7363	-3,3141	306,4	418,7	1207,6	2,2	1,4	1
112,7316	-3,3158	309,4	426,2	1412,4	2,2	1,4	1

111,8427	-3,3775	320,6	463,2	890,6	1	1	1
112,8551	-3,2292	313,9	465,5	967,7	1	1	1
116,4928	-8,4006	319,5	429,4	909,1	1,6	1,2	1
108,9916	0,5641	295	432,4	1804,9	2,4	1,5	-1
110,1422	-1,7502	323,8	500	1603	1	1	-1
105,2266	-3,379	297,1	401,2	1516,4	3,4	1,7	-1

Pada Tabel 4 merupakan hasil pengelompokan data titik panas di Indonesia selama tahun 2023 dengan *confidence* 80% ke atas menggunakan algoritma DBSCAN. Pada klaster 0 yang berisi 12.697 titik panas dengan intensitas suhu dan energi yang dipancarkan oleh api mengindikasikan aktivitas kebakaran yang rendah hingga sedang namun terjadi pada lokasi yang tersebar secara meluas. Kepadatan yang tinggi pada klaster 0 dapat diartikan bahwa pada lokasi tersebut intensitas peristiwa kebakaran cukup sering terjadi. Klaster 1 memiliki titik panas sebanyak 8 sehingga dapat dikatakan sebagai klaster dengan kepadatan rendah. Hal ini menandakan bahwa hanya ada sedikit lokasi kejadian kebakaran. Artinya jarang terjadi kebakaran pada area tersebut. Pada klaster 1 titik panas dengan intensitas suhu dan energi yang terpancar oleh api menggambarkan aktivitas kebakaran yang tinggi. Algoritma DBSCAN dapat dikatakan berhasil karena sesuai dengan fakta data di lapangan yaitu data dari Pantau Gambut tentang analisis Kerentanan Kebakaran Hutan dan Lahan (Karhutla) tahun 2023 yang menyatakan bahwa Provinsi Kalimantan Tengah, Riau, Kalimantan Timur sebagai provinsi dengan kerentanan kebakaran hutan dan lahan kelas tinggi [18] yang selaras dengan karakteristik dan lokasi koordinat titik – titik pada klaster 1.

4. KESIMPULAN

Density-based spatial clustering of application with noise (DBSCAN) merupakan salah satu algoritma pengelompokan yang berfokus pada kepadatan dan memiliki kelebihan yaitu mampu mendeteksi titik kebisingan. Keberhasilan DBSCAN terbukti cukup efektif untuk mengelompokkan titik api. Percobaan dengan parameter terbaik yang ditemukan yaitu Eps 4 dan MinPts 4 menghasilkan nilai SC 0,8022 dengan 3 klaster yaitu cluster 0 yang merupakan wilayah sering terjadi kebakaran, klaster 1 dengan wilayah kebakaran lebih tinggi sehingga harus ditangani segera, dan klaster -1 yang ditandai sebagai noise. Keterbatasan pada penelitian ini yaitu tidak dipertimbangkan dari segi faktor lingkungan. Variabel seperti kelembapan udara, curah hujan, kondisi cuaca dapat berperan dalam pola persebaran titik panas. Rekomendasi untuk penelitian selanjutnya bisa dilakukan eksplorasi lebih lanjut seperti kombinasi DBSCAN dengan optimasi berbasis AI.

DAFTAR PUSTAKA

- [1] Trianawati, Arini, Sari, and Agustina, “Analisis Penerapan Sistem Proteksi Kebakaran Aktif Berdasarkan Peraturan Menteri Kesehatan Republik Indonesia Nomor 52 Tahun 2018 Dan SNI 180-2:2022,” *Dohara Publisher*, vol. 02, no. 09, pp. 825–829, 2023, [Online]. Available: <http://dohara.or.id/index.php/hsk%7C>
- [2] I. P. Anhar, Rina Mardiana, and Rai Sita, “Dampak Kebakaran Hutan dan Lahan Gambut terhadap Manusia dan Lingkungan Hidup (Studi Kasus: Desa Bunsur, Kecamatan Sungai Apit, Kabupaten Siak, Provinsi Riau),” *Jurnal Sains Komunikasi dan Pengembangan Masyarakat [JSKPM]*, vol. 6, no. 1, pp. 75–85, 2022, doi: 10.29244/jskpm.v6i1.967.
- [3] J. Kim, A. Wise, and E. Bowman, “Strong winds pick up, increasing fire danger as firefighters battle LA blazes,” 2025. <https://www.npr.org/2025/01/11/g-s1-42247/la-wildfires-california-containment-evacuation> (accessed Jan. 15, 2025).
- [4] M. R. B. E. Prasetyo, “BNPB: Ada 629 kasus karhutla terjadi di Indonesia sepanjang 2024,” Indonesia, 2025. [Online]. Available: <https://www.antaranews.com/berita/4574894/bnpb-ada-629-kasus-karhutla-terjadi-di-indonesia-sepanjang-2024>
- [5] “Rekapitulasi Luas Kebakaran Hutan dan Lahan(Ha) per Provinsi di Indonesia,” *Sistem Pemantauan Karhutla*, 2024. <https://sipongi.menlhk.go.id/indikasi-luas-kebakaran> (accessed Feb. 11, 2025).
- [6] W. Li, E. Dodwell, and D. Cook, “A Clustering Algorithm to Organize Satellite Hotspot Data for the Purpose of Tracking Bushfires Remotely,” *R Journal*, vol. 15, no. 1, pp. 17–33, 2023, doi: 10.32614/RJ-2023-022.
- [7] P. Y. Utami, S. A. Hudjimartsu, T. A. Viona, and H. Sharfina, “Optimasi Parameter Algoritma DBSCAN untuk Mendeteksi Titik Panas Kebakaran Hutan dan Lahan,” *Jurnal Edukasi dan Penelitian Informatika (JEPIN)*, vol. 9, no. 3, p. 355, Dec. 2023, doi: 10.26418/jp.v9i3.61714.

-
- [8] A. Fauzan, A. Novianti, R. R. M. A. Ramadhani, and M. A. S. Adhiwibawa, "Analysis of Hotels Spatial Clustering in Bali: Density-Based Spatial Clustering of Application Noise (DBSCAN) Algorithm Approach," *EKSAKTA: Journal of Sciences and Data Analysis*, vol. 3, no. 1, pp. 25–38, Mar. 2022, doi: 10.20885/EKSAKTA.vol3.iss1.art4.
- [9] N. Nurhaliza and M. Mustakim, "Pengelompokan Data Kasus Covid-19 di Dunia Menggunakan Algoritma DBSCAN," *Indonesian Journal of Informatic Research and Software Engineering (IJIRSE)*, vol. 1, no. 1, pp. 1–8, 2021, doi: 10.57152/ijirse.v1i1.42.
- [10] S. Putro, M. Syarief, and E. Rochman, "PENERAPAN ALGORITMA K-MEANS, DBSCAN, DAN AHC UNTUK CLUSTERING KUALITAS GARAM PADA PT. GARAM (PERSERO)," *Multitek Indonesia: Jurnal Ilmiah*, vol. 18, no. 2, pp. 114–125, 2024, doi: 10.24269/mtkind.v18i2.8501.
- [11] N. Purnomo and W. Gata, "Pengelompokan Analisis Sentimen Komentar Youtube Terhadap Pengambilalihan Jalan Rusak di Lampung Menggunakan Algoritma Clustering," *Progresif: Jurnal Ilmiah Komputer*, vol. 20, no. 2, p. 701, Aug. 2024, doi: 10.35889/progresif.v20i2.1936.
- [12] A. Kristianto, E. Sedyono, and K. D. Hartomo, "Implementation dbscan algorithm to clustering satellite surface temperature data in indonesia," *Register: Jurnal Ilmiah Teknologi Sistem Informasi*, vol. 6, no. 2, p. 109, Jul. 2020, doi: 10.26594/register.v6i2.1913.
- [13] E. Lidrawati, S. Bahri, U. F. Zubaedi, V. P. Carolina, K. Kusriani, and D. Maulina, "Kebakaran Hutan Implementasi Metode CLARA Clustering Untuk Pengelompokan Data Potensi Kebakaran Hutan/Lahan Berdasarkan Persebaran Titik Panas (Hotspot)," *Journal of Computer System and Informatics (JoSYC)*, vol. 3, no. 4, pp. 507–511, 2022, doi: 10.47065/josyc.v3i4.2006.
- [14] J. Han, J. Pei, and H. Tong, *DATA MINING CONCEPTS AND TECHNIQUES*, Fourth. United States of America: Katey Birtcher, 2023.
- [15] R. R. A. Rahman and A. W. Wijayanto, "Pengelompokan Data Gempa Bumi Menggunakan Algoritma Dbscan," *Jurnal Meteorologi dan Geofisika*, vol. 22, no. 1, p. 31, 2021, doi: 10.31172/jmg.v22i1.738.
- [16] A. Kassambara, *Practical Guide To Cluster Analysis in R (Unsupervised Machine Learning)*, 1st ed. Statistical Tools for High-throughput Data Analysis, 2017.
- [17] O. O. Wijaya and Rushendra, "Analysis of Sulawesi Earthquake Data from 2019 to 2023 using DBSCAN Clustering," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 8, no. 4, pp. 454–465, Aug. 2024, doi: 10.29207/resti.v8i4.5819.
- [18] A. Ramadhi *et al.*, "Kerentanan Kebakaran Hutan dan Lahan (Karhutla) pada Area Kesatuan Hidrologis Gambut (KHG) Tahun," p. 2, 2023.