MASE Media Applied Science and Engineering

Media Journal of Information System and Informatic, **2025, 2, 26-35**
Published Online June 2025 in MASE (http://www.mase.or.id)
DOI: XX.XXXX/mistic.XX.XX.XX

Hotspot Clustering in Bangka Belitung Islands Province Using Agglomerative Hierarchical Clustering Algorithm

Aji Setiawan
Universitas Jambi, Fakultas Sains dan Teknologi, Jambi, Indonesia
E-mail: ajiisetiawann468@gmail.com

Akhiyar Waladi
Universitas Jambi, Fakultas Sains dan Teknologi, Jambi, Indonesia
E-mail: akhiyar.waladi@unja.ac.id
*Corresponding Author

Rahmad Ashar
Universitas Jambi, Fakultas Sains dan Teknologi, Jambi, Indonesia
E-mail: rahmad.ashar@unja.ac.id

Received: March, 2025; Accepted: June, 2025; Published: June, 2025

Abstract: Forest and land fires are common disasters in Indonesia, especially in areas with vast forests and lands, such as the Province of Bangka Belitung Islands. Hotspot data, indicating high-temperature areas, can be used to predict the risk of fires. However, not all hotspots indicate actual fires, making it essential to cluster and analyze hotspot data to identify areas with a high likelihood of fires. This study aims to cluster hotspot data in the Bangka Belitung Islands using the Agglomerative Hierarchical Clustering (AHC) algorithm. The dataset, obtained from NASA's MODIS satellite for the year 2019, consists of 843 records with six attributes, including Brightness, Confidence, and Fire Radiative Power (FRP). Data preprocessing involves cleaning and normalizing the dataset. The AHC algorithm was applied to group hotspots based on their similarity, and Euclidean distance was used to calculate proximity between points. The resulting clusters were evaluated using the Silhouette Coefficient to determine the optimal number of clusters. The analysis revealed that two clusters provided the best result with a Silhouette Coefficient of 0.653. Cluster 1 contained 829 hotspots with an average confidence level of 61.8% (medium risk category), while Cluster 2 comprised 14 hotspots with an average confidence level of 94.1% (high risk category requiring immediate intervention). These findings enable authorities to prioritize fire prevention resources effectively, particularly focusing on the 14 high-risk locations identified in Cluster 2.

Keywords: Hotspot, Agglomerative Hierarchical Clustering, Forest Fire, Silhouette Coefficient, Data Mining, Bangka Belitung

1. Pendahuluan

Kebakaran hutan dan lahan merupakan permasalahan lingkungan yang kompleks dan terus berulang di berbagai wilayah di Indonesia. Wilayah yang paling terdampak umumnya adalah daerah dengan vegetasi yang lebat dan lahan gambut yang luas, yang mudah terbakar terutama pada musim kemarau panjang. Provinsi Kepulauan Bangka Belitung, meskipun secara geografis terdiri dari pulau-pulau, tidak luput dari ancaman kebakaran hutan dan lahan. Data dari Kementerian Lingkungan Hidup dan Kehutanan menunjukkan bahwa pada tahun 2019, lebih dari 4.700 hektare lahan di provinsi ini mengalami kebakaran [1]. Dampak dari kejadian ini tidak hanya berupa kerusakan ekologis seperti hilangnya tutupan hutan, terganggunya habitat satwa liar, dan penurunan kualitas tanah, tetapi juga menyebabkan gangguan kesehatan masyarakat, kerugian ekonomi, serta peningkatan emisi karbon yang berdampak pada perubahan iklim global.

Kebakaran hutan juga seringkali sulit dideteksi secara dini karena luasnya area yang perlu diawasi, keterbatasan sumber daya manusia di lapangan, serta kondisi geografis yang menyulitkan akses ke daerah rawan. Oleh karena itu, diperlukan pendekatan berbasis teknologi untuk memantau dan menganalisis potensi kebakaran secara lebih efisien dan akurat. Salah satu teknologi yang banyak dimanfaatkan adalah penginderaan jauh (remote sensing) yang memanfaatkan data satelit untuk memantau kondisi permukaan bumi secara berkala. Salah satu indikator penting yang dihasilkan dari citra satelit adalah keberadaan titik panas atau *hotspot*. Hotspot merupakan indikasi awal dari adanya anomali suhu di suatu lokasi yang dapat mengarah pada terjadinya kebakaran hutan. Namun demikian, tidak semua hotspot menunjukkan kebakaran yang sebenarnya, karena sinyal panas juga bisa berasal dari aktivitas manusia atau fenomena alam lainnya [2].

Untuk meningkatkan akurasi deteksi dan mengoptimalkan respons, diperlukan analisis lebih lanjut terhadap data hotspot, sehingga dapat dibedakan antara titik-titik yang berisiko tinggi dengan yang berisiko rendah. Salah satu pendekatan yang dapat digunakan adalah teknik *data mining*, khususnya metode *clustering*, yang memungkinkan pengelompokan data berdasarkan kesamaan karakteristiknya. Clustering telah banyak digunakan dalam berbagai aplikasi, termasuk dalam pemetaan risiko bencana, analisis spasial, dan perencanaan tata guna lahan. Metode ini sangat efektif untuk mengenali pola tersembunyi dalam kumpulan data besar, yang sulit dianalisis secara manual [3].

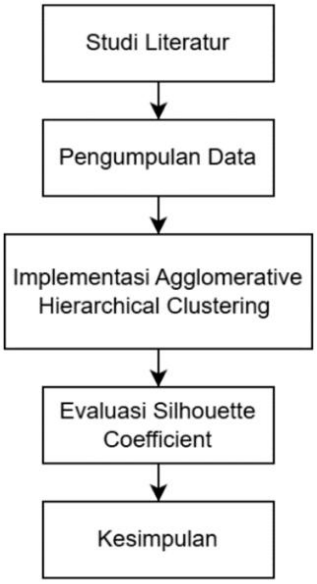
Penelitian-penelitian sebelumnya telah menerapkan berbagai metode clustering seperti K-Means, K-Medoids, dan Self-Organizing Maps (SOM) dalam mengelompokkan data hotspot untuk mengidentifikasi wilayah prioritas dalam penanganan kebakaran. Sebagai contoh, studi oleh Pramesti et al. menggunakan K-Medoids dan menghasilkan dua cluster dengan nilai silhouette coefficient sebesar 0,56745 [4], sedangkan penelitian oleh Sinaga et al. yang menggunakan metode SOM hanya mencapai nilai koefisien maksimum sebesar 0,2489 [5]. Meskipun metode-metode tersebut cukup berhasil dalam memberikan pemetaan awal, pendekatan partisi seperti ini memiliki keterbatasan dalam fleksibilitas, khususnya dalam menentukan jumlah cluster secara optimal sejak awal.

Dalam konteks tersebut, metode *Agglomerative Hierarchical Clustering* (AHC) menawarkan alternatif yang menarik. AHC merupakan salah satu teknik *hierarchical clustering* yang bersifat bottom-up, di mana setiap titik data awalnya dianggap sebagai cluster tersendiri, kemudian secara bertahap digabungkan berdasarkan tingkat kemiripan hingga terbentuk struktur pohon hierarki (*dendogram*) [6]. Keunggulan utama dari pendekatan ini adalah kemampuannya untuk menggambarkan struktur hubungan antar data secara menyeluruh, tanpa perlu menentukan jumlah cluster terlebih dahulu. Hal ini sangat sesuai untuk analisis eksploratori pada data hotspot, yang sering kali memiliki struktur spasial dan variasi risiko yang kompleks.

Penelitian ini bertujuan untuk menerapkan algoritma AHC pada data hotspot di Provinsi Kepulauan Bangka Belitung, yang diperoleh dari citra satelit MODIS tahun 2019. Dataset ini terdiri dari 843 titik panas dengan berbagai atribut, seperti brightness, confidence, dan Fire Radiative Power (FRP), yang menggambarkan intensitas dan kemungkinan terjadinya kebakaran. Dengan melakukan klasterisasi terhadap data ini, penelitian ini bertujuan mengidentifikasi zona dengan risiko kebakaran tinggi dan sedang, sehingga dapat membantu pihak berwenang dalam merumuskan strategi mitigasi dan distribusi sumber daya secara lebih efektif. Selain itu, penelitian ini juga diharapkan dapat memperkaya literatur mengenai penerapan metode *hierarchical clustering* dalam konteks manajemen bencana berbasis data, khususnya kebakaran hutan dan lahan di Indonesia.

2. Bahan dan Metode

Metodologi penelitian menggambarkan proses penelitian secara umum yang dilakukan untuk memperoleh hasil akhir berupa klasterisasi titik api di Provinsi Kepulauan Bangka Belitung. Setiap tahap dalam metodologi ini dirancang secara sistematis, mulai dari pengumpulan data hingga evaluasi hasil klaster, guna memastikan bahwa proses analisis berjalan secara ilmiah dan terstruktur. Tahapan lengkap dari penelitian ini ditunjukkan pada Gambar 1.



Gambar 1. Metodologi Penelitian

2.1. Studi Literatur

Penelitian literatur adalah metode pengumpulan informasi dan data menggunakan jenis literatur berbeda yang ditemukan di perpustakaan, termasuk dokumen, buku, dan jurnal penelitian. Dalam penelitian ini, peneliti mengumpulkan

informasi teoritis terkait penelitian yang akan dilakukan melalui buku, jurnal penelitian sebelumnya, dan artikel di Internet. Hal ini akan membantu mengetahui metode dan algoritma mana yang cocok untuk menyelesaikan masalah yang diteliti dan juga akan menjadi dasar referensi yang kuat untuk menyelesaikan penelitian ini.

2.2. Pengumpulan Data

Pengumpulan data merupakan tahap penting dalam proses penelitian karena kualitas data yang digunakan sangat memengaruhi validitas hasil akhir. Data yang digunakan dalam penelitian ini merupakan data titik panas (*hotspot*) di Provinsi Kepulauan Bangka Belitung, yang diperoleh dari situs resmi NASA dan dapat diakses secara terbuka. Hotspot merujuk pada area yang suhunya lebih tinggi dibandingkan wilayah sekitarnya, dan terdeteksi melalui pengamatan satelit berbasis sensor termal [7]. Setiap titik dicatat dengan koordinat geografis yang menunjukkan lokasi spesifik dari anomali suhu tersebut.

Data yang digunakan adalah rekaman titik api dari bulan Januari hingga Desember 2019, dengan total 843 baris data dan 15 atribut. Namun, hanya 6 atribut yang dianggap relevan untuk proses analisis, yaitu Longitude, Latitude, Brightness, Confidence, Bright_T31, dan FRP. Dari atribut-atribut tersebut, tiga atribut utama—Brightness, Confidence, dan FRP—digunakan dalam proses algoritma *Agglomerative Hierarchical Clustering* karena ketiganya mewakili karakteristik termal dan intensitas api yang paling menentukan.

Tingkat *confidence* pada setiap hotspot menunjukkan seberapa besar kemungkinan bahwa titik tersebut benar-benar merupakan area kebakaran aktif [8]. Oleh karena itu, atribut ini sangat penting untuk membedakan antara hotspot biasa dengan titik yang memiliki risiko kebakaran tinggi. Berdasarkan pedoman MODIS Active Fire Product dari LAPAN, tingkat *confidence* dibagi ke dalam tiga kategori, yaitu rendah (0–30%), sedang (30–80%), dan tinggi (80–100%), masing-masing dengan rekomendasi tindakan berbeda. Klasifikasi lengkap tingkat *confidence* ini dapat dilihat pada **Tabel 1.**

.

Tabel 1. Nilai Tingkat Confidence

Confindence	Kelas	Tindakan
$0\% \leq C \leq 30\%$	Rendah	Perlu diperhatikan
$30\% \leq C \leq 80\%$	Sedang	Waspada
$80\% \leq C \leq 100\%$	Tinggi	Segera Penanggulangan

2.3. Implementasi Agglomerative Hierarchical Clustering

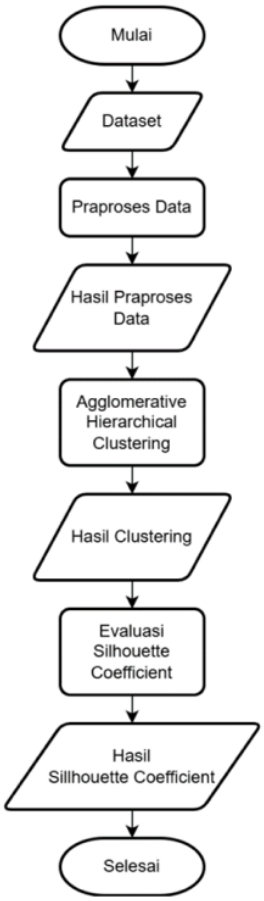
Clustering adalah salah satu metode dalam Data Mining yang bertujuan untuk mengelompokkan objek-objek berdasarkan kemiripan karakteristiknya, sehingga terbentuk kelompok yang homogen di dalamnya namun berbeda dengan kelompok lainnya [9]. Hierarchical Clustering merupakan teknik clustering yang menyusun objek dalam bentuk hirarki, menyerupai struktur pohon. Clustering dilakukan secara bertahap. Algoritma ini memiliki dua pendekatan, yaitu Agglomerative (bottom-up) dan Divisive (top-down). Pada tahap implementasi, dataset yang telah dikumpulkan diproses menggunakan algoritma Agglomerative Hierarchical Clustering dengan bantuan software Notebook pada Google Colab, dengan langkah-langkah berikut:

2.3.1. Praproses Data

Praproses atau Praproses Data adalah langkah awal dalam data mining yang bertujuan untuk mengatasi masalah-masalah yang bisa mempengaruhi hasil dari proses data. Data yang tersedia biasanya masih memiliki masalah seperti nilai yang hilang (missing value), duplikasi, dan inkonsistensi. Dengan melakukan praproses data, data tersebut dibersihkan sehingga yang digunakan adalah data berkualitas tinggi, yang akan menghasilkan pola atau pengetahuan baru yang lebih baik dari proses mining. Salah satu tahap dalam praproses data adalah normalisasi, yang merupakan proses penskalaan ulang nilai atribut data agar berada dalam rentang atau skala tertentu. Normalisasi sering dilakukan dengan metode min-max, yang merupakan metode transformasi linier pada data asli dengan persamaan yaitu [10]:

$$X' = \frac{x - min(x)}{max(x) - min\ (x)} \quad (1)$$

Di mana:
X’ = Data dinormalisasi
min(x) = Nilai minimum dari range data yang akan dinormalisasi
max(x) = Nilai maksimum dari range data yang akan dinormalisasi



Gambar 2. Tahapan Proses Agglomerative Clustering

2.3.2 Perhitungan Jarak Euclidiean

Euclidean Distance merupakan metode yang umum dan sederhana untuk menghitung jarak antara dua titik dalam ruang Euclidean, di mana setiap titik direpresentasikan dalam dimensi yang berbeda [9]. Rumus dari metode ini adalah:

$$D_{Euc}(u,v) = \sqrt{\sum_{i=1}^{n}(u_i, v_i)} \quad (2)$$

Di mana:
u dan v = Dua objek yang akan dihitung jaraknya
ui = Komponen ke I dari u secara berurutan

2.3.3 Proses Agglomerative Hierarchical Clustering

Dalam metode Agglomerative, prosesnya dimulai dengan mengelompokkan setiap objek secara individu, di mana pada awalnya jumlah cluster sama dengan jumlah objek. Kemudian, objek-objek yang memiliki kesamaan atau kedekatan dikelompokkan menjadi cluster baru berdasarkan perhitungan jarak dan parameter kedekatan yang digunakan. Setelah itu, dilakukan perhitungan jarak antar cluster yang baru terbentuk, lalu cluster dengan jarak terdekat digabungkan. Proses ini terus berlanjut hingga semua objek tergabung dalam satu cluster [11].

2.3.4 Evaluasi Sithouette Coefficient

Hasil dari algoritma Hierarchical Clustering yang membentuk beberapa cluster perlu dievaluasi. Evaluasi ini dilakukan untuk menentukan seberapa mirip atau seberapa dekat jarak antara objek-objek dalam satu cluster, serta seberapa jauh cluster tersebut terpisah dari cluster lain [12]. Semakin mirip atau semakin dekat objek-objek dalam sebuah cluster, serta semakin jauh jaraknya dari objek di cluster lain, maka semakin baik kualitas klasterisasi yang dihasilkan.

Salah satu metode yang digunakan untuk mengevaluasi kualitas cluster adalah silhouette coefficient. Silhouette coefficient menggabungkan dua metode, yaitu cohesion dan separation. Metode cohesion mengukur seberapa erat hubungan antara objek dalam satu cluster, sedangkan metode separation mengukur seberapa jauh sebuah cluster terpisah

dari cluster lain. Jumlah cluster yang optimal adalah jumlah yang memiliki rata-rata nilai silhouette coefficient tertinggi atau mendekati 1. Langkah-langkahnya adalah sebagai berikut [13]:

1. Menghitung jarak rata-rata dari suatu data misalkan i dengan semua data lain yang berada dalam klaster yang sama (ai).

$$a(i) = \frac{1}{|A|-1}\sum_{j} \in_{A,j\neq i} d(i,j) \tag{3}$$

Dengan j adalah data lain dalam satu cluster A dan d(i,j) adalah jarak antara data i dan j.

2. Menghitung rata-rata jarak dari data i tersebut dengan semua data di klister lain, lalu ambil nilai terkecilnya.

$$d(i,C) = \frac{1}{|A|-1}\sum_{j} J \in Cd(i,j) \tag{4}$$

Dengan d(i,C) adalah jarak rata-rata data i dengan semua objek pada cluster lain C di mana A ≠ C.

$$b(i) = \min C \neq Ad(i,C) \tag{5}$$

3. Rumus Silhouette Coefficient adalah:

$$s(i) = (b(i) - a(i))/\max(a(i), b(i)) \tag{6}$$

Dengan s(i) adalah semua rata-rata pada semua kumpulan data.

3. Hasil dan Analisis

3.1. Praproses Data

Praproses data adalah langkah awal dalam data mining sebelum melanjutkan ke tahapan berikutnya. Data yang digunakan dalam penelitian ini merupakan data titik api di Provinsi Kepulauan Bangka Belitung, yang dicatat dari 1 Januari 2019 hingga 31 Desember 2019 dan diperoleh dari situs resmi NASA, https://firms.modaps.eosdis.nasa.gov/. Dataset tersebut terdiri dari 843 record dengan 15 atribut. Namun, hanya 6 atribut yang digunakan dalam penelitian ini, yaitu latitude, longitude, brightness, confidence, bright_t31, dan frp. Untuk penerapan algoritma Agglomerative Hierarchical Clustering, hanya atribut brightness, confidence, dan frp yang digunakan.

Dari total 843 data, mungkin terdapat baris yang memiliki nilai atribut yang hilang (missing value), duplikasi data, atau data yang tidak konsisten. Pada penelitian ini, teknik prapemrosesan yang diterapkan adalah pembersihan (cleaning) dan transformasi. Proses cleaning dilakukan dengan menghapus missing value dan duplikasi data, sehingga data tampak seperti ilustrasi berikut:

Tabel 2. Dataset Setelah Cleaning

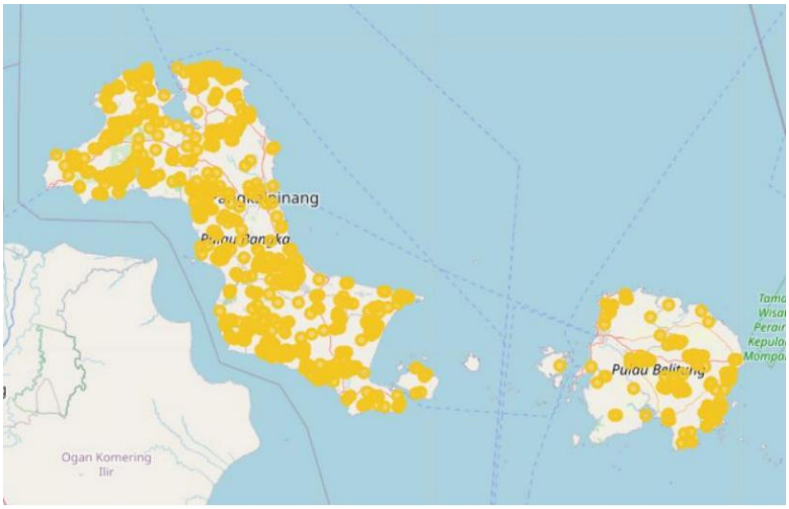
No	Brightness	Confidence	FRP
1	312,3	26	9,5
2	311,9	27	13,9
3	312,4	33	22,9
4	316,9	49	10,3
5	323,9	77	13,2
…	…	…	…
839	312,2	53	5,3
840	313,7	54	7
841	313,5	36	5,5
842	323,8	58	10,1
843	311,1	61	18,6

Setelah melalui proses cleaning, jumlah data tetap sebanyak 843 catatan. Hal ini disebabkan tidak adanya nilai yang hilang atau duplikasi dalam dataset tersebut. Selanjutnya, dilakukan transformasi data dengan teknik normalisasi. Data numerik akan dinormalisasi menggunakan metode min-max, menghasilkan nilai yang berada dalam skala antara 0 hingga 1.

Tabel 3. Dataset Setelah Normalisasi

No	Brightness	Confidence	FRP
1	0,136771	0,26	0,024218
2	0,132287	0,27	0,043593
3	0,137892	0,33	0,083223
4	0,188341	0,49	0,027741
5	0,266816	0,77	0,040511
…	…	…	…
839	0,135650	0,53	0,005724
840	0,152466	0,54	0,013210
841	0,150224	0,36	0,006605
842	0,265695	0,58	0,026860
843	0,123318	0,61	0,064289

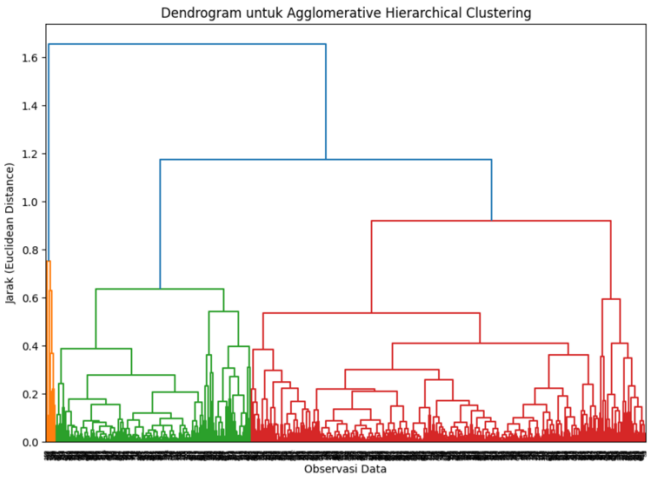
Berikut adalah visualisasi lokasi titik api



Gambar 3. Visualisasi Lokasi Titik Api di Provinsi Kepulauan Bangka Belitung Tahun 2019

3.2. Algoritma AGG

Sebelum memulai proses Clustering, penting untuk menghitung matriks jarak antar data. Jarak antar data dihitung menggunakan metode Euclidean Distance. Proses Agglomerative mengelompokkan data dari N cluster menjadi satu cluster secara bertahap. Dalam hal ini, N merujuk pada jumlah data, yang juga dapat diartikan sebagai jumlah awal cluster pada Agglomerative Hierarchical Clustering, yaitu sebanyak 843 cluster. Hasil dari proses Clustering ini dapat divisualisasikan melalui dendogram yang ditunjukkan dalam Gambar 4.



Gambar 4. Visualisasi Cluster dengan Dendogram

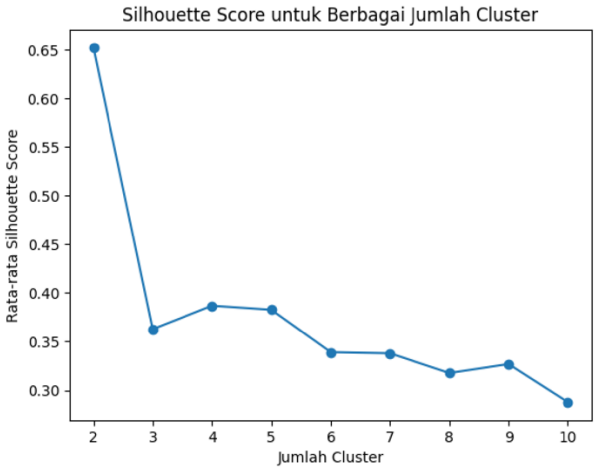
Setelah memotong pohon hierarki menjadi 2 cluster maka didapatkan 2 kelompok data yang telah terpisah dengan Cluster-1 sebanyak 829 data, dan Cluster-2 sebanyak 14 data.

3.3. Evaluasi Hasil Cluster

Evaluasi kluster dilakukan dengan membandingkan nilai rata-rata Silhouette Coefficient untuk setiap jumlah kluster yang berbeda. Pengujian dilakukan dengan jumlah kluster yang bervariasi dari 2 hingga 10. Hasil perhitungan Silhouette Coefficient dapat dilihat dalam Tabel 5.

Tabel 4. Nilai Silhouette Coefficient

Jumlah Cluster	Silhouette Coefficient
2	0.652505292356283
3	0.36272158966916646
4	0.38682498756073325
5	0.3826678922010965
6	0.33906038955396667
7	0.3379106276714015
8	0.31763324672266824
9	0.326830098846762
10	0.2876110972882112



Gambar 5. Diagram nilai Silhouette Coefficient berdasarkan jumlah cluster

Berdasarkan hasil pengujian yang ditampilkan pada Gambar 5, jumlah cluster yang memberikan nilai silhouette coefficient tertinggi adalah 2. Sebaliknya, jumlah cluster yang menghasilkan nilai silhouette coefficient terendah adalah 8. Oleh karena itu, Clustering yang paling sesuai untuk menganalisis potensi kebakaran adalah 2 cluster. Nilai silhouette coefficient yang tinggi menunjukkan bahwa data yang diuji terklaster dengan baik, yang berarti terdapat jarak yang signifikan antar cluster dan jarak yang lebih kecil antar objek dalam cluster yang sama.

3.4. Analisis

Pada koefisien siluet, semakin mendekati nilai 1, semakin baik kualitas kluster yang dihasilkan. Sebaliknya, jika nilainya semakin mendekati 0, kualitas klusternya semakin buruk. Setelah melakukan pengujian kluster dengan koefisien siluet untuk 2 kluster hingga 10 kluster, disimpulkan bahwa 2 kluster merupakan hasil terbaik dalam proses Clustering ini.

Seperti yang telah diketahui, jumlah titik hotspot yang tersebar dalam 2 kluster adalah 829 titik pada kluster 1 dan 14 titik pada kluster 2. Dengan menggunakan atribut longitude dan latitude dari data, lokasi asli titik hotspot dapat ditentukan pada peta, seperti yang ditunjukkan pada Gambar 6.

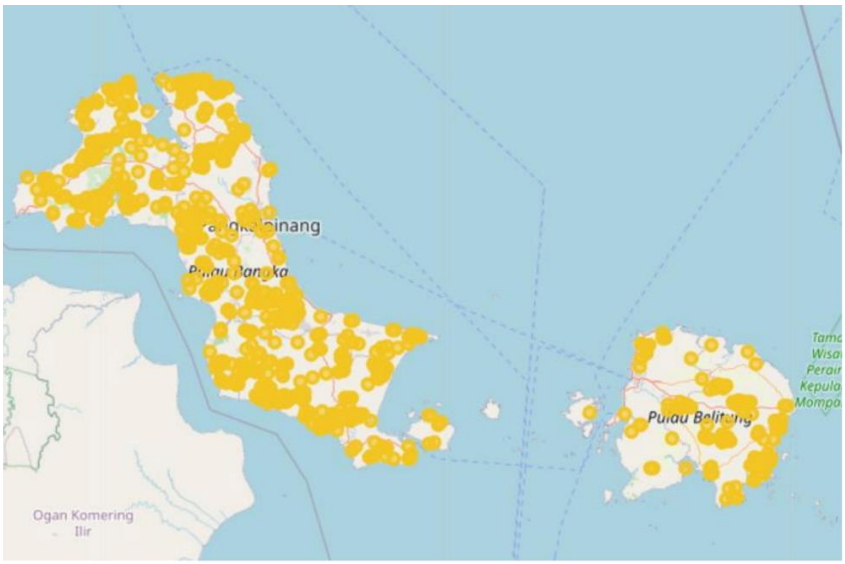
Gambar 6. Lokasi hotspot pada peta

Titik kuning yang terlihat pada gambar 6 menunjukkan hotspot dari cluster 1, yang tersebar di hampir seluruh wilayah provinsi Kepulauan Bangka Belitung. Sementara itu, titik merah mewakili hotspot dari cluster 2. Dengan menghitung rata-rata confidence, brightness, dan FRP untuk masing-masing cluster, diperoleh hasil yang tercantum dalam tabel 5.

Tabel 5. Hasil rata-rata Confidence, Brightness, dan FRP

Cluster	Confidence	Brightness	FRP	Jumlah Data
1	61,796140	320,839083	20,902774	829
2	94,142857	361,407143	156,885714	14

Rata-rata nilai confidence, brightness, dan FRP tertinggi ditemukan pada cluster-2, dengan confidence mencapai lebih dari 94%. Angka ini menunjukkan bahwa titik hotspot di cluster-2 memiliki potensi kebakaran yang lebih tinggi. Confidence di atas 80% termasuk dalam kategori potensi tinggi, sehingga memerlukan penanganan lebih lanjut. Di sisi lain, titik panas pada cluster-1 memiliki nilai confidence sebesar 61%, yang termasuk dalam kategori potensi kebakaran sedang karena nilainya di bawah 80%, sehingga perlu diwaspadai. Lokasi dari cluster 1 dan 2 dapat dilihat pada gambar 7 dan 8.



Gambar 7. Lokasi Persebaran Titik Api Cluster 1

Gambar 8. Lokasi Persebaran Titik Api Cluster 2

Hasil visualisasi lokasi titik api menunjukkan bahwa titik api tersebar hampir di seluruh Provinsi Kepulauan Bangka Belitung. Di cluster-1, rata-rata tingkat confidence mencapai 61,796140% dan dikategorikan sebagai sedang/waspada, sementara cluster-2 menunjukkan rata-rata tingkat confidence 94,142857% dan dikategorikan sebagai tinggi/butuh penanggulangan.

4. Kesimpulan

Berdasarkan pengujian dan analisis algoritma Agglomerative Hierarchical Clustering yang digunakan untuk mengelompokkan titik panas di Provinsi Kepulauan Bangka Belitung pada tahun 2019, dengan metode evaluasi Silhouette Coefficient, dapat disimpulkan bahwa algoritma tersebut berhasil diterapkan untuk Clustering titik panas di Provinsi Kepulauan Bangka Belitung pada tahun 2019. Hasil evaluasi clustering dilakukan menggunakan metode Silhouette Coefficient, yang menunjukkan bahwa jumlah cluster optimal adalah 2. Cluster pertama terdiri dari 829 titik, dengan rata-rata confidence sebesar 61,796140%, rata-rata brightness sebesar 320,839083, dan rata-rata FRP sebesar 20,902774. Sementara itu, cluster kedua memiliki 14 titik, dengan rata-rata confidence mencapai 94,142857%, rata-rata brightness sebesar 361,407143, dan rata-rata FRP sebesar 156,885714. Dengan demikian, potensi terjadinya kebakaran pada cluster kedua lebih tinggi dibandingkan cluster pertama, sehingga memerlukan perhatian dan penanganan yang lebih intensif.

References

[1] Kementerian Lingkungan Hidup dan Kehutanan, "Indikasi Luas Kebakaran," *Sistem Informasi Pemantauan Kebakaran Hutan Indonesia (SIPONGI)*, 2019. [Online]. Available: https://sipongi.menlhk.go.id/indikasi-luas-kebakaran. [Accessed: 20-Dec-2024].

[2] L. Giglio, I. Csiszar, and C. O. Justice, "Global distribution and seasonality of active fires as observed with the Terra and Aqua Moderate Resolution Imaging Spectroradiometer (MODIS) sensors," *Journal of Geophysical Research: Biogeosciences*, vol. 111, no. G2, pp. 1-12, 2006.

[3] J. W. G. Putra, *Pengenalan Konsep Pembelajaran Mesin dan Deep Learning*, ver. 1.4. Tokyo, Japan: Self-published, 2019. [Online]. Available: https://wiragotama.github.io/resources/ebook/intro-to-ml-secured.pdf. [Accessed: 20-Dec-2024].

[4] D. F. Pramesti, M. T. Furqon, and C. Dewi, "Implementasi metode k-medoids clustering untuk pengelompokan data potensi kebakaran hutan/lahan berdasarkan persebaran titik panas (hotspot)," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 1, no. 9, pp. 723-732, Sep. 2017.

[5] D. P. Sinaga, P. P. Adikara, and Y. A. Sari, "Klasterisasi Data Titik Api Menggunakan Metode Self Organizing Map di Wilayah Jawa," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 3, no. 10, pp. 9543-9551, Oct. 2019.

[6] S. Zhou, Z. Xu, and F. Liu, "Method for determining the optimal number of clusters based on agglomerative hierarchical clustering," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, no. 12, pp. 3007-3017, Dec. 2017.

[7] O. Roswintiarti, "Informasi titik panas (hotspot) kebakaran hutan/lahan," *Pusat Pemanfaatan Penginderaan Jauh, Deputi Bidang Penginderaan Jauh-LAPAN*, Jakarta, Indonesia, Technical Report, pp. 1-15, 2016.

[8] M. T. Furqon and A. W. Widodo, "Implementasi Metode Improved K-Means untuk Mengelompokkan Titik Panas Bumi," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 1, no. 11, pp. 1270-1276, Nov. 2017.

[9] M. Nishom and M. Y. Fathoni, "Implementasi Pendekatan Rule-Of-Thumb untuk Optimasi Algoritma K-Means Clustering," *Jurnal Informatika: Jurnal Pengembangan IT*, vol. 3, no. 2, pp. 237-241, May 2018.

[10] K. S. Artha, E. Winarko, and D. Ilmu, "Perbandingan Eros, Euclidean Distance dan Dynamic Time Warping dalam Klasifikasi Data Multivariate Time Series Menggunakan kNN," in *Proc. Seminar Nasional Pendidikan Teknik Informatika*, Singaraja, Indonesia, Aug. 2016, pp. 223-228.

[11] D. P. Kroese, Z. Botev, and T. Taimre, *Data Science and Machine Learning: Mathematical and Statistical Methods*. Boca Raton, FL, USA: Chapman and Hall/CRC, 2019.

[12] B. K. Amijaya, M. T. Furqon, and C. Dewi, "Clustering Titik Panas Bumi Menggunakan Algoritme Affinity Propagation," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 2, no. 10, pp. 3835-3842, Oct. 2018.

[13] R. Handoyo, R. Mangkudjaja, and S. M. Nasution, "Perbandingan metode clustering menggunakan metode Single Linkage dan K-means pada Pengelompokan Dokumen," *Jurnal Sifo Mikroskil*, vol. 15, no. 2, pp. 73-82, Oct. 2014.

Authors' Profiles

Aji Setiawan lahir pada tanggal 2 Februari 2004. Saat ini merupakan mahasiswa Program Studi Sistem Informasi Fakultas Sains dan Teknologi Universitas Jambi. Memiliki minat pada bidang Pengembangan Aplikasi Web dan Data Sains.

Akhiyar Waladi lahir di Bengkulu, pada tanggal 23. Ia menyelesaikan pendidikan S1 di bidang Ilmu Komputer dari Institut Pertanian Bogor, Bogor, Indonesia, pada tahun 2017, dengan fokus utama pada Internet of Things. Kemudian, ia melanjutkan studi S2 dalam bidang Ilmu Komputer di Universitas Indonesia, Depok, Indonesia,dan lulus pada tahun 2020 dengan penelitian utama di bidang Optical Character Recognition..

Rahmad Ashar lahir di Sungai Penuh, 7 Mei 1995. Ia menyelesaikan gelar Sarjana Komputer (S.Kom) Teknik Informatika pada Tahun 2018 dari Universitas Komputer Indonesia Bandung dan Memperoleh gelar Magister Komputer (M.Kom) Teknik Informatika pada Tahun 2022 dari Universitas Putra Indonesia YPTK Padang.