



Hotspot Clustering in Jambi Province Using Agglomerative Hierarchical Clustering Algorithm

Pengelompokkan Titik Api di Provinsi Jambi dengan Algoritma Agglomerative Hierarchical Clustering

Krisman Pratama Simanjuntak^{1*}, Ulfa Khaira²

^{1,2}Fakultas Sains dan Teknologi, Universitas Jambi, Indonesia

E-mail: ¹krismanpratama@gmail.com, ²ulfa.ilkom@gmail.com

Received December 27th 2020; Revised January 23th 2021; Accepted February 20th 2021
Corresponding Author: Krisman Pratama Simanjuntak

Abstract

Forest fires are one of Indonesia's most frequent disasters, especially areas with extensive forests and land such as Jambi Province. In 2019 there were at least 157,137 hectares of forest and land in Jambi Province burned, causing losses of 12 trillion rupiah. We can't let this ever happen again. Early prevention needs to be done by grouping hotspots data into a number of clusters so that it can identify which groups of hotspots are fire emergencies. One method of grouping hotspots data is Agglomerative Hierarchical Clustering. Agglomerative Hierarchical Clustering groups a certain amount of data based on the similarities that make up the hierarchical tree from the bottom up. In this study, clustering was conducted by grouping data into 2 to 10 clusters. Evaluation of clustering results is done by calculating silhouette coefficient and selecting the best coefficient of each number of clusters. By using 6,658 rows of data and then Confidence, Brightness, and FRP as attributes. The best number of clusters is 2 clusters with the Silhouette Coefficient value of 0.5856441. Cluster-1 consists of 6,283 points with an average confidence level of 73.49642% which is classified as moderate, then Cluster-2 consists of 375 points with an average level of confidence of 99.46133% which is classified as high so it needs to be prioritized for handling it.

Keyword: Agglomerative Hierarchical Clustering, Hotspot, Silhouette Coefficient.

Abstrak

Kebakaran adalah salah satu bencana yang kerap kali melanda Indonesia, khususnya daerah dengan hutan dan lahan yang luas seperti Provinsi Jambi. Pada tahun 2019 setidaknya terdapat 157.137 hektare hutan dan lahan di Provinsi Jambi terbakar yang menyebabkan kerugian sebesar Rp.12 triliun. Demi mencegah hal tersebut terulang kembali, perlu dilakukan pencegahan dini dengan cara mengelompokkan data titik panas bumi menjadi sejumlah cluster sehingga bisa mengidentifikasi kelompok titik panas mana yang darurat kebakaran. Salah satu metode mengelompokkan titik panas bumi adalah Agglomerative Hierarchical Clustering. Agglomerative Hierarchical Clustering mengelompokkan sejumlah data berdasarkan kemiripan yang membentuk pohon hierarki dari bawah ke atas. Pada penelitian ini, Clustering dilakukan dengan mengelompokkan data menjadi 2 sampai 10 cluster. Evaluasi hasil clustering dilakukan dengan menghitung Silhouette Coefficient dan memilih koefisien yang terbaik dari tiap jumlah cluster. Dengan menggunakan 6.658 row data lalu Confidence, Brightness, dan FRP sebagai atributnya, hasilnya jumlah cluster 2 adalah jumlah cluster terbaik dengan nilai Silhouette Coefficient sebesar 0,5856441. Cluster-1 terdiri dari 6.283 titik dengan rata - rata tingkat kepercayaan 73,49642% yang tergolong sedang, kemudian Cluster-2 terdiri dari 375 titik dengan rata - rata tingkat kepercayaan 99,46133% yang tergolong tinggi sehingga perlu diprioritaskan penanganannya.

Kata Kunci: Agglomerative Hierarchical Clustering, Hotspot, Silhouette Coefficient.

1. PENDAHULUAN

Kebakaran adalah salah satu bencana yang kerap kali melanda Indonesia, khususnya daerah dengan hutan dan lahan yang luas seperti Provinsi Jambi. Kebakaran hutan dan lahan (karhutla) adalah keadaan terbakarnya hutan atau lahan yang menyebabkan kerusakan populasi vegetasi, lingkungan, dan dampak yang merugikan lainnya. Pada tahun 2019 setidaknya terdapat 157.137 hektare hutan dan lahan di Provinsi Jambi terbakar yang menyebabkan kerugian sebesar Rp.12 triliun. Tak hanya itu, asap hasil kebakaran hutan dan lahan pun dapat mengganggu jarak pandang dan berpotensi menyebabkan penyakit pernapasan. Oleh karena

itu, pencegahan dini sangat penting dilakukan agar kemungkinan hutan terbakar dan lahan dalam skala besar tidak terjadi lagi, salah satunya dengan cara mengetahui kelompok daerah yang rawan terhadap titik api atau titik panas (hotspot).

Titik Panas Bumi atau Hotspot adalah indikasi kemungkinan terjadinya kebakaran pada suatu wilayah. Titik panas (Hotspot) merupakan daerah yang memiliki suhu permukaan relatif lebih tinggi dibandingkan daerah di sekitarnya berdasarkan ambang batas suhu tertentu yang terpantau oleh satelit penginderaan jauh [1]. Semakin banyak titik hotspot yang terpantau oleh satelit, maka semakin banyak pula potensi terjadinya kebakaran pada suatu daerah.

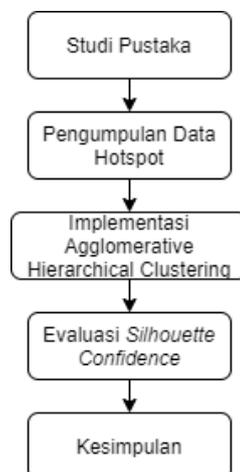
Memang benar bahwa titik panas bumi merupakan salah satu indikasi terjadinya kebakaran, namun tidak semua data titik panas tersebut adalah kebakaran yang benar - benar terjadi di lapangan. Maka dari itu perlu dilakukan validasi dan investigasi langsung di lapangan dan memastikan apakah titik tersebut benar terjadi kebakaran atau hanya sekedar indikasi yang menyebabkan panas sehingga terdeteksi oleh satelit. Namun dalam melakukan hal tersebut akan memakan banyak biaya dan waktu terlebih untuk titik lokasi dalam jumlah besar. Lagipula, belum tentu di wilayah tersebut benar adanya terjadi kebakaran. Oleh sebab itu, diperlukan proses penambangan data untuk menemukan pola dan kelompok data titik panas yang akan berguna untuk penanganan dini pada kebakaran.

Penambangan Data atau Data Mining adalah suatu metode pengolahan data dalam jumlah yang besar. Salah satu teknik dalam Data Mining adalah Clustering. Clustering adalah salah satu cara mengelompokkan data - data dengan karakter mirip [2]. Sehingga objek yang terdapat pada suatu cluster memiliki kemiripan dengan objek lainya didalam cluster yang sama namun memiliki perbedaan dengan objek lain pada cluster yang berbeda. Dengan melakukan clustering pada dataset titik panas di Provinsi Jambi akan didapatkan informasi apakah benar terjadi kebakaran pada suatu wilayah, ataupun informasi status tinggi, sedang, atau rendahnya potensi suatu wilayah mengalami kebakaran.

Beberapa penelitian mengenai pengelompokan titik api dengan teknik data mining telah dilakukan, diantaranya adalah penelitian yang mengimplementasi algoritma K-Medoids menggunakan data titik panas dengan parameter latitude, longitude, brightness, frp (fire radiative power), dan confidence. Algoritma K-Medoids merupakan metode clustering yang berfungsi untuk memecah dataset menjadi kelompok-kelompok. Kelebihan dari metode ini mampu mengatasi kelemahan dari metode K-Means yang sensitive terhadap outlier, proses clustering data titik panas dengan hasil silhouette coefficient terbaik sebesar 0,56745 pada penggunaan 2 cluster [3]. Penelitian lainnya, implemmentasi algoritma Self Organizing Maps (SOM) untuk pengelompokan titik api di Pulau Jawa, berdasarkan hasil pengujian didapatkan silhouette coefficient nilai tertinggi yaitu sebesar 0,2489 untuk penggunaan 3 cluster [4]. Pada penelitian ini akan menggunakan algoritma Agglomerative Hierarchical Clustering dalam mengelompokkan titik panas di Provinsi Jambi. Agglomerative hierarchical clustering (AHC) dengan menggunakan bottom-up, dimulai dari masing-masing data sebagai sebuah cluster, kemudian secara rekursif mencar kelompok terdekat sebagai pasangan yang kemudian akan digabungkan menjadi kelompok yang lebih besar. Proses tersebut dilakukan berulang sampai jumlah cluster 1, sehingga akan tampak bergerak keatas membentuk hirarki. Keunggulan dari algoritma AHC adalah tidak perlu menentukan jumlah klaster yang diinginkan [5].

2. BAHAN DAN METODE

Metode Penelitian menjelaskan alur penelitian secara umum yang dilakukan penelitian untuk mendapatkan hasil akhir berupa Klasterisasi titik panas di Provinsi Jambi. Adapun tahapan tahapan pada penelitian ini dapat digambarkan pada gambar 1



Gambar 1. Metodologi Penelitian

2.1 Studi Pustaka

Studi pustaka merupakan suatu metode dalam mengumpulkan informasi dan data dengan bantuan berbagai macam literatur yang ada di perpustakaan seperti dokumen, buku, dan jurnal penelitian. Pada penelitian ini, peneliti menghimpun informasi teoritis yang relevan dengan penelitian yang akan dilakukan melalui buku, jurnal penelitian sebelumnya, serta artikel di internet. Hal tersebut berguna untuk mengetahui metode dan algoritma apa saja yang tepat digunakan untuk menyelesaikan permasalahan yang diteliti dan sebagai dasar referensi yang kuat dalam menyelesaikan penelitian ini.

2.2 Pengumpulan Data Hotspot

Data yang digunakan pada penelitian ini adalah data Titik Panas atau Hotspot bumi khususnya di Provinsi Jambi. Data diperoleh melalui dokumentasi pada website resmi Nasa yang dapat diunduh secara gratis dan terbuka di websitenya. Hotspot merupakan suatu area yang memiliki suhu lebih tinggi dibandingkan dengan sekitarnya yang dapat deteksi oleh satelit. Area tersebut direpresentasikan dalam suatu titik yang memiliki koordinat tertentu [6]. Data yang didapat adalah Data titik api di provinsi Jambi yang telah direkam oleh Nasa dari Januari 2019 sampai dengan Desember 2019. Terdapat sebanyak 6.658 baris data dengan 28 atribut, namun yang digunakan hanya 5 atribut yaitu *Longitude*, *Latitude*, *Brightness*, *Confidence*, *Bright_T31*, *FRP*, dan Kabupaten. Dengan atribut *Brightness*, *Confidence*, dan *FRP* yang akan digunakan pada proses algoritma *Agglomerative Hierarchical Clustering*.

Titik panas bumi (*hotspot*) memiliki parameter tingkat kepercayaan yang mengindikasikan daerah yang ditandai titik tersebut adalah daerah dimana kebakaran terjadi [7]. Semakin tinggi tingkat kepercayaan, maka semakin tinggi pula potensi bahwa hotspot tersebut adalah benar-benar kebakaran lahan atau hutan yang terjadi di lapangan. LAPAN dalam *MODIS Active Fire Product User's Guide* membagi tiga kelas tingkat kepercayaan yang ditunjukkan pada tabel 1.

Table 1. Nilai Tingkat Kepercayaan

Confidence	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 merupakan salah satu teknik *Data Mining* yang bertujuan untuk mengelompokkan objek berdasarkan kemiripan karakteristiknya dan dapat dipisahkan dari kelompok objek lainnya, sehingga menciptakan kelompok objek yang homogen dan heterogram terhadap kelompok lainnya [9]. *Hierarchical Clustering* adalah teknik *clustering* membentuk hirarki sehingga membentuk struktur pohon. Dengan demikian proses pengelompokkannya dilakukan secara bertingkat atau bertahap. Terdapat 2 metode pada algoritma Hierarchical Clustering yaitu *Agglomerative (bottom-up)* dan *Devisive (top-down)*.

Pada tahap implementasi dilakukan pengolahan dataset yang telah dikumpulkan dan diproses dengan algoritma *Agglomerative Hierarchical Clustering* dengan bantuan *software* Rstudio dengan langkah-langkah sebagai berikut:

2.3.1 Praproses Data

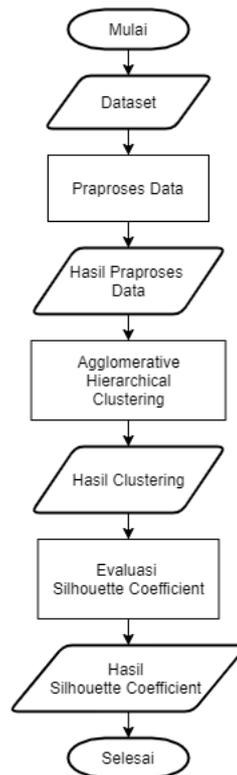
Praproses atau Preprocessing Data adalah tahap awal dalam data mining, yaitu menghilangkan permasalahan-permasalahan yang dapat mengganggu hasil daripada proses data. Data yang tersedia tentunya masih belum bersih dari missing value, duplikasi, serta data yang inkonsisten. Dengan melakukan praproses data, semua data tadi dibersihkan sehingga data yang akan digunakan hanyalah data yang benar benar berkualitas dan memberikan hasil *mining* atau pola dan pengetahuan baru yang lebih baik.

Salah satu tahapan dalam praproses data adalah Normalisasi. Normalisasi merupakan penskalaan kembali nilai atribut dari data sehingga ditetapkan pada range atau skala tertentu. Proses normalisasi dilakukan dengan metode *min-max*. Metode *min-max* merupakan metode normalisasi dengan melakukan transformasi linier terhadap data asli dengan persamaan yaitu [10]:

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

Dimana:

- 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 Hierarchical Clustering*

2.3.2 Perhitungan Jarak Euclidean

Euclidean Distance adalah metode yang umum dan mudah digunakan dalam menghitung jarak antar dua titik dalam Euclidean Space dimana setiap titik direpresentasikan dalam multidimensi [9]. Rumus dari metode ini:

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

Dimana:

- u dan v = dua objek yang akan dihitung jaraknya
- u_i = komponen ke i dari u secara berurutan

2.3.3 Proses *Agglomerative Hierarchical Clustering*

Pada *Agglomerative* cara kerjanya dimulai dari mengelompokkan objek objek individual yang mana pada awalnya jumlah cluster sama dengan banyaknya jumlah objek. Lalu objek objek yang memiliki kemiripan atau kedekatan dikelompokkan membentuk kelompok baru berdasarkan hasil perhitungan jarak dan parameter kedekatan yang digunakan. Selanjutnya dilakukan perhitungan jarak baru antar cluster dan menggabungkan cluster yang memiliki jarak terdekat. Begitu seterusnya hingga seluruh objek membentuk satu cluster [11].

2.3.4 Evaluasi *Silhouette Coefficient*

Hasil dari algoritma Hierarchical Clustering yang menghasilkan sejumlah cluster perlu dilakukan evaluasi. Evaluasi dilakukan guna mengetahui seberapa mirip atau seberapa dekat jarak antar objek dalam suatu cluster dan seberapa jauh jarak cluster terpisah dengan cluster lain [12]. Semakin mirip atau semakin dekat jarak antar objek dalam suatu cluster dan semakin jauh jarak objek terhadap objek di cluster lain maka semakin baik kualitas Klasterisasi tersebut.

Salah satu metode untuk mengetahui kualitas dari suatu cluster adalah *silhouette coefficient*. *Silhouette coefficient* merupakan gabungan dari dua metode yaitu metode cohesion dan separation. Dimana cohesion bertujuan mengukur seberapa dekat hubungan antara objek dalam sebuah cluster, dan metode separation bertujuan mengukur seberapa jauh sebuah cluster terpisah dengan cluster lain. Yang mana, jumlah cluster terbaik adalah jumlah cluster yang memiliki rata rata nilai *silhouette coefficient* tertinggi / mendekati 1. Tahapan - tahapannya adalah sebagai berikut [13]:

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

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

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

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

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

Dengan $d(I,C)$ adalah jarak rata-rata data I dengan semua objek pada cluster lain C dimana $A \neq C$.

$$b(i) = \min_{C \neq A} d(i, C) \tag{5}$$

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.

2. HASIL DAN ANALISIS

2.1 Praproses Data

Praproses data merupakan tahap awal pada data *mining* sebelum melanjutkan ke tahap tahap selanjutnya yang lebih jauh. Data yang telah dikumpulkan adalah data titik api di Provinsi Jambi yang terdata sejak 01 Januari 2019 sampai 31 Desember 2019 didapatkan dari situs resmi NASA <https://firms.modaps.eosdis.nasa.gov/>. Data berjumlah 6.658 *record* dan memiliki 28 atribut. Tidak semua atribut akan digunakan pada penelitian ini. Hanya 7 atribut yang digunakan yaitu *latitude*, *longitude*, *brightness*, *confidence*, *bright_t31*, *frp*, dan kabupaten. Namun untuk proses algoritma *Agglomerative Hierarchical Clustering* hanya digunakan atribut *brightness*, *confidence*, dan *frp* saja.

Dari 6.658 baris data, tanpa disadari mungkin masih terdapat baris data yang kehilangan nilai atribut (*missing value*), terdapat data yang duplikat, ataupun data yang tidak konsisten. Pada penelitian ini, teknik Praproses yang dilakukan adalah *Cleaning* dan *Transformation*. Pada proses *cleaning* data dibersihkan dari *missing value* dan duplikasi data. Sehingga data yang terlihat seperti gambar berikut :

Table 2. Dataset Setelah *Cleaning*

No	Brightness	Confidence	FRP
1	310,7	47	24,5
2	332,8	87	144,7
3	311,1	57	19,4
4	315,3	38	8,8
5	302	46	7,8
...
6654	316,1	63	8,5
6655	316,5	52	9,5
6656	314,8	25	8,0
6657	317,4	68	10,6
6658	315,1	23	7,2

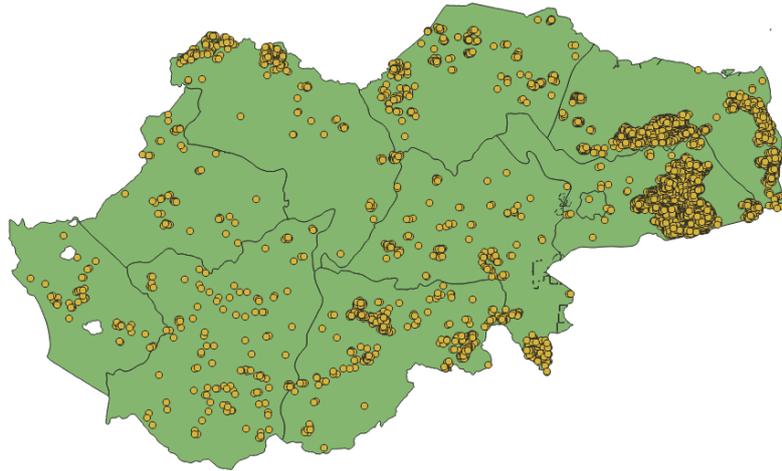
Setelah melewati proses *cleaning*, jumlah data tetap sebanyak 6.658 *record*. Hal ini karena tidak ada *missing value* atau duplikasi pada dataset tersebut. Selanjutnya dilakukan proses transformasi data dengan teknik normalisasi. Data yang merupakan data numerik akan dinormalisasi menggunakan metode *min - max* sehingga menghasilkan nilai yang memiliki skala dari 0 sampai dengan 1.

Table 3. Dataset Setelah Normalisasi

No	Brightness	Confidence	FRP
1	0,052348337	0,47	0,0036855406
2	0,160469667	0,87	0,0217672543

No	Brightness	Confidence	FRP
3	0,054305284	0,57	0,0029183465
4	0,074853229	0,38	0,0013237860
5	0,009784736	0,46	0,0011733558
...
6654	0,078767123	0,63	0,0012786570
6655	0,080724070	0,52	0,0014290872
6656	0,072407045	0,25	0,0012034418
6657	0,085127202	0,68	0,0015945604
6658	0,073874755	0,23	0,0010830977

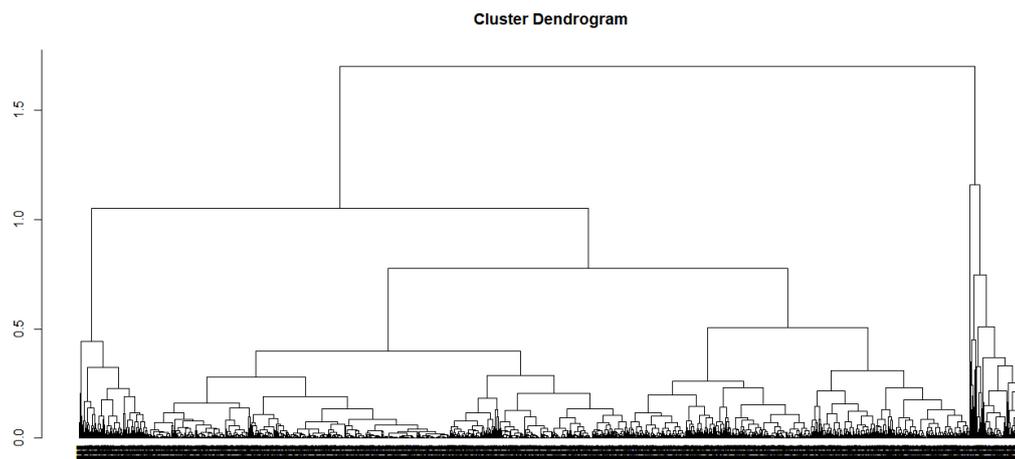
Untuk lebih jelasnya secara visualisasi peta dapat dilihat pada Gambar 3.



Gambar 3. Visualisasi Lokasi Titik Api di Provinsi Jambi Tahun 2019 Menggunakan Software QGIS

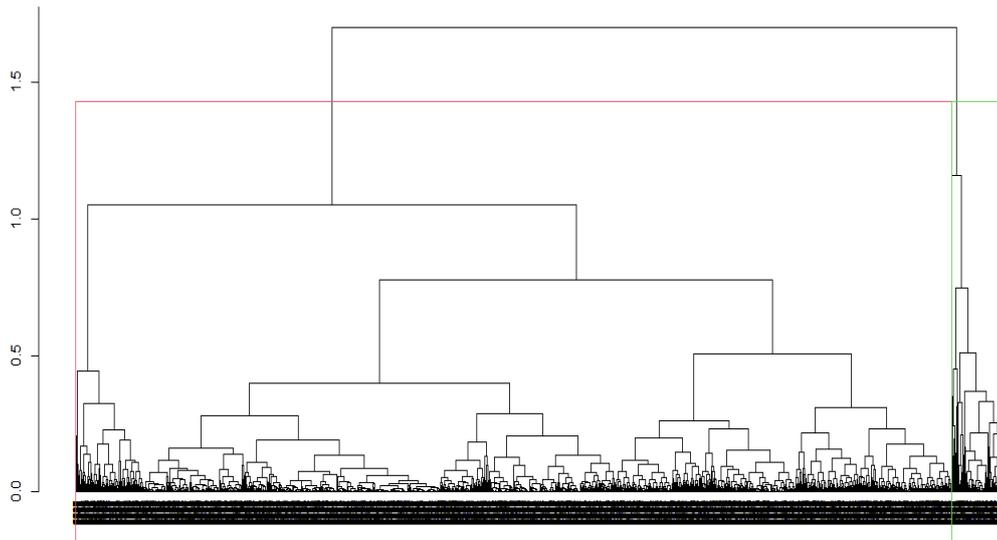
2.2 Algoritma Agglomerative Hierarchical Clustering

Sebelum melakukan proses *clustering* perlu melakukan perhitungan matriks jarak antar data. Perhitungan jarak dilakukan dengan metode atau rumus *Euclidean Distance*. *Agglomerative* mengklusterisasi data dari N cluster menjadi satu cluster secara bertahap. Dengan N adalah jumlah data atau bisa dibilang jumlah awal cluster pada *Agglomerative Hierarchical Clustering* adalah sebanyak jumlah objek/data yaitu 6.658 cluster. Visualisasi hasil *clustering* dapat dilihat pada dendrogram Gambar 4.



Gambar 4. Visualisasi Cluster dengan Dendrogram

Proses *clustering* menggunakan metode *Complete Linkage*. Metode *Complete linkage* menggabungkan cluster menurut jarak antara anggota anggota terjauh diantara 2 cluster. Sehingga jika menetapkan jumlah cluster atau $k = 2$. Maka akan memotong pada titik 1,47.



Gambar. 5 Visualisasi jumlah cluster sebanyak 2 cluster.

Setelah memotong pohon hierarki menjadi 2 cluster maka didapatkan 2 kelompok data yang telah terpisah dengan *Cluster-1* sebanyak 6.283 data, dan *Cluster-2* sebanyak 375 data.

Table 4 . Jumlah Data tiap Cluster

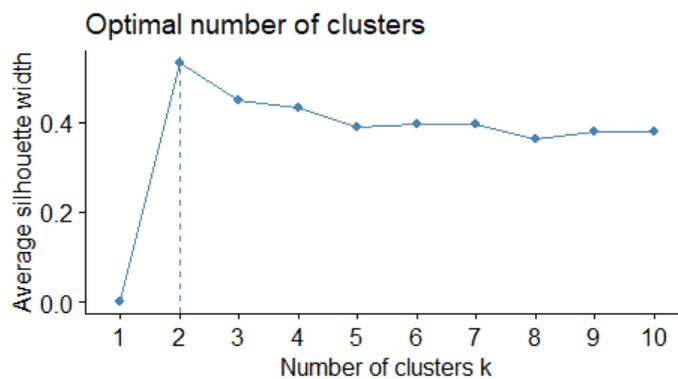
Cluster	Jumlah data
1	6.283
2	375

2.3 Evaluasi Hasil Cluster

Evaluasi *cluster* dilakukan dengan membandingkan nilai rata rata *Silhouette Coefficient* pada tiap tiap masukkan nilai jumlah *cluster* yang berbeda. Pengujian dilakukan dengan jumlah masukan *cluster* dari 2 sampai 10. Hasil perhitungan *silhouette coefficient* dapat dilihat pada tabel 5.

Table 5 . Nilai *Silhouette Coefficient*

Jumlah Cluster	<i>Silhouette Coefficient</i>
2	0.5856441
3	0.4714293
4	0.4269350
5	0.3929206
6	0.3943555
7	0.4292553
8	0.3786495
9	0.3922515
10	0.3823649



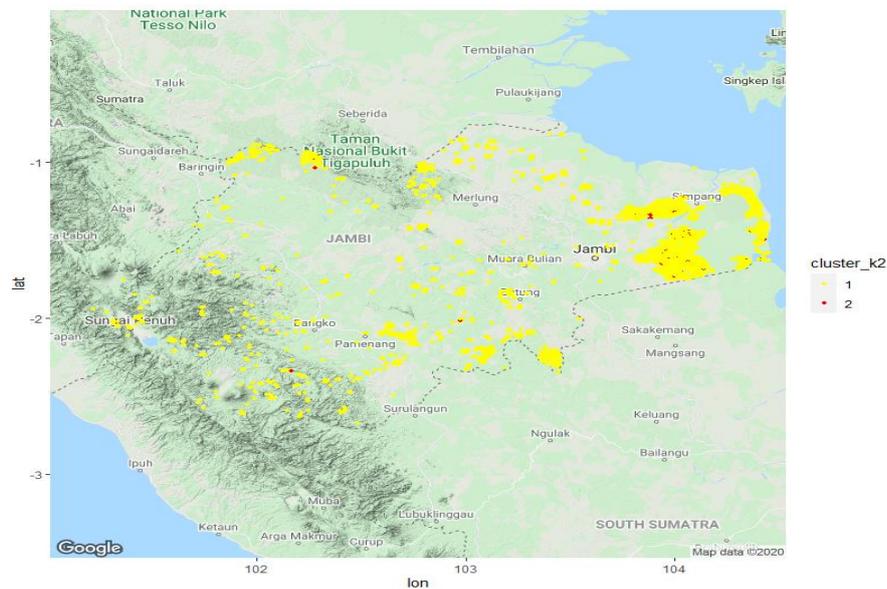
Gambar 6. Diagram nilai *Silhouette Coefficient* berdasarkan jumlah *cluster*

Berdasarkan pengujian yang terlihat pada gambar 5. Terlihat jumlah *cluster* dengan nilai *silhouette coefficient* tertinggi adalah 2. Sedangkan nilai *silhouette coefficient* terendah adalah jumlah *cluster* 8. Sehingga pengelompokan yang paling tepat dalam menganalisis potensi kebakaran adalah 2 *cluster* atau kelompok. Tingginya nilai *silhouette coefficient* mengindikasikan bahwa data yang diuji terkluster dengan baik, yaitu memiliki jarak yang besar atau jauh antar satu *cluster* ke *cluster* lain, dan jarak yang rendah atau dekat antar objek dalam suatu *cluster* yang sama [8].

2.4 Analisis

Pada *silhouette coefficient* apabila hasil semakin mendekati 1 maka kualitas kluster semakin baik, dan sebaliknya apabila hasil semakin mendekati 0 maka semakin buruk. Setelah menguji *cluster* dengan *silhouette coefficient* dari 2 *cluster* terus berturut sampai dengan 10 *cluster*, maka dapat disimpulkan 2 *cluster* adalah hasil terbaik dalam proses *clustering* tersebut.

Seperti yang terlihat pada tabel 3 data titik *hotspot* yang tersebar pada 2 *cluster* ada sebanyak 6.283 titik pada *cluster* 1 dan 375 titik pada *cluster* 2. Dengan menggunakan atribut *longitude* dan *latitude* pada data, bisa didapatkan lokasi asli dari titik *hotspot* pada peta, ditunjukkan pada gambar 7.



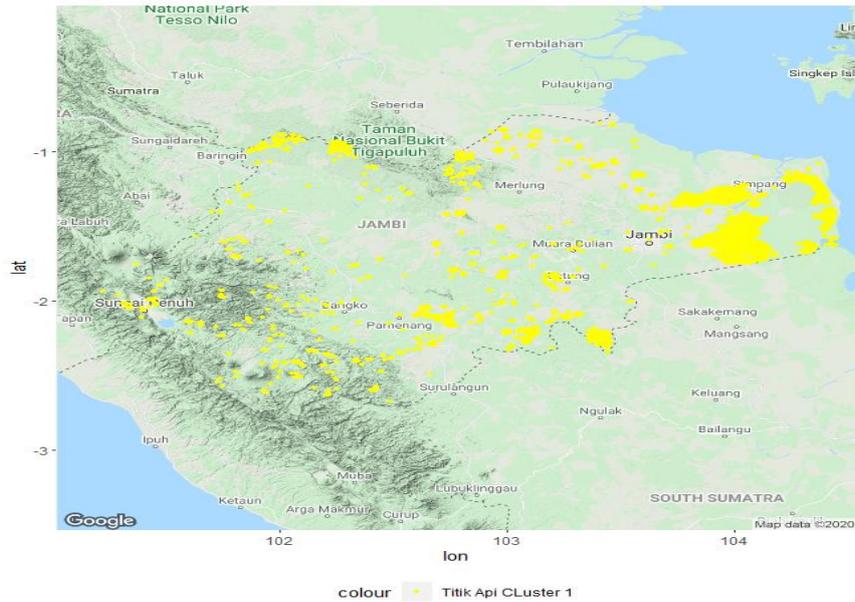
Gambar 7. Lokasi *hotspot* pada peta

Terlihat Titik kuning pada gambar 6 merupakan titik *hotspot cluster* 1 yang tersebar hampir diseluruh wilayah provinsi Jambi. Sedangkan titik merah adalah titik *hotspot cluster* 2. Dengan menghitung rata rata *confidence*, *brightness*, dan *frp* dari masing masing *cluster*, didapat hasil seperti pada tabel 6.

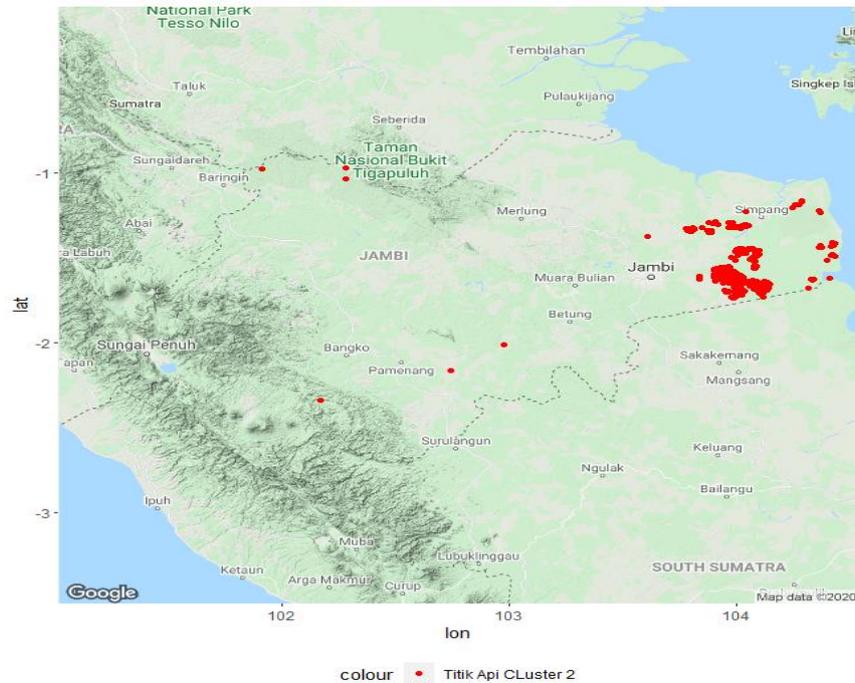
Table 6. Hasil rata - rata *Confidence*, *Brightness*, dan *FRP*

Cluster	Confidence	Brightness	FRP	Jumlah Data
1	73,49642	325,6746	46,64593	6283
2	99,46133	405,0715	679,95973	375

Rata rata nilai *confidence*, *brightness* dan *FRP* tertinggi dimiliki oleh *cluster-2* dengan *confidence* lebih dari 99%. Nilai tersebut menunjukkan bahwa titik *hotspot* yang terdapat pada *cluster-2* memiliki potensi lebih tinggi terjadi kebakaran. *Confidence* dengan nilai lebih dari 80 termasuk dalam kategori berpotensi tinggi sehingga perlu dilakukan penanggulangan lebih lanjut. Sedangkan titik panas pada *cluster-1*, memiliki nilai *confidence* 73,49642. termasuk potensi kebakaran sedang karena nilai *confidence* berada dibawah 80 dan harus di waspadai. Lokasi titik *cluster* 1 dan 2 dapat dilihat pada gambar 8 dan 9.



Gambar 8. Lokasi Persebaran Titik Api Cluster 1



Gambar 9. Lokasi Persebaran Titik Api Cluster 2

Hasil visualisasi lokasi titik api menunjukkan bahwa titik api tersebar hampir di seluruh kabupaten di provinsi Jambi dan yang terbanyak ada di Kabupaten Muaro Jambi baik itu cluster-1 yang memiliki rata rata *confidence* 73,49642% dan berstatus sedang/ waspada maupun pada cluster-2 dengan rata rata *confidence* 99,46133% dan berstatus tinggi/butuh penanggulangan. Hasil ini relevan dengan fakta yang terjadi pada tahun 2019 kabupaten Muaro Jambi adalah daerah memiliki paling banyak titik api/ kebakaran.

4. KESIMPULAN

Berdasarkan pengujian dan analisis dari algoritma *Agglomerative Hierarchical Clustering* untuk mengelompokkan titik panas Provinsi Jambi tahun 2019 yang diujikan dengan metode evaluasi *Silhouette Coefficient* maka didapat kesimpulan bahwa Algoritma *Agglomerative Hierarchical Clustering* berhasil diimplementasikan untuk mengelompokkan titik panas di Provinsi Jambi tahun 2019. Evaluasi hasil *clustering* dilakukan dengan metode *Silhouette Coefficient* dan didapatkan cluster optimal sebanyak 2 cluster dengan

cluster pertama sebanyak 6283 titik, rata - rata *confidence* sebesar 73,49642% , rata - rata *brightness* sebesar 325,6746 , rata - rata *FRP* sebesar 679,95973 dan pada *cluster* kedua sebanyak 375 titik, rata - rata *confidence* sebesar 99,46133% , rata - rata *brightness* sebesar 405,0715 , rata - rata *FRP* sebesar 679,95973. Dengan begitu, potensi terjadinya kebakaran pada *cluster-2* lebih tinggi dibanding *cluster-1* sehingga perlu perhatian dan penanganan lebih. Daerah yang memiliki paling banyak titik api berdasarkan hasil *clustering* adalah kabupaten Muaro Jambi. Hal ini relevan dengan kejadian nyata yang terjadi pada tahun 2019.

REFERENSI

- [1]. 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*, vol. 3.
- [2]. J.W.G. Putra, *Pengenalan Konsep Pembelajaran Mesin dan Deep Learning*. Tokyo: Tokyo Institute of Technology, 2020.
- [3]. PRAMESTI, Dyang Falila; FURQON, M. Tanzil; DEWI, Candra. Implementasi Metode K-Medoids Clustering Untuk Pengelompokan Data Potensi Kebakaran Hutan/Lahan Berdasarkan Persebaran Titik Panas (Hotspot). *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer e-ISSN*, 2017, 2548: 964X.
- [4]. SINAGA, Dika Perdana; ADIKARA, Putra Pandu; SARI, Yuita Arum. Klasterisasi Data Titik Api Menggunakan Metode Self Organizing Map di Wilayah Jawa. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer e-ISSN*, 2020, 2548: 964X.
- [5]. ZHOU, Shibing; XU, Zhenyuan; LIU, Fei. Method for determining the optimal number of clusters based on agglomerative hierarchical clustering. *IEEE transactions on neural networks and learning systems*, 2016, 28.12: 3007-3017.
- [6]. Roswintarti, O., 2016. Informasi Titik Panas (Hotspot) Kebakaran Hutan/Lahan. Pusat Pemanfaatan Penginderaan Jauh Deputi Bidang Penginderaan Jauh – LAPAN, pp. 1-15.
- [7]. Shoolihah, M.T. Furqon, A.W.Widodo, "Implementasi Metode Improved K-Means untuk Mengelompokkan Titik Panas Bumi. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 1, no. 11, pp. 1270-1276, November 2017
- [8]. S. Shalev-Shwartz and S. Ben-David, *Understanding Machine Learning From Theory to Algorithm*, New York: Cambridge University Press, 2014.
- [9]. M. Nishom and M.Y. Fathoni, "Implementasi Pendekatan Rule-Of-Thumb untuk Optimasi Algoritma K-Means Clustering," *Jurnal Pengembangan IT (JPIT)*, vol. 9, no. 2, pp. 237-241, May 2018.
- [10]. K. S. Artha and E. Winarko, "Perbandingan Eros, Euclidean Distance dan Dynamic Time Warping dalam Klasifikasi Data Multivariate Time Series Menggunakan kNN," *Prosiding Seminar Nasional Pendidikan Teknik Informatika*, pp. 223 - 228, August 2016.
- [11]. Dirk P. Kroese, et.al., *Data Science and Machine Learning: Mathematical and Statistical Methods*. Boca Raton: CRC Press, 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, October 2018.
- [13]. R. Handoyo, R. Rumani M., and S. M. Nasution, "Perbandingan Metode Clustering Menggunakan Metode Single Linkage dan K-Means pada Pengelompokan Dokumen," vol. 15, pp. 73-82, October 2014.