

# Grouping Potential Forest and Land Fires in Indonesia Based on Hotspot Distribution Using CLARANS Methods

Silfia Wisa Fitri, Zamahsary Martha\*, Yenni Kurniawati, Zilrahmi

Departemen Statistika, Universitas Negeri Padang, Padang, Indonesia

\*Corresponding author: [zamahsarymartha@fmipa.unp.ac.id](mailto:zamahsarymartha@fmipa.unp.ac.id)

Submitted : 12 Juni 2024  
Revised : 01 Agustus 2024  
Accepted : 12 Agustus 2024

## ABSTRACT

*Forest/land fires in Indonesia still occur every year. In dealing forest and land fires, to develop a system that can predict forest and land fires. In addition to predicting forest fires, anticipation that can be done to help forest fire prevention is to group areas that have fire potential. In this research, hotspot are clustered to divide areas that have the potential to burn. Clustering of areas using the CLARANS algorithm. CLARANS is a variant of k-medoid algorithm and development of previous algorithms, such as PAM and CLARA to handle larger of data and robust to outliers. The data uses in this calculation is the Indonesian region using longitude, latitude, Brightness, frp, and confidence variables to determine the clustering of fire areas. The data in this study is secondary data from the NASA website about Hotspots in 2023. The analysis technique is checking outliers, normalizing the data, performing cluster analysis, and then validating the result cluster. The purpose of this study is to cluster hotspots of forest/land fires based on areas that have the potential for forest/land fires so that have the opportunity to be followed up immediately. From the results obtained the formation of 2 clusters which shows that 2 hotspot groupings in the medium and high classes with a silhouette coefficient evaluation value of 0.896%. The results show cluster 1 is included in high potential with average brightness of 340<sup>0</sup>K and confidence of 95%, and Cluster 2 is included in medium potential with average brightness of 327<sup>0</sup>K.*

**Keywords:** CLARANS, Clustering, Hostpot, Indonesia.



This is an open access article under the Creative Commons 4.0 Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. ©2022 by author and Universitas Negeri Padang.

## I. PENDAHULUAN

Kebakaran hutan/lahan (karhutla) adalah peristiwa penting yang menjadi perhatian masyarakat nasional maupun internasional, karena dapat menimbulkan masalah ekologi, sosial dan perekonomian. Akibat Kebakaran hutan/lahan tersebut membuat negara Indonesia menduduki urutan ketiga di dunia dalam hal penyumbang polusi terbesar, setelah Amerika dan Tiongkok (Fatmala dkk, 2019). Penyebab kebakaran hutan/lahan biasanya terjadi karena dua faktor yaitu faktor alam seperti petir dan faktor yang disebabkan oleh manusia. Berdasarkan hasil studi Adinugroho (2015) menunjukkan bahwa di Indonesia sekitar 99% kebakarakan hutan/lahan disebabkan karena ulah manusia, baik itu karena pembakaran yang disengaja maupun akibat kelalaian saat menyiapkan lahan sehingga terjadi penjaralan api.

Peristiwa kebakaran hutan/lahan dapat dipantau dengan memanfaatkan teknologi penginderaan jauh milik NASA (*National Aeronautics and Space Administration*), seperti satelit Terra/Aqua dengan bantuan sensor MODIS (*Moderate-Resolution Imaging Spectroradiometer*). Satelit akan menemukan titik radiasi api, yang kemudian akan ditunjukkan sebagai titik panas (Suwarsono dkk, 2014). Titik panas dapat diartikan sebagai suatu area yang memiliki suhu permukaan yang lebih tinggi daripada area sekitarnya, yang dipantau oleh satelit penginderaan jauh berdasarkan ambang batas suhu tertentu (LAPAN, 2014). Potensi terjadinya kebakaran hutan/lahan di suatu wilayah akan semakin besar, jika persebaran titik panas yang terpantau oleh satelit semakin banyak.

Di Indonesia kerap terjadi kebakaran hutan/lahan, seperti pada tahun 2020 peristiwa kebakaran hutan/lahan menhanguskan sekitar 296 ribu ha lahan Indonesia. Sepanjang tahun 2020 provinsi yang mengalami kebakaran hutan/lahan paling besar yaitu Provinsi Nusa Tenggara Timur mencapai 114 ribu ha dan berdampak menghasilkan emisi sebesar 6,54 juta ton. Kemudian pada tahun 2021 di beberapa tempat di Indonesia kembali terjadi kebakaran hutan/lahan, seperti di Provinsi Kalimantan Barat terjadi sebanyak 52 kejadian kebakaran hutan/lahan dan 29 kejadian kebakaran hutan/lahan juga tercatat di Provinsi Riau (KLHK, 2021). Untuk mencegah kebakaran hutan/lahan, langkah pertama yang dapat dilakukan adalah memperkirakan wilayah mana saja yang mungkin terjadi kebakaran hutan/lahan di Indonesia. Potensi kebakaran hutan/lahan dapat diketahui berdasarkan informasi pada data titik panas yaitu dengan

cara mengelompokan data tersebut, sehingga diketahui wilayah serta potensi kebakaran hutan/lahan yang terjadi. Kluster merupakan metode yang digunakan dalam proses pengelompokan, salah satu metode kluster ialah *Clustering Large Application based on RANdomized Search* (CLARANS). CLARANS merupakan bagian dari algoritma k-medoid yang dikembangkan berdasarkan algoritma sebelumnya seperti *Partitioning Around Medoids* (PAM) dan *Clustering Large Application* (CLARA). Kedua algoritma itu dirancang untuk menangani data yang lebih besar serta *robust* terhadap *outlier*. Hal ini berdasarkan hasil penelitian Ng & Han (2002) menyatakan bahwa jika dibandingkan dengan PAM, CLARANS lebih efisien karena CLARANS tidak mengecek seluruh objek *non medoid*, dan di sisi lain algoritma CLARANS melakukan pencarian *medoid* yang tidak terbatas pada area lokal saja sehingga menghasilkan kualitas kluster yang lebih baik jika dibandingkan dengan CLARA.

Penelitian tentang titik panas sebelumnya pernah dilakukan oleh Pramesti (2017), dengan menggunakan metode k-medoid terhadap persebaran titik panas di wilayah Asia Tenggara pada tahun 2015. Hasil yang diperoleh data terbagi atas dua kluster dengan nilai *silhouette coefficient* yaitu 0,567. Selanjutnya, terdapat juga penelitian yang melakukan pengelompokan data persebaran titik panas di wilayah Australia pada tahun 2019 sampai 2020 dengan menerapkan algoritma CLARA oleh Lindrawati dkk. (2022). Berdasarkan hasil kluster diperoleh nilai *silhouette coefficient* sebesar 0,89 dengan pembagian dua kluster. Menurut Ng & Han (2002), metode CLARANS menghasilkan kualitas kluster yang lebih baik dibandingkan metode K-Medoid dan CLARA, sehingga penulis tertarik untuk mengkaji metode CLARANS terhadap data titik panas dalam penelitian yang berjudul “Pengelompokan Potensi Kebakaran Hutan/Lahan di Indonesia Berdasarkan Sebaran Titik Panas Menggunakan Metode CLARANS”.

## II. METODE PENELITIAN

### A. Jenis Penelitian dan Sumber Data

Jenis penelitian ini merupakan penelitian terapan dengan menggunakan data sekunder, yang bersumber dari situs resmi NASA LANCE-FIRM MODIS *Active Fire* yang merupakan data titik panas tahun 2023 dengan 12.287 pengamatan. Dalam penelitian ini digunakan variabel-variabel seperti yang terlihat pada Tabel 1.

Tabel 1. Variabel Penelitian

Nama	Keterangan	Skala	Satuan
$X_1$	<i>Brightness</i>	Interval	Kelvin (K)
$X_2$	<i>Confidence</i>	Interval	Persen (%)
$X_3$	<i>Fire Power Radiative</i> (FRP)	Interval	MegaWats (MW)

Sumber: (Pramesti, 2017 dan Lindrawati, 2022)

Data yang digunakan merupakan data dengan tingkat kepercayaan (*confidence*) di atas 80%. Prayoga dkk (2017), menyatakan bahwa pada dasarnya metode perekaman titik api oleh Satelit Terra/Aqua melakukan penentuan tingkat kepercayaan dengan nilai minimal 80%, yang menandakan jika titik api yang diidentifikasi tersebut telah dilewati dan direkam oleh kedua Satelit Terra dan Aqua, sehingga data yang digunakan telah akurat dan memiliki tingkat kepercayaan yang tinggi. Tingkat kepercayaan dalam informasi titik panas dapat dilihat pada Tabel 2.

Tabel 2. Tingkat Kepercayaan Titik Panas

Tingkat Kepercayaan (C)	Kelas
$0\% \leq C < 30\%$	Rendah
$30\% \leq C < 80\%$	Sedang
$80\% \leq C \leq 100\%$	Tinggi

Sumber: (Giglio, 2015)

### B. Teknik Analisis Data

Analisis data pada penelitian dilakukan dengan menggunakan *software R Studio* dan *Tableau*. Setelah variabel dan data dalam analisis dipilih, selanjutnya dilakukan analisis data dengan Langkah-langkah sebagai berikut.

1. Melakukan pra-pemrosesan data dengan memeriksa adanya pencilan multivariat pada data yang sudah dipilih.
2. Melakukan uji asumsi multikolinearitas pada data, berdasarkan hasil studi Hair JR dkk (2009) menjelaskan bahwa dalam analisis kluster hanya ada dua jenis uji yang dapat dilakukan yaitu sampel representatif dan uji multikolinieritas. Dalam penelitian ini data yang digunakan merupakan data populasi, maka pengujian sampel representatif tidak perlu dilakukan.
3. Mengubah data kedalam bentuk standar baku dengan melakukan standarisasi data dengan mengkonversi setiap variabel menjadi nilai standar atau sering disebut *Z-Score* dengan persamaan berikut.

$$Z_i = \frac{x_i - \bar{x}}{s} \tag{1}$$

- Dimana  $x_i$  merupakan data ke- $i$ ,  $\bar{x}$  merupakan rata-rata dari data dan  $s$  merupakan simpangan baku dari data.
4. Menentukan jumlah kluster optimal yang dibentuk pada data dengan menggunakan metode *silhouette* dengan melihat grafik yang terbentuk. Proses penentuan jumlah kluster menggunakan bantuan *software Rstudio* dengan *package library(factoextra)* menggunakan fungsi *fviz\_nbclust()*.
  5. Menentukan nilai untuk parameter *maxneighbor* dan *numlocal*, dimana *maxneighbor* merupakan jumlah maksimal *node* tetangga (*neighbor*) yang diperiksa dari sebuah *node* yang terpilih dan *numlocal* merupakan jumlah minimum lokal yang menentukan berapa kali dilakukannya iterasi atau menentukan berapa kali *node* diperiksa. Nilai *maxneighbor* yang digunakan yaitu  $p\% \times k(n - k)$  dan nilai *numlocal* yang digunakan yaitu 2. Adapun nilai  $p$  antara 0,25 dan 0,5 persen yang menyatakan total kemungkinan *neighbor* memberikan keseimbangan yang baik antara hasil dan waktu pengoperasian yang cepat,  $k$  menyatakan banyaknya kluster yang terbentuk dan  $n$  menyatakan jumlah objek dalam himpunan data.
  6. Memilih secara acak *node* dari  $n$  objek sebagai *medoid* sementara sesuai dengan jumlah  $k$  yang telah ditentukan.
  7. Memilih secara acak *node* tetangga (N) dari *node* yang terpilih (M).
  8. Menghitung tingkat kemiripan kedua *node* tersebut (N dan M) dengan menggunakan ukuran jarak Euclidean berdasarkan persamaan berikut.

$$d_{ij} = \sqrt{\sum_{k=1}^p (x_{ik} - x_{jk})^2} \tag{2}$$

- Dimana  $d_{ij}$  merupakan jarak Euclidean antar objek ke- $i$  dengan objek ke- $j$ ,  $x_{ik}$  merupakan objek ke- $i$  pada variabel ke- $k$ ,  $x_{jk}$  merupakan objek ke- $j$  pada variabel ke- $k$ , dan  $p$  merupakan jumlah variabel.
9. Menghitung nilai *cost* kedua *node* tersebut (N dan M) dengan menggunakan persamaan berikut.

$$E = \sum_{i=1}^k \sum_{p \in C_i} d(p, O_i) \tag{3}$$

- Dimana  $E$  merupakan jumlah kesalahan mutlak/*cost*,  $p$  merupakan objek nonrepresentatif /*non-medoid* dari kluster  $C_i$  dan  $O_i$  merupakan objek representatif /*medoid* dari kluster  $C_i$ .
10. Dari hasil perhitungan tingkat kemiripan dan *cost* tersebut *node* yang terpilih selanjutnya ialah *node* yang memiliki nilai *cost* terendah. Apabila N memiliki nilai *cost* lebih rendah dibanding M, maka N menggantikan M sebagai *node* terpilih.
  11. Setelah menemukan *node* yang terpilih, selanjutnya ulangi langkah 7-10 hingga mencapai nilai *maxneighbor* yang telah ditentukan.
  12. Setelah mencapai nilai *maxneighbor* yang telah ditentukan, kemudian kembali ke langkah 6-11 hingga tahap iterasi mencapai nilai *numlocal* yang telah ditentukan.
  13. Memilih *medoid* berdasarkan *node* yang memiliki nilai *cost* terendah dari setiap iterasi.
  14. Menghitung nilai validasi *coefficient silhouette*, untuk melihat seberapa baik hasil kluster yang digunakan pada algoritma CLARANS dengan menggunakan persamaan berikut.

$$SC(i) = \frac{1}{n_k} \sum_{x_{ij} \in C_k} \frac{b_i - a_i}{\max(a_i, b_i)} \tag{4}$$

- Dengan  $i = 1, 2, \dots, n_k$ ;  $k = 1, 2, \dots$ ;  $x_{ij}$  merupakan objek amatan ke- $i$  variabel ke- $j$ ;  $C_k$  merupakan kluster ke- $k$ ;  $b_i$  merupakan ukuran kohesi objek ke- $i$  dan  $a_i$  merupakan ukuran separasi objek ke- $i$ . Adapun kriteria pengukuran pengelompokan berdasarkan *silhouette coefficient* dapat dilihat pada Tabel 3.

**Tabel 3.** Kriteria *Silhouette Coefficient*

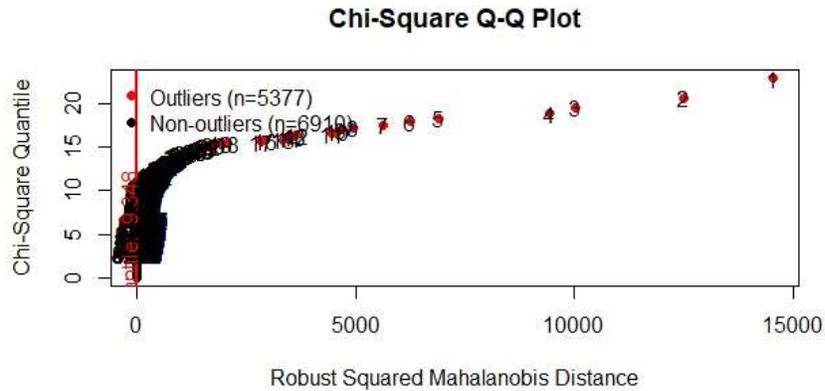
Nilai <i>Silhouette Coefficient</i>	Kriteria
$-1,00 \leq 0,25$	Struktur buruk
$0,26 - 0,50$	Struktur lemah
$0,51 - 0,70$	Struktur baik
$0,71 - 1,00$	Struktur kuat

Sumber: (Kaufman & Roesseeuw, 2005)

### III. HASIL DAN PEMBAHASAN

#### A. Mendeteksi Pencilan (*Outlier*)

Pencilan (*Outlier*) pada data titik panas untuk wilayah Indonesia tahun 2023 dapat dilihat pada Gambar 1.

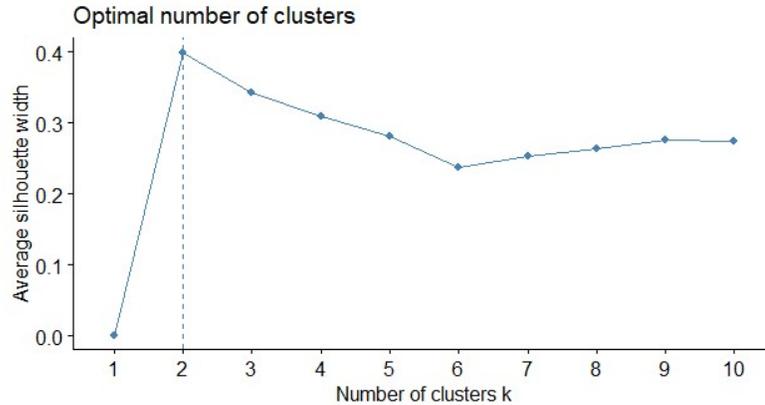


**Gambar 1.** Pencilan Pada Data Titik Panas Tahun 2023

Menurut Gambar 1 hasil perhitungan jarak Mahalanobis pada data penelitian titik panas menunjukkan bahwa sebanyak 5.377 titik merupakan area amatan pencilan. Oleh karena itu, metode CLARANS adalah analisis kluster yang tepat untuk digunakan pada penelitian ini, karena metode CLARANS *robust* terhadap pencilan.

#### B. Analisis Penentuan Potensi Kebakaran Hutan/Lahan

Jumlah kluster yang akan dibentuk pada penelitian ini ditentukan berdasarkan nilai *silhouette coefficient* tertinggi. Gambar 2 menunjukkan jumlah kluster optimal yang dibentuk berdasarkan nilai *silhouette coefficient*.



**Gambar 2.** Plot *Silhouette Coefficient* CLARANS

Berdasarkan Gambar 2 terlihat bahwa saat jumlah kluster = 2 memiliki nilai *silhouette coefficient* tertinggi, sehingga jumlah kluster yang optimal untuk dibentuk dengan metode CLARANS adalah dua kluster. Dengan demikian berdasarkan persebaran data titik panas dapat dilakukan pengelompokan data potensi kebakaran hutan/lahan di Indonesia dengan membentuk dua kluster pada metode CLARANS. Setelah tahap persebaran titik kluster terbentuk selanjutnya perlu dilakukan evaluasi untuk menentukan kualitas hasil kluster dengan melihat nilai *silhouette coefficient*. Nilai *silhouette coefficient* tertinggi menunjukkan hasil pengelompokan terbaik. Tabel 4 menunjukkan nilai *silhouette coefficient* hasil pengelompokan wilayah yang memiliki potensi kebakaran hutan/lahan dalam penelitian ini.

**Tabel 4.** Nilai *Silhouette Coefficient*

Metode	<i>Silhouette Coefficient</i>
CLARANS	0,896

Berdasarkan Tabel 4 nilai *silhouette coefficient* yang dihasilkan dengan menggunakan algoritma CLARANS ialah sebesar 0.896, maka struktur hasil klasterisasi yang dihasilkan sangat baik karena nilainya berada pada rentang 80-100. Dari hasil tersebut terlihat bahwa kualitas hasil klasterisasi yang terbentuk sangat baik dan dapat diterima. Untuk mengetahui potensi kebakaran hutan/lahan, analisis hasil klaster yang telah diuji digunakan untuk mengetahui tingkat potensi kebakaran hutan/lahan. Nilai *brightness* yang lebih dari 330<sup>0</sup>K termasuk dalam potensi kebakaran dengan keyakinan tinggi (*high confidence fire*), yang menunjukkan bahwa api tersebut memiliki kemungkinan yang tinggi terjadinya kebakaran atau pembukaan lahan. Tabel 5 menunjukkan hasil analisis yang diperoleh.

**Tabel 5.** Hasil Analisis Penentuan Potensi Karhutla

Klaster	Rata-Rata <i>Brightness</i> ( <sup>0</sup> K)	Rata-Rata <i>Confidence</i> (%)	Rata-Rata FRP (MW)	Kategori Kebakaran
1	340	95	61	Tinggi
2	327	83	34	Sedang

Berdasarkan Tabel 5 menunjukkan bahwa klaster 1 memiliki nilai rata-rata *brightness* 340<sup>0</sup>K, *confidence* 95% dan FRP 61 MW dimana nilai rata-rata *brightness* tersebut lebih dari 330<sup>0</sup>K dan nilai rata-rata *confidence* lebih dari 80%, serta memiliki nilai FRP cukup tinggi yang menandakan bahwa polusi yang dihasilkan sangat mempengaruhi iklim global. Berdasarkan hasil yang diperoleh, klaster 1 termasuk ke dalam kategori *high confidence fire* sehingga membutuhkan tindakan penanggulangan dengan cepat. Sedangkan klaster 2 memiliki nilai rata-rata *brightness* 327<sup>0</sup>K, *confidence* 83% dan FRP 34 MW, dimana klaster 2 termasuk ke dalam kategori *nominal confidence fire* sehingga memiliki potensi sedang dan dibutuhkan tindakan kewaspadaan terhadap terjadinya kebakaran.

Selanjutnya dapat dilakukan visualisasi pada peta menggunakan variabel longitude dan latitude berdasarkan persebaran titik panas pada masing-masing klaster Adapun hasil klaster wilayah-wilayah potensi kebakaran dengan metode CLARANS ditunjukkan seperti pada Gambar 3.



**Gambar 3.** Visualisasi persebaran titik panas

Gambar 3 menunjukkan klaster pertama dengan titik berwarna merah dan klaster kedua dengan titik berwarna kuning. Berdasarkan gambar terlihat bahwa titik-titik berwarna merah memiliki jumlah yang lebih banyak dibandingkan titik berwarna kuning. Hal tersebut menjelaskan bahwa potensi terjadinya kebakaran hutan/lahan di Indonesia sangatlah tinggi. Titik-titik dengan potensi kebakaran tinggi tersebut tersebar disebagian besar pulau Sumatra dan Kalimantan serta dibeberapa wilayah bagian timur Indonesia yaitu pulau Sulawesi terutama wilayah Poso, Morowali Utara dan Kolaka serta pulau Papua terutama di wilayah Meraoke dan Keppi. Hal ini menunjukkan bahwa potensi terjadinya kebakaran hutan/lahan di Indonesia tidak hanya terjadi di wilayah Sumatera dan Kalimantan namun hampir diseluruh wilayah Indonesia, sehingga pemerintah harus melakukan tindakan lebih lanjut untuk menanggulangi terjadinya peristiwa kebakaran hutan/lahan terhadap wilayah-wilayah tersebut.

#### IV. KESIMPULAN

Sesuai dengan tahap perhitungan CLARANS, metode CLARANS dapat digunakan untuk menangani kluster data yang menunjukkan potensi terjadinya kebakaran hutan/lahan berdasarkan titik panas. Dari hasil metode *silhouette coefficient*, yang memiliki jumlah kluster optimal 2 dan nilai *silhouette coefficient* 0,896, menunjukkan bahwa teknik ini dapat digunakan untuk menangani data sebesar 12.287 amatan. Kluster 1 memiliki potensi kebakaran hutan/lahan yang lebih tinggi, dengan *brightness* rata-rata 340<sup>0</sup>K dan *confidence* rata-rata 95%. Di sisi lain, kluster 2 dengan kecerahan rata-rata 327<sup>0</sup>K dan *confidence* rata-rata 83%, memiliki potensi kebakaran hutan/lahan yang sedang. Dari hasil kluster titik panas yang terbentuk dapat dipertimbangkan pembentukan pos pantau sesuai dengan lokasi yang pemetaan yang memerlukan perhatian khusus untuk pemantauan lebih lanjut.

#### DAFTAR PUSTAKA

- Adinugroho, W. C., I N. N. Suryadiputra, Bambang. H. S., & Labueni, S. (2015). *Panduan Pengendalian Kebakaran Hutan dan Lahan Gambut*. Wetland International-IP Katalog dalam Terbitan (KDT). Bogor.
- Fatmala, S, K., & Syafria, F. (2019). Clustering Data Persebaran Titik Panas (Hotspot) Menggunakan Metode Self Organizing Maps. *Skripsi*. Pekanbaru: Universitas Islam Negeri Sultan Syarif Kasim Riau.
- Giglio, L. (2015). *MODIS Collection 6 Active Fire Product User's Guide Revision A*. Department of Geographical Sciences. University of Maryland.
- Kaufman, L., & Rousseeuw, P. J. (2005). *Finding Groups in Data: An Introduction to Cluster Analysis*. New Jersey: Wiley Publishers.
- Kementerian Lingkungan Hidup dan Kehutanan. (2021). Diakses pada 21 Februari 2024 dari <https://ppid.menlhk.go.id/berita/siaran-pers/5898/empat-pesan-menteri-lhk-pada-peringatan-hari-hutan-internasional>.
- Lembaga Penerbangan dan Antariksa Nasional. (2014). *Lapan Menuju Pusat Keunggulan Sains, Teknologi, Pemanfaatan Dan Kajian Kebijakan Penerbangan Dan Antariksa*. Jakarta: Lembaga Penerbangan dan Antariksa Nasional.
- Lindrawati, E., Saiful, B., Umam, F.Z., Vinnesa, P.C., Kusri., & Dina, M. (2022). Implementasi Metode CLARA Clustering Untuk Pengelompokan Data Potensi Kebakaran Hutan/Lahan Berdasarkan Persebaran Titik Panas (Hostpot). *Journal of Computer System and Informatics*, 3(4), pp: 507-511.
- Ng, R. T., & Han, J. (2002). CLARANS: A Method for Clustering Objects for Spatial Data Mining. *IEEE Transactions on Knowledge and Data Engineering*, 14(5), pp: 1003-1016.
- Pramesti, D.F., M. Tanzil, F., & Candra, D. (2017). Implementasi Metode K-Medoids Clustering Untuk Pengelompokan Data Potensi Kebakaran Hutan/Lahan Berdasarkan Persebaran Titik Panas (Hostpot). *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 1(9), pp: 723-732.
- Prayoga, M. B. R., Yananto, A., & Kusumo, D. A. (2017). Analisis Korelasi Kerapatan Titik Api Dengan Curah Hujan di Pulau Sumatera Dan Kalimantan. *Jurnal Sains & Teknologi Modifikasi Cuaca*, 18(1), pp: 17 – 24.
- Suwarsono, Khomarudin MR, Parwati. (2014). *Aplikasi Penginderaan Jauh untuk Bencana Geologi*. Berita Inderaja, 8(10), Edisi Desember 2014.