

## Implementasi Metode *K-Medoids Clustering* Untuk Pengelompokan Data Potensi Kebakaran Hutan/Lahan Berdasarkan Persebaran Titik Panas (*Hotspot*)

Dyang Falila Pramesti<sup>1</sup>, M. Tanzil Furqon<sup>2</sup>, Candra Dewi<sup>3</sup>

Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Brawijaya  
Email: <sup>1</sup>dyangfpramesti@gmail.com, <sup>2</sup>m.tanzil.furqon@ub.ac.id, <sup>3</sup>dewi\_candra@ub.ac.id

### Abstrak

Kebakaran hutan/lahan merupakan salah satu bencana yang setiap tahunnya terjadi di beberapa negara di dunia. Peristiwa ini mendapat perhatian lebih dari pemerintah karena menimbulkan banyak kerugian baik pada bidang ekonomi, ekologi, dan sosial. Indonesia merupakan negara dengan tingkat bencana kebakaran hutan/lahan yang tinggi. Indonesia mengalami kerugian hingga mencapai Rp 209 Triliun pada tahun 2015. Akibat kerugian yang ditimbulkan dibutuhkan penanggulangan secara dini, salah satu langkah yang dapat dilakukan dengan mengelompokkan wilayah-wilayah dengan potensi kebakaran hutan/lahan dengan memanfaatkan data titik panas. Kebakaran hutan/lahan ditandai dengan dideteksinya titik – titik api oleh satelit yang diindikasikan sebagai titik panas. Dalam penelitian ini menggunakan data titik panas dengan parameter *latitude*, *longitude*, *brightness*, *frp* (*fire radiative power*), dan *confidence* dengan menggunakan metode *K-Medoids*. Metode *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. Hasil dari penelitian ini menunjukkan penggunaan metode *K-Medoids* dapat digunakan untuk proses *clustering* data titik panas dengan hasil *Silhouette Coefficient* terbaik sebesar 0.56745 pada penggunaan 2 *cluster* dengan menggunakan jumlah data sebesar 7352 data. Hasil analisis terhadap hasil *clustering* menunjukkan bahwa dengan penggunaan 2 *cluster* menghasilkan kelompok data dengan potensi yang dimiliki yaitu *cluster* 1 termasuk dalam potensi tinggi dengan hasil rata-rata *brightness* sebesar 344.47<sup>0</sup>K dengan rata-rata *confidence* 87.08% dan *cluster* 2 masuk dalam potensi sedang dengan hasil rata-rata *brightness* sebesar 318.80<sup>0</sup>K dengan rata-rata *confidence* sebesar 58,73%.

**Kata Kunci:** *Clustering, K-Medoids, MODIS, Silhouette Coefficient, titik panas*

### Abstract

Forest / land wildfire is one of the disasters that occur every year in some countries in the world. This incident got more attention from the government because it caused many losses both in the economic, ecological, and social. Indonesia is a country with a high rate of forest / land wildfire disasters. Indonesia suffered losses of up to Rp 209 trillion by 2015. As a result of losses incurred an early prevention is needed, which one can be done by grouping areas with potential forest fires by utilizing hotspot data. Forest wildfires are marked by the detection of fire spots by satellites indicated as hot spots. This research uses hotspot data with parameter of *latitude*, *longitude*, *brightness*, *frp* (*fire radiative power*), and *confidence* by using *K-Medoids* method. *K-Medoids* method is a clustering method that serves to split the dataset into groups. The advantages of this method is able to resolve the weakness of *K-Means* method that is sensitive to outlier. The result of this research shows that the use of *K-Medoids* method can be used for the process of hot spot data clustering with the best silhouette coefficient in amount of 0.56745 on the use of 2 clusters by using 7352 data. The results of the clustering analysis showed that using 2 clusters resulted in a group of data with the potential of high potential with an average brightness of 344.470K with average confidence of 87.18% and medium potential with average brightness of 318.800K with Average confidence of 58.73%.

**Keywords:** *Clustering, K-Medoids, MODIS, Silhouette Coefficient, hotspot*

## 1. PENDAHULUAN

Kebakaran hutan dan lahan merupakan salah satu peristiwa yang menjadi perhatian masyarakat nasional maupun internasional. Berdasarkan Peraturan Menteri Kehutanan Nomor : P.12/Menhut-II/2009, Kebakaran hutan dan lahan adalah salah satu masalah lingkungan utama yang menyebabkan permasalahan pada perekonomian, ekologi dan sosial. Kebakaran hutan/lahan sering terjadi akibat penggunaan api dalam pembukaan hutan/lahan untuk difungsikan sebagai Hutan Tanaman Industri (HTI), perkebunan kelapa sawit, pertanian serta pembalakan liar [FWI, 2013].

Peristiwa kebakaran hutan/lahan dapat dipantau dengan memanfaatkan teknologi penginderaan jarak jauh, salah satunya yaitu satelit Terra/Aqua dengan bantuan sensor MODIS milik NASA [Suwarsono, dkk., 2014]. Kebakaran hutan/lahan dideteksi oleh satelit sebagai titik-titik api yang kemudian diindikasikan sebagai titik panas. Titik panas merupakan permukaan bumi yang memiliki suhu yang relatif lebih tinggi dari permukaan yang lainnya [Lapan, 2014].

Asia Tenggara merupakan kawasan yang termasuk dalam daftar wilayah - wilayah yang terdeteksi terhadap kemunculan titik panas, salah satunya adalah Indonesia [Tempo, 2015]. Pada tahun 2015, tercatat 259 titik panas dengan selang kepercayaan 80% yang dideteksi oleh satelit Terra/Aqua sebagai kebakaran hutan. Indonesia mengalami kerugian mencapai Rp 209 Triliun dan terjangkitnya berbagai penyakit ISPA, mata dan kulit [Kompas, 2016].

Dampak kebakaran hutan/lahan mendorong berbagai pihak untuk melakukan pencegahan secara dini. Mengingat bencana kebakaran hutan/lahan dapat diperparah saat memasuki musim-musim kemarau. Salah satu langkah awal yang dapat dilakukan yaitu dengan memprakirakan wilayah-wilayah yang berpotensi terhadap kebakaran hutan/lahan. Pemanfaatan data titik panas dapat dilakukan untuk proses pengelompokan sesuai dengan informasi yang dimiliki oleh data, sehingga dapat diketahui potensi kebakaran hutan/lahan yang dimiliki. Proses pengelompokan dapat dilakukan dengan mengimplementasikan metode *clustering*, salah satunya yaitu metode *K-Medoids*.

Penelitian sebelumnya telah dilakukan oleh Mohammad Bentar Cahyadarena terhadap

proses *clustering* pada kasus pendeteksian pencilan titik panas pada tahun 2015. Penelitian yang dilakukan menggunakan metode *K-Medoids* atau *PAM* dan *CLARA*. Penelitian yang berjudul “Deteksi Pencilan pada Data Titik Panas Menggunakan *Clustering* berbasis *Medoids*” menyebutkan data yang digunakan adalah data titik panas tahun 2001 hingga 2012 dengan atribut *latitude*, *longitude*, *brightness*, *acq\_date*, *acq\_time*, *bright\_t31* dan *frp* menghasilkan nilai k terbaik terhadap pencilan. Metode *K-Medoids* atau *PAM* menghasilkan pencilan pada k=17, dan metode *CLARA* menghasilkan pencilan pada k=19.

Berdasarkan permasalahan dan paparan penelitian sebelumnya, penulis mengajukan penelitian berjudul “Implementasi Metode *K-Medoids Clustering* Untuk Pengelompokan Data Potensi Kebakaran Hutan/Lahan Berdasarkan Persebaran Titik Panas (*Hotspot*)” sebagai solusi agar memudahkan dalam mengetahui wilayah-wilayah yang memiliki potensi tinggi terjadinya kebakaran hutan/lahan sehingga dapat dilakukan penanggulangan secara dini.

## 2. TINJAUAN PUSTAKA

### 2.1 Kebakaran Hutan/Lahan

Kebakaran hutan didefinisikan sebagai pembakaran yang tidak tertahan dan menyebar secara bebas serta mengkonsumsi bahan bakar yang ada di hutan [Brown A.A, 1973]. Dilihat dari bahan bakar dan perjalanan api, kebakaran hutan/lahan digolongkan dalam 3 tipe [davis, 1959], yaitu :

- Kebakaran Bawah (*Ground Fire*)  
Kebakaran yang membakar bahan-bahan organik yang membusuk yang terdapat pada permukaan.
- Kebakaran Permukaan (*Surface Fire*)  
Kebakaran yang membakar serasah dan tumbuhan yang terdapat pada lantai hutan.
- Kebakaran Atas (*Crown Fire*)  
Kebakaran karena adanya kebakaran permukaan yang menjalar ke arah atas pohon.

### 2.2 Moderate Resolution Imaging Spectroradiometer (MODIS)

MODIS merupakan sensor pada satelit Terra (EOS AM-1) dan satelit Aqua (EOS PM-1) dari NASA yang berfungsi untuk melihat seluruh permukaan bumi seperti suhu permukaan tanah dan laut, indeks vegetasi,

tutupan permukaan lahan, kebakaran hutan/lahan, gunung berapi, awan, aerosol, profil suhu dan uap air [ESA, 2015]. MODIS memiliki informasi terkait dengan titik panas yang mengindikasikan adanya kebakaran. Deteksi kebakaran dilakukan dengan menggunakan algoritma kontekstual dimana memanfaatkan emisi dari radiasi inframerah dari suatu kebakaran. MODIS memanfaatkan *band* 21/22 dan *band* 31 untuk mendeteksi lokasi dan distribusi titik panas.

### 2.3 Titik Panas (*Hotspot*)

Titik panas merupakan suatu indikator penyebab terjadinya kebakaran hutan, dimana suhu permukaan relatif lebih tinggi dibandingkan dengan suhu disekitarnya [LAPAN, 2014]. Suhu titik panas dapat dihasilkan berdasarkan nilai suhu kecerahan (*brightness temperature*) yang ditangkap oleh sensor pada satelit. Satelit akan mengidentifikasi suatu piksel sebagai hotspot apabila suatu piksel memiliki suhu diatas ambang batas. Ambang batas suhu yang diidentifikasi sebagai titik panas yaitu apabila mencapai  $>330^{\circ}\text{K}$  [Roswintarti, 2016]. Data titik panas (*hotspot*) memiliki atribut untuk mendeteksi adanya kebakaran hutan. Atribut pada MODIS yang digunakan di dalam penelitian yaitu [NASA, 2015]:

- Latitude*, garis lintang dari khatulistiwa ( $0^{\circ}$ ) ke kutub selatan, atau khatulistiwa ke kutub utara.
- Longitude*, garis bujur dimana merupakan garis *horizontal* dari garis khatulistiwa.
- Confidence*, tingkat kepercayaan kualitas hotspot dengan skala 0% sampai 100%. *Confidence* menunjukkan tingkat kepercayaan dimana *hotspot* yang dipantau oleh satelit merupakan kejadian kebakaran yang sebenarnya terjadi di lapangan. Semakin tinggi selang kepercayaan, maka semakin tinggi pula potensi bahwa *hotspot* benar-benar merupakan kebakaran hutan atau lahan yang terjadi. Berikut tingkatan *confidence* dalam menggambarkan informasi hotspot ditunjukkan pada Tabel 1.
- Brightness Temperature*, ukuran deskriptif dari pancaran radiasi dalam bentuk suhu yang dipancarkan ke atas pada bagian atmosfer bumi. *Brightness temperature* merupakan fitur dasar pada citra penginderaan jarak jauh yang di deteksi dari lokasi tertentu dan diukur dalam satuan ukur kelvin.

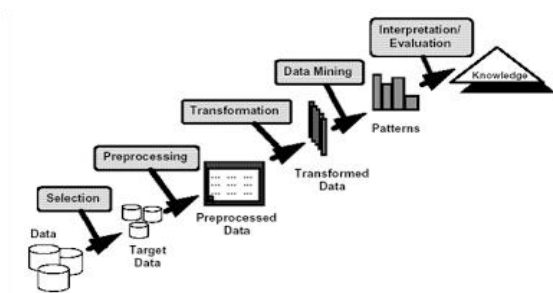
- FRP* (*Fire Radiative Power*), menggambarkan kekuatan radiasi piksel api yang kemudian diintegrasikan dalam MW (*MegaWatts*). *FRP* memberikan informasi pada output radiasi panas dari api yang terdeteksi. Jumlah radiasi energi panas dibebaskan per satuan waktu (*FRP*) dimana diduga berkaitan dengan tingkat bahan bakar yang dikonsumsi.

**Tabel 1.** Makna Selang Kepercayaan Dalam Informasi Hotspot

Tingkat kepercayaan ( C )	Kelas	Tindakan
$0\% \leq C < 30\%$	Rendah	Perlu diperhatikan
$30\% \leq C < 80\%$	Nominal	Waspada
$80\% \leq C \leq 100\%$	Tinggi	Segera penanggulangan

### 2.4 Data Mining

Data mining adalah proses analitik yang dirancang untuk memeriksa sejumlah data yang besar dalam mencari suatu pengetahuan tersembunyi yang berharga dan konsisten [Florin G, 2011]. Tujuan dari data mining yaitu mencari trend atau pola yang diinginkan dalam database besar untuk membantu dalam pengambilan keputusan pada waktu yang akan datang. Berikut Gambar 2.1 menunjukkan proses data mining dalam penemuan pengetahuan berulang dalam database.



**Gambar 1.** Proses Data Mining Dalam Penemuan Pengetahuan Dalam Database

### 2.5 Clustering

*Clustering* merupakan suatu proses pengelompokan record, observasi, atau mengelompokkan kelas yang memiliki kesamaan objek [Larose, 2015]. Perbedaan klustering dengan klasifikasi yaitu tidak adanya variabel target dalam melakukan pengelompokan pada proses *clustering*. *Clustering* sering dilakukan sebagai langkah

awal dalam proses data mining. Terdapat banyak algoritma klastering yang telah digunakan oleh peneliti sebelumnya seperti *K-Means*, *Improved K-Means*, *Fuzzy C-Means*, *DBSCAN*, *K-Medoids (PAM)*, *CLARANS* dan *Fuzzy Subtractive*. Setiap algoritma memiliki kelebihan dan kekurangan masing-masing, namun prinsip algoritma sama, yaitu mengelompokkan data sesuai dengan karakteristik dan mengukur jarak kemiripan antar data dalam satu kelompok.

## 2.6 Normalisasi Data

Normalisasi adalah proses transformasi untuk merubah nilai data. Normalisasi digunakan untuk menyamakan skala atribut data kedalam *range* yang spesifik yang lebih kecil seperti -1 sampai 1 atau 0 sampai 1. *Min-Max Normalization* merupakan teknik normalisasi dengan melakukan transformasi linier pada atribut data asli untuk menghasilkan range nilai yang sama [Junaedi, dkk., 2011]. *Min-Max Normalization* menetapkan sebuah value  $v$  dari atribut  $A$  menjadi  $v'$  kedalam *range* [ $new\_min_A$ ,  $new\_max_A$ ] dengan Persamaan 1.

$$v' = \frac{v - min_A}{max_A - min_A} (new\_max_A - new\_min_A) + new\_min_A \quad (1)$$

## 2.7 Algoritma K-Medoids

*K-Medoids* atau *Partitioning Around Medoids (PAM)* adalah algoritma clustering yang mirip dengan *K-Means*. Perbedaan dari kedua algoritma ini yaitu algoritma *K-Medoids* atau *PAM* menggunakan objek sebagai perwakilan (*medoid*) sebagai pusat *cluster* untuk setiap cluster, sedangkan *K-Means* menggunakan nilai rata-rata (*mean*) sebagai pusat *cluster* [Kaur, dkk., 2014]. Algoritma *K-Medoids* memiliki kelebihan untuk mengatasi kelemahan pada algoritma *K-Means* yang *sensitive* terhadap *noise* dan *outlier*, dimana objek dengan nilai yang besar yang memungkinkan menyimpang dari distribusi data. Kelebihan lainnya yaitu hasil proses *clustering* tidak bergantung pada urutan masuk dataset [Furqon, dkk., 2015]. Langkah-langkah algoritma *K-Medoids*:

1. Inisialisasi pusat *cluster* sebanyak  $k$  (jumlah *cluster*)
2. Alokasikan setiap data (objek) ke *cluster* terdekat menggunakan persamaan ukuran jarak *Euclidian Distance* dengan persamaan:

$$d(x, y) = ||x - y||$$

$$= \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad ; 1, 2, 3, \dots, n \quad (2)$$

3. Pilih secara acak objek pada masing-masing *cluster* sebagai kandidat *medoid* baru.
4. Hitung jarak setiap objek yang berada pada masing-masing *cluster* dengan kandidat *medoid* baru.
5. Hitung total simpangan ( $S$ ) dengan menghitung nilai total *distance* baru – total *distance* lama. Jika  $S < 0$ , maka tukar objek dengan data cluster untuk membentuk sekumpulan  $k$  objek baru sebagai *medoid*.
6. Ulangi langkah 3 sampai 5 hingga tidak terjadi perubahan *medoid*, sehingga didapatkan *cluster* beserta anggota *cluster* masing-masing.

## 2.8 Silhouette Coefficient

*Silhouette coefficient* merupakan metode yang digunakan untuk melihat kualitas dan kekuatan dari cluster. Metode *silhouette coefficient* merupakan gabungan dari dua metode yaitu metode *cohesion* yang berfungsi untuk mengukur seberapa dekat relasi antara objek dalam sebuah cluster, dan metode *separation* yang berfungsi untuk mengukur seberapa jauh sebuah *cluster* terpisah dengan *cluster* lain. Tahapan perhitungan *silhouette coefficient* [Handoyo, dkk., 2014]:

1. Hitung rata-rata jarak objek dengan semua objek lain yang berada di dalam satu cluster dengan persamaan :

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

2. Hitung rata-rata jarak objek dengan semua objek lain yang berada pada cluster lain, kemudian ambil nilai paling minimum dengan persamaan :

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

3. Hitung nilai *silhouette coefficient* dengan persamaan :

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (5)$$

Nilai hasil *silhouette coefficient* terletak pada kisaran nilai -1 hingga 1. Semakin nilai *silhouette coefficient* mendekati nilai 1, maka semakin baik pengelompokan data dalam satu *cluster*. Sebaliknya jika nilai *silhouette coefficient* mendekati nilai -1, maka semakin



buruk pengelompokan data di dalam satu *cluster*.

### 3. METODOLOGI DAN PERANCANGAN

Metodologi penelitian menjelaskan langkah-langkah yang akan digunakan serta perancangan dalam melakukan implementasi metode *K-Medoids clustering* untuk pengelompokan data potensi kebakaran hutan/lahan berdasarkan persebaran titik panas. Tahapan metodologi penelitian ditunjukkan pada Gambar 2.



Gambar 2. Tahapan Metodologi Penelitian

#### 3.1 Pengumpulan Data

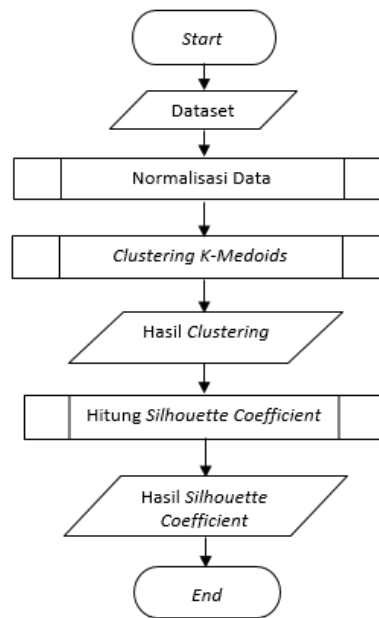
Data yang dibutuhkan dalam penelitian ini adalah data titik panas yang diperoleh dari database sebuah situs resmi dari *NASA LANCE – FIRM MODIS Active Fire*. Data yang digunakan berisikan informasi berupa nilai *brightness temperature*, *frp*, *latitude*, *longitude* dan *confidence*. Data yang dikumpulkan berupa data titik panas wilayah Asia Tenggara tahun 2000 hingga 2015 berjumlah 1.048.576 *record*. Namun pada penelitian ini, data titik panas yang digunakan sebagai data sampel sebanyak 7352 data periode bulan September tahun 2015.

#### 3.2 Analisis Kebutuhan

Analisa kebutuhan dilakukan dengan menentukan kebutuhan apa saja yang dibutuhkan dalam pengimplementasian metode *K-Medoids clustering* untuk pengelompokan data potensi kebakaran hutan/lahan berdasarkan persebaran titik panas (*hotspot*).

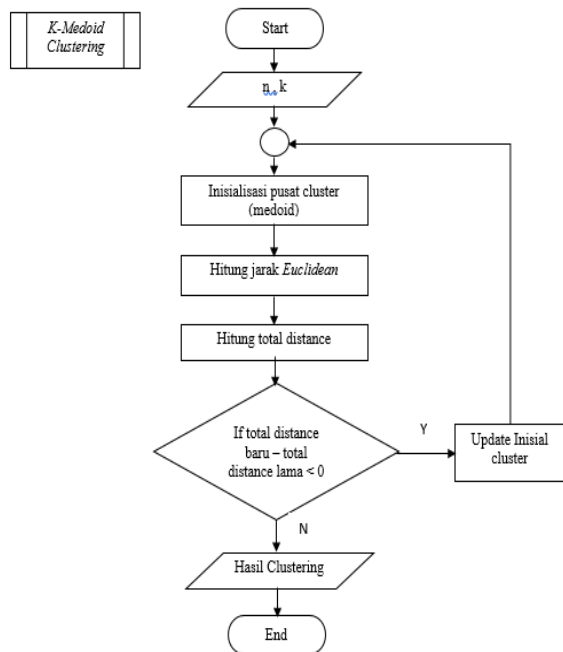
### 3.3 Perancangan

Perancangan sistem dibuat untuk rancangan langkah kerja dari sistem secara keseluruhan untuk mempermudah implementasi dan pengujian. Langkah kerja dalam sistem disesuaikan dengan rancangan perangkat lunak yang terdapat pada Gambar 3.



Gambar 3. Perancangan Perangkat Lunak

Proses pengklasteran akan dilakukan dengan menggunakan algoritma *K-Medoids*. Dimulai dari memasukkan jumlah data yang akan dibutuhkan dan jumlah *cluster* yang ditentukan. Selanjutnya dilakukan proses dari seluruh perhitungan *K-Medoids*. Hasil *cluster* akan ditampilkan pada akhir proses ketika proses perhitungan telah selesai. Diagram alir algoritma *K-Medoids* dapat dilihat pada Gambar 4.



**Gambar 4.** Flowchart Metode *K-Medoids Clustering*

### 3.4 Proses Perhitungan

Dalam proses perhitungan pada metode *K-Medoids* berikut langkah-langkahnya :

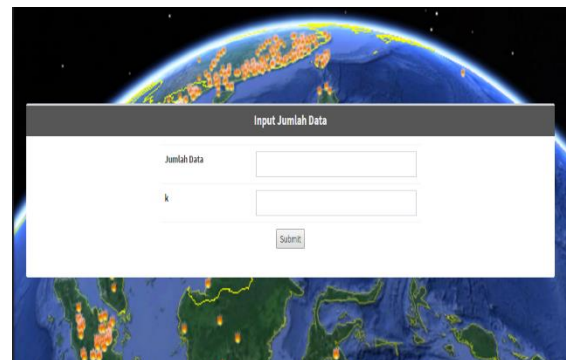
1. Melakukan normalisasi data yang akan digunakan dalam proses perhitungan *K-Medoids*. Perhitungan normalisasi data menggunakan persamaan (1)
2. Menginisialisasikan pusat *cluster* secara acak, kemudian menghitung jarak data (objek) dengan pusat *cluster* dengan *Euclidean Distance*. Perhitungan *Euclidean Distance* menggunakan persamaan (2).
3. Menghitung total *distance* dari keseluruhan data di dalam *cluster*.
4. Menginisialisasikan pusat *cluster* baru secara acak kemudian menghitung jarak data (objek) dengan pusat *cluster* dengan *Euclidean Distance* menggunakan persamaan (2).
5. Menghitung selisih total *distance* dengan mengurangi total *distance* baru – total *distance* lama.
6. Setelah mendapatkan hasil *cluster* akhir, menghitung nilai *average dissimilarity* ( $a_i$ ) untuk evaluasi dengan metode *silhouette coefficient*. Perhitungan  $a(i)$  menggunakan persamaan (3)
7. Menghitung nilai *lowest average dissimilarity* ( $b_i$ ). Perhitungan  $b(i)$  menggunakan persamaan (4).
8. Menghitung nilai *silhouette coefficient* ( $s_i$ ). Perhitungan  $s(i)$  menggunakan persamaan

(5). Hasil nilai *silhouette coefficient* digunakan untuk mengetahui kualitas dari *cluster*.

## 4. IMPLEMENTASI

### 4.1 Antarmuka Masukan

Pada halaman masukkan data dapat dilihat pada Gambar 5., dimana terdapat dua tipe masukan yaitu jumlah data dan jumlah *cluster*.



**Gambar 5.** Antarmuka Masukan

### 4.2 Antarmuka Proses Algoritma

Pada halaman proses algoritma dapat dilihat pada Gambar 6. Terdapat 6 menu pada proses algoritma yang ditampilkan. Data yaitu tampilan hasil load data asli. Normalisasi menampilkan hasil dari normalisasi data. *K-Medoids* menampilkan iterasi terhadap perubahan medoid. Hasil *clustering* menampilkan *cluster* dengan anggota yang dimiliki. *Silhouette coefficient* menampilkan hasil nilai evaluasi terhadap kualitas *cluster*. Hasil Analisis menampilkan hasil perhitungan untuk penentuan potensi kebakaran dari data didalam *cluster*.

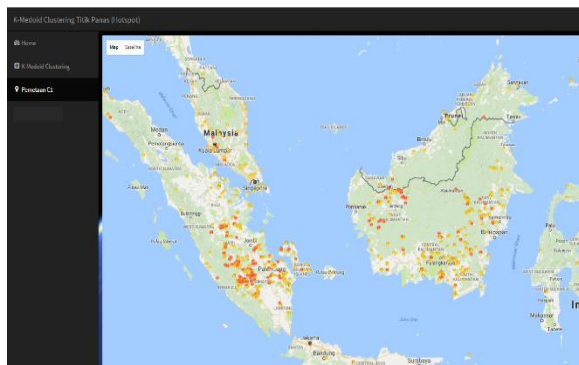


**Gambar 6.** Proses Algoritma

### 4.3 Antarmuka Hasil Pemetaan

Pada halaman hasil pemetaan dapat dilihat pada Gambar 7., dimana merupakan tampilan

titik-titik wilayah dengan potensi kebakaran hutan/lahan dari hasil analisis *cluster*.



**Gambar 7.** Hasil Pemetaan

## 5. PENGUJIAN DAN ANALISIS

Pengujian yang dilakukan pada penelitian ini adalah pengujian terhadap pengaruh jumlah nilai  $k$  dan pengujian terhadap pengaruh jumlah data. Pengujian akan dilakukan berulang sebanyak 5 kali proses percobaan dan diambil nilai rata-rata untuk digunakan sebagai hasil akhir pengujian.

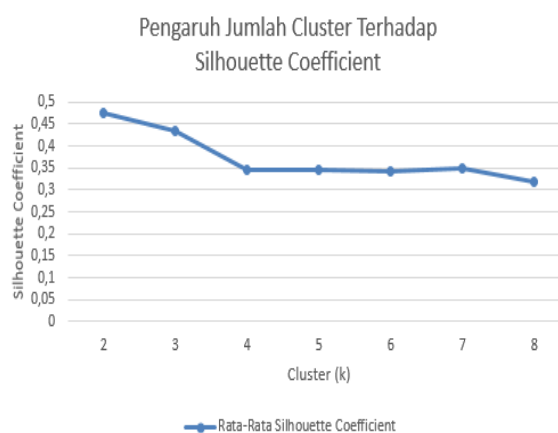
### 5.1 Pengujian Pengaruh Jumlah Cluster ( $k$ )

Pengujian pengaruh jumlah *cluster* dilakukan dengan menggunakan 500 data uji untuk mengetahui jumlah *cluster* terbaik dengan melihat dari nilai *Silhouette Coefficient* yang dihasilkan. Nilai *cluster* ( $k$ ) yang akan diujikan adalah nilai  $k$  2 sampai 8. Hasil pengujian dapat dilihat pada Tabel 2.

**Tabel 2.** Hasil Pengujian Jumlah Cluster

Data Uji	Nilai $k$	Rata-Rata Silhouette Coefficient
500	2	0.474783859
	3	0.433570931
	4	0.345881766
	5	0.345055475
	6	0.342431368
	7	0.347567261
	8	0.319028458

Pada hasil proses pengujian, kualitas kluster terbaik diperoleh pada penggunaan nilai  $k=2$  dengan nilai rata-rata *Silhouette Coefficient* tertinggi 0.474783859. Grafik pengaruh nilai  $k$  ditunjukkan pada Gambar 8.



**Gambar 8.** Grafik Pengaruh Jumlah Cluster Terhadap *Silhouette Coefficient*

Hasil pada grafik terlihat semakin banyak jumlah penggunaan nilai  $k$ , maka nilai *Silhouette Coefficient* yang dihasilkan relatif semakin rendah. Hal ini dikarenakan semakin banyaknya ketetanggaan terdekat yang digunakan pada proses *clustering*, dimana ketetanggaan yang seharusnya tidak dipertimbangkan menjadi dipertimbangkan dan masuk kedalam proses perhitungan. Sehingga jarak antar objek di dalam satu *cluster* semakin besar dan jarak *cluster* dengan *cluster* lain semakin kecil.

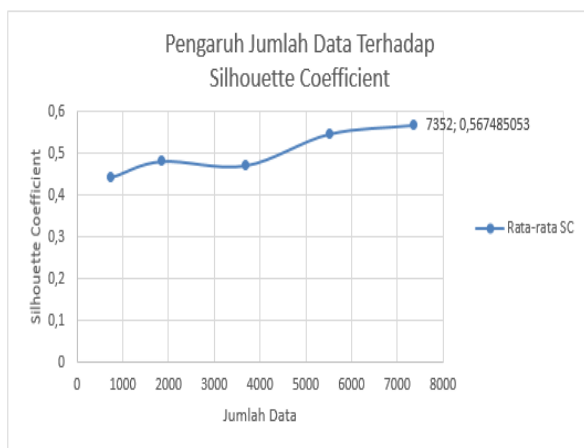
### 5.2 Pengujian Pengaruh Jumlah Data

Pengujian pengaruh jumlah data dilakukan untuk mengetahui pengaruh banyaknya data terhadap hasil dari pengujian sebelumnya. Jumlah data yang digunakan pada pengujian ini yaitu 10%, 25%, 50%, 75% dan 100% dari data sampel yang digunakan pada penelitian ini. Hasil pengujian ditunjukkan pada Tabel 3.

**Tabel 3.** Hasil Pengujian Jumlah Data

Jumlah Cluster	Jumlah Data	Rata-Rata Silhouette Coefficient
2	735	0.442859767
	1838	0.481231816
	3676	0.470980279
	5514	0.546385200
	7352	0.567485053

Pada hasil proses pengujian, jumlah data terbaik diperoleh pada penggunaan 7352 data dengan nilai rata-rata *Silhouette Coefficient* tertinggi 0.567485053. Grafik pengaruh jumlah data ditunjukkan pada Gambar 9.



**Gambar 9.** Grafik Pengaruh Jumlah Data Terhadap *Silhouette Coefficient*

Hasil pada grafik menunjukkan semakin banyak data yang digunakan maka nilai *Silhouette Coefficient* yang dihasilkan relatif semakin tinggi. Hal ini menunjukkan bahwa jumlah data berpengaruh terhadap kualitas dari *cluster*. Ketika jumlah data yang digunakan semakin sedikit, maka nilai *Silhouette Coefficient* yang dihasilkan semakin rendah dikarenakan semakin besarnya jarak intra *cluster* dan semakin kecilnya jarak intra *cluster* dimana menggambarkan tingkat kemiripan data dalam satu *cluster* yang besar dan tingkat kemiripan data antar *cluster* yang kecil.

### 5.3 Analisis Penentuan Potensi Kebakaran Hutan/Lahan

Penentuan potensi kebakaran hutan/lahan dilakukan dengan menganalisis hasil *clustering* yang telah diuji sebelumnya. Dimana hasil pengujian sebelumnya menunjukkan jumlah *cluster* terbaik sebanyak 2 *cluster* dengan jumlah data 7352. Berdasarkan hasil tersebut, akan dilakukan analisis untuk mengetahui kelompok dengan potensi kebakaran tertinggi. Dalam melakukan analisis tingkat potensi kebakaran dapat dilakukan dengan menganalisis nilai dari *brightness temperature* dimana nilai *brightness* dengan nilai lebih dari 330<sup>0</sup>K termasuk dalam potensi kebakaran dengan *high confidence fire* yang berarti mengidentifikasi api dengan tingkat keyakinan tinggi untuk terjadinya kebakaran atau pembukaan lahan. Hasil penentuan potensi kebakaran hutan/lahan juga didukung dengan melihat nilai *confidence* yang telah dijelaskan pada Tabel 1. Hasil analisis ditunjukkan pada Tabel 4.

**Tabel 4.** Hasil Analisis Penentuan Potensi

	Rata-Rata Brightness ( <sup>0</sup> K)	Rata-Rata Confidence (%)
Cluster 1	344.4732240	87.08852459
Cluster 2	318.8059942	58.73741398

Hasil analisis diatas menunjukkan dimana *cluster 1* menghasilkan nilai rata-rata *brightness temperature* lebih dari 330<sup>0</sup>K dan nilai rata-rata *confidence* lebih dari 80%. Berdasarkan hasil yang diperoleh, *cluster 1* termasuk ke dalam kategori *high confidence fire* dimana memiliki potensi tinggi bahwa titik-titik tersebut kebakaran yang sebenarnya dan dibutuhkan tindakan penanggulangan segera. Untuk *cluster 2* menghasilkan nilai rata-rata *brightness temperature* kurang dari 330<sup>0</sup>K dan nilai rata-rata *confidence* dibawah 80% dimana *cluster 2* dapat termasuk ke dalam kategori *nominal confidence fire* dimana memiliki potensi sedang dan dibutuhkan tindakan kewaspadaan terhadap terjadinya kebakaran.

## 6. KESIMPULAN DAN SARAN

### 6.1 Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan dapat diambil kesimpulan sebagai berikut:

1. Algoritma *K-Medoids* dapat diimplementasikan untuk proses pengelompokan data potensi kebakaran hutan/lahan berdasarkan persebaran titik panas (*hotspot*) sesuai dengan tahapan-tahapan perhitungan *K-Medoids Clustering*.
2. Kualitas *clustering* yang dihasilkan berdasarkan proses pengujian yang dilakukan didapatkan hasil nilai *Silhouette Coefficient* tertinggi yaitu 0.567485053 dengan jumlah *cluster* 2 dan jumlah data 7352. Berdasarkan hasil tersebut dapat diambil kesimpulan bahwa jumlah *cluster* dan jumlah data mempengaruhi terhadap hasil kualitas dari *cluster*.

Analisis penentuan potensi diperoleh hasil dimana potensi tinggi terhadap kebakaran hutan/lahan ditunjukkan pada *cluster 1* dengan rata-rata *brightness* diatas batas ambang yaitu 334.47<sup>0</sup>K dan rata-rata *confidence* 87.08% dengan indikasi api termasuk dalam kategori *high confidence fire* atau memiliki potensi tinggi terhadap terjadinya kebakaran yang sebenarnya sehingga dibutuhkan penanggulangan segera. Potensi sedang ditunjukkan pada



*cluster* 2 dengan rata-rata *brightness* dibawah batas ambang yaitu 318.8<sup>0</sup>K dan rata-rata *confidence* 58.73% dengan indikasi api termasuk dalam katagori *nominal confidence fire* dimana dibutuhkan kewaspadaan terjadinya kebakaran hutan/lahan.

## 6.2 Saran

Saran penulisan terkait penelitian “Implementasi Metode *K-Medoids Clustering* Untuk Pengelompokan Data Potensi Kebakaran Hutan/Lahan Berdasarkan Persebaran Titik Panas (*Hotspot*) untuk penelitian selanjutnya sebagai berikut:

1. Dianjurkan melakukan penambahan jumlah data dan parameter, dimana dalam penentuan parameter dilakukan peninjauan ulang dengan menggunakan bantuan sebuah algoritma atau dengan bantuan para ahli untuk prosedur pemilihan parameter, sehingga tingkat kualitas *cluster* yang dihasilkan lebih baik.
2. Penelitian selanjutnya dapat mengembangkan perangkat lunak dengan menggunakan metode *clustering* lain, seperti *CLARA* atau *CLARANS* untuk dapat membandingkan nilai kualitas *cluster* yang terbentuk dengan hasil dari penelitian sebelumnya.

## DAFTAR PUSTAKA

- Abhishek, P. & Purnima, S., 2013. *New Approach for K-means and K-Medoids Algoritma*. International journal of Computer Application Technology and Research. Vol. 2. India.
- Brown, Arthur A. & Kenneth P. Davis, 1973. *Forest Fire Control and Use*. 2nd ed. New York, NY : McGraw-Hill Book Company.
- Chrisnanto, Yulison H. & Abdillah, Gunawan., 2015. *Penerapan Algoritma Partitioning Around Medoids (PAM) Clustering Untuk Melihat Gambaran Umum Kemampuan Akademik Mahasiswa*. [pdf]. ISSN: 2089-9815. 28 Maret 2015. SENTIKA 2015.
- Cahyadarena, Bentar M., 2015. *Deteksi Pencilan pada Data Titik Panas Menggunakan Clustering berbasis Medoids*. S1. [pdf] Institut Pertanian Bogor. Bogor.
- Eko Prasetyo, 2012. *Data Mining : Konsep dan Aplikasi Menggunakan Matlab*. Perpustakaan Nasional Ed. I. Yogyakarta : ANDI.
- Flowrensia, Yanne, 2010. *Perbandingan Penggerombolan K-Means dan K-Medoids Pada Data Yang Mengandung Pencilan*. S1. [pdf] Institut Pertanian Bogor. Bogor.
- FWI., 2013. *Kebakaran Hutan dan Lahan : Bukan Bencana Alam, Awal Bencana Ekologi*. [online] Tersedia di : < <http://fwi.or.id/publikasi/kebakaran-hutan-dan-lahan-bukan-bencana-alam-awal-bencana-ekologi>> [Diakses 17 Januari 2017].
- Gorunescu, Florin., 2011. *Data Mining : Concepts, Models and Techniques*. Intelligent System Reference Library Vol.12. Berlin : Springer-Verlag.
- Handayani, Tri., Santoso, Joko Albertus., Dwiandiyanta, Yudi., 2014. *Pemanfaatan Data Terra Modis Untuk Identifikasi Titik Api Pada Kebakaran Hutan Gambut (Studi Kasus Kota Dumai Provinsi Riau)*. SENTIKA 2014. ISSN: 2089-9813. Yogyakarta. 15 Maret 2014.
- Handoyo, Rendy., M, R. Rumani., Nasution, Surya Michrandi., 2014. *Perbandingan Metode Clustering Menggunakan Metode Single Linkage dan K-Means Pada Pengelompokan Dokumen*. ISSN. 1412-0100 Vol. 15, No. 2. Oktober 2014. Page 1-10.
- Hermawan, Astuti Fajar., 2013. *Data Mining*. Perpustakaan Nasional Ed. I. Yogyakarta : ANDI.
- Indah, P., Rizaldi, B., Ardiyansyah, M., Agus, B., Lailan, S., & Yenni, V., 2012. *Analisis Hubungan Kode-Kode SPBK (Sistem Peringkat Bahaya Kebakaran) dan Hotspot dengan Kebakaran Hutan dan Lahan di Kalimantan Tengah*. [pdf] Institut Pertanian Bogor. Bogor.
- Junaedi, Hartanto., Herman, Budianto., dkk., 2011. *Data Transformation Pada Data Mining*. Posiding Konferensi nasional “Inovasi dalam Desain dan Teknologi”. IDeaTech 2011.
- Kaur, Noor K., Kaur, Usvir., & Singh, Dr.

- Dheerendra., 2014. *K-Medoids Clustering Algorithm – A Review*. [pdf] International Journal of Computer Application and Technology (IJCAT). ISSN. 2349-1841 Vol. 1, Issue 1. April 2014.
- Kodinariya, Trupti M., & Makwana, Dr. Prashant R., 2013. *Review on determining number of Cluster in K-Means Clustering*. [pdf] International Journal of Advance Research in Computer Science and Management Studies. ISSN. 2321-7782 Vol. 1, Issue 6. November 2013. Page 90-95.
- NASA., 2015. [pdf] Tersedia di : <<https://earthdata.nasa.gov/earth-observation-data/near-real-time/firms/active-fire-data> > [ Diakses 10 Februari 2016].
- Oded, Maimon., & Lior Rokach, 2010. *Data Mining and Knowledge Discovery Handbook*. [e-book] 2nd ed. New York, NY : Springer Science + Business Media.
- Roswintiarti, Dr. Orbita, 2016. *Informasi Titik Panas (Hotspot) Kebakaran Hutan/Lahan. Lembaga Penerbangan Dan Antariksa Nasional*. [pdf] Panduan Teknis – V.01. ISBN : 978-602-96532-2-5. Mei 2016.
- Rousseeuw, Petter. J, 1987. *Silhouette : A Graphical Aid to The Interpretation and Validation of Cluster Analysis*. [pdf] Journal of Computational and Applied Mathematics 20 (1987) 53-65. North-Holland.