

# Penentuan Daerah Rawan Titik Api di Provinsi Riau Menggunakan Clustering Algoritma K-Means (*Determination of Fire Point Prone Areas in Riau Province Using Clustering K-Means Algorithm*)

Sukamto<sup>1</sup>, Ibnu Daqiqil Id<sup>2</sup>, T.Rahmilia Angraini<sup>3</sup>

<sup>1,2,3</sup>Program Studi Sistem Informasi, Jurusan Ilmu Komputer, FMIPA Universitas Riau  
Kampus Binawidya Km. 12,5 Simpang Baru, Pekanbaru, 28293, Indonesia

<sup>1</sup>sukamto@lecturer.unri.ac.id

<sup>2</sup>ibnu.daqiqil@lecturer.unri.ac.id

<sup>3</sup>t.rahmilia@student.unri.ac.id

**Abstrak** – Penelitian ini membahas tentang daerah rawan titik api di Provinsi Riau. Kebakaran hutan menjadi ancaman pada hutan di Indonesia. Mengingat faktor timbulnya dan dampak yang akan ditimbulkan dari kebakaran hutan, maka sangat penting untuk mengetahui daerah yang rawan terhadap titik api. Konsep data *mining* sangat cocok diterapkan untuk mengetahui status daerah rawan titik api. Dalam penelitian ini dilakukan pengelompokan data dengan menggunakan *Chebyshev Distance K-Means*. Data yang digunakan adalah data titik api di Provinsi Riau pada tahun 2016. Data dikelompokkan menjadi tiga *cluster*, yaitu 133 titik yang masuk kedalam *cluster* daerah sangat rawan titik api, 101 titik kedalam *cluster* daerah rawan titik api, dan 77 titik kedalam *cluster* daerah yang tidak rawan terhadap titik api, dengan nilai DBI (*Davies Bouldin Index*) 0,361 menandakan bahwa pengklasteran *Chebyshev K-Means* sebanyak 3 *cluster* sudah optimal. Hasil *clustering* divisualisasikan dengan *Google Maps Api*.

**Kata Kunci** – *clustering, hotspot, k-means*.

**Abstract**—*This research discusses about the determination of fire point prone areas in Riau Province. Fire can be a particularly destructive threat to forests. Due to its impact, it is notable to detect the potential hotspots area beforehand. Data mining concept is considerably suitable to be applied on this innovation. During this attempt, the obtained information is based from Riau Province hotspots data (2016) and grouped by using Chebyshev Distance K-Means, which resulted three clusters. The data were grouped into three clusters, namely 133 points that were included in the cluster area that were very hotspots, 101 points into clusters of fire-prone areas, and 77 points into clusters of regions that were not prone to hotspots, so with a DBI (Davies Bouldin Index) value of 0.361 indicating that the cluster of Chebyshev K-Means as many as 3 clusters is optimal. The outcome then are visualized using Google Maps Api.*

**Keywords**— *clustering, hotspot, k-means*

## I. PENDAHULUAN

Salah satu daerah yang sering menghadapi ancaman bencana kebakaran hutan adalah Provinsi Riau. Mengingat faktor timbulnya kebakaran hutan yang terjadi di Indonesia dan terkhususnya di wilayah Provinsi Riau, serta dampak yang akan ditimbulkan dari kebakaran hutan tersebut, maka sangatlah penting untuk mengetahui daerah-daerah yang rawan terhadap titik api, guna pencegahan kebakaran hutan sejak dini.

Data *mining* merupakan metode pengolahan data berskala besar, oleh karena itu data *mining* memiliki peranan penting dalam berbagai bidang. Salah satu kajian data *mining* membahas tentang *clustering*. Pada dasarnya *clustering* merupakan suatu metode untuk mencari dan mengelompokkan data yang memiliki kemiripan karakteristik antara satu data dengan data yang lain. *Clustering* merupakan salah satu metode data *mining* yang bersifat tanpa arahan (*unsupervised*), maksudnya metode ini diterapkan tanpa adanya latihan (*training*) dan tanpa ada guru (*teacher*) serta tidak memerlukan target output.

*K-Means* merupakan salah satu metode data *clustering non hirarki* yang berusaha mempartisi data yang ada ke dalam bentuk satu atau lebih *cluster*/kelompok. Metode ini mempartisi data ke dalam *cluster* sehingga data yang memiliki karakteristik yang sama dikelompokkan ke dalam satu *cluster* yang sama dan data yang mempunyai karakteristik yang berbeda dikelompokkan ke dalam kelompok yang lain. Adapun tujuan dari data *clustering* ini adalah untuk meminimalisasikan *objective function* yang diset dalam proses *clustering*, yang pada umumnya berusaha

meminimalisasikan variasi di dalam suatu *cluster* dan memaksimalkan variasi antar *cluster* [1].

Beberapa penelitian yang berhubungan dengan algoritma *K-Means* antara lain mengembangkan sistem klasterisasi menggunakan metode *K-Means* untuk menentukan posisi *access point* berdasarkan posisi pengguna, yang menyimpulkan bahwa terdapat kemudahan dalam melakukan pengelompokan dengan sistem klasterisasi [2]. Potensi kejahatan berdasarkan *clustering* dengan metode *K-Means* menyimpulkan bahwa *clustering*/pengelompokan potensi kejahatan daerah di kota Banjarbaru diproses berdasarkan perjajaran. Hal ini dilakukan agar hasil dari *clustering* potensi kejahatan daerah di kota Banjarbaru menjadi lebih spesifik [3]. Data-data berupa data karakteristik karies gigi dan kelas-kelas karies gigi berdasarkan anatomi J. V. Black. Dari data-data tersebut dapat dilakukan pengklasifikasian dengan menggunakan metode *Clustering K-Means*. Metode *Clustering K-Means* digunakan karena *K-Means* melakukan pengelompokan data dengan sistem partisi, dimana data dalam satu kelompok memiliki karakteristik yang sama satu sama lainnya, dan memiliki karekteristik berbeda dengan kelompok lainnya [4]. Algoritma *K-Means* digunakan untuk mengetahui pola pengelompokan jurusan siswa berdasarkan nilai *Distance Performance* terhadap variabel-variabel yang diajukan. Hasil klasifikasi menunjukkan bahwa variabel yang paling tinggi pengaruhnya terhadap penjurusan siswa adalah nilai rata-rata IPA dan proses klastering menghasilkan 5 kelompok klaster [5]. Pengolahan citra digital sering digunakan untuk menganalisis suatu citra satelit yang hasilnya dapat dibandingkan dengan kenampakan suatu wilayah yang sebenarnya secara langsung, yang menyimpulkan bahwa algoritma *K-Means* dapat digunakan dalam pengolahan citra digital, yaitu sebagai algoritma untuk mengklasifikasikan citra sesuai nilai pixelnya [6].

Selanjutnya menentukan IPK mahasiswa menggunakan algoritma *K-Means* yang menyimpulkan proses klasterisasi dengan algoritma *K-Means* menggunakan SQL menunjukkan hasil data yang sama untuk setiap klasternya jika dibandingkan dengan klasterisasi data dengan aplikasi diluar DBMS [7]. Pengelompokan ayat Al Quran dengan algoritma *K-Means* yang menyimpulkan bahwa pengelompokan data ayat Al Quran dalam Bahasa Indonesia dengan menggunakan algoritma *K-Means* akan menghasilkan kelompok ayat dengan kata kunci tertentu. Proses *clustering* dengan *K-Means* memiliki hasil pengujian *silhouette* pada Surat Al Fatimah bernilai positif sebesar 0,336 yang artinya data pada kelompok yang tepat [8].

Algoritma *K-Means* yang digunakan pada Sistem Temu Kembali menyimpulkan bahwa Sistem temu kembali citra berbasis konten yang diusulkan dengan menambahkan klasterisasi citra menggunakan algoritma *K-Means Clustering* yang dikombinasikan dengan *Haar Wavelet Transform* dan *progressive retrieval strategy* dapat menambah kecepatan proses temu kembali citra sampai 2,7 kali lipat dan tingkat akurasi (*precision*) 1,4 kali lipat lebih akurat dari pada metode temu kembali citra sebelumnya [9]. Untuk menentukan minat siswa menggunakan algoritma *K-Means* yang menyimpulkan bahwa metode *clustering* dengan algoritma *K-Means* dapat digunakan untuk mengelompokkan siswa sesuai dengan minatnya pada mata pelajaran produktif dikelas XII berdasarkan nilai rapor saat mereka kelas X [10].

Dengan adanya penelitian-penelitian tersebut, peneliti akan melakukan penelitian khususnya algoritma *K-Means* untuk penentuan daerah rawan api di wilayah Provinsi Riau, guna pencegahan kebakaran hutan sejak dini, yang meliputi daerah-daerah sangat rawan api, daerah rawan api, dan daerah tidak rawan api.

## II. METODE PENELITIAN

### A. Pengumpulan Data

Proses pengumpulan data dilakukan dengan cara mengambil data titik api yang tersebar di daerah Provinsi Riau.

### B. Preprocessing Data (Cleaning Data)

Data yang diperoleh masih ada yang mengandung *missing value* (memiliki keterangan yang kurang lengkap), untuk itu dilakukan *cleaning* dan *selection* data. *Preprocessing* adalah penghapusan data *missing value*, normaliasi dan penghapusan data yang tidak digunakan dalam perhitungan *clustering* dengan *K-Means*, dengan normalisasi *Min-Max* dengan persamaan (1) [11], yaitu :

$$X^* = \frac{X - \min(X)}{\max(X) - \min(X)} \dots\dots\dots (1)$$

dimana :

- $X$  : data yang akan dinormalisasi
- $\min(X)$  : nilai minimal dari range data yang akan dinormalisasi
- $\max(X)$  : nilai maksimal dari range data yang akan di normalisasi

### C. Proses Clustering

Merupakan tahap pengolahan *dataset* yang dimodelkan dengan algoritma *K-Means* melalui perhitungan sehingga menghasilkan pengelompokan data. Adapun algoritma *K-Means* adalah:

- 1) Tentukan jumlah *cluster*.
- 2) Tentukan titik *centroid k* (pusat kluster) secara acak.
- 3) Hitung jarak setiap titik ke pusat *cluster*, jarak antar satu data dengan satu *cluster* akan menentukan data tersebut masuk dalam *cluster* mana. Perhitungan jarak dengan persamaan (2) [12], yaitu:

$$d_{ij} = \max_k |x_{ik} - x_{jk}| \dots\dots\dots (2)$$

dimana :

- $d_{ij}$  : jarak *chebyshev*, jarak antara *cluster i* dan *j*
- $x_{ik}$  : *cluster ke-i*
- $x_{jk}$  : *cluster centroid ke-j*

- 4) Hitung pusat *cluster* yang baru dengan keanggotaan *cluster* yang sekarang, dengan cara mencari rata-rata (*mean*) dari semua objek atau data dalam *cluster* dengan persamaan (3) [9], yaitu :

$$c = \frac{\sum_{i=1}^n x_i}{n} \dots\dots\dots (3)$$

dimana :

- $x_i$  : *cluster ke-i*
- $n$  : jumlah objek yang menjadi anggota *cluster*

Ulangi mulai dari langkah ke-3 sampai nilai pusat *cluster* tidak berubah.

**D. Evaluasi Hasil**

Pada tahap ini dilakukan penilaian menggunakan *Davies Bouldin Index* (DBI) untuk menentukan nilai optimal dalam proses *clustering* tersebut. Jika nilai DBI yang diperoleh semakin kecil dan tidak negatif ( $\geq 0$ ) maka semakin baik *cluster* yang diperoleh dari *K-Means* yang digunakan.

Langkah perhitungan DBI :

- 1) Data yang digunakan adalah data dari proses *Chebysev K-Means*, yaitu data yang sudah ter *cluster*.

- 2) Hitung rata-rata dengan persamaan (3) dan *variance* dari masing-masing *cluster* dengan persamaan (4) [12], yaitu :

$$var(x) = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2 \dots\dots\dots (4)$$

- 3) Cari nilai  $R_{ij}$  dan  $R_i$  dengan persamaan (5) dan (6) [12], yaitu :

$$R_i = \max_{j=1, \dots, k; i \neq j} R_{ij} \dots\dots\dots (5)$$

$$R_{ij} = \frac{var(c_i) + var(c_j)}{\|c_i - c_j\|} \dots\dots\dots (6)$$

- 4) Hasil akhir adalah nilai DBI, dengan persamaan (7) [12], yaitu:

$$DBI = \frac{1}{k} \sum_{i=1}^k R_i \dots\dots\dots (7)$$

dimana :

- DBI : validasi *davies bouldin*
- var : *variance* dari data
- N : banyaknya data
- $x_i$  : data ke-*i*
- $\bar{x}$  : rata-rata dari tiap *Cluster*
- $R_{ij}$  : jarak antar *Cluster*
- $c_i$  : *cluster i*

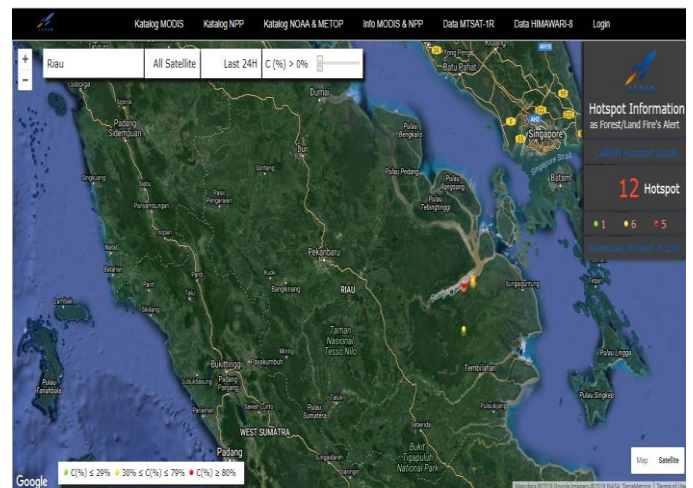
**E. Visualisasi Hasil**

Tahap yang terakhir adalah tahap visualisasi data yang telah dievaluasi yang hasilnya diimplementasikan sebagai rekomendasi ke dinas terkait sehingga dapat digunakan untuk menentukan daerah yang rawan terhadap titik api di Provinsi Riau.

**III. HASIL DAN PEMBAHASAN**

**A. Pengumpulan Data**

Data diperoleh melalui dokumentasi dari Lembaga Penerbangan dan Antariksa Nasional (LAPAN) yang dapat diunduh pada halaman *website* <http://modis-catalog.lapan.go.id/monitoring/>. Halaman pengunduhan data dapat dilihat pada Gambar 1.



**Gambar 1. Halaman Pengunduhan Data**

Data yang digunakan adalah data titik api di Provinsi Riau pada tahun 2016 sebanyak 416 data. Pada data titik api terdapat sembilan atribut, dapat dilihat pada Tabel I.

TABEL I  
KETERENAGAN ATRIBUT DATA TITIK API

Atribut	Keterangan
Lintang	Posisi lintang titik api yang terdeteksi
Bujur	Posisi bujur titik api yang terdeteksi
Tanggal	Tanggal terdeteksinya titik api
Waktu Akuisi	Waktu terdeteksinya titik api
Tingkat Kepercayaan	Tingkat kepercayaan bahwa yang terdeteksi benar-benar titik api (titik yang terdeteksi bahaya)
Satelit	Satelit yang mendeteksi titik api (Aqua,

	Terra, SNPP)
Kecamatan	Kecamatan tempat terdeteksinya titik api
Kabupaten	Kabupaten tempat terdeteksinya titik api
Provinsi	Provinsi tempat terdeteksinya titik api

Dari sembilan atribut hanya lima atribut yang digunakan pada proses *clustering* data, yaitu atribut lintang, bujur, kecamatan, satelit dan kabupaten, dan ditambah satu atribut baru yaitu Id sebagai penanda identitas datanya, dapat dilihat pada Tabel II.

TABEL II  
DATA TITIK API

Id	Kecamatan	Kabupaten	Lintang	Bujur	Satelit
1	Bunguran Barat	Kab. Natuna	3,89837	108,139	snpp
2	Bunguran Barat	Kab. Natuna	3,89837	108,139	snpp
3	Bunguran Barat	Kab. Natuna	3,89837	108,139	snpp
4	Bunguran Barat	Kab. Natuna	3,89837	108,139	snpp
5	Bunguran Barat	Kab. Natuna	3,89837	108,139	snpp
6	Bunguran Barat	Kab. Natuna	3,89837	108,139	snpp
7	Bunguran Barat	Kab. Natuna	3,89837	108,139	snpp
8	Bunguran Barat	Kab. Natuna	3,89837	108,139	snpp
9	Bunguran Barat	Kab. Natuna	3,89837	108,139	snpp
10	Bunguran Barat	Kab. Natuna	3,89837	108,139	snpp
11	Bunguran Barat	Kab. Natuna	3,89837	108,139	snpp
12	Bunguran Barat	Kab. Natuna	3,89837	108,139	snpp
13	Bunguran Barat	Kab. Natuna	3,89837	108,139	snpp
14	Bunguran Barat	Kab. Natuna	3,89837	108,139	snpp
15	Bunguran Barat	Kab. Natuna	3,89837	108,139	snpp
16	Bunguran Barat	Kab. Natuna	3,89837	108,139	snpp
17	Bukit Batu	Bengkalis	1,43849	101,601	aqua
18	Bukit Batu	Bengkalis	1,44892	101,6019	aqua
19	Gaunganakserka	Indragiri Hilir	0	102,954	snpp
20	Tebing Tinggi	Kep Meranti	0,89945	103,072	snpp
21	Bukit Batu	Bengkalis	1,4531	101,604	snpp
22	Kuala Kampar	Pelalawan	0,28744	102,9847	terra
23	Kuala Kampar	Pelalawan	0,28757	102,985	aqua
24	Bukit Batu	Bengkalis	1,41492	101,6092	aqua
25	Mandau	Bengkalis	1,41336	101,5984	aqua
26	Mandau	Bengkalis	1,4118	101,5877	aqua
27	Bukit Batu	Bengkalis	1,42469	101,6077	aqua
28	Mandau	Bengkalis	1,42314	101,597	aqua
29	Bukit Batu	Bengkalis	1,44423	101,6049	aqua
30	Bukit Batu	Bengkalis	1,44269	101,5942	aqua
31	Bukit Batu	Bengkalis	1,45803	101,7454	aqua
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
411	Rangsang	Kep Meranti	1,08385	102,927	snpp
412	Rangsang	Kep Meranti	1,08609	102,918	snpp
413	Rangsang	Kep Meranti	1,09202	102,909	snpp
414	Rangsang	Kep Meranti	1,09917	102,908	snpp
415	Rokan IV Koto	Rokan Hulu	0,41769	100,391	snpp
416	Rokan IV Koto	Rokan Hulu	0,43306	100,403	snpp

**B. Preprocessing Data (Cleaning Data)**

Normalisasi pada data pertama atribut kabupaten pada data titik api tahun 2016 dengan persamaan (1) diperoleh

$$X^* = \frac{1-1}{11-1} = \frac{0}{10} = 0$$

Dengan menggunakan cara yang sama, seluruh data titik api yang akan digunakan pada proses *clustering* dinormalisasikan dengan normalisasi *Min-Max*.

Setelah melewati proses *cleaning*, hanya terdapat 311 data yang digunakan dalam proses *clustering*, dapat dilihat pada Tabel III.

TABEL III  
CLEANING DATASET

Keterangan proses cleaning	Data yang diproses
Data yang di cleaning (di hapus) karena wilayah provinsi bukan termasuk dalam wilayah yang menjadi bagian penelitian	Data dengan id 1-16, 19, 72, 92, 146-176, 179, 180, 182, 183, 191, 221, 223, 225, 227, 228, 230-232, 235, 236, 239-243, 246-250, 252, 253, 256-259, 265, 267-290, 314, 380.
Perubahan angka decimal (pembulatan tiga angka dibelakang koma) pada data titik lintang dan titik bujur	Seluruh data yang sudah di cleaning.

Data hasil *cleaning* yang akan digunakan dalam proses *clustering* dapat dilihat pada Tabel IV.

TABEL IV  
DATA HASIL CLEANING

Id	Kecamatan	Kab.	Lintang	Bujur	Satelit
17	Bukit Batu	0,000	0,702	0,403	0,000
18	Bukit Batu	0,000	0,705	0,404	0,000
19	Gaunganakserka	0,100	0,279	0,777	0,500
20	Tebing Tinggi	0,600	0,544	0,810	0,500
21	Bukit Batu	0,000	0,707	0,404	0,500
22	Kuala kampar	0,700	0,363	0,786	1,000
23	Kuala kampar	0,700	0,363	0,786	0,000
24	Bukit Batu	0,000	0,695	0,406	0,000
25	Mandau	0,000	0,695	0,403	0,000
26	Mandau	0,000	0,695	0,400	0,000
27	Bukit Batu	0,000	0,698	0,405	0,000
28	Mandau	0,000	0,698	0,402	0,000
29	Bukit Batu	0,000	0,704	0,405	0,000
30	Bukit Batu	0,000	0,704	0,402	0,000
31	Bukit Batu	0,000	0,708	0,443	0,000

TABEL V  
HASIL ITERASI PERTAMA

Id	Kecamatan	Kab.	Lintang	Bujur	Satelit	c <sub>1</sub>	c <sub>2</sub>	c <sub>3</sub>
17	Bukit Batu	0,000	0,702	0,403	0,000	0,600	0,002	0,900
18	Bukit Batu	0,000	0,705	0,404	0,000	0,600	0,002	0,900
19	Gaunganakserka	0,100	0,279	0,777	0,500	0,500	0,500	0,800

411	Rangsang	0.600	0.598	0.770	0.500
412	Rangsang	0.600	0.599	0.768	0.500
413	Rangsang	0.600	0.600	0.765	0.500
414	Rangsang	0.600	0.602	0.765	0.500
415	Rokan IV Koto	0.900	0.402	0.069	0.500
416	Rokan IV Koto	0.900	0.406	0.072	0.500

**C. Proses Clustering**

Proses *clustering* menggunakan algoritma *K-Means* dan *Chebyshev Distance* sebagai perhitungan jarak setiap data terhadap pusat *cluster*, yaitu :

1) Jumlah *cluster* 3 (tiga) yang disimbolkan dengan *c<sub>1</sub>* untuk *cluster* 1, *c<sub>2</sub>* untuk *cluster* 2, dan *c<sub>3</sub>* untuk *cluster* 3.

2) Ambil data dengan Id 20, Id 30, dan Id 415 sebagai titik *centroid* maka :

$$c_1 = ( 0,600 ; 0,544 ; 0,810 ; 0,500 )$$

$$c_2 = ( 0,000 ; 0,704 ; 0,402 ; 0,000 )$$

$$c_3 = ( 0,900 ; 0,402 ; 0,069 ; 0,500 )$$

3) Menghitung jarak menggunakan persamaan (2) Jarak data pertama dengan pusat *cluster* pertama

$$d_{11} = \max(|0,000-0,600|,|0,702-0,544|,|0,403-0,810|, |0,000-0,500|)$$

$$= 0,600$$

Jarak data pertama dengan pusat *cluster* kedua

$$d_{12} = \max (|0,000-0,000|,|0,702-0,704|,|0,403-0,402|,|0,000-0,000|)$$

$$= 0,002$$

Jarak data pertama dengan pusat *cluster* ketiga

$$d_{13} = \max (|0,000-0,900|,|0,702-0,402|,|0,403-0,069|,|0,000-0,500|)$$

$$= 0,900$$

Nilai *d<sub>11</sub>* , *d<sub>12</sub>* , dan *d<sub>13</sub>* diletakkan pada baris pertama iterasi pertama di Tabel V.

Pada data pertama, jarak minimal terdapat pada *cluster* 2 sehingga data pertama menjadi anggota *cluster* 2. Selanjutnya dengan cara yang sama dilakukan perhitungan jarak dengan *chebyshev distance* pada seluruh data. Hasil dari iterasi pertama dapat dilihat pada Tabel V.

20	Tebing Tinggi	0,600	0,544	0,810	0,500	0,000	0,600	0,741
21	Bukit Batu	0,000	0,707	0,404	0,500	0,600	0,500	0,900
22	Kuala kampar	0,700	0,363	0,786	1,000	0,500	1,000	0,717
23	Kuala kampar	0,700	0,363	0,786	0,000	0,500	0,700	0,717
24	Bukit Batu	0,000	0,695	0,406	0,000	0,600	0,008	0,900
25	Mandau	0,000	0,695	0,403	0,000	0,600	0,009	0,900
26	Mandau	0,000	0,695	0,400	0,000	0,600	0,009	0,900
27	Bukit Batu	0,000	0,698	0,405	0,000	0,600	0,005	0,900
28	Mandau	0,000	0,698	0,402	0,000	0,600	0,006	0,900
29	Bukit Batu	0,000	0,704	0,405	0,000	0,600	0,003	0,900
30	Bukit Batu	0,000	0,704	0,402	0,000	0,600	0,000	0,900
31	Bukit Batu	0,000	0,708	0,443	0,000	0,600	0,042	0,900
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
411	Rangsang	0,600	0,598	0,770	0,500	0,054	0,600	0,701
412	Rangsang	0,600	0,599	0,768	0,500	0,055	0,600	0,698
413	Rangsang	0,600	0,600	0,765	0,500	0,057	0,600	0,696
414	Rangsang	0,600	0,602	0,765	0,500	0,059	0,600	0,696
415	Rokan IV Koto	0,900	0,402	0,069	0,500	0,741	0,900	0,000
416	Rokan IV Koto	0,900	0,406	0,072	0,500	0,738	0,900	0,005

Cluster 1 Cluster 2 Cluster 3

4) Menghitung pusat cluster yang baru dengan persamaan (3). Untuk cluster 1 atribut kabupaten terdapat 121 data, sehingga :

$$c_{Kab} = \frac{0,100+0,600+0,000+0,700+0,700+\dots+0,600+0,600}{121} = 0,517$$

$$c_{Lintang} = \frac{0,279+0,544+0,707+0,363+0,363+\dots+0,600+0,602}{121} = 0,511$$

$$c_{Bujur} = \frac{0,777+0,810+0,404+0,786+0,786+\dots+0,765+0,765}{121} = 0,627$$

$$c_{Satelit} = \frac{0,500+0,500+0,500+1,000+0,000+\dots+0,500+0,500}{121} = 0,541$$

Untuk mencari nilai centroid baru pada cluster 2 dan cluster 3 dilakukan dengan cara yang sama. Pada cluster 2 terdapat sebanyak 115 data, dan untuk cluster

3 terdapat sebanyak 75 data. Sehingga diperoleh hasil centroid baru untuk melakukan iterasi yang kedua yang dapat dilihat pada Tabel VI.

TABEL VI  
CENTROID BARU UNTUK ITERASI KEDUA

	Kab.	Lintang	Bujur	Satelit
Cluster 1	0,517	0,511	0,627	0,541
Cluster 2	0,064	0,706	0,432	0,487
Cluster 3	0,711	0,725	0,224	0,500

5) Ulangi kembali langkah ke-3 sehingga posisi data pada cluster tidak berubah.

Hasil iterasi pertama dan kedua dapat dilihat pada Tabel VII.

TABEL VII  
HASIL PERUBAHAN PUSAT CLUSTER

Id	Kecamatan	Pusat cluster lama			Pusat cluster baru		
		c <sub>1</sub>	c <sub>2</sub>	c <sub>3</sub>	c <sub>1</sub>	c <sub>2</sub>	c <sub>3</sub>
17	Bukit Batu	0,600	0,002	0,900	0,541	0,487	0,711
18	Bukit Batu	0,600	0,002	0,900	0,541	0,487	0,711
19	Gaunganakserka	0,500	0,500	0,800	0,417	0,427	0,611
20	Tebing Tinggi	0,000	0,600	0,741	0,183	0,536	0,587
21	Bukit Batu	0,600	0,500	0,900	0,517	0,064	0,711
22	Kuala kampar	0,500	1,000	0,717	0,459	0,636	0,562
23	Kuala kampar	0,500	0,700	0,717	0,541	0,636	0,563
24	Bukit Batu	0,600	0,008	0,900	0,541	0,487	0,711
25	Mandau	0,600	0,009	0,900	0,541	0,487	0,711



26	Mandau	0,600	0,009	0,900	0,541	0,487	0,711
27	Bukit Batu	0,600	0,005	0,900	0,541	0,487	0,711
28	Mandau	0,600	0,006	0,900	0,541	0,487	0,711
29	Bukit Batu	0,600	0,003	0,900	0,541	0,487	0,711
30	Bukit Batu	0,600	0,000	0,900	0,541	0,487	0,711
31	Bukit Batu	0,600	0,042	0,900	0,541	0,487	0,711
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
411	Rangsang	0,054	0,600	0,701	0,143	0,536	0,546
412	Rangsang	0,055	0,600	0,698	0,140	0,536	0,544
413	Rangsang	0,057	0,600	0,696	0,138	0,536	0,542
414	Rangsang	0,059	0,600	0,696	0,138	0,536	0,541
415	Rokan IV Koto	0,741	0,900	0,000	0,558	0,836	0,323
416	Rokan IV Koto	0,738	0,900	0,005	0,555	0,836	0,319

cluster 1
  cluster 2
  cluster 3

Karena pada hasil iterasi pertama dan kedua posisi *cluster* berubah, contohnya pada data dengan Id 17 pada iterasi pertama berada pada *cluster 2*, setelah dilakukan iterasi dengan titik *centroid* baru posisi data berubah menjadi anggota dari *cluster 3*. Kemudian dihitung kembali titik *centroid* baru untuk iterasi ketiga seperti langkah ke-4 dengan hasil titik *centroid* baru yang dapat dilihat pada Tabel VIII.

TABEL VIII  
CENTROID BARU UNTUK ITERASI KETIGA

	Kab.	Lintang	Bujur	Satelit
Cluster 1	0,516	0,494	0,612	0,670
Cluster 2	0,106	0,709	0,456	0,391
Cluster 3	0,697	0,721	0,231	0,473

Ulangi langkah ke-3 kembali dengan titik pusat baru pada Tabel IX sampai posisi *cluster* tidak berubah. Karena posisi data pada hasil iterasi kedua dan iterasi ketiga masih berubah maka dihitung kembali *centroid* baru untuk iterasi keempat dengan hasil titik *centroid* baru yang dapat dilihat pada Tabel IX.

TABEL IX  
CENTROID BARU UNTUK ITERASI KEEMPAT

	Kab.	Lintang	Bujur	Satelit
Cluster 1	0,418	0,534	0,593	0,746
Cluster 2	0,133	0,706	0,444	0,250
Cluster 3	0,699	0,708	0,252	0,462

Ulangi langkah ke-3 kembali mencari jarak masing-masing data ke pusat *cluster* dengan titik pusat *cluster* baru pada Tabel X sampai posisi *cluster* tidak berubah. Karena posisi data pada hasil iterasi ketiga dan keempat masih berubah maka dihitung kembali *centroid* baru untuk iterasi kelima dengan hasil titik *centroid* baru yang dapat dilihat pada Tabel X.

TABEL X  
CENTROID BARU UNTUK ITERASI KELIMA

	Kab.	Lintang	Bujur	Satelit
Cluster 1	0,422	0,536	0,585	0,752
Cluster 2	0,135	0,706	0,445	0,243
Cluster 3	0,695	0,711	0,255	0,448

Untuk data titik api di Provinsi Riau tahun 2016 dengan jumlah 311 data titik api, dibutuhkan iterasi sebanyak lima kali untuk mendapatkan hasil klusterisasi akhir. Hasil akhir dari *clustering* data titik api tahun 2016 dapat dilihat pada Tabel XI.

TABEL XI  
HASIL PUSAT CLUSTER OPTIMAL

Id	Kecamatan	Pusat cluster lama			Pusat cluster baru		
		c <sub>1</sub>	c <sub>2</sub>	c <sub>3</sub>	c <sub>1</sub>	c <sub>2</sub>	c <sub>3</sub>
17	Bukit Batu	0,746	0,250	0,699	0,752	0,243	0,695
18	Bukit Batu	0,746	0,250	0,699	0,752	0,243	0,695
19	Gaunganakserka	0,318	0,428	0,599	0,322	0,427	0,595
20	Tebing Tinggi	0,246	0,467	0,558	0,252	0,465	0,555
21	Bukit Batu	0,418	0,250	0,699	0,422	0,257	0,695

22	Kuala kampar	0,282	0,750	0,538	0,278	0,757	0,552
23	Kuala kampar	0,746	0,567	0,534	0,752	0,565	0,531
24	Bukit Batu	0,746	0,250	0,699	0,752	0,243	0,695
25	Mandau	0,746	0,250	0,699	0,752	0,243	0,695
26	Mandau	0,746	0,250	0,699	0,752	0,243	0,695
27	Bukit Batu	0,746	0,250	0,699	0,752	0,243	0,695
28	Mandau	0,746	0,250	0,699	0,752	0,243	0,695
29	Bukit Batu	0,746	0,250	0,699	0,752	0,243	0,695
30	Bukit Batu	0,746	0,250	0,699	0,752	0,243	0,695
31	Bukit Batu	0,746	0,250	0,699	0,752	0,243	0,695
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
411	Rangsang	0,246	0,467	0,518	0,252	0,465	0,515
412	Rangsang	0,246	0,467	0,516	0,252	0,465	0,512
413	Rangsang	0,246	0,467	0,513	0,252	0,465	0,510
414	Rangsang	0,246	0,467	0,513	0,252	0,465	0,510
415	Rokan IV Koto	0,524	0,767	0,306	0,516	0,765	0,309
416	Rokan IV Koto	0,521	0,767	0,301	0,512	0,765	0,305

cluster 1
  cluster 2
  cluster 3

Berdasarkan hasil *clustering* Tabel XI diperoleh hasil analisis sebagai berikut:

1) *Cluster* pertama ( $c_1$ ) mempunyai titik pusat (0,422; 0,536; 0,585; 0,752) dengan jumlah anggota *cluster* sebanyak 133 data titik api. Sehingga dapat disimpulkan bahwa  $c_1$  merupakan kelompok daerah yang sangat rawan terhadap titik api.

2) *Cluster* kedua ( $c_2$ ) mempunyai titik pusat *cluster* (0,135; 0,706; 0,445; 0,243) dengan jumlah anggota *cluster* sebanyak 101 data titik api. Sehingga dapat disimpulkan bahwa  $c_2$  merupakan kelompok daerah yang rawan terhadap titik api.

3) *Cluster* ketiga ( $c_3$ ) mempunyai titik *cluster* (0,695; 0,711; 0,255; 0,448) dengan jumlah anggota *cluster* sebanyak 77 data titik api. Sehingga dapat disimpulkan bahwa  $c_3$  merupakan kelompok daerah yang tidak rawan terhadap titik api.

Berdasarkan Tabel XI maka daerah yang rawan terhadap titik api pada tahun 2016 dapat dilihat pada Tabel XII.

TABEL XII  
DAFTAR DAERAH RAWAN TITIK API DI PROVINSI  
RIAU TAHUN 2016

Status	Kabupaten	Kecamatan	Jumlah Titik api
Tidak Rawan	Rokan Hilir	Tanah Putih	8
		Pasir Limau	1
		Bangko	9
		Kubu	4
		Pujud	15

	Rimba Melintang	2	
	Rokan IV Koto	2	
	Tambusai	1	
Rokan Hulu	Bangun Purba	1	
	Bonai Darussalam	1	
	Rambah	2	
Siak	Kandis	2	
Pelalawan	Kuala Kampar	3	
Dumai	Bukit Kapur	23	
	Sungai Sembilan	2	
Kep.Meranti	Merbau	1	
<b>Jumlah titik api yang tidak rawan</b>		<b>77</b>	
Kep.Meranti	Tebing tinggi	2	
	Rangsang	4	
Siak	Sungai Apit	2	
	Kandis	2	
	Bintan	8	
	Siak Kecil	1	
Rawan	Bengkalis	3	
	Bengkalis	Bukit Batu	28
		Rupat	25
		Pinggir	4
		Mandau	7
	Dumai	Bukit Kapur	9
		Dumai Barat	3
	Rokan Hilir	Pujud	1
	Kampar	Tapung	2
	<b>Jumlah titik api yang rawan</b>		<b>101</b>
Rokan Hilir	Tanah Putih	3	
Rokan Hulu	Rokan IV Koto	1	
Sangat Rawan	Kuantan	Kuantan Mudik	1
	Singingi		
	Siak sri indrapura	2	
Siak	Kandis	3	
	Dayun	4	



	Sungai Apit	16
	Bunga Raya	1
	Pusako	1
	Bukit Batu	8
	Siak Kecil	3
Bengkalis	Rupat	6
	Rupat Utara	1
	Bengkalis	3
	Bintan	4
Kep.Meranti	Rangsang	7
	Rangsang Barat	4
	Tebing Tinggi	10
	Merbau	17
Dumai	Bukit Kapur	8
	Dumai Barat	1
Indragiri	Gaungan Serka	1
Hilir	Mandah	8
Indragiri	Kuala cenaku	2
Hulu	Batang Cenaku	1
	Kuala Kampar	3
	Pangkalan	3
Pelalawan	Ukui	9
	Teluk Meranti	1
Kampar	Kampar Kiri hulu	1

**Jumlah titik api yang sangat rawan 133**

**D. Evaluasi Hasil**

Berdasarkan hasil *clustering* pada Tabel XI. maka dapat diuji hasil *clustering* pada data titik api di Provinsi Riau tahun 2016 dengan menggunakan perhitungan DBI :

1) Menghitung rata-rata dan *variance* data dari masing-masing *cluster* persamaan (4).

Cluster 1 :

$$\bar{x} = \frac{0,322+0,252+0,278+0,422+0,248+\dots+0,252+0,252+0,252}{133}$$

$$= 0,312$$

$$var( c_1 ) = \frac{1}{133-1} (0,322 - 0,312)^2 + (0,252 - 0,312)^2 + (0,248 - 0,312)^2 + (4,218 - 0,312)^2 + \dots + (0,252 - 0,312)^2 + (0,252 - 0,312)^2$$

$$= 0,00635$$

Cluster 2 :

$$\bar{x} = \frac{0,243+0,243+0,257+0,243+0,243+\dots+0,465+0,465+0,465}{130}$$

$$= 0,293$$

$$var( c_2 ) = \frac{1}{101-1} (0,243 - 0,293)^2 + (0,243 - 0,293)^2 + (0,257 - 0,293)^2 + (0,243 - 0,293)^2 + \dots + (0,465 - 0,293)^2 + (0,465 - 0,293)^2$$

$$= 0,00392$$

Cluster 3 :

$$\bar{x} = \frac{0,531+0,195+0,195+0,195+0,195+\dots+0,105+0,309+0,305}{77}$$

$$= 0,228$$

$$var( c_3 ) = \frac{1}{77-1} (0,531 - 0,228)^2 + (0,195 - 0,228)^2 + (0,195 - 0,228)^2 + (0,195 - 0,228)^2 + \dots + (0,309 - 0,228)^2 + (0,305 - 0,228)^2$$

$$= 0,01856$$

2) Menghitung nilai  $R_{ij}$  dengan menggunakan persamaan (6) maka :

$$R_{12} = \frac{var(c_1)+ var(c_2)}{\|c_1- c_2\|}$$

$$= \frac{0,00635 + 0,00392}{\|0,312 - 0,293\|}$$

$$= 0,541$$

$$R_{13} = \frac{var(c_1)+ var(c_3)}{\|c_1- c_3\|}$$

$$= \frac{0,00635 + 0,01856}{\|0,312 - 0,228\|}$$

$$= 0,296$$

$$R_{23} = \frac{var(c_2)+ var(c_3)}{\|c_2- c_3\|}$$

$$= \frac{0,00392 + 0,01856}{\|0,293 - 0,228\|}$$

$$= 0,346$$

3) Menghitung nilai  $R_i$  dan nilai DBI dengan persamaan (5) dan persamaan (7) maka :

$$R_1 = \max ( R_{12} ; R_{13} ; R_{23} )$$

$$= \max ( 0,541 ; 0,296 ; 0,346 )$$

$$= 0,541$$

$$R_2 = \max ( R_{21} ; R_{23} )$$

$$= \max ( 0,541 ; 0,346 )$$

$$= 0,541$$

$$DBI = \frac{1}{k} \times ( R_1 + R_2 )$$

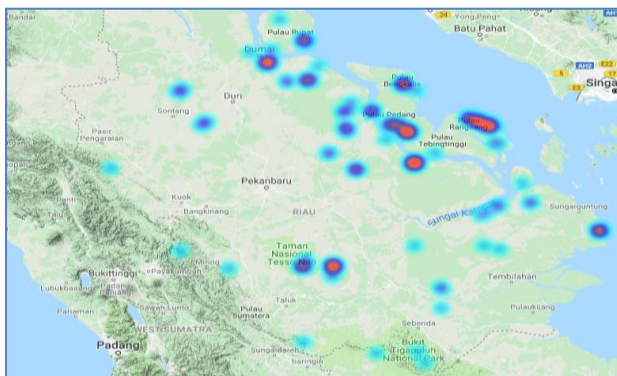
$$= \frac{1}{3} \times 1,082$$

$$= 0,361$$

Jadi, hasil *clustering* pada data titik api di Provinsi Riau pada tahun 2016, memiliki hasil *clustering* yang optimal dengan nilai DBI sebesar 0,361.

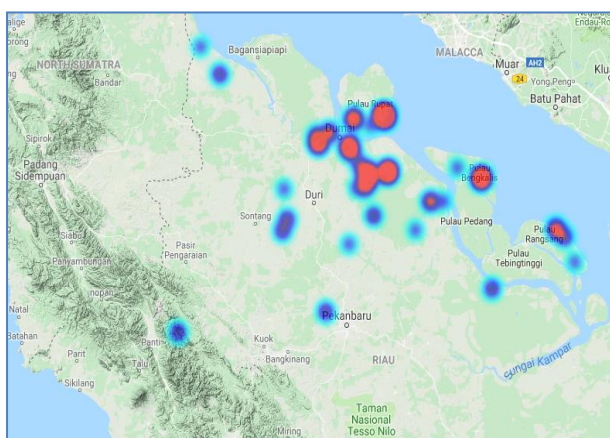
**E. Visualisasi Hasil**

Hasil visualisasi pada *cluster* sangat rawan titik api tahun 2016 dapat dilihat langsung pada titik-titik yang terlihat banyak tersebar, sebagaimana pada Gambar 2.



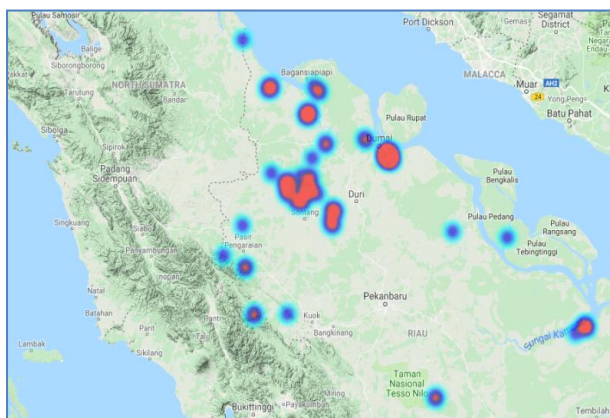
**Gambar 2. Lokasi Cluster sangat Rawan Titik Api**

Hasil visualisasi pada *cluster* rawan titik api tahun 2016 dapat dilihat langsung pada titik-titik yang terlihat daerah persebaran titik lebih banyak di pinggir Provinsi Riau, sebagaimana terlihat pada Gambar 3.



**Gambar 3. Lokasi Cluster Rawan Titik Api**

Hasil visualisasi pada *cluster* tidak rawan titik api tahun 2016 dapat dilihat langsung pada titik-titik yang terlihat tidak banyak tersebar sehingga terlihat bahwa jumlah *hotspot* yang tidak rawan hanya sedikit karena titik yang ada bergerombol namun hanya berapa pada beberapa wilayah saja, sebagaimana terlihat pada Gambar 4.



**Gambar 4. Lokasi Cluster Tidak Rawan Titik Api**

#### IV. KESIMPULAN

Berdasarkan hasil penelitian yang dilakukan, maka dapat disimpulkan:

- A. Pada tahun 2016 terdapat 133 titik yang masuk kedalam *cluster* daerah sangat rawan titik api, 101 titik kedalam *cluster* daerah rawan titik api, dan 77 titik kedalam *cluster* daerah yang tidak rawan terhadap titik api dengan nilai DBI 0,361 yang menandakan bahwa pengklasteran dengan *Chebyshev K-Means* sudah optimal dalam menghasilkan *cluster set* sebanyak 3 *cluster*.
- B. Berdasarkan hasil visualisasi yang telah dilakukan, dapat disimpulkan bahwa daerah yang rawan terhadap titik api adalah daerah yang memiliki beberapa titik-titik api yang posisinya berdekatan dan juga tersebar di banyak daerah, sehingga titik-titik tersebut berpotensi akan menimbulkan kebakaran. Sedangkan daerah yang tidak rawan titik api memiliki pola posisi titik api yang hanya ada satu pada satu daerah (titik api tidak bergerombol) dan tidak tersebar.

#### DAFTAR PUSTAKA

- [1] Agusta, Y., 2007, "K-Means – Penerapan, Permasalahan dan Metode Terkait", *Jurnal Sistem Dan Informatika*, volume 3, halaman 47–60.
- [2] Fauzan, A., Baharudin, A.Y., dan Wibowo, F., 2014, "Sistem Klasterisasi Menggunakan Metode K-Means dalam Menentukan Posisi Access Point Berdasarkan Posisi Pengguna Hotspot di Universitas Muhammadiyah Purwokerto (Clustering System Using K-Means Method in Determining Access Point Position at Muhammadiyah University of Purwokerto)", *Jurnal Informatika (JUITA)*, volume 3, nomor 1, halaman 25–29.
- [3] Rahayu, S., Nugrahadi, D. T., & Indriani, F., 2014, "Clustering Penentuan Potensi Kejahatan Daerah Di Kota Banjarbaru Dengan Metode K-Means", *Kumpulan Jurnal Ilmu Komputer (KLIK)*, volume 1, nomor 1, halaman 33–45.
- [4] Meisida, N., Soesanto, O., & Chandra, H. K., 2014, "K-Means untuk Klasifikasi Penyakit Karies Gigi", *Kumpulan Jurnal Ilmu Komputer (KLIK)*, volume 1, nomor 1, halaman 12–22.
- [5] Nugroho, Y. S., & Haryati, S. N., 2015, "Klasifikasi dan Klastering Penjurusan Siswa SMA Negeri 3 Boyolali", *Khazanah Informatika*, volume 1, nomor 1, halaman 3–8.
- [6] Apriyanti, N. R., Nugroho, R. A., & Soesanto, O., 2015, "Algoritma K-Means Clustering Dalam Pengolahan Citra Digital Landsat", *Kumpulan Jurnal Ilmu Komputer (KLIK)*, volume 2, nomor 2, halaman 1–13.

- [7] Arwani, I., 2015, "Integrasi Algoritma K-Means Dengan Bahasa SQL untuk Klasterisasi IPK Mahasiswa (Studi Kasus: Fakultas Ilmu Komputer Universitas Brawijaya)", *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK)*, volume 2, nomor 2, halaman 143–151.
- [8] Robani, M., & Widodo, A., 2016, "Algoritma K-Means Clustering untuk Pengelompokan Ayat Al Quran Pada Terjemahan Bahasa Indonesia", *Jurnal Sistem Informasi Bisnis*, volume 6, nomor 2, halaman 164–176.
- [9] Adams J, E. M., & Sagita R, D., 2016, "Sistem Temu Kembali Citra Berbasis Konten Menggunakan Haar Wavelet Transform dan K-Means Clustering", *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK)*, volume 3, nomor 1, halaman 1–8.
- [10] Elizawati, N., & Lesmana, L. S., 2017, "Analisis Nilai Rapor Siswa Kelas X Jurusan Multimedia Terhadap Minat pada Pelajaran Produktif Di Kelas XII untuk Menentukan Kompetensi Siswa dengan Metode Clustering Algoritma K-Means ( Studi Kasus Di SMKN 4 Padang )", *Jurnal Komputer Terapan*, volume 3, nomor 2, halaman 133–148.
- [11] Nishom, M., & Fathoni, M. Y., 2018, "Implementasi Pendekatan Rule-Of-Thumb untuk Optimasi Algoritma K-Means Clustering", *Jurnal Pengembangan IT (JPIT)*, volume 3, nomor 2, halaman 237–241.
- [12] Lutfhi, M. A., 2016, "Klasterisasi Daerah Rawan Kecelakaan Lalu Lintas di Kota Semarang Menggunakan Chebyshev Distance K-Means", *Skripsi*, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro, Semarang.

