**Proposal Tugas Akhir**

**Deteksi Pencilan Data Titik panas di Provinsi Riau Menggunakan Algoritme *Local Outlier Factor***

SHOFYAN (G64134009)[[1]](#footnote-1)\*, IMAS SUKAESIH SITANGGANG

**ABSTRAK**

Indonesia mengalami kebakaran hutan yang signifikan. Pada tahun 2013 *World Resources Institute* (WRI) meneliti tren historis titik panas di Pulau Sumatera menggunakan data titik panas aktif *National Aeronautics and Space Administration* (NASA). Pada 13-30 Juni 2013 terjadi 2643 total jumlah peringatan titik panas. Tahun berikutnya Pada  20 Februari hingga 11 Maret tahun 2014 titik panas meningkat menjadi 3101 peringatan titik panas. Salah satu upaya untuk menangani kebakaran hutan ialah dengan menganalisis data titik panas yaitu dengan menganalisis pencilan titik panas sehingga dapat diidentifikasi wilayah yang beresiko terjadinya kebakaran hutan. Beberapa penelitian terkait deteksi pencilan yang sudah dilakukan diantaranya menggunakan algoritme *clustering* k-means dan juga menggunakan algoritme *clustering* berbasis medoids. Kedua penelitian tersebut mendeteksi pencilan berdasarkan frekuensi terjadinya titik panas dan belum mendeteksi pencilan berdasarkan kepadatan penyebaran titik panas. Algoritme yang dapat mendeteksi pencilan berdasarkan kepadatan penyebaran titik panas ialah algoritme *local outlier factor.*  Dengan algoritme *local outlier factor* informasi mengenai wilayah yang berpotensi terjadi kebakaran hutan berdasarkan kepadatan penyebaran titik panas dapat dideteksi sehingga menjadi informasi tambahan untuk pengambilan keputusan oleh pihak terkait.

Kata Kunci**:** kebakaran hutan, *local outlier factor* , titik panas.

***ABSTRACT***

*Indonesia experienced significant forest fires. In 2013 the World Resources Institute (WRI) examined historical trends of hot spots in Sumatra Island using active hot spot data National Aeronautics and Space Administration (NASA). On 13-30 June 2013 occurred in 2643 the total number of hot spots warning. The next year On 20 February to March 11 2014 of hot spots increased to 3101 warning of hot spots. One effort to deal with forest fires is to analyze the data hot spots by analyzing outliers hot spots that can be identified areas at risk of forest fires. Some studies related to the detection of outliers that have been made include using k-means clustering algorithm and also uses a clustering algorithm based medoids. Both studies detecting outliers based on the frequency occurrence of hot spots and yet detect outliers based density of hot spots. Algorithms that can detect outliers based density of hot spots is local outlier factor algorithm. With local outlier factor algorithm information regarding the area of ​​forest fires that could potentially occur based density of hot spots can be detected so that it becomes additional information for decision-making by all relevant parties.*

*Keywords****:*** *forest fire, hot spot, local outlier factor.*

**PENDAHULUAN**

**Latar Belakang**

Indonesia mengalami kebakaran hutan yang signifikan. Pada tahun 2013 World Resources Institute (WRI) meneliti tren historis titik panas di Pulau Sumatera menggunakan data titik panas aktif NASA pada 13-30 Juni 2013 terjadi 2643 total jumlah peringatan titik panas. Tahun berikutnya Pada  20 Februari hingga 11 Maret tahun 2014 titik panas meningkat menjadi 3101 peringatan titik panas (Sizer *et al*. 2014).

Kebakaran hutan dapat mengakibatkan pencemaran kabut asap, emisi karbon, degradasi dan deforesasi hutan yang mengakibatkan hilangnya hasil hutan dan berbagai jasa lingkungan yang diberikan hutan seperti kayu, hasil hutan non- kayu, dan keanekaragaman hayati, serta kerugian di sektor pedesaan contohnya dampak kabut asap pada hasil produksi pertanian (Tacconi 2003).

Data titik panas dapat dijadikan sebagai salah satu indikator tentang kemungkinan terjadinya kebakaran hutan (Adinugroho *et al.* 2005) sehingga dengan menganalisis data titik panas dapat diketahui langkah yang dapat diambil oleh pihak terkait. Diantara analisis yang dapat dilakukan ialah deteksi pencilan titik panas.

Beberapa penelitian terkait deteksi pencilan sudah dilakukan menggunakan algoritme *clustering* k-means (Baehaki 2014) dengan Rata-rata pencilan yang terdeteksi adalah sebesar 481.22 titik panas. Frekuensi titik panas minimum yang terdeteksi sebagai pencilan sebesar 284 titik panas dan terbesar adalah 1118 titik panas. Penelitian kedua menggunakan *clustering* berbasis medoids yaitu PAM dan CLARA (Cahyadarena 2014) . Hasil algoritme PAM pencilan titik panas terjadi pada nilai k=17 dengan cluster ke 13,14,15,16 dan 17. Algoritme CLARA pencilan titik panas terjadi pada nilai k=19 dengan cluster ke 14,15,17 dan 19.

Kedua penelitian tersebut meneliti pencilan titik panas berdasarkan frekuensi terjadinya titik panas, belum mendeteksi pencilan berdasarkan kepadatan penyebaran titik panas. Algoritme yang dapat mendeteksi pencilan denga kriteria tersebut ialah algoritme *local outlier factor* (Beunig 2000). *Local outlier factor* dapat mendeteksi pencilan lokal. Pencilan local ini tidak dapat dideteksi sebagai pencilan jika menggunakan pendekatan *clustering.*

**Perumusan Masalah**

Rumusan masalah dalam penelitian ini adalah bagaimana pencilan diidentifikasi dari data titik panas menggunakan metode *local outlier factor* dan informasi tentang karakteristik pencilan titik panas.

**Tujuan Penelitian**

Tujuan penelitian ini adalah

1. Menentukan pencilan pada data titik panas di Provinsi Riau berdasarkan hasil algoritme *local outlier factor* data titik panas di Provinsi Riau, dan
2. Analisis pencilan data titik panas yang dihasilkan berdasarkan aspek lokasi dan waktu.

**Manfaat Penelitian**

Manfaat dari penelitian ini adalah mendapatkan informasi yang tersembunyi berupa pencilan data titik panas sebagai indikator kebakaran hutan. Penelitian ini juga bermanfaat untuk mengidentifikasi wilayah yang beresiko terjadi kebaharan hutan. Hasil yang diperoleh diharapkan dapat bermanfaat dalam pencegahan kebakaran hutan.

**Ruang Lingkup Penelitian**

Ruang lingkup dari penelitian ini meliputi:

1. Pencilan yang dideteksi adalah pencilan lokal.
2. Implementasi menggunakan *library* DMwR *package* R

**TINJAUAN PUSTAKA**

**Titik Panas**

Titik panas merupakan suatu daerah di permukaan bumi yang memiliki suhu relatif lebih tinggi dibandingkan daerah di sekitarnya berdasarkan ambang batas suhu tertentu (Suwarsono *et al* 2013). Ambang batas yang digunakan oleh Kementrian Kehutanan dan Japan International Cooperation Agency untuk data titik panas satelit NOAA adalah 315o K atau 42o C (Guswanto dan Heriyanto 2009). Pemantauan titik panas merupakan suatu upaya untuk mengendalikan terjadinya kebakaran hutan atau lahan secara dini. Pemantauan titik panas dilakukan dengan remote sensing atau yang biasa disebut sebagai penginderaan jarak jauh dengan menggunakan satelit. Titik panas dipantau menggunakan Satelit Terra (EOS AM) dan satelit Aqua (EOS PM) yang meggunakan sensor *Moderate-resolution Imaging Spectroradiometer* (MODIS). Titik panas merupakan suatu indikasi potensi akan terjadinya kebakaran hutan atau lahan. Pada semua titik panas belum tentu akan terjadi kebakaran, sehingga perlu dilakukan analisis lebih lanjut.

Pemantauan titik panas bertujuan untuk pengendalian kebakaran lahan/hutan dan kebun. Data titik panas bukan jumlah kejadian kebakaran hutan pada wilayah tertentu. Namun, data titik panas dapat digunakan untuk memprediksi kebakaran hutan berdasarkan pola penyebaran titik panas, perubahan koordinat titik panas dan jangka waktu adanya titik panas. Prediksi dengan menggunakan data titik panas bertujuan untuk mengenali asal mula kebakaran dan menelusuri penyebarannya.

**Deteksi Pencilan**

Metode deteksi pencilan terbagi menjadi dua metode (Han *et al.* 2012), yaitu:

1 Deteksi pencilan dengan sampel data untuk analisis diberikan dengan domain *expert*, yaitu domain *expert* memberikan label-label yang dapat digunakan untuk membangun model deteksi pencilan. Metode deteksi pencilan ini, terdiri dari 3 metode, yaitu:

a Metode *supervised*

Domain *expert* menguji dan melabelkan sebuah sampel dari data pokok. Deteksi pencilan dapat dimodelkan sebagai masalah klasifikasi. Tugas dari metode ini adalah untuk mempelajari sebuah *classifier* yang dapat mengenali pencilan. Sampel digunakan sebagai data latih dan data uji.

b Metode *unsupervised*

Metode *unsupervised* digunakan jika objek-objek tidak dilabelkan sebagai “normal” atau “pencilan”. Asumsi implisit metode ini, yaitu objek-objek normal adalah bergerombol.

c Metode *semi-supervised*

Metode semi-supervised digunakan jika suatu kumpulan kecil dari objek-objek normal dan atau pencilan yang diberi label, tetapi sebagian besar data tidak diberi label.

2 Deteksi pencilan dengan membagi metode menjadi kelompok-kelompok sesuai dengan asumsi kelompok-kelompok tersebut mengenai objek-objek yang normal dibandingkan dengan pencilan. Metode ini terdiri atas tiga jenis, yaitu:

a Metode statistik

Metode statistik juga disebut sebagai metode yang berbasis model. Metode ini membuat asumsi normalitas data. Metode ini beranggapan bahwa objek data normal dihasilkan oleh model statistik dan data yang tidak mengikuti model statistik adalah pencilan.

b Metode berbasis kedekatan

Metode berbasis kedekatan berasumsi bahwa sebuah objek adalah pencilan jika kedekatan objek dengan tetangganya secara signifikan menyimpang dari kedekatan sebagian besar objek-objek dengan tetangga-tetangganya pada kumpulan data yang sama atau jika tetangga terdekat dari objek kedekatannya jauh pada ruang fitur.

c Metode berbasis *clustering*

Metode berbasis *clustering* berasumsi bahwa objek-objek data normal milik kelompok yang besar dan padat, sedangkan pencilan milik kelompok kecil atau jarang, atau tidak tergabung dalam kelompok.

***Local Outlier Factor***

*Local outlier factor*  adalah algoritme deteksi *outlier* berdasarkan jarak tetangga lokal (Beunig 2000). *Local outlier factor* tidak menggunakan distribusi data global . Pada Gambar 1 dapat dilihat O1 dan O2 adalah local outlier C1, O3 adalah *global outlier*. Dengan pendekatan *clustering* O1 dan O2 tidak dapat terdeteksi sebagai *outlier*. Ilustrasi *local outlier*  dapat dilihat pada Gambar 2. *Local outlier factor*  menghitung jarak maksimum dari jarak tetangga dengan jumlah tetangga yang didefinisikan oleh pengguna. Notasinya sebagai berikut (Beunig 2000):

*reach-dist* k (p, o) = max { *k-distance*(o), d(p, o) }.

dengan

k = jumlah tetangga

*distance* = fungsi jarak

Sebuah objek *local outlier factor*  adalah rasio rataan dari o degan jarak k-ketetanggan lokal terdekat. Notasinya sebagai berikut (Beunig 2000):

dengan

lrd = *local reachability density*

LOF = *local outlier factor*

O4

O1

C1

C2

O3

Gambar 1. Ilustrasi *local outlier factor.*

**METODE**

**Data Penelitian**

Data yang digunakan dalam penelitian ini adalah data titik panas di Provinsi Riau dari tahun 2002 sampai 2015 yang diperoleh dari *Fire Information for Resource*

*Management System* (FIRMS) MODIS NASA. Aspek yang diamati pada data titik panas adalah aspek spasial kemunculan titik panas bulanan. *Field*  yang digunakan adalah *field* lattidude*,* longitude*,* brightness, scan, track, acq\_date, acq\_time, satellite, confidence, version,bright\_t31, dan frp seperti yang ditunjukan pada Tabel 1.

Tabel 1 *field* data titik panas.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| LATITUDE | LONGITUDE | BRIGHTNESS | SCAN | TRACK | ACQ\_DATE |
| -0.747 | 100.915 | 312.2 | 2.4 | 1.5 | 1/1/2012 |
| -0.74 | 100.92 | 315.5 | 2.4 | 1.5 | 1/1/2012 |
| -2.196 | 101.059 | 312 | 2.5 | 1.5 | 1/1/2012 |
| -0.156 | 101.343 | 317.9 | 1.5 | 1.2 | 1/1/2012 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ACQ\_TIME | SATELLITE | CONFIDENCE | VERSION | BRIGHT\_T31 | FRP |
| 624 | A | 49 | 5.1 | 295 | 29.6 |
| 624 | A | 64 | 5.1 | 295.1 | 41.8 |
| 624 | A | 46 | 5.1 | 291.6 | 27.1 |
| 328 | T | 51 | 5.1 | 305.3 | 22.6 |

Penjelasan *field* Tabel 1 yaitu:

Lattidude : Koordinat lintang

Longitude : Koordinat bujur

Brightness : Suhu kecerahan, diukur dalam Kelvin menggunakan saluran MODIS

21/22 dan saluran 31.

Scan & Track : Resolusi spasial yang sebenarnya dari pixel dipindai. Meskipun algoritma

bekerja pada 1 km resolusi, piksel MODIS mendapatkan lebih besar ke tepi

*scan*.

Acq\_Date : Tanggal akuisisi pixel titik panas aktif.

Acq\_Time : Waktu layang satelit di UTC.

Satellite : Jenis satelit yang mendeteksi. Satelit Terra atau Aqua.

Confidence : Tingkat kepercayaan dari tiap pixel api yang aktif.

Version :Versi yang mengidentifikasikan koleksi data.

Bright\_T31 : Saluran 31 suhu kecerahan (dalam Kelvin) dari pixel titik panas yang aktif.

Frp : *Fire Radiative Power*, Menggambarkan kekuatan radiasi api pixel-

terintegrasi dalam MW (megawat). FRP memberikan informasi mengenai output panas radiasi diukur dari kebakaran terdeteksi. Jumlah energi panas radiasi dibebaskan per satuan waktu (*Fire Radiative Power*) diduga terkait dengan tingkat dimana bahan bakar yang dikonsumsi

**Tahapan Penelitian**

Tahapan-tahapan yang dilakukan pada penelitian ini dapat dilihat pada Gambar 2.

Deteksi pencilan titik panas menggunakan algoritme *local outlier factor*

Praproses data

Data titik panas

Visualisasi data

Analisis pencilan

Gambar 2 Tahapan penelitian

1. **Pengumpulan Data**

Data yang digunakan dalam penelitian ini adalah data titik panas di Provinsi Riau dari tahun 2001 sampai 2014 yang diperoleh dari FIRMS MODIS NASA. Data titik panas terdiri dari data titik panas tahun 2001 hingga tahun 2014 di wilayah Provinsi Riau. Data tersebut terdiri dari atribut *latitude, longitude, brightness, scan, track, acq\_date, acq\_time, satellite, confidence, version, bright\_t31, frp.* Setiap barisnya menjelaskan satu kemunculan titik panas yang diperoleh dari pengindraan jarak jauh menggunakan sensor MODIS.

1. **Praproses Data**

Menurut Han *et al* (2012) “dalam tahap praproses data, terdapat beberapa tahap utama, yaitu pembersihan data, pengintegrasian data, seleksi data, dan transformasi data”. Dalam penelitian ini dilakukan pembersihan dan transformasi data. Pembersihan data dilakukan untuk memilih data titik panas yang berada di Provinsi Riau juga memilih peta Provinsi Riau dari peta kabupaten dan kota se-Indonesia. Langkah ini dilakukan untuk menghilangkan data titik panas yang berada di luar Provinsi Riau. Tahap ini dilakukan menggunakan perangkat lunak PostgreSQL, PostGIS 2.0 *Shapefile and DBF Loader Eksporter*, dan Quantum GIS.

Setelah data bersih, kemudian data titik panas pada tahun 2001 hingga 2012 dipilih menggunakan queri pada DBMS PostgreSQL dan dilakukan transformasi data yaitu agregasi data. Agregasi data adalah operasi penjumlahan jumlah kejadian titik panas menjadi data harian, bulanan ataupun tahunan.

1. **Deteksi Pencilan Titik panas Menggunakan Algoritme *Local Outlier Factor***

Dalam tahapan ini diterapkan fungsi *local outlier factor* pada perangkat lunak R. Fungsi tersebut diberikan masukkan atau argumen berupa data frekuensi titik panas harian dari tahun 2002 hingga 2015 juga nilai k sebesar 2 hingga 10.

1. **Visualisasi Data**

Pada tahapan ini data yang diolah dengan algoritme *local outlier factor* divisualisasikan pada peta sehingga dapat terlihat dengan mudah data mana saja yang termasuk pencilan.

1. **Analisis Pencilan**

Pada tahap ini diperlihatkan objek-objek pencilan dari data penelitian. Data hasil deteksi pencilan dianalisis untuk mengetahui informasi yang terdapat pada data seperti ukuran pemusatan dan tanggal-tanggal yang terdeteksi pencilan.

**Lingkungan Pengembangan**

Spesifikasi perangkat keras dan perangkat lunak yang digunakan dalam penelitian ini adalah sebagai berikut :

1. Perangkat keras berupa komputer personal dengan spesifikasi
   * + Prosesor Intel(R) Core(TM) i7-5500U 2.40GHz dan
     + Memori RAM 12288 MB.
2. Perangkat lunak
   * + Komputasi statistika R versi 3.2.0
     + RStudio versi 0.98.1103,
     + *Database Management System* (DBMS) PostgreSQL dengan ekstensi PostGIS,
     + Pengolah data spatial Quantum GIS 2.6.1, dan
     + Microsoft Excel.
     + PostGIS 2.0 *Shapefile and DBF Loader Eksporter*.
     + Library DMwR pada perangkat lunak R

**JADWAL PENELITIAN**

Jadwal pernelitian dimulai dari bulan April 2015 sampai dengan Desember 2015. Ilustrasi penjadwalan dapat dillihat pada Tabel 2.

Tabel 2 rencana jadwal penelitian.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Tahun 2015** | | | | | | | | |
| **April** | **Mei** | **Juni** | **Juli** | **Agus** | **Sept** | **Okt** | **Nov** | **Des** |
| Penyusunan proposal |  |  |  |  |  |  |  |  |  |
| Pengumpulan data |  |  |  |  |  |  |  |  |  |
| Praproses data |  |  |  |  |  |  |  |  |  |
| Kolokium |  |  |  |  |  |  |  |  |  |
| Implementasi Local outlier factor |  |  |  |  |  |  |  |  |  |
| Evaluasi hasil penelitian |  |  |  |  |  |  |  |  |  |
| Penyusunan skripsi dan makalah seminar |  |  |  |  |  |  |  |  |  |
| Seminar |  |  |  |  |  |  |  |  |  |
| Sidang tugas akhir |  |  |  |  |  |  |  |  |  |
| Revisi Skripsi dan penyelesaian |  |  |  |  |  |  |  |  |  |
| Surat Keterangan Lulus (SKL) |  |  |  |  |  |  |  |  |  |

**DAFTAR PUSTAKA**

Adinugroho WC, Suryadiputra INN, Saharjo BH, Siboro L. 2005. Panduan Pengendalian Kebakaran Hutan dan Lahan Gambut. Proyek Climate Changes, Forests and Peatlands in Indonesia. Bogor(ID): Wetlands International-Indonesia Programme and Wildlife Habibat Canada.

Baehaki D. 2014. Deteksi pencilan data titik panas di provinsi Riau menggunakan algoritme clustering K-Means [skripsi]. Bogor(ID): Insitut Pertanian Bogor.

Beunig Markus M, Kriegel Hans-Peter, Ng Raymond T, Sander J*.* 2000. LOF: Identifying Density-Based Local Outliers. ACM SIGMOD international conference on Management of data; 2, June 2000; New York, USA. New York (USA): ACM SIGMOD Volume 29 Issue Pages 93-104

Cahyadarena M B.2014. Deteksi Pencilan Pada Data Titik Panas Menggunakan Clustering Berbasis Medoids [skripsi]. Bogor(ID): Insitut Pertanian Bogor.

Guswanto, Heriyanto E. 2009. Operational Weather System for National Fire Danger Rating. Jurnal Meteorologi dan Geofisika. 10(2): 77-87

Han J, Kamber M, Pei J. 2012. Data mining: concepts and techniques. Massachusetts (US) : Morgan Kaufmann.

Hasan M I. 1999. Pokok-Pokok Materi Statistik 1: Statistik Deskriptif. Jakarta (ID): Bumi aksara.

Sizer N, Anderson J, Stolle F, Minnemeyer S, Higgins M, Leach A, Alisjahbana A, Utami A. 2014. Kebakaran Hutan di Indonesia Mencapai Tingkat Tertinggi Sejak Kondisi Darurat Kabut Asap Juni 2013 [Internet]. [diunduh 2015 17 Mei]. Tersedia pada http://www.wri.org/blog/2014/03/kebakaran-hutan-di-indonesia-mencapai-tingkat- tertinggi-sejak-kondisi-darurat-kabut.

Suwarsono, Rokhmatuloh, Waryono T. 2013. Pengembangan Model Identifikasi Daerah Bekas Kebakaran Hutan dan Lahan (Burned Area) Menggunakan Citra MODIS di Kalimantan [Model Development of Burned Area Identification Using MODIS Imagery in Kalimantan]. Jurnal Penginderaan Jauh. 10(2): 93-112.

Tacconi L. 2003. Kebakaran Hutan di Indonesia: Penyebab, Biaya dan Implikasi Kebijakan[paper]. Bogor(ID): Center For International Forestry Research

1. Departemen Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Pertanian Bogor, Bogor 16680

   \*Mahasiswa Program Studi S1 Ilmu Komputer, FMIPA-IPB; Surel: shofyanipb@gmail.com [↑](#footnote-ref-1)