



# PROSIDING

## SEMIRATA 2014 Bidang MIPA BKS-PTN-Barat

"Integrasi sains MIPA untuk mengatasi masalah pangan, energi, kesehatan, reklamasi, dan lingkungan"

IPB International Convention Center dan Kampus IPB Baranangsiang, 9-11 Mei 2014

**BUKU 6**

**MATEMATIKA, FISIKA, KIMIA, BIOLOGI,  
STATISTIKA, KOMPUTER, STEM,  
GEOFISIKA DAN METEOROLOGI**

Diterbitkan oleh: Fakultas Matematika dan Ilmu Pengetahuan Alam  
Institut Pertanian Bogor



ISBN 978-602-70491-0-9

## KLASIFIKASI DATA SPASIAL UNTUK KEMUNCULAN HOTSPOT DI PROVINSI RIAU MENGGUNAKAN ALGORITME ID3

### SPATIAL CLASSIFICATION FOR HOTSPOT OCCURRENCES IN RIAU PROVINCE USING ID3 ALGORITHM

Vikhy Fernando <sup>1\*</sup>, Imas Sukaesih Sitanggang <sup>2</sup>

Program Studi S1 Ilmu Komputer, Departemen Ilmu Komputer, FMIPA, Institut Pertanian Bogor, Bogor  
<sup>1</sup>; vikhyfernando10@gmail.com  
Kampus IPB Dramaga, Jl. Meranti Wing 20 Level V, Bogor, Indonesia 16680  
Departemen Ilmu Komputer, FMIPA, Institut Pertanian Bogor, Bogor <sup>2</sup>

#### ABSTRACT

Forest fires is one of environmental issues that occurs almost every year in Indonesia, including in Riau Province. It causes negative impacts for human life. Large spread hotspot data can be analyzed using spatial data mining techniques namely spatial decision tree algorithms. The algorithm results a spatial decision tree from which we can obtain classification rules. This work developed spatial decision trees using the ID3 algorithm. The highest accuracy of the tree is 70.80%. The classification model which consists 125 classification rules can be used to predict hotspots occurrence.

Keywords: Classification, Hotspot, ID3 Algorithm, Spatial decision tree

#### ABSTRAK

Kebakaran hutan merupakan salah satu persoalan lingkungan yang muncul hampir setiap tahun di Indonesia termasuk di Provinsi Riau. Dampak yang ditimbulkan dari kebakaran hutan sangat merugikan. Data persebaran hotspot berukuran besar sebagai indikator terjadinya kebakaran hutan dapat dianalisis menggunakan teknik *spatial data mining* salah satunya adalah pohon keputusan spasial. Dari pohon keputusan yang dibangun dari data spasial dapat dibentuk aturan-aturan klasifikasi. Dalam penelitian ini, pembentukan pohon keputusan spasial dilakukan dengan menggunakan algoritme ID3. Hasilnya adalah pohon keputusan dengan akurasi tertinggi 70.80%. Model klasifikasi ini yang terdiri dari 125 aturan klasifikasi dapat digunakan untuk prediksi kemunculan titik api.

Katakunci: algoritme ID3, hotspot, klasifikasi, pohon keputusan spasial

#### PENDAHULUAN

Kebakaran hutan merupakan salah satu persoalan lingkungan yang muncul hampir setiap tahun di Indonesia termasuk di Provinsi Riau. Padahal, dampak yang ditimbulkan dari kebakaran hutan sangat merugikan. Dampak tersebut tidak hanya dialami oleh masyarakat di Provinsi Riau saja, tetapi juga oleh masyarakat di Provinsi sekitar Riau yakni, Provinsi Sumatera Barat dan Provinsi Sumatera Utara. Selain dekat dengan Provinsi lain di Indonesia, Provinsi Riau juga dekat dengan wilayah negara tetangga, yaitu Singapura dan Malaysia.

Data persebaran *hotspot* yang berukuran besar dapat dianalisis menggunakan teknik *spatial data mining*. Salah satu teknik dalam *spatial data mining* adalah *spatial decision tree*. Hasil dari *spatial decision tree* akan membangun sebuah *decision tree* dari data spasial yang dapat digunakan untuk membentuk aturan-aturan klasifikasi. Klasifikasi spasial bertujuan memberikan sebuah label atau menentukan kelas dari sebuah objek berdasarkan nilai atribut yang ada dalam spasial dataset dengan memperhatikan objek tetangganya.

Tujuan dari penelitian ini untuk membuat model klasifikasi untuk data kebakaran hutan di wilayah Provinsi Riau pada tahun 2005 dan mengevaluasi model klasifikasi untuk prediksi kemunculan titik api. Penelitian ini diharapkan dapat membentuk model klasifikasi dari persebaran *hotspot* di Provinsi Riau pada tahun 2005 dan dapat memprediksi kemunculan titik api di wilayah baru sehingga dapat mengantisipasi kebakaran hutan di wilayah tersebut.

Penelitian ini menggunakan data *hotspot* di wilayah Indonesia yang bersumber dari Direktorat Pengendalian Kebakaran Hutan (DPKH) Departemen Kehutanan Republik Indonesia. Penelitian ini mengklasifikasikan data spasial persebaran *hotspot* di wilayah Riau pada tahun 2005 menggunakan algoritme ID3.

## METODE PENELITIAN

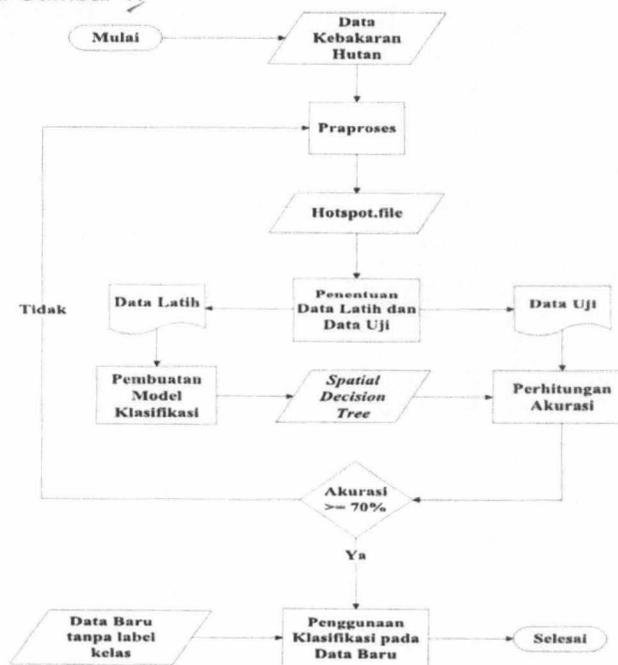
### Data Penelitian dan Area Studi

Luas wilayah Provinsi Riau adalah 107932.71 km<sup>2</sup> yang membentang dari lereng Bukit Barisan hingga Selat Malaka, hal ini membuat Provinsi Riau berada pada jalur yang sangat strategis karena terletak pada jalur perdagangan Regional dan Internasional di kawasan ASEAN. Provinsi Riau memiliki luas daratan 89150.15 km<sup>2</sup> dan luas lautan 18782.56 km<sup>2</sup>, di daratan terdapat 15 (lima belas) sungai di antaranya ada 4 (empat) sungai dapat digunakan sebagai prasarana perhubungan (Pemerintah Provinsi Riau 2013).

Data spasial yang digunakan dalam penelitian ini adalah data sebaran *hotspot* tahun 2005 yang berjumlah 7169 titik *hotspot*, 7200 titik non *hotspot*, 457 segmen aliran sungai, 37656 segmen jalan dan 6 titik pusat kota di wilayah Riau.

### Tahapan Penelitian

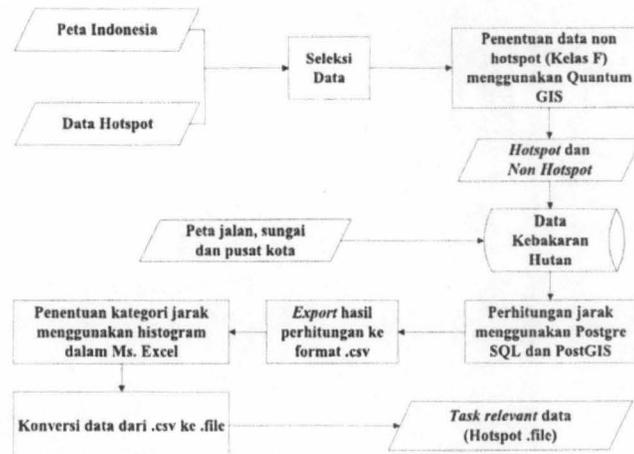
Penelitian ini akan dilakukan dalam beberapa tahap. Tahapan dari metode penelitian ini dapat dilihat pada Gambar 1.



Gambar 10 Tahapan penelitian

### Praproses

Data spasial diolah menggunakan perangkat lunak Quantum GIS 1.8.0 Lisboa untuk mendapatkan data hotspot, non hotspot, pusat kota, sungai dan jalan di Provinsi Riau pada tahun 2005. Selanjutnya, data tersebut di-import ke basis data menggunakan PostgreSQL 9.1 yang didalamnya telah terintegrasi dengan PostGIS 2.0. Kemudian dilakukan kueri pada data yang diperlukan hingga didapatkan data yang dibutuhkan untuk diklasifikasikan. Agar kueri sesuai dengan kode yang telah dibuat di perangkat lunak Python, maka digunakan Microsoft Excel 2010 sebagai perangkat lunak pembantu. Alur dari praproses yang dapat dilihat pada Gambar 2.



Gambar 11 Tahapan praproses

Beberapa tahapan yang dilakukan pada tahap praproses adalah:

1. Seleksi data, pada proses ini dilakukan pemilihan data hotspot dan data non hotspot yang terjadi di Provinsi Riau pada tahun 2005. Selanjutnya, data tersebut juga akan dipilih menjadi data pusat kota, sungai dan jalan di Provinsi Riau.
2. Pemberian label kelas, pemberian label kelas ini digunakan untuk mendapatkan hasil data. Label T digunakan pada data hotspot dan label F digunakan pada data non hotspot.
3. Perhitungan jarak, pada tahap ini dilakukan perhitungan jarak terdekat dari kelas T dan kelas F ke pusat kota, kelas T dan kelas F ke sungai dan kelas T dan kelas F ke jalan di Provinsi Riau.
4. Transformasi jarak terdekat dari hotspot, pada tahap ini dilakukan transformasi hasil data menjadi kategorik agar lebih memudahkan untuk diproses bahasa perograman Python.
5. Transformasi tipe fail, pada penelitian ini digunakan data dengan format .file sebagai fail masukan data latih dan data uji.

#### Klasifikasi Menggunakan Algoritme ID3

Pada tahap ini dibangun model klasifikasi untuk membentuk pohon keputusan spasial. Kemudian dilakukan perhitungan akurasi dari pohon keputusan spasial yang terbentuk. Pembentukan pohon keputusan spasial ini menggunakan algoritme ID3. Algoritme ini menggunakan ukuran entropi informasi dalam membuat pohon keputusan.

#### Pohon Keputusan Spasial

Proses pembentukan pohon keputusan spasial dari data sebaran hotspot di Provinsi Riau menggunakan tabel gabungan yang terdiri dari beberapa hubungan spasial yang terbentuk dari kelas targetnya.

### Perhitungan Akurasi

Pada tahap ini menghitung akurasi dari model klasifikasi yang diperoleh dari proses klasifikasi. Akurasi menunjukkan tingkat kebenaran pengklasifikasian data terhadap kelas yang sebenarnya. Semakin rendah nilai akurasi maka semakin tinggi kesalahan klasifikasi pada data baru. Tingkat akurasi yang baik adalah tingkat akurasi yang mendekati 100%. Dalam penelitian, metode yang digunakan dalam proses perhitungan akurasi ini adalah metode *K-fold cross validation*. *K-fold cross validation* (Stone 1974 diacu dalam Fu 1994) adalah sebuah metode yang membagi himpunan contoh secara acak menjadi *k* himpunan bagian (*subset*). Pada metode ini dilakukan pengulangan sebanyak *k* kali untuk data pelatihan dan pengujian. Pada setiap pengulangan, satu *subset* digunakan untuk pengujian sedangkan *subset* sisanya digunakan untuk pelatihan.

Akurasi diperoleh berdasarkan data pengujian terhadap model klasifikasi. Untuk menghitung akurasi digunakan rumus sebagai berikut:

$$\text{akurasi} = \frac{\sum \text{data uji benar diklasifikasi}}{\sum \text{total data uji}}$$

### Lingkungan Pengembangan

Pembentukan pohon keputusan spasial ini dibangun dengan menggunakan perangkat keras berupa komputer personal dengan spesifikasi antara lain *Processor Intel® Core™ I3* CPU, memori 2048MB RAM, *mouse*, *keyboard* dan perangkat lunak yaitu Sistem Operasi Windows 8, Quantum GIS (1.8.0) untuk analisis dan visualisasi data spasial, Sistem Manajemen Basis Data PostgreSQL, PostGIS 2.0 sebagai ekstensi PostgreSQL untuk analisis data spasial dan bahasa Pemrograman Python 2.7.5.

## HASIL DAN PEMBAHASAN

### Pra proses

#### Seleksi Data

Data yang digunakan dalam penelitian ini merupakan data persebaran *hotspot* pada tahun 2005, data *non hotspot*, data pusat kota, data sungai dan data jalan untuk Provinsi Riau. Secara terperinci dapat dilihat sebagai berikut

#### a. Data persebaran *Hotspot* dan *Non Hotspot*

Data persebaran *hotspot* yang digunakan adalah data *hotspot* tahun 2005. Data ini memiliki atribut *wkt\_geom*, *longitude*, *latitude*, bulan, minggu dan tanggal. Pada awalnya, data persebaran *hotspot* Indonesia berbentuk *shapefile*. Data tersebut di-*clip* menjadi daerah Provinsi Riau saja menggunakan perangkat lunak Quantum GIS. Pada wilayah Provinsi Riau terdapat 7169 titik *hotspot*. Sedangkan, data persebaran *non hotspot* diperoleh dari titik-titik di luar *hotspot*. Data ini didapat dengan cara membangkitkan titik acak di sekitar *hotspot*. Titik non *hotspot* tersebut dibangkitkan di luar *buffer* dengan radius 1 km dari sebuah *hotspot*. *Buffer* dibuat menggunakan operasi *Geo processing Tools* pada Quantum GIS. Data *non hotspot* yang dibangkitkan sebanyak 7200 titik.

#### b. Data Pusat kota Riau, merupakan data yang berisi mengenai ibu kota tiap kota/kabupaten wilayah Riau. Data ini berbentuk *shapefile* yang diolah menggunakan Quantum GIS. Data pusat kota Riau ini memiliki 6 pusat kota yaitu Tembilahan, Selatpanjang, Balai punggut, Duri, Bengkalis dan Dumai.

- c. Data Sungai Riau, merupakan data yang berisi mengenai wilayah Riau yang dilalui oleh aliran sungai. Data ini berbentuk shapefile yang diolah menggunakan Quantum GIS. Terdapat 457 segmen aliran sungai di data sungai Riau.
- d. Data Jalan Riau, data yang berisi seluruh jalan di wilayah Riau. Data ini berbentuk shapefile yang diolah menggunakan Quantum GIS. Data jalan Riau memiliki sebanyak 37657 segmen jalan.

#### Pemberian Label Kelas

Pemberian label kelas ini digunakan untuk mendapatkan hasil data. Setelah semua *shape* di-*import* ke Sistem Manajemen Basis Data postgresQL menjadi beberapa tabel dilakukan pemberian label kelas dengan kueri. Label T (*True*) digunakan pada data hotspot yang berjumlah 7169 titik hotspot dan label F (*False*) digunakan pada data non hotspot yang berjumlah 7200 titik non hotspot. Tabel tersebut diberi nama tabel hotspot.

#### Perhitungan Jarak

Pada tahap ini dilakukan perhitungan jarak dari kelas T dan kelas F ke pusat kota, kelas T dan kelas F ke sungai serta kelas T dan kelas F ke jalan di Provinsi Riau. Setelah setiap tabel diberikan label kelas *True* (T) dan *False* (F) dilakukan perhitungan jarak menggunakan kueri di PostgreSQL. Data sebaran hotspot, non hotspot, pusat kota, sungai dan jalan awalnya diolah di Quantum GIS berbentuk *shapefile*. Agar PostgreSQL dapat membaca *shapefile* yang berbentuk data spasial harus menggunakan PostGIS yang terintegrasi dengan PostgreSQL. Data tersebut di *import* dari Quantum GIS untuk menjadi tabel-tabel di PostgreSQL.

Perhitungan jarak hotspot dan non hotspot ke pusat kota terdekat, ke jalan terdekat dan ke sungai terdekat dilakukan dengan mengeksekusi pernyataan kueri spasial dalam PostgreSQL. *Dataset* jarak harus dikonversi dari *decimal degree* menjadi km untuk menyiapkan dataset akhir (1 degree = 111319.5 m = 111.3195 km). Jarak terdekat beberapa hotspot ke pusat kota, ke sungai dan ke jalan diberikan dalam Tabel 1.

Tabel 7 Jarak terdekat hotspot ke pusat kota, ke sungai dan ke jalan

Id_Hotspot	Jarakterdekat_Pusatkota (km)	Jarakterdekat_Sungai (km)	Jarakterdekat_Jalan (km)	Kelas
1	39.0253843	15.00513534	1.260659936	T
2	39.58416374	18.50195349	1.308696992	T
3	10.50838916	25.17563015	0.019610453	T
.	.	.	.	.
9999	77.74134721	5.021882999	0.681320262	F
10000	107.922778	8.049548723	8.614583682	F

#### Transformasi Jarak Terdekat dari Hotspot

Pada tahap ini dilakukan transformasi data numerik menjadi data kategorik menggunakan histogram dari variabel jarak. Tahapan transformasi ini menggunakan perangkat lunak Microsoft Excel 2010 dengan menggunakan *tools data analyst histogram*. Kategori untuk jarak terdekat ke pusat kota, ke sungai dan ke jalan diberikan berturut-turut dalam Tabel 2, 3 dan 4.

#### Transformasi Tipe Fail

Dalam tahap ini dilakukan perubahan format data latih dan data uji dari format CSV menjadi format *.file* sehingga dapat diolah ke dalam perangkat lunak Python 2.7.5.

Tabel 2 Kategori untuk jarak terdekat ke pusat kota

Bin	Frequency	Kategori	Range (km)
31	2868	M1	(0, 31]
48	2832	M2	(31, 48]
74	2947	M3	(48, 74]
105	2868	M4	(74, 105]
More	2854	M5	> 105

Tabel 3 Kategori untuk jarak terdekat ke sungai

Bin	Frequency	Kategori	Range (km)
2	2913	RI1	(0, 2]
4.5	2823	RI2	(2, 4.5]
8.5	2939	RI3	(4.5, 8.5]
16	2873	RI4	(8.5, 16]
More	2821	RI5	> 16

Tabel 4 Kategori untuk jarak terdekat ke jalan

Bin	Frequency	Kategori	Range(km)
0.22	2910	RO1	(0, 0.22]
0.68	2872	RO2	(0.22, 0.68]
1.5	2877	RO3	(0.68, 1.5]
3.5	2898	RO4	(1.5, 3.5]
More	2812	RO5	> 3.5

### Penentuan Data Latih dan Data Uji

Pemisahan data ke dalam data latih dan data uji untuk membentuk pohon keputusan spasial ini menggunakan *10-folds cross validation*. Data akan dibagi ke dalam 10 bagian (*folds*). Setiap bagian akan digunakan sebagai data uji, 9 bagian (*folds*) akan dijadikan sebagai data latih, dan 1 bagian (*fold*) lainnya akan dijadikan sebagai data uji. Data latih akan digunakan untuk membentuk model klasifikasi. Sedangkan, data uji akan digunakan untuk menghitung akurasi yang diperoleh dari model klasifikasi.

Setelah tahap transformasi tipe fail dilakukan, tahap selanjutnya adalah membagi data menjadi 10 *fold*. *Fold* inilah yang akan digunakan pada tahap klasifikasi sebagai data latih dan data uji. Subset yang terbentuk memiliki jumlah *instance* sebanyak 1000 per *fold* dengan mengabaikan proporsi perbandingan antarkelas.

### Klasifikasi Menggunakan Algoritme ID3

Proses klasifikasi dilakukan dua tahap, yaitu pembentukan pohon keputusan spasial dan perhitungan akurasi dari pohon keputusan spasial yang terbentuk. Pembentukan pohon keputusan spasial dilakukan dengan menggunakan algoritme ID3 yang telah diimplementasikan dalam bahasa pemrograman Python 2.7.5. Modul ID3 dalam Python diperoleh dari Onlamp (2013). Dalam pembuatan pohon keputusan spasial di pemrograman Python ini terdapat 3 modul penting antara lain *decision tree*, ID3 dan modul test. Pada modul *decision tree* adalah proses untuk mendapatkan atribut terbaik dari himpunan atribut yang ada dengan menggunakan fungsi *fitness* yaitu *information gain* sebagai parameter. *Dataset* yang telah terbentuk diolah dan menghasilkan sebuah pohon keputusan. Atribut yang dipilih sebagai atribut uji adalah atribut yang memiliki *information gain* tertinggi sesuai dengan algoritme ID3.

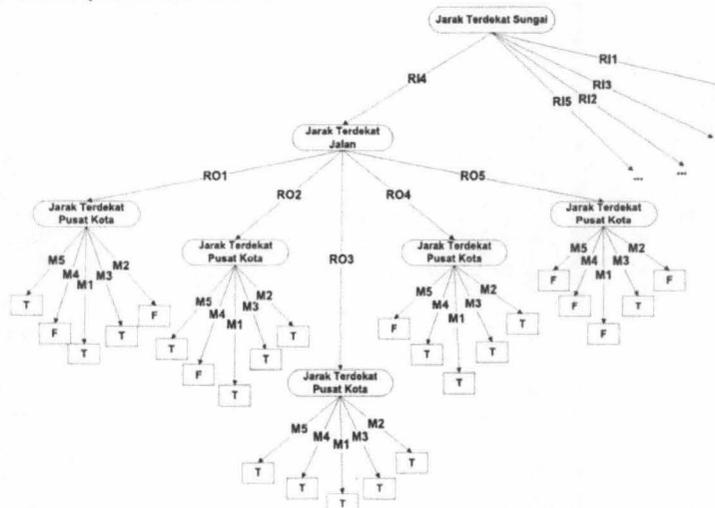
Pada pohon keputusan spasial yang dihasilkan diketahui bahwa semua atribut yang diberikan muncul sebagai *node* pada pohon keputusan spasial. Atribut yang dijadikan label pada akar pohon keputusan adalah jarak terdekat ke sungai. Pembagian *fold* dan akurasi pohon keputusan spasial untuk *data training* dan *data testing* dapat dilihat pada Tabel 5.

Tabel 5 Pembagian *fold* dan akurasi pohon keputusan

Percobaan	Data Training	Data Testing	Pohon Keputusan	Akurasi Data Training (%)	Akurasi Data Testing (%)
1	Fold1 sd fold 9	Fold 10	Tree 1	72.30	66.80
2	fold1 sd fold 8, dan Fold 10	Fold 9	Tree 2	73.80	70.20
3	fold1 sd fold 7, dan Fold 9 sd fold10	Fold 8	Tree 3	72.60	67.90
4	Fold1 sd fold 6, dan fold 8	Fold 7	Tree 4	75.00	70.80

sd fold10						
5	Fold1 sd fold 5, dan fold 7 sd fold10	Fold 6	Tree 5	73.90	69.80	
6	Fold1 sd fold 4, dan fold 6 sd fold10	Fold 5	Tree 6	74.10	70.70	
7	Fold1 sd fold 3, dan fold 5 sd fold10	Fold 4	Tree 7	71.10	67.10	
8	Fold1 sd fold 2, dan fold 4 sd fold10	Fold 3	Tree 8	73.80	69.90	
9	Fold1, dan fold 3 sd fold10	Fold 2	Tree 9	71.60	66.20	
10	Fold2 sd fold 10	Fold 1	Tree 10	72.50	69.10	
<b>Akurasi Rata-Rata</b>				<b>73.07</b>	<b>68.85</b>	

Dari Tabel 5 dapat dilihat bahwa akurasi terbesar terdapat pada percobaan keempat dan akurasi terkecil terdapat pada percobaan pertama dan kesembilan. Hal itu menunjukkan bahwa tingkat kesalahan klasifikasi *subset* 1 dan 9 lebih tinggi dari *subset* yang lainnya. Hal itu dapat disebabkan oleh *instance-instance* yang ada pada data latih dan data uji belum dapat mewakili setiap fitur untuk dapat diklasifikasikan pada kelas tertentu. Akurasi rata-rata untuk *data training* adalah 73.07% dan 68.85% untuk *data testing*. Berikut sub pohon keputusan spasial yang terbentuk untuk jarak terdekat ke sungai dengan interval (8.5 km, 16 km] yang dapat dilihat pada Gambar 3.



Gambar 3 Sub pohon keputusan spasial untuk jarak terdekat ke sungai dengan interval (8.5 km, 16 km]

Berdasarkan pohon keputusan spasial yang terbentuk dapat dibuat 125 aturan klasifikasi. Sebagai contoh, aturan yang terbentuk dari pohon keputusan spasial pada Gambar 3 adalah sebagai berikut:

- Aturan 1: JIKA jarak lokasi ke sungai terdekat pada interval (2 km, 4.5 km] DAN jarak lokasi ke pusat kota terdekat pada interval (48 km, 74 km] DAN jarak lokasi ke jalan terdekat pada interval (0.22 km, 0.68 km] MAKA kemunculan titik api adalah T.
- Aturan 2: JIKA jarak lokasi ke sungai terdekat pada interval (0 km, 2 km] DAN jarak lokasi ke jalan terdekat pada interval (0.22 km, 0.68 km] DAN jarak lokasi ke pusat kota terdekat pada interval (48 km, 74 km] MAKA kemunculan titik api adalah F.
- Aturan 3: JIKA jarak lokasi ke sungai terdekat pada interval (8.5 km, 16 km] DAN jarak lokasi ke pusat kota terdekat pada interval (0 km, 31 km] DAN jarak lokasi ke jalan terdekat pada interval (1.5 km, 3.5 km] MAKA kemunculan titik api adalah T.

- Aturan 4: JIKA jarak lokasi ke sungai terdekat pada interval (2 km, 4.5 km] DAN jarak lokasi ke pusat kota terdekat pada interval (74 km, 105 km] DAN jarak lokasi ke jalan terdekat pada interval (0 km, 0.22 km] MAKA kemunculan titik api adalah F.
- Aturan 5: JIKA jarak lokasi ke sungai terdekat pada interval > 16 km DAN jarak lokasi ke pusat kota terdekat pada interval (31 km, 48 km] DAN jarak lokasi ke jalan terdekat pada interval > 3.5 km MAKA kemunculan titik api adalah T.

### Akurasi Model

*Confusion matrix* dari model yang dihasilkan dapat dilihat pada Tabel 6.

Tabel 6 *Confusion matrix* dari model pohon keputusan

Kelas Hasil Aktual	Kelas Hasil Prediksi		Total
	T	F	
T	324	126	450
F	166	384	550
Total	490	510	1000

Dari Tabel 6 diketahui dari 450 data di kelas hasil aktual T, sebanyak 324 atau 72% data diprediksi secara benar sebagai label kelas T (*True Positif*) dan 126 data diprediksi salah sebagai label kelas F (*False Negative*). Dari 550 data di kelas hasil aktual F, sebanyak 384 data diprediksi benar sebagai label kelas F (*True Negative*) atau 66.12% dan 166 data diprediksi salah sebagai label kelas F (*False Negative*). Akurasi rata-ratanya sebesar 70.91%.

## SIMPULAN DAN SARAN

### Simpulan

Berdasarkan penelitian yang dilakukan dalam membentuk model klasifikasi hotspot di wilayah Riau pada tahun 2005, dapat diambil kesimpulan antara lain pohon keputusan spasial yang telah dibentuk memiliki label akar yaitu jarak terdekat sungai. Pohon keputusan spasial ini terdiri dari 31 *node*, dari pohon keputusan spasial yang dihasilkan dapat diturunkan 125 aturan klasifikasi untuk menentukan kemunculan hotspot, pohon keputusan spasial telah diuji pada *data training* dan *data testing* dengan menggunakan metode *10-fold cross validation*. Akurasi tertinggi dari pohon keputusan spasial pada *data training* adalah 75% , sedangkan pada *data testing* adalah 70.80% dan akurasi rata-rata pada data training adalah 73.07% dan pada data testing adalah 68.85%.

### Saran

Saran-saran yang dapat diberikan untuk pengembangan lebih lanjut adalah penambahan data spasial lainnya terkait kebakaran hutan yang menentukan kemunculan titik api, sehingga pohon keputusan yang dihasilkan lebih akurat.

## DAFTAR PUSTAKA

- [1] Fu L. 1994. *Neural Network in Computers Intelligence*. Singapura: McGraw-Hill.
- [2] Han J, Kamber M. 2006. *Data Mining: Concepts and Techniques*. San Francisco: Morgan Kaufman Publisher.
- [3] Onlamp. 2013. Onlamp [Internet]. [diunduh 2013 Dec 1]. Tersedia pada [http://www.onlamp.com/pub/a/python/2006/02/09/ai\\_decision\\_trees.html?page=2](http://www.onlamp.com/pub/a/python/2006/02/09/ai_decision_trees.html?page=2)
- [4] Pemerintah Provinsi Riau. 2013. Pemerintah Provinsi Riau [Internet]. [diunduh 2013 Dec 1]. Tersedia pada: <http://www.riau.go.id/index.php?/detail/61>.