

APLIKASI KLASIFIKASI BERBASIS WEB UNTUK DATA KEBAKARAN HUTAN MENGGUNAKAN FRAMEWORK **SHINY DAN ALGORITME C5.0**

GITA PUSPITA SIKNUN



DEPARTEMEN ILMU KOMPUTER FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM **INSTITUT PERTANIAN BOGOR BOGOR** 2015





IPB University



PERNYATAAN MENGENAI SKRIPSI DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa skripsi berjudul Aplikasi Klasifikasi Berbasis Web untuk Data Kebakaran Hutan Menggunakan *Framework* Shiny dan Algoritme C5.0 adalah benar karya saya dengan arahan dari komisi pembimbing dan belum diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir skripsi ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Institut Pertanian Bogor.

Bogor, Juni 2015

Gita Puspita Siknun NRP G64110015

ABSTRAK

GITA PUSPITA SIKNUN. Aplikasi Klasifikasi Berbasis Web untuk Data Kebakaran Hutan Menggunakan *Framework* Shiny dan Algoritme C5.0. Dibimbing oleh IMAS SUKAESIH SITANGGANG.

Kebakaran hutan merupakan ancaman bagi ekosistem dan lingkungan karena dampak yang dihasilkan sangat merugikan. Setiap tahun terjadi benambahan titik panas sebagai indikator kebakaran hutan di wilayah Indonesia, salah satunya Provinsi Riau. Untuk memprediksi kemunculan titik panas, dibangun sebuah aplikasi berbasis web berdasarkan karakteristik wilayah menggunakan framework Shiny. Algoritme klasifikasi yang digunakan adalah algoritme C5.0 yang menghasilkan model klasifikasi berupa pohon keputusan dan model berbasis aturan. Framework Shiny dibangun dengan fungsi reactivity yang ketika terjadi perubahan *input*, server akan menghasilkan *output* berdasarkan perubahan tersebut. Klasifikasi dilakukan pada dataset kebakaran hutan di wilayah Rokan Hilir di Provinsi Riau pada tahun 2008. Dataset terdiri atas sepuluh variabel penjelas (data fisik, data cuaca, dan data sosial-ekonomi) dan satu variabel target (data titik panas dan data bukan titik panas). Penerapan algoritme C5.0 pada *dataset* kebakaran hutan menghasilkan model pohon keputusan dengan akurasi 72.72% dan model berbasis aturan dengan akurasi 73.51%. Dari model pohon keputusan diperoleh 16 aturan klasifikasi sedangkan model berbasis aturan terdiri atas 15 aturan klasifikasi.

Kata kunci: C5.0, framework Shiny, kebakaran hutan, klasifikasi, titik panas

ABSTRACT

GITA PUSPITA SIKNUN. Web-Based Classification Application for Forest Fire Data Using the Shiny Framework and the C5.0 Algorithm. Supervised by IMAS SUKAESIH SITANGGANG.

Forest fires are threats for our ecosystems and environment because their impact is very harmful. Every year, the number of hotspots increases, indicating the increase of forest fires in some regions in Indonesia, one of them is Riau Province. To predict the hotspot occurrence, we build a web-based application based on characteristics of area using the framework Shiny. We use the C5.0 algorithm by generating tree and rule-based classification models. The Shiny framework is implemented using reactivity expression, when an input changes, the server will rebuild the output based on the data input. We use the dataset of forest fires for Rokan Hilir district in Riau Province in 2008. The dataset consists of ten explanatory layers (physical, weather, and socio-economic characteristics) and one target layer (hotspot and non hotspot points). The implementation of the C5.0 algorithm on forest fire data yields tree models with accuracy of 72.72% and rule-based models with accuracy of 73.51%. The tree models output 16 classification rules while the rule-based models output 15 classification rules.

Keywords: C5.0, classification, forest fires, framework Shiny, hotspots



APLIKASI KLASIFIKASI BERBASIS WEB UNTUK DATA KEBAKARAN HUTAN MENGGUNAKAN FRAMEWORK SHINY DAN ALGORITME C5.0

GITA PUSPITA SIKNUN

Skripsi sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer pada Departemen Ilmu Komputer

DEPARTEMEN ILMU KOMPUTER
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT PERTANIAN BOGOR
BOGOR
2015

Penguji:
1 Hari Agung Adrianto, SKom MSi
2 Rina Trisminingsih, SKom MT



Judul Skripsi: Aplikasi Klasifikasi Berbasis Web untuk Data Kebakaran Hutan

Menggunakan Framework Shiny dan Algoritme C5.0

Nama

: Gita Puspita Siknun

NIM

: G64110015

Disetujui oleh

Dr Imas Sukaesih Sitanggang, SSi MKom Pembimbing

Diketahui oleh

Dr Ir Agus Buono, MSi MKom Ketua Departemen

Tanggal Lulus: 2 4 JUL 2015

PRAKATA

Puji dan syukur penulis panjatkan kepada Allah subhanahu wa ta'ala atas segala karunia-Nya sehingga karya ilmiah ini berhasil diselesaikan. Tema yang dipilih dalam penelitian yang dilaksanakan sejak bulan September 2014 ini ialah aplikasi berbasis web, dengan judul Aplikasi Klasifikasi Berbasis Web untuk Data Kebakaran Hutan Menggunakan Framework Shiny dan Algoritme C5.0. Penulisan skripsi ini merupakan salah satu syarat untuk memperoleh gelar Sarjana Komputer pada Departemen Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Pertanian Bogor.

Penulis mengucapkan terima kasih kepada semua pihak yang telah berperan dalam penelitian ini, di antaranya:

- Ibu, Ayah, Kakak, kak Isnen dan keluarga lainnya yang telah memberikan dukungan, doa, motivasi, dan semangat untuk keberhasilan studi
- 2 Ibu Dr Imas Sukaesih Sitanggang, SSi MKom selaku dosen pembimbing yang telah memberikan bimbingan, saran, arahan, dan bantuan selama penyelesaian skripsi
- 3 Sahabat-sahabat terbaik penulis, Pha Hy Thah, Nida Zakiya Nurulhaq, serta teman-teman satu bimbingan, terima kasih atas doa, semangat, dan bantuannya
- 4 Teman-teman Core-i48 yang selalu memberikan semangat serta kenangan selama menjalani studi
- 5 Seluruh dosen dan civitas akademika Departemen Ilmu Komputer IPB

Penulis berharap semoga skripsi ini bermanfaat bagi pihak yang membutuhkan.

Bogor, Juni 2015

Gita Puspita Siknun



DAFTAR ISI

DAFTAR TABEL	vi
DAFTAR GAMBAR	vi
DAFTAR LAMPIRAN	vi
PENDAHULUAN	1
Latar Belakang	1
Perumusan Masalah	2
Tujuan Penelitian	2
Manfaat Penelitian	3
Ruang Lingkup Penelitian	3
METODE	3
Data Penelitian	3
Tahapan Penelitian	4
Lingkungan Pengembangan	7
HASIL DAN PEMBAHASAN	8
Klasifikasi Data Menggunakan Algoritme C5.0	8
Implementasi Aplikasi Berbasis Web Menggunakan Shiny	10
Prediksi Titik Panas	16
SIMPULAN DAN SARAN	18
Simpulan	18
Saran	18
DAFTAR PUSTAKA	19
LAMPIRAN	21



DAFTAR TABEL

berbasis aturan Nilai <i>default</i> variabel penjelas	10 17
DAFTAR GAMBAR	
1 Tahapan penelitian	4
2 Algoritme pohon keputusan C5.0	5
3 Aturan klasifikasi dari model pohon keputusan	9
4 Aturan klasifikasi dari model berbasis aturan	10
5 Proses menampilkan model klasifikasi	12
6 Proses menampilkan hasil prediksi titik panas	13
7 Tampilan halaman utama aplikasi berbasis web	14
8 Hasil klasifikasi dengan model pohon keputusan	15
9 Tampilan hasil prediksi pada tabPanel predict new hotspot	15
10 Hasil prediksi kemunculan titik panas kelas <i>hotspot</i>	16
11 Hasil prediksi kemunculan titik panas kelas <i>non-hotspot</i>	17
12 Tampilan halaman <i>tabPanel help</i>	18
DAFTAR LAMPIRAN	
1 Variabel penjelas yang digunakan pada <i>dataset</i> kebakaran hutan	
(Sitanggang et al. 2014)	21
2 Model klasifikasi pohon keputusan dengan algoritme C4.5	
menggunakan perangkat lunak Weka	23
3 Model klasifikasi pohon keputusan dengan algoritme C5.0	2.5
menggunakan perangkat lunak R	26
4 Model klasifikasi berbasis aturan dengan algoritme C5.0 menggunakan	20
perangkat lunak R	28
5 Tampilan ringkasan data dan tabel pada aplikasi berbasis web	31 32
6 Hasil klasifikasi dengan model berbasis aturan 7 Tampilan <i>tah Panal kala</i> pada aplikasi berbasis wah	33
7 Tampilan <i>tabPanel help</i> pada aplikasi berbasis web8 Hasil pengujian fungsi aplikasi berbasis web	34
o masii pengujian tungsi apiikasi berbasis web	34

PENDAHULUAN

Latar Belakang

Kebakaran hutan di Indonesia dipandang sebagai ancaman bagi ekosistem dan lingkungan karena dampak yang dihasilkan sangat merugikan. Menurut studi yang dilakukan Bappenas dan ADB (1999) mengenai kebakaran hutan, luas lahan yang terbakar selama peristiwa bencana El Nino (ENSO) pada tahun 1997/98 adalah sekitar 9.75 juta ha (Tacconi 2003). Faktor pemicu parahnya kebakaran hutan adalah kemarau yang ekstrim pada musim kering ekstrem El-Nino dan/atau penggalian drainase lahan gambut secara berlebihan (Agus dan Subika 2012). Penggalian drainase lahan gambut merupakan faktor kesengajaan terjadinya kebakaran hutan.

Pencegahan terjadinya kebakaran hutan dilakukan dengan pemantauan titik panas melalui satelit penginderaan jauh dan perangkat sistem informasi geografis (SIG) untuk mendapatkan data lokasi terjadinya titik panas. Menurut juru kampanye hutan Greenpeace Asia Tenggara, Yuyun Indradi, 375 dari 11288 titik panas yang terletak di Indonesia pada tahun 2014 dapat ditemukan di lahan gambut (Riswandy 2014). Sebagian besar titik panas yang terjadi di lahan gambut tersebar di sejumlah wilayah Kabupaten Rokan Hilir, Provinsi Riau. Menurut Kepala Badan Penanggulangan Bencana Daerah Provinsi Riau (BPBD Riau), titik panas terbanyak terletak di Kabupaten Rokan Hilir pada tahun 2014, yaitu sebanyak 85 titik panas yang tersebar di sejumlah kecamatan (Ali 2014).

Kebakaran hutan dapat mempengaruhi kondisi lingkungan, ekonomi, dan kesehatan. Pengaruh kondisi lingkungan seperti peningkatan emisi karbon telah menyebabkan peningkatan efek rumah kaca maupun pemanasan bumi secara global. Selain itu, terjadi masalah kesehatan dengan adanya kabut asap kebakaran. Peristiwa kabut asap merupakan pencemaran yang menimbulkan dampak seperti kematian, efek subkronis pada paru-paru, dan penyakit pernapasan. Oleh karena itu, untuk mengurangi kejadian kebakaran hutan perlu dibangun sebuah sistem informasi yang menggunakan teknik data mining. Data mining mampu mengolah dan menganalisis data berukuran besar, salah satunya data kebakaran hutan.

Salah satu teknik data mining ialah klasifikasi. Klasifikasi ialah proses menemukan model atau fungsi yang menggambarkan dan menentukan kelas data. Model diturunkan berdasarkan pada analisis himpunan data latih (objek data yang label kelasnya telah diketahui) dan digunakan untuk memprediksi label kelas yang belum diketahui (Han et al. 2012). Pada penelitian yang dilakukan oleh Sitanggang et al. (2014), dikembangkan model klasifikasi untuk kemunculan titik panas pada lahan gambut di Provinsi Riau dengan melibatkan data spasial berupa titik, garis, dan poligon. Algoritme klasifikasi yang dikembangkan adalah spatial decision tree. Pada penelitian tersebut, algoritme spatial decision tree dibandingkan dengan algoritme ID3 dan C4.5 yang menghasilkan model klasifikasi pada data bukan spasial. Algoritme ID3 dan C4.5 yang digunakan telah tersedia pada aplikasi data mining Weka berbasis desktop. Penelitian Sitanggang et al. (2014) menghasilkan akurasi dari algoritme spatial decision tree sebesar 71.66%. Algoritme spatial decision tree memiliki akurasi yang lebih besar dibandingkan algoritme ID3 dan C4.5. Algoritme ID3 memiliki akurasi sebesar

49.02% dan C4.5 memiliki akurasi sebesar 65.24%. Pada tahun 1980, algoritme C4.5 diperkenalkan oleh JR Quinlan sebagai pengembangan dari algoritme ID3. Kemudian JR Quinlan mengembangkan model klasifikasi berstruktur pohon (*tree*) dan berbasis aturan (*rule-based*) ke dalam algoritme C5.0 dengan hasil pohon keputusan yang lebih kecil serta penggunaan memori yang lebih efisien dari algoritme C4.5. Algoritme C5.0 diperkenalkan secara komersial sampai tahun 2011 ketika versi GPL dihadirkan.

Pada penelitian Sitanggang et al. (2014), model klasifikasi belum diimplementasikan ke dalam aplikasi berbasis web yang dapat diakses dengan mudah tanpa harus melakukan instalasi program. Selain itu, model klasifikasi belum diimplementasikan untuk memprediksi kemunculan titik panas sehingga model klasifikasi tersebut belum dapat digunakan oleh pihak terkait pencegahan kebakaran hutan. Oleh karena itu, pada penelitian ini akan dikembangkan aplikasi berbasis web untuk menampilkan model klasifikasi yang bisa digunakan oleh pengguna tanpa proses instalasi program. Pengembangan aplikasi berbasis web dapat dilakukan dengan menggunakan Shiny. Shiny merupakan package R sebagai framework yang responsif untuk menampilkan hasil analisis data mining ke dalam aplikasi berbasis web. Aplikasi berbasis web ini dikembangkan dengan menggunakan bahasa pemrograman R dengan package yang telah tersedia. Pada package R, terdapat banyak algoritme klasifikasi yang dapat digunakan dengan mudah. Pada penelitian ini, algoritme klasifikasi yang diterapkan adalah algoritme C5.0.

Perumusan Masalah

Memprediksi kemunculan titik panas dengan membangun model klasifikasi merupakan hal yang penting sehingga peristiwa kebakaran hutan dapat dikurangi. Model klasifikasi digunakan untuk menentukan kemunculan titik panas atau bukan titik panas. Kemunculan titik panas bisa disebabkan oleh banyak faktor seperti faktor fisik, cuaca, dan sosial-ekonomi. Faktor fisik bisa terdiri dari jalan, sungai, dan tutupan lahan. Faktor cuaca bisa terdiri dari curah hujan, kecepatan angin, dan suhu. Salah satu faktor sosial-ekonomi adalah sumber pendapatan. Ketiga faktor tersebut direpresentasikan sebagai variabel penjelas. *Dataset* yang digunakan pada penelitian ini merupakan *dataset* kebakaran hutan yang terdiri dari variabel penjelas dan variabel target.

Model klasifikasi telah banyak dikembangkan pada penelitian-penelitian sebelumnya dan dapat digunakan untuk menentukan kemunculan titik panas berdasarkan karakteristik wilayah. Namun, model klasifikasi tersebut belum bisa dimanfaatkan oleh pihak terkait pencegahan kebakaran hutan. Rumusan masalah dari penelitian ini adalah bagaimana membangun dan menerapkan model klasifikasi pada aplikasi berbasis web dengan *framework* Shiny.

Tujuan Penelitian

Penelitian ini bertujuan untuk membangun model klasifikasi pada *dataset* kebakaran hutan dengan menggunakan algoritme C5.0 dan menerapkannya pada

aplikasi berbasis web menggunakan *framework* Shiny untuk memprediksi kemunculan titik panas.

Manfaat Penelitian

Manfaat dari penelitian ini adalah memberikan hasil prediksi kemunculan titik panas berdasarkan karakteristik wilayah yang dapat digunakan oleh pemerintah dan peneliti bidang kehutanan. Selain itu juga diharapkan aplikasi web yang dibangun dapat diakses melalui web browser secara online. Sehingga aplikasi ini dapat digunakan dalam pengambilan keputusan untuk mencegah kebakaran hutan.

Ruang Lingkup Penelitian

Ruang lingkup penelitian ini meliputi:

- 1 Data yang digunakan adalah *dataset* kebakaran hutan wilayah Rokan Hilir Provinsi Riau di Indonesia tahun 2008.
- 2 *Dataset* kebakaran hutan telah dilakukan praproses pada penelitian sebelumnya oleh Sitanggang *et al.* (2014).
- Pengembangan aplikasi berbasis web menggunakan *framework* Shiny dan bahasa pemrograman R.

METODE

Data Penelitian

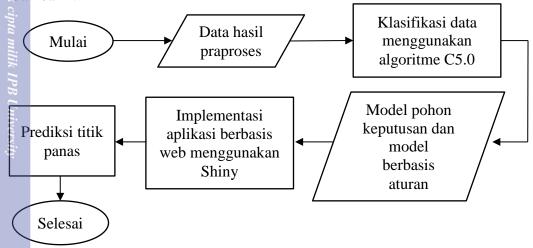
Data yang digunakan dalam penelitian ini adalah *dataset* kebakaran hutan wilayah Rokan Hilir Provinsi Riau pada tahun 2008. *Dataset* kebakaran hutan ini merupakan data dari penelitian Sitanggang *et al.* (2014). *Dataset* terdiri dari sepuluh variabel penjelas (sungai, jalan, pusat kota, sumber pendapatan, curah hujan dalam mm/hari, *screen temperature* dalam K, tutupan lahan, kedalaman lahan gambut, kelajuan angin dalam m/s dan jenis lahan gambut) dan variabel target. Variabel target terdiri dari data titik panas dan data bukan titik panas. Nama dan nilai variabel target dan variabel penjelas dapat dilihat pada Lampiran 1.

Data penelitian terdiri dari data fisik, data sosial-ekonomi, dan data cuaca. Data fisik seperti dari data persebaran dan koordinat titik panas, didapatkan dari FIRMS MODIS *hotspot*, NASA/University of Maryland. Data fisik untuk jalan, sungai, pusat kota, tutupan lahan, dan batas administrasi disediakan dalam bentuk peta digital. Data tersebut didapatkan dari Badan Koordinasi Survei dan Pemetaan Nasional (Bakosurtanal). Sedangkan data fisik untuk kedalaman lahan gambut dan jenis lahan gambut didapatkan dari Wetland International dalam bentuk peta digital. Data sosial-ekonomi terdiri dari data sumber pendapatan yang didapatkan dari Badan Pusat Statistik (BPS). Data cuaca yang terdiri dari variabel, curah hujan dalam mm/hari, kecepatan angin dalam m/s, dan *screen temperature* dalam K didapatkan dari Badan Meteorologi, Klimatologi, dan Geofisika (BMKG). Data

tersebut disediakan dalam format NetCDF. Data penelitian yang didapatkan dari penelitian Sitanggang *et al.* (2014) telah dipraproses dan disimpan dalam bentuk format CSV.

Tahapan Penelitian

Tahapan penelitian yang dilakukan pada penelitian ini ditunjukkan pada Gambar 1.



Gambar 1 Tahapan penelitian

Klasifikasi data menggunakan algoritme C5.0

Klasifikasi adalah tugas pembelajaran sebuah fungsi target f yang memetakan masing-masing atribut set x ke salah satu kelas berlabel y yang telah ditentukan. Fungsi target juga dikenal sebagai model klasifikasi. Model klasifikasi dapat berfungsi sebagai alat penjelas untuk membedakan antara objek dari kelas yang berbeda. Selain itu, model klasifikasi dapat juga digunakan untuk memprediksi label kelas pada data yang belum diketahui (Tan et al. 2005).

Algoritme C5.0 merupakan pengembangan dari algoritme C4.5. Keunggulan algoritme C5.0 adalah menghasilkan model pohon keputusan dan *rulesets* (himpunan aturan) yang lebih kecil dibandingkan C4.5. Menurut Quinlan (1993), pohon keputusan dibangun dengan cara *divide* dan *conquer* dari himpunan data latih. Pembagian himpunan data latih diproses hingga semua sub himpunan kejadian memiliki kelas yang sama.

Model klasifikasi yang digunakan ialah model pohon keputusan dan model berbasis aturan. Model pohon keputusan memiliki struktur pohon seperti flowchart, yang masing-masing simpul internal (simpul bukan daun) menunjukkan pengujian pada atribut, masing-masing cabang mewakili hasil dari pengujian, dan masing-masing simpul daun merupakan label kelas. Simpul paling atas pada struktur tree adalah simpul root. Sedangkan pada model berbasis aturan, dihasilkan kondisi if-then yang diturunkan oleh model pohon keputusan ke dalam kondisi yang lebih sederhana. Pembentukan model klasifikasi pohon keputusan dan model berbasis aturan pada penelitian ini menggunakan algoritme C5.0. Algoritme pohon keputusan C5.0 (Han et al. 2012) dapat dilihat pada Gambar 2.

Algoritme: Generate decision tree input:

- a. Partisi data, D, data latih yang telah ditentukan label
- b. attribute list, himpunan yang terdiri dari atribut kandidat
- c. attribute selection method, prosedur untuk menentukan kriteria pemotongan yang partisi tuple data terbaik ke kelas masing-masing. Kriteria ini terdiri dari splitting attribute dan kemungkinan split point atau splitting subset

output: pohon keputusan

- Metode:
 - 1. Buat simpul N
 - 2. Jika tuple di D memiliki kelas yang sama yaitu C
 - maka jadikan N sebagai simpul daun dan beri label C
 - 4. Jika attribute list kosong, maka
 - jadikan simpul N sebagai simpul daun dengan label = nilai kelas terbanyak pada sampel.
 - 6. Terapkan attribute selection method (D, attribute list) untuk mendapatkan atribut uji terbaik
 - 7. Beri label simpul N dengan atribut uji
 - 8. Jika atribut uji bernilai diskret dan dibolehkan untuk dipisah, maka
 - 9. Attribute list \leftarrow attribute_list atribut uji
 - Untuk setiap nilai j dari atribut uji yang diketahui
 - Buat Dj menjadi kumpulan data tuple untuk memenuhi hasil j
 - Jika Di kosong maka Tambahkan simpul daun dengan label = nilai kelas yang terbanyak pada D ke simpul N
 - iii. Selainnya, tambah cabang baru di bawah yang sekarang dengan memanggil fungsi Generate decision tree (Dj, attribute list) ke simpul N
 - 11. Kembalikan N

Gambar 2 Algoritme pohon keputusan C5.0

Algoritme C5.0 menggunakan information gain sebagai ukuran pemilihan atribut. Atribut dengan information gain terbesar ditentukan sebagai atribut pemisah untuk simpul N. Atribut ini meminimumkan informasi untuk mengklasifikasi tuple dalam partisi yang dihasilkan. Ukuran pemilihan atribut yang didefinisikan pada atribut A adalah sebagai berikut (Han et al. 2012):

$$Info(D) = -\sum_{i=1}^{m} p_i \log_2(p_i)$$
 (1)

Info(D) merupakan informasi yang dibutuhkan untuk mengklasifikasi label kelas sebuah tuple di D. p_i adalah peluang bukan nol dengan sebuah tuple acak di D memiliki kelas C_i dan ditentukan dengan $|C_{i,D}|/|D|$. Info(D) juga dikenal sebagai entropy.

 a_2, \ldots, a_v dari data latih. Atribut A digunakan untuk memisahkan D ke dalam v partisi atau sub himpunan $\{D_1, D_2, ..., D_v\}$. $\frac{|D_j|}{|D|}$ merupakan bobot partisi ke-j. Nilai entropy yang dihasilkan untuk mengklasifikasi tuple dari D berdasarkan partisi oleh A adalah:

 $Info_A(D) = \sum_{j=1}^{\nu} \frac{|D_j|}{D} \times Info(D_j)$ (2)

Information gain yang diperoleh pada atribut *A* adalah:

$$Gain(A) = Info(D) - Info_A(D)$$
 (3)

Algoritme C5.0 terdapat pada *package* C50 yang telah disediakan oleh perangkat lunak R. Variabel target yang digunakan adalah titik panas MODIS pada tahun 2008 dan data bukan titik panas. Data dibagi menjadi 75% data latih dan 25% data uji secara acak. Hasil keluarannya adalah model pohon keputusan dan model berbasis aturan untuk klasifikasi kemunculan titik panas. Variabel target dari hasil klasifikasi adalah kelas dengan label *hotspot* (titik panas) dan *non-hotspot* (bukan titik panas). Label *hotspot* menunjukkan kemunculan titik panas sedangkan label *non-hotspot* menunjukkan ketidakmunculan titik panas. Model klasifikasi yang digunakan untuk memprediksi kemunculan titik panas diimplementasikan pada tahapan selanjutnya.

Implementasi aplikasi berbasis web menggunakan Shiny

Shiny merupakan *framework* untuk pengembangan aplikasi berbasis web yang interaktif pada sisi klien. Shiny dikembangkan dengan menggunakan bahasa pemrograman R. R merupakan sebuah paket dan bahasa pemrograman untuk analisis data dan grafik. R merupakan bahasa pemrograman tingkat tinggi yang bersifat *open source* (Venables dan Smith 2014). R dilengkapi dengan *packages* sebagai fitur-fitur tambahan yang dapat dikembangkan sesuai kebutuhan. R sangat efektif dalam mengolah data berukuran besar dengan berbagai teknik *data mining*, terutama klasifikasi. Shiny sangat mudah dikembangkan dan diintegrasikan dengan konten web menggunakan HTML dan CSS. Javascript dan jQuery juga bisa digunakan untuk memperluas pengembangan aplikasi menggunakan Shiny (Beeley 2013).

Aplikasi klasifikasi berbasis web dibangun menggunakan perangkat lunak R dengan memanfaatkan package Shiny. Model klasifikasi yang digunakan, yaitu algoritme C5.0, dapat diterapkan pada framework Shiny dengan pemanggilan fungsi yang tersedia pada package C5.0. Pengembangan model klasifikasi membutuhkan partisi data latih dan uji dengan memanggil fungsi yang tersedia pada package Caret. Ketiga package tersebut telah tersedia pada perangkat lunak R yang dapat digunakan dan disebarkan secara bebas mengacu pada lisensi dokumen pada setiap package atau software library. Package Shiny dan C5.0 dilisensi dibawah GPL versi 3 sedangkan package Caret dilisensi dibawah GPL versi 2 atau diatasnya. Package tersebut diperbolehkan untuk disalin dan disebarluaskan selain diubah isinya.

Pengembangan aplikasi berbasis web dapat dilakukan pada perangkat lunak selain R, salah satunya adalah OpenGeo Suite. Untuk mengembangkan aplikasi klasifikasi berbasis web dengan perangkat lunak OpenGeo Suite, diperlukan penerapan model klasifikasi berdasarkan aturan yang dihasilkan pada perangkat lunak data mining ke perangkat lunak OpenGeo Suite. Pengembangan ini memerlukan waktu yang lebih lama dibandingkan dengan perangkat lunak R yang telah menyediakan packages untuk pengembangan aplikasi berbasis web dan tools untuk data mining. Selain R, perangkat lunak Weka dapat digunakan untuk melakukan tugas data mining. Pada perangkat lunak Weka, terdapat banyak

metode klasifikasi, salah satunya algoritme C4.5 yang diimplementasikan di dalam modul J48. Pengembangan dari algoritme C4.5 adalah algoritme C5.0 tetapi model klasifikasi algoritme C4.5 menghasilkan ukuran pohon keputusan yang lebih besar dibandingkan dengan algoritme C5.0. Algoritme C5.0 terdapat pada *packages* perangkat lunak R yang dikembangkan untuk keperluan analisis *data mining*.

Analisis *data mining* dapat ditampilkan ke dalam aplikasi berbasis web yang interaktif, salah satunya analisis untuk model klasifikasi. Pengembangan aplikasi berbasis web dengan Shiny dilakukan dengan menerapkan aturan yang telah dihasilkan dari model klasifikasi data titik panas. Pengembangan aplikasi ini menggunakan *package* Shiny yang telah di-*install* di perangkat lunak R.

Prediksi titik panas

Pada tahapan ini, dilakukan prediksi kemunculan titik panas berdasarkan karakteristik wilayah pada aplikasi berbasis web. Hasil prediksi ditampilkan untuk setiap model klasifikasi pohon keputusan dan berbasis aturan.

Lingkungan Pengembangan

Pengembangan aplikasi klasifikasi berbasis web dengan *framework* Shiny dan algoritme C5.0 menggunakan beberapa perangkat keras dan perangkat lunak sebagai berikut:

Perangkat keras:

- *Processor* Intel Core i7 @3.5GHz
- RAM 4 GB

Perangkat lunak:

- Sistem operasi Windows 8.1
- Bahasa pemrograman R-3.1.2
- RStudio versi 0.98.1091 dengan *package* Shiny untuk implementasi aplikasi berbasis web, *package* C50 untuk klasifikasi data, *package* Caret untuk partisi data latih dan uji serta membuat antarmuka untuk model dan prediksi
- Browser Mozilla Firefox untuk menampilkan aplikasi berbasis web

HASIL DAN PEMBAHASAN

Klasifikasi Data Menggunakan Algoritme C5.0

Algoritme C5.0 diterapkan pada *dataset* kebakaran hutan dengan menghasilkan model klasifikasi berupa model pohon keputusan dan model berbasis aturan. Algoritme C5.0 telah tersedia pada perangkat lunak R dengan *package* C50 (Kuhn *et al.* 2015a). Pada algoritme C5.0, model berbasis aturan dihasilkan dari model pohon keputusan yang memiliki kondisi *if-then* dengan kondisi yang tidak sama pada masing-masing aturan. Model berbasis aturan yang memiliki banyak kondisi *if-then* dapat disederhanakan dan dipangkas sehingga aturan yang diturunkan dapat berjumlah sedikit dari aturan yang dihasilkan oleh model pohon keputusan. Sebelum dibentuk model klasifikasi dengan *package* C50, dilakukan partisi data latih dan data uji dengan *package* Caret. Jumlah *record* pada *dataset* kebakaran hutan sebesar 1013 data. Data kebakaran hutan dipartisi secara acak dengan 75% *dataset* dijadikan sebagai data latih. Sehingga terdapat sebanyak 760 data latih yang digunakan untuk menghasilkan model klasifikasi.

Model klasifikasi dengan menggunakan algoritme C5.0 pada dataset kebakaran hutan menghasilkan pohon keputusan dengan banyaknya aturan sebesar 16. Sedangkan pada model berbasis aturan, didapatkan hasil klasifikasi dengan banyaknya aturan sebesar 15. Model klasifikasi dengan algoritme C4.5 menghasilkan pohon keputusan dengan ukuran daun yang lebih banyak dibandingkan dengan algoritme C5.0 (Lampiran 2). Hasil klasifikasi dengan algoritme C4.5 menghasilkan banyaknya aturan sebesar 87. Model klasifikasi yang dihasilkan oleh model pohon keputusan secara lengkap dapat dilihat pada Lampiran 3 dan model klasifikasi yang dihasilkan oleh model berbasis aturan secara lengkap dapat dilihat pada Lampiran 4. Pada data uji, akurasi yang dihasilkan oleh model pohon keputusan adalah 72.72% dengan 69 dari 253 data titik panas yang diklasifikasikan tidak benar oleh model pohon keputusan. Pada model model berbasis aturan, akurasi yang dihasilkan lebih tinggi, yaitu sebesar 73.51% dengan 67 dari 253 data titik panas yang diklasifikasikan tidak benar oleh model berbasis aturan. Beberapa aturan yang dihasilkan oleh pohon keputusan dapat dilihat pada Gambar 3.

Pada aturan ke-1, jika jenis lahan gambut memiliki karakteristik (hemists/saprists(60/40), moderat) atau (saprists/min(50/50), moderat) atau (hemists/min(30/70), moderat) atau bukan lahan gambut atau (saprists/min(50/50), dangkal) atau (saprists/hemists(60/40), moderat) dan jarak ke jalan terdekat kurang dari atau sama dengan 2.5 km atau lebih dari 5 km dan jarak ke sungai terdekat kurang dari atau sama dengan 1.5 km atau di antara 1.5 km dan 3 km maka wilayah tersebut diprediksi tidak terdapat kemunculan titik panas. Pada aturan ke-5, jika jenis lahan gambut (hemists/saprists(60/40), sangat dalam) atau (saprists/min(90/10), moderat) atau (hemists/saprists(60/40), dalam) atau (saprists(100), moderat) dan sumber pendapatan berasal dari perkebunan atau pertanian lainnya atau tidak ada data atau kehutanan dan jarak ke jalan terdekat kurang dari atau sama dengan 2.5 km atau di antara 2.5 km dan 5 km maka wilayah tersebut diprediksi terdapat kemunculan titik panas (Gambar 3). Masingmasing aturan memiliki nilai statistik (n/m), salah satunya pada aturan ke-5. Nilai

statistik (234/43) adalah 234 data tergolong kelas *hotspot* dan 43 adalah data yang bukan kelas *hotspot* tetapi terprediksi sebagai kelas. Aturan yang dihasilkan oleh model berbasis aturan dapat dilihat pada Gambar 4.

- 1. IF peatland_type in { (Hemists/Saprists(60/40), Moderate),
 (Saprists/min(50/50), Moderate), (Hemists/min(30/70),
 Moderate), (non_peatland), (Saprists/min(50/50), Shallow),
 (Saprists/Hemists(60/40), Moderate) } AND dist_road in
 { (Low: <= 2.5 km), (High: > 5 km) } AND dist_river in { (Low: <= 1.5 km), (Medium: (1.5 km, 3]) } THEN Class Non Hotspot
 (214/41)</pre>
- 2. IF peatland_type in { (Hemists/Saprists(60/40), Moderate),
 (Saprists/min(50/50), Moderate), (Hemists/min(30/70),
 Moderate), (non_peatland), (Saprists/min(50/50), Shallow),
 (Saprists/Hemists(60/40), Moderate) } AND dist_road in
 { (Low: <= 2.5 km), (High: > 5 km) } AND dist_river in
 { (High: > 3 km) } AND wind_speed in { [1 m/s, 2 m/s) } THEN
 Class Non Hotspot (61/16)
- 3. IF peatland_type in {(Hemists/Saprists(60/40), Moderate),
 (Saprists/min(50/50), Moderate), (Hemists/min(30/70),
 Moderate), (non_peatland), (Saprists/min(50/50), Shallow),
 (Saprists/Hemists(60/40), Moderate)} AND dist_road in
 {(Medium: (2.5 km, 5 km])} AND peatland_type in
 {(non_peatland)} AND land_cover in
 {(Unirrigated_agri_field)} AND income_source in
 {(Forestry)} THEN Class Hotspot (5/1)
- 4. IF peatland_type in {(Hemists/Saprists(60/40), Very_deep),
 (Saprists/min(90/10), Moderate), (Hemists/Saprists(60/40),
 Deep), (Saprists(100), Moderate)} AND income_source in
 {(Others), (Trading_restaurant), (Agriculture)} THEN Class
 Non Hotspot (18/5)
- 5. IF peatland_type in {(Hemists/Saprists(60/40), Very_deep), (Saprists/min(90/10), Moderate), (Hemists/Saprists(60/40), Deep), (Saprists(100), Moderate)} AND income_source in {(Plantation), (Other_agriculture), (No_data), (Forestry)} AND dist_road in {(Low: <= 2.5 km), (Medium: (2.5 km, 5 km])} THEN Class Hotspot (234/43)</p>

Gambar 3 Aturan klasifikasi dari model pohon keputusan

Pada aturan ke-2 (Gambar 4), jika jarak ke sungai terdekat lebih dari 3 km dan jarak ke jalan terdekat kurang dari atau sama dengan 2.5 km atau lebih dari 5 km dan jenis lahan gambut memiliki karakteristik (hemists/saprists(60/40), moderat) atau bukan jenis lahan gambut atau (saprists/min(50/50), dangkal) maka wilayah tersebut diprediksi terdapat kemunculan titik panas. Pada aturan ke-4, jika sumber pendapatan berasal dari pertanian atau perdagangan atau yang lainnya maka wilayah tersebut diprediksi tidak terdapat kemunculan titik panas (Gambar 4). Masing-masing aturan memiliki nilai statistik (n/m) dan *lift*. Pada aturan ke-5, nilai statistik (165/27) adalah 165 data tergolong kelas *hotspot* dan 27 adalah data yang bukan kelas *hotspot* tetapi terprediksi sebagai kelas. Nilai *lift* merupakan hasil pembagian akurasi rasio Laplace dengan frekuensi relatif kelas yang diprediksi pada data latih. Hasil klasifikasi kedua model, baik model pohon keputusan dan model berbasis aturan, memiliki perbedaan penggunaan variabel penjelas pada aturan yang dihasilkan. Penggunaan variabel tersebut dapat dilihat pada Tabel 1.

DR Ilniversity

1. IF dist_river in {(Low: <= 1.5 km), (Medium: (1.5 km, 3])}
 AND dist_road in {(Low: <= 2.5 km), (High: > 5 km)} AND
 peatland_type in {(Hemists/Saprists(60/40), Moderate),
 (Saprists/Hemists(60/40), Moderate), (Saprists/min(50/50),
 Moderate), (Hemists/min(30/70), Moderate), (non_peatland),
 (Saprists/min(50/50), Shallow)} THEN Class Non Hotspot
 (214/41, lift 1.6)

2. IF dist_river in {(High: > 3 km)} AND dist_road in {(Low: <=
2.5 km), (High: > 5 km)} AND peatland_type in
{(Hemists/Saprists(60/40), Moderate), (non_peatland),
 (Saprists/min(50/50), Shallow)} THEN Class Hotspot (15/2,
 lift 1.6)

3. IF dist_road in {(Medium: (2.5 km, 5 km])} AND income_source
 in {(Forestry)} THEN Class Hotspot (40/10, lift 1.5)

4. IF income_source in {(Others), (Trading_restaurant), (Agriculture)} THEN Class Non Hotspot (138/35, lift 1.5)

5. IF income_source in {(Medium: (2.5 km, 5 km]), (Other_agriculture), (Forestry)} AND peatland_type in {(Saprists/min(90/10), Moderate), (Hemists/Saprists(60/40), Deep), (Saprists(100), Moderate)} THEN Class Hotspot (165/27, lift 1.6)

Gambar 4 Aturan klasifikasi dari model berbasis aturan

Tabel 1 Penggunaan variabel penjelas pada model pohon keputusan dan model berbasis aturan

Penggunaan variabel				
Model pohon keputusan	Model berbasis aturan			
100% Jenis lahan gambut	93.16% Jenis lahan gambut			
97.63% Jarak ke jalan terdekat	82.50% Jarak ke jalan terdekat			
54.61% Jarak ke sungai terdekat	63.95% Sumber pendapatan			
51.71% Sumber pendapatan	42.50% Jarak ke sungai terdekat			
10.39% Tutupan lahan	33.16% Tutupan lahan			
10.00% Kelajuan angin	31.58% Kelajuan angin			

Tabel 1 menunjukkan bahwa pada kedua model, baik model pohon keputusan dan model berbasis aturan, keenam variabel penjelas yang digunakan pada model klasifikasi adalah sama. Keenam variabel tersebut adalah variabel jenis lahan gambut, jarak ke jalan terdekat, jarak ke sungai terdekat, sumber pendapatan, tutupan lahan, dan kelajuan angin. Variabel jenis lahan gambut memiliki penggunaan variabel yang paling besar pada kedua model klasifikasi. Pada model pohon keputusan, variabel jenis lahan gambut digunakan 100%. Penggunaan variabel ini menandakan bahwa pada setiap aturan model pohon keputusan, terdapat variabel jenis lahan gambut. Sedangkan pada model berbasis aturan, variabel jenis lahan gambut digunakan sebesar 93.16% yang menandakan bahwa tidak semua aturan mengandung variabel jenis lahan gambut.

Implementasi Aplikasi Berbasis Web Menggunakan Shiny

Pembangunan aplikasi berbasis web dilakukan di perangkat lunak R dengan package Shiny (Chang et al. 2015). Aplikasi Shiny memiliki dua komponen dasar, yaitu user-interface dan server. Masing-masing komponen

disimpan dalam fail ui.r dan server.r. Komponen *user-interface* berfungsi untuk mengatur tampilan aplikasi dan server berfungsi untuk menyimpan instruksi yang dijalankan pada aplikasi.

Pada aplikasi ini, input dataset adalah fail csv yang berisi sepuluh variabel penjelas (jarak ke sungai terdekat, jarak ke jalan terdekat, jarak ke pusat kota terdekat, sumber pendapatan, curah hujan dalam mm/hari, screen temperature dalam K, tutupan lahan, kedalaman lahan gambut, kelajuan angin dalam m/s dan jenis lahan gambut) dan variabel target (data titik panas dan data bukan titik panas). Input dataset kebakaran hutan yang telah dimasukkan dibentuk sebagai data frame dan dijadikan sebagai objek reactivity yang ketika input-nya dimasukkan berubah, server akan menghasilkan output berdasarkan perubahan input tersebut. Fungsi reactivity yang digunakan pada penelitian ini adalah reactive dan renderPrint. Fungsi reactive berfungsi untuk menghasilkan objek yang akan digunakan oleh lebih dari satu output. Pada fungsi renderPrint, output akan ter-update ketika terdapat perubahan input di dalam fungsi renderPrint. Setiap perubahan input akan menyebabkan pemanggilan ulang model klasifikasi yang terbentuk. Fungsi renderPrint menampilkan hasil klasifikasi dataset kebakaran hutan dan prediksi titik panas. Fungsi ini menghasilkan argumen tunggal berupa output tercetak. Hasil output tersebut ditampilkan tanpa perlu memuat ulang halaman browser. Proses antara komponen server dan *user-interface* untuk menampilkan model klasifikasi dapat dilihat pada Gambar 5.

Pada komponen *user-interface*, *user* memasukkan *dataset* kebakaran hutan sebagai fail csv dan diterima oleh server sebagai *data frame*. *Data frame* yang telah dibentuk oleh fungsi reactive, disimpan dalam fungsi dataset *Input*. Fungsi dataset *Input* digunakan sebagai objek untuk membangun data partisi dengan *package* Caret (Kuhn *et al.* 2015b) dan untuk membangun model klasifikasi dengan *package* C50. *Output* model klasifikasi dihasilkan dengan fungsi renderPrint pada komponen server dan ditampilkan dengan fungsi verbatimText*Output* pada komponen *user-interface*.

Model klasifikasi yang telah terbentuk dapat digunakan untuk memprediksi kemunculan titik panas. Prediksi titik panas ini menggunakan kedua model klasifikasi, model pohon keputusan dan model berbasis aturan. Fungsi prediksi disesuaikan dengan fungsi klasifikasi yang telah diimplementasikan ke dalam model secara reaktif. Sehingga jika terdapat perubahan *input* oleh pengguna, maka model klasifikasi akan berubah dan fungsi prediksi menyesuaikan dengan model klasifikasi tersebut. Fungsi prediksi telah terdapat pada *package* C50 dan dapat menghasilkan nilai *model confidence* serta kelas yang diprediksi. Proses antara komponen server dan *user-interface* untuk menampilkan hasil prediksi titik panas dapat dilihat pada Gambar 6.

2 1 #reading dataset #reading dataset fileInput("file", label datasetInput <- reactive({</pre> = h3("Choose CSV obj<-File"), as.data.frame(read.csv(inFile accept=c('text/csv','te \$datapath, xt/comma-separatedheader=input\$header,sep=input values,text/plain', \$sep, quote=input\$quote)) }) '.csv')) #creating data partition, 3 data training, and data testing dataPartition <- reactive({</pre> set.seed(123) inTrain<createDataPartition(datasetIn put()\$target_attr, p = .75, list = FALSE) }) dataTraining <- reactive({</pre> training <datasetInput()[dataPartition(),]}) dataTesting <- reactive({</pre> testing <- datasetInput()[-</pre> dataPartition(),]}) #creating tree #displaying tree output\$trees <- renderPrint({</pre> verbatimTextOutput("tre trees <- C5.0 (target attr ~., es") data = dataTraining()) summary(trees) }) #creating rule-based output\$rules<- renderPrint({</pre> #displaying rule-based 7 rules <- C5.0(target attr ~., verbatimTextOutput("rul data = dataTraining(), rules es") = TRUE) summary(rules) })

User-interface

Server

Gambar 5 Proses menampilkan model klasifikasi

Server User-Interface #new object to predict 2 #selecting attribute selectInput("dist city","Ci testing <- reactive({</pre> obj2<-data.frame(ty distance", choices = 10 = input\$dist city, ("Low: <= 7km" = "l0v1",11 = input\$dist river, 12 = input\$dist road, "Medium: (7 km, 14 km]" = 13 = input\$income source, 14 = input\$land cover, "High: > 14 km" = "10v2")), 15 = input peatland type16 = input\$precipitation, selectInput("dist river", 17 = input\$peatland depth, "River distance", choices = 18 = input\$screen temp, 19 = input\$wind speed, ("Low: <= 1.5 km" = "llv0",target attr = \overline{NA})}) "Medium: (1.5 km, 3 km]" = "l1v1", "High: > 3 km" = "l1v2")), #predict by rules output\$predictNewObjectByR ules <- renderPrint({</pre> #displaying the predicting 5 rules <- C5.0(target attr result by rule-based ~., data = dataTraining(), verbatimTextOutput("predict rules = TRUE)rulesClass<-predict(rules,</pre> NewObjectByRules") testing(), type="class") rulesClass}) #predict by Trees output\$predictNewObjectByT #displaying the predicting rees <- renderPrint({</pre> result by trees trees <- C5.0(target attr</pre> verbatimTextOutput("predict ~., data = dataTraining()) NewObjectByTrees") treesClass<-predict(trees, testing(), type="class") treesClass})

Gambar 6 Proses menampilkan hasil prediksi titik panas

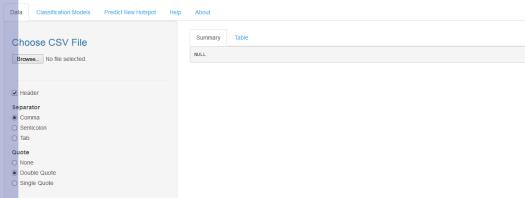
Pada aplikasi ini, hasil prediksi kebakaran hutan yang diberikan adalah berupa label kelas hotspot maupun non-hotspot. Sebelum menampilkan hasil prediksi, pengguna perlu memasukkan *input* untuk setiap parameter yang telah disediakan. Pada user-interface, parameter disimpan dalam fungsi selectInput. Parameter tersebut merupakan variabel penjelas (jarak ke sungai terdekat, jarak ke jalan terdekat, jarak ke pusat kota terdekat, sumber pendapatan, curah hujan dalam mm/hari, screen temperature dalam K, tutupan lahan, kedalaman lahan gambut, kelajuan angin dalam m/s dan jenis lahan gambut). Sepuluh parameter ini digabungkan dan dibentuk sebagai data frame yang baru yang akan digunakan sebagai input fungsi prediksi. Data frame yang telah dibentuk oleh fungsi reactive, disimpan dalam fungsi testing. Objek tersebut disimpan dalam bentuk objek yang reaktif untuk menanggapi setiap perubahan yang diminta oleh pengguna.

Pada bagian komponen user-interface, terdapat fungsi fluidPage untuk menghasilkan tampilan yang disesuaikan dengan ukuran browser pengguna. Setiap elemen untuk membangun aplikasi ditempatkan pada fungsi fluidPage. Pada aplikasi ini, elemen yang digunakan adalah titlePanel, tabsetPanel, dan sidebarLayout. titlePanel merupakan judul aplikasi yang ditampilkan pada web. sidebarLayout merupakan elemen dasar yang terdiri dari sidebarPanel dan mainPanel untuk menampilkan fungsi output. Fungsi sidebarPanel ditampilkan pada sisi kiri aplikasi dan fungsi mainPanel ditampilkan pada sisi kanan aplikasi secara default. tabsetPanel merupakan elemen untuk mengatur layout aplikasi untuk membagi antarmuka ke dalam beberapa bagian. Pembagian antarmuka di dalam tabsetPanel dapat dilakukan dengan menggunakan elemen tabPanel. Pada aplikasi ini, tabPanel terdiri dari data, classification models, predict new hotspot, about, dan help.

tabPanel data merupakan halaman utama aplikasi berbasis web untuk data kebakaran hutan (Gambar 7). Pada tabPanel tersebut, pengguna dapat memilih dataset kebakaran hutan yang tersimpan di dalam komputer pribadinya dan mengunduh dataset tersebut untuk diproses. Dataset tersebut harus terdiri dari sepuluh variabel penjelas (jarak ke sungai terdekat, jarak ke jalan terdekat, jarak ke pusat kota terdekat, sumber pendapatan, curah hujan dalam mm/hari, screen temperature dalam K, tutupan lahan, kedalaman lahan gambut, kelajuan angin dalam m/s dan jenis lahan gambut) dan variabel target. Dataset yang diunduh harus memiliki kode variabel sebagai nilai dari variabel penjelas. Kode variabel penjelas telah didapatkan dari penelitian sebelumnya oleh Sitanggang et al. (2014). Kode variabel penjelas dapat dilihat pada Lampiran 1.

Setelah pengguna memasukkan *dataset*, aplikasi menampilkan ringkasan data dan tabel (Lampiran 5). Selain itu, aplikasi dapat menghasilkan model klasifikasi yang dapat dilihat pada tabPanel *classification models* (Gambar 8). Gambar 8 merupakan tampilan hasil klasifikasi dengan model pohon keputusan. Tampilan hasil klasifikasi dengan model berbasis aturan dapat dilihat pada Lampiran 6.

Hotspot classification using C5.0 Algorithm

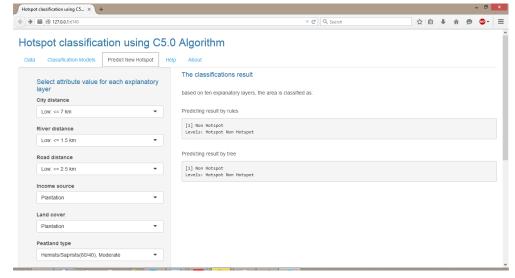


Gambar 7 Tampilan halaman utama aplikasi berbasis web

. 🗇 🗦

Gambar 8 Hasil klasifikasi dengan model pohon keputusan

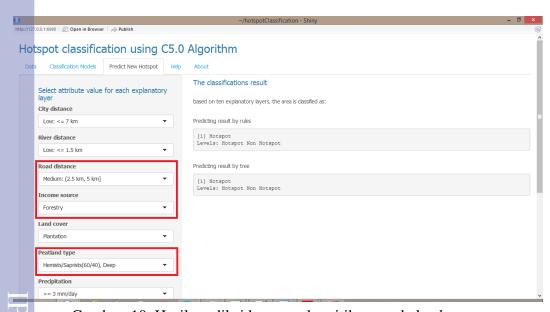
Model klasifikasi pada dataset kebakaran hutan dengan menggunakan algoritme C5.0 yang telah terbentuk dapat digunakan untuk memprediksi kemunculan titik panas. Fungsi prediksi ini dapat digunakan pada bagian tabPanel predict new hotspot. Pada tampilan predict new hotspot, pengguna dapat memilih variabel penjelas yang terdiri dari jarak ke sungai terdekat, jarak ke jalan terdekat, jarak ke pusat kota terdekat, sumber pendapatan, curah hujan dalam mm/hari, screen temperature dalam K, tutupan lahan, kedalaman lahan gambut, kelajuan angin dalam m/s dan jenis lahan gambut. Kemudian, pengguna memilih button classify untuk memprediksi kemunculan titik panas tersebut. Output yang dihasilkan adalah kelas hotspot atau non-hotspot. Kelas hotspot mengindikasikan bahwa pada wilayah tersebut berpotensi munculnya titik panas dan kelas nonhotspot mengindikasikan tidak adanya potensi kemunculan titik panas. Tampilan predict new hotspot dapat dilihat pada Gambar 9.



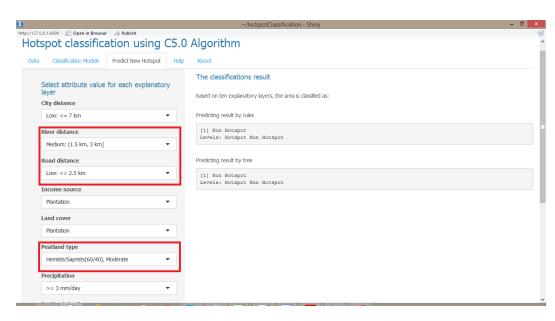
Gambar 9 Tampilan hasil prediksi pada tabPanel predict new hotspot

Prediksi Titik Panas

Pada tampilan tabPanel predict new hotspot, dilakukan prediksi kemunculan titik panas dengan karakteristik wilayah tertentu. Pada tahapan ini, diprediksi kemunculan titik panas untuk wilayah dengan karakteristik jenis lahan gambut adalah (Hemists/Saprists(60/40), Deep) dan jarak lokasi tertentu ke jalan terdekat adalah 3 km serta sumber pendapatan berasal dari hutan. Pada variabel jenis lahan gambut, data yang dimasukkan adalah (Hemists/Saprists(60/40), Deep). Pada jarak terdekat dengan jalan (road distance), data yang dimasukkan adalah medium yang jarak lokasi ke jalan terdekat di antara 2.5 km dan 5 km. Sedangkan pada variabel sumber pendapatan berasal dari hutan, data yang dimasukkan adalah forestry. Variabel penjelas selain jenis lahan gambut, jalan, dan sumber pendapatan dibiarkan secara default. Nilai default untuk variabel penjelas yang digunakan pada aplikasi ini, dapat dilihat pada Tabel 2. Hasil prediksi kemunculan lahan panas dengan nilai parameter jenis gambut berkategori Deep), jalan (Hemists/Saprists(60/40), berkategori medium dan sumber pendapatan berkategori forestry adalah kelas hotspot (Gambar 10). Prediksi kemunculan titik panas juga dilakukan untuk wilayah dengan karakteristik jarak lokasi tertentu ke sungai terdekat adalah 3 km serta jarak lokasi tertentu ke jalan terdekat adalah 1 km dan tipe lahan gambut adalah (Hemists/Saprists(60/40), Moderate). Masing-masing nilai parameter yang dimasukkan pada aplikasi dengan jarak lokasi tertentu ke sungai terdekat, jarak lokasi tertentu ke jalan terdekat, serta jenis lahan gambut adalah Medium, Low, dan (Hemists/Saprists(60/40), Moderate). Hasil prediksi kemunculan titik panas berdasarkan karakteristik wilayah tersebut adalah kelas *non-hotspot* (Gambar 11). Berdasarkan hasil prediksi, output yang diberikan oleh aplikasi telah sesuai dengan output yang dihasilkan oleh algoritme C5.0, baik model pohon keputusan maupun model berbasis aturan.



Gambar 10 Hasil prediksi kemunculan titik panas kelas hotspot



Gambar 11 Hasil prediksi kemunculan titik panas kelas non-hotspot

Tabel 2 Nilai default variabel penjelas

Variabel penjelas	Nilai <i>default</i>
Jarak ke pusat kota terdekat	<i>Low</i> : <= 7 km
Jarak ke sungai terdekat	<i>Low</i> : <= 1.5 km
Jarak ke jalan terdekat	<i>Low</i> : <= 2.5 km
Sumber pendapatan	Plantation
Tutupan lahan	Plantation
Jenis lahan gambut	Hemists/saprists(60/40), moderate
Curah hujan dalam mm/hari	>= 3 mm/day
Kedalaman lahan gambut	Shallow/thin 50-100 cm
Screen temperature dalam K	[298 K, 299 K)
Kelajuan angin dalam m/s	[1 m/s, 2 m/s)

Pengguna yang awam dengan aplikasi berbasis web ini dapat melihat tabPanel help untuk membaca panduan dan keterangan mengenai aplikasi ini. Tampilan tabPanel help terdiri dari sub tabPanel get started, data, classification models, dan predict new hotspot. Tampilan tabPanel help dapat dilihat pada Gambar 13. Tampilan tabPanel help secara lengkap dapat dilihat pada Lampiran 7. Fungsi-fungsi yang terdapat pada aplikasi berbasis web telah diuji menggunakan metode black box dengan hasil setiap fungsi berjalan dengan baik. Hasil pengujian aplikasi dapat dilihat pada Lampiran 8.



Gambar 12 Tampilan halaman tabPanel help

SIMPULAN DAN SARAN

Simpulan

Penelitian ini menerapkan algoritme C5.0 pada data kebakaran hutan dan telah berhasil dibangun aplikasi berbasis web dengan framework Shiny. Dataset kebakaran hutan terdiri dari data fisik, data cuaca, dan data sosial-ekonomi yang mempengaruhi kemunculan titik panas di wilayah Rokan Hilir Provinsi Riau di Indonesia. Akurasi dari hasil klasifikasi untuk memprediksi kemunculan titik panas pada model pohon keputusan adalah sebesar 72.72% dan pada model berbasis aturan adalah sebesar 73.51%. Jumlah aturan yang dihasilkan oleh model pohon keputusan adalah 16 sedangkan jumlah aturan yang dihasilkan oleh model berbasis aturan adalah 15. Berdasarkan besarnya penggunaan variabel penjelas pada hasil klasifikasi, ditunjukkan bahwa jenis lahan gambut, jarak ke jalan terdekat, jarak ke sungai terdekat, sumber pendapatan, tutupan lahan, dan kelajuan angin mempengaruhi kemunculan titik panas. Model klasifikasi pohon keputusan dan model berbasis aturan telah berhasil diimplementasikan untuk memprediksi kemunculan titik panas berdasarkan penggunaan variabel tertinggi yaitu jenis lahan gambut dan jarak lokasi tertentu ke jalan terdekat. Pada framework Shiny, fungsi reactivity telah diimplementasikan untuk menghasilkan output berdasarkan perubahan *input* oleh pengguna.

Saran

Saran yang dapat dilakukan untuk penelitian selanjutnya agar menghasilkan pengembangan yang lebih baik adalah:

Pengembangan aplikasi klasifikasi berbasis web dengan memprediksi kemunculan titik panas tanpa nilai variabel penjelas yang diisi secara *default*.

- Pada penelitian ini, diperlukan input nilai variabel pada fungsi prediksi dan jika salah satu nilai variabel kosong, maka aplikasi tidak dapat melakukan fungsi prediksi.
- Selain itu, diperlukan pengembangan fungsi *reactive* pada fungsi prediksi titik panas sehingga aplikasi dapat membangun fungsi yang disesuaikan dengan perubahan variabel penjelas.
- Data yang digunakan diintegrasikan dalam database server sehingga pengguna dapat menyimpan data tersebut dalam kapasitas besar.
- Diperlukan validasi model klasifikasi yang dihasilkan dengan kondisi riil di
- 5 Pengujian *usability* aplikasi yang telah dibuat perlu dilakukan oleh pengguna.
- Pengembangan fungsi aplikasi berbasis web dengan menampilkan visualisasi hasil klasifikasi dalam bentuk peta.

DAFTAR PUSTAKA

- Agus F, Subiksa IGM. 2012. Lahan Gambut: Potensi untuk Pertanian dan Aspek lingkungan. Bogor (ID): Balai Penelitian Tanah dan World Agroforestry Centre (ICRAF).
- Ali M. 2014. Jumlah titik api di Riau bertambah jadi 130 [Internet]. [diunduh 2014 Okt 25]. Tersedia pada: http://news.liputan6.com/read/2069482/jumlahtitik-api-di-riau-bertambah-jadi-130.
- Bappenas, ADB. 1999. Causes, Extent, Impact and Costs of 1997 / 1998 Fires and Drought. Jakarta (ID): Bappenas dan ADB.
- Beeley C. 2013. Web Application Development with R Using Shiny. Birmingham (GB): Packt.
- Chang W, Cheng J, Allaire JJ, Xie Y, McPherson J, Otto M, Thornton J, Farkas A, Jehl S, Petre S, Rowls A, et al. 2015. Package 'shiny' [Internet]. [diunduh 2015] Jan 12]. Tersedia pada: http://shiny.rstudio.com.
- Han J, Kamber M, Pei J. 2012. Data Mining: Concepts and Techniques. 3rd ed. Massachusetts (US): Morgan Kaufmann.
- Kuhn M, Weston S, Coulter N, Culp M. 2015a. Package 'C50' [Internet]. [diunduh 2015 Feb 4]. Tersedia pada: http://cran.rproject.org/web/packages/C50/index.html.
- Kuhn M, Wing J, Weston S, Williams A, Keefer C, Engelhardt A, Cooper T, Mayer Z, Kenkel B, Benesty M, et al. 2015b. Package 'caret' [Internet]. [diunduh 2015 Feb 4]. Tersedia pada: http://caret.r-forge.r-project.org/
- Quinlan JR. 1993. C4.5: Programs for Machine Learning. Vol ke-1. California (US): Morgan Kaufmann.
- Riswandy A. 2014. Greenpeace: pengeringan lahan gambut penyebab utama kebakaran hutan [Internet]. [diunduh 2014 Okt 9]. Tersedia pada: http://news.detik.com/read/2014/05/28/150422/2594055/10/greenpeacepengeringan-lahan-gambut-penyebab-utama-kebakaran-hutan?nd771104bcj.
- Sitanggang IS, Yaakob R, Mustapha N, Ainuddin AN. 2014. A decision tree based on spatial relationships for predicting hotspots in peatlands. Telkomnika. 12(2):511-518.doi:10.12928/TELKOMNIKA.v12i2.2036.

20

Tacconi L. 2003. Kebakaran Hutan di Indonesia: Penyebab, Biaya, dan Implikasi Kebijakan. Bogor (ID): Center for International Forestry Research (CIFOR).

Tan PN, Steinbach M, Kumar V. 2005. Introduction to Data Mining. Boston (US): Pearson Addison Wesley.

Venables WN, Smith DM. 2014. An Introduction to R [Internet]. [diunduh 2014] Okt 25]. Tersedia pada: http://cran.r-project.org/doc/manuals/r-release/Rintro.html



LAMPIRAN

Lampiran 1 Variabel penjelas yang digunakan pada dataset kebakaran hutan (Sitanggang et al. 2014)

Variabel penjelas (nama)	Nilai variabel (nama)
dist_city (jarak ke pusat kota terdekat)	low: <= 7 km Medium: (7 km, 14 km] High: > 14 km
dist_river (jarak ke sungai terdekat)	Low: <= 1.5 km Medium: (1.5 km, 3 km] High: > 3 km
dist_road (jarak ke jalan terdekat)	Low: <= 2.5 km Medium: (2.5 km, 5 km] High: > 5 km
income_source (sumber pendapatan)	Plantation Others Other_agriculture No_data Trading_restaurant Forestry Agriculture
land_cover (tutupan lahan)	Plantation Dryland_forest Mangrove Settlement Water_body Natural_forest Bare_land Shrubs Paddy_field Unirrigated_agri_field Swamp Mix_garden
peatland_type (jenis lahan gambut)	Hemists/Saprists(60/40),Moderate Saprists(100),Deep Saprists/min(50/50),Moderate Hemists/min(30/70),Moderate Hemists/Saprists(60/40),Very_deep non_peatland Saprists/min(90/10),Moderate Hemists/Saprists(60/40),Deep



Lampiran 1 Lanjutan

22

	_		3
	-		1
	L	L	J
		3	
	<		
	(D	
	9	3	
	-	4	
	_	1	
6	<		

Variabel penjelas (nama)	Nilai variabel (nama)		
peatland_type (jenis lahan gambut)	Saprists/min(50/50),Shallow Saprists(100),Moderate Saprists/Hemists(60/40),Moderate		
Precipitation (curah hujan)	3: >=3 mm/day 2: [2 mm/day, 3 mm/day)		
peatland_depth (kedalaman lahan gambut)	D1: (Shallow/Thin 50-100 cm) D4: (very deep/very thick > 400 cm) D3: (Deep/Thick 200-400 cm), non_peatland D2: (Moderate 100-200 cm)		
screen_temp (screen temperature)	298: [298 K, 299 K) 299: >=299 K 297: [297 K, 298 K)		
wind_speed (kelajuan angin)	1: [1 m/s, 2 m/s) 2: >=2 m/s 0: [0 m/s, 1 m/s)		



Lampiran 2 Model klasifikasi pohon keputusan dengan algoritme C4.5 menggunakan perangkat lunak Weka

```
peatland_type = non_peatland
    income source = Plantation
        land cover = Mix_garden
            wind_speed = [1m/s, 2m/s)
                dist road = Low <= 2.5 km: Non Hotspot (28.0/10.0)
                dist road = High>5km: Hotspot (4.0/1.0)
                dist road = Medium(2.5km,5km]
                    dist city = High>14km
                dist river = Medium(1.5km,3km]: Hotspot
(1.0)
                        dist_river = Low<=1.5km: Hotspot (7.0)</pre>
                        dist_river = High>3km: Non Hotspot (2.0)
                dist city = Medium(7km,14km]: Non Hotspot
                (2.0)
                    dist city = Low < = 7km: Hotspot (7.0)
                wind speed = [0m/s, 1m/s): Hotspot (4.0)
        land cover = Dryland forest: Non Hotspot (5.0)
        land cover = Bare land: Hotspot (10.0/3.0)
        land cover = Plantation
            wind speed = [1m/s, 2m/s): Non Hotspot (45.0/7.0)
            wind speed = [0m/s, 1m/s)
                dist river = Medium(1.5km,3km]: Hotspot (3.0)
                dist_river = Low <= 1.5 km: Non Hotspot (6.0/1.0)
                dist_river = High>3km: Hotspot (11.0/4.0)
            land_cover = Shrubs: Non Hotspot (20.0/6.0)
        land cover = Unirrigated agri field: Non Hotspot
(24.0/9.0)
        land_cover = Paddy_field: Non Hotspot (4.0)
        land_cover = Settlement: Non Hotspot (6.0/1.0)
        land_cover = Mangrove: Non Hotspot (1.0)
        land_cover = Swamp: Non Hotspot (3.0/1.0)
        land_cover = Water_body: Non Hotspot (0.0)
    income source = Forestry
        dist road = Low <= 2.5 km: Non Hotspot (23.0/5.0)
        dist road = High>5km: Non Hotspot (16.0/1.0)
        dist road = Medium(2.5km, 5km)
           precipitation = >=3mm/day: Hotspot (12.0/1.0)
           precipitation = [2mm/day, 3mm/day): Non Hotspot (2.0)
    income_source = Other_agriculture: Non Hotspot (18.0/1.0)
    income source = Agriculture: Non Hotspot (122.0/22.0)
    income source = No data: Non Hotspot (0.0)
    income source = Others: Non Hotspot (5.0)
    income source = Trading restaurant: Non Hotspot (0.0)
peatland type = (Hemists/Saprists(60/40), Deep): Hotspot (93.0/9.0)
peatland type = (Saprists(100), Moderate): Hotspot (41.0/10.0)
peatland type = (Saprists/min(90/10), Moderate): Hotspot
(90.0/17.0)
peatland type = (Hemists/Saprists(60/40), Very deep)
    screen temp = [298K, 299K): Non Hotspot (14.0)
    screen temp = [297K, 298K)
        dist road = Low <= 2.5 km
            wind speed = [1m/s, 2m/s): Hotspot (81.0/11.0)
            wind speed = [0m/s, 1m/s)
                dist city = High>14km
```

Lampiran 2 Lanjutan

```
dist river = Medium(1.5km,3km]: Non Hotspot
(2.0)
                    dist river = Low<=1.5km: Hotspot (0.0)</pre>
                    dist river = High>3km: Hotspot (14.0/4.0)
                dist_city = Medium(7km, 14km]: Hotspot (36.0/12.0)
                dist city = Low < = 7 km
                    peatland depth = non peatland: Hotspot (0.0)
                    peatland depth = Deep/Thick 200-400cm: Non
Hotspot (10.0)
                    peatland depth = Moderate 100-200cm: Hotspot
   (0.0)
                    peatland depth = VeryDeep/VeryThick>400cm:
Hotspot (14.0/2.0)
peatland depth = Shallow/Thin 50-100cm:
Hotspot (0.0)
        dist road = High>5km
            income source = Plantation
                dist river = Medium(1.5km, 3km]: Hotspot (4.0)
                dist river = Low<=1.5km</pre>
                    land cover = Mix garden: Hotspot (0.0)
                    land cover = Dryland forest: Hotspot (0.0)
                    land cover = Bare land: Hotspot (0.0)
                    land cover = Plantation: Hotspot (4.0/1.0)
                    land cover = Shrubs: Non Hotspot (2.0)
                    land_cover = Unirrigated_agri_field: Hotspot
(0.0)
                    land cover = Paddy field: Hotspot (0.0)
                    land cover = Settlement: Hotspot (0.0)
                    land cover = Mangrove: Hotspot (0.0)
                    land cover = Swamp: Hotspot (0.0)
                    land cover = Water body: Hotspot (0.0)
                dist river = High>3km
                    dist city = High>14km: Hotspot (22.0/10.0)
                    dist city = Medium(7km,14km]: Non Hotspot
(7.0)
                    dist city = Low<=7km: Non Hotspot (1.0)
            income source = Forestry: Non Hotspot (13.0)
            income source = Other agriculture: Non Hotspot (1.0)
            income source = Agriculture: Non Hotspot (0.0)
            income_source = No_data: Non Hotspot (0.0)
            income_source = Others: Non Hotspot (0.0)
            income source = Trading restaurant: Non Hotspot (0.0)
        dist road = Medium(2.5km, 5km)
            peatland_depth = non_peatland: Hotspot (0.0)
            peatland depth = Deep/Thick 200-400cm: Non Hotspot
(9.0/2.0)
            peatland depth = Moderate 100-200cm: Hotspot (0.0)
            peatland depth = VeryDeep/VeryThick>400cm: Hotspot
(39.0/11.0)
            peatland depth = Shallow/Thin 50-100cm: Hotspot (0.0)
peatland type = (Hemists/Saprists(60/40), Moderate)
    income source = Plantation: Non Hotspot (35.0/12.0)
    income_source = Forestry
        dist river = Medium(1.5km, 3km]: Hotspot (4.0)
```

Lampiran 2 Lanjutan

```
dist river = Low <= 1.5 km: Non Hotspot (9.0/1.0)
        dist river = High>3km: Non Hotspot (1.0)
    income source = Other agriculture: Hotspot (4.0)
    income source = Agriculture
        dist_road = Low<=2.5km: Non Hotspot (4.0)</pre>
        dist_road = High>5km: Hotspot (3.0)
        dist_road = Medium(2.5km,5km]: Hotspot (15.0/1.0)
    income_source = No_data: Non Hotspot (0.0)
    income_source = Others: Non Hotspot (6.0)
    income_source = Trading_restaurant: Non Hotspot (2.0)
peatland type = (Saprists/min(50/50), Moderate): Non Hotspot
(8.0/1.0)
peatland type = (Saprists/min(50/50), Shallow)
    dist river = Medium(1.5km, 3km]: Non Hotspot (5.0/1.0)
    dist river = Low<=1.5km: Non Hotspot (9.0)</pre>
    dist river = High>3km: Hotspot (5.0)
peatland type = (Saprists(100), Deep): Non Hotspot (13.0)
peatland type = (\text{Hemists/min}(30/70), \text{Moderate}): Non Hotspot (2.0)
Catatan:
Non Hotspot (9.0/1.0) merupakan statistik (n/m) dengan n adalah
sejumlah data yang tergolong kelas dan m adalah sejumlah data yang
```

bukan kelas tetapi diprediksi sebagai kelas.

Lampiran 3 Model klasifikasi pohon keputusan dengan algoritme C5.0 menggunakan perangkat lunak R

```
peatland_type in {(Hemists/min(30/70), Moderate),
                   (Hemists/Saprists(60/40), Moderate),
                           (Saprists(100), Deep),
                   (Saprists/min(50/50), Moderate),
                           (Saprists/min(50/50), Shallow),
                   non_peatland):
  ..dist road in {High>5km, Low<=2.5km}:
    :...dist river in {Low<=1.5km, Medium(1.5km, 3km]}:
            Non Hotspot (214/41)
        dist_river = High>3km:
        :...wind_speed = [0m/s, 1m/s): Hotspot (15/2)
            wind_speed = [1m/s, 2m/s): Non Hotspot (61/16)
    dist road = Medium(2.5km,5km]:
    :...peatland type = (Hemists/Saprists(60/40), Moderate):
                               Hotspot (22/6)
        peatland type in {(Hemists/min(30/70), Moderate),
                                       (Saprists (100), Deep),
                          (Saprists/min(50/50), Moderate),
                          (Saprists/min(50/50), Shallow) }:
                                       Non Hotspot (3)
        peatland_type = non peatland:
        :...land_cover in {Bare_land, Mangrove, Mix_garden, :
                                         Paddy field, Swamp}:
Hotspot (24/7)
            land cover in {Dryland forest, Plantation,
                                         Settlement, Shrubs,
Water body }:
                                        Non Hotspot (28/6)
            land_cover = Unirrigated_agri_field:
            :...income_source in {Agriculture, No_data,
                                   Other_agriculture, Others,
                                   Plantation,
                                   Trading restaurant):
                                   Non Hotspot (11/2)
                income source = Forestry: Hotspot (5/1)
peatland type in {(Hemists/Saprists(60/40), Deep),
                   (Hemists/Saprists(60/40), Very deep),
                   (Saprists (100), Moderate),
                   (Saprists/min(90/10), Moderate) }:
:...income source in {Agriculture,Others,
                       Trading_restaurant): Non Hotspot (18/5)
    income source in {Forestry, No data, Other agriculture,
                       Plantation }:
    :...dist road in \{Low \le 2.5km, Medium(2.5km, 5km]\}:
                       Hotspot (234/43)
        dist road = High>5km:
        :...dist river = Low<=1.5km: Non Hotspot (5/2)
            dist river = High>3km: [S1]
            dist river = Medium(1.5km,3km]:
            :...land cover in {Dryland forest,
                                Unirrigated agri field }:
```

Lampiran 3 Lanjutan

```
Non Hotspot (4)
                land_cover in {Bare_land, Mangrove,
                              Mix_garden, Paddy_field,
                               Plantation, Settlement, Shrubs,
                               Swamp, Water_body}: Hotspot (7/1)
SubTree [S1]
peatland_type in {(Hemists/Saprists(60/40), Deep),
                  (Saprists(100), Moderate),
                  (Saprists/min(90/10), Moderate) }:
                   Hotspot (72/14)
peatland_type = (Hemists/Saprists(60/40), Very_deep):
                 Non Hotspot (37/7)
Catatan:
Hotspot (72/14) merupakan statistik (n/m) dengan n adalah
sejumlah data yang tergolong kelas dan m adalah sejumlah data
yang bukan kelas tetapi diprediksi sebagai kelas.
```

Lampiran 4 Model klasifikasi berbasis aturan dengan algoritme C5.0 menggunakan perangkat lunak R

```
Rule 1: (165/27, lift 1.6)
      income_source in {Forestry, Other_agriculture,Plantation}
      peatland_type in {(Hemists/Saprists(60/40), Deep),
                         (Saprists (100), Moderate),
                         (Saprists/min(90/10), Moderate)}
                         [0.832]
          class Hotspot
Rule 2: (15/2, lift 1.6)
      dist river = High>3km
      dist_road in {High>5km, Low<=2.5km}</pre>
      peatland_type in {(Hemists/Saprists(60/40), Moderate),
                         (Saprists/min(50/50), Shallow),
                         non peatland}
      wind speed = [0m/s, 1m/s)
      -> class Hotspot [0.824]
Rule 3: (234/43, lift 1.6)
      dist road in {Low<=2.5km, Medium(2.5km,5km)}</pre>
      income source in {Forestry, No data, Other agriculture,
                         Plantation }
      peatland type in {(Hemists/Saprists(60/40), Deep),
                         (Hemists/Saprists(60/40), Very deep),
                         (Saprists/min(90/10), Moderate)}
         class Hotspot
                         [0.814]
Rule 4: (45/9, lift 1.6)
      dist river = Medium(1.5km,3km]
      land_cover in {Bare_land, Mix_garden, Paddy_field,
                     Plantation, Shrubs}
      peatland_type in {(Hemists/Saprists(60/40), Deep),
                         (Hemists/Saprists(60/40), Very_deep),
                         (Saprists (100), Moderate),
                         (Saprists/min(90/10), Moderate)}
      -> class Hotspot
                         [0.787]
Rule 5: (40/10, lift 1.5)
      dist road = Medium(2.5km,5km]
      income source = Forestry
      -> class Hotspot [0.738]
Rule 6: (35/9, lift 1.4)
      dist road = Medium(2.5km,5km]
      land cover in {Bare land, Mangrove, Mix garden,
                     Paddy field, Swamp}
      -> class Hotspot [0.730]
Rule 7: (22/6, lift 1.4)
      dist road = Medium(2.5km, 5km)
      peatland type = (Hemists/Saprists(60/40), Moderate)
      -> class Hotspot [0.708]
```

Lampiran 4 Lanjutan

```
Rule 8: (7, lift 1.8)
      dist river = Medium(1.5km, 3km)
      dist road = High>5km
      income source in {Forestry, Plantation}
      land_cover in {Dryland_forest, Unirrigated_agri_field}
      -> class Non Hotspot [0.889]
Rule 9: (214/41, lift 1.6)
      dist_river in {Low<=1.5km, Medium(1.5km,3km)}</pre>
      dist road in {High>5km, Low<=2.5km}</pre>
      peatland type in {(Hemists/min(30/70), Moderate),
                         (Hemists/Saprists(60/40), Moderate),
                         (Saprists (100), Deep),
                         (Saprists/min(50/50), Moderate),
                         (Saprists/min(50/50), Shallow),
                         non peatland}
      ->
          class Non Hotspot [0.806]
Rule 10: (38/7, lift 1.6)
      dist river = High>3km
      dist road = High>5km
      peatland type = (Hemists/Saprists(60/40), Very deep)
      -> class Non Hotspot [0.800]
Rule 11: (46/9, lift 1.6)
      dist river = Low<=1.5km</pre>
      dist road = High>5km
      -> class Non Hotspot [0.792]
Rule 12: (30/6, lift 1.6)
      income source in {Agriculture, Others, Plantation}
      land cover = Unirrigated agri field
      peatland_type in {(Saprists/min(50/50), Shallow),
                          non peatland}
      -> class Non Hotspot [0.781]
Rule 13: (225/49, lift 1.6)
      dist road in {High>5km, Low<=2.5km}</pre>
      peatland type in {(Hemists/min(30/70), Moderate),
                         (Hemists/Saprists(60/40), Moderate),
                         (Saprists (100), Deep),
                         (Saprists/min(50/50), Moderate),
                         (Saprists/min(50/50), Shallow),
                          non peatland}
      wind speed = [1m/s, 2m/s)
      -> class Non Hotspot [0.780]
Rule 14: (138/31, lift 1.6)
      land cover in {Dryland forest, Plantation, Settlement,
                     Shrubs, Water body}
      peatland type = non peatland
      -> class Non Hotspot [0.771]
```



Lampiran 4 Lanjutan

30

```
Rule 15: (138/35, lift 1.5)
      income source in {Agriculture, Others,
                        Trading_restaurant}
          class Non Hotspot [0.743]
```

Catatan:

Class Hotspot [0.889] [0.889] merupakan akurasi Rule 1: (165/27, lift 1.6)

- 165 (n) jumlah data yang tergolong kelas
- 27 (m) jumlah data yang bukan kelas tetapi diprediksi sebagai kelas
 - Lift 1.6 adalah hasil pembagian akurasi (n-m+1)/(n+2)dengan frekuensi relatif kelas yang diprediksi

Lampiran 5 Tampilan ringkasan data dan tabel pada aplikasi berbasis web

Tampilan ringkasan data Hotspot classification using C5.0 Algorithm Summary Table Choose CSV File Choose File Forestry
No_data
Other_agriculture **✓** Header peatland type
non_peatland :391
(Hemists/Saprists(60/40), Very_dep):273
(Hemists/Saprists(60/40), Deep): 273
(Saprists/min(90/30), Moderate): 93
(Hemists/Saprists(60/40), Moderate): 83
(Saprists/Saprists(60/40), Moderate): 83
(Saprists(00), Moderate): 41
(Other): 42
ppth Comma Semicolon ○ Tab O None : '12
screen_temp wind_speed target_attr
:737 [0m/s,lm/s):257 Hotspot :512
:276 [1m/s,2m/s):756 Non Hotspot:501 Double Quote

Tampilan data dalam format tabel Hotspot classification using C5.0 Algorithm Data Classification Models Predict New Hotspot Help Summary Table Choose CSV File Show 25 entries dist_city dist_river dist_road Plantation High>14km Medium(1.5km,3km] Low<=2.5km Mix_garden non_peatland Low<=1.5km Mix_garden non_peatland Separator Medium(1.5km,3km] Low<=2.5km Forestry Dryland_forest (Hemists/Saprists(60/40),Deep) Comma Plantation ○ Tab High>14km High>3km High>5km Plantation (Saprists/min(90/10),Moderate) ○ None High>14km High>3km Medium(2.5km,5km) Plantation Plantation (Hemists/Saprists(60/40),Deep) Double Quote Medium(1.5km,3km] Medium(2.5km,5km] Other_agriculture Dryland_forest (Hemists/Saprists(60/40),Very_d High>14km High>14km High>3km Medium(2.5km,5km] Plantation Mix_garden Medium(7km,14km) High>3km Low<=2.5km Plantation Plantation (Hemists/Saprists(60/40),Very_d High>14km Low<=2.5km Forestry Bare_land (Hemists/Saprists(60/40),Very_d v



32

Lampiran 6 Hasil klasifikasi dengan model berbasis aturan



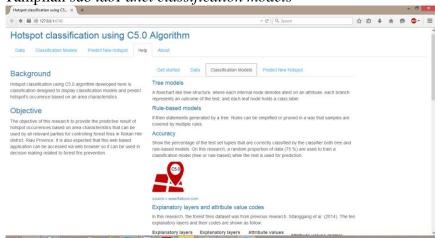
IFB Offiversity

Lampiran 7 Tampilan tabPanel help pada aplikasi berbasis web

Tampilan sub *tabPanel* data



Tampilan sub tabPanel classification models



Tampilan sub tabPanel predict new hotspot





Lampiran 8 Hasil pengujian fungsi aplikasi berbasis web

No	Kelas uji	Deskripsi uji	Kondisi awal	Skenario uji	Hasil yang diperoleh	Hasil uji
@Hak	tabPanel data	Menampilkan halaman utama	Halaman utama	Pilih tabPanel data	Tampilan halaman utama	OK
cipta milik 1		Mengunduh dataset kebakaran hutan	Halaman utama	Pilih choose file button	Dataset yang telah berhasil diunduh	OK
@Hak cipta milik IPB University		Menampilkan ringkasan data	Halaman utama dan dataset yang telah diunduh	Pilih tabPanel summary	Tampilan ringkasan data	OK
ity		Menampilkan data dalam tabel	Halaman utama dan dataset yang telah diunduh	Pilih tabPanel table	Tampilan data dalam tabel	OK
2	tabPanel classifica tion models	Menampilkan hasil klasifikasi model pohon keputusan	Halaman utama dan dataset yang telah diunduh	Pilih tabPanel tree models	Tampilan model pohon keputusan	OK
		Menampilkan hasil klasifikasi model berbasis aturan	Halaman utama dan dataset yang telah diunduh	Pilih tabPanel rule-based models	Tampilan model berbasis aturan	ОК
3	tabPanel predict new hotspot	Menampilkan hasil prediksi titik panas	Halaman utama dan dataset yang telah diunduh	Pilih variabel penjelas sesuai dengan karakteristik wilayah dan tekan predict button	Tampilan hasil prediksi titik panas dengan model pohon keputusan dan model berbasis	OK
				button	berbasis aturan	

RIWAYAT HIDUP

Penulis dilahirkan di Tangerang pada tanggal 9 Maret 1993 dari pasangan Bapak Lopiyanto Siknun dan Ibu Maryani. Penulis merupakan anak terakhir dari dua bersaudara. Pada tahun 2011, penulis lulus dari SMA Negeri 3 Kabupaten Tangerang dan di tahun yang sama, penulis diterima sebagai mahasiswa di Departemen Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Pertanian Bogor melalui jalur Undangan (SNMPTN). Pada tahun 2013, penulis mendapatkan beasiswa Bank Indonesia.

Selama menjalani studi, penulis pernah menjadi asisten praktikum mata kuliah Penerapan Komputer semester genap tahun 2013 dan menjadi asisten praktikum Rekayasa Perangkat Lunak semester ganjil tahun 2014. Penulis pernah melaksanakan praktik kerja lapangan (PKL) di PT Zenith Solusi Global pada tahun 2014 dengan mengembangkan Web API menggunakan framework .NET. Selain itu penulis pernah mengikuti pekan kreativitas mahasiswa bidang karsa cipta (PKM-KC) pada tahun 2013/2014 yang lolos didanai DIKTI.