

PENERAPAN ALGORITME POHON KEPUTUSAN C5.0 UNTUK KLASIFIKASI LAHAN GAMBUT YANG TERBAKAR DI KABUPATEN OGAN KOMERING ILIR

Meliana O. Meo

Program Studi Teknik Informatika, STIKOM Uyelindo Kupang
E-mail: meliana.oktavia.g@gmail.com

ABSTRACT

Lahan gambut merupakan salah satu tipe lahan basah yang memiliki peranan yang sangat penting dalam kehidupan makhluk hidup. Selain sebagai media penyimpanan air dimusim hujan juga merupakan habitat bagi hewan dan ikan. Luasan lahan gambut di Indonesia saat ini mengalami degradasi salah satunya penyebabnya dikarenakan banyaknya kebakaran yang terjadi. Kebakaran dilahan gambut jauh lebih berbahaya dibandingkan kebakaran non gambut selain itu dampak dari kebakaran lahan gambut sangat merugikan masyarakat. Salah satu upaya yang dapat dilakukan untuk mengamati lokasi, luasan dan dampak dari kebakaran gambut yaitu dengan memanfaatkan teknologi *remote sensing*. Dari satelit, data lokasi kebakaran, intensitas kebakaran dan luas area yang terbakar dapat diketahui dengan mudah dan efektif. Citra satelit yang dihasilkan dari *remote sensing* dapat dimanfaatkan melalui proses klasifikasi. Tujuan dari penelitian ini adalah membangun model klasifikasi menggunakan algoritme C5.0 untuk mengklasifikasi citra satelit lahan gambut sebelum terbakar, terbakar, dan setelah terbakar. Area yang digunakan adalah Kabupaten Ogan Komering Ilir, Provinsi Sumatera Selatan. Hasil dari penelitian ini menunjukkan bahwa algoritme C5.0 memiliki akurasi sebesar 97.10% dan nilai Kappa sebesar 0.96. Selain itu, hasil penelitian menunjukkan bahwa total estimasi luasan lahan gambut di kabupaten Ogan Komering Ilir, Sumatera Selatan pada tanggal 6 September 2015 dengan menggunakan algoritme C5.0 adalah 7.119,995 km² pada kelas sebelum terbakar, 689.895 km² pada kelas terbakar dan 2.155,300 km² pada kelas setelah terbakar.

Kata Kunci: C5.0, Lahan Gambut, Kebakaran Hutan, Klasifikasi, Pohon Keputusan

Pendahuluan

Indonesia merupakan salah satu negara tropis di dunia yang memiliki lahan gambut terluas yaitu seluas 17-27 juta hektar yang tersebar di beberapa Pulau diantaranya pulau Sumatera, Kalimantan dan Papua [1]. Lahan gambut memiliki peranan yang sangat penting dalam menunjang kehidupan makhluk hidup. Selain sebagai media penyimpanan air dimusim hujan juga merupakan habitat bagi hewan dan ikan.

Saat ini luasan lahan gambut di Indonesia mengalami degradasi, salah satunya di wilayah Sumatera Selatan. Hal ini dikarenakan banyaknya kebakaran yang terjadi pada lahan gambut. Di wilayah Sumatera Selatan sendiri pada tahun 2015, kebakaran hutan mencapai 30.984,98 ha. [2]. Hal ini memberikan dampak yang buruk bagi lingkungan diantaranya hilangnya biomassa dan keanekaragaman hayati, terjadinya proses subsiden, hilangnya fungsi penyerapan karbon serta timbulnya kabut asap yang menyebabkan gangguan kesehatan dan transportasi [1].

Kebakaran yang terjadi dilahan gambut termasuk dalam tipe *ground fire* dimana api menjalar di bawah permukaan dengan pembakaran yang tidak menyala (*smoldering*). *Ground fire* menyebabkan pemadaman kebakaran sulit dilakukan karena api dapat bertahan lama dan menghasilkan asap yang tebal. Berdasarkan hal tersebut dapat disimpulkan bahwa kebakaran gambut lebih berbahaya dibandingkan dengan kebakaran non gambut.

Besarnya kerugian akibat kebakaran lahan gambut menuntut kita untuk melakukan usaha pencegahan dan pengendalian kebakaran secara terus menerus. Salah satu upaya yang dapat dilakukan untuk mengamati lokasi, luasan dan dampak dari kebakaran gambut yaitu dengan memanfaatkan teknologi *remote sensing*. Dari satelit, data lokasi kebakaran, intensitas kebakaran dan luas area yang terbakar dapat diketahui dengan mudah dan efektif [3]. Metode yang dapat digunakan antara lain dengan mengklasifikasikan data citra satelit. Salah satu algoritme yang digunakan untuk mengklasifikasikan citra satelit adalah pohon keputusan. Terdapat berbagai algoritme dalam pohon

keputusan diantaranya algoritme *Random Forest*, algoritme C5.0, algoritme CART dan lainnya.

Algoritme pohon keputusan dapat digunakan untuk mengklasifikasikan data yang besar, namun hasil klasifikasi pohon keputusan cenderung tidak stabil, karena perubahan-perubahan kecil pada data *learning* akan mempengaruhi hasil akurasi prediksi [4]. Untuk memperbaiki stabilitas dan kekuatan akurasi prediksi pohon klasifikasi dapat digunakan metode *boosting*. *Boosting* adalah teknik untuk menghasilkan dan menggabungkan beberapa pengklasifikasi untuk meningkatkan akurasi prediksi [5]. Salah satu algoritme yang telah menerapkan metode ini adalah algoritme C5.0

Berdasarkan pemaparan tersebut, maka pada penelitian ini akan dilakukan pengklasifikasian area lahan gambut yang terbakar di kabupaten Ogan Komering Ilir, Sumatera Selatan dari data citra Landsat 7. Klasifikasi citra sendiri dilakukan dengan menggunakan algoritme C5.0. Hasil klasifikasi dari algoritme ini akan digunakan untuk mengklasifikasikan dan menghitung estimasi luasan lahan atau wilayah yang belum terbakar, terbakar dan telah terbakar.

1. LANDASAN TEORI

1.1 Citra Satelit

Citra dalam bidang *remote sensing* merupakan gambaran sebagian permukaan bumi yang diperoleh dari sistem perekaman melalui sensor yang dipasang pada pesawat terbang ataupun satelit [6]. Citra digital adalah *array* angka-angka dalam bentuk dua dimensi. Citra digital dapat memiliki dimensi ketiga yang disebut dengan layer. Layer adalah suatu citra yang sama tetapi memiliki informasi yang berbeda dengan informasi pada layer-layer lainnya. Pada citra satelit layer yang divisualisasikan secara multispektral, layer berupa saluran atau band yang berbeda dari citra yang sama [7].

1.2 Konsep Dasar Satelit Landsat

Satelit Landsat merupakan satelit hasil program sumberdaya bumi milik Amerika Serikat yang dikembangkan oleh NASA (*the National Aeronautical and Space Administration*). Satelit ini pertama kali diluncurkan pada tanggal 23 Juli 1972 dengan nama ERTS-I (*Earth Resources Technology Satellite*) yang kemudian berganti nama menjadi Landsat I. Hingga saat ini, seri Landsat yang diluncurkan telah sampai pada seri Landsat 8. Kelebihan sensor pada Landsat ETM+ adalah menggunakan delapan saluran, enam saluran dititikberatkan untuk studi vegetasi, satu saluran untuk studi geologi dan satu saluran untuk sensor pankromatik seperti tampak pada Tabel 1.

Tabel 1. Karakteristik sensor Landsat 7 ETM+

| Band | Resolusi | Keterangan | Spektral (μm) |
|------|----------|------------|----------------------------|
| 1 | 30 m | Blue | 0.441 - 0.514 |
| 2 | 30 m | Green | 0.519 - 0.601 |
| 3 | 30 m | Red | 0.631 - 0.692 |
| 4 | 30 m | NIR | 0.772 - 0.898 |
| 5 | 30 m | SWIR-1 | 1.547 - 1.749 |
| 6 | 60 m | Thermal IR | 10.31 - 12.36 |
| 7 | 30 m | SWIR-2 | 2.064 - 2.345 |
| 8 | 15 m | Pan | 0.515 - 0.896 |

1.3 Klasifikasi Data Citra Satelit

Klasifikasi citra satelit merupakan salah satu teknik yang digunakan untuk mengekstrak informasi dari sejumlah besar gambar satelit. Pada klasifikasi citra dilakukan pemberian label atau kelas pada setiap piksel citra berdasarkan karakteristik spektral pada berbagai *bands* [8].

Terdapat dua metode klasifikasi citra satelit yaitu klasifikasi tak terbimbing dan klasifikasi terbimbing. Klasifikasi tak terbimbing akan mencari kelompok-kelompok piksel dari citra satelit kemudian menandai setiap piksel ke dalam sebuah kelas berdasarkan parameter-parameter pengelompokan awal yang didefinisikan oleh penggunaannya. Sedangkan klasifikasi terbimbing memerlukan inputan atau informasi yang dikenal sebagai data latih untuk memulai klasifikasi. Piksel atau kelompok piksel yang sesuai dijadikan sebagai data latih, piksel-piksel tersebut digunakan untuk melatih dan mengenali piksel serupa lainnya [9].

1.4 Algoritma C5.0

Algoritme C5.0 merupakan perbaikan dari C4.5. C5.0 lebih baik dari C4.5 dalam hal kecepatan, efisiensi penggunaan memori, ukuran pohon keputusan dan kesalahan klasifikasi [10]. Pada algoritme C5.0, data latih awal dianggap sebagai *node root* dari pohon keputusan, kemudian dihitung setiap *gain ratio* setiap atribut. Untuk menghitung entropi informasi, digunakan Persamaan 1 [11].

$$E(S) = -\sum_{i=1}^m P_i \log_2(P_i) \quad (1)$$

dimana S set data yang terdiri atas n data sampel. C adalah kategori atribut yang memiliki m nilai berbeda. p_i adalah proporsi yang bisa dihitung dengan $p_i = \frac{n_i}{|S|}$, n_i adalah jumlah data yang termasuk kelas ke- i dan $|S|$ adalah banyaknya data pada data set S (maka $|S|=n$). Untuk menghitung entropi bersyarat atribut A digunakan persamaan 2 [11].

$$E(S|A) = -\sum_{j=1}^v P_j' \sum_{i=1}^m p_{ij} \log_2(P_{ij}) \quad (2)$$

dimana P_j' adalah proporsi yang dapat dihitung dengan $P_j' = \frac{|S_j|}{S} = \frac{\sum_{i=1}^m n_{ij}}{n}$; P_j' adalah peluang

bersyarat yang dapat dihitung dengan $p_{ij} = \frac{n_{ij}}{|S_j|}$; dan $|S_j|$ adalah jumlah data dengan atribut A. Maka, nilai *gain* dari atribut A dapat dihitung dengan persamaan 3 [11].

$$Gain = E(A) - E(S|A) \quad (3)$$

Nilai *gain ratio* dari atribut A dihitung dengan Persamaan 4 [11].

$$Gain\ Ratio(A) = \frac{Gain(A)}{Split(A)} \quad (4)$$

dimana $Split(A) = -\sum_{j=1}^p p_j' \log_2(p_j')$. Algoritme C5.0 memecah data latih berdasarkan atribut yang memiliki nilai informasi *gain* terbesar. Prosedur *split* terus dilakukan hingga tidak ada lagi *subset* data yang dapat di-*split*.

1.5 Area Studi dan data penelitian

- Area studi yang akan digunakan dalam penelitian ini adalah lahan gambut di Kabupaten Ogan Komering Ilir, Provinsi Sumatera Selatan.
- Data penelitian berupa :
 - Data citra satelit Landsat 7 ETM+ di kabupaten Ogan Komering Ilir yang diperoleh dari website <http://earthexplorer.usgs.gov>.
 - Data peta lahan gambut di pulau Sumatera tahun 2002 diperoleh dari Wetlands International Programme Indonesia
 - Data titik panas bulan September tahun 2015 diperoleh dari FIRMS MODIS Fire/Hotspot, NASA/University of Maryland

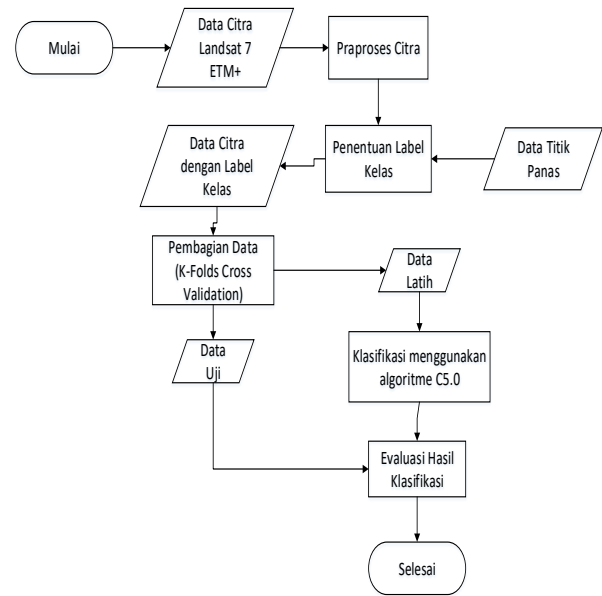
1.6 Peralatan Penelitian

Dalam penelitian ini, perangkat lunak yang digunakan yaitu :

- R sebagai bahasa pemrograman yang digunakan untuk klasifikasi data
- ILWIS dan Quantum GIS 2.6.0 untuk praproses citra satelit.
- PostgreSQL 9.1 digunakan untuk mengelola data spasial

1.7 Tahapan Penelitian

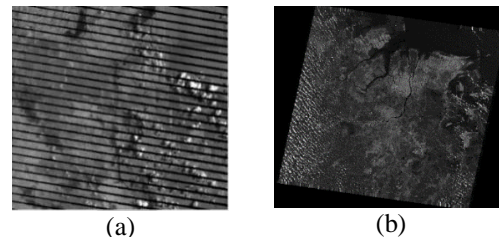
Penelitian ini terdiri dari 3 tahapan utama yaitu praproses citra, klasifikasi citra satelit dan evaluasi hasil klasifikasi.



Gambar 1 Tahapan penelitian

a. Praproses Citra Satelit

Pada tahap ini dilakukan proses pengisian *gap*, proses kombinasi *band*, georeferensi, proses *overlay* citra satelit dengan peta gambut. Proses pengisian *gap* dilakukan untuk memperbaiki data citra yang mengalami kerusakan akibat terdapatnya *gap* pada citra satelit.



Gambar 2 Proses pengisian *gap* (a) Citra satelit sebelum pengisian *gap* (b) Citra satelit setelah pengisian *gap*

Proses kombinasi *band* dilakukan untuk mendapatkan warna RGB dan kombinasi *band* yang digunakan. Dalam penelitian ini kombinasi band yang digunakan adalah *band 7*, *band 4*, dan *band 2* yang mengacu pada standar dari NASA. *Band 7* direpresentasikan dengan warna merah, *band 4* direpresentasikan dengan warna hijau, dan *band 2* direpresentasikan dengan warna biru.

Proses georeferensi ini dilakukan dengan tujuan untuk menyamakan sistem koordinat yang digunakan pada citra satelit yang memiliki sistem proyeksi koordinat UTM Zone 48S with WGS84, hal ini berarti Kabupaten Ogan Komering Ilir, Provinsi Sumatera

Selatan terletak pada zona 48S dalam sistem proyeksi UTM (*Universal Transverse Mercator*) dengan sistem referensi geospasial WGS84.

b. Klasifikasi Citra Satelit

Beberapa tahapan yang dilakukan sebelum dilakukan klasifikasi, yaitu proses penentuan kelas citra satelit, pembagian data latih dan data uji serta proses klasifikasi. Proses penentuan kelas citra satelit dilakukan dengan cara meng-*overlay* titik panas dengan citra satelit untuk mendapatkan label kelas yang dibutuhkan.

Titik panas yang digunakan dalam penelitian ini diambil dari tanggal 1 September 2015 sampai tanggal 6 September 2015 yang disesuaikan dengan tanggal akuisisi citra yaitu tanggal 6 September 2015. Terdapat empat buah kelas label yang akan dihasilkan, yaitu: sebelum terbakar, terbakar, setelah terbakar dan awan.

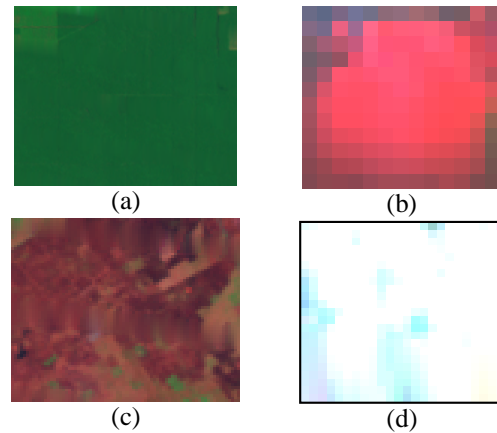
Tahapan selanjutnya adalah pengambilan sampel atau area contoh. Sampel piksel yang diambil berdasarkan kenampakan visual pada *display monitor*, dimana kelas sebelum terbakar ditandai dengan warna hijau, kelas terbakar ditandai dengan merah terang, kelas setelah terbakar ditandai dengan merah kecoklatan dan kelas awan ditandai dengan warna Putih. Jumlah sampel yang diambil dari masing-masing kelas label yaitu sebanyak 10 sampel. Tabel 2 menunjukkan banyaknya jumlah piksel setiap kelas yang menjadi area contoh.

Tabel 2. Jumlah piksel area contoh

| Kelas Tutupan Lahan | Jumlah Sampel | Jumlah Piksel |
|---------------------|---------------|---------------|
| Sebelum Terbakar | 10 | 2.500 |
| Terbakar | 10 | 2.343 |
| Setelah Terbakar | 10 | 2.502 |
| Awan | 10 | 2.400 |

Sampel piksel yang ada diubah ke dalam nilai digital setelah itu dibagi menjadi data latih dan data uji. Pembagian data dilakukan dengan menggunakan *K-fold cross validation*. Data dibagi menjadi 10 kelompok percobaan dengan porsi 9/10 data digunakan sebagai data latih dan 1/10 data digunakan sebagai data uji. Artinya dari 9.745 piksel dataset diambil 8.770 piksel atau 8.771 piksel sebagai data latih dan sisanya sebanyak 975 piksel atau 974 piksel sebagai data uji

Gambar 3 merupakan contoh kelas untuk kebakaran lahan.



Gambar 3 Contoh kelas tutupan lahan (a) Kelas sebelum terbakar (b) kelas terbakar (c) Kelas setelah terbakar (d) Kelas awan

Setelah proses penentuan label kelas, tahapan selanjutnya adalah pembagian data latih dan data uji. Pembagian data dilakukan dengan menggunakan metode *K-fold cross validation*. Metode ini akan membagi data menjadi *K* kelompok dan akan mengulang percobaan sebanyak *K*-kali. Data akan dibagi menjadi 10 kelompok percobaan dengan proporsi 90% data latih dan 10% data uji. Tahapan akhir setelah pembagian data yaitu proses klasifikasi. Proses klasifikasi dilakukan dengan menerapkan algoritme *C5.0*.

c. Evaluasi Hasil Klasifikasi

Evaluasi hasil klasifikasi dilakukan dengan melihat nilai akurasi keseluruhan dan koefisien Kappa yang diperoleh dari perhitungan *confusion matrix*.

Akurasi keseluruhan menunjukkan banyaknya jumlah piksel yang terklasifikasi secara benar pada tiap kelas dibanding jumlah sampel yang digunakan untuk uji akurasi pada semua kelas sedangkan Kappa dapat digunakan untuk mengukur kebenaran antara model dengan kenyataan atau menghitung jumlah nilai yang ada dalam perhitungan *Confusion matrix* [12]

$$Akurasi = \frac{\sum_{k=1}^q n_{kk}}{n} \times 100 \tag{5}$$

$$Kappa = \frac{n \sum_{k=1}^q n_{kk} - \sum_{k=1}^q n_{k+} \times n_{+k}}{n^2 - \sum_{k=1}^q n_{k+} \times n_{+k}} \tag{6}$$

dimana:

n_{kk} = nilai diagonal dari matrik kontingensi baris ke-k dan kolom ke- n_{+k}

n_{+k} = jumlah piksel dalam kolom ke-k

n_{k+} = jumlah piksel dalam baris ke-k

n = banyaknya piksel dalam contoh

2. HASIL DAN PEMBAHASAN

Algoritme C5.0 menghasilkan model klasifikasi berupa model pohon keputusan dan model berbasis aturan. Model berbasis aturan memiliki banyak aturan yang dapat disederhanakan dan dipangkas sehingga aturan yang diturunkan dapat berjumlah sedikit dari aturan yang dihasilkan oleh model berbasis pohon keputusan. Dari tahapan implementasi menggunakan algoritme C5.0 dengan menggunakan perangkat lunak R, diperoleh model berbasis pohon keputusan untuk *dataset* tiap *fold*. Akurasi model berbasis pohon keputusan dapat dilihat pada Tabel 3 sedangkan akurasi model berbasis aturan dapat dilihat pada Tabel 4.

Tabel 3. Akurasi rata-rata model berbasis pohon keputusan

| <i>Fold</i> | Ukuran Pohon | Koefisien Kappa | Akurasi Keseluruhan (%) |
|-------------|--------------|-----------------|-------------------------|
| 1 | 63 | 0.95 | 96.41 |
| 2 | 63 | 0.95 | 96.31 |
| 3 | 59 | 0.97 | 97.54 |
| 4 | 63 | 0.97 | 97.54 |
| 5 | 60 | 0.96 | 97.33 |
| 6 | 55 | 0.96 | 97.13 |
| 7 | 58 | 0.96 | 97.23 |
| 8 | 54 | 0.96 | 96.92 |
| 9 | 60 | 0.97 | 97.64 |
| 10 | 45 | 0.96 | 97.33 |
| Rata-Rata | | 0.96 | 97.14 |

Akurasi rata-rata yang diperoleh dari model pohon keputusan dari 10-*fold* pada Tabel 3 diatas sebesar 97.14% dan nilai koefisien Kappa sebesar 0.96. Sedangkan akurasi rata-rata model berbasis aturan sebesar 97.10% dan nilai koefisien Kappa sebesar 0.96 seperti diberikan pada Tabel 4.

Tabel 4. Akurasi rata-rata model berbasis aturan

| <i>Fold</i> | Jumlah Aturan | Koefisien Kappa | Akurasi Keseluruhan (%) |
|-------------|---------------|-----------------|-------------------------|
| 1 | 32 | 0.95 | 96.21 |
| 2 | 31 | 0.95 | 96.21 |
| 3 | 28 | 0.97 | 97.44 |
| 4 | 28 | 0.96 | 97.13 |
| 5 | 33 | 0.96 | 97.33 |
| 6 | 26 | 0.96 | 97.33 |
| 7 | 31 | 0.97 | 97.53 |
| 8 | 32 | 0.96 | 97.02 |
| 9 | 27 | 0.97 | 97.64 |
| 10 | 27 | 0.96 | 97.12 |
| Rata-Rata | | 0.96 | 97.10 |

Kecepatan dari suatu algoritme dapat dilihat dari penerapan model pada data baru yang akan diklasifikasi. Model yang dihasilkan dari pohon

keputusan berupa sejumlah aturan. Semakin besar jumlah pohon dan aturan yang dihasilkan, maka semakin lama proses pengklasifikasian dilakukan.

Berdasarkan hasil pada Tabel 3 dan Tabel 4, dapat dilihat bahwa ukuran pohon keputusan dan jumlah aturan yang dihasilkan pada *fold* ke-10 lebih sedikit dibandingkan dengan *fold* lainnya yaitu dengan ukuran pohon sebesar 45 pohon dan jumlah aturan sebanyak 27 aturan.

Klasifikasi dengan algoritme C5.0 menghasilkan 27 aturan. Aturan ini didapat dari pohon keputusan yang dihasilkan dari model klasifikasi. Berikut 27 aturan yang dihasilkan dari algoritme C5.0:

1. JIKA Band4 > 94 DAN Band2 > 83 MAKA Awan
2. JIKA Band7 ≤ 86 DAN Band4 > 94 DAN Band2 ≤ 71 MAKA Sebelum Terbakar
3. JIKA Band7 ≤ 40 DAN Band2 ≤ 82 MAKA Sebelum Terbakar
4. JIKA Band4 > 119 DAN Band2 ≤ 81 MAKA Sebelum Terbakar
5. JIKA 86 < Band7 ≤ 101 DAN Band4 > 94 DAN 69 < Band2 ≤ 83 MAKA Sebelum Terbakar
6. JIKA Band7 ≤ 82 DAN Band4 > 149 DAN Band2 ≤ 123 MAKA Sebelum Terbakar
7. JIKA 40 < Band7 ≤ 89 DAN Band4 > 73 DAN 61 < Band2 ≤ 66 MAKA Sebelum Terbakar
8. JIKA 40 < Band7 ≤ 73 DAN Band4 ≤ 87 Setelah Terbakar
9. JIKA Band7 ≤ 79 DAN Band4 ≤ 94 DAN Band2 > 82 Setelah Terbakar
10. JIKA Band7 > 85 DAN Band4 ≤ 61 DAN Band2 ≤ 82 Setelah Terbakar
11. JIKA 40 < Band7 ≤ 78 DAN Band4 ≤ 82 MAKA Setelah Terbakar
12. JIKA Band7 ≤ 44 DAN Band2 > 83 MAKA Setelah Terbakar
13. JIKA Band7 > 86 DAN Band2 ≤ 69 MAKA Setelah Terbakar
14. JIKA 101 < Band7 ≤ 166 DAN 94 < Band4 ≤ 119 DAN Band2 ≤ 81 MAKA Setelah Terbakar
15. JIKA Band4 ≤ 94 MAKA Setelah Terbakar
16. JIKA Band7 > 161 DAN Band4 ≤ 94 MAKA Terbakar
17. JIKA Band7 > 166 DAN Band4 ≤ 149 MAKA Terbakar

18. JIKA Band7 > 79 DAN Band4 ≤ 70 DAN Band2 > 82 MAKA Terbakar
19. JIKA Band7 > 79 DAN Band4 ≤ 94 DAN Band2 > 87 MAKA Terbakar
20. JIKA Band7 > 79 DAN Band4 ≤ 73 DAN Band2 > 83 MAKA Terbakar
21. JIKA Band7 > 101 DAN 94 < Band4 ≤ 149 DAN Band2 > 81 MAKA Terbakar
22. JIKA Band7 > 44 DAN 94 < Band4 ≤ 130 DAN Band2 > 83 MAKA Terbakar
23. JIKA Band7 > 78 DAN Band4 ≤ 69 DAN Band2 > 78 MAKA Terbakar
24. JIKA 78 < Band7 ≤ 89 DAN 61 < Band4 ≤ 66 DAN Band2 > 61 MAKA Terbakar
25. JIKA 78 < Band7 ≤ 103 DAN 69 < Band4 ≤ 94 DAN Band2 > 75 MAKA Terbakar
26. JIKA 78 < Band7 ≤ 89 DAN 55 < Band4 ≤ 73 DAN 61 < Band2 ≤ 69 MAKA Terbakar
27. JIKA Band7 ≤ 86 DAN 94 < Band4 ≤ 122 DAN 71 < Band2 ≤ 83 MAKA Terbakar

Dari aturan-aturan yang dihasilkan oleh algoritme C5.0 dapat disimpulkan bahwa kelas sebelum terbakar memiliki nilai band 4 yang lebih besar dari band 7, kelas terbakar memiliki nilai band 7 yang lebih besar dari band 4, kelas setelah terbakar berada di pertengahan nilai band dan kelas awan memiliki nilai band 2 yang lebih besar dari band lainnya.

Aturan yang dihasilkan dari algoritme ini dapat dimanfaatkan untuk mendeteksi adanya kebakaran di lahan gambut pada citra satelit. Aturan ini dapat memberikan kemudahan dalam mendapatkan informasi mengenai area yang terbakar, sebelum terbakar, dan telah terbakar. Kemudahan tersebut berupa waktu yang singkat untuk mendeteksi terjadinya kebakaran lahan gambut pada citra satelit.

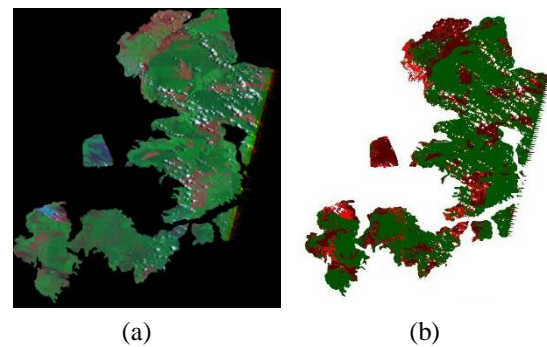
Tabel 5 merupakan *confusius matrix* dari algoritme C5.0 yang diperoleh dari salah satu data uji yang memiliki akurasi yang terbaik dari 10-fold.

Tabel 5. Confusion matrix untuk classifier dari algoritme C5.0

| Aktual | Prediksi | | | |
|------------------|------------------|----------|------------------|------|
| | Sebelum Terbakar | Terbakar | Setelah Terbakar | Awan |
| Sebelum Terbakar | 248 | 0 | 1 | 1 |
| Terbakar | 0 | 227 | 7 | 0 |
| Setelah Terbakar | 1 | 5 | 244 | 0 |

| | | | | |
|------|---|---|---|-----|
| Awan | 3 | 0 | 0 | 237 |
|------|---|---|---|-----|

Hasil confusion matrix yang didapatkan pada proses klasifikasi dari algoritme C5.0 menunjukkan bahwa terdapat kemiripan antara kelas setelah terbakar dengan kelas terbakar. Kemiripan antara kelas setelah terbakar dan kelas terbakar terletak pada warna yang ditunjukkan pada citra satelit. Lahan bekas terbakar memiliki warna merah kecoklatan dan warna merah identik dengan kelas terbakar. Berikut ini merupakan tampilan Hasil klasifikasi citra algoritme C5.0



Gambar 4 hasil klasifikasi citra (a) Citra asli (b) menggunakan algoritme C5.0

Berdasarkan hasil klasifikasi yang dilakukan dengan algoritme C5.0, dapat diketahui jumlah piksel dari masing-masing kelas. Jumlah piksel yang diperoleh dari masing-masing kelas dapat digunakan untuk mengestimasi luasan dari lahan gambut pada citra yang digunakan. Estimasi luasan lahan gambut dari masing-masing kelas diperoleh dengan cara mengalikan nilai piksel dengan resolusi spasial citra Landsat 7. Resolusi spasial citra Landsat 7 adalah $30 \times 30 m$ ($900 m^2$). Tabel 6 berikut ini merupakan estimasi luasan lahan gambut yang terbakar dengan menggunakan algoritme C5.0

Tabel 6. Estimasi luasan lahan gambut dengan algoritme C5.0

| Kelas | Jumlah Piksel | Estimasi Luasan (km^2) | Estimasi Luasan (%) |
|------------------|---------------|----------------------------|---------------------|
| Sebelum Terbakar | 7.911.106 | 7.119,995 | 68.16 |
| Terbakar | 766.550 | 689.895 | 6.60 |
| Setelah Terbakar | 2.394.778 | 2.155,300 | 20.63 |
| Awan | 53.976 | 481.478 | 4.61 |

3. SIMPULAN

Pengklasifikasi area lahan gambut yang terbakar di kabupaten Ogan Komering Ilir, Sumatera Selatan dengan menggunakan algoritme C5.0 menghasilkan nilai akurasi rata-rata sebesar 97.10% dan nilai Kappa sebesar 0.96. Selain itu hasil penelitian menunjukkan bahwa estimasi luasan lahan gambut pada kelas terbakar dengan mencapai 689.895 km² (6.60%). Sedangkan estimasi luasan lahan gambut untuk kelas setelah terbakar mencapai 2.155,300 km² (20.63%). Perbedaan tersebut terlihat karena adanya piksel-piksel pada kelas terbakar yang terdapat di antara kelas setelah terbakar dan sebaliknya.

International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol.XXXIX-B7. pp:422-426.

- [12] Foody GM. 2002. Status of Land Cover Classification Accuracy Assesment. *Remote Sensing of Environment*. 8:185-20

DAFTAR PUSTAKA

- [1] Syaufina L.. 2008. Kebakaran Hutan dan Lahan di Indonesia : Perilaku Api, Penyebab dan Dampak Kebakaran. Malang (ID) : Bayumedia Publishing
- [2] [KLH] Kementrian Lingkungan Hidup dan Kehutanan. 2015. Statistik Kementrian Lingkungan Hidup dan Kehutanan Tahun 2014. Jakarta (ID): Kementrian Lingkungan Hidup dan Kehutanan.
- [3] Justice CO, Malingreau JP, Setzer AW. 1993. Satellite remote sensing of fires: potential and limitations. P. Crutzen, & J. Goldammer (Eds.). pp:77-88.
- [4] Sutton CD. 2005. Classification and regression trees, Bagging, and Boosting, Handbook of statistics, 24(1): 303-329
- [5] Rulequest. 2012. C5.0: An Informal Tutorial. [Internet]. [diunduh 2016 Maret 23]. Tersedia pada: <https://www.rulequest.com/see5-unix.html>
- [6] Danoedoro P. 2012. *Pengantar Penginderaan Jauh Digital*. Yogyakarta (ID) : Penerbit Andi.
- [7] Liu JG, Mason PJ. 2009. Essential Image Processing and GIS for Remote Sensing. Chichester : John Wiley and Sons.
- [8] Sharma R, Ghosh A, Joshi PK. 2013. Decision Tree Approach for Classification of Remotely Sensed Satellite Data Using Open Source Support. *Journal Earth System Indian Academy of Sciences*. 122(5):1237-1247
- [9] Lillesand TM, Kiefer RW, Chipman JW. 2004. *Remote Sensing and Image Interpretation*, USA: John Wiley & Sons, Inc
- [10] Pandya R, Pandya J. 2015. C5.0 Algorithm to Improved Decision Tree with Feature Selection and Reduced Error Pruning. *International Journal of Computer Application*. 117(16): 18-21.
- [11] Zhai L, Sun J, Sang H, Yang G, Jia Y. 2012. Learning spatial decision tree for geographical classification: a summary of results.