

Klasifikasi Tutupan Lahan Menggunakan Supervised Machine Learning Pada Citra Satelit Menggunakan Google Earth Engine

1st Dhova Afasel

Fakultas Teknik Elektro
Universitas Telkom
Bandung, Indonesia

dhovaafasel@students.telkomuniversity.ac.id

2nd Rita Purnamasari

Fakultas Teknik Elektro
Universitas Telkom
Bandung, Indonesia

ritapurnamasari@telkomuniversity.ac.id

3rd Edwar

Fakultas Teknik Elektro
Universitas Telkom
Bandung, Indonesia

edwarm@telkomuniversity.ac.id

Abstrak— Informasi tutupan lahan telah digunakan berbagai pihak seperti pemerintah, peneliti, dan lain-lain. Permukaan bumi tidak homogen sama sekali, mengandung variasi seperti air, tanah kosong, rumput, aspal, dan lain-lain. Tutupan lahan paling umum biasanya adalah kategori urban (perkotaan) dan keperluan pertanian. Dan dari aktivitas maupun perilaku tersebut akan mengakibatkan berkurangnya lahan hijau sehingga menjadi tidak terkontrol. Perencanaan dan pemantauan penggunaan lahan masih menjadi masalah dibanyak negara berkembang salah satunya seperti Indonesia. Sistem yang dirancang adalah dengan metode supervised machine learning dengan metode klasifikasi Minimum Distance, CART, dan Gradient Boost pada daerah Kec. Cihampelas, Kabupaten Bandung Barat, Jawa Barat. Penelitian ini melakukan perhitungan uji akurasi pada tutupan lahan yang diambil dari citra satelit Sentinel 2, dengan menggunakan parameter perhitungan user accuracy, consumer accuracy, overall accuracy dan kappa statistic. Klasifikasi tutupan lahan yang diterapkan pada penelitian ini terdiri dari empat kelas yaitu badan air, lahan terbangun, lahan terbuka, dan vegetasi. Pada hasil pengujian didapatkan pada dari tanggal 1 Januari 2020 sampai 12 Desember 2020 menunjukkan hasil overall accuracy pada algoritma minimum distance yaitu 96.66%, pada algoritma CART yaitu 91.66%, dan pada algoritma gradient boost 90%. Algoritma minimum distance adalah algoritma terbaik dibandingkan metode lainnya untuk klasifikasi tutupan lahan, dengan tingkat overall accuracy tertinggi.

Kata kunci— tutupan lahan, supervised learning, minimum distance, CART, gradient boost.

Abstract— Land cover information has been used by various parties such as the government, researchers, and others. The earth's surface is not homogeneous at all, containing variations such as water, bare land, grass, asphalt, and others. The most common land cover is usually in the category of urban (urban) and agricultural purposes. And from these activities and behavior will result in reduced green land so that it becomes uncontrollable. Land use planning and monitoring is still a problem in many developing countries, one of which is Indonesia. The system designed is the supervised machine learning method with the Minimum Distance, CART, and Gradient Boost classification methods in the Kec. Cihampelas, West Bandung Regency, West Java. This study calculates the accuracy test on land cover taken from Sentinel 2 satellite imagery, using the calculation parameters of user accuracy, consumer accuracy, overall accuracy and kappa statistics. The land cover classification applied in this study consisted of four classes, namely water bodies, built up land, open land, and vegetation. The test results obtained from January 1, 2020 to December 12, 2020 show the overall accuracy results on the minimum distance algorithm, which is 96.66%, on the CART algorithm, which is 91.66%, and on the gradient boost algorithm,

90%. The minimum distance algorithm is the best algorithm compared to other methods for land cover classification, with the highest overall accuracy level.

Keywords: land cover, supervised learning, minimum distance, CART, gradient boost.

I. PENDAHULUAN

A. Latar Belakang

Informasi tutupan lahan telah banyak digunakan berbagai pihak seperti pemerintah, peneliti, dan lain-lain. Permukaan bumi tidak homogen sama sekali, melainkan mengandung variasi seperti air, tanah kosong, rumput, aspal, dan lain-lain. Tutupan lahan paling umum biasanya adalah kategori urban (perkotaan) dan keperluan pertanian [1]. Untuk mengidentifikasi tutupan lahan, teknologi observasi lapangan yang masih konvensional untuk observasi daerah yang luas memerlukan waktu yang lama. Sedangkan pengidentifikasi secara otomatis dapat dilakukan dengan bantuan citra satelit dan deep learning. Menurut Wahyunto, Pemanfaatan citra landsat sudah banyak digunakan untuk beberapa kegiatan survei maupun penelitian. Cara ini dirasa lebih efektif dan juga tidak memerlukan waktu yang lama. Keuntungan menggunakan deep learning dengan data satelit sebagai masukannya adalah dapat memfasilitasi akuisisi data yang lebih cepat dan cakupan datanya yang lebih luas dibandingkan dengan teknologi konvensional atau observasi lapangan [2]. Penginderaan jauh dengan menggunakan citra satelit Sentinel-2 dapat dimanfaatkan untuk mengklasifikasikan jenis penggunaan lahan di area yang cukup luas dengan resolusi sedang. Melalui citra satelit, dapat dilakukan beberapa klasifikasi untuk penggunaan tutupan lahan. Diantaranya hutan, bukit, pemukiman, sungai, pertanian, dan lahan terbuka dengan area penelitian di area Kec. Cihampelas, Kabupaten Bandung Barat, Jawa Barat. Pada penelitian sebelumnya, dari Navavit Pongan yang menerapkan dan menguji metode Pixel-Based dan klasifikasi tutupan lahan Object-Oriented. menggunakan segmentasi superpixel berdasarkan algoritma Simple Non-Iterative Clustering (SNIC) untuk mengklasifikasikan kluster spasial, dan tiga algoritma umum digunakan yaitu Random Forest (RF), Classification and Regression Trees (CART), dan Gradient Boost (GB) untuk diterapkan di sawah wilayah Laos, kecamatan Luang, bagian dari Thung Kula Rong Hai, Thailand. Dalam penelitian ini, Object-Oriented lebih efisien daripada Pixel-Based secara keseluruhan dengan nilai nilai

overall accuracy dari random forest 83%, CART 79%, dan gradient boost 87% [3]. Kemudian penelitian berjudul dari Indrayani Jayanti yang melakukan evaluasi tutupan lahan dengan menggunakan dua metode klasifikasi terbimbing yaitu Maximum Likelihood Classification (MLC) dan Minimum Distance Classification (MDC). Kajian ini melakukan perhitungan uji akurasi klasifikasi menggunakan tabel error matrix dan confusion matrix. Kesimpulan dari kajian ini yaitu Maximum Likelihood Classification (MLC) lebih baik digunakan dalam memetakan tutupan lahan di Kota Langsa dibandingkan dengan metode Minimum Distance Classification (MDC) dengan nilai akurasi keseluruhan 86% [4]. Penelitian dari Eka Miranda melakukan metode klasifikasi berdasarkan Convolutional Network Neural Network (CNN) dan citra satelit Sentinel-2. Metode CNN yang digunakan mampu mengklasifikasi dengan baik yaitu overall accuracy 98.4% untuk data pelatihan dan data pengujian, serta rata-rata nilai akurasi user's accuracy dan producer's accuracy lebih besar dari 95% untuk tiap kelas tutupan lahan [5].

B. Rumusan Masalah

1. Bagaimana cara mendekripsi jenis tutupan lahan dan klasifikasi tutupan lahan melalui citra satelit Sentinel-2?
2. Bagaimana cara melakukan klasifikasi dengan algoritma *Minimum Distance*, *CART*, dan *Gradient Boost* menggunakan *software Google Earth Engine*?
3. Bagaimana hasil uji akurasi dan analisis klasifikasi tutupan lahan dengan algoritma *Minimum Distance*, *CART*, dan *Gradient Boost*?
4. Algoritma manakah yang memiliki tingkat akurasi paling bagus?

C. Tujuan Penelitian

1. Merancang dan mengimplementasikan suatu sistem untuk mengklasifikasi tutupan lahan menggunakan algoritma *Minimum Distance*, *CART*, dan *Gradient Boost*.
2. Mengetahui hasil dan analisis seberapa besar akurasi sistem klasifikasi tutupan lahan berdasarkan algoritma *Minimum Distance*, *CART*, dan *Gradient Boost*.
3. Membandingkan hasil akurasi sistem klasifikasi tutupan lahan berdasarkan algoritma *Minimum Distance*, *CART*, dan *Gradient Boost*.

D. Manfaat Penelitian

1. Mendapatkan informasi tutupan lahan dengan akurat pada wilayah tertentu yang ingin diteliti.
2. Proses mendapatkan informasi tutupan lahan yang memakan waktu yang singkat.
3. Mengolah data yang luas tanpa memakan memori penyimpanan pada perangkat keras.
4. Memberikan informasi mengenai metode algoritma terbaik untuk klasifikasi terbimbing.

E. Batasan Masalah

1. Kawasan yang akan dijadikan objek identifikasi adalah Kec. Cihampelas, Kab. Bandung Barat, Jawa Barat.
2. Citra yang digunakan pada penelitian ini adalah citra dari satelit Sentinel 2.
3. Jenis klasifikasi tutupan lahan yang menjadi objek penelitian adalah lahan air, lahan terbangun, lahan terbuka, dan vegetasi.
4. Data yang digunakan dari setiap kelas tutupan lahan yaitu 30 data latih dan 15 data uji.
5. File citra yang digunakan bersumber dari *software Google Earth Engine*.

II. KAJIAN TEORI

A. Tutupan Lahan

Tutupan lahan atau yang disebut *land cover* memiliki beberapa pengertian oleh para ahli, berikut definisi tentang tutupan lahan antara lain, menurut Lillesland dan Kiefer, Penutupan lahan berkaitan dengan jenis kenampakan yang ada dipermukaan bumi. Menurut Townshend dan Justice, Tutupan lahan adalah perwujudan secara fisik (visual) dari vegetasi, benda alam, dan unsur budaya yang ada dipermukaan bumi tanpa memperhatikan kegiatan manusia terhadap obyek tersebut. Menurut Barret dan Curtis, Permukaan bumi sebagian terdiri dari kenampakan alamiah (penutupan lahan) contohnya antara lain; vegetasi, salju.

B. Penginderaan Jauh

Penjelasan penginderaan jauh dari Lillesand dan Kiefer, Penginderaan jauh adalah ilmu dan seni untuk memperoleh informasi tentang suatu objek, daerah, atau fenomena melalui analisis data yang diperoleh tanpa kontak langsung dengan objek, daerah, atau fenomena yang dikaji. Konsep dasar penginderaan jauh terdiri atas beberapa elemen atau komponen yang meliputi sumber tenaga, atmosfer, interaksi tenaga dengan objek dipermukaan bumi, sensor, sistem pengolahan data, dan berbagai penggunaan data.

C. Satelit Sentinel-2

Satelit Sentinel-2 adalah satelit yang terdiri dari dua satelit konstelasi yaitu Sentinel-2A dan Sentinel-2B yang mengorbit kutub pada orbit sun-synchronous pada ketinggian 786 km. Sentinel -2 merupakan citra satelit dengan resolusi spasial sedang dengan swath yang lebar, revisit di lokasi yang sama setiap 5 hari. Satelit ini diluncurkan untuk memantau kondisi permukaan bumi, sehingga mampu memberikan informasi kondisi terkini bumi dari angkasa untuk aplikasi lingkungan dan keamanan. Citra yang dihasilkan oleh satelit Sentinel-2 memiliki resolusi spasial sebesar 10 meter untuk 4 band, 20 meter untuk 6 band, dan 3 band sisanya memiliki resolusi spasial sebesar 60 meter. Citra satelit Sentinel-2 juga memiliki 13 band multispektral, yang dibagi atas spektrum visible (coastal aerosol , merah, hijau) near infrared, dan sortwave infrared.

D. Pengolahan Citra Digital

Pengolahan citra digital merupakan suatu metode untuk mengolah citra yang berfungsi meningkatkan kualitas citra sehingga dapat dengan mudah diinterpretasikan oleh manusia ataupun komputer dalam berupa foto atau gambar bergerak [11]. Citra tersebut merupakan representasi dari citra dua dimensi yang umumnya direpresentasikan dengan nilai biner 8 bit. Nilai biner ini disimpan pada elemen citra yang sering disebut sebagai pixel [12].

E. Google Earth Engine

Google Earth Engine adalah platform berbasis cloud yang memudahkan untuk mengakses sumber daya komputasi berkinerja tinggi untuk memproses kumpulan data geospasial yang sangat besar, tanpa harus mengalami kesulitan IT saat ini [13]. Selain itu, dan tidak seperti pusat superkomputer lainnya, Earth Engine juga dirancang untuk membantu peneliti menyebarkan hasil mereka dengan mudah kepada peneliti lain, pembuat kebijakan, LSM, pekerja lapangan, dan bahkan masyarakat umum. Setelah algoritma dikembangkan di Earth Engine, pengguna dapat menghasilkan produk data sistematis atau menerapkan aplikasi interaktif yang didukung oleh sumber daya Earth Engine, tanpa perlu menjadi ahli dalam pengembangan aplikasi, pemrograman web, atau HTML.

F. Supervised Learning

Menurut Danoedoro, Supervised classification melibatkan interaksi analis secara intensif, dimana dilakukan proses identifikasi objek pada citra (training area). Pengambilan setiap sampel perlu dilakukan dengan mempertimbangkan pola spektral pada setiap panjang gelombang tertentu, sehingga diperoleh daerah acuan tertentu.

G. Algoritma Minimum Distance

Menurut Mukhaiyar, metode Minimum Distance merupakan metode klasifikasi terbimbing yang cara klasifikasinya berdasarkan dengan melihat dengan jarak terpendek antara nilai kecerahan suatu piksel dengan rata-rata kelas yang ingin diklasifikasikan. Nilai kecerahan suatu piksel didapat dengan cara menghitung nilai dari spektral rata-rata untuk setiap kategori yang dipilih, kemudian dapat dilihat jarak piksel yang ingin ditentukan kelasnya, dengan kata lain jarak yang terdekat terhadap nilai rata-rata tersebut akan menjadi kelas yang sama. Apabila jaraknya > dari jarak rata-rata dari setiap kategori maka piksel tersebut akan tidak terklasifikasi.

H. Algoritma Classification and Regression Trees

Classification and Regression Trees (CART) merupakan salah satu metode atau algoritma dari salah satu teknik eksplorasi data yaitu teknik pohon keputusan. CART terbilang sederhana namun merupakan metode yang kuat. CART bertujuan untuk mendapatkan suatu kelompok data yang akurat sebagai penciri dari suatu pengklasifikasian, selain itu CART digunakan untuk menggambarkan hubungan antara variable respon (variabel dependen atau tak bebas) dengan satu atau lebih variabel prediktor (variabel independen atau bebas). Model pohon yang dihasilkan bergantung pada skala variable respon, jika variabel respon

data berbentuk kontinu maka model pohon yang dihasilkan adalah regression trees (pohon regresi) sedangkan bila variabel respon mempunyai skala kategorik maka pohon yang dihasilkan adalah classification trees (pohon klasifikasi) [14].

I. Algoritma Gradient Boost

Gradient Boost biasanya digunakan dengan *decision tree* (terutama *Decision Tree CART*) dengan ukuran tetap sebagai *base learner*. Untuk kasus khusus ini Friedman mengusulkan modifikasi metode peningkatan gradien yang meningkatkan kualitas kecocokan setiap *base learner*.

III. METODE

A. Gambaran Umum Sistem

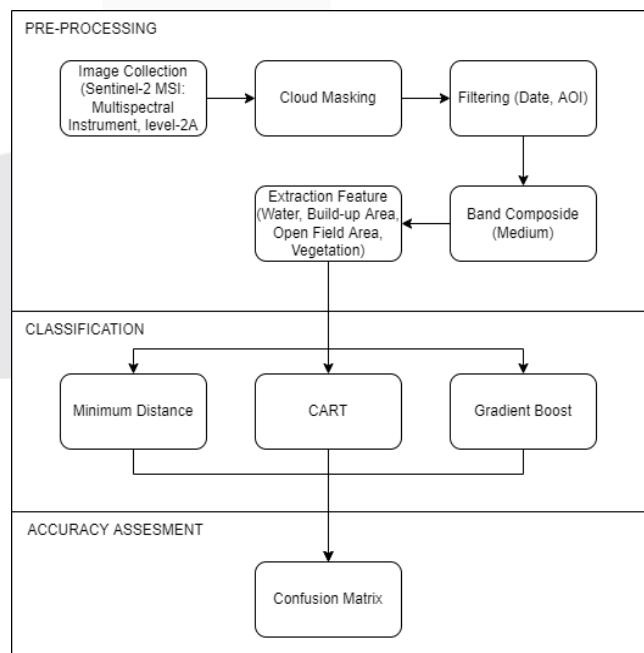
Klasifikasi pada tutupan lahan dilakukan pada gambar peta yang sudah disediakan *software Google Earth Engine* dimana gambar tersebut diproses seperti gambar 3.1.



GAMBAR 3.1

DIAGRAM BLOK SISTEM KESELURUHAN.

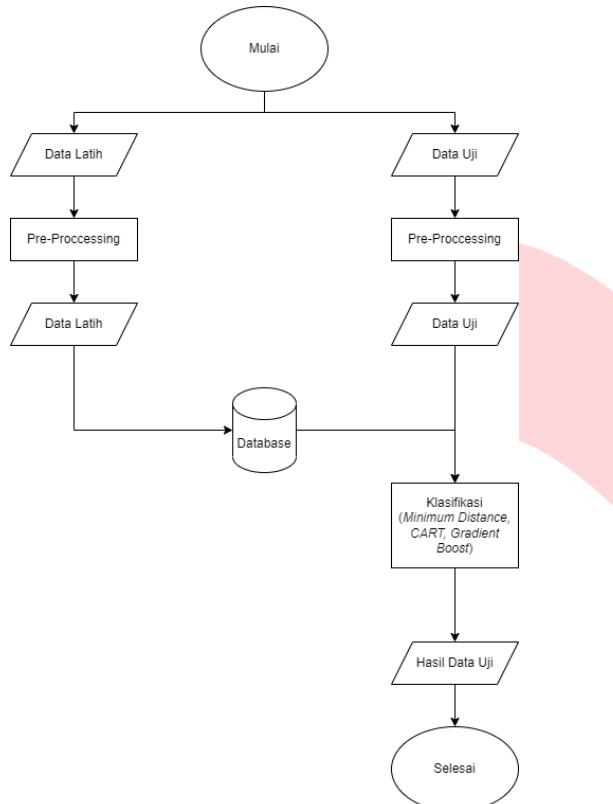
Tahap pertama adalah pemilihan data, yaitu data yang digunakan adalah jenis tutupan lahan berupa badan air, lahan terbuka, lahan terbangun, dan vegetasi. Kemudian jenis tutupan lahan dibagi menjadi dua tahapan, yaitu data latih dan data uji. Citra yang digunakan adalah citra hasil pemantauan satelit Sentinel-2. Data dilanjutkan kedalam proses *pre-processing*, yaitu pemilihan satelit, *cloud masking*, *filtering*, *band composite*, dan *extraction feature*. Setelah itu dilanjutkan dengan proses klasifikasi dengan algoritma *minimum distance*, algoritma *CART*, dan algoritma *gradient boost*. Dan yang terakhir hasil klasifikasi dalam bentuk *confusion matrix* untuk menghitung tingkat keakuratan klasifikasi.



GAMBAR 3.2
FLOWCHART PENELITIAN.

B. Diagram Alir Perancangan sistem

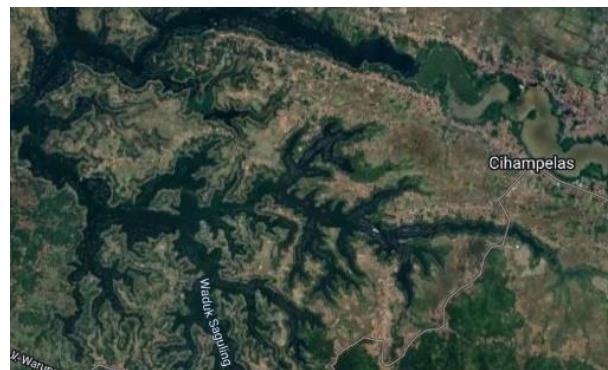
Pada penelitian ini, identifikasi sistem deteksi tutupan lahan dibagi menjadi dua tahapan, tahapan data latih dan tahapan data uji. Tahapan data latih berfungsi untuk melatih sistem untuk mendapatkan ciri dari area kelas yang dideteksi dan dijadikan database latih untuk tahap klasifikasi pada tahapan uji. Sedangkan tahapan data uji untuk menguji citra yang kemudian diklasifikasi oleh sistem.



GAMBAR 3. 3
DIAGRAM SISTEM.

C. Akuisisi Citra

Akuisisi citra merupakan proses awal yang dilakukan untuk mendapatkan citra digital sebagai alat data latih dan data uji. Proses akuisisi berupa pengambilan data citra yang kemudian diproses pada sistem. Citra yang digunakan adalah citra satelit diwilayah Kec. Cihampelas, Kabupaten Bandung Barat, Jawa Barat. Dimana daerah tersebut memenuhi parameter data yang akan dilatih dan diuji pada penelitian ini, yaitu vegetasi, badan air, lahan terbuka, dan lahan terbangun.



GAMBAR 3. 4
LOKASI STUDI AREA

D. Data Latih dan Data Uji

Pengambilan citra diambil pada sepanjang tahun 2020. Pengumpulan data latih dan data uji dikumpulkan dengan memasukkan input interpretasi visual melalui GEE menggunakan gambar asli sentinel-2. Pada penelitian ini pengambilan data terdiri dari empat kelas tutupan lahan (vegetasi, badan air, lahan terbuka, dan lahan terbangun).

TABEL 3. 1
JUMLAH POIN DATA SAMPLE DAN DATA UJI UNTUK SETIAP TUTUPAN LAHAN.

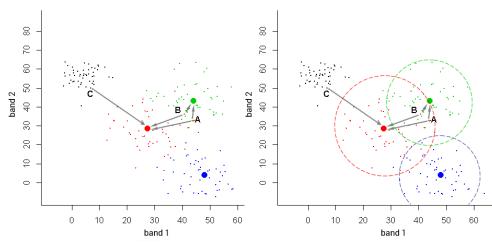
Jenis Tutupan Lahan	Jumlah data latih	Jumlah data uji
Vegetasi	30	15
Badan Air	30	15
Lahan Terbuka	30	15
Lahan Terbangun	30	15
Total	120	60

E. Pre-Processing

Citra satelit untuk klasifikasi tutupan lahan ini dibagi menjadi tiga langkah yang terdiri dari preparasi data, klasifikasi, dan confusion matrix. Nilai akurasi dari setiap klasifikasi yang diperoleh akan dibandingkan dan akan dipilih klasifikasi yang terbaik untuk penelitian ini. Proses preparasi data disini yaitu dimulai dengan citra satelit yang digunakan untuk klasifikasi diambil menggunakan Satelit Sentinel-2 Multispectral Instrument Level 2A (Sentinel-2 MSI). Resolusi spasial untuk penelitian ini adalah 10 m. Kemudian dilanjutkan dengan metode could masking, metode ini berfungsi meminimalisir hingga menghilangkan awan yang tertangkap dari satelit. Langkah selanjutnya adalah tanggal citra yang diinput dan area of interest (AOI), dimana proses ini adalah melakukan penentuan tanggal penangkapan citra yaitu mulai dari tanggal 01 Januari 2020 sampai 12 Desember 2020. Citra yang dipilih adalah tumpukan citra sepanjang tanggal yang sudah ditentukan. Proses mengambil citra terbaik dapat digunakan menggunakan instruksi memilih nilai median dalam tumpukan citra. Manfaat dari median ini menghilangkan awan yang memiliki nilai yang tinggi dan bayangan yang memiliki nilai yang tinggi.

F. Klasifikasi

1. Algoritma Minimum Distance



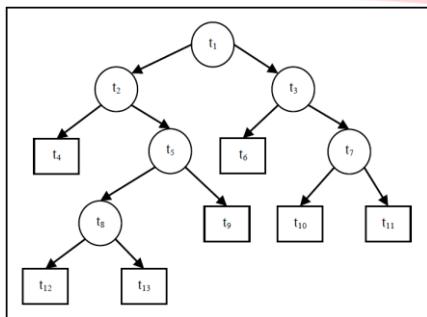
GAMBAR 3.5
ALGORITMA MINIMUM DISTANCE.

Gambar diatas adalah contoh imajiner dari algoritma *minimum distance* yang akan digunakan untuk membedakan kelas. Perintah untuk memasukkan *code* klasifikasi adalah sebagai berikut:

```
ee.Classifier.minimumDistance (metric, kNearest)
```

Membuat pengklasifikasi jarak minimum untuk metrik jarak yang diberikan. Dalam mode klasifikasi, kelas terdekat dikembalikan. Dalam mode regresi, jarak ke pusat kelas terdekat dikembalikan. Dalam mode RAW, jarak ke setiap pusat kelas dikembalikan.

2. Algoritma CART



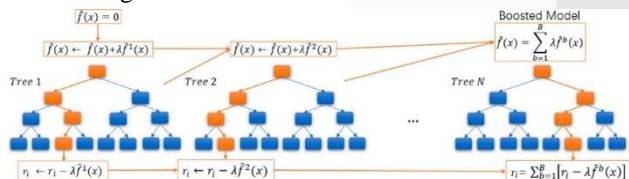
GAMBAR 3.6
ALGORITMA CART.

Gambar diatas menjelaskan contoh algoritma CART Struktur pohon klasifikasi. Perintah untuk memasukkan *code* klasifikasi adalah sebagai berikut:

```
ee.Classifier.smileCart (maxNodes, minLeafPopulation)
```

Membuat pengklasifikasi CART kosong.

3. Algoritma Gradient Boost



GAMBAR 3.7
ALGORITMA GRADIENT BOOST.

Gambar diatas adalah contoh algoritma *gradient boost*. Perintah untuk memasukkan *code* klasifikasi adalah sebagai berikut:

```
ee.Classifier.smileGradientTreeBoost (numberOfTrees, shrinkage, samplingRate, maxNodes, loss, seed)
```

Membuat pengklasifikasi Gradient tree Boost yang kosong.

Setiap kelas dilambangkan dengan bilangan integer sesuai dengan proses klasifikasi di GEE dan diawali dengan 0. Jumlah data latih dan data uji ditunjukkan pada tabel 3.2.

TABEL 3.2
LABEL AREA PELATIHAN.

Jenis tutupan lahan	Angka Area Pelatihan
Badan Air	0
Lahan Terbuka	1
Lahan Terbangun	2
Vegetasi	3

G. Pengujian

Uji akurasi hasil klasifikasi dilakukan untuk melihat tingkat akurasi yang dihasilkan dari klasifikasi tutupan lahan menggunakan GEE. Hasil dari pengujian dapat dilihat menggunakan *confusion matrix*. *Confusion matrix* dihitung dengan membandingkan antar basis kategori, hubungan antara data referensi yang menjadi data latih dan data uji. Dengan demikian matriks berbentuk *square* dengan angka pada baris dan kolom sama dengan angka pada kategori tingkat akurasi klasifikasi yang dinilai.

TABEL 3.3
CONTOH TABEL CONFUSION MATRIX.

KELAS TUTUPAN LAHAN		DATA UJI				JUMLAH
		A	B	C	D	
DATA LATIH	A	X _{jj}	X _{K+}
	B
	C
	D	X _{KK}	...
	JUMLAH	X _{K+}	X ₂₊	X ₃₊	X ₄₊	N

Overall Accuracy

$$\frac{\sum_{i=1}^r x_{ii}}{N} \cdot 100\% \quad (3.1)$$

User Accuracy

$$\frac{x_{jj}}{x_{k+}} \cdot 100\% \quad (3.2)$$

Consumer Accuracy

$$\frac{x_{kk}}{x_{k+}} \cdot 100\% \quad (3.3)$$

Kappa Statistic

$$\frac{N \sum_{i=1}^r x_{ii} - \sum_{i=1}^r x_{i+} \cdot x_{+i}}{N^2 - \sum x_{i+} \cdot x_{i+}} \cdot 100\% \quad (3.4)$$

Keterangan:

X_{ii} = Nilai diagonal dari matrix kontingensi baris ke-i dan kolom k-i

X_{i+} = Jumlah piksel dalam kolom ke-i

X_{+i} = jumlah piksel dalam baris ke-i

N = Banyaknya piksel dalam contoh.

IV. HASIL DAN PEMBAHASAN

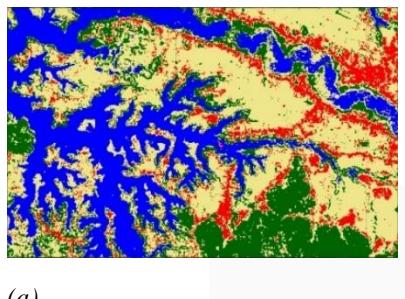
Perbandingan hasil performansi yang digunakan pada sistem ini adalah akurasi, *overall accuracy*, *produser accuracy*, *consumer accuracy*, dan *kappa statistic*. Pengujian skenario dilakukan untuk mendapatkan nilai-nilai hasil performansi terbaik. Proses klasifikasi citra dilakukan dengan menggunakan *software google earth engine code editor*, citra

satelit Sentinel-2 MultiSpectral Instrument Level 2A (Sentinel-2 MSI) dari tanggal 01 Januari 2020 sampai 12 Desember 2020 dengan resolusi spasial yang digunakan adalah 10 m. Pada gambar 4.1 adalah tampilan studi area yang akan diklasifikasi dari hasil satelit Sentinel-2. Pada gambar 4.2, dibawah adalah hasil pengujian pada setiap masing-masing algoritma klasifikasi yang sudah dipilih.

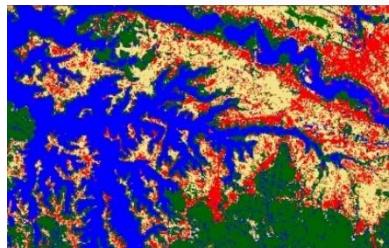


GAMBAR 4.1

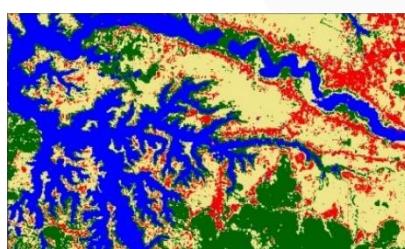
TAMPILAN STUDI AREA DARI CITRA SATELIT SENTINEL-2.



(a)



(b)



(c)

GAMBAR 4.2

TAMPILAN KLASIFIKASI BERDASARKAN STUDI AREA (A)
KLASIFIKASI CART (B) KLASIFIKASI MINIMUM DISTANCE (C)
KLASIFIKASI GRADIENT BOOST.

Keterangan warna dari hasil tampilan klasifikasi adalah pada badan air menggunakan warna biru, pada lahan

terbuka menggunakan warna khaki, pada lahan terbangun menggunakan warna merah, dan pada vegetasi menggunakan warna hijau gelap.

Identifikasi luas dari tiap-tiap kelas tutupan lahan yang diklasifikasi dapat dilihat pada tabel berikut.

TABEL 4. 1
TABEL LUAS TUTUPAN LAHAN ALGORITMA MINIMUM DISTANCE

Kode tutupan lahan	Jenis tutupan lahan	Luas tutupan lahan
0	Badan Air	16.59 km ²
1	Lahan Terbuka	26.67 km ²
2	Lahan Terbangun	6.41 km ²
3	Vegetasi	13.12 km ²
TOTAL		62.79 km ²

Hasil luasan jenis tutupan lahan di studi area kecamatan Cihampelas menggunakan algoritma *minimum distance* mendapatkan luasan terbesar pada jenis tutupan lahan lahan terbuka dengan luas 26.67 km². Jenis tutupan lahan terkecil lahan terbangun dengan luas 6.41 km².

TABEL 4. 2
TABEL LUAS TUTUPAN LAHAN ALGORITMA CART

Kode tutupan lahan	Jenis tutupan lahan	Luas tutupan lahan
0	Badan Air	23.95 km ²
1	Lahan Terbuka	14.97 km ²
2	Lahan Terbangun	11.53 km ²
3	Vegetasi	18.63 km ²
TOTAL		69.08 km ²

Hasil luasan jenis tutupan lahan di studi area kecamatan Cihampelas menggunakan algoritma CART mendapatkan luasan terbesar pada jenis tutupan lahan badan air dengan luas 23.95 km². Jenis tutupan lahan terkecil lahan terbangun dengan luas 11.53 km².

TABEL 4. 3
TABEL LUAS TUTUPAN LAHAN ALGORITMA GRADIENT BOOST

Kode tutupan lahan	Jenis tutupan lahan	Luas tutupan lahan
0	Badan Air	19.54 km ²
1	Lahan Terbuka	28.78 km ²
2	Lahan Terbangun	8.45 km ²
3	Vegetasi	19.46 km ²
TOTAL		76.23 km ²

Hasil luasan jenis tutupan lahan di studi area kecamatan Cihampelas menggunakan algoritma *gradient boost* mendapatkan luasan terbesar pada jenis tutupan lahan lahan terbuka dengan luas 28.78 km². Jenis tutupan lahan terkecil lahan terbangun dengan luas 8.45 km².

- A. Analisa Hasil Pengujian Confusion Matrix
- 1. Klasifikasi Minimum Distance

TABEL 4. 4
CONFUSION MATRIX MENGGUNAKAN ALGORITMA MINIMUM DISTANCE.

	Badan Air	Lahan Terbuka	Lahan Terbangun	Vegetasi
Badan Air	13	2	0	0
Lahan Terbuka	0	15	0	0
Lahan Terbangun	0	0	15	0
Vegetasi	0	0	0	15

Dari tabel diatas memberikan informasi bahwa nilai diagonal dengan warna yang ditebalkan adalah nilai dari piksel gambar yang berhasil diklasifikasi dengan benar, sedangkan untuk nilai dari tabel yang tidak diwarnai tebal adalah jumlah piksel yang salah atau mengalami *error*.

Kesalahan klasifikasi yang terjadi pada hasil klasifikasi terjadi dikarenakan kelas tersebut memiliki warna dan rona yang hampir sama dengan warna dan rona kelas lainnya. Seperti pada badan air, dimana terjadi kesalahan yaitu lahan terbuka. Dan juga dikarenakan kesalahan penempatan titik kelas tutupan lahan, contohnya jenis tutupan lahan terbuka ditandai dengan jenis tutupan lahan vegetasi. terdapat 13 titik benar dan 2 titik yang salah diklasifikasikan. Kelas dari badan air dan kelas lahan terbuka berdekatan sehingga memiliki warna dan rona yang hampir sama. Kemudian tidak terjadi kesalahan sama sekali pada kelas kelas lainnya.

TABEL 4. 5

TABEL NILAI KESELURUHAN PADA ALGORIMA MINIMUM DISTANCE.

Tutupan Lahan	Producer Accuracy (%)	Consumer Accuracy (%)	Overall Accuracy (%)	Kappa Statistic (%)
Badan air	86.66	100		
Lahan Terbuka	100	88.23		
Lahan Terbangun	100	100		
Vegetasi	100	100		

Dari tabel diatas dapat diketahui bahwa akurasi keseluruhan dari hasil klasifikasi *minimum distance* adalah 96.66 %.

2. Klasifikasi CART

TABEL 4. 6

CONFUSION MATRIX MENGGUNAKAN ALGORITMA CART.

	Badan Air	Lahan Terbuka	Lahan Terbangun	Vegetasi
Badan Air	15	0	0	0
Lahan Terbuka	0	12	2	1
Lahan Terbangun	0	1	14	0
Vegetasi	1	0	0	14

Dari tabel diatas memberikan informasi bahwa nilai diagonal dengan warna yang ditebalkan adalah nilai dari piksel gambar yang berhasil diklasifikasi dengan benar, sedangkan untuk nilai dari tabel yang tidak diwarnai tebal adalah jumlah piksel yang salah atau mengalami *error*.

Kesalahan klasifikasi yang terjadi pada hasil klasifikasi terjadi dikarenakan kelas tersebut memiliki warna

dan rona yang hampir sama dengan warna dan rona kelas lainnya. Dan juga dikarenakan kesalahan penempatan titik kelas tutupan lahan, contohnya jenis tutupan lahan terbuka ditandai dengan jenis tutupan lahan vegetasi. Seperti pada lahan terbuka dimana terdapat 2 kesalahan dengan lahan terbangun dan 1 kesalahan dengan vegetasi. Lahan terbangun yang memiliki 14 nilai yg benar tetapi memiliki 1 kesalahan dengan lahan terbuka. Dan vegetasi yang miliki 14 nilai yang benar tetapi memiliki 1 kesalahan dengan badan air. Kemudian tidak terjadi kesalahan sama sekali pada kelas badan air.

TABEL 4. 7
TABEL NILAI KESELURUHAN PADA ALGORIMA CART.

Tutupan Lahan	Producer Accuracy (%)	Consumer Accuracy (%)	Overall Accuracy (%)	Kappa Statistic (%)
Badan air	100	93.75	91.66	88.88
Lahan Terbuka	80	92.3		
Lahan Terbangun	93.33	87.5		
Vegetasi	93.33	93.3		

Dari tabel diatas dapat diketahui bahwa akurasi keseluruhan dari hasil klasifikasi CART adalah 91.66 %.

3. Klasifikasi Gradient Boost

Berikut ini adalah tabel dari *confusion matrix* dari algoritma gradient boost. Dengan memasukkan perintah mencari *confusion matrix* dari *software google earth engine* pada algoritma gradient boost, maka didapatkan hasil sebagai berikut.

TABEL 4. 8
CONFUSION MATRIX ALGORITMA GRADIENT BOOST.

	Badan Air	Lahan Terbuka	Lahan Terbangun	Vegetasi
Badan Air	13	0	0	2
Lahan Terbuka	0	12	3	0
Lahan Terbangun	0	1	14	0
Vegetasi	0	0	0	15

Dari tabel diatas memberikan informasi bahwa nilai diagonal dengan warna yang ditebalkan adalah nilai dari piksel gambar yang berhasil diklasifikasi dengan benar, sedangkan untuk nilai dari tabel yang tidak diwarnai tebal adalah jumlah piksel yang salah atau mengalami *error*.

Kesalahan klasifikasi yang terjadi pada hasil klasifikasi terjadi dikarenakan kelas tersebut memiliki warna dan rona yang hampir sama dengan warna dan rona kelas lainnya. Dan juga dikarenakan kesalahan penempatan titik kelas tutupan lahan, contohnya jenis tutupan lahan terbuka ditandai dengan jenis tutupan lahan vegetasi. Kelas badan air mendapatkan 13 titik yang benar diklasifikasikan dan 2 kesalahan klasifikasi. Pada kelas lahan terbuka mendapatkan

12 titik yang benar dan 3 kesalahan klasifikasi. Pada kelas lahan terbangun mendapatkan 14 klasifikasi benar dan 1 kesalahan klasifikasi. Pada kelas vegetasi tidak terjadi kesalahan klasifikasi.

TABEL 4. 9 TABEL NILAI KESELURUHAN PADA ALGORITMA GRADIENT BOOST.

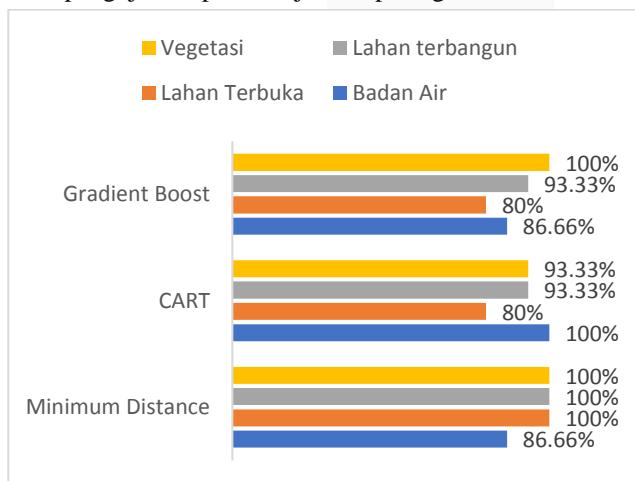
Tutupan Lahan	Producer Accuracy (%)	Consumer Accuracy (%)	Overall Accuracy (%)	Kappa Statistic (%)
Badan air	86.66	100	90	86.66
Lahan Terbuka	80	92.3		
Lahan Terbangun	93.33	82.35		
Vegetasi	100	88.23		

Dari tabel diatas dapat diketahui bahwa akurasi keseluruhan dari hasil klasifikasi *gradient boost* adalah 90 %.

B. Perbandingan Hasil dan Analisis Pengujian Sistem

1. Producer Accuracy

Pada skenario ini dilakukan pengujian perbandingan hasil *producer accuracy* agar didapatkan jenis klasifikasi terbaik untuk *supervised machine learning* menggunakan *google earth engine*. Pengujian dilakukan dengan memasukkan data latih sebanyak 30 data latih dan 15 data uji dari setiap jenis tutupan lahan, kemudian dilanjutkan memasukkan perintah untuk mendapatkan hasil *producer accuracy* dan dilanjutkan dengan menjalankan simulasi. Hasil pengujian dapat ditunjukkan pada gambar 4.5.



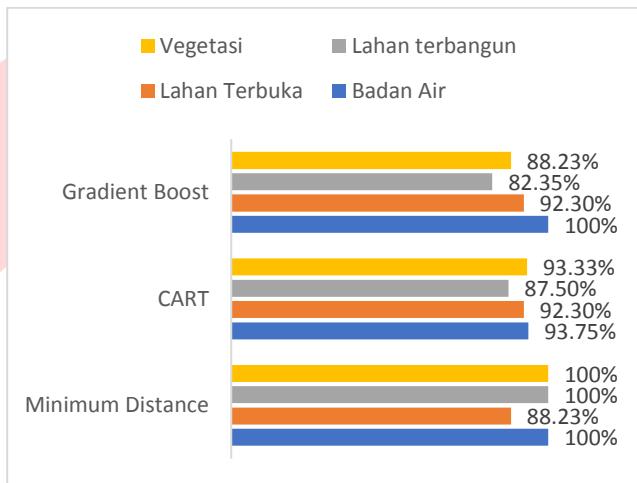
GAMBAR 4. 1
PERSENTASE PRODUCER ACCURACY.

Hasil pengujian *producer accuracy* pada masing-masing klasifikasi menunjukkan hasil terendah didapat ketika menggunakan klasifikasi *CART*, bisa dilihat pada tingkat akurasi vegetasi 93.33%, lahan terbangun 93.33%, lahan terbuka 80%, dan badan air 100%. Sedangkan untuk hasil tertinggi didapat ketika menggunakan *Minimum Distance* yang hampir sempurna yaitu pada tingkat akurasi vegetasi 100%, lahan terbangun 100%, lahan terbuka 100%, dan badan air 86.66%.

100%, lahan terbangun 100%, lahan terbuka 100%, dan badan air 86.66%.

2. Consumer Accuracy

Pada skenario ini dilakukan pengujian perbandingan hasil *Consumer Accuracy* agar didapatkan jenis klasifikasi terbaik untuk *supervised machine learning* menggunakan *google earth engine*. Pengujian dilakukan dengan memasukkan data latih sebanyak 30 data latih dan 15 data uji dari setiap jenis tutupan lahan, kemudian dilanjutkan memasukkan perintah untuk mendapatkan hasil *Consumer Accuracy* dan dilanjutkan dengan menjalankan simulasi. Hasil pengujian dapat ditunjukkan pada gambar 4.6.

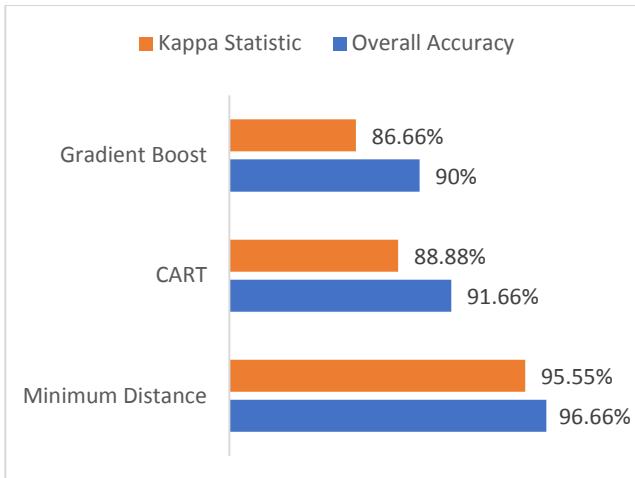


GAMBAR 4. 2
PERSENTASE CONSUMER ACCURACY.

Hasil pengujian *Consumer Accuracy* pada masing-masing klasifikasi menunjukkan hasil terendah didapat ketika menggunakan klasifikasi *Gradient Boost*, bisa dilihat pada tingkat akurasi vegetasi 88.23%, lahan terbangun 82.35%, lahan terbuka 92.30%, dan badan air 100%. Sedangkan untuk hasil tertinggi didapat ketika menggunakan *Minimum Distance* yang hampir sempurna yaitu pada tingkat akurasi vegetasi 100%, lahan terbangun 100%, lahan terbuka 88.23%, dan badan air 100%.

C. Hasil Klasifikasi Terbaik

Pada skenario terakhir, dilakukan pengujian perbandingan hasil *overall accuracy* dan *kappa statistic* agar didapatkan jenis klasifikasi terbaik untuk *supervised machine learning* menggunakan *google earth engine*. Pengujian dilakukan dengan memasukkan data latih sebanyak 30 data latih dan 15 data uji dari setiap jenis tutupan lahan, kemudian dilanjutkan memasukkan perintah untuk mendapatkan hasil *overall accuracy* dan *kappa statistic* dan dilanjutkan dengan menjalankan simulasi. Hasil pengujian dapat ditunjukkan pada gambar 4.5.



GAMBAR 4. 3

PERSENTASE OVERALL ACCURACY DAN KAPPA STATISTIC.

Hasil pengujian *overall accuracy* dan *kappa statistic* pada masing-masing klasifikasi menunjukkan akurasi terendah didapat ketika menggunakan klasifikasi *Gradient Boost* sebesar 90% untuk *overall accuracy* dan 86.66% *kappa statistic*. Sedangkan untuk hasil tertinggi didapat ketika menggunakan *Minimum Distance* sebesar 96.66% untuk *overall accuracy* dan 95.55% untuk *kappa statistic*.

V. KESIMPULAN DAN SARAN

A. Kesimpulan

1. Teknologi klasifikasi tutupan lahan pada *supervised machine learning* citra satelit menggunakan *google earth engine* menggunakan satelit sentinel 2 mendapatkan tingkat keakuratan yang tinggi dengan rata-rata akurasi 92.77 %.
2. Hasil klasifikasi terbimbing yang dilakukan pada citra satelit sentinel 2 diperoleh nilai *overall accuracy* pada algoritma minimum distance yaitu 96.66%, pada algoritma CART yaitu 91.66%, dan pada algoritma gradient boost 90%.
3. Dari segala aspek penelitian algoritma minimum distance adalah algoritma terbaik dari algoritma lainnya dengan tingkat *overall accuracy* dan *kappa statistic* tertinggi.

B. Saran

1. Melakukan perbandingan lain dengan algoritma klasifikasi terbimbing lainnya.
2. Menambahkan proses ekstraksi ciri pada data latih dan data uji.
3. Menambahkan proses segmentasi pada penelitian.
4. Melakukan perbandingan dengan algoritma tidak terbimbing.
5. Menggunakan perbandingan dengan citra satelit yang lainnya.

REFERENSI

- [1] Rudrapal, Dhriti, Subhedar, Mansi. Land Cover Classification using Support Vector Machine. International Journal of Engineering Research and. V4.10.17577/IJERTV4IS090611. 2015.
- [2] Q. Weng, Z. Mao, J. Lin, X. Liao. *Land-Use Scene Classification Based on A CNN Using A Constrained Extreme Learning Machine*. International Journal of Remote Sensing, pp. 1-18. 2018.
- [3] Ponganen, N., Horanont, T., Artlert, K., Nuallaong, P. Land Cover Classification using Google Earth Engine's Object-oriented and Machine Learning Classifier. 2021 2nd International Conference on Big Data Analytics and Practices, IBDAP 2021, 33–37. <https://doi.org/10.1109/IBDAP52511.2021.9552099>. 2021.
- [4] Indrayani, J. Perbandingan Metode Klasifikasi Maximum Likelihood Dan Minimum Distance Pada Pemetaan Tutupan Lahan Di Kota Langsa. Tugas Akhir. Fakultas Matematika dan Ilmu Pengetahuan Alam. Universitas Syiah Kuala Darussalam: Banda Aceh. 2017.
- [5] Mulyoutami E, Sakuntaladewi N, Agus F, van Noordwijk M. *Perubahan Pola Perladangan: Pergeseran persepsi mengenai para peladang di Indonesia*. Bogor, Indonesia. World Agroforestry Centre - ICRAF, SEA Regional Office. 101p. 2010.
- [6] Sutoyo, T., Mulyanto, E., Suhartono, V., Nurhayanti, D. dan Wijanarto. *Pengolahan Citra Digital*. Penerbit Andi. Yogyakarta. 2009.
- [7] E. Mandanici, G. Bitelli, "Preliminary comparison of sentinel-2 and landsat 8 imagery for a combined use", Remote Sensing vol. 8, issue.12, pp. 1014, 2016.
- [8] A Canada Centre for Remote Sensing, Fundamentals of Remote Sensing, A Canada Centre for Remote Sensing, 2016.
- [9] Sentinel-2 EO. "Sentinel Hub Data", [Online]. Available: <https://www.sentinel-hub.com/explore/data/> [Diakses 31 Agustus 2022].
- [10] H. Saadat, J. Adamowski, R. Bonnell, F. Sharifi, M. Namdar, S. Ale-Ebrahim, "Land use and land cover classification over a large area in Iran based on single date analysis of satellite imagery", ISPRS Journal of Photogrammetry and Remote Sensing, vol. 66, issue. 5, pp. 608–619, 2011.
- [11] Sutoyo, T., Mulyanto, E., Suhartono, V., Nurhayanti, D. dan Wijanarto. *Pengolahan Citra Digital*. Penerbit Andi. Yogyakarta. 2009.
- [12] Widodo, Sri. *Ekstraksi Fitur Citra Biomedik: Dilengkapi dengan Source Code Matlab*. (n.d.). (n.p.): Penerbit Pustaka Rumah Cinta. 2021.
- [13] Hamdi A. Zurqani, Christopher J. Post, Elena A. Mikhailova, Mark A. Schlautman, Julia L. Sharp, Geospatial analysis of land use change in the Savannah River Basin using Google Earth Engine, International Journal of Applied Earth Observation and Geoinformation, Volume 69, 2018, Pages 175-185, ISSN 1569-8432, <https://doi.org/10.1016/j.jag.2017.12.006>.

- [14] Breiman L., Friedman J.H Olshen R.A & Stone C.J. ClassificationAnd Regression Tree. New York, NY: Chapman And Hall. 1984.
- [15] Lewis, M.D dan Roger, J. An Introduction to Classification and Regression Tree (CART) Analysis. Presented at the 2000 Anual Meeting of Society For Academy Emergency Medicine in San Fransisco, California. 2000.
- [16] J. H. Friedman. Greedy Function Approximation: A Gradient Boosting Machine, 1999.
- [17] Jerome H. Friedman, Stochastic gradient boosting, Computational Statistics & Data Analysis, Volume 38, Issue 4, Pages 367-378, ISSN 0167-9473, [https://doi.org/10.1016/S0167-9473\(01\)00065-2](https://doi.org/10.1016/S0167-9473(01)00065-2). 2002.
- [18] Jiao, Z., Hu, P., Xu, H., & Wang, Q. Machine learning and deep learning in chemical health and safety: A systematic review of techniques and applications. In *Journal of Chemical Health and Safety* (Vol. 27, Issue 6, pp. 316–334). American Chemical Society. <https://doi.org/10.1021/acs.chas.0c00075>. 2020.
- [19] Pavel, U. (2017, June, 14) *SUPERVISED IMAGE CLASSIFICATION USING MINIMUM DISTANCE ALGORITHM* [online]. Available : <http://www.50northspatial.org/supervised-image-classification-using-minimum-distance-algorithm/> [Diakses : 13 Sepetember 2022].
- [20] James, G.; Witten, D.; Hastie, T.; Tibshirani, R. An Introduction to Statistical Learning: with Applications in R; Springer: New York, 2017.