

## KLASIFIKASI DATA TINGKAT KERAWANAN KEBAKARAN MENGGUNAKAN ALGORITMA CART

Nurahman<sup>1</sup>, Rika Sumanti<sup>2</sup>

<sup>1,2</sup>Program Studi Sistem Informasi, Universitas Darwan Ali

Jln. Batu Berlian No.10 Sampit – Kalimantan Tengah

<sup>1</sup>[nurahman@unda.ac.id](mailto:nurahman@unda.ac.id), <sup>2</sup>[rika.sumanti02@gmail.com](mailto:rika.sumanti02@gmail.com)

### Abstract

*Forest and land fires are urgent issues in the East Kotawaringin District. Their negative impacts include the destruction of forest and land ecosystems, leading to air pollution that significantly disrupts community activities. The objective of this research is to develop a classification model capable of predicting the level of fire vulnerability based on specific variables. The development of this classification model to predict the level of fire vulnerability is carried out using the CART algorithm. The CART algorithm is capable of classifying data on fire vulnerability by generating a decision tree. The results of this research conclude that the classification data using the CART algorithm calculation in the Rapidminer application shows an accuracy level of 79.89%, which is adequate in predicting the level of fire vulnerability.*

**Keywords :** *Fire Vulnerability, CART Algorithm, Decision Trees*

### Abstrak

Kebakaran hutan dan lahan merupakan permasalahan yang mendesak di wilayah Kabupaten Kotawaringin timur. Dampak negatifnya meliputi kerusakan ekosistem hutan dan lahan. Hal itu menimbulkan polusi udara yang sangat mengganggu aktivitas Masyarakat. Tujuan penelitian ini adalah untuk mengembangkan model klasifikasi yang dapat memprediksi tingkat kerawanan kebakaran berdasarkan variabel-variabel tertentu. Pengembangan model klasifikasi agar dapat memprediksi tingkat kerawanan kebakaran dilakukan dengan menggunakan algoritma CART. Algoritma CART merupakan algoritma yang mampu mengklasifikasikan data tingkat kerawanan kebakaran dengan menghasilkan sebuah pohon keputusan. Hasil dari penelitian ini menyimpulkan nilai dari klasifikasi data menggunakan perhitungan Algoritma CART pada aplikasi Rapidminer menunjukkan tingkat akurasi 79.89% yang berarti memadai dalam memprediksi tingkat kerawanan kebakaran.

**Kata kunci :** *Kerawanan Kebakaran, Algoritma CART, Pohon Keputusan*

### 1. PENDAHULUAN

Kabupaten Kotawaringin Timur memiliki kondisi lingkungan yang cocok untuk pertumbuhan vegetasi yang subur dan beragam. Ini disebabkan oleh tanah yang subur, curah hujan yang memadai, dan iklim yang mendukung perkembangan flora. Meskipun lingkungan mendukung pertumbuhan vegetasi, terdapat faktor risiko yang signifikan terhadap kebakaran hutan dan lahan di wilayah tersebut. Kemungkinan adanya kondisi cuaca ekstrem, seperti suhu tinggi dan kurangnya hujan, yang dapat menciptakan kekeringan dan meningkatkan risiko kebakaran. Kabupaten ini memiliki lahan gambut

yang mudah terbakar. Gambut merupakan material organik yang rentan terhadap kebakaran, dan apabila terbakar, dapat menyebabkan kebakaran yang sulit dikendalikan. Aktivitas manusia, seperti perambahan hutan dan perladangan, juga diidentifikasi sebagai kontributor signifikan terhadap meningkatnya kerawanan terhadap kebakaran. Perambahan hutan dapat menciptakan area yang lebih terbuka dan rentan terhadap kebakaran, sedangkan perladangan dengan metode pembakaran juga dapat meningkatkan risiko kebakaran[1].

Sebagai Masyarakat kotawaringin timur yang tinggal di daerah subur dan indah, semestinya dapat merasakan ketenangan namun terkadang juga dihantui ancaman kebakaran hutan yang tak terduga. Masalah ini menjadi perhatian serius, dan pemerintah setempat perlu untuk mengatasi risiko kebakaran dengan pendekatan yang inovatif. Pada artikel ini tingkat kerawanan kebakaran adalah parameter atau skala yang digunakan untuk mengukur sejauh mana suatu daerah atau lokasi dapat menjadi tempat yang rentan terhadap kebakaran. Ini dapat mencakup sejumlah faktor seperti kepadatan tanaman atau vegetasi, jenis tanah, jarak ke sumber air, Kelembaban Udara, Curah Hujan dan lain sebagainya. Data yang diolah menjadi pengetahuan bermanfaat bagi banyak orang, mengubah manusia melakukan prediksi dan estimasi yang terjadi[2].

Proses pengembangan sistem dimulai dengan pengumpulan data historis tentang kebakaran, termasuk lokasi, cuaca, jenis vegetasi, dan aktivitas manusia. Setelah memiliki dataset yang mencakup berbagai variabel, peneliti menerapkan algoritma CART untuk mengklasifikasikan tingkat kerawanan kebakaran pada setiap wilayah di Kabupaten Kotawaringin Timur. Kebakaran hutan dan lahan menjadi satu permasalahan bagi lingkungan dan kehidupan yang dapat menimbulkan kerugian yang sangat besar. Jumlah yang secara drastis semakin meningkat, sehingga menyebabkan munculnya kerusakan yang cepat bahkan Kotawaringin Timur salah satu berkontribusi terhadap awal munculnya bahaya[3].

Kebakaran hutan dan lahan yang terjadi di Kabupaten Kotawaringin Timur adalah salah satu dari banyak faktor yang berkontribusi terhadap kekeringan dan banjir di Dataran Tinggi Tengah adalah perubahan iklim[4]. Untuk mengatasi situasi ini upaya untuk mengurangi bencana terkait hutan dan lahan harus dilakukan, Salah satu langkah pertama yang harus dilakukan adalah menentukan lokasi erosi dan menganalisis penyebab terjadinya erosi pada hutan dan lahan. Selain itu, tantangan lain adalah kompleksitas faktor-faktor yang mempengaruhi kebakaran. Cuaca, misalnya, dapat berubah secara drastis, dan aktivitas manusia, seperti pembukaan lahan pertanian dengan cara membakar, dapat sulit diprediksi. Oleh karena itu, peneliti harus terus memperbarui dan mengkalibrasi model mereka agar tetap relevan dengan kondisi terkini Salah satu penyebab kebakaranlahanadalah kelalaian ataupun kesengajaan dari manusia terkait dengan deforestasi hutan[5].

Meningkatnya perubahan dalam proses sosial ekonomi dan iklim yang mengarah pada modifikasi ekstensif terhadap lingkungan alam dan periode kekeringan yang berkepanjangan telah

menempatkan tuntutan yang kuat pada pihak berwenang dan pembuat keputusan untuk membatasi kawasan hutan secara temporal dan spasial dalam hal kerentanan terhadap kebakaran. Di kawasan hutan, yang luasnya mencapai ribuan hektar setiap tahun dan sangat melindungi terhadap penurunan kualitas lahan, dan infrastruktur hutan serta pembangunan manusia di sekitarnya[6]. Meskipun ada kesulitan,pada penelitian ini menunjukkan potensi besar dalam meningkatkan manajemen kebakaran di Kabupaten Kotawaringin Timur. Dengan klasifikasi yang akurat tentang tingkat kerawanan kebakaran, pihak berwenang dapat lebih efektif mengalokasikan sumber daya untuk pencegahan dan penanggulangan kebakaran. Seiring berjalannya waktu, harapannya adalah bahwa solusi ini dapat diadopsi secara luas dan memberikan dampak positif pada kelestarian lingkungan dan keselamatan masyarakat lokal[7].

Klasifikasi data tingkat kerawanan kebakaran menggunakan algoritma CART (Classification and Regression Trees) merupakan suatu metode analisis dalam data yang digunakan untuk mengelompokkan atau mengkategorikan data berdasarkan atribut-atribut tertentu. Pada artikel ini membahas mengenai atribut-atribut yang berkaitan dengan faktor-faktor yang mempengaruhi tingkat kerawanan terhadap kebakaran. Untuk melakukan klasifikasi data tingkat kerawanan kebakaran dengan algoritma CART perlu dilakukan mengumpulkan dan menyiapkan data. Data yang dikumpulkan merupakan data yang relevan yang mampu menunjukkan karakteristik dari tingkat kerawanan kebakaran di kabupaten Kotawaringin Timur[8].

Data mining memiliki peran penting dalam klasifikasi tingkat kebakaran di suatu daerah seperti Kabupaten Kotawaringin Timur. Kabupaten ini memiliki keanekaragaman geografis dan lingkungan yang dapat mempengaruhi risiko kebakaran. Pengujian model digunakan untuk mengukur akurasi sistem dalam melakukan klasifikasi dengan hasil yang dikeluarkan berdasarkan dengan tujuan yang diinginkan. Penelitian ini melakukan pengumpulan data untuk memperoleh informasi yang akan digunakan selama proses penelitian Untuk itu, perlu adanya proses untuk memperoleh hasil penelitian yaitu dengan menggunakan pendekatan pada Algoritma CART[9].

Pada Data Mining juga melibatkan beberapa Langkah dalam penggalian data, analisis data historis untuk memahami pola, dan hubungan dalam kumpulan data yang besar. Hal ini Bertujuan untuk mengidentifikasi dan mengekstrak informasi yang bermanfaat dari data[10].Proses pada data mining juga mencakup pola tersembunyi, tren, hubungan, atau pengetahuan yang dapat membantu dalam

pengambilan keputusan. Mengenai factor-faktor yang mempengaruhi banyaknya titik api dan jenis kelimpahan yang mudah ditebus, bahkan kabupaten Kotawaringin Timur mempunyai potensi hutan kebakaran dan lahan gambut yang tinggi dan mudah terbakar pada musim kemarau. Untuk itu, analisis pada metode data mining akan membantu dalam penyelesaian permasalahan ini[11].

Analisis dengan menggunakan metode data mining terkait Kabupaten Kotawaringin Timur merupakan wilayah yang mempunyai kekhawatiran utama terhadap risiko perubahan iklim pada kebakaran hutan dan lahan. Untuk itu, maka pada artikel ini atribut atau variabel yang dapat digunakan terkait perubahan iklim adalah atribut cuaca. Kemudian perubahan cuaca ekstrem, keberadaan lahan gambut yang mudah terbakar, dan aktivitas manusia merupakan elemen-elemen kritis yang perlu dipertimbangkan dalam pengelolaan risiko kebakaran di wilayah Kabupaten Kotawaringin Timur. Cuaca ekstrim biasanya terjadi pada bulan-bulan tertentu sehingga atribut luas wilayah, bulan, dan penyebab kebakaran juga diperlukan dalam penelitian ini[12].

Tujuan utamanya pada penelitian terkait kebakaran hutan dan lahan adalah untuk merinci dan mengelompokkan faktor-faktor yang berkontribusi pada kebakaran, memungkinkan pihak berwenang untuk mengambil tindakan preventif secara lebih efektif. Agar hasil penelitian memberikan kontribusi terhadap sebuah tujuan penelitian sangat perlu mendapatkan data yang relevan dengan topik penelitian. Bahkan dalam penelitian Data mining tahapan Tokenizing Pemrosesan juga dapat dilakukan untuk menghasilkan kata-kata dari kalimat[13]. Data terhadap bencana dengan potensi kejadian tinggi mungkin diperlukan sebagai informasi untuk mengakomodasi masyarakat dan siaga dalam mengenal lingkungan.[14].

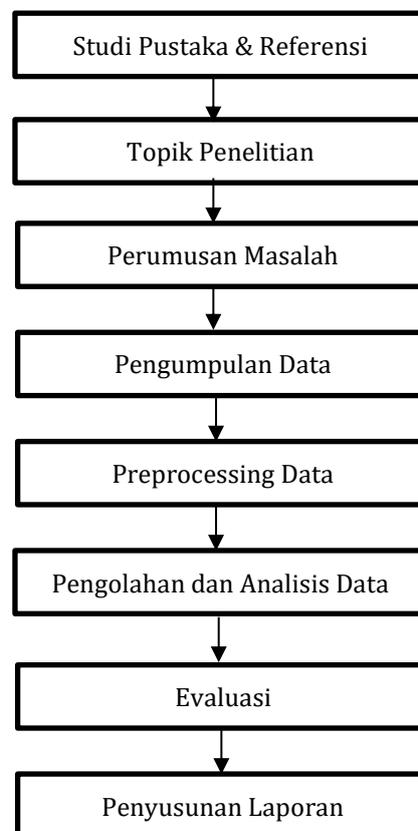
Dalam upaya mengatasi masalah tersebut, penelitian ini bertujuan untuk mengklasifikasikan tingkat kerawanan kebakaran di Kabupaten Kotawaringin Timur menggunakan Algoritma Classification and Regression Trees (CART). Algoritma CART (Classification and Regression Trees) adalah algoritma yang digunakan dalam data mining, statistik, dan pembelajaran mesin untuk membuat pohon keputusan dengan mengklasifikasikan data dengan tingkat akurasi yang tinggi Dengan menerapkan algoritma ini pada data kerawanan kebakaran, kita dapat memahami faktor-faktor utama yang berkontribusi terhadap kebakaran dan dapat meningkatkan upaya pencegahan dan mitigasi yang lebih efektif[15].

Tujuan penelitian ini melakukan klasifikasi Data Tingkat Kerawanan Kebakaran Menggunakan

Algoritma CART untuk mengembangkan model klasifikasi seperti pohon keputusan yang dapat memprediksi tingkat kerawanan kebakaran berdasarkan variabel-variabel tertentu. Dengan mencapai tujuan-tujuan ini, penelitian ini diharapkan dapat memberikan kontribusi pada pemahaman kita tentang faktor-faktor yang mempengaruhi tingkat kerawanan kebakaran dan membantu dalam mengembangkan strategi pencegahan yang lebih efektif[16]. Tujuan utamanya adalah untuk merinci dan mengklasifikasikan faktor-faktor yang berkontribusi terhadap kebakaran, sehingga memungkinkan pihak berwenang mengambil tindakan pencegahan yang lebih efektif. penelitian ini bertujuan untuk mengklasifikasikan tingkat kerawanan kebakaran di Kabupaten Kotawaringin Timur menggunakan Algoritma Classification and Regression Trees (CART).

## 2. METODOLOGI PENELITIAN

Metode penelitian adalah suatu pendekatan sistematis atau langkah-langkah yang digunakan peneliti untuk merancang, melakukan, dan mengevaluasi suatu penelitian atau penelitian. Metode ini membantu peneliti merancang kerangka penelitian, mengumpulkan data, menganalisis informasi, dan mensintesis hasil atau kesimpulan.



Gambar 1. Tahapan Penelitian

## 2.1 Data Sekunder

Data sekunder adalah data yang diperoleh dari sumber yang telah ada sebelumnya, seperti laporan, dokumen, basis data, atau sumber lainnya yang sudah ada sebelum penelitian dimulai. Data sekunder biasanya dikumpulkan oleh entitas atau organisasi lain untuk tujuan mereka sendiri. Dalam penelitian, data sekunder dapat digunakan untuk tujuan analisis, evaluasi, atau penelitian lebih lanjut. Dengan menggunakan data sekunder dapat menghemat waktu dan sumber daya, tetapi juga memerlukan keterampilan dalam mengakses, mengevaluasi, dan mengolah data tersebut untuk mencapai tujuan penelitian yang diinginkan [17]. Pada penelitian ini data sekunder diperoleh dari dinas Badan Penanggulangan Bencana Daerah (BPBD) Kabupaten Kotawaringin Timur. Data diperoleh dengan melakukan komunikasi dengan berbagai sub bidang pada dinas BPBD.

## 2.2 Pengumpulan Data

Dalam penelitian ini pengumpulan data dilakukan dengan komunikasi aktif dari pihak-pihak terkait informasi laporan kejadian dari kantor pemerintah tersebut. Melalui komunikasi tersebut, peneliti dapat mengkaji data dalam bentuk laporan rekap yang telah diurutkan ke dalam beberapa format table. Proses pengalihan data dari sumbernya memastikan keakuratan dan konsistensi data yang digunakan dalam penelitian. Pengumpulan data dengan pihak-pihak terkait memperoleh beberapa atribut yang dapat mendukung dalam penelitian ini. Beberapa atribut yang diperoleh dari pengumpulan data mencakup data cuaca (misalnya, suhu, kelembaban, dan angin), karakteristik topografi (seperti kemiringan tanah), jenis vegetasi, jarak ke sumber air, dan variabel-variabel lain yang diperlukan untuk membangun model klasifikasi. Penting untuk memastikan bahwa data yang dikumpulkan mencakup representasi yang baik dari variasi kondisi lingkungan yang dapat memengaruhi kerawanan kebakaran. Proses pengumpulan data ini merupakan langkah awal yang krusial untuk memberikan dasar yang kuat bagi model klasifikasi menggunakan algoritma CART. Data yang baik dan relevan akan memperkuat kemampuan model untuk memahami dan memprediksi tingkat kerawanan kebakaran dengan akurat.

## 2.3 Dataset

Data yang digunakan pada penelitian ini adalah data kerawanan kebakaran di kotawaringin timur dengan jumlah 184 record data. sebuah data set adalah kumpulan data yang terorganisir dan disusun

sehingga dapat digunakan untuk tujuan analisis atau pengujian. Data set terdiri dari sejumlah observasi atau entitas, dan setiap observasi memiliki beberapa atribut atau fitur yang menggambarkan karakteristiknya. Dataset pada penelitian klasifikasi data tingkat kerawanan kebakaran menggunakan algoritma CART adalah kumpulan data yang dikumpulkan dan dipersiapkan untuk melakukan analisis dan pemodelan tingkat risiko kebakaran. Dataset ini berisi informasi mengenai berbagai faktor yang mempengaruhi kebakaran, seperti kondisi cuaca, topografi, kepadatan vegetasi, atau variabel lain yang dianggap relevan. pada tahap ini, data telah melalui proses preprocessing, termasuk penanganan nilai yang hilang, pembersihan data, dan transformasi variabel agar sesuai dengan persyaratan algoritma CART. Setelah persiapan, dataset ini digunakan sebagai input untuk melatih model klasifikasi menggunakan algoritma CART. Dataset ini terdiri dari beberapa observasi atau entitas (mungkin mewakili area geografis tertentu), dan setiap observasi memiliki atribut atau fitur yang dapat digunakan oleh model untuk memprediksi tingkat kerawanan kebakaran. Setiap entitas ini kemudian diklasifikasikan ke dalam kategori tingkat risiko kebakaran tertentu oleh model CART. Hasil dari penelitian ini akan memberikan wawasan tentang faktor-faktor apa yang memiliki pengaruh signifikan terhadap kerawanan kebakaran, dan bagaimana algoritma CART dapat digunakan untuk memahami dan memprediksi tingkat risiko tersebut. Kesimpulan dan temuan dari dataset ini dapat digunakan untuk meningkatkan pemahaman tentang distribusi risiko kebakaran dan mengarah pada tindakan mitigasi yang lebih efektif.

## 2.4 Preprocessing

*Preprocessing* adalah tahap penting dalam analisis data yang dilakukan untuk mengubah data mentah dari proses pengambilan data menjadi format yang lebih efisien untuk digunakan dalam proses klasifikasi. Setelah melalui tahap preprocessing, data tersebut disimpan sebagai dataset baru yang terdiri dari 184 *record* yang akan digunakan dalam proses pemodelan. Selain itu atribut yang akan digunakan dalam proses pengolahan data setelah preprocessing diperoleh adalah atribut Luas Km<sup>2</sup>, Penyebab Kebakaran, Cuaca, dan Bulan sebagai atribut input atau disebut *variable independent*. Sedangkan atribut Prosentase Tingkat Kerawanan dinyatakan sebagai atribut label atau *variable dependen*. Kemudian kolom desa/kelurahan pada dataset dinyatakan dengan status id atau kolom yang tidak dihitung dalam proses klasifikasi. Langkah-langkah dalam preprocessing melibatkan penanganan data yang

hilang, pembersihan data dari noise atau outlier, transformasi data untuk memenuhi persyaratan model, pengkodean variabel kategorikal, normalisasi atau standarisasi data numerik, dan langkah-langkah lainnya yang diperlukan untuk meningkatkan kualitas data. Dataset yang dihasilkan dari proses preprocessing ini menjadi dasar untuk membangun model klasifikasi. Pemilihan dan penerapan langkah-langkah preprocessing yang tepat dapat membantu memastikan bahwa model yang dibangun dapat bekerja dengan baik, menghasilkan prediksi yang akurat, dan memberikan wawasan yang berharga dari data yang digunakan.

## 2.5 Klasifikasi data

Klasifikasi data pada rawan kebakaran melibatkan penggunaan algoritma atau model machine learning untuk memisahkan atau mengelompokkan data menjadi kategori tingkat risiko kebakaran. Proses ini biasanya melibatkan langkah-langkah seperti pengumpulan data yang relevan (seperti kondisi cuaca, topografi, atau vegetasi), pemrosesan data, pemilihan variabel penting, dan penerapan algoritma klasifikasi. Hasilnya adalah model yang dapat memprediksi sejauh mana suatu area atau wilayah rawan terhadap kebakaran berdasarkan karakteristik tertentu, membantu dalam upaya mitigasi risiko dan pengambilan keputusan.

## 2.6 Analisa Data

Analisis data pada penelitian "Klasifikasi Data Tingkat Kerawanan Kebakaran Menggunakan Algoritma CART" melibatkan penerapan metode analisis klasifikasi untuk memahami dan memprediksi tingkat kerawanan kebakaran berdasarkan data yang dikumpulkan. Dalam penelitian ini, algoritma CART (Classification and Regression Trees) digunakan sebagai pendekatan untuk mengklasifikasikan area atau wilayah ke dalam kategori tingkat risiko tertentu. Analisis data merupakan proses penyelidikan dan interpretasi data untuk mendapatkan informasi yang berguna, menarik kesimpulan, dan mendukung pengambilan keputusan. Tujuan utama dari analisis data adalah untuk mengidentifikasi pola, dan hubungan di antara variabel-variabel yang diamati. Analisis data adalah proses sistematis untuk memeriksa, membersihkan, mengorganisir, dan menginterpretasikan data dengan tujuan mendapatkan wawasan, mengidentifikasi pola, serta mendukung pengambilan keputusan. Dalam

konteks klasifikasi tingkat kerawanan kebakaran, analisis data mencakup penilaian terhadap informasi yang terkumpul mengenai faktor-faktor yang memengaruhi kerawanan kebakaran.

## 2.7 Pengolahan Data dan Performa

Pengolahan data di penelitian ini dilakukan dengan menggunakan algoritma CART (Classification and Regression Trees). Penggunaan prosedur penyelesaian CART ini memungkinkan peneliti untuk mendapatkan wawasan yang mendalam dari data dan memahami karakteristik serta faktor-faktor yang berpengaruh dalam penelitian ini. Dengan demikian, algoritma CART menjadi salah satu pendekatan yang kuat untuk mengolah data dalam penelitian ini. Perhitungan pada Algoritma CART dilakukan dua tahap yaitu mencari nilai index gini dan mencari nilai gini gain. Hal ini dilakukan berulang hingga terbentuk pohon keputusan. Adapun persamaan pada algoritma CART disajikan pada persamaan (1) dan (2).

Mencari nilai index Gini;

$$Index\ Gini = 1 - \sum_{i=1}^k p_i^2 \quad (1)$$

Mencari nilai Index gini terlebih dahulu dilakukan dengan mengetahui kasus dalam setiap fitur. Kemudian pada kasus akan memiliki beberapa pilihan kondisi. Setelah semua kasus dan kondisi diketahui maka langkah berikutnya melakukan perhitungan dengan persamaan pada index gini gini  $1 - ((konisi/kasus)^2 + ((konisi/kasus)^2)$

Mencari nilai Gini Gain:

$$Gini\ Gain = Gini(A, S) - \sum_{i=1}^n \frac{|S_i|}{|S|} \times Gini(S_i) \quad (2)$$

Nilai pada Gini Gain dapat diperoleh dengan melakukan perhitungan sesuai pada persamaan 2. Perhitungan persamaan 2 yaitu nilai index gini label dikurang dengan nilai jumlah kasus dibagi nilai seluruh record pada dataset kemudian dikali dengan nilai gini index setiap kasus. Jika diperhatikan persamaan bahwa terdapat symbol sigma ( $\sum$ ) artinya hal ini akan menjumlah bagian-bagian pada kasus yang lain juga yang terdapat pada fitur tersebut. Sehingga nilai Gini Gain pada suatu fitur telah diketahui.

Evaluasi performa pada algoritma CART dilakukan dengan memperhatikan nilai tingkat performa yang dihasilkan. Salah satu metrik evaluasi yang penting adalah akurasi. Semakin tinggi persentase akurasi yang dihasilkan, semakin baik performa dari model pohon keputusan tersebut dalam

mengklasifikasikan data. Nilai akurasi (acc) adalah proporsi jumlah prediksi yang benar. Dapat dihitung dengan menggunakan persamaan (3):

$$acc = \frac{tp + tn}{tp + tn + fp + fn}$$

Untuk memahami tingkat akurasi, penelitian ini akan menyajikan tabel Confusion Matrix. nilai-nilai true positif (tp), true negative (tn), benar positif(fp), dan negatif (fn) akan ditampilkan pada tabel Confusion Matrix. (tp), true negative (tn), positif (fp), dan negatif (fn). Setelah mengacu pada tabel Confusion Matrix, maka dapat ditemukan nilai keakuratannya pada data.

Algoritma CART (Classification and Regression Trees) adalah algoritma pembelajaran yang digunakan untuk membuat pohon keputusan. Berikut adalah tahapan-tahapan dari algoritma CART: Menentukan Variabel Pemisah (Splitting Variable) yaitu dengan Pilih variabel dari dataset yang paling baik memisahkan data ke dalam kelas atau kelompok yang berbeda. Kriteria pemisahan dapat berupa Gini impurity untuk klasifikasi atau reduksi varians untuk regresi. Menghitung Kriteria Pemisahan yaitu dengan Hitung nilai Gini impurity atau reduksi varians untuk setiap variabel pemisah dan setiap nilai ambang yang mungkin. Memilih Variabel Pemisah Terbaik yaitu Pilih variabel pemisah dan nilai ambang yang menghasilkan Gini impurity atau reduksi varians yang paling baik.

Membagi Data (Splitting Data) yaitu Pisahkan data menjadi dua subset berdasarkan variabel pemisah dan nilai ambang yang telah dipilih. Pembentukan Pohon secara Rekursif yaitu Terapkan langkah-langkah 1 dan 2 secara rekursif untuk setiap subset data yang dihasilkan pada tahap sebelumnya. Lanjutkan proses ini hingga mencapai kriteria penghentian, seperti kedalaman maksimum pohon atau jumlah sampel minimum di setiap cabang. Menetapkan Label pada Daun pohon Setiap daun pohon mewakili suatu kelas atau nilai regresi. Tetapkan label kelas atau nilai regresi yang sesuai berdasarkan mayoritas kelas atau nilai di setiap daun. Pemangkasan Pohon yaitu Lakukan pruning untuk mengurangi kompleksitas pohon dan mencegah overfitting. Pemangkasan dilakukan dengan menghapus cabang-cabang yang tidak memberikan kontribusi signifikan pada peningkatan akurasi atau mengurangi varians. Menggunakan Model yaitu Gunakan model pohon keputusan yang telah dibentuk untuk mengklasifikasikan data baru atau memprediksi nilai regresi.

Algoritma CART memiliki kelebihan dalam interpretabilitas dan dapat menangani baik data kategorikal maupun numerik. Namun, dapat menjadi rentan terhadap overfitting, yang dapat diatasi dengan teknik pruning. Selain itu, model ini juga dapat digunakan untuk regresi dengan penyesuaian langkah-langkah yang sesuai

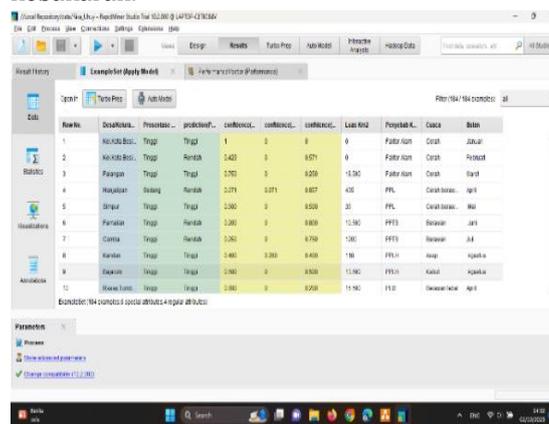
### 3. HASIL DAN PEMBAHASAN

Pada penelitian ini dilakukan berdasarkan tahapan yang telah dilakukan mulai dari studi literature hingga penyajian hasil laporan. Sumber data yang digunakan dalam penelitian ini berasal dari laporan dari instansi. Dari komunikasi yang dilakukan dalam pengumpulan data, ditemukan sebanyak 184 record yang akan digunakan sebagai sampel dalam penelitian ini. Tujuan utama adalah untuk menganalisis pola data dalam dataset tersebut dengan menggunakan Algoritma CART (Classification and Regression Trees). Data penelitian diperoleh dengan beberapa jenis type data. Tabel I menunjukkan jenis Type data yang digunakan pada penelitian.

TABEL I. TYPE DATA

Atribut	Type Data
Desa/Kelurahan	Polynomial (Id)
Luas Km2	Real
Penyebab Kebakaran	Polynomial
Cuaca	Polynomial
Bulan	Polynomial
Prosentase Tingkat Kerawanan	Polynomial (Label)

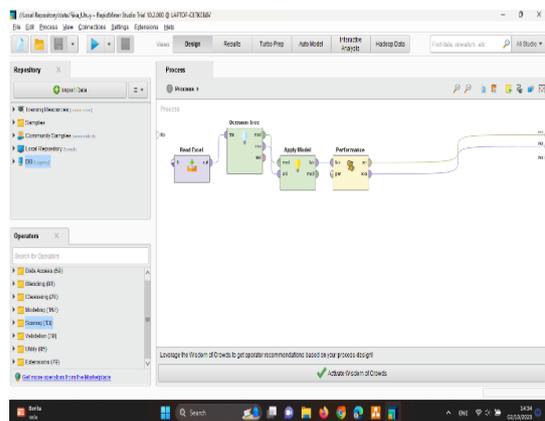
Tabel I menunjukkan pada dataset terdapat dua jenis type data yang digunakan dalam penelitian. Penelitian ini memiliki atribut desa/kelurahan yang digunakan sebagai atribut id. Atribut id tidak akan dihitung dalam penerapan rumus algoritma CART karena atribut ini dipertimbangkan sebagai atribut yang memiliki record sebagai objek yang dipertimbangkan titik kerawanan lokasi kebakaran.



Gambar 2. Type Data Pada Dataset

Dalam penelitian ini, Gambar 1 menunjukkan bahwa terdapat sebuah exampleSet yang terdiri dari 184 record. ExampleSet ini menjadi fokus utama penelitian, mencakup data yang akan

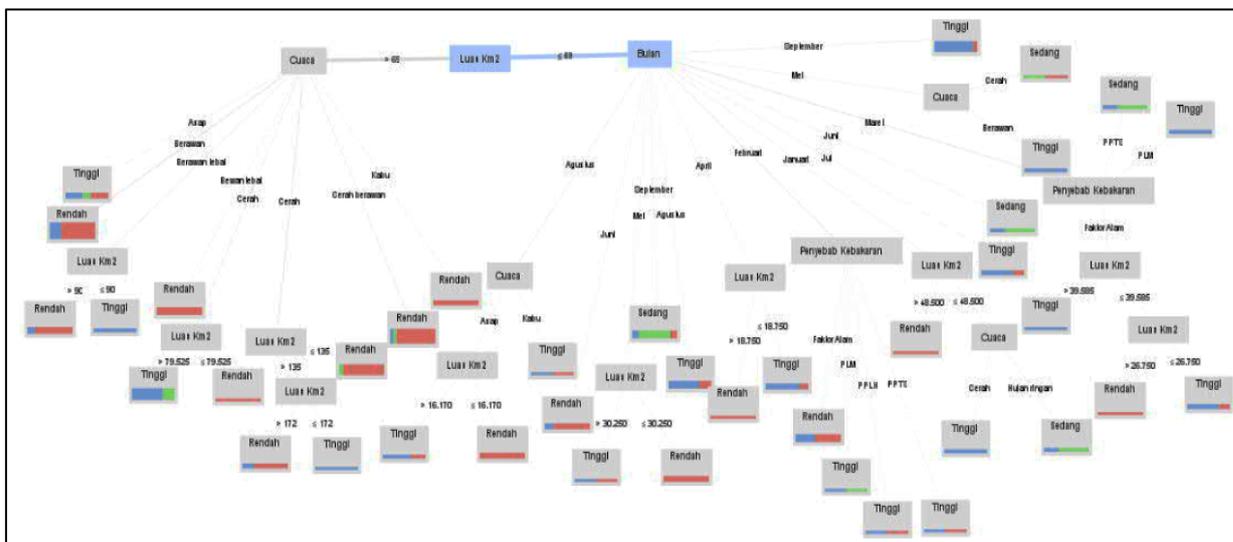
digunakan dalam analisis dan pengembangan model klasifikasi menggunakan algoritma CART. Dengan adanya struktur data yang telah terdefinisi dengan baik, diharapkan dapat mempermudah proses analisis dan meningkatkan akurasi model yang dikembangkan pada tahap selanjutnya.



Gambar 3. Process Design

Pada gambar 3 menjelaskan bahwa proses Desain adalah serangkaian langkah untuk mengembangkan sistem yang dapat memprediksi

tingkat kerawanan kebakaran berdasarkan berbagai faktor seperti cuaca, kelembapan, dan kecepatan angin. Langkah-langkahnya melibatkan pemahaman tujuan, pengumpulan dan pemrosesan data, pembagian dataset, pelatihan model CART, evaluasi model, dan implementasi. Hasil akhirnya adalah model yang dapat digunakan untuk mengklasifikasikan wilayah-wilayah berdasarkan risiko kebakaran, membantu dalam perencanaan mitigasi dan respons cepat. Pada penelitian ini struktur desain dalam pengolahan data menggunakan tools rapid minner terdiri dari beberapa operator seperti operator read model, Decision Tree, Apply Model, dan operator performance. Keempat operator ini memiliki fungsi yang berbeda-beda mulai dari penyimpanan dataset hingga pengujian performa algoritma. Algoritma CART terdapat pada operator Decision tree. Sebagaimana diketahui bahwa operator decision tree terdapat beberapa pilihan algoritma seperti, ID3, C45, Cart dan beberapa algoritma lainnya. Karena pada penelitian ini menggunakan algoritma CART maka parameter pada operator Decision tree dapat kita pilih criterion Gini\_Index. Criterion Gini\_Index akan mengarahkan perhitungan dalam pengolahan data untuk menggunakan algoritma CART.

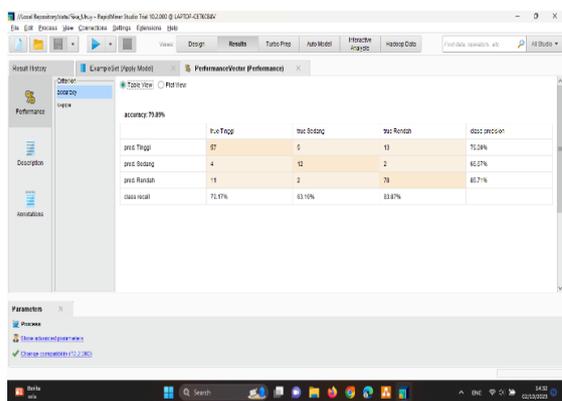


Gambar 4. Pohon Keputusan

Pada gambar 4 menjelaskan bahwa pohon keputusan adalah model prediktif yang digunakan dalam analisis data untuk menggambarkan dan memprediksi keputusan berdasarkan serangkaian aturan atau kondisi. Pohon tersebut terdiri dari simpul-simpul yang merepresentasikan keputusan atau tes pada suatu fitur, cabang-cabang yang merepresentasikan hasil dari tes tersebut, dan daun-daun yang merepresentasikan hasil akhir atau prediksi. Setiap cabang di pohon mengarah ke simpul atau daun berikutnya berdasarkan hasil tes

pada fitur tertentu.pohon keputusan digunakan untuk mengelompokkan data ke dalam kategori atau kelas yang sesuai berdasarkan fitur-fitur tertentu. Algoritma CART (Classification and Regression Trees) adalah salah satu metode yang sering digunakan untuk membangun pohon keputusan dalam konteks klasifikasi.pohon keputusan menggambarkan serangkaian pertanyaan atau kondisi (misalnya, variabel cuaca, Luas Km2, Penyebab Kebakaran, Bulan) yang digunakan untuk membagi data ke dalam kelas tingkat kerawanan

kebakaran (rendah, sedang, tinggi). Setiap cabang dan simpul pada pohon mewakili suatu keputusan berdasarkan kondisi tertentu, dan daunnya menunjukkan kelas hasil atau tingkat kerawanan kebakaran yang telah diprediksi. Pohon keputusan menyediakan cara intuitif untuk memahami faktor-faktor yang paling berpengaruh dalam klasifikasi tingkat kerawanan kebakaran.



Gambar 5. PerformanceVector (Performance)

Gambar 2 menguraikan konsep Performance Vector dalam konteks analisis data atau evaluasi model. Performance Vector ini merujuk pada sejumlah metrik atau ukuran yang digunakan untuk mengukur kinerja, performa, model, dan sistem tertentu. Dalam konteks khusus ini, metrik yang digunakan adalah accuracy, yang menunjukkan seberapa tepat model dapat memprediksi hasil dengan membandingkan jumlah prediksi yang benar dengan total jumlah prediksi. Hasil evaluasi menunjukkan bahwa model yang dikembangkan memiliki tingkat akurasi sebesar 79.89%, yang menunjukkan seberapa baik model tersebut dapat melakukan prediksi dengan tepat. Angka ini memberikan gambaran kualitatif terkait sejauh mana model berhasil dalam melakukan klasifikasi data dan dapat menjadi dasar untuk peningkatan atau pengembangan model lebih lanjut.

#### 4. Kesimpulan dan Saran

Berdasarkan hasil penelitian "Klasifikasi Data Tingkat Kerawanan Kebakaran menggunakan Algoritma CART pada kabupaten Kotawaringin Timur" bahwa algoritma CART dapat membentuk pohon keputusan yang mampu digunakan untuk memprediksi tingkat kerawanan kebakaran di kabupaten Kotawaringin Timur. Hal ini memberikan wawasan penting tentang faktor-faktor yang berkontribusi terhadap potensi risiko kebakaran. Evaluasi model pohon keputusan menggunakan data yang uji menunjukkan bahwa tingkat akurasi 79.89% yang memadai dalam memprediksi tingkat kerawanan kebakaran. Penggunaan algoritma CART

secara spesifik memberikan hasil yang dapat diandalkan dalam mengklasifikasikan area atau wilayah ke dalam kategori tingkat risiko tertentu. Hasil dari analisis menunjukkan pola dan hubungan yang dapat digunakan untuk memahami karakteristik area atau wilayah yang rentan terhadap kebakaran. Informasi ini dapat menjadi dasar untuk perencanaan mitigasi risiko dan pengembangan strategi pencegahan.

Saran Untuk meningkatkan kehandalan model, memperkaya dataset dengan data yang lebih lengkap dan relevan. Penambahan atribut-atribut tertentu atau data spasial dapat memberikan informasi tambahan untuk meningkatkan kualitas prediksi. Melakukan optimasi parameter pada algoritma CART, seperti penyetelan kedalaman pohon atau kriteria pemisahan, dapat membantu meningkatkan performa model. Meningkatkan pemahaman tentang faktor-faktor lingkungan yang mempengaruhi tingkat kerawanan kebakaran. Mendorong kerjasama antara pemangku kepentingan, termasuk pihak berwenang, lembaga pemadam kebakaran, dan masyarakat, untuk mengembangkan strategi mitigasi risiko yang lebih efektif dan Mengingat dinamika lingkungan dan perubahan dalam faktor-faktor yang mempengaruhi risiko kebakaran, disarankan untuk melakukan pemantauan dan pembaruan secara periodik terhadap model ini. Ini memastikan model tetap relevan dan efektif seiring berjalannya waktu.

#### Daftar Pustaka:

- [1] Maruli Tua Silaen, "Klasifikasi Karakteristik Kepribadian Siswa Berdasarkan the Big Five Personality Dengan Menggunakan Metode K-Nearest Neighbor (Knn)," *J. Inform. dan Rekayasa Elektron.*, vol. 6, no. 1, pp. 121–129, 2023, doi: 10.36595/jire.v6i1.860.
- [2] M. A. Hasanah, S. Soim, and A. S. Handayani, "Implementasi CRISP-DM Model Menggunakan Metode Decision Tree dengan Algoritma CART untuk Prediksi Curah Hujan Berpotensi Banjir," *J. Appl. Informatics Comput.*, vol. 5, no. 2, 2021, doi: 10.30871/jaic.v5i2.3200.
- [3] C. A. Sugiarto, "Analisis Komparasi Algoritma Klasifikasi Untuk Menangani Data Tidak Seimbang Pada Data Kebakaran Hutan," *Techno.Com*, vol. 14, no. 4, 2019.
- [4] I. Guntoro, D. M. Midyanti, and R. Hidayati, "Penerapan Dropout Pada Jaringan Saraf Tiruan Backpropagation Dalam Mengklasifikasi Tingkat Fine Fuel Moisture Code (Ffmc) Untuk Kebakaran Hutan Dan Lahan," *J. Komput. dan Apl.*, vol. 10, no. 01, pp.

- 114–123, 2022.
- [5] A. Primajaya, B. N. Sari, and A. Khusaeri, "Prediksi Potensi Kebakaran Hutan dengan Algoritma Klasifikasi C4.5 Studi Kasus Provinsi Kalimantan Barat," *J. Edukasi dan Penelit. Inform.*, vol. 6, no. 2, 2020, doi: 10.26418/jp.v6i2.37834.
- [6] K. Kurniati, "Penerapan Algoritma Classification and Regression Tree (Cart) Untuk Klasifikasi Jurusan Siswa Baru Man 1 Oku Timur," *Bina Darma Conf. Comput. ...*, pp. 331–339, 2022, [Online]. Available: <https://conference.binadarma.ac.id/index.php/BDCCS/article/view/3123%0Ahttps://conference.binadarma.ac.id/index.php/BDCCS/article/download/3123/1359>.
- [7] D. I. Kalimantan, "Karakteristik Kebakaran Permukaan Gambut Berbasis," vol. 18, no. 1, pp. 17–29, 2021.
- [8] R. T. Sirenden, Y. A. Silvianningsih, and ..., "Pengelolaan Sumberdaya Hutan dan Lahan Gambut yang Berkelanjutan Berbasis Kearifan Lokal dan Pengakuan Hak Kelola Masyarakat di Desa Kalumpang ...," *Diteksi J. ...*, vol. 1, pp. 51–59, 2023, [Online]. Available: <https://e-journal.upr.ac.id/index.php/diteksi/article/view/8980>.
- [9] F. Sodik, A. S. Mashuri, S. Huda, K. Kusri, and K. Khusnawi, "Implementasi Algoritma Convolutional Neural Network Dan Linear Regresi Untuk Memprediksi Kebakaran Hutan," *CogITo Smart J.*, vol. 9, no. 1, pp. 28–37, Jun. 2023, doi: 10.31154/cogito.v9i1.401.28-37.
- [10] A. L. D. tavares Duarte and Eddy Nurraharjo, "Analisis Sentimen Dan Klasifikasi Tweet Terkait Naiknya Kasus Omicron Menggunakan Naive Bayes Classifier," *J. Inform. dan Rekayasa Elektronik.*, vol. 6, no. 1, pp. 1–8, 2023, doi: 10.36595/jire.v6i1.779.
- [11] R. O. Mardiyanto, F. Fitriani, R. J. Purnomo, K. Kusri, and D. Maulina, "PEMETAAN LOKASI KEBAKARAN HUTAN DAN LAHAN DI NTB DENGAN MENGGUNAKAN ALGORITMA NAIVE BAYES," *Tek. Teknol. Inf. dan Multimed.*, vol. 2, no. 2, 2021, doi: 10.46764/teknimedia.v2i2.44.
- [12] C. A. Sugianto, "Analisis Komparasi Algoritma Klasifikasi Untuk Menangani Data Tidak Seimbang Pada Data Kebakaran Hutan," *Techno.Com*, vol. 14, no. 4, pp. 336–342, 2019, [Online]. Available: <http://publikasi.dinus.ac.id/index.php/technoc/article/view/992>.
- [13] F. N. Dhewayani, D. Amelia, D. N. Alifah, B. N. Sari, and M. Jajuli, "Implementasi K-Means Clustering untuk Pengelompokan Daerah Rawan Bencana Kebakaran Menggunakan Model CRISP-DM," *J. Teknol. dan Inf.*, vol. 12, no. 1, pp. 64–77, 2022, doi: 10.34010/jati.v12i1.6674.
- [14] R. Melia, N. Mutiah, and S. Rahmayuda, "SISTEM INFORMASI GEOGRAFIS PEMETAAN KAWASAN RAWAN KEBAKARAN MENGGUNAKAN METODE SEQUENTIAL PATTERN MINING," *INTI Nusa Mandiri*, vol. 17, no. 2, pp. 55–63, 2023, doi: 10.33480/inti.v17i2.4005.
- [15] T. F. Dicelebica, A. A. Akbar, and D. R. Jati, "Identifikasi dan Pencegahan Daerah Rawan Bencana Kebakaran Hutan dan Lahan Gambut Di Kalimantan Barat," *J. Ilmu Lingkungan.*, vol. 20, no. 1, pp. 115–126, 2022, doi: 10.14710/jil.20.1.115-126.
- [16] H. Haryani, C. I. Agustyaningrum, A. Surniandari, S. Sahara, and R. K. Sari, "Algoritma Klasifikasi Multilayer Perceptron Dalam Analisa Data Kebakaran Hutan," *J. Infortech*, vol. 5, no. 1, 2023, doi: 10.31294/infortech.v5i1.15792.
- [17] M. Bakri, "Penerapan Data Mining untuk Clustering Kualitas Batu Bara dalam Proses Pembakaran di PLTU Sebalang Menggunakan Metode K-Means," *J. Teknoinfo*, vol. 11, no. 1, 2017, doi: 10.33365/jti.v11i1.3.