

# Klasifikasi Kebakaran Hutan dan Lahan Menggunakan Algoritma Naïve Bayes Classifier

**Deni\*<sup>1</sup>, Barry Ceasar Octariadi<sup>2</sup>, Putri Yuli Utami<sup>3</sup>**

Program Studi Teknik Informatika, Fakultas Teknik dan Ilmu Komputer,  
Universitas Muhammadiyah Pontianak

Email: [\\*191220054@unmuhpnk.ac.id](mailto:*191220054@unmuhpnk.ac.id), [barry.ceasaro@unmuhpnk.ac.id](mailto:barry.ceasaro@unmuhpnk.ac.id), [putriyuli@unmuhpnk.ac.id](mailto:putriyuli@unmuhpnk.ac.id)

(Naskah masuk: 28 Januari 2025, diterima untuk diterbitkan: 13 Oktober 2025)

**Abstrak:** Salah satu bencana alam yang sering terjadi di Indonesia adalah kebakaran. Pada tahun 2019, terdapat 1.124 titik panas di Kalimantan Barat. Kebakaran hutan menyebabkan banyak kerugian, termasuk deforestasi hutan dan hasilnya, pencemaran kabut asap dan emisi dan turunnya kualitas udara, penelitian ini menerapkan algoritma Naive Bayes Classifier untuk menentukan tingkat status rawan kebakaran hutan dan lahan di Kabupaten Kubu Raya. Hasil akurasi klasifikasi status kebakaran hutan dan lahan dengan algoritma Naïve Bayes 92% pada data latih dan 96% pada data uji. Dapat ditarik kesimpulan bahwa Gaussian Naïve Bayes dapat mengklasifikasi status low, medium, high pada kebakaran hutan dan lahan dengan baik.

**Kata Kunci** – data mining; kebakaran hutan; klasifikasi; naïve bayes classifier

## Classification of Forest and Land Fires Using the Naïve Bayes Classifier Algorithm

**Abstract:** One of the most recent natural disasters in Indonesia is wildfires. In 2019, there were 1,124 hotspots in West Kalimantan. Forest fires cause extensive damage, including reducing forest area and yields, air pollution from haze and emissions, decreasing agricultural potential, and overall forest depletion. The aim of this research is to apply the Naive Bayes Classifier algorithm to determine the level of forest and land fire vulnerability status in Kubu Raya District and assess the accuracy of the Naïve Bayes Classifier method in classifying forest and land fire status. The method employed in this study is Gaussian Naïve Bayes. The classification accuracy of forest and land fire status using Naive Bayes algorithm is 92% on training data and 96% on test data. It can be concluded that Gaussian Naïve Bayes effectively classifies forest and land fire status.

**Keywords** – data mining; wildfires; classification; naïve bayes classifier

### 1. PENDAHULUAN

Salah satu bencana alam yang paling baru terjadi di Indonesia adalah kebakaran. Pada tahun 2019, terdapat 1.124 titik panas di Kalimantan Barat yang diperoleh dari Lembaga Penerbangan dan Antariksa Nasional (LAPAN). Kebakaran hutan merujuk pada peristiwa terbakarnya suatu wilayah yang menimbulkan risiko atau dapat menyebabkan bencana. Kejadian ini bisa disebabkan oleh pembakaran yang tidak terkendali, proses alamiah yang terjadi secara spontan, atau tindakan yang disengaja [1]. Kebakaran hutan menyebabkan banyak kerugian, termasuk penurunan jumlah hutan dan hasilnya, pencemaran kabut asap dan emisi, penurunan kualitas potensi pedesaan, dan pengurangan jumlah hutan secara keseluruhan [2].

Berdasarkan laporan Kementerian Kehutanan dan Lingkungan Hidup, kawasan hutan Provinsi Kalimantan Barat adalah yang paling luas di Indonesia, dengan luas 8.200.000 ha hutan dan 1.600.000 ha lahan gambut [3]. Tingkat curah hujan yang rendah saat musim kemarau dan banyaknya titik *hotspot* menyebabkan rentan terjadinya kebakaran. Berdasarkan laporan Kementerian Kehutanan dan Lingkungan Hidup (KLHK) 2022-2023, total luas kebakaran hutan dan lahan di provinsi Kalimantan Barat mencapai 113.684,43 ha. Berdasarkan data titik panas (*hotspot*)

di Kalimantan Barat diperoleh dari Dinas Lingkungan Hidup dan Kehutanan (DLHK) sebanyak 2.436 pada tahun 2022 sampai tahun 2023 dengan rincian tingkat kerentanan sebanyak 103 titik untuk *low*, 1849 titik untuk *medium*, dan 86 titik untuk *high*. Dan data yang saya peroleh dari Dinas Lingkungan Hidup Dan Kehutanan (DLHK).provinsi Kalimantan Barat dengan jumlah keseluruhan 1096 data berdasarkan variabel suhu, kelembapan, curah hujan, penyinaran matahari, dan kecepatan angin pada tahun 2021-2023. Oleh karena itu, pemerintah terus berupaya dalam menanggulangi kebakaran hutan dan lahan di Kalimantan Barat [4]

Prediksi dapat dibuat melalui klasifikasi memungkinkan variabel untuk memprediksi apakah akan terjadi kebakaran hutan atau tidak, ada beberapa metode yang dapat digunakan untuk mengklasifikasi seperti *Naive Bayes Classifier*, *K-Nearest Neighbour (KNN)*, *Support Vector Machine*, dan lain-lain, pada penelitian ini menggunakan algoritma *Naive bayes Classifier*. Metode klasifikasi *naive bayes* mampu membuat model dan memprediksi peluang di masa depan dengan melakukan prediksi untuk memastikan bahwa informasi bergantung pada kumpulan jenis data saat ini *naive bayes* menghitung probabilitas berdasarkan standar yang sudah ditetapkan metode ini menggunakan klasifikasi objek yang diawasi dengan menetapkan label kelas berdasarkan kategori [5].

Penelitian pertama penelitian yang dilakukan oleh Sari pada tahun (2022) dengan judul "Klasifikasi Penyakit Jantung Menggunakan Metode *Naive Bayes*" Berdasarkan hasil penelitian yang telah dilakukan terhadap dataset UCI mengenai klasifikasi penyakit jantung menggunakan Metode *Neural Network* memiliki kemampuan dalam meningkatkan akurasi pada algoritma *Naive Bayes* untuk menentukan klasifikasi resiko Penyakit Jantung dengan hasil nilai akurasi terbaik sebesar 85.33% presisi sebesar 94,84% dan *recall* sebesar 88,55% menggunakan data dari UCI dataset. Metode *Neural Network* yang mendapatkan hasil lebih rendah yaitu 78,55% [7]. Penelitian kedua yang dilakukan oleh Indrayuni tahun (2019) dengan judul "Klasifikasi *Text Mining Review* Produk Kosmetik Untuk Teks Bahasa Indonesia Menggunakan Algoritma *Naive Bayes*" Berdasarkan pengujian model menggunakan algoritma *Naive Bayes* pada eksperimen yang telah dilakukan terbukti bahwa algoritma *Naive Bayes* merupakan algoritma paling sederhana yang terbukti menghasilkan nilai akurasi tinggi hingga 90.50% dengan nilai AUC sebesar 0.715 [8]. Penelitian ketiga yang dilakukan oleh Pratiwi tahun (2021) dengan judul "Klasifikasi Kebakaran Hutan Dan Lahan Menggunakan Algoritma *Naive Bayes* Di Kabupaten Pelalawan" Penelitian ini berhasil mengklasifikasikan kemunculan titik panas kebakaran hutan di Kabupaten Pelalawan pada tahun 2015 hingga 2019 menggunakan algoritma *Naive Bayes* [4]. Penelitian keempat yang dilakukan oleh Fauzia dan Dana (2023) dengan judul "Implementasi Algoritma *Naive bayes* dalam Klasifikasi Status Kesejahteraan Masyarakat Desa Gunungsari" Berdasarkan hasil pengujian dalam penilaian akurasi status kesejahteraan masyarakat menggunakan algoritma *naive bayes*, *variable* data yang digunakan yaitu tempat tinggal, status lahan, jenis lantai, jenis dinding, penerangan rumah, energi untuk memasak, tempat pembuangan sampah, fasilitas MCK, sumber air mandi, sumber air minum, rumah berada di bantaran sungai, rumah di lereng bukit, BLT, PKH, BST, BANPRES, Bantuan UMKM, bantuan pendidikan anak. Metode Algoritma *Naive bayes* dapat diimplementasikan dan mampu mengidentifikasi class kesejahteraan masyarakat dengan akurasi hasil prediksi sebesar 93,69% [8]. Penelitian kelima yang dilakukan oleh Putri Yuli Utami, Sahid Agustian Hudjimartsu, Tiara Aurilia Viona, Hulwan Sharfina pada tahun 2023 dengan judul " Optimasi Parameter Algoritma DBSCAN Untuk Mendeteksi Titik Panas Kebakaran Hutan Dan Wana " yang meneliti tentang pendeteksi kebakaran agar membantu para instansi terkait dengan mendeteksi tentang titik sebaran panas agar bisa menemukan titik kebakaran didalam hutan dan lahan[10].

Berdasarkan penelitian sebelumnya yang dilakukan oleh Asnawi dkk tahun 2021 dengan judul "Perbandingan Algoritma *Naive Bayes*, *K-NN*, dan *SVM* dalam Pengklasifikasian Sentimen Media Sosial" Dari penelitian ini diketahui bahwa algoritma *Naive Bayes Classifier* memiliki nilai akurasi lebih tinggi sebesar 79.8% dibandingkan algoritma *K-Nearest Neighbor* sebesar 50.23% dan *Support Vector Machine* sebesar 75.29%, sehingga algoritma *Naive Bayes Classifier* lebih tepat digunakan untuk mengklasifikasi ulasan aplikasi media sosial [6]. Berdasarkan permasalahan di atas penulis bertujuan melakukan penelitian dengan judul "Klasifikasi Kebakaran Hutan dan Lahan Menggunakan Algoritma *Naive Bayes Classifier*".

## 2. METODE PENELITIAN

### 1) Identifikasi Masalah

Langkah yang dapat dilakukan adalah mengidentifikasi permasalahan yang ada di Kabupaten Kubu Raya. Salah satu permasalahan yang ada di Kabupaten Kubu Raya sering terjadinya kebakaran hutan dan lahan sehingga Kabupaten Kubu Raya dijuluki primadona kebakaran oleh Dinas Lingkungan Hidup dan Kehutanan. Oleh karena itu, perlu dilakukan klasifikasi status kebakaran hutan dan lahan.

### 2) Studi Pustaka

Pada tahap ini akan dilakukan pencarian mengenai penelitian sebelumnya. Studi pustaka dilakukan untuk mengumpulkan bahan referensi. Studi pustaka yang digunakan dapat berupa jurnal ilmiah terdahulu, buku, dan bahan-bahan lain yang dapat digunakan untuk mendukung penyelesaian penelitian.

### 3) Pengumpulan Data

Pada fase pengumpulan data, tujuannya adalah mendapatkan dataset yang relevan dengan tujuan penelitian. Pengumpulan data yang dilakukan pada penelitian ini menggunakan data dari dinas Lingkungan Hidup Dan Kehutanan Pontianak yang diambil dari ibu Lidia, S.T. selaku Kepala Bidang Umum dan Aparatur Dinas Lingkungan Hidup Dan Kehutanan Kalimantan Barat mengenai data kebakaran hutan di Kabupaten Kubu Raya.

### 4) Pengolahan Data

pada fase ini Pengolahan data tujuannya yaitu proses mengumpulkan, mengatur, menganalisis, dan menyajikan data untuk menghasilkan informasi yang berguna. Proses ini melibatkan serangkaian langkah yang bertujuan untuk mengubah data mentah menjadi informasi yang dapat diinterpretasikan dan digunakan untuk pengambilan keputusan.

### 5) Implementasi

Pada fase ini penulis akan menggunakan hasil pemodelan *naïve bayes classifier* yang telah disimpan dalam bentuk file pickle, file pickle kemudian akan diimport sebagai perhitungan dari website yang dibangun menggunakan Framework streamlit untuk pengklasifikasian status kebakaran hutan dan lahan pada tingkatan *low, medium, high*.

### 6) Kesimpulan

Kesimpulan didapat berdasarkan dari analisis hasil pengujian yang telah dilakukan dan juga dari hasil analisis selama pembangunan aplikasi. Kesimpulan berupa hasil dari penelitian sesuai yang tergambar pada diagram alir metode penelitian.

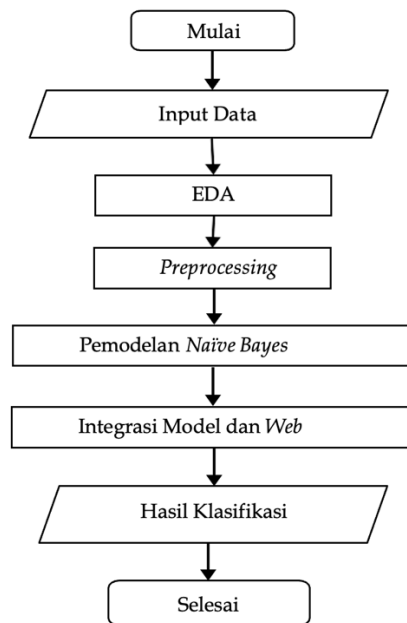
## 3. HASIL DAN PEMBAHASAN

Perancangan Proses Membangun Model digunakan untuk mendesain proses pembuatan model klasifikasi kebakaran hutan dan lahan menggunakan metode *naïve bayes* agar dapat berjalan dengan maksimal.

Data yang digunakan pada penelitian ini data primer yang diperoleh dari Badan Meteorologi Klimatologi dan Geofisika kabupaten Kubu Raya dan data sekunder yang diperoleh dari Dinas Lingkungan Hidup dan Kehutanan (DLHK) 2022-2023.

Exploratory Data Analysis (EDA) adalah salah satu tahapan penting dalam proses analisis data. Tahapan ini bertujuan untuk memahami karakteristik dalam *dataset*, seperti total keseluruhan data, nilai rata-rata, nilai kurtosis bawah, dan nilai kuartil atas.

*Preprocessing* data adalah tahapan mempersiapkan data yang akan digunakan dalam proses *data mining*. Proses ini dilakukan agar data yang akan digunakan sesuai kebutuhan. Tahapan-tahapan *preprocessing* meliputi mengubah tipe data dari *string* menjadi numerikal, menangani nilai yang hilang dalam *dataset*, mengecek apakah dalam *dataset* terdapat data yang terduplikat, mengecek *outlier*, dan memilih atribut yang akan digunakan dalam melakukan klasifikasi.



Gambar 1. Diagram Alir Penelitian

Pemodelan *naïve bayes* adalah proses penerapan metode untuk melakukan klasifikasi pada *dataset*. Hasil dari pengujian metode *Gaussian Naïve Bayes* serta tahapan model dievaluasi untuk mengetahui pengaruh setiap tahapan terhadap kemampuan model dalam melakukan klasifikasi. Metrik evaluasi yang digunakan berupa hasil dari *Confusion Matrix* dengan akurasi yang sangat baik yaitu 95%.

#### 4. KESIMPULAN

Dari hasil penelitian dapat ditarik kesimpulan bahwa dalam mengklasifikasi tingkat kebakaran hutan dan lahan melalui beberapa tahapan, yaitu mulai dari *Input Data*, menangani *Missing Value*, mengecek *Duplicate Data*, melakukan *label encoding*, menghilangkan outlier, menangani *Data Imbalance* dengan menggunakan teknik *Smote*, pemodelan dengan metode *Gaussian Naïve Bayes* setelah itu pengecekan akurasi. Metrik evaluasi yang digunakan berupa hasil dari *Confusion Matrix* dengan akurasi yang sangat baik yaitu 95%.

#### DAFTAR PUSTAKA

- [1] S. Marlina and R. Rahmaniati, "DESA SIAGA API DI KABUPATEN PULANG PISAU Instruments For Village Preparedness Level In Forest And Land Fire Control In Pulang Pisau Distric," *J. Penelit. Ekosist. Dipterokarpa*, vol. 8, No 2, p. 12, 2022.
- [2] 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, p. 188, 2020, doi: 10.26418/jp.v6i2.37834.
- [3] A. Rudiyan, A. E. Dzul kifli, and K. Munazar, "Klasifikasi Kebakaran Hutan Menggunakan Metode K-Nearest Neighbor : Studi Kasus Hutan Provinsi Kalimantan Barat," *JTIM J. Teknol. Inf. dan Multimed.*, vol. 3, no. 4, pp. 195–202, 2022, doi: 10.35746/jtim.v3i4.177.
- [4] T. A. Pratiwi, M. Irsyad, R. Kurniawan, S. Agustian, and B. S. Negara, "Klasifikasi Kebakaran Hutan Dan Lahan Menggunakan Algoritma Naïve Bayes Di Kabupaten Pelalawan," *CESS (Journal Comput. Eng. Syst. Sci.)*, vol. 6, no. 1, p. 139, 2021, doi: 10.24114/cess.v6i1.22555.
- [5] A. Piqram, "Analisis dampak kebakaran hutan dan lahan di kalimantan barat terhadap kualitas udara di kecamatan pontianak tenggara kota pontianak," pp. 1–164, 2020.
- [6] M. H. Asnawi, I. Firmansyah, R. Novian, and R. S. Pontoh, "Perbandingan Algoritma Naive Bayes , K-NN , dan SVM dalam Pengklasifikasian Sentimen Media Sosial," *Semin. Nas. Stat.*, no. January, pp. 1–12, 2021, doi: 10.1234/pns.v10i.85.

- [7] A. N. Sari and S. Alfionita, "Klasifikasi Penyakit Jantung Menggunakan Metode Naïve Bayes," *AMRI (Analisa Metod. Rekayasa Inform.,* vol. 1, no. 1, pp. 22-26, 2022, doi: 10.12487/AMRI.v1i1.xxxxx.
- [8] E. Indrayuni, "Klasifikasi Text Mining Review Produk Kosmetik Untuk Teks Bahasa Indonesia Menggunakan Algoritma Naive Bayes," *J. Khatulistiwa Inform.,* vol. 7, no. 1, pp. 29-36, 2019, doi: 10.31294/jki.v7i1.1.
- [9] N. S. Fauziah and R. D. Dana, "Implementasi Algoritma Naive bayes dalam Klasifikasi Status Kesejahteraan Masyarakat Desa Gunungsari," *Blend Sains J. Tek.,* vol. 1, no. 4, pp. 295-305, 2023, doi: 10.56211/blendsains.v1i4.234.
- [10] P. Y. Utami, S. A. Hudjimartsu, T. A. Viona, and H. Sharfina, "Optimasi Parameter Algoritma DBSCAN untuk Mendeteksi Titik Panas Kebakaran Hutan dan," vol. 9, no. 3, pp. 355-361, 2023.