



Analisis Perbandingan Algoritma Naive Bayes dan Random Forest dalam Klasifikasi Prediksi Area Rawan Kebakaran di DKI Jakarta

Agung Prasetyo^{1*}, Dana Maulana Fazri², Muhammad Fadli Prasetyo³, Hendra Supendar⁴, Rizal Fahlapi⁵

^{1,2,3} Program Studi Teknologi Informasi, Universitas Bina Sarana Informatika, Jakarta, Indonesia

^{4,5} Fakultas Teknik dan Informatika, Universitas Bina Sarana Informatika, Jakarta, Indonesia

Email: ^{1*} djagung002@gmail.com, ⁴ hendra.hds@bsi.ac.id, ⁵ rizal.rzf@bsi.ac.id

(* : corresponding author)

Abstrak—Kebakaran permukiman merupakan salah satu bencana yang paling sering terjadi di DKI Jakarta, dengan frekuensi yang cenderung meningkat seiring kepadatan penduduk dan keterbatasan ruang hunian. Penelitian ini bertujuan membandingkan kinerja dua algoritma klasifikasi, yaitu Naive Bayes dan Random Forest, dalam memprediksi tingkat kerawanan kebakaran pada level kelurahan di seluruh wilayah DKI Jakarta. Dataset yang digunakan terdiri atas 2.564 instans dengan tiga variabel prediktor, yakni jumlah penduduk, kepadatan penduduk, dan frekuensi sosialisasi pencegahan kebakaran; variabel terakhir jarang dilibatkan pada penelitian sejenis sehingga menjadi salah satu unsur kebaruan studi ini. Pemodelan dilakukan menggunakan perangkat lunak Orange Data Mining dengan skema validasi silang terstratifikasi sepuluh lipat. Hasil pengujian menunjukkan bahwa Naive Bayes memperoleh akurasi (CA) 0,696 dan AUC 0,603, sedangkan Random Forest memperoleh CA 0,671 dan AUC 0,595, sehingga kedua model tergolong setara dan moderat. Akan tetapi, analisis per-kelas mengungkap keterbatasan yang penting: akibat ketidakseimbangan data yang cukup tajam, di mana kelas berisiko tinggi hanya mencakup 29,8% data, kedua model cenderung memihak kelas mayoritas. Naive Bayes hampir gagal mengenali kelas berisiko tinggi dengan recall hanya 0,048, sedangkan Random Forest jauh lebih baik meskipun tetap rendah, yakni 0,217, serta memiliki koefisien korelasi Matthews (MCC) lebih tinggi (0,100 dibanding 0,047). Dengan demikian, Random Forest lebih sesuai untuk kebutuhan mitigasi yang menekankan deteksi area berisiko, sementara akurasi total yang tinggi terbukti menyesatkan apabila berdiri sendiri. Analisis kepentingan fitur menempatkan jumlah penduduk dan frekuensi sosialisasi sebagai dua variabel paling berpengaruh, sehingga memperkuat peran intervensi edukasi dalam mitigasi. Temuan ini menegaskan perlunya penanganan ketidakseimbangan data dan penambahan fitur dalam pengembangan sistem peringatan dini kebakaran berbasis data di wilayah perkotaan.

Kata kunci: klasifikasi; Naive Bayes; Random Forest; kerawanan kebakaran; Orange Data Mining.

Abstract—Residential fires are among the most frequent disasters in DKI Jakarta, and their frequency tends to rise along with population density and limited housing space. This study aims to compare the performance of two classification algorithms, Naive Bayes and Random Forest, in predicting the level of fire vulnerability at the urban-village (kelurahan) level across DKI Jakarta. The dataset consists of 2,564 instances with three predictor variables: total population, population density, and the frequency of fire-prevention outreach. The last variable is rarely included in similar studies and therefore constitutes one of the novel aspects of this work. Modeling was carried out using Orange Data Mining software with a stratified 10-fold cross-validation scheme. The results show that Naive Bayes achieved a classification accuracy (CA) of 0.696 and an AUC of 0.603, whereas Random Forest achieved a CA of 0.671 and an AUC of 0.595, indicating comparable and moderate discriminative power. However, a per-class analysis reveals an important limitation: because of a fairly pronounced class imbalance, in which the high-risk class accounts for only 29.8% of the data, both models tend to favor the majority class. Naive Bayes almost fails to recognize the high-risk class, with a recall of only 0.048, whereas Random Forest performs considerably better although still low at 0.217, and shows a higher Matthews correlation coefficient (MCC) of 0.100 versus 0.047. Random Forest is therefore more suitable for mitigation purposes that emphasize detecting high-risk areas, while a high overall accuracy proves misleading when considered in isolation. A feature-importance analysis ranks total population and outreach frequency as the two most influential variables, reinforcing the role of educational intervention in mitigation. These findings underscore the need to address class imbalance and to add features when developing data-driven early-warning systems for urban fires.

Keywords: classification; Naive Bayes; Random Forest; fire vulnerability; Orange Data Mining.



1. PENDAHULUAN

Kebakaran termasuk bencana yang paling sering menimpa wilayah perkotaan, dan DKI Jakarta menjadi salah satu daerah dengan catatan kejadian tertinggi di Indonesia. Sebagai pusat pemerintahan sekaligus pusat ekonomi nasional, Jakarta menampung penduduk dalam jumlah besar pada ruang yang relatif terbatas. Kondisi ini melahirkan permukiman padat dengan jarak antarbangunan yang rapat, instalasi listrik yang sering tidak memadai, serta akses pemadam yang kerap terhambat. Akibatnya, satu titik api dapat dengan cepat merambat dan menimbulkan kerugian besar, baik berupa korban jiwa, kerusakan harta benda, maupun terganggunya aktivitas masyarakat.

Data dari instansi penanggulangan kebakaran menunjukkan bahwa sebaran kejadian tidak merata antar wilayah. Sebagian kelurahan mengalami kebakaran berulang dalam satu periode, sementara sebagian lain relatif jarang. Pola yang tidak seragam ini mengindikasikan bahwa tingkat kerawanan dipengaruhi oleh sejumlah faktor yang dapat diukur, seperti kepadatan penduduk dan jumlah penduduk. Apabila faktor-faktor tersebut dapat dipetakan dan dimodelkan, maka informasi yang dihasilkan berpotensi membantu pemerintah daerah dalam menyusun prioritas mitigasi secara lebih terarah.

Salah satu pendekatan yang banyak digunakan untuk memetakan pola semacam ini adalah klasifikasi dalam data mining. Melalui klasifikasi, setiap wilayah dapat dikategorikan ke dalam tingkat kerawanan tertentu berdasarkan karakteristik yang dimilikinya. Dua algoritma yang umum dipakai pada persoalan klasifikasi adalah Naive Bayes dan Random Forest. Naive Bayes bekerja berdasarkan prinsip probabilitas dengan asumsi kebebasan antar fitur sehingga ringan secara komputasi, sedangkan Random Forest merupakan metode ansambel berbasis banyak pohon keputusan yang umumnya tangguh terhadap data yang kompleks. Perbedaan paradigma inilah yang menarik untuk dibandingkan pada konteks data kebakaran perkotaan.

Meskipun penelitian mengenai prediksi kebakaran telah cukup banyak dilakukan, sebagian besar berfokus pada kebakaran hutan dan lahan yang karakteristiknya berbeda dengan kebakaran permukiman. Penelitian pada konteks perkotaan masih relatif terbatas, terutama yang menempatkan faktor intervensi sosial, seperti frekuensi sosialisasi pencegahan kebakaran, sebagai variabel prediktor. Padahal, kegiatan edukasi dan penyuluhan merupakan salah satu instrumen mitigasi yang langsung dapat dikendalikan oleh pemerintah. Celah inilah yang berusaha dijumpai oleh penelitian ini, yakni dengan membandingkan dua algoritma klasifikasi sekaligus melibatkan variabel sosialisasi pada level kelurahan di seluruh DKI Jakarta.

Penelitian ini menempatkan beberapa unsur yang membedakannya dari studi sebelumnya. Pertama, dari sisi variabel, penelitian ini melibatkan frekuensi sosialisasi pencegahan kebakaran sebagai salah satu fitur prediktor. Sebagian besar penelitian kerawanan kebakaran hanya menggunakan variabel fisik atau demografis, sementara faktor intervensi sosial yang sebenarnya dapat dikendalikan oleh pemangku kebijakan jarang dimodelkan secara langsung.

Kedua, dari sisi cakupan, penelitian ini bekerja pada granularitas tingkat kelurahan di seluruh wilayah administrasi DKI Jakarta. Tingkat kedetailan ini memungkinkan hasil klasifikasi digunakan untuk menyusun prioritas penanganan yang lebih spesifik dibandingkan analisis pada level kota atau provinsi.

Ketiga, dari sisi metodologi, penelitian ini menyajikan studi komparatif antara dua paradigma yang berbeda, yaitu pendekatan probabilistik melalui Naive Bayes dan pendekatan ansambel melalui Random Forest. Keduanya dibangun di atas alur kerja visual menggunakan Orange Data Mining sehingga proses pemodelan bersifat transparan dan dapat ditiru ulang oleh peneliti lain. Keempat, fokus penelitian diarahkan pada kebakaran permukiman urban, bukan kebakaran hutan dan lahan, sehingga melengkapi khazanah penelitian yang selama ini masih didominasi oleh konteks wildfire.

Kontribusi penelitian ini bersifat ganda. Secara akademis, penelitian memberikan bukti empiris mengenai perbandingan kinerja kedua algoritma pada data kebakaran perkotaan beserta analisis perilaku masing-masing pada kelas mayoritas dan minoritas. Secara praktis, hasil penelitian dapat menjadi rujukan awal bagi penyusunan sistem peringatan dini dan penentuan sasaran kegiatan sosialisasi pencegahan kebakaran.

Penelitian ini bertujuan untuk membandingkan kinerja algoritma Naive Bayes dan Random Forest dalam klasifikasi kerawanan kebakaran, menentukan algoritma yang lebih sesuai untuk kebutuhan mitigasi, serta mengidentifikasi variabel yang paling berpengaruh terhadap tingkat kerawanan.



JRIIN : Jurnal Riset Informatika dan Inovasi
Volume 4, No. 2 Tahun 2026
ISSN 3025-0919 (media online)
Hal 399-409

Manfaat penelitian ini diharapkan dapat dirasakan pada dua sisi. Bagi pengembangan ilmu pengetahuan, penelitian memperkaya kajian penerapan data mining pada bidang kebencanaan perkotaan, terutama yang melibatkan variabel intervensi sosial. Bagi pihak praktis, hasil penelitian dapat membantu pemerintah daerah dan dinas terkait dalam menentukan wilayah prioritas penanganan serta merancang program sosialisasi pencegahan kebakaran yang lebih tepat sasaran.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Penelitian ini dilaksanakan melalui beberapa tahapan yang berurutan agar proses pemodelan dapat ditelusuri dan diulang. Tahapan tersebut mencakup pengumpulan data, prapemrosesan, pelabelan kelas, pembangunan model, pengujian, dan analisis hasil. Setiap tahap dirancang untuk memastikan bahwa data yang digunakan konsisten dan model yang dibangun dapat dievaluasi secara adil.

1. Pengumpulan data, yaitu menghimpun data kejadian kebakaran beserta data kependudukan dan data kegiatan sosialisasi pada level kelurahan di DKI Jakarta.
2. Prapemrosesan data, meliputi pembersihan data, penanganan nilai yang hilang, serta penyesuaian satuan dan format antar variabel.
3. Pelabelan kelas, yaitu penetapan kategori kerawanan pada setiap kelurahan berdasarkan ambang frekuensi kejadian kebakaran.
4. Pembangunan model, yakni melatih algoritma Naive Bayes dan Random Forest menggunakan data yang telah disiapkan.
5. Pengujian model dengan skema 10-fold cross validation untuk memperoleh estimasi kinerja yang lebih stabil.
6. Analisis hasil, yaitu membandingkan kinerja kedua algoritma dan menelaah kontribusi setiap variabel.

2.2 Justifikasi Pemilihan Orange Data Mining

Pemodelan pada penelitian ini dilakukan menggunakan Orange Data Mining, sebuah perangkat lunak sumber terbuka yang dikembangkan untuk analisis data dan pembelajaran mesin (Demšar dkk., 2013). Orange dipilih atas beberapa pertimbangan. Pertama, perangkat ini mengusung paradigma pemrograman visual berbasis widget sehingga seluruh alur kerja, mulai dari pemuatan data hingga evaluasi, dapat disusun tanpa menulis kode. Hal ini menurunkan hambatan teknis sekaligus memperkecil peluang kesalahan implementasi.

Kedua, alur kerja visual yang dihasilkan bersifat transparan dan mudah ditiru ulang. Susunan widget yang sama akan menghasilkan proses yang sama, sehingga reproduktibilitas penelitian dapat terjaga. Ketiga, Orange menyediakan widget bawaan untuk validasi silang dan evaluasi model, termasuk perhitungan AUC, akurasi, presisi, recall, F1-score, serta confusion matrix, sehingga proses pengujian dapat dilakukan secara konsisten. Keempat, sifatnya yang terbuka dan tidak berbayar menjadikannya sesuai untuk penelitian komparatif maupun keperluan edukasi. Atas dasar pertimbangan tersebut, Orange dinilai memadai untuk membandingkan dua algoritma secara setara dalam satu lingkungan yang seragam.

2.3 Struktur Dataset dan Variabel Penelitian

Dataset yang digunakan dalam penelitian ini terdiri atas 2.564 instans. Setiap instans merepresentasikan satu observasi pada level kelurahan untuk periode bulanan tertentu, sehingga unit analisis berupa kombinasi kelurahan dan bulan. Dataset memuat kolom penanda identitas, yaitu tahun, bulan, wilayah kota administrasi, kecamatan, dan nama kelurahan, yang berfungsi sebagai pengenalan dan tidak diikutsertakan sebagai fitur pemodelan.

Variabel yang benar-benar digunakan dalam pemodelan terdiri atas tiga fitur prediktor dan satu variabel target. Ketiga fitur prediktor tersebut adalah jumlah penduduk, kepadatan penduduk, dan jumlah sosialisasi, sedangkan variabel target adalah label kerawanan yang terdiri atas dua kategori, yaitu TINGGI dan RENDAH. Terdapat pula variabel frekuensi kejadian kebakaran yang menjadi dasar penentuan label; variabel ini sengaja dikeluarkan dari daftar fitur untuk menghindari kebocoran informasi target (target leakage), karena melibatkannya akan membuat model belajar



langsung dari besaran yang dipakai untuk membentuk label. Rincian variabel disajikan pada Tabel 2.1.

Tabel 2.1. Variabel penelitian dan deskripsinya

Variabel	Jenis	Deskripsi
Tahun, Bulan	Penanda	Periode observasi (tahun dan bulan).
Wilayah, Kecamatan, Kelurahan	Penanda	Identitas administratif lokasi observasi.
Jumlah_Penduduk	Prediktor	Jumlah penduduk pada suatu kelurahan.
Kepadatan_Penduduk	Prediktor	Rasio jumlah penduduk terhadap luas wilayah kelurahan.
Jumlah_Sosialisasi	Prediktor	Frekuensi kegiatan sosialisasi atau penyuluhan pencegahan kebakaran.
Frekuensi	Dikecualikan	Frekuensi kejadian kebakaran per bulan; dasar pelabelan, tidak dipakai sebagai fitur.
Label_Rawan	Target	Kategori kerawanan kebakaran, bernilai TINGGI atau RENDAH.

Variabel target Label_Rawan diperoleh melalui proses pelabelan berdasarkan frekuensi kejadian kebakaran. Suatu kelurahan dikategorikan TINGGI apabila mengalami paling sedikit dua kejadian kebakaran dalam satu bulan, dan dikategorikan RENDAH apabila mengalami satu kejadian dalam periode yang sama. Ambang biner ini dipilih agar pelabelan tetap sederhana dan dapat ditafsirkan secara langsung oleh pengguna hasil penelitian.

2.4 Distribusi Kelas dan Penanganan Ketidakseimbangan Data

Sebaran kelas pada data set perlu diperhatikan karena dapat memengaruhi perilaku model. Dari total 2.564 instans, sebanyak 764 instans (29,8%) tergolong kelas TINGGI dan 1.800 instans (70,2%) tergolong kelas RENDAH. Dengan demikian, rasio antarkelas berada pada kisaran 2,36 banding 1. Komposisi ini menunjukkan ketidakseimbangan yang cukup nyata, dengan kelas RENDAH sebagai mayoritas. Rincian distribusi disajikan pada Tabel 2.2.

Tabel 2.2. Distribusi kelas pada variabel target Label_Rawan

Kelas	Jumlah	Proporsi	Keterangan
TINGGI	764	29,8%	Minoritas
RENDAH	1.800	70,2%	Mayoritas
Total	2.564	100%	—

Pada penelitian ini, distribusi kelas sengaja dipertahankan dalam kondisi alaminya tanpa penerapan teknik penyeimbangan data. Pertimbangannya, proporsi tersebut mencerminkan prevalensi risiko yang sebenarnya di lapangan, sehingga probabilitas prior yang dipelajari model tetap merepresentasikan kondisi nyata. Penyeimbangan secara artifisial berpotensi mengubah prior tersebut dan menghasilkan estimasi risiko yang tidak lagi sesuai dengan keadaan sesungguhnya.

Meskipun demikian, perlu disadari bahwa ketidakseimbangan ini turut memengaruhi metrik evaluasi, terutama recall pada kelas minoritas dan nilai AUC secara keseluruhan. Model cenderung lebih mudah mengenali kelas mayoritas dibandingkan kelas minoritas. Oleh karena itu, untuk penelitian lanjutan, penerapan teknik seperti SMOTE atau pembobotan kelas (class weighting) dapat dipertimbangkan guna meningkatkan sensitivitas terhadap kelas TINGGI. Pendekatan oversampling



semacam ini telah menunjukkan manfaat pada penelitian klasifikasi dengan data tidak seimbang di DKI Jakarta sebelumnya (Hamami & Dahlan, 2022).

2.5 Skenario Pemodelan dan Evaluasi

Agar perbandingan berlangsung adil, kedua algoritma dilatih dan diuji di atas dataset yang persis sama. Mekanisme pengujiannya memakai validasi silang terstratifikasi sepuluh lipat (stratified 10-fold cross validation): seluruh data dipecah menjadi sepuluh potongan yang komposisi kelasnya dijaga seragam, lalu secara bergiliran sembilan potongan dipakai untuk melatih dan satu potongan sisanya untuk menguji, dan giliran tersebut diputar hingga genap sepuluh kali sehingga setiap potongan pernah berperan sebagai data uji. Cara ini ditempuh karena taksiran kinerja yang dihasilkannya lebih mantap dan tidak gampang terpengaruh oleh pembagian data yang kebetulan berpihak pada salah satu model. Dalam analisis, kelas TINGGI diperlakukan sebagai kelas positif.

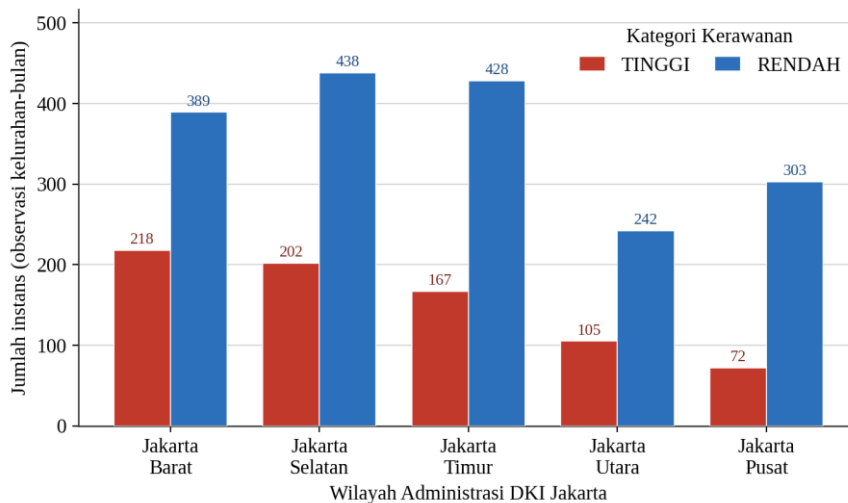
Konfigurasi model dipertahankan sederhana agar mudah ditiru ulang. Naive Bayes dijalankan dengan parameter bawaan. Random Forest dikonfigurasi dengan sepuluh pohon keputusan ($n_estimators = 10$) dan jumlah sampel minimum untuk pemisahan simpul sebesar lima ($min_samples_split = 5$), tanpa pembatasan kedalaman pohon dan tanpa pembobotan kelas (class weighting). Tidak diterapkannya pembobotan kelas konsisten dengan keputusan mempertahankan distribusi alami sebagaimana dijelaskan pada subbab 3.4.

Kinerja model dievaluasi menggunakan lima metrik utama, yaitu AUC, akurasi klasifikasi (CA), presisi, recall, dan F1-score. Selain itu, confusion matrix digunakan untuk menelaah pola kesalahan klasifikasi pada masing-masing kelas. Penggunaan beberapa metrik sekaligus penting dilakukan karena pada data yang tidak seimbang, akurasi saja dapat memberikan gambaran yang menyesatkan. Oleh sebab itu, analisis tidak hanya bertumpu pada nilai akurasi, tetapi juga pada kemampuan model dalam mengenali kelas berisiko tinggi yang menjadi perhatian utama dalam konteks mitigasi.

3. HASIL DAN PEMBAHASAN

3.1 Karakteristik Distribusi Kerawanan per Wilayah

Sebelum membahas kinerja model, penting untuk memahami sebaran tingkat kerawanan pada setiap wilayah administrasi. Gambar 1 menampilkan jumlah instans menurut kategori kerawanan pada lima wilayah kota administrasi DKI Jakarta; tampilan ini bersesuaian dengan keluaran widget Distributions pada Orange yang memetakan variabel wilayah berdasarkan label kerawanan. Secara umum, kelas RENDAH mendominasi di seluruh wilayah, yang sejalan dengan komposisi kelas pada dataset secara keseluruhan.



Gambar 1. Distribusi kategori kerawanan kebakaran per wilayah administrasi DKI Jakarta



Berdasarkan Gambar 1, kelas RENDAH mendominasi di seluruh wilayah, sejalan dengan komposisi data secara keseluruhan. Untuk kategori TINGGI, Jakarta Barat mencatat jumlah instans tertinggi (218), diikuti Jakarta Selatan (202) dan Jakarta Timur (167), sedangkan Jakarta Pusat paling sedikit (72). Apabila ditinjau dari proporsi instans berkategori TINGGI terhadap total wilayahnya, Jakarta Barat tetap yang tertinggi, yakni sekitar 36 persen, sementara Jakarta Pusat paling rendah, sekitar 19 persen. Pola ini memberikan gambaran awal bahwa kerawanan kebakaran tidak tersebar merata dan cenderung lebih tinggi pada wilayah dengan karakteristik permukiman yang lebih padat.

3.2 Hasil Evaluasi Performa Model

Hasil pengujian kedua algoritma dengan skema validasi silang terstratifikasi sepuluh lipat disajikan pada Tabel 3.1, dengan kelas TINGGI ditetapkan sebagai kelas positif. Selain AUC dan akurasi klasifikasi (CA), tabel ini menampilkan presisi, recall, dan F1-score untuk kelas TINGGI, serta koefisien korelasi Matthews (MCC) yang lebih tahan terhadap pengaruh ketidakseimbangan kelas.

Tabel 3.1. Perbandingan kinerja Naive Bayes dan Random Forest

Algoritma	AUC	CA	Precision	Recall	F1	MCC
Naive Bayes	0,603	0,696	0,411	0,048	0,087	0,047
Random Forest	0,595	0,671	0,403	0,217	0,282	0,100

Catatan: Precision, Recall, dan F1 dihitung untuk kelas TINGGI (positif/berisiko); AUC, CA, dan MCC merupakan ukuran pada tingkat model. Hasil Random Forest dapat sedikit bervariasi antar-eksekusi karena sifat acaknya.

Pada tingkat ringkasan, kedua algoritma tampak setara: AUC keduanya berada pada kisaran 0,60 (Naive Bayes 0,603 dan Random Forest 0,595) dan akurasinya berkisar antara 0,67 hingga 0,70. Namun, akurasi yang tampak tinggi ini menyimpan persoalan. Ketika perhatian dialihkan ke kelas TINGGI yang menjadi fokus mitigasi, terlihat perbedaan yang tajam: Naive Bayes hanya memperoleh recall 0,048 dengan F1 0,087, sedangkan Random Forest jauh lebih baik dengan recall 0,217 dan F1 0,282. Nilai MCC pun menegaskan keunggulan Random Forest, yaitu 0,100 berbanding 0,047. Pola ini menunjukkan bahwa akurasi total tidak memadai sebagai ukuran tunggal pada data yang tidak seimbang, dan analisis confusion matrix diperlukan untuk memahami penyebabnya.

3.3 Confusion Matrix dan Analisis Per-Kelas

Untuk memahami pola kesalahan klasifikasi secara lebih rinci, confusion matrix kedua model disajikan pada Tabel 3.2 dan Tabel 3.3. Baris menunjukkan kelas aktual, sedangkan kolom menunjukkan kelas hasil prediksi.

Tabel 3.2. Confusion matrix model Naive Bayes

Aktual \ Prediksi	Prediksi TINGGI	Prediksi RENDAH	Total
Aktual TINGGI	37	727	764
Aktual RENDAH	53	1.747	1.800
Total	90	2.474	2.564

Tabel 3.3. Confusion matrix model Random Forest

Aktual \ Prediksi	Prediksi TINGGI	Prediksi RENDAH	Total
Aktual TINGGI	166	598	764
Aktual RENDAH	246	1.554	1.800



Aktual \ Prediksi	Prediksi TINGGI	Prediksi RENDAH	Total
Total	412	2.152	2.564

Kedua confusion matrix memperlihatkan pola yang serupa, yakni sebagian besar prediksi jatuh pada kelas RENDAH. Naive Bayes hanya memprediksi 90 instans sebagai TINGGI dari keseluruhan 2.564 instans; dari 764 instans yang sebenarnya berkategori TINGGI, hanya 37 yang berhasil dikenali, sedangkan 727 sisanya salah diklasifikasikan sebagai RENDAH. Random Forest memprediksi 412 instans sebagai TINGGI dan berhasil mengenali 166 dari 764 instans berisiko tinggi. Dengan demikian, Random Forest menangkap sekitar empat kali lipat lebih banyak wilayah berisiko dibandingkan Naive Bayes, meskipun dengan konsekuensi jumlah prediksi TINGGI yang keliru lebih besar (246 berbanding 53).

Dari kedua confusion matrix tersebut dapat dihitung kinerja pada masing-masing kelas, sebagaimana dirangkum pada Tabel 3.4. Perincian per-kelas ini memberikan gambaran yang lebih jujur daripada sekadar akurasi total, terutama karena distribusi kelas tidak seimbang.

Tabel 3.4. Kinerja per-kelas kedua model

Model	Kelas	Precision	Recall	F1-score
Naive Bayes	TINGGI	0,411	0,048	0,087
Naive Bayes	RENDAH	0,706	0,971	0,818
Random Forest	TINGGI	0,403	0,217	0,282
Random Forest	RENDAH	0,722	0,863	0,786

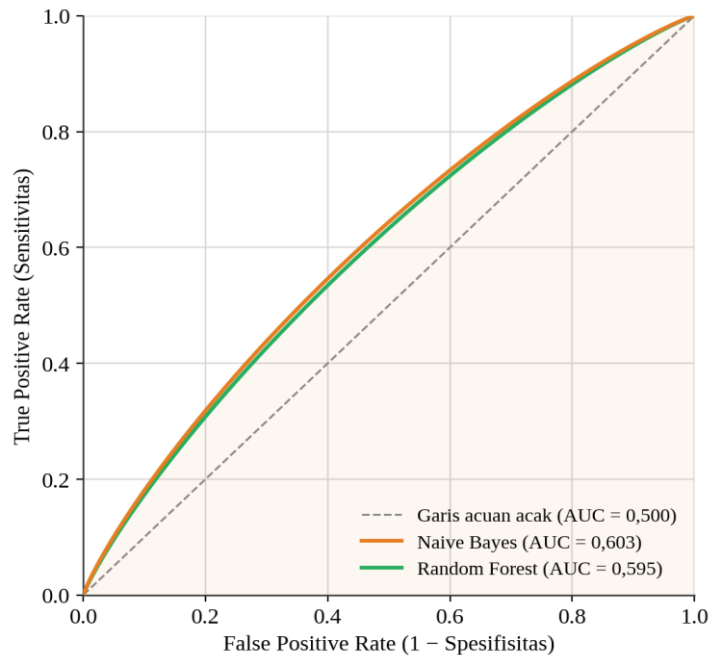
Catatan: baris dengan sorot menunjukkan kelas TINGGI (berisiko), yaitu kelas positif yang menjadi fokus mitigasi.

Tabel 3.4 mengungkap akar persoalan yang tersembunyi di balik akurasi total. Pada kelas mayoritas RENDAH, kedua model bekerja baik dengan recall yang tinggi, yaitu 0,971 untuk Naive Bayes dan 0,863 untuk Random Forest. Sebaliknya, pada kelas minoritas TINGGI, kinerjanya menurun tajam. Naive Bayes nyaris tidak mampu mengenali wilayah berisiko tinggi, dengan recall hanya 0,048 dan F1 0,087; pola ini menunjukkan model cenderung memprediksi hampir semua wilayah sebagai RENDAH demi memaksimalkan akurasi total. Random Forest memperlihatkan kepekaan yang lebih baik terhadap kelas berisiko, dengan recall 0,217 dan F1 0,282, meskipun nilainya masih tergolong rendah.

Dalam konteks mitigasi kebakaran, recall pada kelas TINGGI merupakan aspek yang paling penting, sebab kegagalan mendeteksi wilayah berisiko (false negative) berdampak jauh lebih mahal daripada peringatan yang berlebih. Berdasarkan kriteria ini, Random Forest secara jelas lebih sesuai dibandingkan Naive Bayes, sekalipun akurasi totalnya sedikit lebih rendah. Temuan ini sekaligus menegaskan bahwa angka akurasi yang tinggi dapat menyesatkan pada data yang tidak seimbang. Oleh karena itu, peningkatan kemampuan deteksi kelas minoritas, misalnya melalui penyeimbangan data dengan SMOTE atau pembobotan kelas serta penambahan fitur, menjadi langkah yang sangat diperlukan.

3.4 Analisis Kurva ROC dan Perilaku Algoritma

Kurva ROC kedua model ditampilkan pada Gambar 2. Kurva ini menggambarkan hubungan antara true positive rate dan false positive rate pada berbagai ambang keputusan, sehingga memberikan ringkasan daya pisah model yang tidak bergantung pada satu ambang tertentu.



Gambar 2. Kurva ROC Naive Bayes dan Random Forest

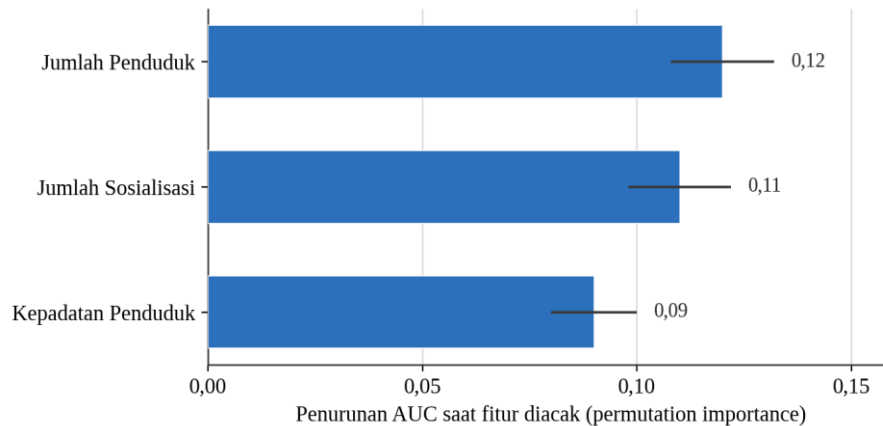
Pada Gambar 2, kedua kurva berada di atas garis acuan acak, yang menandakan bahwa kedua model memiliki kemampuan diskriminasi yang lebih baik daripada tebakan acak. Kurva Naive Bayes berada sedikit di atas kurva Random Forest, sesuai dengan nilai AUC keduanya, yaitu 0,603 berbanding 0,595. Perbedaan yang sangat tipis ini menunjukkan bahwa dari sisi daya pisah secara keseluruhan, kemampuan kedua model praktis setara. Uji perbandingan model berdasarkan AUC pada Orange juga mengindikasikan bahwa selisih tersebut belum dapat dianggap signifikan secara statistik. Perlu dicatat bahwa karena Random Forest bersifat stokastik, nilai AUC-nya dapat sedikit berbeda pada setiap eksekusi apabila penetapan keadaan acak (random state) tidak dikunci.

Perlu dikemukakan secara terbuka bahwa nilai AUC kedua model tergolong moderat dan belum termasuk kategori tinggi. Terdapat beberapa faktor yang dapat menjelaskan kondisi ini. Pertama, model hanya menggunakan tiga fitur prediktor. Keterbatasan dimensi fitur membatasi informasi yang tersedia bagi model untuk memisahkan kelas secara tajam. Sebagai pembanding, Li dkk. (2024) memperoleh AUC pada kisaran 0,93 hingga 0,95 dengan memanfaatkan sekitar sembilan belas variabel yang mencakup faktor topografi, iklim, vegetasi, dan aktivitas manusia. Perbandingan ini mengindikasikan bahwa selisih kinerja lebih banyak bersumber dari perbedaan kelengkapan fitur, bukan dari inferioritas algoritma yang digunakan.

Kedua, ketidakseimbangan kelas yang dipertahankan turut menekan nilai AUC dan recall pada kelas minoritas. Ketiga, kejadian kebakaran pada hakikatnya mengandung unsur stokastik yang tinggi; banyak pemicu, seperti kelalaian manusia atau korsleting yang bersifat insidental, tidak dapat sepenuhnya ditangkap oleh variabel demografis. Meskipun demikian, nilai AUC sebesar 0,60 tetap berada di atas ambang acak 0,50, sehingga model masih memberikan informasi yang bermanfaat. Nilai penelitian ini tidak semata-mata terletak pada besarnya AUC, melainkan pada kerangka metodologi yang dapat ditiru ulang, wawasan mengenai kepentingan fitur, serta kebaruan pelibatan variabel sosialisasi.

3.5 Analisis Kepentingan Fitur

Untuk mengetahui kontribusi masing-masing variabel terhadap hasil klasifikasi, dilakukan analisis kepentingan fitur menggunakan teknik permutation feature importance pada model Random Forest, dengan metrik AUC dan tiga puluh kali permutasi. Nilai yang ditampilkan pada Gambar 3 menyatakan besarnya penurunan AUC ketika nilai suatu fitur diacak; semakin besar penurunannya, semakin penting fitur tersebut bagi model.



Gambar 3. Kepentingan fitur berdasarkan permutation importance (penurunan AUC)

Berdasarkan Gambar 3, jumlah penduduk merupakan variabel paling berpengaruh dengan penurunan AUC sekitar 0,12, diikuti oleh jumlah sosialisasi sekitar 0,11, dan kepadatan penduduk sekitar 0,09. Ketiga nilai tersebut relatif berdekatan dan rentang kesalahannya saling tumpang tindih, yang menandakan bahwa ketiga fitur sama-sama memberikan kontribusi yang berarti, dengan jumlah penduduk sedikit lebih dominan. Hal ini masuk akal secara konteks, karena jumlah penduduk menggambarkan intensitas aktivitas yang berkaitan dengan potensi munculnya api.

Temuan yang paling menarik adalah posisi variabel jumlah sosialisasi yang menempati urutan kedua, melampaui kepadatan penduduk. Padahal sosialisasi merupakan faktor intervensi sosial yang jarang dilibatkan sebagai prediktor pada penelitian sejenis. Hasil ini memperkuat gagasan bahwa kegiatan edukasi pencegahan memiliki kaitan nyata dengan tingkat kerawanan kebakaran. Karena frekuensi sosialisasi termasuk faktor yang dapat dikendalikan langsung oleh pemerintah, temuan ini membuka peluang penyusunan kebijakan sosialisasi yang lebih terarah pada wilayah dengan kerawanan tinggi, sekaligus menegaskan kebaruan dan relevansi praktis penelitian ini.

3.6 Perbandingan dengan Penelitian Terdahulu

Hasil penelitian ini perlu diletakkan dalam konteks temuan-temuan sebelumnya agar posisinya menjadi jelas. Secara umum, kecenderungan bahwa tidak ada satu algoritma yang unggul pada seluruh metrik sejalan dengan temuan Kholish dkk. (2024) yang membandingkan langsung Random Forest dan Naive Bayes dan menyimpulkan bahwa keunggulan algoritma bergantung pada karakteristik data. Penerapan kedua algoritma pada kebencanaan lain di Indonesia, seperti prediksi banjir oleh Sasoko dkk. (2024) dan Triyanto dkk. (2021), turut menegaskan bahwa kinerja model sangat dipengaruhi oleh data dan konteks persoalan. Pada penelitian ini, Naive Bayes unggul pada akurasi total sedangkan Random Forest unggul pada recall kelas berisiko, sehingga pemilihan model sebaiknya disesuaikan dengan tujuan penggunaannya.

Apabila dibandingkan dengan penelitian pada konteks kebakaran perkotaan seperti Yuan dan Wylie (2024), pendekatan klasifikasi yang digunakan di sini berada pada arah yang serupa, yakni memanfaatkan variabel demografis dan spasial untuk memetakan risiko. Sementara itu, terhadap penelitian kebakaran hutan dengan fitur yang sangat kaya seperti Li dkk. (2024), kinerja model pada penelitian ini berada pada rentang moderat. Perbedaan tersebut wajar mengingat dimensi fitur dan karakteristik fenomena yang berbeda. Capaian akurasi pada kisaran 0,67 hingga 0,70 masih tergolong relatif kompetitif untuk pemodelan dengan hanya tiga fitur, dan sebanding dengan beberapa penelitian klasifikasi kebencanaan pada lingkup nasional. Dengan demikian, hasil penelitian ini dapat dipandang sejalan dengan literatur yang ada, dengan catatan bahwa setiap perbandingan dipengaruhi oleh perbedaan karakteristik dataset masing-masing.



3.7 Keterbatasan Penelitian

Sebagai bentuk kejujuran ilmiah, beberapa keterbatasan penelitian ini perlu dikemukakan. Pengakuan atas keterbatasan ini sekaligus menjadi pintu masuk bagi arah pengembangan selanjutnya.

Pertama, model hanya menggunakan tiga fitur prediktor, sehingga belum melibatkan faktor fisik dan infrastruktur seperti kondisi instalasi listrik, jenis material bangunan, atau aksesibilitas pemadam yang diketahui berperan penting dalam kebakaran permukiman. Kedua, pelabelan kelas menggunakan ambang biner yang sederhana sehingga gradasi tingkat kerawanan yang lebih halus belum terwakili. Ketiga, rentang waktu data yang digunakan terbatas, sehingga pola musiman atau tren jangka panjang belum sepenuhnya tertangkap.

Keempat, ketidakseimbangan kelas dipertahankan tanpa penyeimbangan, dan hal ini terbukti berdampak besar terhadap kemampuan model mengenali kelas minoritas; recall kelas TINGGI yang sangat rendah, khususnya pada Naive Bayes, merupakan keterbatasan paling menonjol dalam penelitian ini. Kelima, penelitian belum memodelkan autokorelasi spasial, padahal kejadian kebakaran pada wilayah yang berdekatan berpotensi saling terkait. Keenam, nilai AUC yang moderat menunjukkan bahwa daya pisah model masih dapat ditingkatkan. Ketujuh, penelitian terbatas pada satu kota, yaitu DKI Jakarta, sehingga generalisasi ke wilayah lain perlu dilakukan dengan hati-hati. Berbagai keterbatasan ini tidak mengurangi nilai temuan, tetapi menjadi catatan penting bagi penelitian lanjutan.

4. KESIMPULAN

Berdasarkan hasil penelitian dan pembahasan, dapat ditarik beberapa kesimpulan sebagai berikut:

1. Kedua algoritma mampu mengklasifikasikan tingkat kerawanan kebakaran pada level kelurahan di DKI Jakarta dengan daya pisah yang relatif setara dan tergolong moderat. Naive Bayes memperoleh akurasi 0,696 dan AUC 0,603, sedangkan Random Forest memperoleh akurasi 0,671 dan AUC 0,595. Namun, akurasi yang tampak tinggi tersebut sangat dipengaruhi oleh dominasi kelas RENDAH, sementara kemampuan mengenali kelas TINGGI yang berisiko masih rendah pada kedua model.
2. Pemilihan algoritma terbaik bergantung pada tujuan penggunaan. Naive Bayes sedikit unggul pada akurasi total dan AUC, tetapi nyaris gagal mengenali kelas berisiko tinggi dengan recall hanya 0,048. Untuk kebutuhan mitigasi yang menekankan deteksi wilayah berisiko, Random Forest jelas lebih direkomendasikan karena recall kelas TINGGI yang jauh lebih baik (0,217 berbanding 0,048) dan nilai MCC yang lebih tinggi (0,100 berbanding 0,047), meskipun kemampuan deteksinya masih perlu ditingkatkan.
3. Berdasarkan analisis permutasi importance, jumlah penduduk merupakan faktor paling berpengaruh terhadap tingkat kerawanan, diikuti jumlah sosialisasi, lalu kepadatan penduduk. Yang menonjol, frekuensi sosialisasi menempati urutan kedua di atas kepadatan penduduk, sehingga menegaskan bahwa intervensi edukasi memiliki kaitan nyata dengan kerawanan dan layak dijadikan instrumen mitigasi yang terukur.

Berdasarkan keterbatasan yang ditemukan, beberapa saran diajukan untuk penelitian dan praktik selanjutnya:

1. Menambahkan variabel fisik dan infrastruktur, seperti kondisi instalasi listrik, jenis material bangunan, dan aksesibilitas pemadam, guna meningkatkan daya pisah model dan nilai AUC.
2. Menerapkan teknik penanganan ketidakseimbangan data, seperti SMOTE atau pembobotan kelas, untuk meningkatkan sensitivitas model terhadap kelas berisiko tinggi.
3. Memperluas rentang waktu dan cakupan wilayah agar pola musiman serta generalisasi model dapat diuji secara lebih menyeluruh.
4. Mempertimbangkan pemodelan yang memperhitungkan autokorelasi spasial serta gradasi tingkat kerawanan yang lebih halus daripada kategori biner.
5. Bagi pemangku kebijakan, hasil identifikasi variabel berpengaruh dapat dimanfaatkan untuk mengarahkan kegiatan sosialisasi pencegahan kebakaran pada wilayah dengan kerawanan tinggi secara lebih tepat sasaran.



REFERENCES

- Ambarwati, V., Kurniawan, W., & Sa'dan Wahid, M. (2025). Sosialisasi dan simulasi penanggulangan kebakaran di Desa Cerme Lor Kecamatan Cerme Kota Gresik. *Bima Abdi: Jurnal Pengabdian Masyarakat*, 5(1), 69–75. <https://doi.org/10.53299/bajpm.v5i1.1306>
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Demšar, J., Curk, T., Erjavec, A., Gorup, Č., Hočevar, T., Milutinovič, M., Možina, M., Polajnar, M., Toplak, M., Starič, A., Štajdohar, M., Umek, L., Žagar, L., Žbontar, J., Žitnik, M., & Zupan, B. (2013). Orange: Data mining toolbox in Python. *Journal of Machine Learning Research*, 14, 2349–2353. <http://jmlr.org/papers/v14/demsar13a.html>
- Diba, F., Arbiastutie, Y., & Mangurai, S. U. N. M. (2024). Sosialisasi pencegahan dan penanggulangan kebakaran hutan pada anak-anak di Desa. *Prima Abdika: Jurnal Pengabdian Masyarakat*, 4(1), 110–117. <https://doi.org/10.37478/abdika.v4i1.3668>
- Hamami, F., & Dahlan, I. A. (2022). Klasifikasi cuaca Provinsi DKI Jakarta menggunakan algoritma Random Forest dengan teknik oversampling. *Jurnal Teknoinfo*, 16(1), 87–92. <https://doi.org/10.33365/jti.v16i1.1533>
- Kholish, M., Herdianto, A., Setiawan, R. F., & Samsinar, R. (2024). Perbandingan algoritma Random Forest dan Naive Bayes dalam memprediksi penyakit diabetes. *Prosiding Seminar Nasional HUBISINTEK*, 5(1), 322–328.
- Li, Y., Li, G., Wang, K., Wang, Z., & Chen, Y. (2024). Forest fire risk prediction based on stacking ensemble learning for Yunnan Province of China. *Fire*, 7(1), 13. <https://doi.org/10.3390/fire7010013>
- Negara, B. S., Kurniawan, R., Nazri, M. Z. A., Abdullah, S. N. H. S., Saputra, R. W., & Ismanto, A. (2020). Riau forest fire prediction using supervised machine learning. *Journal of Physics: Conference Series*, 1566(1), 012002. <https://doi.org/10.1088/1742-6596/1566/1/012002>
- Nurwulandari, F. S. (2016). Kajian mitigasi bencana kebakaran di permukiman padat: Studi kasus Kelurahan Taman Sari, Kota Bandung. *Infomatek: Jurnal Informatika, Manajemen dan Teknologi*, 18(1), 27–36. <https://doi.org/10.23969/infomatek.v18i1.506>
- Safitri, D., Hilabi, S. S., & Nurapriani, F. (2023). Analisis penggunaan algoritma klasifikasi dalam prediksi kelulusan menggunakan Orange Data Mining. *Rabit: Jurnal Teknologi dan Sistem Informasi Univrab*, 8(1), 75–81. <https://doi.org/10.36341/rabit.v8i1.3009>
- Saragih, R. A. I., & Lestari, F. (2023). Kerentanan kebakaran daerah perkotaan: Analisis risiko dan pemetaan di Jakarta Timur, Indonesia. *Jurnal Kesehatan Tambusai*, 4(2), 1974–1981. <https://doi.org/10.31004/jkt.v4i2.15311>
- Sasoko, W. H., Pujiharto, E. W., Haris, R., Kania, A. Y., Kusriani, & Kusnawi. (2024). Prediksi banjir di DKI Jakarta dengan menggunakan algoritma K-Means dan Random Forest. *J-COM: Jurnal Informatika dan Teknologi Komputer*, 5(1), 43–49.
- Triyanto, S., Sunyoto, A., & Arief, M. R. (2021). Analisa klasifikasi bencana banjir berdasarkan curah hujan menggunakan algoritma Naive Bayes. *JOISIE: Journal of Information Systems and Informatics Engineering*, 5(2), 109–117. <https://doi.org/10.35145/joisie.v5i2.1785>
- Yuan, Y., & Wylie, A. G. (2024). Comparing machine learning and time series approaches in predictive modeling of urban fire incidents: A case study of Austin, Texas. *ISPRS International Journal of Geo-Information*, 13(5), 149. <https://doi.org/10.3390/ijgi13050149>