

Implementasi Metode *Random Forest* untuk Peningkatan Efisiensi Penilaian Status Uji Kelayakan Kendaraan Bermotor di Kota Malang

*Hamidah Lutfiyanti Maharani¹, Syahiduz Zaman²

^{1,2}Teknik Informatika, Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang

Jl. Gajayana No. 50, Kota Malang, Jawa Timur

Email: ¹hamidahalem@gmail.com, ²syahid@ti.uin-malang.ac.id

ABSTRACT

The growth of vehicle volume in Malang City presents challenges in the form of increased accident risks, especially if the technical condition of the vehicles does not meet standards. As a preventive measure, the Motor Vehicle Feasibility Testing (KIR Test) is conducted to ensure that vehicles comply with safety standards. However, manual assessments in this process are prone to human error, necessitating a more efficient and accurate system. This study implements the Random Forest method to classify the eligibility status of motor vehicles, focusing on two main categories: public and private vehicles. This implementation is expected to improve the efficiency and accuracy of the KIR test process. Among the data split ratios tested, a 60% training data and 40% test data ratio yielded the best results with an accuracy of 86.94% and an OOB error rate of 13.03%, indicating the model's error rate on data not used during training. These results indicate that the Random Forest method effectively identifies the eligibility status of motor vehicles with an optimal data configuration.

Keywords : machine learning; random forest; vehicle inspection

ABSTRAK

Pertumbuhan volume kendaraan di Kota Malang menghadirkan tantangan berupa peningkatan risiko kecelakaan, terutama jika kondisi teknis kendaraan tidak memenuhi standar. Sebagai upaya *preventif*, Uji Kelayakan Kendaraan Bermotor (Uji KIR) dilakukan untuk memastikan kendaraan memenuhi standar keselamatan. Namun, penilaian manual dalam proses ini rentan terhadap kesalahan manusia, sehingga diperlukan sistem yang lebih efisien dan akurat. Penelitian ini mengimplementasikan metode *Random Forest* untuk mengklasifikasikan status kelayakan kendaraan bermotor, dengan fokus pada dua kategori utama, yaitu kendaraan umum dan pribadi. Implementasi ini diharapkan dapat meningkatkan efisiensi dan akurasi dalam pelaksanaan uji KIR. Dari beberapa rasio pembagian data yang diuji, rasio 60% data latih dan 40% data uji memberikan hasil terbaik dengan akurasi 86,94% dan *OOB error rate* sebesar 13,03% menunjukkan tingkat kesalahan model pada data yang tidak digunakan saat pelatihan. Hasil ini menyatakan bahwa metode *Random Forest* secara efektif mengidentifikasi status kelayakan kendaraan bermotor dengan konfigurasi data yang optimal.

Kata kunci : *machine learning; random forest; uji KIR*

1. PENDAHULUAN

Pesatnya pertumbuhan Kota Malang, baik dari segi jumlah penduduk maupun volume kendaraan, menghadirkan tantangan besar dalam menjaga keselamatan lalu lintas di jalan-jalan yang semakin padat. Mengacu Badan Pusat Statistika (BPS), jumlah kendaraan bermotor di Kota Malang dari tahun 2022 meningkat sebesar 20,9% dari tahun sebelumnya. Peningkatan jumlah kendaraan, terutama jika tidak diimbangi dengan kondisi teknis yang memadai, dapat meningkatkan risiko kecelakaan (Susanto *et al.*, 2019). Sedangkan, kecelakaan lalu lintas yang terjadi di Kota Malang pada tahun 2022 sebanyak 371 kejadian, angka ini menunjukkan perlunya perhatian lebih terhadap masalah tersebut (Dwi Rahayuning Surastia *et al.*, 2023). Sebagai langkah *preventif*, pemerintah melalui Dinas Perhubungan memberlakukan uji KIR, atau Uji Kelayakan Kendaraan Bermotor, untuk memastikan bahwa kendaraan yang beroperasi di jalan raya memenuhi standar keselamatan. Uji KIR ini diwajibkan setiap enam bulan sekali untuk jenis kendaraan tertentu, seperti mobil barang dan mobil penumpang,

sebagaimana diatur dalam Undang-Undang No. 22 Tahun 2009 tentang Lalu Lintas dan Angkutan Jalan (Mentiri, 2022).

Uji KIR bertujuan untuk memastikan bahwa kendaraan berada dalam kondisi aman dan layak jalan (Subhan *et al.*, 2024). Proses ini mencakup pemeriksaan berbagai aspek teknis, seperti sistem pengereman, emisi gas buang, dan kondisi fisik kendaraan secara keseluruhan (Purba & Azwar, 2024). Kendaraan yang memenuhi seluruh kriteria teknis akan dinyatakan lulus dan diberikan sertifikat kelayakan yang berlaku selama enam bulan. Kendaraan yang tidak lulus harus diperbaiki dan diuji ulang sebelum diizinkan kembali beroperasi.

Meskipun sistem uji KIR telah memenuhi standar keselamatan, proses penilaian kelayakan masih dilakukan secara manual oleh petugas uji, yang membuka peluang terjadinya kesalahan manusia. Penelitian ini menggunakan data dari kendaraan yang telah melalui proses uji KIR dan dinyatakan lulus. Data tersebut dianalisis untuk mengidentifikasi faktor-faktor yang memengaruhi kelulusan kendaraan. Oleh karena itu, diperlukan analisis lebih

mendalam guna meningkatkan efisiensi proses uji KIR serta menghasilkan sistem otomatis yang lebih andal dalam menilai kelayakan kendaraan.

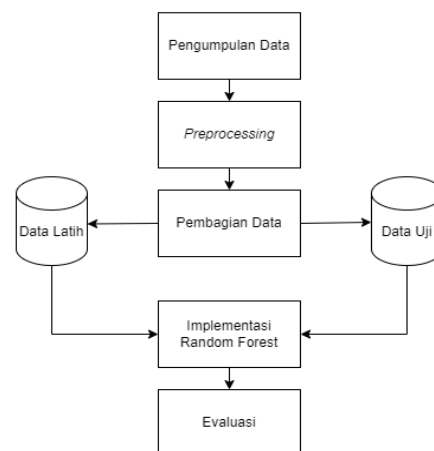
Guna meningkatkan efisiensi dan akurasi penelitian uji KIR, digunakan metode machine learning, yaitu *Random Forest*. Metode ini dipilih karena kemampuannya menangani data yang kompleks dan beragam serta menghasilkan model yang *robust* meskipun data tidak seimbang. Penerapan *Random Forest* diharapkan dapat mempercepat proses uji KIR, meningkatkan akurasi, dan mengurangi kesalahan penilaian manusia. Analisis dilakukan untuk mengkaji pengaruh fitur-fitur seperti jenis kendaraan, bahan bakar, dan tipe uji terhadap kelulusan uji KIR di Kota Malang. Selain itu, perhatian diberikan pada perbedaan status antara kendaraan umum dan kendaraan pribadi, yang dapat dibedakan berdasarkan warna plat nomor yang digunakan.

Penelitian sebelumnya dilakukan dalam mengklasifikasi tingkat kepadatan lalu lintas dengan menggunakan metode *Random Forest* yang menghasilkan akurasi sebesar 95,6%, tingkat *error* 4,4%, dan nilai *recall* 100% (Kusumah,

2020). (Irwansyah *et al.*, 2023) juga melakukan penelitian mengklasifikasikan uji emisi sepeda motor dengan metode *Naïve Bayes* yang menghasilkan akurasi 91,49%. Meskipun penelitian terdahulu menggunakan metode *Random Forest* yang sama, penelitian tersebut berbeda dari studi ini dalam hal kasus yang dianalisis dan fitur-fitur yang digunakan. Oleh karena itu, hasil akurasi yang diperoleh dalam penelitian ini mungkin berbeda, mengingat variasi pada konteks dan variabel yang digunakan dalam penilaian uji KIR kendaraan.

2. METODE

Penelitian ini terdiri dari beberapa tahapan dari penerapan metode *Random Forest* dengan menggunakan *dataset*. Alur penelitian tersebut dapat dilihat pada Gambar 1.



Gambar 1. Tahap Penelitian

2.1. Pengumpulan Data

Tahap pengumpulan data bertujuan untuk memperoleh informasi yang relevan dan akurat sebagai dasar analisis dan pemodelan prediksi menggunakan teknik pembelajaran mesin (Agung *et al.*, 2023). Data yang digunakan berasal dari situs Data Malang Kota berjudul "Data Lulus Uji dan Wajib Uji Pengujian Kendaraan Bermotor Kota Malang Tahun 2022" (<https://data.malangkota.go.id/>), dengan fokus pada data lulus uji KIR. Meskipun terdiri dari dua kategori, yaitu lulus uji dan wajib uji, penelitian ini hanya menganalisis data lulus uji KIR. Untuk mempermudah proses analisis, data yang awalnya berformat xlsx diubah menjadi csv. Data tersebut mencakup 12.285 entri dengan 10 atribut yang berkaitan dengan uji kelayakan kendaraan bermotor seperti pada Tabel 1.

Tabel 1. Detail *Dataset*

Atribut	Keterangan
no	Nomor urut data dalam dataset
tgl_uji	Tanggal pengujian kendaraan bermotor
tgl_jt_tempo	Tanggal jatuh tempo uji KIR berikutnya
no_uji	Nomor identifikasi pengujian
no_kend	Plat Kendaraan
nama_pemilik	Nama pemilik kendaraan
jenis	Jenis kendaraan yang diuji, seperti mobil barang, mobil penumpang, dll.

nm_b_bakar	Jenis bahan bakar yang digunakan.
nm_status	Status penggunaan kendaraan, apakah Umum atau Tidak Umum.
tipe_uji	Jenis uji kelayakan yang dilakukan

2.2. Preprocessing

Dataset yang digunakan dalam sebuah penelitian melalui tahap *preprocessing* yang mencakup pembersihan data dari noise, nilai yang hilang, dan penghapusan atribut tidak relevan seperti `tgl_uji`, `tgl_jatuh_tempo`, `nama`, `no_kend`, dan `no_uji`. Atribut ini dihapus karena bersifat spesifik dan unik, sehingga dapat menambah noise dan mengganggu proses pembelajaran mesin. Dengan adanya *preprocessing* ini memastikan hanya data relevan yang digunakan untuk meningkatkan kinerja model (Muttaqin *et al.*, 2023). Setelah pembersihan, dilakukan encoding untuk mengonversi fitur kategorikal menjadi numerik agar dapat diproses dengan optimal.

2.3. Pembagian Data

Setelah proses pembersihan dan konversi, dataset dibagi menjadi dua bagian yaitu fitur dan target yang akan diprediksi. Data kemudian dipisahkan dengan beberapa rasio, yaitu 80:20,

60:40, dan 50:50 (Reza, 2023). Rasio 80% untuk data latih dan 20% untuk data uji sering digunakan dalam pembelajaran mesin karena dianggap memberikan keseimbangan yang optimal antara pelatihan dan pengujian model. Meski begitu, rasio 60:40 dan 50:50 juga diuji dalam studi ini. Eksperimen dengan berbagai rasio pembagian ini bertujuan untuk mengevaluasi stabilitas model. *Preprocessing* dan pembagian data yang tepat sangat penting untuk mencapai hasil prediktif yang optimal (Mediansyah *et al.*, 2024).

2.4. Data Latih

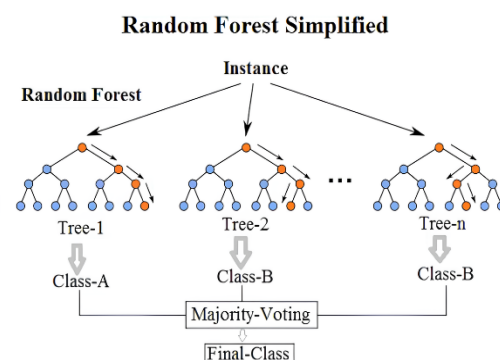
Data yang telah di *preprocessing* dibagi menjadi beberapa rasio untuk data latih, guna membangun dan mengoptimalkan model prediktif. Proses pelatihan melibatkan iterasi berulang, di mana model belajar mengenali pola dalam data untuk menghasilkan prediksi akurat pada data uji. Evaluasi berkala dilakukan untuk mencegah *overfitting* dan memastikan model dapat menggeneralisasi dengan baik.

2.5. Data Uji

Setelah proses pelatihan model selesai, langkah berikutnya adalah menguji kinerjanya. Sebanyak 20%,

40%, dan 50% dari dataset dialokasikan sebagai data uji untuk mengevaluasi kinerja model yang telah dilatih. Data uji digunakan untuk memastikan model dapat menggeneralisasi pola ke data baru. Hasil prediksi dibandingkan dengan nilai sebenarnya untuk mengukur akurasi, presisi, recall, dan metrik evaluasi lainnya (Nainggolan *et al.*, 2024). Penggunaan berbagai rasio data uji memastikan model bekerja baik tidak hanya pada data latih tetapi juga pada data baru.

2.6. Implementasi Metode *Random Forest*



Gambar 2. 1 Random Forest Simplified
(Sumber : learningbox.coffeecup.com, 2024)

Dalam penelitian ini menggunakan metode *Random Forest*, sebuah teknik pembelajaran mesin berbasis ensemble yang menggabungkan beberapa pohon keputusan (*decision trees*) untuk meningkatkan akurasi dan stabilitas prediksi (Febtiawan *et al.*, 2024). Setiap pohon dibangun dengan subset acak dari

data latih dan fitur yang tersedia, yang membantu mengurangi korelasi antar pohon serta memastikan model tidak bergantung pada fitur tertentu. Hasil prediksi dari setiap pohon digabungkan melalui voting mayoritas untuk menghasilkan prediksi akhir (Sinambela *et al.*, 2023). Dalam *Decision Tree*, *Indeks Gini* digunakan untuk memilih atribut terbaik dengan Persamaan 1.

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

Di mana p_i adalah proporsi kelas ke- i pada node. *Random Forest* dipilih karena kemampuannya dalam mengatasi variabilitas data dan meningkatkan kemampuan generalisasi model, menjadikannya metode yang efektif untuk tugas klasifikasi dalam penelitian ini.

2.7. Evaluasi

Kinerja hasil latih *Random Forest* dievaluasi menggunakan *metrix* akurasi, *precision*, *recall*, dan laporan klasifikasi. Akurasi memberikan gambaran umum, sementara *precision* dan *recall* menilai kemampuan model mendeteksi kelas tertentu (Sahmony & Rianto, 2024). *Confusion matrix* digunakan untuk menganalisis prediksi benar dan salah, serta pola kesalahan. *OOB error* juga dihitung untuk

mengestimasi kinerja tanpa data uji tambahan.

Hasil evaluasi didukung dengan visualisasi, termasuk barplot untuk menyoroti fitur penting dan *ROC curve* untuk mengukur kemampuan model membedakan kelas positif dan negatif menggunakan AUC. Visualisasi ini membantu memperjelas hasil numerik secara grafis.

3. HASIL DAN PEMBAHASAN

Hasil penelitian ini diawali dengan memproses data menggunakan *Google Collaboratory* dan *Python*. Dataset terdiri dari 12.285 entri dengan 3 variabel prediktor dan 1 variabel target, yaitu status, yang telah dikonversi ke format numerik.

3.1 Preprocessing Data

Tahap preprocessing dalam penelitian ini terdiri dari dua tahapan, yaitu *Feature Encoding* dan *Feature Selection*. Dataset yang digunakan memiliki 11 variabel yang perlu diproses terlebih dahulu agar siap digunakan dalam model. *Feature Encoding* dilakukan untuk mengubah data yang berupa teks menjadi format numerik, sehingga memudahkan model dalam mengenali dan memproses data tersebut.

Beberapa atribut dalam data mengandung informasi yang terlalu spesifik dan unik untuk uji kelayakan kendaraan, yang tidak hanya kurang berkontribusi dalam memprediksi pola umum. Oleh karena itu, atribut-atribut tersebut dihapus agar data yang digunakan relevan dan bersih dari elemen yang mengganggu proses klasifikasi seperti pada Gambar 2.

0	1	Mobil Barang	Solar	Umum	Berkala
1	2	Mobil Barang	Solar	Umum	Berkala
2	3	Mobil Barang	Bensin	Tidak Umum	Berkala
3	4	Mobil Barang	Solar	Tidak Umum	Berkala
4	5	Mobil Barang	Solar	Tidak Umum	Berkala

Gambar 2. Hasil *Preprocessing* Data

Gambar 2 menunjukkan lima entri pertama dari dataset setelah dilakukan *preprocessing*, sehingga, hanya atribut yang relevan yang tersisa, seperti 'jenis', 'nm_b_bakar', 'nm_status', dan 'tipe_uji'. Proses ini memastikan bahwa dataset siap digunakan untuk diolah oleh mesin.

3.2 *Random Forest (RF) OOB estimate of error rate*

Proses klasifikasi menggunakan metode *Random Forest* untuk klasifikasi dengan tiga skenario pembagian data: 50:50, 60:40, dan 80:20, untuk mengevaluasi kinerja model. Salah satu metrik yang dievaluasi adalah OOB

error rate, yang mengukur kesalahan model tanpa memerlukan data uji eksternal, karena dihitung dari data yang tidak digunakan oleh beberapa pohon dalam *Random Forest*.

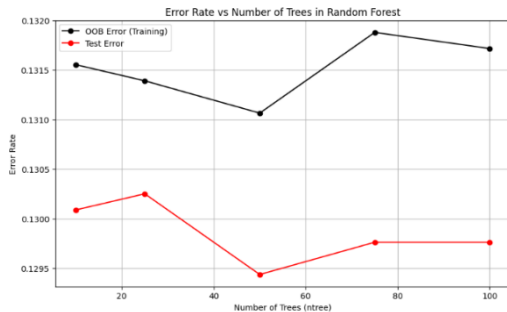
Tabel 2. *Error* OOB Setiap Rasio Data

<i>Training</i>	<i>Error OOB</i>
50%	13.17%
60%	13.03%
80%	12.94%

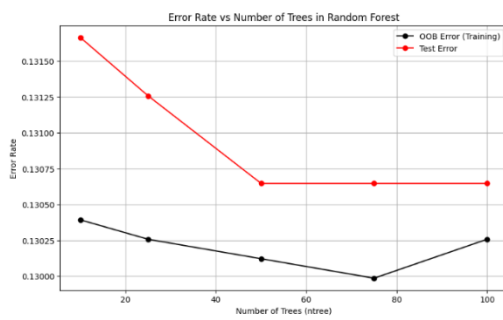
Tabel 2 menunjukkan hasil evaluasi *Out-of-Bag (OOB) error rate* pada model *Random Forest* yang diterapkan untuk uji kelayakan kendaraan bermotor dengan tiga skenario pembagian data latih. Hal ini menunjukkan bahwa dengan meningkatnya proporsi data latih, model semakin mampu memprediksi status uji kelayakan kendaraan bermotor baik untuk status umum (kendaraan komersial) maupun tidak umum (kendaraan pribadi) dengan lebih akurat. Oleh karena itu, pembagian data yang paling optimal adalah 80% data latih, karena menghasilkan error OOB terendah, menandakan kinerja model yang paling baik dalam penelitian ini.

Visualiasi grafik jumlah pohon (*n tree*) juga dilakukan dalam penelitian ini, guna melihat kinerja model dalam

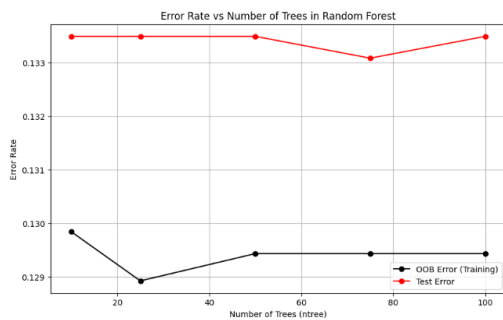
berbagai rasio data seperti pada Gambar 3, 4, dan 5.



Gambar 3. Visualisasi *Error Rate* Rasio 50:50



Gambar 4. Visualisasi *Error Rate* Rasio 60:40



Gambar 5. Visualisasi *Error Rate* Rasio 80:20

Dari ketiga gambar visualisasi model *Random Forest* diatas dibandingkan berdasarkan tiga rasio pembagian data latih dan uji terdapat variasi jumlah pohon dari 10 hingga 100. Analisis dari ketiga visualisasi menunjukkan bahwa rasio 60:40

memberikan hasil yang paling konsisten, di mana test error menurun hingga stabil pada sekitar 50 pohon, mencerminkan kemampuan model yang baik dalam memprediksi data uji tanpa indikasi *overfitting*. Sebaliknya, rasio 50:50 dan 80:20 menunjukkan fluktuasi yang lebih besar pada *test error*, khususnya pada rasio 80:20, di mana *test error* tetap tinggi dan tidak mengalami penurunan signifikan dengan peningkatan jumlah pohon.

Dapat disimpulkan bahwa proses klasifikasi sudah stabil pada ntree ke-50, sehingga model yang didapatkan pada analisis RF dapat dilanjutkan ke tahap prediksi dengan menggunakan *confusion matrix*.

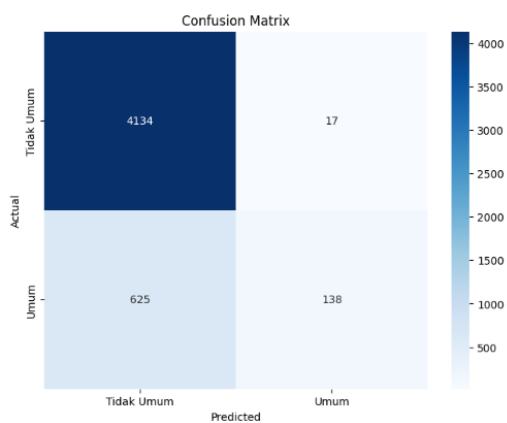
3.3 Evaluasi Model

Setelah melatih model *Random Forest* dengan berbagai rasio data, dilakukan evaluasi kinerja berdasarkan akurasi untuk menilai kemampuan model dalam mengklasifikasikan status kelayakan kendaraan bermotor. Akurasi menunjukkan proporsi prediksi benar dari total prediksi, memberikan gambaran umum tentang kemampuan model membedakan kendaraan yang layak dan tidak, berikut Tabel 3 hasil akurasi setiap rasio.

Tabel 3. Hasil Akurasi Setiap Rasio

<i>Rasio Data</i>	<i>Hasil Akurasi</i>
50 : 50	0.8702
60 : 40	0.8694
80 : 20	0.8665

Hasil evaluasi menunjukkan rasio 50:50 memberikan akurasi tertinggi (87,02%). Meski perbedaannya kecil, stabilitas dan generalisasi model penting dalam menentukan rasio optimal. Rasio 50:50 menunjukkan kinerja baik, tetapi berisiko *overfitting* jika data latih kurang representatif. Rasio 60:40 menawarkan keseimbangan lebih baik antara data latih dan uji seperti pada Gambar 6, sementara rasio 80:20 kurang menangkap variabilitas data sehingga menurunkan kemampuan generalisasi model.



Gambar 6. Confusion Matrix Rasio 60:40

Confusion matrix dengan rasio data 60:40, yang dipilih sebagai model terbaik, menunjukkan bahwa dari 4.134

kendaraan "Tidak Umum", model mengklasifikasikan semuanya dengan benar, dan hanya 17 yang salah menjadi "Umum". Namun, dari 763 kendaraan "Umum", 625 salah diklasifikasikan sebagai "Tidak Umum" (false negatives) dan hanya 138 yang benar. Ini menunjukkan model sangat baik dalam mendeteksi "Tidak Umum" dengan sedikit *false positives*, namun kesulitan mengklasifikasikan "Umum" dengan *false negatives* yang lebih tinggi. Hal ini disebabkan ketidakseimbangan data, namun secara keseluruhan model tetap menunjukkan performa baik seperti pada Gambar 7.

	precision	recall	f1-score	support
0	0.87	1.00	0.93	4151
1	0.89	0.18	0.30	763
accuracy			0.87	4914
macro avg	0.88	0.59	0.61	4914
weighted avg	0.87	0.87	0.83	4914

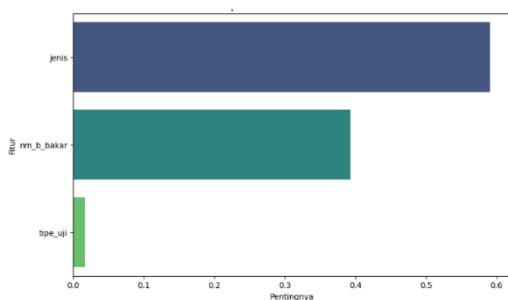
Gambar 7. Hasil Evaluasi Rasio 60:40

Evaluasi model *Random Forest* melalui *precision*, *recall*, dan *f1-score* menunjukkan performa yang baik pada kategori "Tidak Umum" (label 0). *Precision* mencapai 87%, *recall* 100%, dan *f1-score* 0.93. Namun, performa untuk kategori "Umum" (label 1) kurang memuaskan, dengan *recall* hanya 18%, meskipun *precision* mencapai 89%. *F1-score* rendah (0.30) menunjukkan model

sering gagal mengenali "Umum" dengan benar. Secara keseluruhan, akurasi model 87% dan weighted average f1-score 0.83, menunjukkan bahwa meskipun kuat dalam mendeteksi "Tidak Umum," model perlu peningkatan dalam mengidentifikasi "Umum" agar lebih seimbang.

3.4 Visualisasi

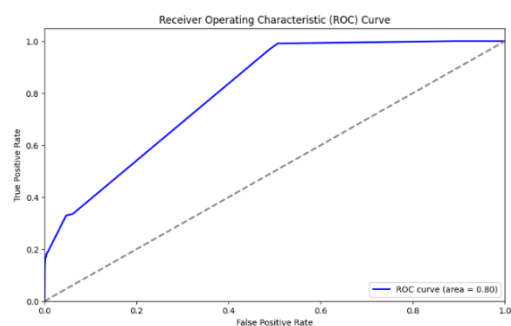
Penelitian ini menyajikan visualisasi yang menunjukkan pengaruh fitur-fitur terhadap hasil model *Random Forest* pada uji kelayakan kendaraan bermotor. Fitur seperti jenis kendaraan, jenis bahan bakar, dan jenis uji berkala dievaluasi untuk menilai seberapa besar kontribusinya terhadap keputusan akhir model.



Gambar 8. Feature Importance dari Model Random Forest

Gambar 8 di atas menampilkan tingkat pentingnya tiga fitur utama yang digunakan oleh model *Random Forest* untuk memprediksi status kelayakan kendaraan bermotor (`nm_status`).

Feature importance dihitung dari pengurangan impuritas pada setiap pohon keputusan, dengan fitur yang sering digunakan untuk pemisahan besar memiliki *importance* lebih tinggi. Fitur `jenis`, yang mengacu pada jenis kendaraan seperti mobil barang atau penumpang, memiliki pengaruh terbesar. Fitur `nm_b_bakar`, terkait jenis bahan bakar, berkontribusi cukup signifikan dengan nilai sekitar 0,35. Sebaliknya, fitur `tipe_uji`, yang mencerminkan jenis uji berkala, memiliki pengaruh paling kecil dengan tingkat pentingnya di bawah 0,1. Temuan ini menunjukkan bahwa jenis kendaraan dan bahan bakar adalah faktor utama dalam prediksi, sementara jenis uji memiliki peran yang lebih terbatas.



Gambar 9. Visualisasi ROC

Gambar 9 menunjukkan kinerja model *Random Forest* dengan rasio data 60:40 dalam memprediksi status kelayakan kendaraan bermotor. Kurva

ini menampilkan hubungan antara *True Positive Rate* (TPR) dan *False Positive Rate* (FPR), dengan AUC sebesar 0,80. AUC ini menunjukkan bahwa model cukup baik dalam membedakan kelas "Umum" dan "Tidak Umum." Meskipun performanya kuat, masih ada ruang untuk perbaikan, terutama dalam mengurangi *false positives* pada FPR yang lebih tinggi.

4. KESIMPULAN

Dari hasil penelitian yang telah dilakukan, *Random Forest* berhasil mengklasifikasikan status uji kelayakan kendaraan bermotor, dengan dua kategori utama yaitu umum (kendaraan komersial) dan tidak umum (kendaraan pribadi). Perbandingan rasio pembagian data dilakukan agar mengetahui efektivitas masing-masing rasio. Evaluasi menggunakan *confusion matrix* menunjukkan bahwa rasio 60:40 memberikan hasil yang paling optimal, dengan akurasi 86,94% dan OOB error rate sebesar 13,03%. Hasil ini lebih baik dibandingkan dengan rasio 80:20 yang menghasilkan akurasi 86,65% dan OOB error rate 12,94%, meskipun rasio 80:20 awalnya dianggap yang terbaik berdasarkan metrik *Error OOB*. Faktor yang menyebabkan model kurang andal

dalam mengklasifikasikan kategori kendaraan umum adalah ketidakseimbangan data, yang menyebabkan tingkat *false negatives* yang tinggi dalam kategori tersebut. Namun, secara keseluruhan, performa model dengan rasio 60:40 terbukti lebih efektif dalam penelitian ini.

DAFTAR PUSTAKA

- Agung, A., Daniswara, A., Kadek, I., & Nuryana, D. (2023). Data Preprocessing Pola Pada Penilaian Mahasiswa Program Profesi Guru. *Journal of Informatics and Computer Science*, *05*, 97–100.
- Dwi Rahayuning Surastia, Moch. Yunus, Anita Sulistyorini, & Marji Marji. (2023). Hubungan Pengetahuan, Sikap dan Kelelahan Kerja dengan Perilaku *Safety Riding* pada Pengendara Go-Jek di Kota Malang. *Jurnal Rumpun Ilmu Kesehatan*, *3*(3), 201–219. <https://doi.org/10.55606/jrik.v3i3.2663>
- Febtiawan, E. P., Akbar, L. A. S. I., & Rachman, A. S. (2024). Forecasting Produksi Energi *Photovoltaic* Menggunakan Algoritma *Random Forest Classification*. *Journal of Information System Research (JOSH)*, *5*(4), 1053–1062. <https://doi.org/10.47065/josh.v5i4.5514>
- Irwansyah, Dittyata, R., Rizal, & Wiyono. (2023). Optimalisasi Klasifikasi Uji Emisi Sepeda Motor Menggunakan Algoritma *Naive Bayes*. *Infotech: Journal Of Technology Information*, *9*(1), 67–76.
- Kusumah, H. (2020). Klasifikasi Tingkat Kepadatan Lalu Lintas Oleh

- Pengguna Kendaraan Bermotor Dengan Menggunakan *Random Forest* (Studi Kasus Jalan Raya By Pass Jomin Cikampek). Sekolah Tinggi Manajemen Informatika dan Komputer Likmi Bandung.
- Mediansyah, I., Septian, F., & Zikry, A. (2024). Penerapan *Whale Optimization Algorithm* dalam Pengoptimalan Portofolio Investasi Menggunakan Model *Prediktif Artificial Intelligence*. *Jurnal Software Engineering and Computational Intelligence*, 2(01), 50–58. <https://doi.org/10.36982/jseci.v2i01.4147>
- Mentiri, S. A. (2022). Tinjauan Terhadap Tugas dan Wewenang Dinas Perhubungan Kabupaten Poso dalam Pelaksanaan Uji Laik Jalan Angkutan Umum dan Angkutan Barang Menurut Undang-Undang Nomor 22 Tahun 2009.
- Muttaqin, Widiyanto, W. W., Munsarif, M., Mandias, G. F., Pungus, Stenly Richard Widarman, A., Hapsari, W. K., Hardiyanti, S. A., Fatkhudin, A., Pasnur, Bisono, E. F., Anshori, M., Suryani, & Saputra, N. (2023). Pengenalan *Data Mining* (R. W. & J. Simarmata (ed.); Issue July). Yayasan Kita Menulis.
- Nainggolan, E. S., Nasir, C., Fatoni, M., & Udariansyah, D. (2024). Perbandingan Klasifikasi Jenis Sampah Menggunakan *Convolutional Neural Network* Dengan *Arsitektur ResNet18* dan *ResNet50*. *CSRID Journal*, 16(1), 76. <https://www.doi.org/10.22303/csrid.1.1.2022.01-10>
- Purba, A. T., & Azwar, T. K. D. (2024). Perbandingan Kasus Pertanggungjawaban Pidana Korporasi Terhadap Kecelakaan Lalu Lintas Di Indonesia Dan Amerika Serikat. 6(2), 82–97.
- Reza, S. F. S. (2023). Implementasi Algoritma *Random Forest* Terhadap *Prediksi Good Loan/Bad Loan* Kredit Nasabah Bank Di Jakarta. *Prosiding Seminar Nasional Mahasiswa Bidang Ilmu Komputer Dan Aplikasinya*, 4(2), 535–543. <https://repository.upnvj.ac.id/25253/2/AWAL.pdf>
- Sahmony, N. F., & Rianto, H. (2024). Analisis Perbandingan Kinerja Model *Machine Learning* untuk Memprediksi Risiko Stunting pada Pertumbuhan Anak. *Institut Riset Dan Publikasi Indonesia (IRPI)*, 4(2), 413–422.
- Sinambela, D. P., Naparin, H., Zulfadhilah, M., & Hidayah, N. (2023). Implementasi Algoritma *Decision Tree* dan *Random Forest* dalam Prediksi Perdarahan Pascasalin. *Jurnal Informasi Dan Teknologi*, 5(3), 58–64. <https://doi.org/10.60083/jidt.v5i3.393>
- Subhan, A., Ritawaty, N., Nisvie, M. R., & Darmawandi, A. (2024). Analisis Kebijakan Overdimensi Kendaraan Uji KIR Terkait Keselamatan Jalan dan Efisiensi Muatan di Hulu Sungai Utara. *INNOVATIVE: Journal Of Social Science Research Volume*, 4, 4806–4827.
- Susanto, E., Samad, S., Gaus, A., & Sultan, M. A. (2019). Analisis Kemacetan Dan Gangguan Lalu Lintas Di Kota Ternate Tengah Ditinjau Dari Aspek, Perilaku Dan Budaya Masyarakat. *STABILITA : Jurnal Ilmiah Teknik Sipil*, 9(November), 119-.