

**KLASIFIKASI PRIORITAS PERBAIKAN JALAN DI KABUPATEN MALANG
MENGUNAKAN ALGORITMA *RANDOM FOREST***

SKRIPSI

Oleh :

IZZA SYAHRI MUHARRAM
NIM. 220605110073



**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2026**

**KLASIFIKASI PRIORITAS PERBAIKAN JALAN DI KABUPATEN
MALANG MENGGUNAKAN ALGORITMA RANDOM FOREST**

SKRIPSI

Diajukan kepada:
Universitas Islam Negeri Maulana Malik Ibrahim Malang
Untuk memenuhi Salah Satu Persyaratan dalam
Memperoleh Gelar Sarjana Komputer (S.Kom)

Oleh :
IZZA SYAHRI MUHARRAM
NIM. 220605110073

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2026**

HALAMAN PERSETUJUAN

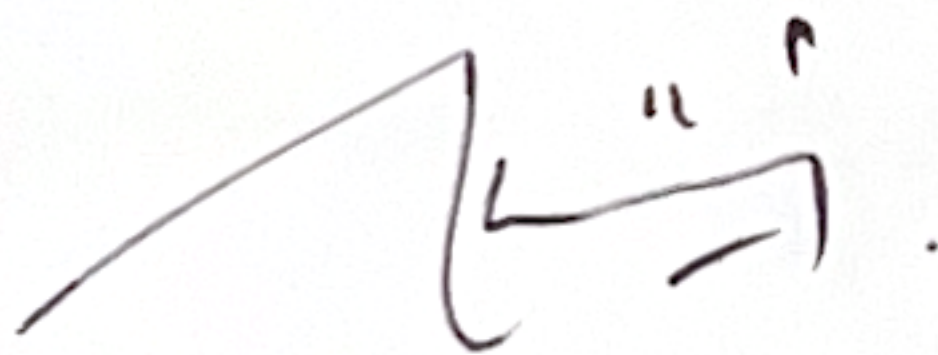
KLASIFIKASI PRIORITAS PERBAIKAN JALAN DI KABUPATEN MALANG MENGGUNAKAN ALGORITMA RANDOM FOREST

SKRIPSI

Oleh :
IZZA SYAHRI MUHARRAM
NIM. 220605110073

Telah Diperiksa dan Disetujui untuk Diuji:
Tanggal: 25 Juni 2026

Pembimbing I,



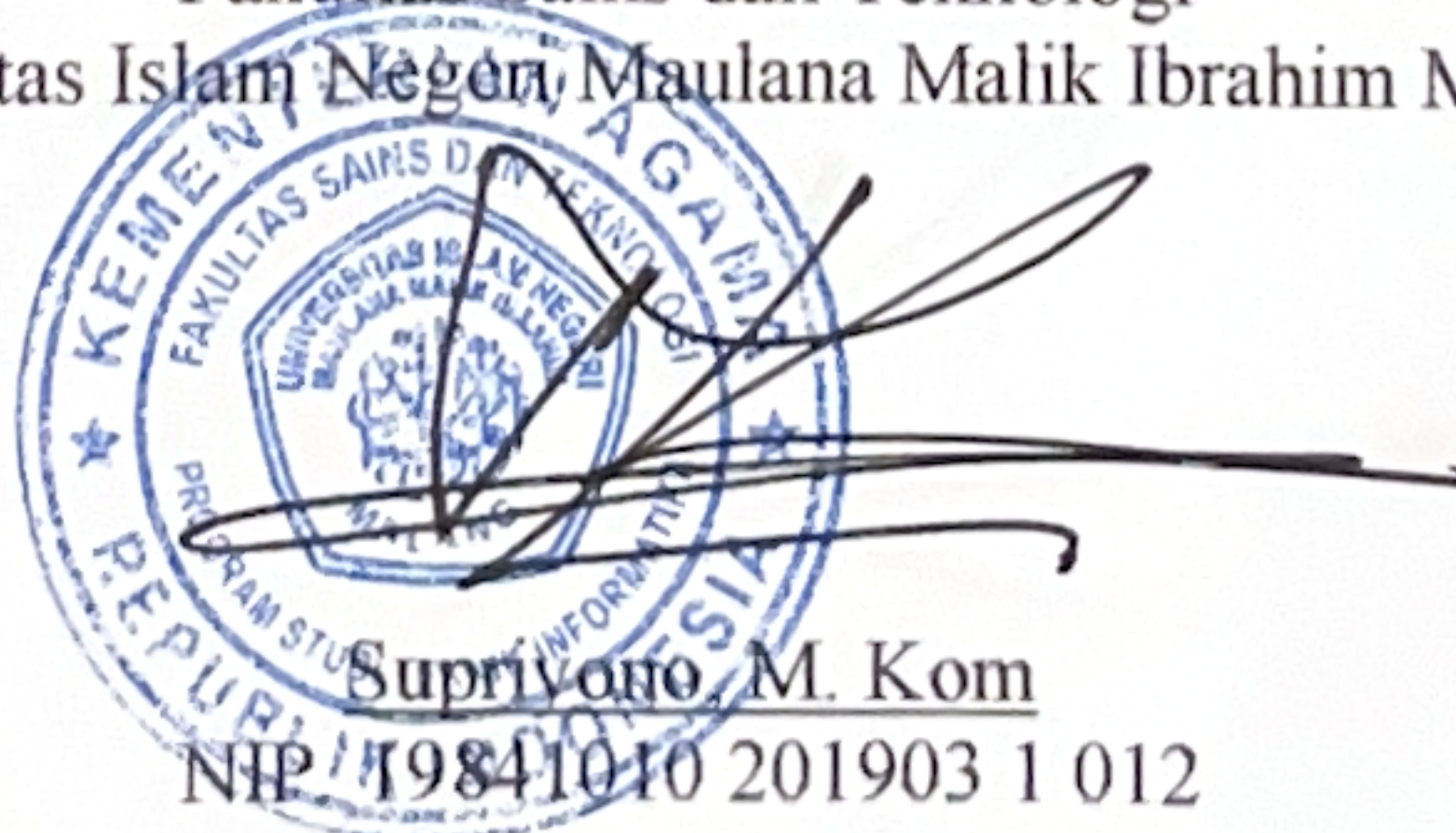
Ashri Shabrina Afrah, M.T
NIP. 19900430 202012 2 003

Pembimbing II,



Okta Qomaruddin Aziz, M.Kom
NIP. 19911019 201903 1 013

Mengetahui,
Ketua Program Studi Teknik Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang



Supriyono, M. Kom
NIP. 19841010 201903 1 012

HALAMAN PENGESAHAN

KLASIFIKASI PRIORITAS PERBAIKAN JALAN DI KABUPATEN MALANG MENGGUNAKAN ALGORITMA RANDOM FOREST

SKRIPSI

Oleh :

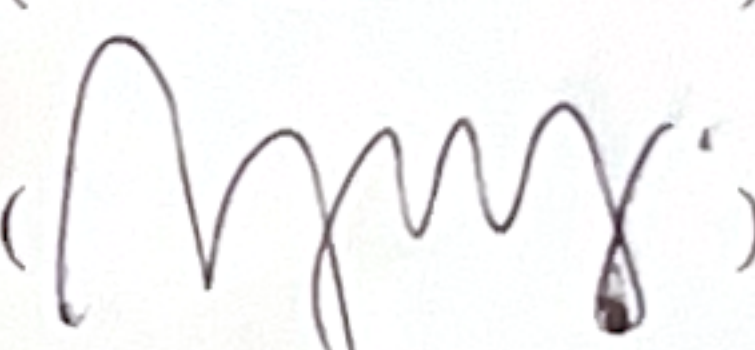
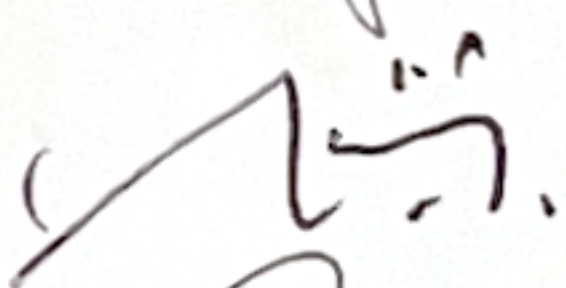
IZZA SYAHRI MUHARRAM

NIM. 220605110073

Telah Dipertahankan di Depan Dewan Penguji Skripsi
dan Dinyatakan Diterima Sebagai Salah Satu Persyaratan
Untuk Memperoleh Gelar Sarjana Komputer (S.Kom)
Tanggal: 25 Juni 2026

Susunan Dewan Penguji

Ketua Penguji	: <u>Syahiduz Zaman, M.Kom</u> NIP. 19700502 200501 1 005
Anggota Penguji I	: <u>Dr. Agung Teguh Wibowo Almais, M.T</u> NIP. 198603012 02321 1 016
Anggota Penguji II	: <u>Ashri Shabrina Afrah, M.T</u> NIP. 19900430 202012 2 003
Anggota Penguji III	: <u>Okta Qomaruddin Aziz, M.Kom</u> NIP. 19911019 201903 1 013

()
()
()
()

Mengetahui dan Mengesahkan,
Ketua Program Studi Teknik Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang



Supriyono, M. Kom
NIP. 19841010 201903 1 012

PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini:

Nama : Izza Syahri Muharram

NIM : 220605110073

Fakultas / Jurusan : Sains dan Teknologi / Teknik Informatika

Judul Skripsi : Klasifikasi Prioritas Perbaikan Jalan Di Kabupaten Malang
Menggunakan Algoritma *Random Forest*

Menyatakan dengan sebenarnya bahwa Skripsi yang saya tulis ini benar-benar merupakan hasil karya saya sendiri, bukan merupakan pengambil alihan data, tulisan, atau pikiran orang lain yang saya akui sebagai hasil tulisan atau pikiran saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar pustaka.

Apabila dikemudian hari terbukti atau dapat dibuktikan skripsi ini merupakan hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 25 Juni 2026
Yang membuat pernyataan,



Izza Syahri Muharram
NIM.220605110073

MOTTO

... Start now, perfection will follow....

HALAMAN PERSEMBAHAN

Segala puji bagi Allah Subhanahu Wa Ta'ala atas rahmat, karunia, dan hidayah-

Nya, sehingga penulis dapat menuntaskan skripsi ini.

Dengan penuh rasa syukur, penulis persembahkan kepada:

Orang tua saya,

Yang senantiasa memberikan doa, semangat, dan kasih sayang yang tiada henti
dan tak terhitung nilainya.

Dosen pembimbing dan penguji,

Atas ilmu dan bimbingan yang diberikan selama proses penyusunan skripsi.

Teman-teman seperjuangan,

Atas kebersamaan, tenaga, dan waktu selama masa perkuliahan.

KATA PENGANTAR

Puji dan syukur ke hadirat Allah SWT atas limpahan rahmat, taufik, dan hidayah-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul “Klasifikasi Prioritas Perbaikan Jalan Menggunakan *Random Forest* di Kabupaten Malang” sebagai salah satu syarat untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang. Shalawat serta salam semoga senantiasa tercurah kepada Nabi Muhammad SAW beserta keluarga dan para sahabatnya.

Penyusunan skripsi ini merupakan proses akademik yang tidak terlepas dari dukungan, arahan, serta kontribusi berbagai pihak. Oleh karena itu, pada kesempatan ini penulis menyampaikan penghargaan dan terima kasih yang sebesar-besarnya kepada:

1. Prof. Dr. Hj. Ilfi Nur Diana, M.Si., CAHRM., CRMP selaku rektor Universitas Islam Negeri Maulana Malik Ibrahim Malang atas kepemimpinan dan kebijakan yang mendukung terselenggaranya proses pendidikan dengan baik.
2. Dr. H. Agus Mulyono, M.Si, selaku Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang atas fasilitas dan dukungan akademik yang diberikan.
3. Supriyono, M.Kom, selaku Ketua Program Studi Teknik Informatika Universitas Islam Negeri Maulana Malik Ibrahim Malang atas arahan dan kebijakan akademik selama masa studi penulis.

4. Ashri Shabrina Afrah, M.T, selaku Dosen Pembimbing I yang telah memberikan bimbingan, masukan, koreksi, serta motivasi secara konsisten dalam proses penyusunan skripsi ini.
5. Okta Qomaruddin Aziz, M.Kom, selaku Dosen Pembimbing II yang telah memberikan bimbingan, masukan, koreksi, serta motivasi secara konsisten dalam proses penyusunan skripsi ini.
6. Syahiduz Zaman, M.Kom, selaku Ketua Penguji yang telah memberikan bimbingan, masukan, koreksi, serta motivasi secara konsisten sejak tahap seminar proposal hingga sidang skripsi.
7. Dr. Teguh Agung Wibowo Almais, M.T, selaku Anggota Penguji yang telah memberikan bimbingan, masukan, koreksi, serta motivasi sejak tahap seminar proposal hingga sidang skripsi.
8. Nia Faricha S, Si., selaku Admin Program Studi Teknik Informatika yang selalu sabar memberikan informasi, membantu, dan memberikan arahan selama perkuliahan dan proses penulisan skripsi ini.
9. Seluruh dosen dan tenaga kependidikan di lingkungan Fakultas Sains dan Teknologi yang telah memberikan ilmu, pelayanan, serta pengalaman akademik yang berharga.
10. Kedua orang tua tercinta, Mama Nurul dan Ayah Hendro, yang selalu menjadi sumber doa, semangat, dan kasih sayang yang tak pernah padam. Segala pengorbanan, dukungan, dan ketulusan cinta yang telah diberikan menjadi kekuatan utama bagi penulis dalam menyelesaikan skripsi ini.

11. Sahabat terbaik sejak awal perkuliahan Sania Nabila Salsabila. Terima kasih atas kebersamaan, dukungan, dan segala tawa maupun tangis yang telah kita lalui bersama sejak semester pertama. Persahabatan kalian menjadi salah satu bagian terindah dalam perjalanan penulis selama menempuh studi ini.
12. Teman-teman saya, Zulfa, Sella, Faza, yang telah menjadi bagian dari perjalanan penulis selama masa perkuliahan. Terima kasih atas kebersamaan, kerja sama, serta dukungan yang diberikan, baik dalam kegiatan akademik maupun dalam berbagai momen kebersamaan lainnya yang membuat perjalanan studi ini menjadi lebih berkesan.
13. Seluruh teman-teman Angkatan 2022 Teknik Informatika yang telah banyak membantu, mendukung, dan memotivasi dalam menyelesaikan penyusunan skripsi ini.
14. Teman saya, Noval yang selalu ada ketika penyusunan penelitian ini.

Penulis menyadari bahwa skripsi ini masih memiliki keterbatasan, baik dari segi substansi maupun penyajiannya. Oleh karena itu, kritik dan saran yang konstruktif sangat diharapkan demi penyempurnaan karya ilmiah ini di masa mendatang. Semoga skripsi ini dapat memberikan kontribusi bagi pengembangan ilmu pengetahuan, khususnya dalam bidang rekayasa perangkat lunak dan penerapan machine learning.

Malang, 19 Juni 2026

Penulis

DAFTAR ISI

HALAMAN PERSETUJUAN	iii
HALAMAN PENGESAHAN	iv
PERNYATAAN KEASLIAN TULISAN	v
MOTTO	vi
HALAMAN PERSEMBAHAN	vii
KATA PENGANTAR.....	viii
DAFTAR ISI.....	xi
DAFTAR GAMBAR.....	xii
DAFTAR TABEL	xiii
ABSTRAK	xiv
ABSTRACT	xv
مستخلص البحث	xvi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	8
1.3 Batasan Masalah	8
1.4 Tujuan Penelitian	9
1.5 Manfaat Penelitian	9
BAB II STUDI PUSTAKA	10
2.1 Penelitian Terkait	10
2.2 Prioritas Perbaikan Jalan	22
2.3 Klasifikasi	24
2.4 Preprocessing	25
2.5 Random Forest	28
2.6 Evaluasi Model	33
BAB III DESAIN DAN IMPLEMENTASI.....	38
3.1 Desain Penelitian	38
3.2 Pengumpulan Data	38
3.3 Design Sistem	42
3.4 Implementasi <i>Random Forest</i>	48
3.5 Evaluasi Model	58
3.6 Skenario Uji Coba	62
BAB IV HASIL DAN PEMBAHASAN	64
4.1 Hasil Penerapan Model	64
4.2 Evaluasi Performa Model	69
4.3 Hasil <i>Feature Importance</i>	84
4.4 Analisis Hasil Model.....	87
BAB V PENUTUP	107
5.1 Kesimpulan	107
5.2 Saran	108
DAFTAR PUSTAKA	110
LAMPIRAN.....	114

DAFTAR GAMBAR

Gambar 2. 1 Algoritma Random Forest.....	31
Gambar 3.1 Alur Penelitian.....	38
Gambar 3.2 Diagram Tahapan Design Sistem.....	42
Gambar 3. 3 Contoh Visualisasi Decission Tree	56
Gambar 4. 1 Distribusi Data	67
Gambar 4. 2 Hasil Akurasi Random Forest Berdasarkan n- estimators	71
Gambar 4. 3 Hasil Evaluasi Tiap Skenario Uji Coba.....	81
Gambar 4. 4 Perbandingan Prediksi dan Aktual	92
Gambar 4. 5 UX n_estimator	94
Gambar 4. 6 Antarmuka Sistem	95

DAFTAR TABEL

Tabel 2. 1 Penelitian Terkait	18
Tabel 3. 1 Contoh Dataset Tingkat Kerusakan Jalan	40
Tabel 3. 2 Contoh Fitur Untuk Normalisasi.....	45
Tabel 3. 3 Contoh Hasil Normalisasi	46
Tabel 3. 4 Contoh Bootstrap Sampling	49
Tabel 3. 5 Contoh Pemilihan Fitur Secara Acak.....	50
Tabel 3. 6 Contoh Data Untuk Split Terbaik	51
Tabel 3. 7 Contoh Data Untuk Membuat Decision Tree	53
Tabel 3. 8 Contoh Hasil Decission Tree	55
Tabel 3. 9 Contoh Dari Lima Pohon Keputusan	57
Tabel 3. 10 Contoh Hasil Agregrasi Prediksi.....	57
Tabel 3. 11 Contoh Confussion Matrix.....	59
Tabel 4. 1 Pengelompokan Kategori Prioritas	65
Tabel 4. 2 Skenario Uji Coba n-estimators 2	74
Tabel 4. 3 Skenario Uji Coba n-estimators 4	75
Tabel 4. 4 Skenario Uji Coba n-estimators 6	75
Tabel 4. 5 Skenario Uji Coba n-estimators 8	76
Tabel 4. 6 Skenario Uji Coba n-estimators 10	76
Tabel 4. 7 Skenario Uji Coba n-estimators 12	77
Tabel 4. 8 Skenario Uji Coba n-estimators 14	77
Tabel 4. 9 Skenario Uji Coba n-estimators 16	78
Tabel 4. 10 Skenario Uji Coba n-estimators 18	78
Tabel 4. 11 Skenario Uji Coba n-estimators 20	78
Tabel 4. 12 Skenario Uji Coba Tiap n_estimators	79
Tabel 4. 13 Fitur yang Sering Muncul	85
Tabel 4. 14 Hasil Prediksi Prioritas	90

ABSTRAK

Muharram, Izza Syahri. 2026. **Klasifikasi Prioritas Perbaikan Jalan Di Kabupaten Malang Menggunakan Algoritma Random Forest**. Skripsi. Jurusan Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing: (I) Ashri Shabrina Afrah, M.T (II) Okta Qomaruddin Aziz, M.Kom.

Kata kunci: *Random Forest*, klasifikasi, prioritas jalan, LHR, machine learning.

Kerusakan jalan merupakan salah satu permasalahan infrastruktur yang dapat menghambat mobilitas masyarakat, meningkatkan risiko kecelakaan, serta menurunkan efektivitas aktivitas ekonomi. Di Kabupaten Malang, keterbatasan anggaran dan sumber daya menyebabkan tidak seluruh ruas jalan yang mengalami kerusakan dapat diperbaiki secara bersamaan, sehingga diperlukan penentuan prioritas penanganan jalan yang tepat dan objektif. Penelitian ini bertujuan untuk mengklasifikasikan tingkat prioritas penanganan jalan berdasarkan kondisi jalan dan volume lalu lintas harian (LHR) menggunakan algoritma Random Forest. Dataset yang digunakan diperoleh dari Dinas Pekerjaan Umum dan Penataan Ruang (PUPR) yang mencakup atribut panjang jalan, kondisi jalan (baik, sedang, rusak ringan, rusak berat), serta nilai LHR yang merepresentasikan tingkat kepadatan lalu lintas. Metode Random Forest diterapkan dengan variasi parameter *n_estimators* untuk mengetahui pengaruh jumlah pohon terhadap kinerja model. Hasil penelitian menunjukkan bahwa peningkatan jumlah pohon mampu meningkatkan performa model secara signifikan hingga mencapai nilai optimal, dengan accuracy, precision, recall, dan F1-score mendekati atau mencapai 1,0. Analisis confusion matrix juga menunjukkan bahwa sebagian besar data berhasil diklasifikasikan dengan benar dengan tingkat kesalahan yang sangat minimal. Selain itu, hasil analisis feature importance menunjukkan bahwa kondisi jalan merupakan faktor yang paling berpengaruh dalam menentukan prioritas penanganan jalan, diikuti oleh panjang jalan dan nilai LHR. Meskipun demikian, distribusi data yang tidak seimbang serta jumlah data yang terbatas menjadi faktor yang perlu diperhatikan dalam interpretasi hasil. Secara keseluruhan, algoritma Random Forest terbukti efektif dalam mengklasifikasikan prioritas penanganan jalan dan dapat dijadikan sebagai dasar dalam mendukung pengambilan keputusan terkait perbaikan infrastruktur jalan di Kabupaten Malang.

ABSTRACT

Muharram, Izza Syahri. 2026. **Klasifikasi Prioritas Perbaikan Jalan Di Kabupaten Malang Menggunakan Algoritma Random Forest**. Skripsi. Jurusan Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing: (I) Ashri Shabrina Afrah, M.T (II) Okta Qomaruddin Aziz, M.Kom.

Road damage is one of the infrastructure problems that can hinder community mobility, increase the risk of traffic accidents, and reduce the effectiveness of economic activities. In Malang Regency, limited budgets and resources make it impossible to repair all damaged road segments simultaneously, thereby requiring an appropriate and objective road maintenance priority determination process. This study aims to classify road maintenance priority levels based on road conditions and Average Daily Traffic (ADT) using the Random Forest algorithm. The dataset was obtained from the Department of Public Works and Spatial Planning (PUPR) and includes road length, road condition (good, moderate, lightly damaged, and heavily damaged), and ADT values representing traffic density levels. The Random Forest method was implemented with various $n_estimators$ parameter values to analyze the effect of the number of trees on model performance. The results indicate that increasing the number of trees significantly improves model performance, achieving optimal values with accuracy, precision, recall, and F1-score close to or equal to 1.0. Confusion matrix analysis also shows that most data were correctly classified with very minimal classification errors. Furthermore, feature importance analysis reveals that road condition is the most influential factor in determining road maintenance priorities, followed by road length and ADT values. However, data imbalance and the limited size of the dataset should be considered when interpreting the results. Overall, the Random Forest algorithm proved effective in classifying road maintenance priorities and can be utilized as a decision-support tool for road infrastructure management in Malang Regency.

Key words: Random Forest, classification, road priority, traffic volume, machine learning.

مستخلص البحث

محرم، عزة شحري. 2026. تصنيف أولويات صيانة الطرق في منطقة مالانغ باستخدام خوارزمية الغابة العشوائية رسالة جامعية. قسم تقنية المعلومات، كلية العلوم والتكنولوجيا، جامعة مولانا مالك إبراهيم (Random Forest) الإسلامية الحكومية مالانغ. المشرفان: (الأول) أشري شيرينا أفراح، ماجستير هندسة، (الثاني) أوكتا قمر الدين عزيز، ماجستير.

تُعدّ أضرار الطرق من المشكلات الأساسية في البنية التحتية التي قد تعيق حركة المجتمع، وتزيد من مخاطر الحوادث المرورية، وتؤثر سلبيًا على الأنشطة الاقتصادية. وفي محافظة مالانج، تؤدي محدودية الميزانية والموارد المتاحة إلى عدم إمكانية إصلاح جميع الطرق المتضررة في الوقت نفسه، مما يستلزم تحديد أولويات صيانة الطرق بطريقة دقيقة وموضوعية. تهدف هذه الدراسة إلى تصنيف مستويات أولوية *Random Forest* باستخدام خوارزمية الغابة العشوائية (LHR) صيانة الطرق بناءً على حالة الطريق وحجم المرور اليومي المتوسط ، وتشمل طول الطريق، وحالة الطريق (PUPR) تم الحصول على البيانات من دائرة الأشغال العامة والتخطيط العمراني (*Forest* جيد، متوسط، تضرر خفيف، تضرر شديد)، بالإضافة إلى قيم حجم المرور اليومي المتوسط التي تمثل كثافة حركة المرور. تم تطبيق لدراسة تأثير عدد الأشجار على أداء النموذج. أظهرت النتائج $n_estimators$ خوارزمية الغابة العشوائية مع عدة قيم للمعامل ، والدقة أن زيادة عدد الأشجار تسهم بشكل ملحوظ في تحسين أداء النموذج، حيث حقق النموذج قيمًا مثالية في مقاييس الدقة (*Confusion*) ، والتي اقتربت من 1.0 أو وصلت إليها. كما أظهر تحليل مصفوفة الالتباس ، ودرجة والاستدعاء الإيجابية أن معظم البيانات صُنفت بشكل صحيح مع وجود نسبة ضئيلة جدًا من الأخطاء. بالإضافة إلى ذلك، بين تحليل أهمية (*Matrix* الخصائص أن حالة الطريق تُعد العامل الأكثر تأثيرًا في تحديد أولوية الصيانة، تليها طول الطريق وقيم حجم المرور اليومي المتوسط. ومع ذلك، ينبغي أخذ عدم توازن البيانات ومحدودية حجم مجموعة البيانات بعين الاعتبار عند تفسير النتائج. وبشكل عام، أثبتت خوارزمية الغابة العشوائية فعاليتها في تصنيف أولويات صيانة الطرق، ويمكن الاستفادة منها كأداة داعمة لاتخاذ القرار في إدارة البنية التحتية للطرق في محافظة مالانج.

الغابة العشوائية، التصنيف، أولوية الطرق، حجم حركة المرور، تعلم الآلة :الكلمات المفتاحية

BAB I

PENDAHULUAN

1.1 Latar Belakang

Sebagai salah satu wilayah di Provinsi Jawa Timur, Kabupaten Malang didukung oleh infrastruktur jaringan jalan yang tergolong luas dan tersebar di berbagai kecamatan. Berdasarkan data Pemerintah Kabupaten Malang, panjang jalan kabupaten mencapai lebih dari 1.668 kilometer yang berfungsi sebagai sarana penghubung antarwilayah, pusat perekonomian, kawasan pendidikan, serta fasilitas pelayanan publik. Keberadaan jalan yang baik mempunyai peran krusial guna menjamin kelancaran pergerakan penduduk sekaligus efisiensi penyaluran komoditas barang dan jasa. Oleh karena itu, kondisi jalan yang terpelihara dengan baik merupakan salah satu fokus utama yang mesti diprioritaskan dalam program pembangunan daerah.

Seiring dengan perkembangan wilayah dan pertumbuhan penduduk, kebutuhan mobilitas masyarakat di Kabupaten Malang juga terus meningkat. Peningkatan aktivitas ekonomi, perdagangan, pendidikan, serta sektor pariwisata menyebabkan intensitas penggunaan jalan semakin tinggi dari tahun ke tahun (Fitri et al., 2024). Selain itu, jumlah kendaraan bermotor yang terus bertambah berdampak pada peningkatan volume beban jalan di sejumlah rute. Fenomena ini memicu kebutuhan pemeliharaan dan perbaikan jalan menjadi semakin penting guna menjaga kualitas pelayanan infrastruktur transportasi bagi masyarakat (Az-Zahriyah et al., 2025).

Meskipun demikian, berbagai ruas jalan di Kabupaten Malang masih menunjukkan derajat kerusakan yang bervariasi. Degradasi ini mencakup rincian kondisi dari yang masih berkategori baik, rusak ringan, rusak sedang, hingga rusak berat. Faktor pemicu kerusakan ini sangat beragam, di antaranya curah hujan yang tinggi, usia perkerasan jalan, kualitas konstruksi, buruknya saluran pembuangan air (drainase), hingga muatan armada transportasi yang melampaui batas tonase jalan (Fitri et al., 2024). Kerusakan tersebut dapat berupa retak-retak pada permukaan jalan, lubang, amblas, maupun deformasi lainnya yang berpotensi menurunkan aspek kenyamanan sekaligus mengancam keselamatan para pengendara yang melintas.

Tidak stabilnya kondisi jalan tidak hanya mengakibatkan pada menurunnya kualitas pelayanan transportasi, tetapi juga dapat menimbulkan berbagai kerugian bagi masyarakat. Akses jalan yang tidak layak ini rawan memicu insiden kecelakaan di jalan raya, memperpanjang waktu perjalanan, meningkatkan biaya operasional kendaraan, serta menghambat aktivitas ekonomi masyarakat (Putri & Susanto, 2025). Apabila kondisi tersebut tidak segera ditangani, maka dapat memengaruhi produktivitas masyarakat dan mengurangi efektivitas pembangunan daerah secara keseluruhan.

Dalam pelaksanaannya, pemerintah daerah sering menghadapi keterbatasan anggaran, tenaga kerja, serta sumber daya dalam melakukan perbaikan jalan. Kondisi ini menyebabkan program perbaikan terhadap seluruh titik jalan yang rusak tidak mungkin dilaksanakan secara serentak dalam satu waktu. Atas dasar situasi ini, diperlukan sebuah rumusan formula atau metode khusus guna memetakan

segmentasi jalan mana saja yang wajib didahulukan supaya pemanfaatan dana belanja daerah beserta seluruh potensi yang ada bisa direalisasikan secara lebih efisien serta terarah.

Penentuan prioritas perbaikan jalan pada praktiknya umumnya dilakukan melalui survei lapangan dan penilaian langsung terhadap kondisi jalan. Metode tersebut mampu memberikan gambaran mengenai kondisi aktual di lapangan, akan tetapi mekanisme tersebut menguras durasi pengerjaan yang relatif panjang sekaligus menyerap anggaran yang cukup besar. Di sisi lain, hasil penilaian juga berpotensi dipengaruhi oleh subjektivitas penilai sehingga dapat menimbulkan perbedaan dalam menentukan tingkat prioritas suatu ruas jalan (Fitri et al., 2024). Kondisi tersebut menunjukkan bahwa diperlukan penerapan sebuah metode yang lebih netral serta bertumpu pada validitas data demi menyokong efektivitas penarikan kesimpulan atau kebijakan.

Selain kondisi kerusakan jalan, terdapat faktor lain yang perlu diperhatikan dalam menentukan prioritas perbaikan, yaitu lalu lintas harian rata-rata (LHR). Ruas jalan yang memiliki tingkat kerusakan yang sama belum tentu memiliki tingkat prioritas yang sama apabila volume kendaraan yang melintasinya berbeda. Jalan yang dilalui oleh kendaraan dalam jumlah besar memiliki tingkat urgensi perbaikan yang lebih tinggi dibandingkan jalur yang sepi kendaraan, mengingat kendala fisik pada rute yang padat secara otomatis bakal memengaruhi kenyamanan cakupan mobilitas masyarakat yang jauh lebih luas (Putri & Susanto, 2025).

Di samping itu, panjang ruas jalan juga menjadi faktor yang perlu dipertimbangkan dalam proses penentuan prioritas. Ruas jalan yang lebih panjang

umumnya membutuhkan biaya pemeliharaan yang lebih besar dibandingkan ruas akses jalan yang bentangannya cenderung pendek. Maka dari itu, penentuan prioritas perbaikan jalan tidak bisa hanya berpedoman pada satu indikator saja, tetapi diperlukan peninjauan menyeluruh terhadap beragam aspek karakteristik jalan secara simultan, sehingga kebijakan yang diambil bisa lebih akurat sekaligus relevan dengan fakta riil di lapangan.

Konsep penentuan prioritas berdasarkan kondisi dan kebutuhan tersebut selaras dengan tuntunan syariat Islam. Dalam pandangan Islam, terdapat urgensi besar untuk memposisikan segala sesuatu secara proporsional sesuai porsi dan kondisinya masing-masing, sebagaimana diisyaratkan oleh Allah SWT melalui QS. Al-Isra' ayat 84:

قُلْ كُلٌّ يَعْمَلُ عَلَىٰ شَاكِلَتِهِ ۖ فَرَبُّكُمْ أَعْلَمُ بِمَنْ هُوَ أَهْدَىٰ سَبِيلًا

"Katakanlah (Nabi Muhammad), 'Setiap orang berbuat sesuai dengan pembawaannya masing-masing.' Maka Tuhanmu lebih mengetahui siapa yang lebih benar jalannya." (QS. Al-Isra': 84)

Kandungan ayat di atas memberikan pemahaman bahwa seluruh makhluk diciptakan dengan keadaan, karakteristik, dan kondisi yang berbeda-beda. Jika ditarik ke dalam ruang lingkup kajian ini, tiap-tiap jalur jalan mempunyai spesifikasi yang berlainan, baik dari segi derajat kerusakan fisik, bentang panjangnya, hingga kapasitas beban lalu lintasnya. Konsekuensinya, setiap ruas jalan memerlukan penanganan dan tingkat prioritas yang berbeda sesuai dengan tingkat kebutuhannya.

Islam juga mengajarkan pentingnya memberikan kemanfaatan dan kemudahan bagi hajat hidup orang banyak. Landasan teologis mengenai hal ini dapat dirujuk pada firman Allah SWT dalam QS. Al-Maidah ayat 2:

وَتَعَاوَنُوا عَلَى الْبِرِّ وَالتَّقْوَىٰ وَلَا تَعَاوَنُوا عَلَى الْإِثْمِ وَالْعُدْوَانِ وَاتَّقُوا اللَّهَ إِنَّ اللَّهَ شَدِيدُ الْعِقَابِ

Artinya: *"Dan tolong-menolonglah kamu dalam (mengerjakan) kebajikan dan takwa, dan jangan tolong-menolong dalam berbuat dosa dan permusuhan."* (QS. Al-Maidah: 2)

Isyarat dari ayat tersebut menegaskan bahwa segala bentuk kontribusi kepada masyarakat merupakan bentuk kebajikan yang dianjurkan dalam Islam. Pemeliharaan sarana jalan merepresentasikan wujud nyata dari pelayanan publik dengan daya jangkau aktivitas ekonomi, pendidikan, kesehatan, maupun kegiatan sosial masyarakat (Erlinawati, 2024).

Selain itu, Allah SWT juga berfirman dalam QS. Al-Baqarah ayat 185:

يُرِيدُ اللَّهُ بِكُمُ الْيُسْرَ وَلَا يُرِيدُ بِكُمُ الْعُسْرَ

Artinya: *"Allah menghendaki kemudahan bagimu dan tidak menghendaki kesukaran bagimu."* (QS. Al-Baqarah: 185)

Ayat di atas menggarisbawahi komitmen ajaran Islam yang senantiasa menghendaki adanya kelonggaran dan kemudahan bagi hajat hidup manusia. Dalam konteks pembangunan infrastruktur, penyediaan jalan yang layak dan aman merupakan salah satu bentuk upaya demi mempermudah mobilitas masyarakat dalam menuntaskan agenda keseharian mereka (Abdur Rahim et al., 2024). Oleh karena itu, penentuan prioritas perbaikan jalan yang tepat dapat menjadi bagian dari implementasi nilai-nilai kemaslahatan dalam Islam.

Rasulullah SAW juga bersabda:

مَنْ نَفَسَ عَنْ مُؤْمِنٍ كُرْبَةً مِنْ كُرْبِ الدُّنْيَا نَفَسَ اللَّهُ عَنْهُ كُرْبَةً مِنْ كُرْبِ يَوْمِ الْقِيَامَةِ

"Barang siapa menghilangkan satu kesulitan dunia dari seorang mukmin, maka Allah akan menghilangkan darinya satu kesulitan pada hari kiamat." (HR. Muslim)

Hadis tersebut menunjukkan pentingnya membantu mengurangi kesulitan yang dihadapi masyarakat. Jalan yang rusak sering kali menjadi penyebab munculnya berbagai kesulitan, mulai dari hambatan transportasi hingga meningkatnya risiko kecelakaan (Tafsir Al-Muyasar, n.d.). Oleh karena itu, upaya menentukan prioritas perbaikan jalan secara tepat dapat menjadi salah satu bentuk ikhtiar dalam memberikan kemudahan dan kemanfaatan bagi masyarakat luas.

Perkembangan teknologi informasi saat ini memungkinkan penerapan pendekatan berbasis data untuk mendukung proses pengambilan keputusan. Salah satu teknologi yang banyak digunakan ialah teknologi *machine learning*, sebuah sistem komputasi yang dirancang untuk mengenali tren atau karakteristik dari rekam data masa lalu guna melakukan estimasi maupun pengelompokan secara mandiri. Penerapan kecerdasan buatan ini dinilai mampu mereduksi bias dalam proses penentuan prioritas serta meningkatkan kecepatan dan konsistensi dalam pengolahan data (Kim et al., 2020).

Di antara sekian banyak opsi algoritma pemelajaran mesin untuk urusan kategorisasi data, *Random Forest* merupakan model yang sangat populer. Sistem ini beroperasi lewat perancangan sekumpulan pohon keputusan (*decision tree*) secara masif, lalu mengintegrasikan proyeksi dari tiap-tiap komponen pohon tersebut demi melahirkan sebuah konklusi akhir dengan tingkat akurasi yang lebih

optimal (Rahman et al., 2025). Keunggulan utama dari *Random Forest* terletak pada reliabilitasnya saat mengolah himpunan data berdimensi besar (multi-variabel), ketahanannya dalam meminimalkan gejala *overfitting*, serta konsistensinya dalam menyajikan hasil pemetaan yang presisi di beragam ranah studi ilmiah.

Merujuk pada sejumlah literatur terdahulu, metode Random Forest ini secara konsisten mendemonstrasikan efisiensi dan keandalan yang sangat mumpuni untuk memecahkan problem penataan atau pengelompokan data. Keberhasilan algoritma tersebut dalam mengolah data yang kompleks membuktikan adanya peluang besar bagi implementasi metode *Random Forest* dalam memetakan urgensi pemeliharaan prasarana jalan (Rahman et al., 2025). Dengan memanfaatkan data keadaan fisik jalur, bentang panjangnya, serta volume lalu lintas harian rata-rata (LHR), algoritma ini diharapkan mampu menghasilkan klasifikasi prioritas perbaikan jalan ke dalam level tinggi, sedang, maupun rendah secara lebih netral dan terukur.

Berdasarkan uraian diatas, penelitian mengenai klasifikasi prioritas perbaikan jalan di Kabupaten Malang menggunakan algoritma *Random Forest* menjadi penting untuk dilakukan. Penelitian ini ditargetkan mampu mempermudah jajaran pemerintah daerah untuk merumuskan kebijakan pemeliharaan jalan secara akurat, efektif, dan berbasis data. Selain memberikan kontribusi dalam bidang machine learning, penelitian ini juga sejalan dengan nilai-nilai Islam yang menekankan keadilan, kemaslahatan, serta urgensi memfasilitasi hajat hidup orang banyak melalui tata kelola layanan publik yang prima.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, maka rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Seberapa baik tingkat akurasi dan performa Algoritma *Random Forest* dalam mengklasifikasikan prioritas perbaikan jalan di Kabupaten Malang berdasarkan data kondisi jalan dan LHR?
2. Fitur apa yang paling berpengaruh dalam proses klasifikasi prioritas jalan pada model *Random Forest*, serta bagaimana kontribusi masing-masing fitur dalam membentuk keputusan model?

1.3 Batasan Masalah

Untuk menjaga fokus penelitian agar lebih terarah, maka penelitian ini memiliki batasan masalah sebagai berikut:

1. Data yang digunakan terbatas pada data kondisi jalan di Kabupaten Malang yang diperoleh secara resmi dari Dinas Pekerjaan Umum dan Penataan Ruang (PUPR) Kabupaten Malang melalui surat izin pengambilan data, dan tidak tersedia secara publik.
2. Metode klasifikasi yang digunakan dibatasi pada Algoritma *Random Forest* tanpa melakukan perbandingan dengan algoritma *machine learning* lain.
3. Hasil penelitian difokuskan pada klasifikasi prioritas perbaikan jalan ke dalam beberapa kategori (misalnya: tinggi, sedang, rendah) dan tidak mencakup estimasi biaya atau penjadwalan perbaikan.

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah ditetapkan, maka tujuan dari penelitian ini adalah sebagai berikut:

1. Untuk mengetahui tingkat akurasi dan performa Algoritma *Random Forest* dalam melakukan klasifikasi prioritas perbaikan jalan di Kabupaten Malang berdasarkan data kondisi jalan dan Lalu Lintas Harian Rata-Rata (LHR).
2. Untuk mengidentifikasi fitur atau faktor yang paling berpengaruh dalam menentukan prioritas perbaikan jalan berdasarkan hasil analisis Algoritma *Random Forest*.

1.5 Manfaat Penelitian

- a. Menambah wawasan dan literatur di bidang ilmu komputer serta penerapan metode *machine learning* dalam menyelesaikan permasalahan nyata.
- b. Memberikan gambaran penerapan metode klasifikasi yang dapat digunakan sebagai solusi dalam bidang perencanaan dan pengelolaan infrastruktur.
- c. Memberikan manfaat bagi masyarakat luas melalui peningkatan efektivitas pelayanan publik.

BAB II

STUDI PUSTAKA

Pada bab ini peneliti menguraikan sejumlah studi terdahulu yang relevan untuk dijadikan sebagai landasan komparasi dan telaah kritis. Di samping itu, dipaparkan pula kerangka teori fundamental yang mendasari analisis penentuan skala prioritas pemeliharaan jalan melalui implementasi model *Random Forest*.

2.1 Penelitian Terkait

Penelitian yang dilakukan oleh Rahman et al. (2025) membahas penerapan metode machine learning untuk mendukung klasifikasi kerusakan jalan berdasarkan data lapangan di Kabupaten Bantaeng. Pada eksperimen tersebut, skema algoritma K-Means diterapkan untuk memetakan kualitas fisik jalan menjadi empat tingkatan, yang meliputi kelas *Baik-Baik Saja*, *Rusak Kecil*, *Rusak Sedang*, hingga *Rusak Parah*. Hasil klasterisasi menunjukkan bahwa kategori *Baik-Baik Saja* mendominasi dengan jumlah 185 ruas jalan, sementara kategori lainnya memiliki jumlah yang lebih sedikit, yakni 59 ruas *Rusak Kecil*, 24 ruas *Rusak Parah*, dan 21 ruas *Rusak Sedang*. Distribusi data ini kemudian dimanfaatkan sebagai label tambahan dalam proses klasifikasi menggunakan Algoritma *Random Forest*, di mana kolaborasi tersebut menorehkan tingkat akurasi total di angka 95%. Perolehan skor precision paling optimal menyentuh 100% untuk kelompok jalan berkategori Rusak Parah serta Rusak Sedang, walaupun parameter recall di sejumlah kelas lain cenderung fluktuatif akibat adanya ketimpangan sebaran data sampel. Output dari studi tersebut membuktikan bahwa integrasi model K-Means bersama Random

Forest tergolong andal untuk mendeteksi derajat kerusakan fasilitas jalan, kendati demikian penelitian tersebut belum mempertimbangkan faktor lalu lintas serta belum mengembangkan sistem implementatif yang berorientasi pada penentuan prioritas perbaikan jalan.

Pendekatan klasifikasi yang berfokus pada pemetaan dan visualisasi kondisi jalan juga dikaji dalam penelitian lain, salah satunya oleh Justam et al. (2024). Kajian tersebut mengolah basis data berupa 6.100 rekam kerusakan jalan pada tiga rute berbeda sepanjang kurun waktu 2021 hingga 2023. Melalui seleksi awal, diperoleh tujuh fitur utama yang dikelompokkan ke dalam tiga rumpun kelas, meliputi *Rutin Biasa* (1001 data), *Rutin Kondisi* (3957 data), dan *Holding* (1142 data). Selanjutnya, kumpulan data ini didistribusikan ke dalam subsampel training dan testing menggunakan perbandingan proporsi sebesar 80:20. Berdasarkan tahapan evaluasi, ditemukan bahwa arsitektur C4.5 menorehkan tingkat precision di angka 87.9%, recall 82.6%, F1-score 87.8%, dengan nilai akurasi total menyentuh 88%. Di sisi lain, pemodelan menggunakan Random Forest mencatatkan skor precision sebesar 86.6%, recall 86.8%, F1-score 86.6%, serta tingkat akurasi pada kisaran 87%. Meskipun akurasi C4.5 sedikit lebih tinggi, *Random Forest* dinilai lebih konsisten dalam proses klasifikasi. Penelitian ini menegaskan bahwa integrasi algoritma *machine learning* dengan sistem WebGIS mampu membantu pemerintah daerah dalam memetakan dan menentukan prioritas pemeliharaan fasilitas jalan secara lebih transparan, responsif, sekaligus efisien.

Sementara itu, penggunaan Algoritma *Random Forest* tidak hanya terbatas pada klasifikasi kerusakan jalan, tetapi juga telah diterapkan dalam konteks

prioritas pembangunan infrastruktur secara lebih luas. Urgensi tersebut selaras dengan pembuktian empiris yang dipaparkan dalam studi Maulindar & Guterres (2025) yang mana menggunakan data sekunder sebanyak 50 entri yang diperoleh dari pemerintah kota, dengan sepuluh variabel utama, antara lain kepadatan penduduk, pertumbuhan ekonomi, kepadatan lalu lintas, kebutuhan energi, ketersediaan infrastruktur, kualitas udara, frekuensi bencana alam, kebutuhan air bersih, indeks kepuasan publik, serta distribusi layanan publik. Data diproses melalui tahap pembersihan, normalisasi, beserta pendistribusian sampel yang dikelompokkan ke dalam 70% sebagai data training dan 30% untuk data testing. Selanjutnya, penyetelan parameter secara optimal dieksekusi lewat penerapan skema *grid search*. Hasil penelitian membuktikan keandalan model Random Forest yang berhasil menembus angka akurasi 90% sewaktu mengolah data uji, serta 88% saat diuji pada sampel validasi sekunder. Konklusi ini mengindikasikan tingkat generalisasi model yang tergolong prima. Di samping itu, pengujian melalui aspek *feature importance* memaparkan bahwa variabel kepadatan lalu lintas, ketersediaan infrastruktur, dan kebutuhan air bersih menjadi variabel yang mendominasi arah kebijakan pembangunan fasilitas publik. Fakta tersebut semakin menguatkan argumen bahwa Algoritma *Random Forest* memiliki potensi besar sebagai alat pendukung keputusan berbasis data dalam menentukan prioritas pembangunan infrastruktur, sehingga sangat kontekstual untuk diterapkan dalam studi sekarang yang menitikberatkan pada pemetaan skala prioritas pemeliharaan jalan.

Sebagai pengembangan dari berbagai penelitian sebelumnya yang menekankan performa model *Random Forest* dalam mengategorikan rumpun data

spasial maupun *non-spasial*, penelitian yang dilakukan oleh Yuliana & Rizqiana (2024) mengkaji pemanfaatan algoritma Random Forest (RF) untuk mengolah citra satelit demi memetakan sebaran tutupan lahan di wilayah perkotaan. Penelitian ini menggunakan data citra Sentinel-2 dan Landsat 8 yang diproses di platform *Google Earth Engine* (GEE), dengan empat kelas utama tutupan lahan, yaitu permukiman, lahan hijau, perairan, dan lahan terbuka. Data sampel diambil berupa data *ground truth* beserta peta acuan resmi yang diposisikan sebagai basis data latih untuk pemodelan *Random Forest*. Rekam eksperimen tersebut membuktikan bahwa penerapan metode ini mampu menguraikan tingkat akurasi total pada angka 89%, dengan akurasi per kelas masing-masing *permukiman* 92%, *lahan hijau* 85%, *perairan* 80%, dan *lahan terbuka* 87%. Validasi silang menghasilkan rata-rata akurasi 88.5%, *precision* 0.91, *recall* 0.88, dan *F1-score* 0.89, yang menegaskan keandalan model. Kesalahan klasifikasi yang muncul terutama disebabkan oleh tumpang tindih spektral antara lahan hijau dengan lahan terbuka, serta perairan dengan kelas lainnya. Penelitian ini menyimpulkan bahwa penggunaan *Random Forest* pada GEE sangat efisien, akurat, dan praktis untuk pemetaan tutupan lahan perkotaan, serta memiliki manfaat strategis dalam mendukung perencanaan tata ruang, pengelolaan lingkungan, dan pemantauan perubahan lahan secara berkelanjutan.

Pendekatan spasial dalam analisis data juga banyak dimanfaatkan dalam konteks infrastruktur jalan, khususnya dalam penentuan prioritas pemeliharaan. Aspek tersebut selaras dengan temuan ilmiah yang dipaparkan oleh Sitanggang (2022), studi tersebut mengkaji penataan skala prioritas pemeliharaan

jalan di wilayah Kabupaten Dairi lewat kolaborasi antara pendekatan *Analytical Hierarchy Process* (AHP) bersama ekosistem Sistem Informasi Geografis (SIG/GIS). Penelitian ini menggunakan beberapa parameter utama, yang meliputi keadaan fisik jalur, bentang panjang jalan, serta volume Lalu Lintas Harian Rata-rata (LHR) guna merepresentasikan derajat kepadatan arus kendaraan. Hasil penelitian membuktikan bahwa segmentasi jalan yang menunjukkan tingkat kerusakan fisik parah disertai angka LHR yang tinggi memiliki tingkat prioritas pemeliharaan yang lebih besar. Penggunaan GIS memungkinkan hasil prioritas divisualisasikan secara spasial sehingga memudahkan pengambilan keputusan. Temuan ini menegaskan bahwa LHR merupakan parameter penting yang dapat meningkatkan ketepatan penentuan prioritas pemeliharaan jalan.

Peran LHR sebagai indikator beban lalu lintas juga diperkuat dalam penelitian lain yang menggunakan pendekatan pembobotan prioritas. Penelitian oleh Aprillianto (2022) berfokus pada perancangan model tata urutan pemeliharaan prasarana jalan di wilayah Kota Tarakan lewat pengaplikasian teknik pembobotan nilai urgensi. Sederet indikator yang dilibatkan dalam pengujian tersebut mencakup derajat kerusakan fisik jalan, LHR, klasifikasi fungsi jalan, tingkat pelayanan jalan, serta riwayat pemeliharaan. Hasil penelitian tersebut mengindikasikan bahwa segmentasi jalur yang berada pada rentang kerusakan berskala sedang sampai berat disertai nilai LHR yang tinggi memperoleh skor prioritas tertinggi untuk dilakukan penanganan. Hal ini menunjukkan bahwa intensitas lalu lintas berpengaruh besar terhadap urgensi perbaikan jalan. Temuan penelitian ini memperkuat peran LHR

sebagai indikator beban lalu lintas yang relevan dalam penentuan prioritas penanganan jalan.

Sejalan dengan penelitian tersebut, pendekatan lain yang menggabungkan kondisi jalan dan kelas LHR juga dikaji oleh Priyanto (2023) mengkaji penentuan skala prioritas dan pola penanganan jalan dengan menggabungkan kelas LHR dan nilai kondisi jalan melalui suatu formula perhitungan prioritas. LHR diklasifikasikan berdasarkan volume lalu lintas harian, sedangkan kondisi jalan dinilai dari tingkat kerusakan perkerasan. Hasil penelitian menunjukkan bahwa ruas jalan dengan kelas LHR tinggi dan kondisi jalan yang buruk memiliki tingkat prioritas penanganan yang lebih besar. Pendekatan ini menegaskan bahwa kombinasi antara data kondisi jalan dan LHR dapat menghasilkan urutan prioritas yang logis dan konsisten. Temuan tersebut menunjukkan bahwa LHR merupakan faktor kunci yang tidak dapat dipisahkan dalam proses penentuan prioritas perbaikan jalan.

Selain banyak diterapkan pada klasifikasi infrastruktur dan data spasial, integrasi Algoritma *Random Forest* dengan metode lain juga digunakan dalam analisis data lingkungan dan iklim. Hal ini ditunjukkan dalam penelitian yang dilakukan oleh Rahman (2024) berfokus pada analisis data cuaca dari Stasiun Meteorologi Tunggul Wulung Cilacap periode 1975–2023 dengan mengombinasikan metode *K-Means clustering* dan Algoritma *Random Forest*. Metode *K-Means* digunakan untuk mengelompokkan data cuaca berdasarkan parameter utama seperti suhu, kelembaban, curah hujan, dan kecepatan angin, dengan jumlah cluster optimal ditentukan melalui *Elbow Method* dan *Silhouette*

Score yang pada beberapa tahun menunjukkan nilai cukup baik, misalnya 0.65–0.70. Hasil pengelompokan ini kemudian dijadikan dasar klasifikasi menggunakan *Random Forest*, yang terbukti mampu memprediksi kondisi cuaca dengan tingkat akurasi yang tinggi pada sebagian besar periode, meskipun mengalami penurunan performa pada tahun 2018–2023 akibat kompleksitas pola cuaca. Eksperimen parameter *Random Forest* dengan variasi jumlah pohon (*n_estimators* 100, 500, 1000) dan kedalaman pohon (*max_depth* 0, 5, 10) menunjukkan bahwa penyesuaian *hyperparameter* berpengaruh signifikan terhadap kualitas prediksi. Secara keseluruhan, penelitian ini menyimpulkan bahwa integrasi *K-Means* dan *Random Forest* efektif dalam meningkatkan akurasi prakiraan cuaca dan memberikan kontribusi penting bagi pemahaman pola iklim di wilayah Cilacap.

Keunggulan Algoritma *Random Forest* dalam menangani data kompleks dan multivariat juga terlihat pada penelitian yang memanfaatkan data geospasial untuk analisis risiko bencana. Penelitian yang dilakukan oleh Lemenkova (2024) membahas penerapan algoritma *machine learning* untuk menilai risiko banjir pesisir di kawasan pantai utara Jawa. Studi ini memanfaatkan data geospasial yang mencakup faktor-faktor seperti elevasi, jarak ke garis pantai, tata guna lahan, curah hujan, serta kepadatan penduduk, yang kemudian dianalisis menggunakan empat algoritma klasifikasi, yaitu *Random Forest* (RF), *Support Vector Machine* (SVM), *Gradient Boosting Machine* (GBM), dan *K-Nearest Neighbors* (KNN) dengan pembagian data 70% untuk training dan 30% untuk testing. Hasil pengujian menunjukkan bahwa *Random Forest* memberikan performa terbaik dengan akurasi 92%, *F1-score* 0.91, dan AUC 0.95, menjadikannya lebih unggul dibandingkan

SVM dan GBM (akurasi 88–90%) maupun KNN (82%). Analisis *feature importance* juga menegaskan bahwa elevasi, jarak ke pantai, dan kepadatan penduduk merupakan faktor paling signifikan dalam menentukan tingkat kerentanan banjir pesisir. Penelitian ini menyimpulkan bahwa integrasi *machine learning*, khususnya *Random Forest*, mampu menghasilkan pemetaan risiko banjir pesisir yang akurat dan dapat dijadikan dasar untuk strategi mitigasi bencana dan perencanaan tata ruang wilayah pesisir secara lebih adaptif dan berkelanjutan .

Pendekatan serupa dengan memanfaatkan data multi-sumber juga dikaji dalam penelitian lain yang berfokus pada pemetaan banjir. Penelitian oleh Albertini et al. (2024) meneliti pemetaan wilayah banjir dengan Algoritma *Random Forest* (RF) menggunakan data multi-sumber, yaitu citra optik Sentinel-2, radar Sentinel-1, dan fitur morfologi dari DEM. Penelitian ini menunjukkan bahwa indeks MNDWI, NDMI, dan RSWIR merupakan prediktor paling stabil, sementara fitur geomorfologi seperti GFI dan elevasi relatif (H) dapat meningkatkan akurasi bila data DEM berkualitas baik. Hasilnya, akurasi pemetaan banjir mencapai 92.67% di Italia (Sesia River), 94.61% di Australia (Wee Waa), 96.02% di Malawi, dan 95.82% di Italia (Emilia-Romagna). Kesimpulan ini membuktikan bahwa integrasi antara rumpun data optik, sensor radar, serta karakteristik morfologi tanah yang dipadukan dengan pemodelan RF efektif untuk menyajikan visualisasi sebaran banjir secara presisi, sekaligus menyokong efisiensi program mitigasi bencana alam.

Selain pada bidang lingkungan dan kebencanaan, keunggulan *Random Forest* juga terbukti pada kasus klasifikasi data nonspasial. Hal ini ditunjukkan

dalam penelitian Helmud et.al (2024) yang menguji komparasi performa antara arsitektur pembelajaran mesin *Random Forest* dengan *Decision Tree*. Kajian tersebut diarahkan untuk mengategorikan basis data ketenagakerjaan di India untuk memprediksi apakah karyawan akan keluar atau tetap bekerja. *Dataset* yang digunakan berjumlah 4654 data, dengan pembagian 90% sebagai data latih dan 10% sebagai data uji, kemudian dianalisis menggunakan *confusion matrix*. Catatan hasil pengujian membuktikan bahwa pemodelan *Random Forest* mencapai tingkat akurasi yang lebih dominan di angka 86.45%, mengungguli metode *Decision Tree* yang tertahan pada level 84.30%. Keunggulan skema *Random Forest* ini juga diperkuat oleh capaian parameter evaluasi sekunder lainnya juga lebih baik pada *Random Forest* (*Precision* 96.7%, *Recall* 84.7%, *Specificity* 91.4%, dan *F1-Score* 90.2%). Temuan ini mengindikasikan keandalan model *Random Forest* yang jauh lebih responsif untuk mendeteksi status kepegawaian daripada arsitektur *Decision Tree*. Konsekuensinya, algoritma ini sangat direkomendasikan untuk memecahkan problem klasterisasi dengan rumpun kasus serupa.

Tabel 2. 1 Penelitian Terkait

No	Judul Referensi	Metode	Hasil	Relevansi Penelitian
1.	Sistem Klasifikasi Kerusakan Jalan Metode <i>Machine Learning</i> dengan Algoritma <i>K-Means</i> dan <i>Random Forest</i> (Rahman et al., 2025)	<i>K-Means</i> <i>Random Forest</i>	- Akurasi keseluruhan 95%. - <i>Precision</i> tertinggi 100% (Rusak Parah & Rusak Sedang). - <i>Recall</i> tertinggi 100% (Baik), namun rendah pada Rusak Parah (75%) & Rusak Sedang (67%).	- Menjadi rujukan penggunaan <i>Random Forest</i> dalam klasifikasi data jalan, khususnya tingkat kerusakan. - Penelitian ini dikembangkan lebih lanjut dengan fokus pada klasifikasi prioritas perbaikan jalan dan penambahan parameter pendukung seperti LHR.
2.	Penerapan Algoritma C4.5 dan <i>Random Forest</i>	<i>C4.5</i> & <i>Random Forest</i> .	- C4.5: <i>Precision</i> 87,9%, <i>Recall</i> 82,6%,	- Membuktikan keandalan dan kapabilitas model

	untuk Pemetaan Kerusakan Jalan dengan WebGIS. (Justam et al., 2024)		<i>F1-score</i> 87,8%, Akurasi 88%. - <i>Random Forest</i>: <i>Precision</i> 86,6%, <i>Recall</i> 86,8%, <i>F1-score</i> 86,6%, Akurasi 87%. - <i>Random Forest</i> lebih konsisten meski akurasi C4.5 sedikit lebih tinggi.	Random Forest sewaktu melakukan klasterisasi kerusakan infrastruktur jalan bertumpu pada basis data numerik. - Diferensiasi dengan studi sekarang terletak pada orientasi output; penelitian ini mengarah pada perumusan skala urutan perbaikan jalan, bukan sebatas menyajikan visualisasi sebaran tingkat kerusakan.
3.	<i>The Implementation of Random Forest Classification for Identifying Infrastructure Priorities in Smart Cities</i> (Maulindar & Guterres, 2025)	<i>Random Forest</i>	- Akurasi 90% pada data uji. - Validasi dengan data tambahan akurasi 88%. - Variabel paling berpengaruh: kepadatan lalu lintas, ketersediaan infrastruktur, kebutuhan air bersih. - Model dapat membantu pemerintah kota menentukan prioritas pembangunan infrastruktur.	- Menunjukkan kemampuan Random Forest dalam menentukan prioritas infrastruktur berbasis data. - Kajian tersebut kemudian diadaptasikan ke dalam riset saat ini untuk mengukur tingkat kepentingan pemeliharaan jalan, dengan memfokuskan variabel pada keadaan fisik rute serta intensitas arus lalu lintas.
4.	Analisis Metode Klasifikasi Pemetaan Tutupan Lahan di Area Kota Bandung Menggunakan Algoritma <i>Random Forest</i> Pada <i>Google Earth Engine</i> . (Yuliana & Rizqiana, 2024)	<i>Random Forest</i>	- Akurasi keseluruhan 89%. - Akurasi per kelas: Permukiman (92%), Lahan Hijau (85%), Perairan (80%), Lahan Terbuka (87%). - Validasi silang: akurasi rata-rata 88,5%, <i>precision</i> 0,91, <i>recall</i> 0,88, <i>F1-score</i> 0,89.	- Membuktikan keandalan Random Forest dalam klasifikasi data berbasis spasial dengan akurasi tinggi. - Berbeda dengan penelitian ini yang berfokus pada klasifikasi prioritas perbaikan jalan, bukan pemetaan tutupan lahan.
5	Analisis Prioritas Pemeliharaan Jalan Kabupaten Dairi Menggunakan AHP dan GIS Sitanggang (2022)	<i>Analytical Hierarchy Process</i> (AHP) dan Sistem Informasi Geografis (GIS)	Berdasarkan temuan riset, tingkat urgensi pemeliharaan yang paling tinggi ditemukan pada segmen jalan yang mengalami kerusakan fisik parah serta memiliki volume LHR yang padat. LHR terbukti menjadi parameter penting	- Menegaskan LHR dan kondisi jalan sebagai penentu prioritas pemeliharaan. - Dikembangkan dalam skripsi ini dengan Random Forest agar penentuan prioritas lebih objektif.

			dalam penentuan prioritas pemeliharaan jalan.	
6	Analisis Pemetaan Skala Prioritas Penanganan Jalan Kota Tarakan Aprillianto (2022)	Metode Pembobotan Skala Prioritas	Hasil penelitian mengonfirmasi bahwa akumulasi nilai prioritas utama untuk program perbaikan diberikan kepada rute jalan dengan derajat kerusakan di fase sedang sampai berat yang dibarengi angka LHR tinggi.	- Membuktikan pengaruh kerusakan jalan dan LHR terhadap prioritas penanganan. - Skripsi ini menggunakan Random Forest untuk memodelkan hubungan antarvariabel.
7	Penentuan Skala Prioritas dan Pola Penanganan Jalan Berdasarkan Kelas LHR dan Nilai Kondisi Jalan Priyanto (2023)	Perhitungan Matematis (Formula Prioritas)	Integrasi antara klasifikasi kelompok LHR dengan parameter mutu jalan mampu membentuk sebuah skema urutan prioritas yang rasional, di mana ruas jalan dengan LHR tinggi dan kondisi buruk menjadi prioritas utama.	- Menjadi dasar penggunaan kelas LHR dan kondisi jalan dalam penentuan prioritas. - Dikembangkan melalui klasifikasi Random Forest untuk hasil yang lebih akurat.
8.	<i>Assessing multi-source random forest classification and robustness of predictor variables in flooded areas mapping.</i> (Albertini et al., 2024)	<i>Random Forest</i>	- Akurasi pemetaan banjir sangat tinggi (>92%) - Variabel paling stabil: MNDWI, NDMI, RSWIR - Integrasi optik, radar, dan morfologi paling efektif	- Membuktikan Random Forest unggul dan stabil dalam klasifikasi berbasis data kompleks. - Menguatkan penggunaan Random Forest pada skripsi ini untuk klasifikasi prioritas perbaikan jalan.
9.	Perbandingan Algoritma <i>Random Forest</i> dan <i>Decision Tree</i> untuk Klasifikasi Data Karyawan di India (Helmud et al., 2024)	<i>Random Forest</i> dan <i>Decision Tree</i>	- Pemodelan menggunakan <i>Random Forest</i> mencatatkan performa yang lebih optimal (dengan tingkat akurasi 86.45%) jika dibandingkan dengan skema <i>Decision Tree</i> (yang berada di angka 84.30%). - Metrik lain juga lebih tinggi pada RF	- Menjadi bukti empiris mengenai keunggulan serta stabilitas algoritma Random Forest atas model Decision Tree untuk menyelesaikan masalah klasterisasi data. - Berperan sebagai landasan kuat di balik keputusan mengadopsi Random Forest sebagai metode fundamental dalam penelitian sekarang.
10.	<i>Clustering dan Klasifikasi Data Cuaca Kota Cilacap Menggunakan K-</i>	<i>K-Means dan Random Forest</i>	- <i>K-Means</i> efektif mengelompokkan data cuaca (nilai <i>Silhouette Score</i> 0,65–0,70)	- Menunjukkan efektivitas K-Means sebagai pendukung Random Forest dalam

	<i>Means dan Random Forest</i> (Rahman et al., 2024)		- <i>Random Forest</i> mencapai akurasi tinggi, meski menurun di periode 2018–2023	pengolahan data kompleks. - Pendekatan ini diadaptasi pada skripsi untuk klasifikasi prioritas perbaikan jalan.
11.	<i>Random Forest Classifier Algorithm of GRASS GIS for Satellite Image Processing: Case Study of Bight of Sofala, Mozambique</i> (Lemenkova, 2024)	<i>Random Forest</i>	- Berhasil memetakan perubahan pesisir Mozambik - Ditemukan penurunan mangrove dan peningkatan lahan terbangun - RF lebih akurat daripada metode tradisional	- Menunjukkan akurasi tinggi Random Forest dalam pengolahan data spasial. - Pendekatan ini relevan untuk diterapkan pada data jalan dalam penentuan prioritas perbaikan.
12.	Klasifikasi Prioritas Perbaikan Jalan di Kabupaten Malang Menggunakan Random Forest	<i>Random Forest</i>	- Klasifikasi prioritas perbaikan jalan menjadi tiga kelas: Tinggi, Sedang, Rendah. - Pengujian performa model bersandarkan pada instrumen <i>confusion matrix</i> dengan mengukur parameter akurasi, tingkat <i>precision</i> , nilai <i>recall</i> , serta capaian <i>f1-score</i> .	- Mengimplementasikan Random Forest untuk klasifikasi prioritas perbaikan jalan menjadi tiga kelas. - Menjadi pengembangan dari penelitian sebelumnya dengan fokus langsung pada pengambilan

Berdasarkan tabel 2.1, terlihat jelas bahwa arsitektur *Random Forest* mendominasi sebagai pilihan instrumen utama yang kerap diandalkan dalam berbagai studi kategorisasi pada berbagai bidang, mulai dari pemetaan kerusakan jalan, klasifikasi lahan, hingga pemetaan bencana. Secara umum, *Random Forest* mampu memberikan hasil akurasi yang tinggi (di atas 85%), konsistensi performa yang stabil, serta lebih unggul dibandingkan metode lain seperti *K-Means*, *C4.5*, maupun *Decision Tree*. Namun, mayoritas penelitian masih memiliki keterbatasan pada jumlah data, keterbatasan area uji, serta belum diuji secara luas pada berbagai konteks wilayah.

Titik pembeda mendasar antar-literatur tersebut berada pada corak algoritma yang diadopsi serta lokus penerapannya. Sejumlah riset terdahulu tampak mengkombinasikan beberapa rumpun metode (misalnya *K-Means* dengan *Random Forest*) untuk meningkatkan hasil *clustering* maupun klasifikasi, sedangkan yang lain membandingkan *Random Forest* dengan algoritma lain (seperti C4.5 atau *Decision Tree*). Dari hasil komparatif tersebut, model *Random Forest* dinilai mendominasi dalam aspek presisi prediksi, konsistensi performa, hingga kapabilitasnya dalam mengolah karakteristik basis data yang rumit, meskipun tetap memiliki tantangan terkait *preprocessing* data dan keterbatasan variabel input.

Dengan berbagai perbedaan tersebut, kajian ini diproyeksikan mampu menyumbang kontribusi nyata terhadap formulasi model kategorisasi urutan pemeliharaan jalan yang lebih presisi, transparan, dan dapat digunakan secara praktis sebagai bahan pertimbangan dalam pengelolaan sekaligus perancangan prasarana jalan secara efisien dan jangka panjang.

2.2 Prioritas Perbaikan Jalan

Jalan merupakan salah satu infrastruktur vital yang memegang andil besar dalam menunjang perputaran sektor sosial, roda ekonomi, hingga pergerakan mobilitas publik. Ketersediaan jalan dengan kondisi baik mampu memperlancar arus transportasi, distribusi barang, serta meningkatkan aksesibilitas menuju pusat kegiatan ekonomi, pendidikan, maupun pelayanan publik (Putri & Susanto, 2025). Sebaliknya, kerusakan jalan dapat menimbulkan berbagai permasalahan seperti terganggunya kelancaran lalu lintas, meningkatnya risiko kecelakaan, hingga bertambahnya biaya operasional kendaraan.

Kerusakan jalan umumnya terjadi karena faktor tingginya volume beban kendaraan yang melintas, fluktuasi cuaca dan kondisi iklim, buruknya saluran pembuangan air (drainase), maupun umur konstruksi jalan yang sudah melewati masa layan. Permasalahan tersebut seringkali menimbulkan kesenjangan antara kebutuhan perbaikan dengan keterbatasan anggaran yang dimiliki pemerintah daerah (Rahman et al., 2025). Karena itu, tidak semua jalan dapat diperbaiki secara bersamaan, sehingga dibutuhkan suatu sistem prioritas perbaikan jalan untuk memastikan penanganan dilakukan secara tepat sasaran.

Pada prinsipnya, perancangan urutan urgensi pemeliharaan jalan ditentukan oleh sederet indikator krusial. Faktor-faktor tersebut meliputi derajat kerusakan fisik jalur (baik skala ringan, sedang, maupun berat), intensitas arus lalu lintas harian, letak strategis rute, hingga implikasi sosial serta dampak sosial dan ekonomi yang ditimbulkan (Putri & Susanto, 2025). Selama ini, penentuan prioritas biasanya dilakukan secara manual melalui survei lapangan dan analisis subjektif oleh instansi terkait. Metode tersebut cenderung memerlukan waktu yang lama, biaya besar, dan rentan menimbulkan ketidaktepatan dalam pengambilan keputusan. Atas dasar permasalahan tersebut, integrasi metode bertumpu pada teknik *data mining* ataupun *algoritma machine learning* yang mampu melakukan klasifikasi prioritas perbaikan jalan secara lebih akurat, objektif, dan efisien.

Dalam konteks Kabupaten Malang, luasnya wilayah dan beragamnya jaringan jalan menyebabkan permasalahan kerusakan jalan cukup sering terjadi. Kondisi ini menuntut adanya sistem yang mampu membantu pengelolaan dan perencanaan infrastruktur secara lebih terukur. Dengan penerapan metode

klasifikasi yang tepat, prioritas perbaikan jalan di Kabupaten Malang dapat ditentukan secara lebih sistematis sehingga mendukung pembangunan infrastruktur yang berkelanjutan.

2.3 Klasifikasi

Klasifikasi diposisikan sebagai salah satu metode fundamental guna memetakan sekumpulan data ke dalam kompartemen atau klaster spesifik berdasarkan pada karakteristik atributnya. Target utama dari proses ini berkaitan dengan penyusunan suatu kerangka model atau fungsi yang mampu memprediksi data anyar agar masuk ke dalam rumpun kelas yang sudah didefinisikan pada tahap awal (Rahman et al., 2024). Dengan demikian, metode klasifikasi digolongkan ke dalam ranah pembelajaran terpandu (*supervised learning*), karena dalam proses pengkajian model membutuhkan data latih (*training data*) yang sudah memiliki label kelas.

Proses klasifikasi secara umum terdiri dari dua fase krusial, yang meliputi tahap latih (*training*) serta tahapan uji (*testing*). Sewaktu fase latih berlangsung, sistem akan merumuskan struktur model berdasarkan data latih yang berisi atribut serta label kelas. Model yang dihasilkan selanjutnya divalidasi memanfaatkan sekumpulan data uji (*testing data*) demi mengukur parameter akurasi sekaligus performa komputasinya (Yuliana & Rizqiana, 2024). Hasil peninjauan ini memegang peranan vital untuk mengukur kelayakan model dalam memprediksi data baru.

Terdapat beragam jenis sistem komputasi yang bisa diadaptasikan untuk kebutuhan klasifikasi, seperti *Decision Tree*, *Naïve Bayes*, *k-Nearest Neighbor* (*k-*

NN), *Support Vector Machine* (SVM), hingga *Random Forest*. Berhubung masing-masing rumpun model mengantongi keunggulan serta batasannya tersendiri, maka penentuan opsi algoritma yang hendak dipakai akan sangat bergantung pada tipologi basis data serta target luaran dari studi tersebut. Misalnya, Algoritma *Decision Tree* menawarkan kemudahan dari segi pemahaman alur kerja dan visualisasi grafis, tetapi rentan terjebak dalam masalah *overfitting* jika dihadapkan pada basis data yang rumit (Helmud et al., 2024). Sebagai solusinya, arsitektur *Random Forest* hadir sebagai bentuk modifikasi tingkat lanjut dari *Decision Tree* yang didesain untuk mampu mengurangi risiko *overfitting* melalui teknik pengintegrasian kelompok struktur pohon keputusan (*ensemble learning*), yang pada akhirnya mampu memicu capaian akurasi secara lebih optimal.

Klasifikasi tercatat telah diimplementasikan secara luas pada bermacam sektor, seperti kesehatan (misalnya prediksi penyakit), transportasi (klasifikasi kondisi jalan), keuangan (analisis kredit), hingga lingkungan (pemetaan penggunaan lahan). Dalam penelitian ini, klasifikasi digunakan guna merumuskan skala urutan urgensi pemeliharaan jalan di wilayah Kabupaten Malang. Dengan menerapkan metode klasifikasi, data terkait kondisi jalan dapat diolah untuk menghasilkan kategori prioritas tertentu, alur penentuan kebijakan strategis mampu dieksekusi secara lebih sistematis, presisi, serta imparsial.

2.4 Preprocessing

Preprocessing merupakan tahap penyiapan basis data awal sebelum diinput ke dalam sistem komputasi *machine learning*. Tahapan ini memegang andil yang sangat krusial mengingat validitas luaran hasil klasifikasi akan berkorelasi langsung

dengan mutu data yang diolah. Preprocessing dalam penelitian ini meliputi beberapa langkah utama, yaitu pembersihan data, penggabungan dataset, normalisasi, dan pembagian data (split data).

1. Pembersihan Data (Data Cleaning)

Data yang diperoleh dari Dinas Pekerjaan Umum dan Penataan Ruang (PUPR) Kabupaten Malang berpotensi mengandung masalah seperti duplikasi, nilai kosong (*missing values*), atau nilai tidak logis (*outlier*). Sebagai contoh, pada data LHR dapat ditemukan nilai yang terlalu tinggi dan tidak sesuai dengan kondisi riil lalu lintas. Pada tahap ini dilakukan pemeriksaan data untuk memastikan konsistensi dan validitas. Data yang tidak sesuai dapat dibuang agar tidak memengaruhi performa model.

2. Normalisasi Data (Data Normalization)

Karena fitur dalam dataset memiliki rentang nilai yang berbeda (misalnya Panjang Jalan dalam kilometer dan LHR dalam jumlah kendaraan), maka dilakukan normalisasi agar skala data seragam. Normalisasi digunakan untuk mencegah fitur dengan nilai besar mendominasi proses pembelajaran model. Dalam penelitian ini digunakan metode *Min-Max Normalization*.

Salah satu metode yang umum digunakan adalah *Min-Max Normalization*, yaitu mengubah setiap nilai ke dalam rentang [0-1]. Rumus normalisasi *Min-Max* adalah sebagai berikut:

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (2.1)$$

Keterangan :
x = nilai asli

x' = nilai hasil normalisasi
 x_{min} = nilai terkecil dalam kolom data
 x_{max} = nilai terbesar dalam kolom data

Pada Rumus (2.1) ditunjukkan bahwa normalisasi *Min-Max* dilakukan dengan mengubah setiap nilai asli (x) ke dalam rentang $[0-1]$ berdasarkan nilai minimum (x_{min}) dan maksimum (x_{max}) dalam suatu kolom data. Hasil normalisasi (x') membuat skala data menjadi seragam, sehingga mencegah perbedaan skala antar fitur memengaruhi proses pembelajaran model.

3. Pembagian Data (Split Data)

Setelah data siap digunakan, dilakukan pembagian dataset menjadi dua bagian, yaitu data latih (*training set*) dan data uji (*testing set*). Data latih digunakan untuk membangun model *Random Forest*, sedangkan data uji digunakan untuk mengevaluasi performa model. Pembagian dilakukan dengan proporsi tertentu, misalnya 70% data untuk latih dan 30% data untuk uji.

Secara matematis, jika D adalah *dataset* keseluruhan dengan jumlah data n , maka dapat dituliskan sebagai berikut:

$$D = D_{Train} \cup D_{Test} \quad (2.2)$$

Keterangan

D = *dataset* keseluruhan

D_{Train} = data latih

D_{Test} = data uji

Persamaan (2.2) menunjukkan bahwa dataset keseluruhan (D) dibagi menjadi dua bagian, yaitu data latih (D_{train}) dan data uji (D_{test}). Segmen data latih

difungsikan agar arsitektur *Random Forest* dapat mendeteksi pola kelakuan serta korelasi antar-variabel atribut, sementara data uji disiapkan sebagai instrumen untuk menguji kapabilitas model dalam mengategorikan objek anyar yang belum terekam pada fase latihan. Merujuk pada eksperimen ini, distribusi data dikondisikan dengan rasio komparasi 70% untuk kebutuhan pemelajaran model dan 30% untuk proses pengujian. Hal tersebut ditujukan agar model dapat dilatih menggunakan sebagian besar data namun tetap memiliki data independen yang dioptimalkan guna menakar performa komputasi model secara imparsial.

Lewat intervensi fase pengondisian awal (*preprocessing*) ini, sekumpulan informasi yang awalnya bersifat mentah dapat dikonversi menjadi lebih higienis, saling terhubung, dan siap digunakan oleh Algoritma *Random Forest*. Proses ini memastikan bahwa model dapat belajar dari data yang valid, konsisten, serta tersusun sistematis demi memicu output klasterisasi yang jauh lebih presisi sekaligus akuntabel.

2.5 Random Forest

Random Forest merupakan salah satu arsitektur machine learning yang bersandar pada konsep pemelajaran kolektif (*ensemble learning*). Kerangka kerja algoritma ini bertumpu pada perumusan sekumpulan struktur pohon keputusan (*decision tree*) secara acak, yang kemudian mengombinasikan akumulasi luaran prediksi dari setiap unit pohon guna menetapkan kesimpulan akhir lewat mekanisme pemungutan suara terbanyak atau *majority voting* (pada rumpun klasifikasi) maupun skema rata-rata atau *averaging* (pada rumpun regresi). *Random Forest* pertama kali diperkenalkan oleh Leo Breiman pada tahun 2001 sebagai

pengembangan dari algoritma *bagging* dan pohon keputusan, dengan tujuan meningkatkan akurasi dan mengurangi permasalahan *overfitting* pada model.

Konsep dasar *Random Forest* berasal dari metode *bagging* (Bootstrap Aggregating) yang diperkenalkan oleh Breiman pada tahun 1996. Aplikasi metode *Bagging* diproyeksikan untuk mendongkrak stabilitas kerangka kerja model lewat penyusunan beragam variasi model dari ekstraksi sampel data yang heterogen, lalu menggabungkan hasil prediksinya. Selanjutnya, Breiman bersama Ho mengembangkan ide penggunaan fitur acak dalam membangun pohon keputusan. Gabungan dari konsep *bagging* dengan pemilihan fitur acak inilah yang kemudian menghasilkan Algoritma *Random Forest* yang dikenal luas hingga saat ini.

Tahapan utama dalam membangun model *Random Forest* adalah sebagai berikut:

1. *Bootstrap Sampling*, Membuat sampel data secara acak (*sampling with replacement*) dari data latih untuk membangun setiap pohon.
2. *Random Feature Selection*, Pada setiap node pohon, dipilih subset fitur secara acak, lalu dipilih atribut terbaik berdasarkan kriteria tertentu (misalnya Gini Index atau Information Gain).

Jumlah fitur yang dipilih secara acak pada setiap node ditentukan dengan rumus:

$$M = \sqrt{n} \quad (2.4)$$

Keterangan :

M = jumlah maksimum fitur yang dipilih secara acak pada setiap node
 n = jumlah total fitur dalam *dataset*

Pada Rumus (2.4) ditunjukkan bahwa jumlah fitur acak yang dipilih pada setiap node ditentukan dengan $M = \sqrt{n}$, di mana n adalah jumlah total fitur pada dataset.

Dalam penelitian ini, kriteria yang digunakan adalah Gini Index, karena perhitungannya lebih sederhana dibandingkan *Entropy* namun tetap efektif untuk mengukur *impurity* (ketidakmurnian) data. Rumus Gini Index adalah sebagai berikut:

$$Gini(t) = 1 - \sum_{i=1}^c p_i^2 \quad (2.5)$$

Keterangan

C = jumlah kelas target (misalnya: prioritas tinggi, sedang, rendah)

p_i = proporsi data kelas ke- i pada node t

Pada Rumus (2.5), nilai Gini Index digunakan untuk mengukur tingkat *impurity* atau ketidakmurnian suatu node. Semakin kecil nilai Gini, semakin homogen data dalam node tersebut, sehingga pemisahan (*split*) dianggap lebih baik.

3. *Tree Construction*, Membangun pohon keputusan hingga maksimal, tanpa dilakukan pemangkasan (*pruning*).
4. *Voting/Averaging*, Hasil dari semua pohon digabungkan dengan menggunakan *majority voting* (untuk klasifikasi) atau *averaging* (untuk regresi) guna menghasilkan prediksi akhir.

Secara matematis, prediksi klasifikasi dengan *Random Forest* dapat dirumuskan sebagai berikut:

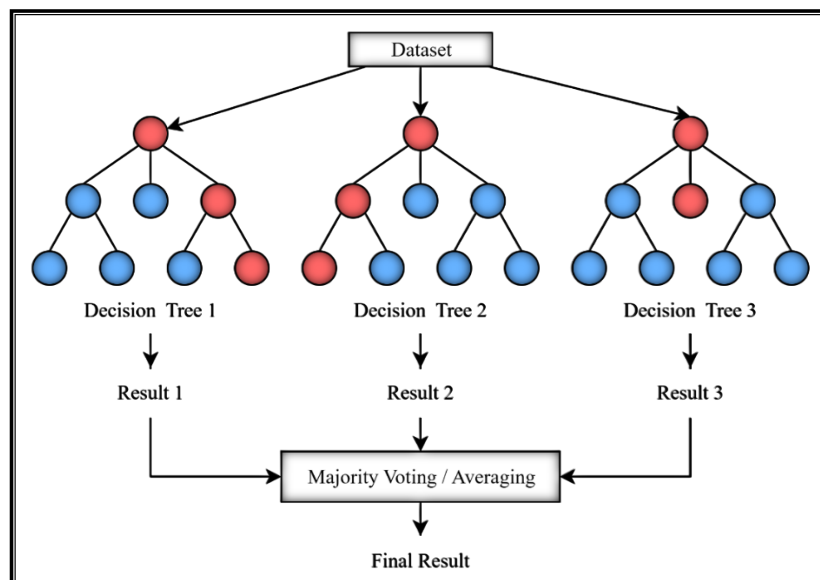
$$\hat{f}(x) = mode \{T_1(x), T_2(x), \dots, T_B(x)\}, \quad (2.6)$$

Keterangan :

$\hat{f}(x)$ = Hasil prediksi akhir *Random Forest*

$T_B(x)$ = Prediksi dari pohon keputusan ke-b

Pada Rumus (2.6) ditunjukkan bahwa hasil prediksi akhir $\hat{f}(x)$ pada Algoritma *Random Forest* diperoleh dengan cara majority voting dari seluruh pohon keputusan $T_B(x)$. Artinya, kelas yang paling banyak dipilih oleh pohon-pohon tersebut akan menjadi hasil klasifikasi akhir.



Gambar 2. 1 Algoritma Random Forest

Gambar 2.1 memperlihatkan alur kerja Algoritma *Random Forest* yang dibangun dari kumpulan pohon keputusan (*decision tree*). *Dataset* awal dibagi menjadi beberapa subset yang berbeda, baik dari sisi data maupun fitur, kemudian digunakan untuk membentuk pohon keputusan secara acak. Setiap pohon menghasilkan prediksi secara independen, yang selanjutnya digabungkan melalui mekanisme *majority voting* pada kasus klasifikasi atau *averaging* pada kasus regresi. Pendekatan ini membuat *Random Forest* lebih unggul dibandingkan

penggunaan satu pohon keputusan karena mampu meningkatkan akurasi, mengurangi risiko *overfitting*, serta memberikan hasil prediksi yang lebih stabil.

Secara matematis, agregasi prediksi pada *Random Forest* klasifikasi dapat dituliskan sebagai berikut:

$$H(x) = \text{mode} \{h_1(x), h_2(x), \dots, h_k(x)\} \quad (2.7)$$

Keterangan

$H(x)$ = hasil prediksi akhir untuk data input x

$h_i(x)$ = Hasil Prediksi Pohon ke i

k = jumlah pohon dalam *Random Forest*

mode = nilai kelas yang paling sering muncul

Rumus (2.7) menjelaskan bahwa hasil akhir prediksi pada Algoritma *Random Forest* diperoleh melalui proses agregasi dengan menggunakan metode *majority voting* atau pemilihan kelas yang paling sering muncul. Setiap pohon keputusan $h_i(x)$ menghasilkan prediksi terhadap data input x , kemudian seluruh hasil prediksi tersebut dikumpulkan. Fungsi agregasi $H(x)$ akan memilih kelas yang paling dominan sebagai hasil akhir prediksi.

Dalam penelitian ini, Algoritma *Random Forest* diterapkan untuk melakukan klasifikasi prioritas perbaikan jalan di Kabupaten Malang. *Dataset* yang digunakan memuat informasi terkait kondisi jalan, tingkat kerusakan, dan variabel pendukung lainnya yang relevan. Setiap pohon keputusan pada *Random Forest* memberikan hasil klasifikasi prioritas, misalnya tinggi, sedang, atau rendah, yang kemudian digabungkan dengan metode *majority voting* sehingga diperoleh hasil klasifikasi akhir. Dengan demikian, penggunaan *Random Forest* diharapkan mampu menghasilkan model klasifikasi prioritas perbaikan jalan yang lebih

objektif, transparan, serta dapat mendukung pengambilan keputusan secara sistematis.

2.6 Evaluasi Model

Evaluasi kinerja model merupakan tahapan penting dalam penelitian berbasis *machine learning* untuk mengetahui seberapa baik model yang dibangun dalam melakukan prediksi atau klasifikasi. Pada penelitian ini, evaluasi dilakukan untuk mengukur kinerja Algoritma *Random Forest* dalam mengklasifikasikan prioritas perbaikan jalan. Evaluasi diperlukan agar hasil prediksi yang dihasilkan tidak hanya akurat secara hitungan algoritma, tetapi juga relevan dengan tujuan praktis penelitian.

Salah satu metode evaluasi yang paling umum digunakan dalam klasifikasi adalah *confusion matrix*. *Confusion matrix* merupakan representasi tabel yang menunjukkan perbandingan antara hasil prediksi model dengan kondisi aktual pada *dataset* (Helmud et al., 2024). Dari *confusion matrix* dapat dihitung berbagai metrik evaluasi seperti akurasi, presisi, *recall*, dan *F1-score*.

1. Akurasi

Akurasi merupakan metrik paling umum yang digunakan untuk mengevaluasi model klasifikasi. Akurasi menunjukkan seberapa besar proporsi prediksi yang benar (baik kelas positif maupun negatif) dibandingkan dengan jumlah keseluruhan data uji. Semakin tinggi nilai akurasi, semakin baik kinerja model dalam mengklasifikasikan data secara keseluruhan.

$$Akurasi = \frac{TP + TN}{TP + TN + FP + FN} \quad (2.8)$$

Keterangan:

TP (*True Positive*) = jumlah data kelas positif yang diprediksi benar oleh model

TN (*True Negative*) = jumlah data kelas negatif yang diprediksi benar oleh model

FP (*False Positive*) = jumlah data kelas negatif yang salah diprediksi sebagai positif

FN (*False Negative*) = jumlah data kelas positif yang salah diprediksi sebagai negative

Pada Rumus (2.8) terlihat bahwa nilai akurasi dihitung dari perbandingan jumlah prediksi benar ($TP + TN$) dengan keseluruhan data uji. Semakin besar nilai akurasi, semakin baik kinerja model dalam melakukan klasifikasi secara keseluruhan.

2. Presisi

Presisi mengukur tingkat ketepatan prediksi positif yang diberikan model. *Precision* merupakan metrik evaluasi yang digunakan untuk mengukur tingkat ketepatan model dalam memprediksi suatu kelas. Dengan kata lain, *precision* menunjukkan seberapa besar proporsi data yang diprediksi sebagai kelas tertentu memang benar-benar termasuk dalam kelas tersebut.

Precision sangat penting ketika konsekuensi dari prediksi positif yang salah (*false positive*) perlu diminimalisasi. Dalam konteks penelitian ini, *precision* dapat menunjukkan seberapa tepat model dalam memprediksi ruas jalan dengan prioritas tinggi, sedang, maupun rendah.

$$Presisi = \frac{TP}{TP + FP} \quad (2.9)$$

Keterangan :

TP (*True Positive*) = jumlah data kelas positif yang diprediksi benar oleh model

FP (*False Positive*) = jumlah data kelas negatif yang salah diprediksi sebagai positif

Pada Rumus (2.9) terlihat bahwa nilai *precision* diperoleh dari perbandingan jumlah prediksi positif yang benar (TP) terhadap seluruh data

yang diprediksi positif ($TP + FP$). Semakin tinggi nilai precision, semakin tepat model dalam mengklasifikasikan data ke dalam kelas yang benar tanpa banyak kesalahan prediksi positif.

3. Recall

Recall, atau sering juga disebut *Sensitivity* atau *True Positive Rate*, merupakan metrik evaluasi yang digunakan untuk mengukur kemampuan model dalam menemukan seluruh data yang benar-benar termasuk dalam suatu kelas tertentu. Dengan kata lain, *recall* menunjukkan seberapa banyak data kelas positif yang berhasil teridentifikasi dengan benar oleh model.

Recall sangat penting ketika konsekuensi dari *false negative* perlu diminimalisasi. Dalam konteks penelitian ini, *recall* memberikan gambaran seberapa baik model dapat mengenali ruas jalan yang seharusnya diprioritaskan untuk perbaikan. Jika *recall* pada kelas *Prioritas Tinggi* rendah, artinya masih ada banyak ruas jalan dengan kondisi buruk yang tidak terdeteksi sebagai prioritas tinggi.

$$Recall = \frac{TP}{TP + FN} \quad (2.10)$$

Keterangan:

TP (*True Positive*) = jumlah data kelas positif yang diprediksi benar oleh model

FN (*False Negative*) = jumlah data kelas positif yang salah diprediksi sebagai negative

Pada Rumus (2.10) terlihat bahwa recall dihitung dari perbandingan jumlah prediksi positif yang benar (TP) terhadap seluruh data aktual positif ($TP + FN$). Nilai recall yang tinggi menunjukkan bahwa model mampu mengenali sebagian besar data yang seharusnya termasuk dalam kelas positif, sehingga semakin sedikit data positif yang terlewat.

4. *F1- Score*

F1-Score merupakan metrik evaluasi yang mengombinasikan *Precision* dan *Recall* ke dalam satu nilai tunggal melalui rata-rata harmonis. Metrik ini sangat berguna ketika diperlukan keseimbangan antara ketepatan prediksi (*precision*) dan kemampuan mendeteksi seluruh data positif (*recall*).

Dalam penelitian ini, *F1-Score* digunakan untuk menilai sejauh mana model *Random Forest* mampu melakukan klasifikasi prioritas perbaikan jalan dengan mempertimbangkan kedua aspek tersebut secara bersamaan. Hal ini penting karena model tidak hanya harus tepat dalam memprediksi prioritas tertentu, tetapi juga harus mampu mengenali seluruh ruas jalan yang benar-benar membutuhkan perbaikan. Secara matematis, *F1-Score* dirumuskan sebagai berikut:

$$Recall = 2 \times \frac{Presisi \times Recall}{Presisi + Recall} \quad (2.11)$$

Keterangan:

Precision = proporsi prediksi positif yang benar dibandingkan semua prediksi positif
Recall = proporsi data positif yang berhasil diprediksi dibandingkan seluruh data positif actual

Pada Rumus (2.11) terlihat bahwa *F1-Score* merupakan rata-rata harmonis antara *precision* dan *recall*. Nilai ini memberikan ukuran yang seimbang dengan mempertimbangkan ketepatan prediksi (*precision*) sekaligus kemampuan model dalam mengenali data positif secara menyeluruh (*recall*). Semakin tinggi nilai *F1-Score*, semakin baik kinerja model dalam klasifikasi, terutama ketika terdapat ketidakseimbangan data antar kelas.

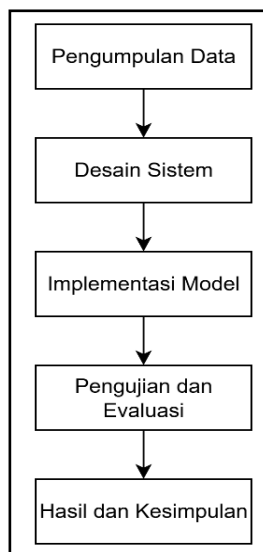
Dengan menggunakan metrik-metrik evaluasi tersebut, penelitian ini dapat menilai sejauh mana Algoritma *Random Forest* mampu melakukan klasifikasi prioritas perbaikan jalan dengan baik. Evaluasi ini juga memberikan gambaran mengenai kelebihan maupun kelemahan model, sehingga hasil penelitian dapat dipertanggungjawabkan secara akademis maupun praktis.

BAB III

DESAIN DAN IMPLEMENTASI

3.1 Desain Penelitian

Desain penelitian ini merupakan tahap menjelaskan secara rinci metode yang digunakan dalam penelitian untuk melakukan klasifikasi prioritas perbaikan jalan di Kabupaten Malang dengan menggunakan Algoritma *Random Forest*. Alur penelitian ini disusun secara sistematis mulai dari pengumpulan data, perancangan sistem, pengujian serta evaluasi, hingga penarikan kesimpulan. Adapun alur penelitian dapat dilihat pada Gambar 3.1.



Gambar 3.1 Alur Penelitian

3.2 Pengumpulan Data

Data yang digunakan dalam penelitian ini diperoleh secara resmi dari Dinas Pekerjaan Umum dan Penataan Ruang (PUPR) Kabupaten Malang. Data ini bersifat internal dan tidak tersedia secara umum ataupun secara daring, sehingga hanya dapat diakses melalui permohonan resmi kepada instansi terkait. Terdapat dua jenis

data utama yang digunakan, yaitu data kondisi jalan serta data Lalu Lintas Harian Rata-rata (LHR).

Data kondisi jalan memuat informasi mengenai ruas jalan di Kabupaten Malang, yang mencakup nama ruas jalan, panjang jalan hasil survei dalam kilometer, lebar jalan dalam meter, serta kondisi jalan berdasarkan *Surface Distress Index (SDI)* yang terbagi menjadi kategori baik, sedang, rusak ringan, dan rusak berat.

Sementara itu, data LHR berisi informasi mengenai volume lalu lintas rata-rata harian pada tiap ruas jalan. *Dataset* ini memuat nomor ruas, nama ruas jalan, fungsi jalan, panjang dan lebar jalan, serta LHR total yang menggambarkan jumlah kendaraan yang melintasi ruas tersebut setiap harinya. Data LHR menjadi penting karena menunjukkan tingkat kepadatan lalu lintas yang melengkapi informasi kondisi jalan, sehingga dapat digunakan sebagai salah satu indikator dalam menentukan tingkat prioritas perbaikan jalan.

Dengan adanya data mengenai kondisi fisik jalan sekaligus volume lalu lintas, proses klasifikasi menggunakan Algoritma *Random Forest* dapat dilakukan dengan lebih akurat, karena mempertimbangkan aspek kualitas infrastruktur maupun intensitas penggunaan jalan. Rincian lebih lanjut mengenai *dataset* dapat dilihat pada Table 3.1.

Tabel 3. 1 Contoh Dataset Tingkat Kerusakan Jalan

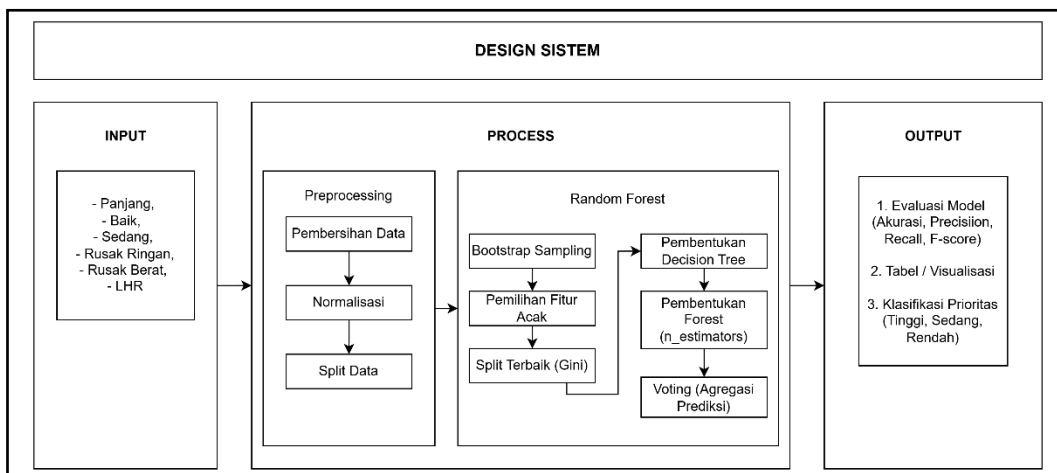
Nama Ruas	Panjang (km)	Baik (km)	Sedang (km)	Rusak Ringan (km)	Rusak Berat (km)	LHR	Kategori Jalan
Kepanjen - Pagak	161.0	161.0	0.0	0.0	0.0	412.90	Rusak Sedang
Ngadilangkung - Ngajum	345.0	345.0	0.0	0.0	0.0	1028.80	Rusak Sedang
Mojosari - Dilem	343.0	343.0	0.0	0.0	0.0	527.40	Rusak Sedang
Talangagung - Ngajum	45.0	39.0	6.0	0.0	0.0	392.20	Rusak Berat
Kepanjen - Kendalpayak	1178.0	1178.0	0.0	0.0	0.0	1288.50	Rusak Sedang
Jalan Lingkar Barat Kepanjen	454.0	454.0	0.0	0.0	0.0	1172.40	Rusak Ringan
Pakisaji - Karangduren	31.0	31.0	0.0	0.0	0.0	337.20	Rusak Sedang
Pakisaji - Maduarjo	1382.0	1192.0	0.0	0.0	0.0	191.80	Rusak Sedang
Permanu - Babadan	31.0	31.0	0.0	0.0	0.0	143.50	Rusak Sedang
Kebonagung - Sidorahayu	405.0	325.0	0.0	0.0	0.0	367.90	Rusak Sedang
Kapurono - Permanu	353.0	164.0	0.0	0.0	0.0	169.80	Rusak Sedang
Mendalanwangi - Pakisaji	63.0	61.0	1.0	0.0	0.0	131.00	Rusak Ringan
Wagir - Gunung Kawi	192.0	192.0	0.0	0.0	0.0	71.80	Rusak Sedang
Jedong (Bts. Kota Malang) - Pandanrejo	151.0	143.0	2.0	0.0	0.0	162.40	Rusak Sedang
Pandanlandung (Bts. Kota Malang) - Kalisongo	34.0	34.0	0.0	0.0	0.0	241.30	Rusak Sedang
Pandanlandung - Gondowangi	35.0	35.0	0.0	0.0	0.0	172.40	Rusak Sedang
Mendalanwangi - Sumbersuko	722.0	331.0	4.0	0.0	0.0	103.50	Rusak Berat
Kesamben - Balesari	96.0	93.0	0.0	0.0	0.0	110.70	Rusak Ringan

Dalam penelitian ini, tidak seluruh atribut yang terdapat pada *data* digunakan dalam proses klasifikasi. Peneliti hanya memilih atribut-atribut yang dianggap paling relevan dengan tujuan penelitian, yaitu menentukan prioritas perbaikan jalan. Dari dataset kondisi jalan, fitur yang digunakan meliputi panjang jalan (km), serta distribusi kondisi jalan berdasarkan kategori baik, sedang, rusak ringan, dan rusak berat. Keempat kategori kondisi ini memberikan gambaran tingkat kerusakan ruas jalan yang menjadi indikator utama dalam menentukan kebutuhan perbaikan.

Selain itu, dari data LHR (Lalu Lintas Harian Rata-rata) dipilih fitur LHR total sebagai representasi intensitas lalu lintas pada tiap ruas jalan. Fitur ini penting karena semakin tinggi volume kendaraan yang melintas, semakin besar pula tingkat beban yang memengaruhi kecepatan kerusakan jalan. Dengan demikian, total terdapat enam fitur yang digunakan dalam pembangunan model *Random Forest*, yaitu panjang jalan, kondisi baik, kondisi sedang, kondisi rusak ringan, kondisi rusak berat, serta LHR. Keenam fitur ini diolah untuk menghasilkan model klasifikasi prioritas perbaikan jalan dengan kategori tinggi, sedang, dan rendah.

3.3 Design Sistem

Desain sistem dalam penelitian ini mencakup beberapa tahapan yang telah digambarkan pada Gambar 3.2.



Gambar 3.2 Diagram Tahapan Design Sistem

Pada gambar 3.2 menunjukkan Desain sistem pada penelitian ini dimana menggambarkan alur kerja klasifikasi prioritas perbaikan jalan dengan menggunakan Algoritma *Random Forest*. Sistem dibangun berdasarkan tiga komponen utama, yaitu input, proses, dan output. Pada bagian input, terdapat dua sumber data yang digunakan, yaitu data kondisi jalan yang mencakup informasi panjang, kondisi jalan (baik, sedang, rusak ringan, dan rusak berat). Selain itu, digunakan pula data Lalu Lintas Harian Rata-rata (LHR) yang memberikan gambaran mengenai intensitas kendaraan pada setiap ruas jalan.

Tahap proses meliputi serangkaian langkah *preprocessing* seperti pembersihan data, penggabungan *dataset*, normalisasi, serta pembagian data menjadi data latih dan data uji (Dineshkumar et al., 2024). Setelah data siap, proses dilanjutkan dengan penerapan Algoritma *Random Forest* untuk melakukan

klasifikasi prioritas. Model yang dihasilkan kemudian disimpan dalam format *file* Python. Bagian output dari sistem ini berupa hasil klasifikasi prioritas perbaikan jalan dengan tiga kategori, yaitu tinggi, sedang, dan rendah. Selain itu, sistem juga menghasilkan tabel atau visualisasi distribusi prioritas untuk memudahkan analisis, serta laporan evaluasi model berupa metrik akurasi, *precision*, *recall*, dan *f-score*.

Dengan desain sistem ini, penelitian tidak hanya menghasilkan model *machine learning*, tetapi juga membangun prototipe aplikasi berbasis web yang dapat berfungsi sebagai sistem pendukung keputusan bagi Dinas PUPR Kabupaten Malang dalam menentukan prioritas perbaikan jalan.

Tahap *preprocessing* merupakan langkah awal yang dilakukan sebelum data digunakan dalam pemodelan. Tujuan utama *preprocessing* adalah mempersiapkan *dataset* agar dalam kondisi bersih, konsisten, dan sesuai format yang dapat diproses oleh Algoritma *Random Forest* (Dineshkumar et al., 2024). Dengan *preprocessing*, kualitas data dapat ditingkatkan sehingga model tidak bias dan hasil klasifikasi menjadi lebih akurat.

Preprocessing pada penelitian ini terdiri dari beberapa tahap, yaitu pembersihan data, normalisasi, dan pembagian data (*split data*). Tahapan tersebut disusun untuk memastikan setiap data yang masuk ke dalam model benar-benar terolah dengan baik dan siap untuk digunakan pada proses klasifikasi.

3.3.1 Pembersihan Data

Pembersihan data merupakan tahap awal dalam *preprocessing* yang bertujuan untuk memastikan bahwa data yang digunakan tidak mengandung

kesalahan ataupun ketidaksesuaian. Pada tahap ini, dilakukan identifikasi dan penanganan terhadap beberapa permasalahan umum, seperti:

1. Data kosong (*missing values*) → misalnya ada ruas jalan yang tidak memiliki catatan LHR atau panjang jalan yang tidak tercatat. Nilai kosong dapat ditangani dengan cara dihapus jika jumlahnya sedikit dan tidak signifikan.
2. Data duplikat → terkadang satu ruas jalan tercatat lebih dari sekali. Hal ini dapat menyebabkan bias dalam model sehingga harus dihapus.
3. Data tidak konsisten → misalnya perbedaan format penulisan nama ruas jalan, satu ditulis “Kepanjen – Pagak” dan lainnya “Kepanjen-Pagak”. Data semacam ini perlu diseragamkan agar tidak menimbulkan masalah saat penggabungan *dataset*.
4. *Outlier* → Outlier merupakan data yang nilainya jauh berbeda dari pola umum, misalnya nilai LHR yang sangat tinggi tidak sesuai realitas. Dalam penelitian ini, outlier yang ditemukan akan dihapus agar tidak memengaruhi hasil klasifikasi.

Dengan adanya pembersihan data, *dataset* yang digunakan dalam penelitian ini menjadi lebih akurat, konsisten, dan dapat diandalkan. Tahap ini sangat penting karena kualitas data input akan sangat mempengaruhi kualitas output yang dihasilkan oleh Algoritma *Random Forest* (Borrouh et al., 2023).

3.3.2 Normalisasi

Normalisasi merupakan proses skala ulang data numerik agar berada dalam rentang tertentu sehingga nilai antar variabel menjadi lebih seimbang. Normalisasi

diperlukan karena pada *dataset* penelitian ini terdapat beberapa atribut numerik dengan skala yang berbeda, misalnya panjang jalan dalam kilometer, lebar jalan dalam meter, dan volume lalu lintas dalam satuan kendaraan per hari. Perbedaan skala ini dapat menyebabkan salah satu atribut mendominasi proses perhitungan pada metode *machine learning*, meskipun secara logis tidak selalu lebih penting.

Salah satu metode yang umum digunakan adalah *Min-Max Normalization*, yaitu mengubah setiap nilai asli (x) ke dalam rentang $[0-1]$ berdasarkan nilai minimum (x_{min}) dan maksimum (x_{max}) dalam suatu kolom data. Hasil normalisasi (x') membuat skala data menjadi seragam, sehingga mencegah perbedaan skala antar fitur memengaruhi proses pembelajaran model.

Contoh perhitungan normalisasi dapat dilihat dari tabel 3.2 dimana mengambil fitur panjang dan LHR.

Tabel 3. 2 Contoh Fitur Untuk Normalisasi

Nama Ruas	Panjang (km)	LHR
Jalan Lingkar Barat Kepanjen	4.54	12,885
Permanu – Babadan	3.10	1,435
Kepanjen – Kendalpayak	11.78	3,922
Gondanglegi – Bululawang	8.65	6,780
Pagelaran – Bantur	5.12	2,145

Setelah mengetahui nilai fitur panjang dan LHR, kemudian menghitung x_{min} dan x_{max} .

$$\text{Panjang} = x_{min} = 3.10$$

$$x_{max} = 11.78$$

$$\text{LHR} = x_{min} = 1.435$$

$$x_{max} = 12.885$$

Kemudian menghitung semua fitur menggunakan rumus normalisasi.

Sebagai contoh untuk Jalan Lingkar Barat Kepanjen.

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}}$$

$$x' = \frac{454 - 3,10}{11,78 - 3,10} = 0,166$$

Perhitungan tersebut menjadi contoh normalisasi fitur Panjang untuk Jalan Lingkar Barat Kepanjen. Berikut adalah contoh perhitungan untuk fitur LHR.

$$x' = \frac{12,885 - 1,435}{12,885 - 1,435} = 1,0$$

Dari kedua contoh perhitungan normalisasi tersebut dapat disimpulkan bahwa hasil normalisasi Jalan Lingkar Barat Kepanjen yaitu pada fitur panjang menghasilkan 1.166, sementara pada fitur LHR hasil normalisasinya adalah 1.0, begitu juga untuk jalan yang lain. Berikut contoh hasil normalisasi fitur panjang dan LHR.

Tabel 3. 3 Contoh Hasil Normalisasi

Nama Ruas	Panjang (km)	Panjang Normalisasi	LHR	LHR Normalisasi
Jalan Lingkar Barat Kepanjen	4.54	0.166	12.885	0.989
Permanu – Babadan	3.10	0.000	1.435	0.324
Kepanjen – Kendalpayak	11.78	1.000	3.922	0.218
Gondanglegi – Bululawang	8.65	0.639	6.780	0.467
Pagelaran – Bantur	5.12	0.233	2.145	0.062

Pada tabel 3.3 menunjukkan bahwa semua nilai dari fitur Panjang Jalan dan LHR telah diubah ke dalam rentang $[0,1]$ menggunakan metode *Min-Max Normalization*. Proses ini membuat skala antar fitur menjadi seragam sehingga tidak ada fitur yang mendominasi model akibat perbedaan satuan.

3.3.3 Split Data

Tahap terakhir dalam preprocessing adalah split data, yaitu proses membagi *dataset* menjadi dua bagian: data latih (*training set*) dan data uji (*testing set*). Tujuan dari pembagian ini adalah agar model dapat dilatih menggunakan sebagian data, kemudian diuji menggunakan data yang berbeda untuk mengevaluasi kinerjanya secara objektif. Dengan demikian, hasil evaluasi tidak bias karena model diuji pada data yang belum pernah dilihat sebelumnya.

Proporsi yang digunakan dalam penelitian ini adalah 70% untuk data latih dan 30% untuk data uji. Data latih digunakan untuk membangun dan menyesuaikan Algoritma *Random Forest*, sedangkan data uji digunakan untuk mengukur performa model dengan menghitung metrik evaluasi seperti akurasi, *precision*, *recall*, dan *f1-score*.

Dengan adanya split data, proses pelatihan dan pengujian dapat berjalan secara terstruktur, sehingga hasil yang diperoleh dari Algoritma *Random Forest* benar-benar menggambarkan kemampuan model dalam melakukan klasifikasi prioritas perbaikan jalan.

3.4 Implementasi *Random Forest*

Random Forest merupakan algoritma *ensemble learning* berbasis *Decision Tree* yang bekerja dengan cara membangun sejumlah pohon keputusan, kemudian menggabungkan hasil prediksi dari setiap pohon untuk menghasilkan klasifikasi akhir. Keunggulan utama *Random Forest* adalah kemampuannya mengurangi risiko overfitting yang sering terjadi pada *Decision Tree* tunggal, serta menghasilkan prediksi yang lebih stabil dan akurat.

Dalam penerapannya, terdapat beberapa parameter penting yang perlu ditentukan sebelum model dibangun, antara lain jumlah pohon (*n_estimators*), kedalaman maksimum pohon (*max_depth*), jumlah fitur yang dipilih pada setiap pemisahan (*max_features*), dan jumlah sampel minimum pada setiap node (*min_samples_split* dan *min_samples_leaf*). Penentuan parameter ini berpengaruh terhadap performa model, sehingga perlu dilakukan penyesuaian atau *tuning* agar diperoleh hasil terbaik (Rahman et al., 2025).

3.5.1 *Bootstrap Sampling*

Tahap pertama dalam penerapan *Random Forest* adalah *bootstrap sampling*, yaitu teknik pengambilan sampel data secara acak dengan pengembalian (*sampling with replacement*). Artinya, setiap data yang sudah terpilih tetap memiliki kemungkinan untuk terpilih kembali pada proses pengambilan berikutnya (Karabadji et al., 2023). Proses ini bertujuan untuk menghasilkan beberapa subset data yang berbeda dari *dataset* latih, sehingga setiap pohon keputusan dalam *Random Forest* dapat dilatih pada kombinasi data yang bervariasi.

Dalam penelitian ini, data gabungan antara data kondisi jalan dan data lalu lintas harian rata-rata (LHR) digunakan sebagai dasar pembentukan subset. Melalui *bootstrap sampling*, sebagian data dipilih secara acak dengan jumlah yang sama dengan data asal, sehingga pada subset hasil sampling terdapat kemungkinan adanya duplikasi data maupun data yang tidak terpilih sama sekali.

Sebagai contoh, Tabel 3.4 berikut memperlihatkan hasil *bootstrap sampling* dari lima data pertama pada *dataset* gabungan. Dari tabel dapat terlihat bahwa ruas jalan “permanu – babadan” muncul lebih dari sekali akibat mekanisme pengambilan dengan pengembalian. Hal ini menunjukkan bahwa setiap subset hasil *bootstrap sampling* memiliki karakteristik yang unik, yang kemudian akan digunakan untuk membangun satu pohon keputusan dalam *Random Forest*.

Tabel 3. 4 Contoh Bootstrap Sampling

Nama Ruas	Panjang (km)	Lebar (m)	Baik (km)	Sedang (km)	Rusak Ringan (km)	Rusak Berat (km)	No Ruas	LHR Total
Jalan lingkar barat kepanjen	4.54	12.00	4.54	0	0	0	35	12885.0
permanu - babadan	3.10	3.50	3.10	0	0	0	35	1435.0
kepanjen - kendalpayak	11.78	6.03	11.78	0	0	0	35	3922.0
permanu - babadan	3.10	3.50	3.10	0	0	0	35	1435.0
permanu - babadan	3.10	3.50	3.10	0	0	0	35	1435.0

3.5.2 Pemilihan Fitur Secara Acak

Setelah terbentuk subset data melalui *bootstrap sampling*, tahap selanjutnya adalah melakukan pemilihan fitur secara acak (*random feature selection*). Pada

tahap ini, Algoritma *Random Forest* tidak langsung menggunakan seluruh fitur yang ada dalam *dataset* untuk membangun node pada *decision tree*. Sebaliknya, hanya sebagian fitur yang dipilih secara acak pada setiap node untuk mencari pemisahan (split) terbaik. Pendekatan ini dikenal dengan *feature randomness*, yang bertujuan agar setiap pohon memiliki keragaman dan tidak hanya bergantung pada fitur tertentu saja.

Sebagai contoh pada *dataset* terdapat 36 fitur, dimana 36 fitur ini yang menjadi n pada rumus diatas, dengan begitu di peroleh hasil sebagai berikut:

$$M = \sqrt{n}$$

$$M = \sqrt{36}$$

$$M = 6$$

Hasil dari contoh perhitungan diatas dapat disimpulkan bahwa dari 36 fitur dalam dataset, hanya 6 saja yang terpilih secara acak. Contoh hasil pemilihan fitur secara acak dapat dilihat pada Tabel 3.5.

Tabel 3. 5 Contoh Pemilihan Fitur Secara Acak

Posisi	Fitur yang Terpilih
1	Panjang (km)
2	LHR Total
3	Baik
4	Sedang
5	Rusak Ringan
6	Rusak Berat

3.5.3 Pemilihan Split Terbaik

Setelah fitur dipilih secara acak, tahap berikutnya adalah menentukan split terbaik pada node *decision tree*. Split terbaik dalam klasifikasi adalah titik pemisahan yang menghasilkan node anak dengan tingkat kemurnian (homogenitas) paling tinggi. Dengan kata lain, data dalam node anak hasil pemisahan sebaiknya memiliki kelas target yang lebih jelas dan dominan.

Dalam penelitian ini, kriteria yang digunakan adalah Gini Index, karena perhitungannya lebih sederhana dibandingkan *Entropy* namun tetap efektif untuk mengukur *impurity* (ketidakmurnian) data. Nilai Gini Index digunakan untuk mengukur tingkat impurity atau ketidakmurnian suatu node. Semakin kecil nilai Gini, semakin homogen data dalam node tersebut, sehingga pemisahan (split) dianggap lebih baik. Perhitungan dilakukan dengan menjumlahkan kuadrat proporsi setiap kelas pada node, kemudian dikurangkan dari 1.

Sebagai contoh perhitungan pada Tabel 3.6 yaitu memkai 5 potongan baris data dengan dikelompokkan menjadi 3 kelas tinggi, sedang dan rendah, dimana menggunakan 2 fitur yaitu LHR dan Rusak Berat.

Tabel 3. 6 Contoh Data Untuk Split Terbaik

Ruas	LHR	Rusak Berat	Kelas
A (Jln Lingkar Barat)	12 885	0	Rendah
B (Permanu–Babadan)	1 435	0	Rendah
C (Kepanjen–Kendalpayak)	3 922	0	Sedang
D (Talangagung–Ngajum)	6 120	1	Tinggi
E (Singosari–Lawang)	4 500	0	Sedang

Kandidat 1: Fitur LHR (threshold contoh $t = 5\,000$)

Left ($\leq 5\ 000$): B(R), C(S), E(S) $\rightarrow$ R=1, S=2, T=0, $N_L = 3$

$$Gini(L) = 1 - ((\frac{1}{3})^2 + (\frac{2}{3})^2) = 1 - (\frac{1}{9} + \frac{4}{9}) = \frac{4}{9} \approx 0.444$$

Right ($> 5\ 000$): A(R), D(T) $\rightarrow$ R=1, T=1, $N_R = 2$

$$Gini(R) = 1 - ((\frac{1}{2})^2 + (\frac{1}{2})^2) = 0.5$$

$$Gini(t) = 1 - \sum_{i=1}^c p_i^2$$

Gini tertimbang:

$$Gini_{split} = \frac{3}{5} \times 0.444 + \frac{2}{5} \times 0.5 = 0.2664 + 0.2 = 0.4664$$

Kandidat 2: Fitur Rusak Berat (biner: 1/0)

Left (RusakBerat = 0): A(R), B(R), C(S), E(S) $\rightarrow$ R=2, S=2, $N_L = 4$

$$Gini(L) = 1 - ((\frac{2}{4})^2 + (\frac{2}{4})^2) = 1 - (\frac{1}{4} + \frac{1}{4}) = 0.5$$

Right (RusakBerat = 1): D(T) saja $\rightarrow$ T=1, $N_R = 1$

$$Gini(R) = 0$$

Gini tertimbang:

$$Gini_{split} = \frac{4}{5} \times 0.5 + \frac{1}{5} \times 0 = 0.4$$

Pada contoh ini, fitur “Rusak Berat” memberikan split terbaik (*impurity* terendah), sehingga dipilih untuk node tersebut. Semakin kecil nilai Gini, semakin murni node tersebut, artinya data dalam node sudah lebih terpisah berdasarkan

kelasnya. Langkah yang sama diulang pada node-node berikutnya (dengan subset fitur acak sesuai aturan Random Forest).

3.5.4 Membangun *Decision Tree*

Setelah dilakukan *bootstrap sampling*, pemilihan fitur secara acak, dan penentuan split terbaik, tahap berikutnya adalah membangun *decision tree*. Pada tahap ini, data subset hasil bootstrap digunakan untuk membentuk struktur pohon keputusan berdasarkan fitur yang dipilih secara acak dan split terbaik yang telah dihitung.

Proses pembangunan decision tree dilakukan secara rekursif, dimulai dari node akar (*root node*) hingga terbentuk node daun (*leaf node*). Pada setiap node, algoritma memilih subset fitur acak, lalu menentukan split terbaik menggunakan kriteria Gini Index. Node kemudian dibagi menjadi dua atau lebih node anak sesuai hasil split. Proses ini terus diulang sampai salah satu kondisi berhenti terpenuhi,

Hasil akhir dari proses ini adalah sebuah decision tree yang memiliki jalur keputusan dari *root* hingga *leaf*. Setiap *leaf node* merepresentasikan hasil klasifikasi, misalnya *Prioritas Tinggi*, *Prioritas Sedang*, atau *Prioritas Rendah*.

Sebagai contoh dengan lima baris data di peroleh pada tabel 3.7.

Tabel 3. 7 Contoh Data Untuk Membuat Decision Tree

Ruas	LHR	Rusak Berat	Panjang (km)	Kelas
A	12 885	0	4.54	Rendah
B	1 435	0	3.10	Rendah
C	3 922	0	11.78	Sedang
D	6 120	1	7.8	Tinggi
E	4 500	0	10.00	Sedang

Pilih Split Akar

Kandidat A: LHR (ambang 5 000)

Left ($\leq 5\ 000$): B(R), C(S), E(S) $\rightarrow R=1, S=2 \rightarrow Gini = 1 - (\frac{1^2}{3} + \frac{2^2}{3}) = 0.444$

Right ($> 5\ 000$): A(R), D(T) $\rightarrow R=1, T=1 \rightarrow Gini = 0.5$

Gini tertimbang: $\frac{3}{5} \cdot 0.444 + \frac{2}{5} \cdot 0.5 = 0.4664$

Kandidat B: Rusak Berat (biner)

Left (RB=0): A(R), B(R), C(S), E(S) $\rightarrow R=2, S=2 \rightarrow Gini = 0.5$

Right (RB=1): D(T) $\rightarrow Gini = 0$

Gini tertimbang: $\frac{4}{5} \cdot 0.5 + \frac{1}{5} \cdot 0 = 0.4$

Pilih “Rusak Berat” di akar karena $0.4 < 0.4664$.

Split Anak Kiri (Rb=0)

Subdata: A(R), B(R), C(S), E(S) (4 baris).

Coba fitur Panjang (km) dengan ambang 6.0

Left (≤ 6.0): A(4.54,R), B(3.10,R) $\rightarrow R=2 \rightarrow Gini = 0$

Right (> 6.0): C(11.78,S), E(10.00,S) $\rightarrow S=2 \rightarrow Gini = 0$

Gini tertimbang: $\frac{2}{4} \cdot 0 + \frac{2}{4} \cdot 0 = 0 \leftarrow \text{sempurna}$

(Alternatif LHR $\leq 5\ 000$ pada subdata menghasilkan $Gini_{split} = \frac{3}{4} \cdot 0.444 + \frac{1}{4} \cdot$

$0 = 0.333 \rightarrow$ masih lebih buruk daripada Panjang.)

Pilih “Panjang ≤ 6.0 km” untuk anak kiri.

Hasil node & label daun (mayoritas)

Root: jika Rusak Berat = 1 $\rightarrow$ Tinggi (node kanan sudah murni).

Jika Rusak Berat = 0 (anak kiri):

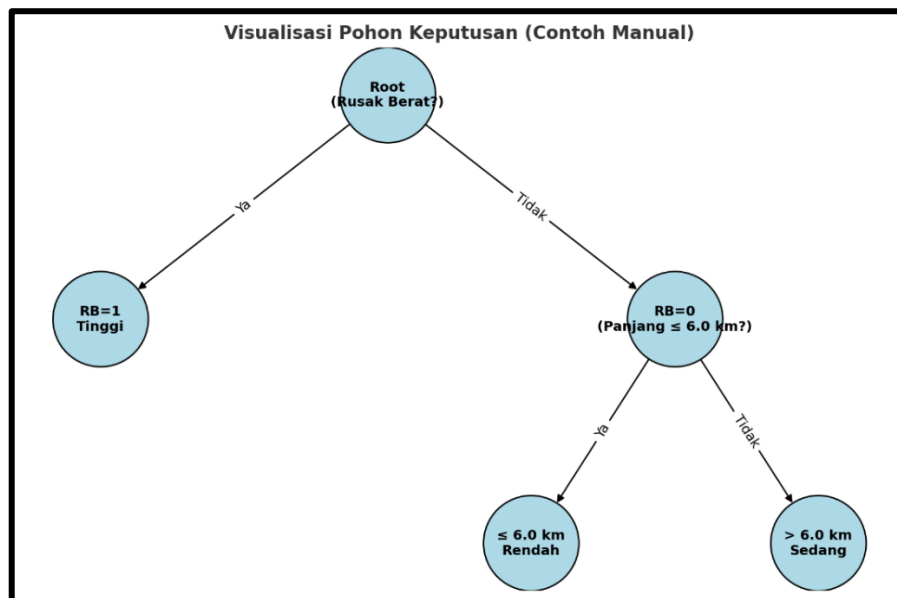
- Panjang ≤ 6.0 km $\rightarrow$ Rendah (A,B)
- Panjang > 6.0 km $\rightarrow$ Sedang (C,E)

Berdasarkan perhitungan manual dengan menggunakan kriteria Gini Index, pada tahap awal seluruh data memiliki nilai impurity sebesar 0.64. Dari enam fitur yang tersedia, secara acak dipilih dua fitur kandidat yaitu *Rusak Berat* dan *LHR*. Hasil perhitungan menunjukkan bahwa fitur *Rusak Berat* menghasilkan nilai Gini Split sebesar 0.40, lebih kecil dibandingkan dengan fitur *LHR* yang memiliki nilai 0.4664, sehingga dipilih sebagai split terbaik pada root node. Selanjutnya, pada cabang *Rusak Berat* = 0, pemisahan kembali dilakukan dengan fitur *Panjang Jalan* dengan ambang 6 km. Hasil perhitungan menunjukkan kedua subset data menjadi homogen sehingga nilai Gini mencapai 0. Dengan demikian, terbentuklah struktur pohon sederhana di mana ruas jalan dengan *Rusak Berat* = 1 diklasifikasikan sebagai Prioritas Tinggi, *Rusak Berat* = 0 dan *Panjang* ≤ 6 km dikategorikan sebagai Prioritas Rendah, sedangkan *Rusak Berat* = 0 dan *Panjang* > 6 km dikategorikan sebagai Prioritas Sedang. Berdasarkan hasil perhitungan tersebut dapat disimpulkan pada tabel 3.8 berikut.

Tabel 3. 8 Contoh Hasil Decission Tree

Node	Fitur & Ambang	Komposisi Kiri	Gini Kiri	Komposisi Kanan	Gini Kanan	Gini Tertimbang	Keputusan
Root	Rusak Berat (0/1)	R=2,S=2	0.50	T=1	0.00	0.40	Dipilih
Anak kiri (RB=0)	Panjang ≤ 6.0	R=2	0.00	S=2	0.00	0.00	Dipilih

Lebih mudahnya dapat dilihat pada gambar 3.3.



Gambar 3. 3 Contoh Visualisasi Decision Tree

3.5.5 Agregasi Prediksi

Tahap terakhir dalam *Random Forest* adalah agregasi prediksi, yaitu menggabungkan hasil prediksi dari seluruh *decision tree* yang telah dibangun. Setiap pohon pada *Random Forest* memberikan prediksi kelas berdasarkan data input yang diterimanya. Namun, karena setiap pohon dilatih dengan subset data bootstrap dan fitur yang berbeda, hasil prediksi antar pohon bisa bervariasi.

Dalam klasifikasi, agregasi dilakukan dengan metode *majority voting*. Artinya, kelas yang paling banyak diprediksi oleh pohon-pohon dalam *Random Forest* akan dipilih sebagai hasil akhir (Kim et al., 2020). Misalnya, jika terdapat 100 pohon dalam model dan 60 pohon memprediksi Prioritas Tinggi, 30 pohon

memprediksi Prioritas Sedang, dan 10 pohon memprediksi Prioritas Rendah, maka hasil klasifikasi akhir adalah Prioritas Tinggi.

Sebagai ilustrasi, Tabel 3.9 berikut memperlihatkan contoh dari lima pohon keputusan (T1–T5) untuk tiga ruas jalan yang diambil dari data yang digunakan.

Tabel 3. 9 Contoh Dari Lima Pohon Keputusan

Ruas Jalan	T1	T2	T3	T4	T5
Jalan Lingkar Barat Kepanjen	Tinggi	Tinggi	Sedang	Tinggi	Sedang
Permanu – Babadan	Rendah	Rendah	Sedang	Rendah	Rendah
Kepanjen – Kendalpayak	Sedang	Sedang	Sedang	Rendah	Sedang

Dari kelima pohon keputusan dapat diambil hasil voting atau keputusan yang memiliki suara terbanyak. Dilihat pada Tabel 3.9 pada data “Jalan Lingkar Barat Kepanjen”, pohon T1, T2, dan T4 memberikan prediksi kelas *Tinggi*, sedangkan T3 dan T5 memberikan prediksi kelas *Sedang*. Dengan demikian jumlah suara untuk *Tinggi* = 3, *Sedang* = 2, dan *Rendah* = 0, sehingga hasil akhir ditentukan oleh kelas dengan suara terbanyak, yaitu *Tinggi*. Proses serupa berlaku pada data lain, misalnya “Permanu – Babadan” yang memperoleh mayoritas suara Rendah, serta “Kepanjen – Kendalpayak” yang mayoritas pada *Sedang*. Berikut adalah tabel hasil agregrasi prediksi.

Tabel 3. 10 Contoh Hasil Agregrasi Prediksi

Ruas Jalan	Suara Tinggi	Suara Sedang	Suara Rendah	Kelas Akhir (mode)
Jalan Lingkar Barat Kepanjen	3	2	0	Tinggi
Permanu – Babadan	0	1	4	Rendah
Kepanjen – Kendalpayak	0	4	1	Sedang

Berdasarkan hasil klasifikasi pada tabel 3.10, Jalan Lingkar Barat Kepanjen termasuk dalam kategori prioritas tinggi, Permanu–Babadan termasuk kategori prioritas rendah, dan Kepanjen–Kendalpayak termasuk kategori prioritas sedang. Perbedaan hasil klasifikasi tersebut menunjukkan bahwa setiap ruas jalan memiliki karakteristik yang berbeda berdasarkan data kondisi jalan, panjang jalan, dan nilai LHR yang digunakan dalam proses klasifikasi. Dengan demikian, model Random Forest mampu membedakan tingkat prioritas perbaikan jalan sehingga dapat membantu proses pengambilan keputusan secara lebih objektif.

3.5 Evaluasi Model

Evaluasi model merupakan tahap penting untuk menilai sejauh mana Algoritma *Random Forest* yang dibangun mampu melakukan klasifikasi prioritas perbaikan jalan secara akurat. Tahap ini dilakukan dengan menguji model menggunakan data uji yang telah dipisahkan sebelumnya pada tahap *split data*. Melalui evaluasi, dapat diketahui keunggulan maupun kelemahan model dalam memprediksi kelas target (prioritas tinggi, sedang, dan rendah).

Pada penelitian ini, pengujian dilakukan menggunakan *Confusion Matrix*, yang berfungsi untuk membandingkan hasil prediksi model dengan data aktual. Melalui *Confusion Matrix*, diperoleh informasi mengenai jumlah prediksi benar dan salah pada setiap kelas, yang kemudian digunakan sebagai dasar perhitungan berbagai metrik evaluasi model. Adapun metrik yang digunakan meliputi akurasi, *precision*, *recall*, dan *f1-score*, yang umum diterapkan dalam kasus klasifikasi untuk mengukur tingkat performa dan ketepatan model.

Sebagai contoh perhitungan evaluasi model menggunakan confusion matrix dapat dilihat pada tabel 3.11 berikut

Tabel 3. 11 Contoh Confussion Matrix

	Pred T	Pred S	Pred R	Total Aktual
Akt T	8	1	1	10
Akt S	2	10	3	15
Akt R	1	2	12	15
Total Pred	11	13	16	40

Diagonal tebal = TP per kelas: $TP_T = 8$, $TP_S = 10$, $TP_R = 12$.

Jumlah baris = data aktual per kelas (support).

Jumlah kolom = data diprediksi per kelas.

Akurasi

$$\text{Accuracy} = \frac{\text{jumlah diagonal}}{\text{total}} = \frac{8 + 10 + 12}{40} = 0.75$$

Precision, Recall, F1 per kelas

Rumus:

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$F1 = \frac{2 \cdot P \cdot R}{P + R}$$

Kelas Tinggi (T)

- $P_T = 8/11 = 0.7273$
- $R_T = 8/10 = 0.8000$
- $F1_T = 0.7619$

Kelas Sedang (S)

- $P_S = 10/13 = 0.7692$
- $R_S = 10/15 = 0.6667$
- $F1_S = 0.7141$

Kelas Rendah (R)

- $P_R = 12/16 = 0.7500$
- $R_R = 12/15 = 0.8000$
- $F1_R = 0.7742$

Berdasarkan Confusion matrix pada Tabel 3.11 menunjukkan perbandingan antara data aktual dan hasil prediksi model untuk tiga kelas, yaitu Tinggi (T), Sedang (S), dan Rendah (R). Nilai pada diagonal utama (8, 10, dan 12) merupakan jumlah prediksi yang benar untuk masing-masing kelas (True Positive/TP), sedangkan nilai di luar diagonal menunjukkan kesalahan klasifikasi. Jumlah seluruh data adalah 40, dengan total prediksi benar sebanyak 30, sehingga diperoleh nilai akurasi sebesar 0.75. Secara logika, akurasi ini menunjukkan bahwa 75% data berhasil diklasifikasikan dengan benar oleh model, sedangkan 25% lainnya mengalami kesalahan prediksi. Distribusi kesalahan juga dapat dilihat, misalnya pada kelas Sedang yang memiliki kesalahan cukup tinggi ke kelas lain (2 ke Tinggi

dan 3 ke Rendah), yang mengindikasikan adanya kemiripan karakteristik antar kelas tersebut dalam data.

Perhitungan precision, recall, dan F1-score dilakukan untuk mengevaluasi performa model secara lebih spesifik pada tiap kelas. Precision mengukur ketepatan prediksi suatu kelas dengan membandingkan TP terhadap total prediksi pada kolom tersebut, misalnya pada kelas Tinggi diperoleh $8/11 = 0.7273$ yang berarti dari seluruh prediksi “Tinggi”, sekitar 72.73% benar. Recall mengukur kemampuan model dalam mendeteksi seluruh data aktual suatu kelas, misalnya kelas Sedang memiliki recall $10/15 = 0.6667$ yang menunjukkan masih ada data Sedang yang salah diklasifikasikan ke kelas lain. F1-score merupakan rata-rata harmonis antara precision dan recall, sehingga memberikan keseimbangan antara keduanya, seperti pada kelas Rendah dengan F1 sebesar 0.7742. Secara keseluruhan, nilai-nilai ini menunjukkan bahwa model memiliki performa yang cukup baik, namun masih terdapat kesalahan klasifikasi terutama pada kelas dengan distribusi data yang lebih kompleks.

Dari hasil tersebut dapat disimpulkan bahwa Algoritma *Random Forest* mampu melakukan klasifikasi prioritas perbaikan jalan dengan tingkat akurasi yang baik serta distribusi performa yang relatif seimbang antar kelas. Meskipun demikian, masih terdapat beberapa kesalahan prediksi, khususnya pada kelas *Sedang*, yang menunjukkan perlunya perhatian lebih pada data dengan karakteristik menengah agar model dapat lebih optimal.

3.6 Skenario Uji Coba

Peneliti melakukan skenario uji coba untuk mengevaluasi performa model Random Forest yang telah dibangun dengan memvariasikan parameter *n_estimators* atau jumlah pohon keputusan. Parameter ini dipilih karena memiliki pengaruh penting terhadap kinerja model dalam proses klasifikasi. Secara umum, semakin banyak jumlah pohon yang digunakan dalam Random Forest, maka hasil prediksi akan menjadi lebih stabil karena keputusan akhir diperoleh dari hasil voting lebih banyak pohon. Namun demikian, peningkatan jumlah pohon juga akan berdampak pada bertambahnya waktu komputasi yang dibutuhkan dalam proses pelatihan model.

Dalam penelitian ini, dilakukan pengujian dengan beberapa variasi jumlah pohon yaitu 2, 4, 6, 8, 10, 12, 14, 16, 18, dan 20. Variasi tersebut dipilih untuk melihat bagaimana perubahan performa model pada rentang jumlah pohon yang masih relatif kecil hingga menengah. Dengan menggunakan rentang tersebut, peneliti dapat mengamati pola peningkatan atau stabilitas performa model secara bertahap, serta menentukan apakah penambahan jumlah pohon masih memberikan dampak signifikan terhadap peningkatan akurasi atau justru sudah mencapai titik stabil.

Pada skenario pertama dengan 2 pohon, model dilatih dengan jumlah yang sangat sedikit sehingga proses komputasi menjadi sangat cepat, namun hasil prediksi masih cenderung kurang stabil karena keputusan hanya dihasilkan dari sedikit pohon. Selanjutnya pada skenario dengan 4 dan 6 pohon, mulai terlihat adanya peningkatan stabilitas dalam hasil prediksi meskipun masih terdapat variasi

dalam performa model. Hal ini menunjukkan bahwa penambahan jumlah pohon mulai memberikan dampak terhadap kualitas klasifikasi.

Pada skenario berikutnya yaitu 8 hingga 10 pohon, model mulai menunjukkan keseimbangan antara kecepatan komputasi dan stabilitas hasil prediksi. Pada tahap ini, Random Forest sudah mampu menghasilkan keputusan yang lebih konsisten dibandingkan skenario sebelumnya. Kemudian pada rentang 12 hingga 16 pohon, performa model semakin stabil dengan tingkat akurasi yang cenderung meningkat, meskipun waktu komputasi mulai bertambah secara bertahap.

Selanjutnya pada skenario 18 hingga 20 pohon, model berada pada kondisi yang paling stabil dalam rentang pengujian ini. Hasil prediksi menjadi lebih konsisten karena jumlah pohon yang digunakan sudah cukup banyak untuk melakukan voting keputusan. Namun demikian, peningkatan akurasi yang dihasilkan tidak terlalu signifikan jika dibandingkan dengan skenario sebelumnya, sehingga dapat disimpulkan bahwa pada titik tertentu penambahan jumlah pohon tidak lagi memberikan peningkatan performa yang besar, tetapi hanya meningkatkan beban komputasi.

Seluruh skenario pengujian dilakukan dengan tahapan yang sama, dimulai dari proses preprocessing data, pembagian data menjadi data latih dan data uji, pelatihan model Random Forest sesuai dengan masing-masing nilai $n_estimators$, hingga tahap evaluasi menggunakan Confusion Matrix. Dari hasil Confusion Matrix tersebut diperoleh nilai akurasi, precision, recall, dan f1-score yang digunakan untuk membandingkan performa model pada setiap skenario yang diuji.

BAB IV

HASIL DAN PEMBAHASAN

4.1 Hasil Penerapan Model

Dalam studi ini, pengujian model diaplikasikan lewat teknik *Random Forest*. Algoritma tersebut dipakai guna mengelompokkan kondisi jalan memakai basis data yang sebelumnya sudah melewati tahapan pra-pemrosesan serta normalisasi. Tahapan klasifikasi ini diawali dengan mengelompokkan data terlebih dahulu menggunakan metode *K-Means Clustering* untuk melihat pola distribusi awal data berdasarkan karakteristik yang dimiliki. Hasil dari proses clustering ini kemudian digunakan sebagai gambaran awal sebelum data masuk ke tahap klasifikasi.

Selain itu, pada bagian ini juga ditampilkan distribusi data untuk memberikan gambaran mengenai sebaran masing-masing kelas dalam dataset yang digunakan. Di samping itu, disajikan pula hasil klasifikasi dari model *Random Forest* yang memuat rangkuman prediksi sekaligus komparasi realita di lapangan dengan estimasi sistem. Seluruh capaian data ini dimanfaatkan demi menyajikan performa utuh dari model yang dirancang pada riset ini.

Sementara itu, Tabel 4.1 menyajikan data dari Dinas PUPR Kabupaten Malang yang disiapkan sebelum nantinya masuk ke dalam tahapan pengelompokkan memakai *Random Forest*.

Tabel 4. 1 Pengelompokan Kategori Prioritas

Nama Ruas	Panjang	Baik	Sedang	Rusak Ringan	Rusak Berat	Lhr	Kategori Jalan
Kepanjen - Pagak	161.0	1610000038 14697.0	0.0	0.0	0.0	4129.0	Rusak Sedang
Ngadilangkung - Ngajum	345.0	3450000047 68372.0	0.0	0.0	0.0	10288.0	Rusak Sedang
Mojosari - Dilem	343.0	3430000066 7572.0	0.0	0.0	0.0	5274.0	Rusak Sedang
Talangagung - Ngajum	45.0	3900000095 36743.0	6000000238418 58.0	0.0	0.0	3922.0	Rusak Berat
Kepanjen - Kendalpayak	1178.0	1177999973 29712.0	0.0	0.0	0.0	12885.0	Rusak Sedang
Jalan Lingkar Barat Kepanjen	454.0	4539999961 85303.0	0.0	0.0	0.0	11724.0	Rusak Ringan
Pakisaji - Karangduren	31.0	3099999904 63257.0	0.0	0.0	0.0	3372.0	Rusak Sedang
Pakisaji - Maduarjo	1382.0	1192000007 62939.0	0.0	0.0	0.0	1918.0	Rusak Sedang
Permanu - Babadan	31.0	3099999904 63257.0	0.0	0.0	0.0	1435.0	Rusak Sedang

Tabel 4.1 menampilkan informasi mengenai jalur transportasi dari Dinas PUPR. Data tersebut dipakai sebagai basis data utama pada studi ini. Di dalamnya, termuat sejumlah parameter seperti nama lajur, jarak tempuh, situasi jalan (mulai dari kondisi prima, normal, hingga kerusakan ringan maupun parah), serta angka LHR (Lalu Lintas Harian Rata-rata). Nilai Lalu Lintas Harian Rata-rata ini menggambarkan total kendaraan yang melintas di area jalan terkait sepanjang 24 jam untuk semua tipe transportasi, sehingga bisa memberikan proyeksi umum terkait kepadatan lalu lintas. Lebih lanjut, informasi ini ditunjang dengan kolom klasifikasi jalan sebagai label kelas guna memetakan skala prioritas perbaikan, misalnya untuk area dengan tingkat kerusakan kecil, menengah, hingga parah.

Dari informasi tersebut, tampak bahwa sebaran kondisi jalan dan prioritasnya tidak merata. Mayoritas angka sampel berkumpul pada kelompok

kerusakan sedang jika dikomparasikan dengan kelompok lainnya. Fenomena tersebut mengindikasikan adanya isu imbalance data yang bisa memengaruhi performa sistem dalam mengelompokkan objek. Selain itu, dalam studi ini variabel situasi jalan dan LHR diposisikan sebagai parameter masukan (*input*). Sementara itu, kelompok jalan dijadikan acuan luaran (*output*) yang nantinya diolah via Random Forest guna menciptakan sistem pengelompokan yang akurat dalam memproyeksikan skala prioritas perbaikan jalan.

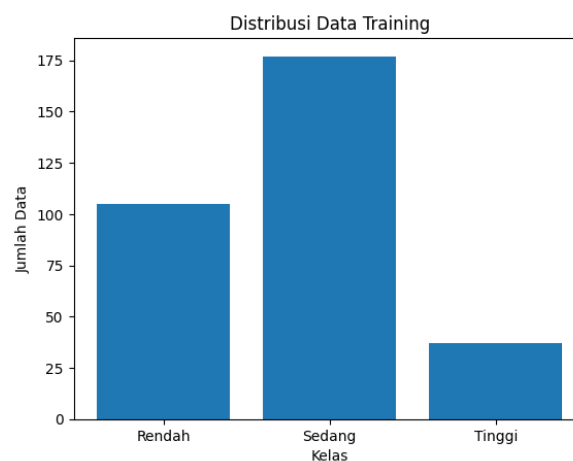
Tahap berikutnya yaitu menjalankan pengujian data memakai metode *Random Forest*. Rekapitan hasil tebakan sistem lalu disajikan untuk menggambarkan cara kerja algoritma dalam memetakan situasi jalan berdasarkan basis data yang telah diproses pada tahap awal.

Secara konsep, metode ini beroperasi lewat pembuatan banyak decision tree sekaligus menentukan keputusan akhir lewat sistem pemilihan suara terbanyak (*majority voting*). Melalui pendekatan matematis, estimasi target kelas $\hat{y}$ ditentukan oleh kelompok data yang paling dominan muncul dari seluruh pohon yang terbentuk, yaitu : $\hat{y} = \text{mode}(y_1, y_2, \dots, y_n)$. Dengan demikian, ketika sebagian besar pohon menghasilkan prediksi yang sama, maka output akhir akan konsisten. Kondisi ini terlihat pada data, di mana nilai aktual 0 (prioritas rendah), 1 (prioritas sedang), dan 2 (prioritas tinggi) seluruhnya berhasil diprediksi dengan tepat oleh model.

Keberhasilan prediksi ini juga dipengaruhi oleh karakteristik fitur yang digunakan, yaitu kondisi jalan dan nilai LHR, yang secara implisit membentuk pola pemisahan kelas pada setiap pohon keputusan. Dalam pembentukan pohon,

pemilihan atribut dilakukan berdasarkan pengukuran seperti *Gini Index* yang dirumuskan sebagai $Gini = 1 - \sum p_i^2$, di mana p_i adalah proporsi kelas pada suatu node. Atribut dengan nilai Gini terkecil akan dipilih karena mampu memisahkan data dengan lebih baik. Oleh karena itu, kombinasi antara kondisi jalan (baik, sedang, rusak) dan LHR memungkinkan model untuk membedakan tingkat prioritas secara jelas. Namun demikian, hasil yang sempurna ini juga dapat dipengaruhi oleh jumlah data uji yang terbatas atau distribusi data yang relatif mudah dipisahkan, sehingga perlu dilakukan evaluasi lanjutan untuk memastikan bahwa model tidak mengalami overfitting dan tetap memiliki kemampuan generalisasi yang baik.

Untuk memberikan gambaran yang lebih jelas mengenai sebaran data pada setiap kelas, maka dilakukan visualisasi dalam bentuk grafik distribusi data. Grafik distribusi data tersebut dapat dilihat pada Gambar 4.1.



Gambar 4. 1 Distribusi Data

Jika memperhatikan diagram 4.1, tampak jelas bahwa kuantitas sampel untuk kategori sedang mendominasi apabila dibandingkan dengan kelompok

rendah maupun tinggi. Ditinjau secara kuantitatif, sebaran ini mengonfirmasi fenomena ketidakseimbangan kelas (*imbalanced data*). Kondisi tersebut terjadi lantaran peluang kemunculan dari tiap-tiap kategori tidak selaras atau timpang. Jika dinyatakan dalam bentuk probabilitas, maka $p_i = \frac{n_i}{N}$, dengan n_i merepresentasikan kuantitas sampel pada kategori ke- i , sedangkan variabel N menunjukkan keseluruhan jumlah sampel yang ada. Besarnya nilai peluang pada kategori sedang ini mengakibatkan sistem lebih intensif mempelajari pola dari kelompok tersebut ketimbang kelompok lainnya. Dampaknya, proses penyusunan model akan terpengaruh karena algoritma cenderung berfokus menekan angka kekeliruan global tanpa memedulikan apakah sebaran kategorinya sudah proporsional atau belum.

Dalam konteks Algoritma *Random Forest*, ketidakseimbangan ini mempengaruhi proses pemilihan split pada setiap pohon keputusan yang menggunakan ukuran seperti *Gini Index* dengan rumus $Gini = 1 - \sum p_i^2$. Karena nilai p_i untuk kelas sedang lebih besar, maka kontribusinya dalam perhitungan Gini juga lebih dominan, sehingga pohon cenderung membentuk aturan yang mengarah pada pemisahan kelas sedang. Efeknya, sistem punya kecenderungan kuat untuk menebak kelompok dominan secara lebih akurat ketimbang kelompok yang jumlahnya sedikit. Implikasi dari situasi tersebut adalah angka akurasi bisa saja tampak melonjak tinggi, padahal parameter ukur lain layaknya precision, recall, maupun f1-score untuk kategori rendah dan tinggi justru anjlok. Maka dari itu, ketidakseimbangan sebaran sampel ini krusial sekali buat dijadikan bahan pertimbangan utama sewaktu mengukur kinerja sistem.

Meskipun demikian, model *Random Forest* masih mampu melakukan proses klasifikasi berdasarkan pola yang terdapat pada data yang tersedia. Guna mengukur tingkat akurasi estimasi yang dihasilkan oleh sistem secara lebih mendalam, maka dicari tahu tingkat kecocokan lewat komparasi antara angka tebakan sistem dengan realita data di lapangan.

Berdasarkan seluruh rangkaian ulasan mulai dari output pengelompokan, sebaran data, hingga komparasi target nyata dengan estimasi bisa ditarik kesimpulan jika metode *Random Forest* yang diterapkan pada riset ini sanggup memetakan situasi jalan dengan performa yang cukup oke. Temuan tersebut nantinya dijadikan landasan dalam melangkah ke tahapan peninjauan lanjutan di sub-bab berikutnya. Langkah ini diambil demi membedah efektivitas arsitektur model secara mendetail bersandarkan parameter pengujian berupa tingkat akurasi, *precision*, *recall*, serta *f1-score*.

4.2 Evaluasi Performa Model

Pada tahap ini dilakukan pengujian dilakukan untuk menakar kualitas dari model yang sebelumnya sudah dirancang memakai algoritma *Random Forest Classifier*. Agenda pengujian ini diorientasikan untuk memetakan seberapa tangguh kapabilitas sistem dalam mengelompokkan objek pada kategori prioritas kondisi jalan dengan mengacu pada basis data yang sudah melewati tahapan pra-pemrosesan di awal.

Proses peninjauan ini tak semata-mata bertumpu pada indikator akurasi sistem, namun turut melibatkan parameter ukur lain layaknya *precision*, *recall*, serta *f1-score*. Langkah ini diambil lantaran karakteristik sampel yang dipakai

mempunyai sebaran yang timpang (*imbalanced data*). Oleh sebab itu, tolok ukur akurasi tunggal dinilai kurang representatif untuk memetakan performa arsitektur model secara utuh dan menyeluruh.

Selain itu, pengujian model dilakukan melalui beberapa skenario percobaan untuk melihat pengaruh perubahan parameter terhadap hasil klasifikasi. Dengan adanya beberapa skenario ini, diharapkan dapat diperoleh konfigurasi model yang paling optimal dalam melakukan prediksi kelas prioritas jalan.

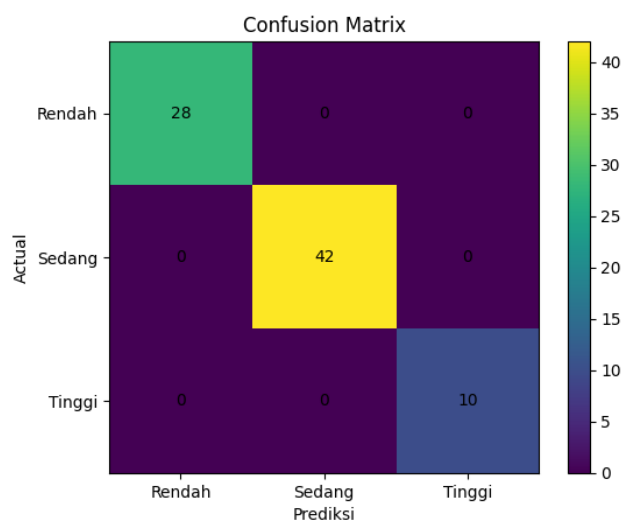
4.2.1 Evaluasi Kinerja Model Random Forest

Menguji performa sistem menjadi sebuah tahapan krusial dalam studi pengelompokan demi memetakan kapabilitas algoritma sewaktu membuat perkiraan dari sampel yang diuji. Lewat riset ini, pengujian difokuskan pada model *Random Forest* yang diaplikasikan untuk memetakan skala prioritas pembenahan jalan di wilayah Kabupaten Malang. Agenda peninjauan tersebut ditargetkan guna mengukur seberapa baik sistem dalam menangkap serta mengenali pola keterkaitan antara kondisi kerusakan jalan dan volume kendaraan terhadap penentuan kelas prioritasnya.

Setelah mengeksplorasi simulasi konfigurasi parameter *n_estimators*, agenda berikutnya yaitu membedah kualitas model paling optimal secara lebih mendalam. Proses ini memanfaatkan instrumen *confusion matrix* beserta parameter penilaian untuk tiap-tiap kelas yang ada. Melalui pengujian tersebut, diarahkan untuk mengukur kapabilitas sistem dalam memetakan tiap-tiap tingkatan prioritas jalan, meliputi skala Rendah, Sedang, hingga Tinggi

Penerapan instrumen ini krusial sekali mengingat potret capaian akurasi tunggal belum representatif dalam menyajikan kinerja sistem secara komprehensif. Lewat bantuan matriks ini, kuantitas tebakan yang tepat beserta titik kekeliruan pengelompokan pada tiap-tiap kategori dapat dipetakan secara gamblang. Dengan begitu, kita bisa menganalisis apakah performa model cenderung unggul pada kelompok tertentu atau justru kerap keliru saat memisahkan sejumlah tingkat prioritas yang ada.

Berikut merupakan hasil confusion matrix dari model *Random Forest* terbaik yang digunakan pada penelitian ini.



Gambar 4. 2 Hasil Akurasi Random Forest Berdasarkan n- estimators

Berdasarkan confusion matrix pada tabel 4.2, tampak jelas bahwa sistem menunjukkan kualitas pengelompokan yang sangat prima. Hal ini dibuktikan oleh dominasi angka estimasi yang tepat (*True Positive*) pada seluruh kelompok data. Pada kelas 0 (prioritas rendah), total 28 sampel sanggup dipetakan secara akurat oleh algoritma tanpa ditemukannya kekeliruan sedikit pun, baik berupa nilai FP (*False Positive*) ataupun FN (*False Negative*). Begitu pula pada kelas 1 (prioritas

sedang), di mana seluruh 42 data juga (FP maupun FN). Pada kelas 1 (prioritas sedang), seluruh 42 data juga diprediksi dengan tepat. Sedangkan pada kelas 2 (prioritas tinggi), tercatat sebanyak 9 sampel sanggup ditebak secara tepat, sementara hanya ada 1 sampel yang keliru dimasukkan ke dalam kelompok 1. Secara umum, kekeliruan pengelompokan hanya ditemukan sekali dari keseluruhan data, yang mengindikasikan bahwa sistem mempunyai derajat ketepatan yang amat tinggi.

Berdasarkan grafik *confusion matrix*, dominasi pola garis diagonal mengonfirmasi jika mayoritas sampel berkumpul pada area pengelompokan yang tepat. Kondisi ini terjadi lantaran angka paling signifikan menempati lajur diagonal utama matriks. Logikanya, apabila angka pada lajur utama tersebut kian mendominasi elemen lainnya, artinya kualitas prediksi sistem juga semakin tangguh. Sebaliknya, angka-angka yang melenceng di luar garis diagonal mengindikasikan adanya miscalculasi penempatan kelas, yang dalam situasi tersebut hanya terjadi satu kali pada transisi kategori 2 ke kategori 1. Fakta ini membuktikan jika arsitektur model hampir tidak menemukan kendala berarti dalam memisahkan tiap kelompok data, khususnya saat memetakan kategori rendah serta sedang.

Dengan demikian, berdasarkan analisis *confusion matrix* dan perhitungan metrik evaluasi, bisa ditarik kesimpulan bahwa algoritma Random Forest mempunyai kapabilitas pengelompokan yang luar biasa prima. Minimnya angka kekeliruan ini menjadi indikator kuat bahwa variabel-variabel yang dilibatkan sudah representatif dalam mengidentifikasi karakteristik tiap-tiap kelas. Namun,

tetap perlu diperhatikan bahwa performa tinggi ini juga harus dikaitkan dengan distribusi data dan jumlah data uji, sehingga interpretasi hasil tetap dilakukan secara objektif dan tidak hanya berdasarkan nilai metrik yang tinggi.

Dapat ditegaskan bahwa metode *Random Forest* memperlihatkan rekam jejak kinerja yang sangat optimal sewaktu memetakan urutan perbaikan jalan. Capaian angka *accuracy* yang mendekati poin 1 untuk seluruh kelompok data membuktikan bahwa model mampu mengidentifikasi kecenderungan data secara sangat akurat. Di samping itu, tingginya angka *precision* dan *recall* turut memperkuat bukti jika tingkat kekeliruan sistem tergolong amat minim. Oleh sebab itu, implementasi algoritma *Random Forest* dinilai sangat layak diaplikasikan demi mempermudah perumusan skala prioritas pembenahan jalan di wilayah Kabupaten Malang secara objektif sekaligus akurat.

4.2.1 Hasil Skenario Uji Coba Tiap *n*-estimator

Langkah ini berfokus pada pengujian arsitektur *Random Forest Classifier* lewat simulasi beberapa nilai parameter *n_estimators*. Indikator *n_estimators* tersebut merepresentasikan kuantitas dari *decision tree* yang dirancang sewaktu tahapan pengelompokan berjalan. Secara teoretis, bertambahnya kuantitas pohon yang dipakai sanggup meningkatkan kapabilitas sistem dalam menggeneralisasikan data, sebab penentuan *output* akhir bertumpu pada sistem voting dari kumpulan pohon tersebut.

Pengujian dilakukan melalui nilai *n_estimators* mulai dari 2 hingga 20 dengan interval kelipatan dua. Agar perbandingan hasil pengujian bersifat objektif, basis data training beserta data testing yang seragam diaplikasikan pada tiap

simulasi. Penilaian efektivitas sistem mengacu pada parameter berupa accuracy, precision, recall, serta F1-score. Dengan begitu, fokus peninjauan tidak Cuma melihat keakuratan estimasi saja, melainkan ikut memantau kestabilan kualitas pengelompokan di semua kategori data.

Berdasarkan pendekatan matematis, kalkulasi tingkat akurasi dicari melalui komparasi total tebakan yang tepat terhadap keseluruhan sampel pengujian, di mana nilai kombinasi *true positive* serta *true negative* dibagi dengan total data. Di sisi lain, nilai *precision* memetakan derajat ketepatan untuk perkiraan positif, sedangkan recall mencerminkan kapabilitas sistem dalam mengidentifikasi data positif riil. Sementara itu, parameter *f1-score* memuat nilai rata-rata harmonik dari kombinasi precision serta *recall*. Melalui karakteristik tersebut, kombinasi keempat indikator ini diimplementasikan demi menakar kualitas model secara menyeluruh.

a. Skenario Uji coba dengan n-estimators 2

Tabel 4. 2 Skenario Uji Coba n-estimators 2

n_estimators	Accuracy	Precision	Recall	F1-Score
2	0.85	0.8917	0.7778	0.8020

Pada tabel 4.2 skenario dengan jumlah pohon yang sangat sedikit, model belum mampu merepresentasikan keseluruhan pola dalam data secara optimal. Hal ini terlihat dari nilai akurasi yang masih relatif rendah serta adanya selisih yang cukup besar antara precision dan recall. Kondisi ini menunjukkan bahwa model cenderung belum stabil dalam mengenali seluruh kelas, khususnya dalam mengidentifikasi data yang seharusnya masuk ke kelas tertentu.

Ketidakseimbangan antara precision dan recall mengindikasikan bahwa model masih menghasilkan kesalahan dalam bentuk ketidakmampuan

menangkap seluruh variasi data. Hal ini wajar terjadi karena dengan jumlah pohon yang terbatas, variasi keputusan yang dihasilkan juga terbatas, sehingga model belum mampu menggeneralisasi pola data secara menyeluruh.

b. Skenario Uji coba dengan n-estimators 4

Tabel 4. 3 Skenario Uji Coba n-estimators 4

n_estimators	Accuracy	Precision	Recall	F1-Score
4	0.8875	0.9170	0.8151	0.8489

Pada tabel 4.3 $n_estimators = 4$, terlihat adanya peningkatan performa yang menunjukkan bahwa model mulai menangkap pola data dengan lebih baik. Penambahan jumlah pohon memberikan variasi keputusan yang lebih banyak, sehingga hasil prediksi menjadi lebih stabil dibandingkan sebelumnya.

Meskipun demikian, perbedaan antara precision dan recall masih terlihat, yang menandakan bahwa model belum sepenuhnya seimbang dalam mengklasifikasikan semua kelas. Hal ini menunjukkan bahwa model masih dalam tahap transisi menuju kondisi optimal, di mana peningkatan jumlah pohon masih sangat berpengaruh terhadap performa.

c. Skenario Uji coba dengan n-estimators 6

Tabel 4. 4 Skenario Uji Coba n-estimators 6

n_estimators	Accuracy	Precision	Recall	F1-Score
2	0.85	0.8917	0.7778	0.8020

Pada tabel 4.4 $n_estimators = 6$, model menunjukkan peningkatan yang signifikan dan mulai mendekati kondisi optimal. Hampir seluruh data berhasil diklasifikasikan dengan benar, yang menunjukkan bahwa model sudah mampu memahami pola utama dalam dataset.

Keseimbangan antara precision dan recall juga mulai terlihat, yang berarti model tidak hanya akurat tetapi juga konsisten dalam mengenali seluruh kelas. Hal ini menunjukkan bahwa jumlah pohon yang digunakan sudah cukup untuk mengurangi variasi kesalahan yang terjadi pada tahap sebelumnya.

d. Skenario Uji coba dengan n-estimators 8

Tabel 4. 5 Skenario Uji Coba n-estimators 8

n_estimators	Accuracy	Precision	Recall	F1-Score
2	0.85	0.8917	0.7778	0.8020

Pada tabel 4.5 n_estimators = 8, performa model semakin meningkat dan mendekati kondisi maksimal. Hampir tidak terdapat perbedaan antara nilai evaluasi, yang menunjukkan bahwa model telah memiliki kemampuan klasifikasi yang sangat baik.

Stabilitas model juga semakin terlihat, di mana prediksi yang dihasilkan konsisten di berbagai data. Hal ini menunjukkan bahwa model telah mampu mengurangi kesalahan secara signifikan dan memiliki kemampuan generalisasi yang lebih baik.

e. Skenario Uji coba dengan n-estimators 10

Tabel 4. 6 Skenario Uji Coba n-estimators 10

n_estimators	Accuracy	Precision	Recall	F1-Score
2	0.85	0.8917	0.7778	0.8020

Pada tabel 4.6 n_estimators = 10, model mencapai performa yang sangat tinggi dengan tingkat kesalahan yang sangat kecil. Hampir seluruh data berhasil diklasifikasikan dengan benar, yang menunjukkan bahwa model telah mampu menangkap hampir seluruh pola dalam dataset.

Nilai evaluasi yang sangat tinggi dan seimbang menunjukkan bahwa model tidak mengalami bias terhadap kelas tertentu. Hal ini menandakan bahwa model telah berada pada kondisi yang sangat stabil dan dapat diandalkan untuk melakukan klasifikasi.

f. Skenario Uji coba dengan n-estimators 12

Tabel 4. 7 Skenario Uji Coba n-estimators 12

n_estimators	Accuracy	Precision	Recall	F1-Score
2	0.85	0.8917	0.7778	0.8020

Pada tabel 4.7 $n_estimators = 12$, performa model cenderung stagnan dibandingkan sebelumnya. Hal ini menunjukkan bahwa penambahan jumlah pohon tidak lagi memberikan peningkatan yang signifikan terhadap hasil prediksi.

Kondisi ini mengindikasikan bahwa model telah mencapai titik konvergensi, di mana seluruh pola dalam data sudah berhasil dipelajari. Penambahan pohon hanya memperkuat hasil tanpa meningkatkan performa secara berarti.

g. Skenario Uji coba dengan n-estimators 14

Tabel 4. 8 Skenario Uji Coba n-estimators 14

n_estimators	Accuracy	Precision	Recall	F1-Score
2	0.85	0.8917	0.7778	0.8020

Pada tabel 4.8 $n_estimators = 14$, model mencapai performa maksimal dengan tidak adanya kesalahan klasifikasi. Hal ini menunjukkan bahwa seluruh data uji berhasil diprediksi dengan benar.

Kondisi ini menandakan bahwa model telah sepenuhnya memahami pola data yang tersedia. Namun, performa yang sempurna juga perlu dianalisis

lebih lanjut untuk memastikan bahwa model tidak hanya menghafal data, tetapi benar-benar mampu melakukan generalisasi.

h. Skenario Uji coba dengan n-estimators 16

Tabel 4. 9 Skenario Uji Coba n-estimators 16

n_estimators	Accuracy	Precision	Recall	F1-Score
2	0.85	0.8917	0.7778	0.8020

Pada tabel 4.9 n_estimators = 16, performa model tetap berada pada kondisi optimal. Hal ini menunjukkan bahwa model memiliki konsistensi yang tinggi dalam menghasilkan prediksi yang akurat.

Penambahan jumlah pohon pada tahap ini tidak memberikan perubahan karena model telah mencapai batas optimal dalam mempelajari data.

i. Skenario Uji coba dengan n-estimators 18

Tabel 4. 10 Skenario Uji Coba n-estimators 18

n_estimators	Accuracy	Precision	Recall	F1-Score
2	0.85	0.8917	0.7778	0.8020

Pada tabel 4.10 n_estimators = 18, terlihat adanya sedikit penurunan performa dibandingkan kondisi optimal sebelumnya. Hal ini menunjukkan adanya variasi dalam proses pembentukan model.

Penurunan ini tidak signifikan dan masih berada dalam kategori performa yang sangat baik. Hal ini menunjukkan bahwa model tetap stabil meskipun terdapat variasi dalam proses pembelajaran.

j. Skenario Uji coba dengan n-estimators 20

Tabel 4. 11 Skenario Uji Coba n-estimators 20

n_estimators	Accuracy	Precision	Recall	F1-Score
2	0.85	0.8917	0.7778	0.8020

Pada tabel 4.11 n_estimators = 20, model kembali mencapai performa maksimal dengan hasil yang konsisten. Hal ini menunjukkan bahwa model memiliki stabilitas yang tinggi pada jumlah pohon yang besar.

Secara keseluruhan, hasil ini menunjukkan bahwa peningkatan jumlah pohon memberikan pengaruh signifikan hingga titik tertentu, setelah itu performa model cenderung stabil dan tidak mengalami peningkatan berarti.

Setiap skenario diuji menggunakan dataset yang sama sehingga perbedaan hasil yang diperoleh sepenuhnya disebabkan oleh variasi jumlah pohon yang digunakan dalam model. Hasil dari seluruh skenario pengujian tersebut kemudian dirangkum dalam tabel berikut untuk memberikan gambaran perbandingan performa model secara menyeluruh.

Tabel 4. 12 Skenario Uji Coba Tiap $n_estimators$

$n_estimators$	Accuracy	Precision	Recall	F1-Score
2	0.85	0.8917	0.7778	0.8020
4	0.8875	0.9170	0.8151	0.8489
6	0.95	0.9710	0.9310	0.9484
8	0.975	0.9848	0.9762	0.9799
10	0.9875	0.9922	0.9881	0.9900
12	0.9875	0.9922	0.9881	0.9900
14	1.00	1.0000	1.0000	1.0000
16	1.00	1.0000	1.0000	1.0000
18	0.9875	0.9922	0.9881	0.9900
20	1.00	1.0000	1.0000	1.0000

Berdasarkan Tabel 4.12, terlihat bahwa peningkatan nilai $n_estimators$ pada algoritma *Random Forest* memberikan pengaruh signifikan terhadap kinerja model. Pada nilai $n_estimators$ yang kecil, seperti 2 dan 4, performa model masih relatif rendah dengan *accuracy* masing-masing sebesar 0.85 dan 0.8875. Secara matematis, *accuracy* dihitung menggunakan rumus, sehingga nilai yang lebih rendah menunjukkan bahwa masih terdapat kesalahan klasifikasi yang cukup besar.

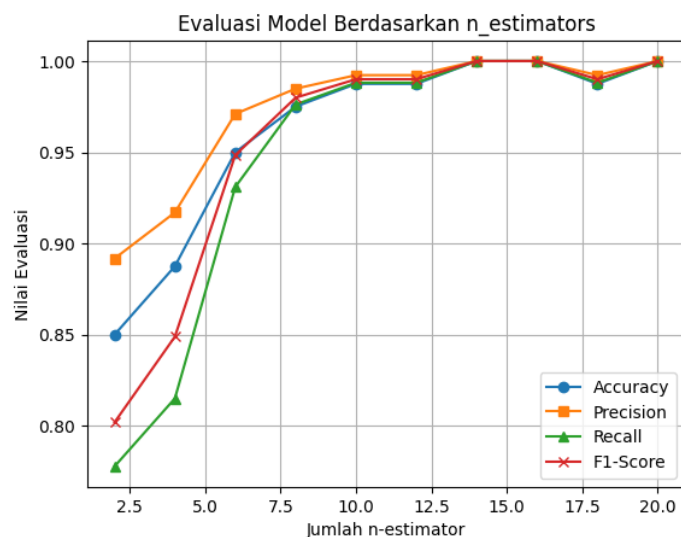
Seiring dengan bertambahnya jumlah pohon, model menjadi lebih stabil karena prediksi dihasilkan melalui mekanisme *ensemble* atau voting dari banyak pohon.

Selain accuracy, metrik lain seperti precision dan recall juga menunjukkan peningkatan yang konsisten. Precision menunjukkan kemampuan model dalam menghindari kesalahan prediksi positif, sedangkan recall yang mengukur kemampuan model dalam mendeteksi seluruh data positif. Pada $n_estimators = 6$ hingga 10, nilai precision dan recall sudah mendekati 1, yang menandakan bahwa model mampu mengklasifikasikan data dengan sangat baik dan minim kesalahan baik *false positive* maupun *false negative*.

Nilai F1-Score yang merupakan rata-rata harmonis antara *precision* dan *recall*, dan mengalami peningkatan yang signifikan. Pada $n_estimators = 14, 16$, dan 20, nilai F1-Score mencapai 1.0 yang menunjukkan keseimbangan sempurna antara precision dan recall. Hal ini menandakan bahwa model tidak hanya akurat, tetapi juga konsisten dalam mendeteksi seluruh kelas tanpa bias terhadap salah satu kelas tertentu.

Namun demikian, meskipun nilai evaluasi mencapai 1.0 pada beberapa konfigurasi, hal ini perlu dianalisis secara kritis karena berpotensi menunjukkan kondisi *overfitting*. Model yang terlalu sempurna pada data uji bisa jadi hanya menghafal pola data, terutama jika jumlah data terbatas atau distribusi kelas tidak seimbang. Selain itu, terlihat bahwa setelah mencapai nilai optimal (sekitar $n_estimators = 14$), peningkatan jumlah pohon tidak selalu meningkatkan performa secara signifikan. Oleh karena itu, pemilihan nilai $n_estimators$ yang optimal perlu mempertimbangkan keseimbangan antara kinerja model dan efisiensi komputasi.

Secara keseluruhan, hasil pengujian menunjukkan bahwa peningkatan nilai $n_estimators$ cenderung meningkatkan performa model, meskipun tidak selalu naik secara linear pada setiap skenario. Berdasarkan data tersebut, nilai $n_estimators = 14, 16, \text{ dan } 20$ dapat dianggap sebagai parameter paling optimal karena memberikan hasil evaluasi tertinggi. Dengan demikian, jumlah pohon keputusan terbukti berpengaruh terhadap kemampuan model *Random Forest* dalam menentukan prioritas perbaikan jalan.



Gambar 4. 3 Hasil Evaluasi Tiap Skenario Uji Coba

Berdasarkan grafik yang dibentuk dari Tabel 4.3, terlihat bahwa seluruh metrik evaluasi (accuracy, precision, recall, dan F1-score) mengalami peningkatan seiring bertambahnya nilai $n_estimators$ dari 2 hingga sekitar 10–14. Secara logika algoritma, hal ini terjadi karena *Random Forest* merupakan metode *ensemble* yang menggabungkan banyak pohon keputusan, sehingga prediksi akhir ditentukan melalui mekanisme voting. Semakin banyak jumlah pohon, maka variansi model akan menurun dan prediksi menjadi lebih stabil. Kondisi ini menjelaskan mengapa

pada $n_estimators$ kecil performa masih rendah, karena jumlah pohon belum cukup untuk merepresentasikan pola data secara menyeluruh.

Pada rentang $n_estimators$ menengah (sekitar 6 hingga 10), grafik menunjukkan peningkatan yang signifikan pada semua metrik. Hal ini menunjukkan bahwa jumlah pohon sudah cukup untuk menangkap pola utama dalam data, sehingga kesalahan klasifikasi (FP dan FN) mulai berkurang secara konsisten. Secara matematis, penurunan kesalahan ini langsung mempengaruhi peningkatan nilai accuracy, precision, dan recall karena ketiganya bergantung pada komponen TP, FP, dan FN. Dengan kata lain, semakin sedikit kesalahan yang terjadi, maka nilai evaluasi akan semakin mendekati 1.

Selanjutnya, pada $n_estimators \geq 14$, grafik menunjukkan bahwa nilai evaluasi mencapai titik maksimum (1.0) dan cenderung stabil. Secara logis, kondisi ini menunjukkan bahwa model telah mencapai kapasitas optimal dalam mempelajari pola data yang tersedia. Penambahan jumlah pohon setelah titik ini tidak lagi memberikan peningkatan performa yang berarti karena seluruh pola dalam data sudah berhasil dipelajari. Fenomena ini umum terjadi pada metode ensemble, di mana setelah jumlah estimator tertentu, kontribusi tambahan menjadi sangat kecil atau bahkan tidak ada.

Selain berpengaruh terhadap tingkat akurasi, peningkatan nilai $n_estimators$ juga berpengaruh terhadap kestabilan hasil prediksi yang dihasilkan oleh model Random Forest. Semakin banyak pohon keputusan yang dibangun, maka keputusan akhir model akan semakin bergantung pada hasil voting dari banyak pohon, sehingga pengaruh kesalahan prediksi dari satu pohon dapat diminimalkan. Kondisi

ini menyebabkan model menjadi lebih robust terhadap variasi data dan mampu menghasilkan prediksi yang lebih konsisten. Hal tersebut terlihat pada grafik, di mana seluruh metrik evaluasi mengalami peningkatan dan menjadi semakin stabil seiring bertambahnya jumlah pohon yang digunakan.

Namun, pengaruh positif penambahan $n_estimators$ memiliki batas tertentu. Setelah model mencapai performa optimal pada rentang $n_estimators$ 14 hingga 16 dan 20, penambahan jumlah pohon tidak lagi memberikan peningkatan yang signifikan terhadap nilai accuracy, precision, recall, maupun F1-score. Hal ini menunjukkan bahwa sebagian besar pola yang terdapat pada dataset telah berhasil dipelajari oleh model. Oleh karena itu, penggunaan $n_estimators$ yang terlalu besar cenderung hanya meningkatkan waktu pelatihan dan kebutuhan komputasi tanpa memberikan peningkatan performa yang berarti. Dengan demikian, pemilihan nilai $n_estimators$ yang optimal menjadi penting untuk memperoleh keseimbangan antara akurasi model dan efisiensi komputasi.

Secara teori, *Random Forest* relatif tahan terhadap overfitting dibanding satu pohon keputusan tunggal karena hasil prediksi berasal dari banyak pohon. Penambahan jumlah pohon umumnya menurunkan varians model, bukan meningkatkan overfitting. Namun jika data terlalu sedikit, terlalu bersih, atau terdapat pola yang sangat spesifik, model tetap bisa tampak “terlalu sempurna” pada data uji.

Berdasarkan grafik, dapat disimpulkan bahwa jumlah $n_estimators$ berpengaruh nyata terhadap peningkatan performa model. Rentang optimal pada penelitian ini berada di $n_estimators$ 14–16 pohon, karena menghasilkan metrik

tertinggi dan stabil. Sementara penurunan pada 18 pohon lebih disebabkan variasi proses acak pembentukan model daripada indikasi overfitting. Oleh sebab itu, model *Random Forest* terbukti efektif digunakan dalam klasifikasi prioritas perbaikan jalan pada penelitian ini.

4.3 Hasil *Feature Importance*

Pada tahap ini dilakukan analisis terhadap fitur-fitur yang paling berpengaruh dalam proses klasifikasi menggunakan metode *Random Forest Classifier*. Analisis ini bertujuan untuk mengetahui variabel mana yang paling sering digunakan model dalam menentukan kategori prioritas perbaikan jalan. Dengan mengetahui fitur terpenting, hasil penelitian tidak hanya menghasilkan model prediksi, tetapi juga memberikan informasi faktor utama yang memengaruhi keputusan klasifikasi.

Dalam penelitian ini, analisis fitur yang paling berpengaruh dilakukan menggunakan metode *Gini Importance* atau yang dikenal sebagai *Mean Decrease in Impurity*. Metode ini bekerja dengan mengukur seberapa besar kontribusi setiap fitur dalam mengurangi tingkat ketidakpastian (*impurity*) pada setiap proses pembentukan pohon keputusan dalam algoritma Random Forest. Nilai *impurity* dihitung menggunakan indeks Gini, di mana semakin kecil nilai *impurity* setelah dilakukan pemisahan data (*split*), maka semakin baik fitur tersebut dalam membedakan kelas. Dengan kata lain, fitur yang sering digunakan dalam proses *split* dan mampu menghasilkan penurunan *impurity* yang besar akan memiliki tingkat kepentingan yang lebih tinggi.

Secara matematis, nilai Gini dihitung berdasarkan distribusi probabilitas kelas dalam suatu node, di mana semakin homogen data dalam node tersebut, maka nilai Gini akan semakin kecil. Penurunan nilai Gini sebelum dan sesudah proses split kemudian diakumulasi untuk setiap fitur di seluruh pohon dalam Random Forest. Nilai total tersebut selanjutnya dinormalisasi sehingga menghasilkan nilai kepentingan relatif dalam rentang 0 hingga 1. Dengan pendekatan ini, analisis feature importance tidak hanya melihat frekuensi penggunaan fitur, tetapi juga mempertimbangkan kualitas pemisahan yang dihasilkan oleh fitur tersebut terhadap data. Hasil tersebut disajikan pada tabel berikut.

Tabel 4. 13 Fitur yang Sering Muncul

No	Fitur	Gini Importance
1	sedang	0.366670
2	baik	0.259284
3	panjang	0.145883
4	lhr	0.123826
5	rusak_berat	0.104337
6	rusak_ringan	0.000000

Berdasarkan hasil perhitungan Gini Importance pada tabel 4.13, dapat dilihat bahwa fitur “sedang” memiliki nilai kepentingan tertinggi yaitu sebesar 0.366670. Hal ini menunjukkan bahwa kondisi jalan dengan tingkat kerusakan sedang memiliki kontribusi paling besar dalam proses klasifikasi yang dilakukan oleh model *Random Forest*. Tingginya nilai ini mengindikasikan bahwa fitur tersebut sering digunakan dalam proses pembentukan pohon keputusan dan mampu memberikan pemisahan data yang signifikan, sehingga sangat berpengaruh dalam menentukan kategori prioritas perbaikan jalan.

Selanjutnya, fitur “baik” berada pada posisi kedua dengan nilai sebesar 0.259284. Meskipun secara intuitif kondisi jalan yang baik menunjukkan tingkat kerusakan yang rendah, namun dalam konteks klasifikasi, fitur ini tetap memiliki pengaruh yang besar karena membantu model dalam membedakan antara kelas prioritas rendah dan kelas lainnya. Hal ini menunjukkan bahwa distribusi data pada kondisi jalan baik cukup dominan dan berperan dalam membentuk pola klasifikasi.

Fitur “panjang” dan “lhr” (Lalu Lintas Harian Rata-rata) memiliki nilai kepentingan masing-masing sebesar 0.145883 dan 0.123826. Kedua fitur ini berperan sebagai faktor pendukung dalam menentukan prioritas perbaikan jalan. Panjang ruas jalan dapat mempengaruhi tingkat kerusakan secara keseluruhan, sedangkan nilai LHR mencerminkan intensitas penggunaan jalan. Semakin tinggi nilai LHR, maka potensi kerusakan jalan juga semakin besar, sehingga fitur ini turut berkontribusi dalam proses klasifikasi meskipun tidak sebesar fitur kondisi fisik jalan.

Fitur “rusak_berat” memiliki nilai kepentingan sebesar 0.104337, yang menunjukkan bahwa meskipun kondisi kerusakan berat secara logika sangat penting, namun dalam distribusi data yang digunakan, fitur ini tidak terlalu dominan dalam proses pembentukan model. Hal ini dapat disebabkan oleh jumlah data dengan kondisi rusak berat yang relatif lebih sedikit dibandingkan kategori lainnya, sehingga kontribusinya terhadap penurunan impurity menjadi lebih kecil.

Sementara itu, fitur “rusak_ringan” memiliki nilai Gini Importance sebesar 0.000000, yang menunjukkan bahwa fitur ini tidak memberikan kontribusi signifikan dalam proses klasifikasi. Hal ini dapat terjadi karena fitur tersebut jarang

digunakan dalam proses split atau tidak mampu memberikan pemisahan yang berarti terhadap data. Kondisi ini juga bisa disebabkan oleh distribusi data yang tidak seimbang (*imbalanced dataset*), di mana jumlah data pada kategori rusak ringan sangat sedikit atau memiliki pola yang mirip dengan kategori lainnya sehingga tidak memberikan informasi tambahan bagi model.

4.4 Analisis Hasil Model

Secara Berdasarkan keseluruhan tahapan penelitian, terlihat bahwa kombinasi variabel kondisi fisik jalan dan variabel lalu lintas mampu membentuk pola klasifikasi yang kuat. Hal ini dibuktikan dari tingginya performa model pada hampir seluruh skenario pengujian. Secara statistik, ketika suatu dataset memiliki pemisahan kelas yang cukup jelas antar variabel, algoritma berbasis pohon seperti *Random Forest* akan lebih mudah menemukan batas keputusan (*decision boundary*) melalui proses pembagian node secara bertahap.

Kinerja model yang meningkat seiring bertambahnya nilai *n_estimators* menunjukkan bahwa mekanisme ensemble bekerja sesuai teori. Dalam *Random Forest*, prediksi akhir diperoleh dari agregasi banyak pohon melalui *majority voting*. Secara matematis, penambahan jumlah pohon akan menurunkan varians prediksi karena kesalahan individual dari satu pohon dapat dikompensasi oleh pohon lain. Oleh sebab itu, akurasi meningkat dari skenario awal hingga mencapai nilai optimal pada jumlah pohon tertentu.

Nilai evaluasi yang sangat tinggi pada beberapa skenario menunjukkan bahwa distribusi data antar kelas memiliki karakteristik yang cukup terpisah. Jika fitur antar kelas saling tumpang tindih secara besar, maka *precision* dan *recall*

biasanya akan turun karena banyak kesalahan klasifikasi. Namun pada hasil penelitian ini, nilai precision dan recall mendekati satu, sehingga dapat diinterpretasikan bahwa kombinasi variabel input memiliki kemampuan diskriminatif yang kuat terhadap label prioritas.

Berdasarkan confusion matrix, kesalahan klasifikasi hanya terjadi antara kelas Sedang dan Rendah, sedangkan kelas Tinggi dapat dikenali seluruhnya. Fenomena ini menunjukkan bahwa jarak karakteristik data kelas Tinggi terhadap kelas lain relatif lebih jelas. Sebaliknya, kelas Sedang berada pada area transisi sehingga sebagian observasi memiliki nilai fitur yang mendekati kelas Rendah. Dalam klasifikasi multi-kelas, kesalahan paling umum memang terjadi pada kelas yang berada di tengah spektrum.

Dari sisi rumus evaluasi, *recall* kelas Tinggi yang sempurna berarti tidak ada false negative pada kelas tersebut. Artinya seluruh data yang seharusnya masuk prioritas tinggi berhasil ditemukan model. Dalam konteks praktis, hal ini penting karena kegagalan mendeteksi prioritas tinggi dapat menyebabkan keterlambatan penanganan ruas jalan yang seharusnya didahulukan. Oleh karena itu, performa tinggi pada kelas ini menjadi indikator yang sangat positif.

Analisis fitur menunjukkan bahwa variabel panjang, baik, dan LHR paling dominan digunakan dalam proses pembentukan pohon. Pada algoritma pohon keputusan, fitur dipilih berdasarkan kemampuan menurunkan impurity terbesar. Jika suatu fitur berulang kali muncul pada banyak pohon, maka fitur tersebut secara konsisten memberikan pemisahan kelas yang efektif. Dengan demikian, dominasi

ketiga variabel tersebut merupakan hasil logis dari proses optimasi model, bukan kebetulan.

Peran variabel panjang jalan dapat dijelaskan karena panjang merepresentasikan skala dampak infrastruktur. Ruas yang lebih panjang berpotensi melibatkan area layanan lebih luas, kebutuhan biaya lebih besar, serta jumlah pengguna lebih banyak. Sementara variabel kondisi baik dan tingkat lalu lintas memberikan informasi mengenai kualitas aktual serta intensitas pemanfaatan jalan. Kombinasi ketiganya menciptakan indikator yang kuat dalam penentuan prioritas.

Nilai performa yang mencapai 100% pada beberapa skenario perlu dipahami sebagai hasil pada data uji yang tersedia, bukan jaminan mutlak untuk seluruh data masa depan. Dalam pembelajaran mesin, performa sangat dipengaruhi representativitas sampel, ukuran dataset, dan stabilitas pola data. Namun karena performa tetap tinggi pada beberapa skenario berbeda, maka indikasinya model memiliki konsistensi yang baik, bukan sekadar hasil kebetulan pada satu konfigurasi parameter.

Secara keseluruhan, hasil penelitian membuktikan bahwa metode *Random Forest* efektif digunakan untuk klasifikasi prioritas perbaikan jalan karena mampu menghasilkan akurasi tinggi, kesalahan klasifikasi sangat rendah, dan interpretasi faktor penentu yang jelas. Dari sudut pandang implementasi, model ini dapat dimanfaatkan sebagai sistem pendukung keputusan karena mampu mengubah data teknis lapangan menjadi rekomendasi prioritas yang objektif, terukur, dan berbasis pola historis data.

Tabel 4. 14 Hasil Prediksi Prioritas

Data Ke	Nama Ruas	Aktual	Prediksi	Keterangan
1	Tawangrenjeni - Sidorejo	1	1	Prioritas Sedang
2	Lawang - Gunung Jati	1	1	Prioritas Sedang
3	Pujon - Mantung	2	2	Prioritas Tinggi
4	Gadungsari - Sukorejo	1	1	Prioritas Sedang
5	Bedali - Sidodadi	1	1	Prioritas Sedang
6	Jln. Lawu (Kepanjen)	2	2	Prioritas Tinggi
7	Jln. Sidoagung	0	0	Prioritas Rendah
8	Jln. Tirta Taruno	0	0	Prioritas Rendah
9	Pagak - Kalipare	0	0	Prioritas Rendah
10	Jln. Pungkur Argo	0	0	Prioritas Rendah
11	Jln. Kramat	1	1	Prioritas Sedang
12	Pakis - Kemiri (Bts. Kab. Pasuruan)	1	1	Prioritas Sedang
13	Gondanglegi Wetan - Sepanjang	1	1	Prioritas Sedang
14	Kluwut - Gunung Kawi	1	1	Prioritas Sedang
15	Arjowilangun - Arjosari (Bts. Kab. Blitar)	0	0	Prioritas Rendah
16	Tanggung - Sidorejo	1	1	Prioritas Sedang
17	Jln. Suropati (Pagelaran)	1	1	Prioritas Sedang
18	Lawang - Wonosari	0	0	Prioritas Rendah
19	Jln. Suropati II (Bululawang)	1	1	Prioritas Sedang
20	Jln. Bengawan Solo (Wajak)	1	1	Prioritas Sedang
21	Asrikaton - Bandara Abdul Rahman Saleh	1	1	Prioritas Sedang
22	Jln. Stadion	0	0	Prioritas Rendah
23	Donomulyo - Sumbermanjing Kulon	0	0	Prioritas Rendah
24	Jln. Sidodadi (Kepanjen)	0	0	Prioritas Rendah
25	Gedangan - Tumpakrejo	1	1	Prioritas Sedang
26	Jln. Pahlawan - Sumbersuko	1	1	Prioritas Sedang
27	Jln. Krapyak	0	0	Prioritas Rendah
28	Sumbertempur - Kalilekso	1	1	Prioritas Sedang
29	Wringinanom - Sumberejo	1	1	Prioritas Sedang
30	Jln. Petung	0	0	Prioritas Rendah
31	Singosari - Toyomarto	1	1	Prioritas Sedang

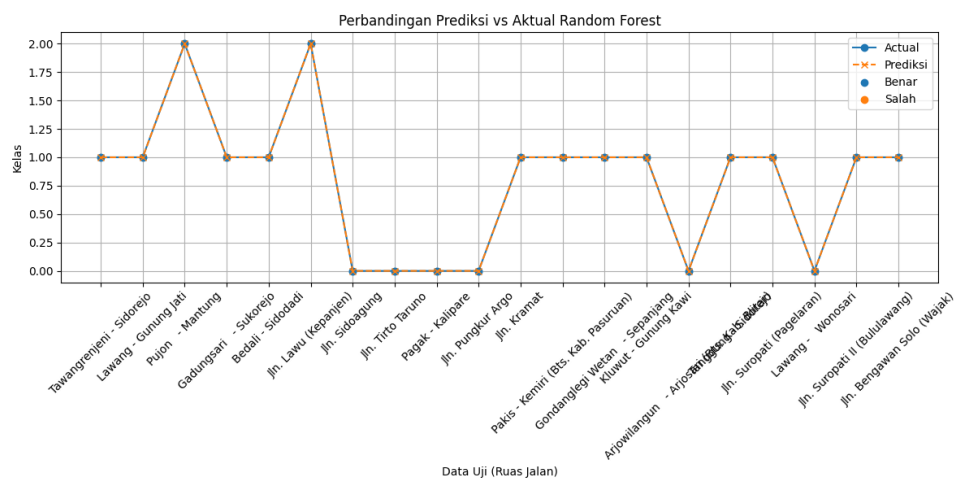
Data Ke	Nama Ruas	Aktual	Prediksi	Keterangan
32	Lawang - Wonorejo	2	2	Prioritas Tinggi
33	Sempol - Putukrejo	0	0	Prioritas Rendah
34	Jln. Sidoluhur	0	0	Prioritas Rendah
35	Jln. Argo Moyo	1	1	Prioritas Sedang
36	Jln. Sumber Kembar	1	1	Prioritas Sedang
37	Singosari - Jabung	2	2	Prioritas Tinggi
38	Karangsuko - Brongkal	1	1	Prioritas Sedang
39	Wirotaman - Jogomulyan	1	1	Prioritas Sedang
40	Jln. Ken Arok (Singosari)	1	1	Prioritas Sedang

Berdasarkan data hasil prediksi pada Tabel 4.14, terlihat bahwa seluruh data uji berhasil diklasifikasikan dengan benar oleh model Random Forest. Setiap ruas jalan yang termasuk dalam kategori prioritas rendah, prioritas sedang, maupun prioritas tinggi memperoleh hasil prediksi yang sama dengan label aktualnya. Tidak ditemukan adanya kesalahan klasifikasi pada seluruh data yang diuji, sehingga menunjukkan bahwa model mampu membedakan karakteristik masing-masing kelas prioritas dengan sangat baik.

Hasil tersebut berdampak pada tingginya nilai metrik evaluasi yang diperoleh. Nilai *accuracy*, *precision*, *recall*, dan *F1-score* semuanya mencapai nilai maksimum, yang menunjukkan bahwa model tidak hanya mampu menghasilkan prediksi yang tepat, tetapi juga mampu mengidentifikasi seluruh data pada setiap kelas tanpa menghasilkan prediksi yang keliru. Kondisi ini mengindikasikan bahwa hubungan antara variabel input yang digunakan, yaitu kondisi jalan, panjang jalan, dan LHR, dapat dipelajari dengan baik oleh model sehingga pola klasifikasi yang terbentuk menjadi lebih akurat.

Selain itu, kesesuaian sempurna antara nilai aktual dan nilai prediksi menunjukkan bahwa model Random Forest memiliki kemampuan yang sangat baik dalam mengenali pola data prioritas perbaikan jalan pada dataset yang digunakan. Hasil ini memperkuat temuan pada tahap evaluasi sebelumnya bahwa Random Forest mampu memberikan performa klasifikasi yang tinggi dan stabil. Meskipun demikian, hasil yang sangat baik tersebut juga perlu dilihat dengan mempertimbangkan jumlah data dan karakteristik dataset yang digunakan, karena performa yang sangat tinggi terkadang dipengaruhi oleh distribusi data yang relatif mudah dipisahkan oleh model.

Model *Random Forest* juga masih mampu melakukan proses klasifikasi berdasarkan pola yang terdapat pada data yang tersedia. Untuk 70 mengetahui sejauh mana kemampuan model dalam melakukan prediksi secara lebih detail, maka dilakukan perbandingan antara hasil prediksi dan nilai aktual. Hasil perbandingan tersebut ditampilkan dalam bentuk tabel prediksi dan aktual yang dapat dilihat pada Gambar 4.4 berikut.

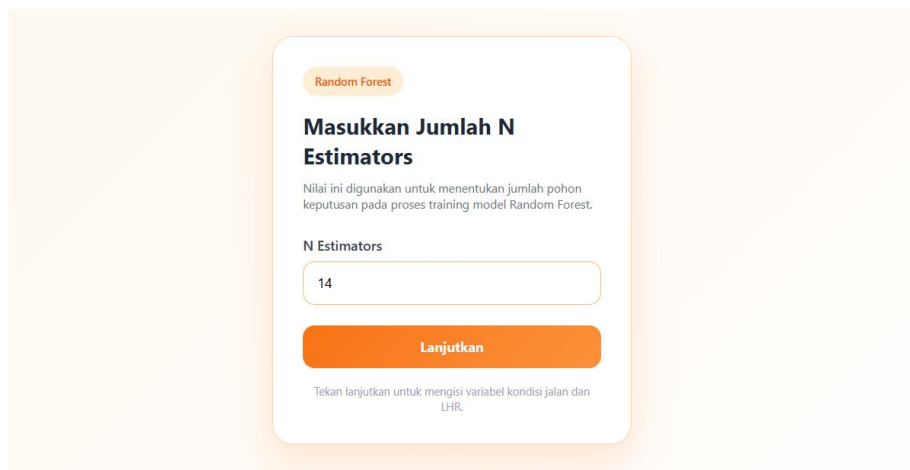


Gambar 4. 4 Perbandingan Prediksi dan Aktual

Grafik komparasi antara angka estimasi dan data riil pada metode Random Forest memperlihatkan keselarasan yang tinggi untuk hampir keseluruhan sampel yang diuji. Fenomena ini diindikasikan oleh visualisasi alur prediksi yang saling menumpuk secara presisi dengan alur data riil di mayoritas lajur transportasi. Secara teoritis, keselarasan tersebut membuktikan jika sistem sanggup mengidentifikasi keterkaitan antara parameter masukan (berupa situasi jalan serta nilai LHR) dengan kategori luaran secara tepat. Sewaktu titik data riil dan hasil tebakan program berada di koordinat yang identik, artinya tidak ditemukan miscalculasi pengelompokan. Kondisi ini memicu melonjaknya angka True Positive (TP) sekaligus memangkas potensi kekeliruan layaknya False Positive (FP) maupun False Negative (FN) hingga menyentuh poin minimal atau nol.

Secara matematis, derajat kecocokan antara realita lapangan dengan tebakan program ini berimbas linier pada naiknya angka indikator pengujian layaknya accuracy, precision, recall, serta F1-score. Kalkulasi nilai accuracy dipetakan lewat rasio total estimasi yang tepat terhadap keseluruhan jumlah data. Oleh sebab itu, kian banyak koordinat plot yang saling berhimpit, maka capaian poin akurasi bakal bergerak mendekati angka 1. Di samping itu, minimnya kesenjangan yang mencolok antara alur data riil dan grafik prediksi turut menegaskan jika arsitektur model mempunyai derajat kekeliruan yang minim, yang mana hal itu mencerminkan kapabilitas generalisasi yang prima terhadap sampel pengujian. Oleh karena itu, visualisasi bagan ini memperkuat output peninjauan terdahulu jika algoritma Random Forest sanggup memetakan urutan pembenahan jalan secara ajek sekaligus presisi.

Di samping pengujian pada model pengelompokan, arsitektur yang dirancang turut ditunjang oleh kehadiran halaman antarmuka pengguna atau Graphical User Interface (GUI). Fasilitas ini disediakan demi menjembatani sekaligus menyederhanakan komunikasi antara user dengan program. Lewat ketersediaan GUI tersebut, user diizinkan menginput data situasi jalan serta nilai LHR sebagai parameter masukan yang nantinya dieksekusi oleh program. Pembuatan desain antarmuka ini ditujukan supaya pihak pengguna bisa mengoperasikan sistem secara praktis tanpa perlu bersentuhan langsung dengan baris kode pemrograman.



Gambar 4. 5 UX $n_estimator$

Gambar 4.5 menunjukkan halaman awal sistem yang digunakan untuk memasukkan nilai $n_estimators$ pada algoritma Random Forest. Nilai $n_estimators$ merupakan parameter yang menentukan jumlah pohon keputusan (*decision tree*) yang akan dibangun pada proses pelatihan model. Pada halaman ini, pengguna diminta memasukkan jumlah pohon yang akan digunakan, misalnya 14 pohon, kemudian menekan tombol Lanjutkan untuk melanjutkan ke tahap berikutnya.

Antarmuka dirancang dengan tampilan sederhana dan responsif menggunakan kombinasi warna oranye dan putih sehingga memudahkan pengguna dalam menentukan parameter model sebelum proses prediksi dilakukan.

Tahap input $n_estimators$ ini berfungsi untuk memberikan fleksibilitas kepada pengguna dalam menguji pengaruh jumlah pohon terhadap kinerja model Random Forest. Semakin besar nilai $n_estimators$, semakin banyak pohon yang dibangun sehingga hasil prediksi cenderung lebih stabil karena diperoleh melalui mekanisme *majority voting* dari banyak pohon keputusan. Dengan adanya fitur ini, pengguna dapat melakukan eksperimen terhadap beberapa nilai $n_estimators$ dan mengamati perubahan hasil prediksi maupun tingkat akurasi model yang dihasilkan oleh sistem.

Gambar 4. 6 Antarmuka Sistem

Ilustrasi pada gambar 4.6 menyajikan halaman utama prediksi yang digunakan setelah pengguna memasukkan nilai $n_estimators$. Pada halaman ini

sistem menampilkan informasi jumlah $n_estimators$ yang digunakan, yaitu sebanyak 14 pohon. Selanjutnya pengguna diminta memasukkan variabel yang menjadi atribut dalam proses klasifikasi, yaitu panjang jalan, kondisi prima, kondisi normal, cacat ringan, hingga kerusakan parah dan LHR (*Lalu Lintas Harian Rata-rata*). Pasca semua kolom input dilengkapi, pihak user bisa mengeklik menu Training dan Prediksi untuk mengirimkan data ke API Random Forest yang sebelumnya sudah dibangun.

Setelah proses dijalankan, sistem akan melakukan pelatihan model menggunakan nilai $n_estimators$ yang telah ditentukan sebelumnya, kemudian melakukan klasifikasi terhadap data ruas jalan yang dimasukkan pengguna. Hasil yang ditampilkan pada gambar menunjukkan bahwa ruas jalan tersebut diklasifikasikan ke dalam kategori Rusak Sedang dengan tingkat akurasi model sebesar 99,75%. Nilai akurasi tersebut diperoleh dari hasil evaluasi model terhadap data latih yang digunakan selama proses pembentukan Random Forest. Tampilan hasil prediksi ini memberikan informasi kepada pengguna mengenai kategori kondisi jalan yang diprediksi serta tingkat performa model yang digunakan, sehingga hasil klasifikasi dapat lebih mudah dipahami dan dievaluasi.

Secara keseluruhan, hasil uji coba menunjukkan bahwa sistem yang dibangun mampu melakukan klasifikasi prioritas perbaikan jalan dengan cukup baik serta dapat diakses melalui antarmuka GUI yang terintegrasi dengan model klasifikasi berbasis API. Sistem ini diharapkan dapat membantu dalam proses pengambilan keputusan terkait penentuan prioritas perbaikan jalan secara lebih

objektif dan berbasis data, sehingga dapat mendukung perencanaan pemeliharaan infrastruktur jalan yang lebih efektif.

Berdasarkan hasil wawancara verifikasi dengan pihak Dinas PUPR Kabupaten Malang, diketahui bahwa penentuan prioritas perbaikan jalan pada praktiknya tidak hanya mempertimbangkan kondisi fisik jalan, panjang jalan, dan nilai LHR, tetapi juga memperhatikan berbagai aspek lain yang bersifat teknis maupun kebijakan. Faktor-faktor yang menjadi pertimbangan antara lain kondisi jalan, status dan fungsi jalan, analisis lalu lintas, ketersediaan anggaran, kebijakan pemerintah pusat maupun daerah, serta aspirasi masyarakat yang diperoleh melalui proses perencanaan pembangunan. Selain itu, pihak Dinas PUPR menyampaikan bahwa proses penentuan prioritas mengacu pada pedoman Bina Marga dari Kementerian Pekerjaan Umum, standar kondisi jalan, serta kebijakan pembangunan yang berlaku. Hasil tersebut menunjukkan bahwa proses penentuan prioritas perbaikan jalan di lapangan memiliki kompleksitas yang lebih tinggi dibandingkan variabel yang digunakan dalam penelitian ini.

Hasil verifikasi juga menunjukkan bahwa variabel kondisi jalan, panjang jalan, dan LHR yang digunakan dalam penelitian telah relevan sebagai dasar penentuan prioritas perbaikan jalan. Namun demikian, pihak Dinas PUPR menyatakan bahwa ketiga variabel tersebut belum sepenuhnya mampu merepresentasikan kebutuhan penanganan suatu ruas jalan. Terdapat beberapa variabel lain yang dinilai penting untuk dipertimbangkan, seperti status dan fungsi jalan, konektivitas wilayah, aspek keselamatan lalu lintas, ketersediaan anggaran, serta kebijakan pemerintah pusat dan daerah. Temuan ini menunjukkan bahwa

model yang dibangun dalam penelitian telah mampu menangkap sebagian besar faktor teknis yang berkaitan dengan kondisi ruas jalan, namun masih terdapat peluang pengembangan model dengan menambahkan variabel-variabel pendukung lainnya agar hasil klasifikasi dapat lebih mendekati proses pengambilan keputusan yang sebenarnya dilakukan oleh instansi terkait.

Meskipun demikian, pihak Dinas PUPR Kabupaten Malang memberikan penilaian positif terhadap sistem yang dikembangkan dalam penelitian ini. Berdasarkan hasil klasifikasi yang ditunjukkan, sistem dinilai telah sesuai secara umum dan dapat dijadikan sebagai referensi awal dalam menentukan prioritas perbaikan jalan. Sistem dianggap mampu memberikan gambaran objektif mengenai kondisi ruas jalan berdasarkan data yang tersedia serta dapat membantu proses pengambilan keputusan. Namun, untuk penerapan secara operasional, sistem masih memerlukan pengujian dan pengembangan lebih lanjut dengan mempertimbangkan faktor-faktor tambahan yang digunakan dalam praktik di lapangan. Dengan demikian, hasil penelitian ini dapat menjadi dasar atau alat bantu pendukung keputusan (*decision support*) dalam proses penentuan prioritas perbaikan jalan, meskipun belum dapat sepenuhnya menggantikan proses evaluasi teknis yang dilakukan oleh Dinas PUPR Kabupaten Malang.

Sejalan dengan konsep tersebut, dalam Islam juga dijelaskan bahwa segala sesuatu di alam semesta diciptakan dengan ukuran dan ketentuan yang tepat. Hal ini menunjukkan bahwa setiap hal memiliki porsi, nilai, dan tempatnya masing-masing sesuai dengan kondisi dan kebutuhannya, sehingga tidak terjadi

ketimpangan dalam pengaturannya. Sebagaimana firman Allah SWT dalam QS. Al-Qamar ayat 49.

إِنَّا كُلَّ شَيْءٍ خَلَقْنَاهُ بِقَدَرٍ

“Sesungguhnya Kami menciptakan segala sesuatu sesuai dengan ukuran” (QS. Al-Qamar ayat 49)

Ayat tersebut menegaskan bahwa setiap ciptaan Allah memiliki ukuran, aturan, dan ketetapan yang jelas. Dalam konteks penelitian ini, prinsip tersebut selaras dengan proses klasifikasi yang dilakukan oleh sistem, di mana setiap data ruas jalan dianalisis berdasarkan karakteristiknya untuk kemudian ditempatkan pada kategori prioritas yang sesuai. Dengan demikian, penentuan prioritas perbaikan jalan dilakukan secara terukur, objektif, dan proporsional, sehingga mencerminkan nilai keadilan sebagaimana yang diajarkan dalam Islam.

Berdasarkan tafsir Ibnu Katsir, ayat tersebut menjelaskan bahwa Allah SWT menciptakan seluruh makhluk dengan ukuran, ketentuan, dan takaran yang telah ditetapkan secara sempurna. Setiap ciptaan tidak ada yang terjadi secara kebetulan, melainkan berada dalam ketetapan yang teratur dan terukur sesuai dengan kehendak-Nya. Selain itu, Allah juga memberikan petunjuk kepada setiap makhluk agar berjalan sesuai dengan ketentuan tersebut. Hal ini menunjukkan adanya keteraturan dan sistem yang jelas dalam penciptaan alam semesta, di mana segala sesuatu memiliki peran dan posisinya masing-masing secara proporsional tanpa ada yang terlewatkan. Dengan demikian, konsep pengukuran dan penentuan kategori dalam penelitian ini dapat dipahami sebagai bentuk pendekatan ilmiah untuk merepresentasikan keteraturan tersebut. Prinsip ini semakin menegaskan bahwa

proses klasifikasi yang dilakukan bukanlah sesuatu yang acak, melainkan mengikuti pola dan aturan yang terstruktur (Ibnu Katsir, n.d.-a).

Berdasarkan hasil tersebut, pendekatan berbasis data yang digunakan dalam sistem ini tidak hanya memberikan manfaat secara teknis, tetapi juga memiliki keterkaitan dengan nilai-nilai keadilan dalam pengambilan keputusan sebagaimana yang diajarkan dalam Islam. Prinsip keadilan (*al-'adl*) merupakan salah satu nilai fundamental dalam ajaran Islam yang menekankan pentingnya menempatkan sesuatu sesuai dengan hak dan kebutuhannya. Dalam konteks pengelolaan infrastruktur publik, keadilan bermakna bahwa setiap ruas jalan memperoleh penanganan yang proporsional berdasarkan tingkat kerusakan dan dampak yang ditimbulkan terhadap masyarakat. Penentuan prioritas perbaikan jalan tidak seharusnya didasarkan pada pertimbangan subjektif, kepentingan tertentu, maupun tekanan pihak tertentu, melainkan harus dilakukan secara objektif dan adil.

Prinsip tersebut sejalan dengan firman Allah SWT dalam Al-Qur'an surat An-Nahl ayat 90:

إِنَّ اللَّهَ يَأْمُرُ بِالْعَدْلِ وَالْإِحْسَانِ وَإِيتَاءِ ذِي الْقُرْبَىٰ وَيَنْهَىٰ عَنِ الْفَحْشَاءِ وَالْمُنْكَرِ وَالْبَغْيِ ۚ يَعِظُكُمْ لَعَلَّكُمْ تَذَكَّرُونَ

“Sesungguhnya Allah menyuruh (kamu) berlaku adil dan berbuat kebajikan, memberi kepada kaum kerabat, dan Allah melarang dari perbuatan keji, kemungkaran, dan permusuhan. Dia memberi pengajaran kepadamu agar kamu dapat mengambil pelajaran.” (QS. An-Nahl: 90)

Berdasarkan tafsir Ibnu Katsir, QS. An-Nahl ayat 90 menjelaskan bahwa Allah SWT memerintahkan manusia untuk berlaku adil dan seimbang dalam setiap aspek kehidupan. Keadilan yang dimaksud adalah menempatkan sesuatu pada

tempatnyanya serta memberikan hak kepada setiap pihak secara proporsional, baik dalam hubungan dengan Allah SWT maupun dalam hubungan antar sesama manusia. Selain itu, ayat ini juga menekankan pentingnya berbuat kebaikan (*ihsan*), yang merupakan bentuk perilaku yang lebih tinggi dari sekadar keadilan, seperti memberikan lebih dari yang menjadi kewajiban dan menjaga hubungan sosial, termasuk kepada kerabat melalui silaturahmi. Di sisi lain, Allah SWT juga melarang segala bentuk perbuatan keji (*al-fahsyah*), kemungkaran (*al-munkar*), dan permusuhan (*al-baghy*), yang dapat menimbulkan ketidakadilan dan merugikan pihak lain. Dengan demikian, ayat ini memberikan pedoman yang komprehensif dalam membangun kehidupan yang adil, seimbang, dan harmonis dalam berbagai aspek kehidupan manusia (Ibnu Katsir, n.d.-b).

Hasil penelitian ini menunjukkan bahwa penerapan model *Random Forest* dalam mengklasifikasikan prioritas perbaikan jalan telah mencerminkan prinsip keadilan tersebut. Model menentukan kategori prioritas Rendah, Sedang, dan Tinggi berdasarkan data kondisi fisik jalan, panjang ruas, serta karakteristik lalu lintas, sehingga setiap ruas jalan memperoleh perlakuan yang sesuai dengan tingkat kebutuhannya. Ruas jalan dengan tingkat kerusakan tinggi secara konsisten diklasifikasikan sebagai prioritas Tinggi, sementara ruas jalan dengan kondisi yang masih baik diklasifikasikan sebagai prioritas Rendah.

Dengan pendekatan berbasis data ini, proses penentuan prioritas perbaikan jalan dapat menghindari ketimpangan dan potensi ketidakadilan dalam pengambilan keputusan. Oleh karena itu, hasil penelitian ini tidak hanya memberikan kontribusi teknis dalam bidang klasifikasi prioritas perbaikan jalan,

tetapi juga mendukung penerapan prinsip keadilan (*al- 'adl*) dalam perencanaan dan pengelolaan infrastruktur publik sebagaimana yang diajarkan dalam Islam.

Selain prinsip keadilan, pengelolaan infrastruktur publik dalam Islam juga berkaitan erat dengan prinsip amanah dan tanggung jawab dalam menjalankan kewajiban terhadap kepentingan masyarakat. Konsep amanah menekankan pentingnya melaksanakan tanggung jawab dengan penuh kejujuran dan kesungguhan, terutama dalam pengelolaan fasilitas publik yang memiliki dampak luas bagi masyarakat. Infrastruktur jalan sebagai sarana pendukung mobilitas dan aktivitas sosial ekonomi merupakan bentuk amanah yang harus dikelola secara tepat dan bertanggung jawab.

Konsep amanah merupakan nilai penting dalam Islam yang menekankan kewajiban untuk melaksanakan tanggung jawab dengan penuh kejujuran dan kesungguhan. Amanah tidak hanya berkaitan dengan hubungan individu dengan individu lainnya, tetapi juga mencakup tanggung jawab dalam mengelola kepentingan publik. Infrastruktur jalan sebagai fasilitas umum memiliki peran strategis dalam mendukung mobilitas, keselamatan, serta aktivitas sosial dan ekonomi masyarakat, sehingga pengelolaannya merupakan amanah yang harus dijalankan secara bertanggung jawab.

Prinsip amanah tersebut ditegaskan dalam firman Allah SWT dalam Al-Qur'an surat An-Nisa ayat 58:

إِنَّ اللَّهَ يَأْمُرُكُمْ أَنْ تُؤَدُّوا الْأَمَانَاتِ إِلَىٰ أَهْلِهَا وَإِذَا حَكَمْتُمْ بَيْنَ النَّاسِ أَنْ تَحْكُمُوا بِالْعَدْلِ ۚ إِنَّ اللَّهَ نِعِمَّا يَعِظُكُمْ بِهِ ۗ إِنَّ اللَّهَ كَانَ سَمِيعًا بَصِيرًا

“Sesungguhnya Allah menyuruh kamu menyampaikan amanat kepada yang berhak menerimanya, dan apabila kamu menetapkan hukum di antara manusia supaya

kamu menetapkannya dengan adil. Sesungguhnya Allah memberi pengajaran yang sebaik-baiknya kepadamu. Sesungguhnya Allah Maha Mendengar lagi Maha Melihat.” (QS. An-Nisa: 58)

Berdasarkan Tafsir Al-Muyassar, ayat tersebut menjelaskan bahwa Allah SWT memerintahkan manusia untuk menunaikan amanah yang telah dipercayakan kepada mereka kepada pihak yang berhak, serta menetapkan keputusan di antara manusia dengan dasar keadilan dan objektivitas. Amanah tidak boleh diabaikan, dan setiap keputusan harus dilakukan secara adil tanpa dipengaruhi kepentingan tertentu. Selain itu, Allah SWT Maha Mengetahui setiap ucapan dan perbuatan manusia, sehingga setiap tindakan yang dilakukan akan dimintai pertanggungjawaban (Tafsir Al-Muyassar, n.d.).

Hasil penelitian ini menunjukkan bahwa pemanfaatan model Random Forest dalam menentukan prioritas perbaikan jalan dapat mendukung pelaksanaan amanah tersebut. Dengan adanya sistem klasifikasi yang objektif dan berbasis data, pengambilan keputusan perbaikan jalan dapat dilakukan secara lebih bertanggung jawab dan transparan. Ruas jalan yang memiliki tingkat kerusakan tinggi dapat diidentifikasi secara jelas sehingga penanganannya dapat diprioritaskan sesuai dengan tingkat urgensi dan dampaknya terhadap masyarakat.

Pendekatan ini membantu mengurangi potensi penyalahgunaan wewenang maupun keputusan yang bersifat subjektif dalam pengelolaan infrastruktur jalan. Dengan demikian, penerapan hasil penelitian ini tidak hanya berkontribusi pada peningkatan efektivitas perencanaan perbaikan jalan, tetapi juga mencerminkan pelaksanaan amanah dan tanggung jawab sebagaimana yang diajarkan dalam Islam. Keputusan yang diambil berdasarkan hasil klasifikasi model dapat

dipertanggungjawabkan secara ilmiah sekaligus secara moral, sehingga selaras dengan nilai-nilai keislaman dalam pengelolaan fasilitas publik.

Selain prinsip keadilan dan amanah, tujuan utama dari penerapan nilai-nilai Islam dalam pengelolaan infrastruktur publik adalah tercapainya kemaslahatan umat (*maslahah*), yaitu terwujudnya manfaat yang luas bagi masyarakat serta terhindarnya dari berbagai risiko dan kerugian. Dalam konteks pengelolaan infrastruktur publik, kemaslahatan umat tercermin dalam upaya menyediakan fasilitas jalan yang aman, nyaman, dan layak digunakan oleh seluruh lapisan masyarakat. Oleh karena itu, penentuan prioritas perbaikan jalan harus diarahkan untuk memberikan manfaat sebesar-besarnya bagi masyarakat serta meminimalkan risiko yang dapat membahayakan pengguna jalan.

Tujuan utama dari penerapan nilai-nilai Islam dalam kehidupan bermasyarakat adalah terwujudnya kemaslahatan umat (*maslahah*), yaitu terciptanya manfaat yang luas serta terhindarnya masyarakat dari kerugian dan kemudatan. Dalam konteks pengelolaan infrastruktur publik, kemaslahatan umat tercermin dalam upaya menyediakan fasilitas jalan yang aman, nyaman, dan layak digunakan oleh seluruh lapisan masyarakat. Oleh karena itu, penentuan prioritas perbaikan jalan harus diarahkan untuk memberikan manfaat sebesar-besarnya bagi masyarakat serta meminimalkan risiko yang dapat membahayakan pengguna jalan.

Konsep kemaslahatan ini sejalan dengan firman Allah SWT dalam Al-Qur'an surat Al-Baqarah ayat 195:

وَأَنْفِقُوا فِي سَبِيلِ اللَّهِ وَلَا تُلْقُوا بِأَيْدِيكُمْ إِلَى التَّهْلُكَةِ ۚ وَأَحْسِنُوا ۚ إِنَّ اللَّهَ يُحِبُّ الْمُحْسِنِينَ

“Dan belanjakanlah (harta bendamu) di jalan Allah, dan janganlah kamu menjatuhkan dirimu sendiri ke dalam kebinasaan, dan berbuat baiklah, karena sesungguhnya Allah menyukai orang-orang yang berbuat baik.” (QS. Al-Baqarah: 195)

Berdasarkan Tafsir Al-Muyassar, ayat tersebut menjelaskan bahwa Allah SWT memerintahkan manusia untuk berinfak dan berusaha di jalan-Nya serta melarang perbuatan yang dapat menjerumuskan diri ke dalam kebinasaan. Kebinasaan dalam ayat ini dapat dimaknai sebagai segala tindakan yang merugikan diri sendiri maupun orang lain, termasuk kelalaian dalam menjaga keselamatan dan kesejahteraan masyarakat. Selain itu, manusia juga diperintahkan untuk senantiasa berbuat baik (*ihsan*) dan melaksanakan setiap perbuatan dengan penuh keikhlasan, karena Allah SWT menyukai orang-orang yang berbuat baik (Tafsir Al-Muyassar, n.d.).

Hasil penelitian ini menunjukkan bahwa klasifikasi prioritas perbaikan jalan menggunakan model *Random Forest* dapat menjadi sarana untuk mewujudkan kemaslahatan umat. Dengan mengidentifikasi ruas jalan yang memiliki tingkat kerusakan tinggi sebagai prioritas utama, potensi risiko kecelakaan dan gangguan mobilitas dapat diminimalkan. Sebaliknya, ruas jalan dengan kondisi yang masih baik atau tingkat kerusakan rendah dapat direncanakan perbaikannya secara bertahap, sehingga penggunaan sumber daya dapat dilakukan secara lebih efisien dan tepat sasaran.

Secara keseluruhan, integrasi nilai kemaslahatan umat dalam penelitian ini menegaskan bahwa penggunaan metode ilmiah dan teknologi, seperti *Random Forest*, dapat berjalan seiring dengan nilai-nilai keislaman. Hasil penelitian tidak

hanya memberikan solusi teknis dalam penentuan prioritas perbaikan jalan, tetapi juga berkontribusi pada upaya mewujudkan kesejahteraan dan keselamatan masyarakat sebagai tujuan utama pembangunan infrastruktur menurut perspektif Islam.

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan mengenai penerapan algoritma Random Forest untuk klasifikasi prioritas perbaikan jalan di Kabupaten Malang, dapat disimpulkan bahwa metode yang digunakan mampu memberikan hasil klasifikasi yang akurat serta mendukung proses pengambilan keputusan dalam penentuan prioritas perbaikan jalan. Adapun kesimpulan yang diperoleh dari penelitian ini adalah sebagai berikut.

1. Algoritma Random Forest berhasil digunakan untuk melakukan klasifikasi prioritas perbaikan jalan di Kabupaten Malang berdasarkan data kondisi jalan, panjang jalan, dan Lalu Lintas Harian Rata-Rata (LHR). Hasil evaluasi menunjukkan bahwa model memiliki performa yang sangat baik dengan nilai accuracy berkisar antara 0,9875 hingga 1,0, serta didukung oleh nilai precision, recall, dan F1-score yang tinggi pada nilai *n_estimators* optimal. Hasil tersebut menunjukkan bahwa Random Forest mampu mengenali pola data prioritas perbaikan jalan secara efektif dan menghasilkan klasifikasi yang akurat untuk kategori prioritas rendah, sedang, dan tinggi.
2. Berdasarkan analisis feature importance menggunakan metode Gini Importance, faktor yang paling berpengaruh dalam menentukan prioritas perbaikan jalan adalah kondisi jalan, terutama pada kategori jalan sedang

dan jalan baik. Selain itu, panjang jalan, nilai LHR, dan kondisi rusak berat juga memberikan kontribusi dalam proses klasifikasi. Hasil ini menunjukkan bahwa kondisi jalan merupakan faktor utama yang dipertimbangkan oleh model dalam menentukan tingkat prioritas perbaikan jalan di Kabupaten Malang, sehingga informasi tersebut dapat digunakan sebagai dasar pendukung pengambilan keputusan yang lebih objektif dan tepat sasaran. Berdasarkan hasil evaluasi menggunakan confusion matrix, model Random Forest memperoleh performa yang tinggi pada seluruh metrik evaluasi. Nilai accuracy mencapai 0,9875 hingga 1,0, sedangkan nilai precision, recall, dan F1-score juga menunjukkan hasil yang sangat baik pada nilai $n_estimators$ optimal. Hasil tersebut menunjukkan bahwa Random Forest mampu mengenali pola data prioritas perbaikan jalan secara efektif dan menghasilkan prediksi yang akurat.

5.2 Saran

1. Penelitian selanjutnya disarankan untuk menambahkan variabel lain yang berpotensi memengaruhi prioritas perbaikan jalan, seperti tingkat kecelakaan, kondisi drainase, faktor lingkungan, serta kepadatan lalu lintas, agar model mampu menangkap pola data secara lebih komprehensif.
2. Perlu dilakukan perbandingan dengan metode machine learning lain, seperti Gradient Boosting, XGBoost, maupun algoritma ensemble lainnya untuk memperoleh model dengan performa yang lebih optimal.

3. Disarankan untuk menerapkan teknik penanganan data tidak seimbang (imbalanced dataset), seperti oversampling (SMOTE), undersampling, atau class weighting, guna meningkatkan kemampuan model dalam mengenali kelas minoritas.
4. Pengembangan sistem berbasis machine learning ini diharapkan dapat terus dilanjutkan guna mendukung pengambilan keputusan yang lebih objektif, transparan, serta meningkatkan efektivitas alokasi anggaran dan keselamatan pengguna jalan.
5. Hasil verifikasi dengan pihak Dinas PUPR menunjukkan bahwa penentuan prioritas perbaikan jalan tidak hanya dipengaruhi oleh kondisi fisik jalan, tetapi juga mempertimbangkan faktor strategis lainnya, seperti struktur dan fungsi jalan, ketersediaan anggaran pemeliharaan, kebijakan pemerintah pusat maupun daerah, serta peran jalan dalam mendukung konektivitas antarwilayah. Oleh karena itu, penelitian selanjutnya disarankan untuk mengintegrasikan variabel-variabel tersebut ke dalam model klasifikasi agar sistem yang dikembangkan mampu merepresentasikan proses pengambilan keputusan secara lebih komprehensif dan realistis.

DAFTAR PUSTAKA

- Abdur Rahim, Aulia Unnisa, Salwa Haifa, & Harun Al-Rasyid. (2024). Analisis Surah Al-Baqarah Ayat 185 dan Surah Al-A'raf Ayat 26 dalam Kajian Al-Istikhdam dan Al-Istithrad. *Perspektif: Jurnal Pendidikan Dan Ilmu Bahasa*, 3(1), 27–33. <https://doi.org/10.59059/perspektif.v3i1.2043>
- Albertini, C., Gioia, A., Iacobellis, V., Petropoulos, G. P., & Manfreda, S. (2024). Assessing multi-source random forest classification and robustness of predictor variables in flooded areas mapping. *Remote Sensing Applications: Society and Environment*, 35, 101239. <https://doi.org/10.1016/j.rsase.2024.101239>
- Aprillianto, A. D. (2022). *Analisis Pemetaan Skala Prioritas Penanganan Jalan Dengan Metode Analytical Hierarchy Process Di Kota Tarakan*. Universitas Borneo Tarakan. (Perpustakaan UBT). [//repository.ubt.ac.id/index.php/Fid%3D9402%26keywords%3D%26p%3Dshow_detail%26utm_source%3Dchatgpt.com](http://repository.ubt.ac.id/index.php/Fid%3D9402%26keywords%3D%26p%3Dshow_detail%26utm_source%3Dchatgpt.com)
- Ashfania, G. A. M., Prahasto, T., Widodo, A., & Warsokusumo, T. (2022). Penggunaan Algoritma Random Forest untuk Klasifikasi berbasis Kinerja Efisiensi Energi pada Sistem Pembangkit Daya. *ROTASI*, 24(3), 14–21. <https://doi.org/10.14710/rotasi.24.3.14-21>
- Az-Zahriyah, F. A., Sunariyanto, S., & Putra, L. R. (2025). Evaluasi Kebijakan Pemerintah Kota Malang Dalam Target Perbaikan Jalan Rusak Dan Berlubang (Studi Pada Dinas Pekerjaan Umum, Penataan Ruang, Perumahan dan Kawasan Permukiman Kota Malang). *Respon Publik*, 19(2), 41–49.
- Borrouhou, S., Fissoune, R., & Badir, H. (2023). Data cleaning survey and challenges – improving outlier detection algorithm in machine learning. *Journal of Smart Cities and Society*, 2, 125–140. <https://doi.org/10.3233/SCS-230008>
- Dineshkumar, R., Aluvala, S., Srinath, S., Alsalami, Z., & Krishnaprasath, V. T. (2024). A Semiconductor Manufacturing Final Test Yield Classification Using Random Forest. *2024 Second International Conference on Data Science and Information System (ICDSIS)*, 1–4. <https://doi.org/10.1109/ICDSIS61070.2024.10594688>
- Erlinawati, E. (2024). Kerja Sama Dalam Lembaga Pendidikan Analisi Tafsir Q.S. Al-Maidah:2. *Jurnal Sosial Humaniora Sigli*, 7(2). <https://doi.org/10.47647/jsh.v7i2.2916>

- Fitri, A., Marjono, M., & Raharjo, N. D. (2024). Analisis Kerusakan Perkerasan Jalan Dengan Metode Bina Marga Pada Jalan Raya Gondanglegi-Turen Kabupaten Malang. *Jurnal Online Skripsi Manajemen Rekayasa Konstruksi (JOS-MRK)*, 5(2), 159–165. <https://doi.org/10.33795/jos-mrk.v5i2.4131>
- Helmud, E., Helmud, E., Fitriyani, F., & Romadiana, P. (2024). Classification Comparison Performance of Supervised Machine Learning Random Forest and Decision Tree Algorithms Using Confusion Matrix. *Jurnal Sisfokom (Sistem Informasi Dan Komputer)*, 13(1), 92–97. <https://doi.org/10.32736/sisfokom.v13i1.1985>
- Ibnu Katsir. (n.d.-a). Tafsir QS. Al-Qamar Ayat 49. *TafsirWeb*. Retrieved March 28, 2026, from <https://tafsirweb.com/10287-surat-al-qamar-ayat-49.html>
- Ibnu Katsir. (n.d.-b). Tafsir QS. An-Nahl Ayat 90. *TafsirWeb*. Retrieved March 28, 2026, from <https://tafsirweb.com/4438-surat-an-nahl-ayat-90.html>
- Justam, J., Jamilah, N., Umar, S. M., Erlita, E., & Ramba, J. (2024). Penerapan Algoritma C4.5 dan Random Forest untuk Pemetaan Kerusakan Jalan dengan WebGIS. *Jurnal Ilmiah Sistem Informasi dan Teknik Informatika (JISTI)*, 7(2), 326–339. <https://doi.org/10.57093/jisti.v7i2.270>
- Karabadji, N. E. I., Korba, A. A., Assi, A., Seridi-Bouchelaghem, H., Aridhi, S., & Dhifli, W. (2023). Accuracy and diversity-aware multi-objective approach for random forest construction. *Expert Syst. Appl.*, 225. <https://doi.org/10.1016/j.eswa.2023.120138>
- Kim, A., Myung, J., & Kim, H. (2020). *Random forest ensemble using a weight-adjusted voting algorithm*. 31, 427–438. <https://doi.org/10.7465/jkdi.2020.31.2.427>
- Lemenkova, P. (2024). Random Forest Classifier Algorithm of Geographic Resources Analysis Support System Geographic Information System for Satellite Image Processing: Case Study of Bight of Sofala, Mozambique. *Coasts*, 4(1), 127–149. <https://doi.org/10.3390/coasts4010008>
- Maulindar, J., & Guterres, J. X. (2025). The Implementation of Random Forest Classification for Identifying Infrastructure Priorities in Smart Cities. *Jurnal Pekommas*, 10(1). <https://doi.org/10.56873/jpkm.v9i1.5841>
- Naufal, A., & Prasetyo, E. (n.d.). *Sistem Pendukung Keputusan Penerima Bantuan Zakat Menggunakan Random Forest dan Fuzzy Analytical Hierarchy Process | Jurnal Komputer Teknologi Informasi Sistem Informasi (JUKTISI)*. Retrieved September 22, 2025, from <https://ejurnal.lkparyaprima.id/index.php/juktisi/article/view/510>

- Priyanto, T., Karyawan, I. D. M. A., & Mahendra, M. (2023). Penentuan Skala Prioritas Dan Pola Penanganan Kerusakan Jalan Lingkungan Kawasan Permukiman Kumuh Di Kota Mataram. *GANEC SWARA*, 17(2), 726. <https://doi.org/10.35327/gara.v17i2.479>
- Putri, M. Y., & Susanto, E. R. (2025). Prioritas Perbaikan Jalan Menggunakan Metode Ahp Pada Dinas Pupr Kabupaten Way Kanan. *JUPI (Jurnal Ilmiah Penelitian Dan Pembelajaran Informatika)*, 10(1), 599–610. <https://doi.org/10.29100/jipi.v10i1.5546>
- Rahman, F. D., Mulki, M. I. Z., & Taryana, A. (2024). Clustering Dan Klasifikasi Data Cuaca Cilacap Dengan Menggunakan Metode K-Means Dan Random Forest. *Jurnal SINTA: Sistem Informasi Dan Teknologi Komputasi*, 1(2), 90–97. <https://doi.org/10.61124/sinta.v1i2.15>
- Rahman, Lukman, & Hildayati. (2025). Sistem Klasifikasi Kerusakan Jalan Metode Machine Learning dengan Algoritma K-Means dan Random Forest. *Arus Jurnal Sains dan Teknologi*, 3(1), 116–126. <https://doi.org/10.57250/ajst.v3i1.1212>
- Sitanggang, P. B., Mulia, A. P., & Nasution, Z. P. (2022). Analisis Prioritas Pemeliharaan Jalan Kabupaten Dairi dengan Metode AHP dan GIS. *Jurnal Syntax Admiration*, 3(2), 388–403. <https://doi.org/10.46799/jsa.v3i2.391>
- Tafsif Al-Muyassar. (n.d.). Tafseer al-Muyassar Arabic King Fahad Complex. *Alhamdolillah*. Retrieved October 13, 2025, from <https://alhamdolillah.com/book/tafseer-al-muyassar-arabic-king-fahad-complex/>
- Yuliana, H., & Rizqiana, Z. C. H. (2024). Analisis Metode Klasifikasi Pemetaan Tutupan Lahan (Land Cover) di Area Kota Bandung Menggunakan Algoritma Random Forest Pada Google Earth Engine. *EPSILON: Journal of Electrical Engineering and Information Technology*, 22(2), 97–106. <https://doi.org/10.55893/epsilon.v22i2.119>

LAMPIRAN

LAMPIRAN

Lampiran 1 - Verifikasi Data Ke Dinas PUPR Kab Malang



Link wawancara verifikasi penelitian:

<https://drive.google.com/file/d/1oIB0SfSCjjZfDIbyOVQKrAqFg6-fvlpV/view?usp=sharing>

<https://youtu.be/ftag51oIHsE?si=li1pkJkxhhcnlpkf>



VERIFIKASI HASIL PENELITIAN
KLASIFIKASI PRIORITAS PERBAIKAN JALAN DI
KABUPATEN MALANG MENGGUNAKAN ALGORITMA
RANDOM FOREST

DINAS PUPR KAB MALANG

Izza Syahri Muharram

220605110073

Q	1. Bagaimana Dinas PUPR Kabupaten Malang menentukan prioritas perbaikan jalan yang akan ditangani setiap tahunnya? - Faktor apa yang paling berpengaruh dalam menentukan prioritas?
A	- Kondisi jalan - Status jalan (jalan nasional, provinsi, kabupaten) - Analisis lalu lintas - Ketersediaan Anggaran - Kegiatan pusat dan daerah - Aspirasi Masyarakat.
Q	2. Apakah terdapat pedoman atau standar tertentu yang digunakan dalam menentukan prioritas perbaikan jalan? - Apakah mengacu pada pedoman Bina Marga?
A	- Pedoman Bina Marga dan kementerian PU - Standar kondisi jalan - Peraturan daerah dan kegiatan pusat. - Proses perencanaan daerah pembangunan Daerah.
Q	3. Dalam penelitian ini digunakan variabel kondisi jalan, panjang jalan, dan LHR. Menurut Bapak/Ibu, apakah ketiga variabel tersebut sudah relevan untuk menentukan prioritas perbaikan jalan? - Apakah ada variabel penting lain yang seharusnya dipertimbangkan? - Jika ada, variabel apa saja?
A	1. Status dan fungsi jalan. 2. Ketersediaan Anggaran. 3. konektivitas wilayah. 4. kegiatan pusat dan Daerah.
Q	4. Menurut pengalaman Bapak/Ibu, apakah kondisi jalan dan LHR sudah cukup merepresentasikan kebutuhan penanganan suatu ruas jalan?
A	Belum Cukup. masih membutuhkan Variabel lain untuk

	Penilaian keputusan prioritas perbaikan. - Struktur & fungsi jalan. - Kejangkauan pusat dan Daerah. - Ketersediaan Anggaran. - Aspek keselamatan lalu lintas.
Q	5. Berdasarkan hasil klasifikasi yang telah dihasilkan sistem, apakah pengelompokan prioritas pada ruas jalan tersebut sudah sesuai dengan kondisi yang biasa digunakan oleh Dinas PUPR?
A	1. Struktur dan fungsi jalan. 2. Ketersediaan Anggaran. 3. Kejangkauan pusat / Daerah. Sudah sesuai Apabila kepermih.
Q	6. Apakah hasil yang diberikan sistem dapat dijadikan referensi awal dalam menentukan prioritas perbaikan jalan?
A	Referensi Sekunder dan masih perlu ditambahkan Variabel lagi. Karena sistem kami akan diterapkan pada survey. Jalan Nasional (IRMS) dan jalan kab/pwu (PKRMS)
Q	7. Apakah hasil klasifikasi yang dihasilkan sistem secara umum sudah sesuai dengan pertimbangan teknis yang digunakan oleh Dinas PUPR Kabupaten Malang dalam menentukan prioritas perbaikan jalan?
A	Secara Umum bisa dijadikan untuk referensi juga. untuk penerapan di PUPR masih perlu data lain. dan pengujian.

Q	8. Berdasarkan hasil yang telah ditunjukkan, bagaimana pendapat Bapak/Ibu mengenai kelayakan sistem klasifikasi prioritas perbaikan jalan yang dibangun dalam penelitian ini?
A	Secara garis besar bagus → Sebagai referensi tambahan. tapi untuk penerapan perlu di coba dan diteliti lebih. lebih lanjut menggunakan saran saya di awal.

Malang, 15 Juni 2026

Mahasiswa Ybs.



Izza Syahri Muharram
NIM.220605110073

Penata Laksana Jalan & Jembatan Terampil
Dinas PUPR Kabupaten Malang



Haris Sumarta, A.Md
NIPK. 198709282025211012

Lampiran 2 – Hasil Klasifikasi Prioritas Perbaikan Jalan Di Kabupaten Malang

Data Ke	Nama Ruas	Aktual	Prediksi	Keterangan
1	Jln. Imam Soedjono (Wonosari)	1	1	Prioritas Sedang
2	Bumirejo - Baturetno	1	1	Prioritas Sedang
3	Sumbersekar - Dalisodo	0	0	Prioritas Rendah
4	Karangploso - Giripurno (Bts. Kota Batu)	1	1	Prioritas Sedang
5	Jln. Sitirejo - Bts Kota Malang	2	2	Prioritas Tinggi
6	Jln. Cepoko	1	1	Prioritas Sedang
7	Kaweden - Sekar	0	0	Prioritas Rendah
8	Krebet - Wajak	1	1	Prioritas Sedang
9	Jln. Hasanudin (Donomulyo)	1	1	Prioritas Sedang
10	Jln. Teuku Umar	0	0	Prioritas Rendah
11	Selorejo - Krisik	0	0	Prioritas Rendah
12	Jln. Murcoyo (Gondanglegi)	0	0	Prioritas Rendah
13	Kalipare - Sumberpetung	1	1	Prioritas Sedang
14	Jln. Suropati I (Bululawang)	1	1	Prioritas Sedang
15	Jln. Sumber Sekar	2	2	Prioritas Tinggi
16	Jln. Tumapel IV (Singosari)	0	0	Prioritas Rendah
17	Donomulyo - Jalan Lintas Selatan	1	1	Prioritas Sedang
18	Jln. KH. Ahmad Dahlan (Kepanjen)	0	0	Prioritas Rendah
19	Talok - Tulusbesar	1	1	Prioritas Sedang
20	Wonokerto - Bantur	1	1	Prioritas Sedang
21	Jln. Pajajaran	1	1	Prioritas Sedang
22	Singosari - Klampok	0	0	Prioritas Rendah
23	Jln. Karangrejo	2	2	Prioritas Tinggi
24	Tirtomarto - Mulyoasri	1	1	Prioritas Sedang
25	Kebonagung - Sidorahayu	1	1	Prioritas Sedang
26	Harjokuncaran - Gedokwetan	1	1	Prioritas Sedang
27	Jln. Sidoluhur II (Kepanjen)	1	1	Prioritas Sedang
28	Mojosari - Dilem	1	1	Prioritas Sedang
29	Jln. W. Monginsidi	1	1	Prioritas Sedang
30	Jln. Tunggal Ika	1	1	Prioritas Sedang
31	Jln. Sumber Pacar	1	1	Prioritas Sedang
32	Jln. Kauman (Sumberpucung)	1	1	Prioritas Sedang
33	Jln. Sutojoyo	2	2	Prioritas Tinggi
34	Jln. Punten	0	0	Prioritas Rendah
35	Bantur Krajan - Jalan Lintas Selatan	1	1	Prioritas Sedang

36	Sukowilangun - Kalirejo (Bts. Kab. Blitar)	0	0	Prioritas Rendah
37	Jln. Sekar Putih	0	0	Prioritas Rendah
38	Jln. Gajahmada Pagak	1	1	Prioritas Sedang
39	Srigonco - Pantai Kondang Merak	1	1	Prioritas Sedang
40	Jln. Raya Sutojayan - Raya Wonokerso	1	1	Prioritas Sedang
41	Jln. Dorowati	1	1	Prioritas Sedang
42	Jln. Salak II (Turen)	0	0	Prioritas Rendah
43	Jln. Wijaya Kusuma	1	1	Prioritas Sedang
44	Jln. Sumber Waras	1	1	Prioritas Sedang
45	Ngebruk - Senggreng	0	0	Prioritas Rendah
46	Jln. Kudus	1	1	Prioritas Sedang
47	Jln. Dukuh Mangir	1	1	Prioritas Sedang
48	Sukodono - Pantai Sidoasri	1	1	Prioritas Sedang
49	Jln. Ahmad Yani (Jabung)	1	1	Prioritas Sedang
50	Buring (Bts. Kota Malang) - Kidal	2	2	Prioritas Tinggi
51	Toyomarto - Ketindan	1	1	Prioritas Sedang
52	Druju - Wonokerto	1	1	Prioritas Sedang
53	Jln. Margo Besuki	0	0	Prioritas Rendah
54	Dawung - Jalan Lintas Selatan	1	1	Prioritas Sedang
55	Jln. Regulo	1	1	Prioritas Sedang
56	Pandanajeng - Slamet	1	1	Prioritas Sedang
57	Jln. Efendi	1	1	Prioritas Sedang
58	Sitiarjo - Sidodadi	1	1	Prioritas Sedang
59	Jln. Slamet Riyadi	0	0	Prioritas Rendah
60	Jln. Ronggowuni (Tumpang)	1	1	Prioritas Sedang
61	Rembun - Jambangan	0	0	Prioritas Rendah
62	Jln. Sidomulyo (Singosari)	1	1	Prioritas Sedang
63	Slorok - Jambuwer	1	1	Prioritas Sedang
64	Tambakrejo - Pantai Tamban	2	2	Prioritas Tinggi
65	Mulyoarjo - Sumbergepuh	1	1	Prioritas Sedang
66	Jatiguwi - Trenggeng	0	0	Prioritas Rendah
67	Jln. Nusa Indah (Tumpang)	1	1	Prioritas Sedang
68	Codo - Garotan	1	1	Prioritas Sedang
69	Tajinan - Kidangbang	0	0	Prioritas Rendah
70	Sudimoro - Pringu	2	2	Prioritas Tinggi
71	Jln. Lingkar Ngajum	1	1	Prioritas Sedang
72	Jln. Untung Suropati	0	0	Prioritas Rendah
73	Kapurono - Perمانو	1	1	Prioritas Sedang
74	Jalan Lintas Selatan-Pantai Wonogoro	1	1	Prioritas Sedang

75	Jln. Sidorejo Raya	0	0	Prioritas Rendah
76	Jln. Kenongosari IV	1	1	Prioritas Sedang
77	Jln. Kenongosari	1	1	Prioritas Sedang
78	Gedangan - Bajulmati	0	0	Prioritas Rendah
79	Kepuharjo - Tegalondo (Bts. Kota Malang)	1	1	Prioritas Sedang
80	Tangsi-Sipelot	1	1	Prioritas Sedang
81	Kedungrejo - Kambingan	1	1	Prioritas Sedang
82	Bululawang - Sumbersuko	1	1	Prioritas Sedang
83	Jln. Tunggulametung	0	0	Prioritas Rendah
84	Jln. Ketanen Penarukan (Kepanjen)	1	1	Prioritas Sedang
85	Jln. Pasuruan	1	1	Prioritas Sedang
86	Jalan Lintas Selatan - Pantai Sendang Biru	1	1	Prioritas Sedang
87	Gedog Wetan - Undaan	2	2	Prioritas Tinggi
88	Jln. Adi Utomo	0	0	Prioritas Rendah
89	Arjosari - Tumpakrejo	1	1	Prioritas Sedang
90	Jln. Timur Pasar Kepanjen	1	1	Prioritas Sedang
91	Jln. Demak	0	0	Prioritas Rendah
92	Jln. Nusa Indah (Kepanjen)	0	0	Prioritas Rendah
93	Jln. Patimura	1	1	Prioritas Sedang
94	Ardimulyo - Toyomarto	0	0	Prioritas Rendah
95	Ngijo-Ngenep	2	2	Prioritas Tinggi
96	Dampit - Sukodono	0	0	Prioritas Rendah
97	Rembun - Gedokwetan	1	1	Prioritas Sedang
98	Jln. Tegalondo	1	1	Prioritas Sedang
99	Jln. Suropati III (Bululawang)	0	0	Prioritas Rendah
100	Banjararum - Kepuharjo	0	0	Prioritas Rendah
101	Sedayu - Sidorejo	0	0	Prioritas Rendah
102	Karangwidoro - Candi Badut (Bts. Kota Malang)	1	1	Prioritas Sedang
103	Jln. Hayam Wuruk II (Gondanglegi)	1	1	Prioritas Sedang
104	Jln. Imam Bonjol (Pagelaran)	0	0	Prioritas Rendah
105	Jln. Anggrek	0	0	Prioritas Rendah
106	Pandesari - Jurangrejo	2	2	Prioritas Tinggi
107	Pakisaji - Maduarjo	1	1	Prioritas Sedang
108	Jln. Cerme	2	2	Prioritas Tinggi
109	Jeru - Jabung	0	0	Prioritas Rendah
110	Jln. Adiwijaya	1	1	Prioritas Sedang
111	Jln. Hasanudin (Pagelaran)	2	2	Prioritas Tinggi
112	Jln. Ampeldento - Kepuharjo	1	1	Prioritas Sedang

113	Jln. Hasyim Ashari	2	2	Prioritas Tinggi
114	Jln. Sidodadi (Jabung)	1	1	Prioritas Sedang
115	Sukolilo - Pringu	1	1	Prioritas Sedang
116	Baturetno - Sumbertangkil	1	1	Prioritas Sedang
117	Mendalanwangi - Pakisaji	0	0	Prioritas Rendah
118	Jln. Lingkar Pasar Ampelgading	0	0	Prioritas Rendah
119	Jln. Diponegoro IV (Bululawang)	1	1	Prioritas Sedang
120	Klepu - Sumbersuko	0	0	Prioritas Rendah
121	Jln. Sidomulyo III (Bululawang)	0	0	Prioritas Rendah
122	Pakisaji - Karangduren	1	1	Prioritas Sedang
123	Jln. Pujon Lor - Pujon Kidul	0	0	Prioritas Rendah
124	Jln. Bayangkara Turen	1	1	Prioritas Sedang
125	Banjarejo - Sengguruh	0	0	Prioritas Rendah
126	Asrikaton - Banjararum	2	2	Prioritas Tinggi
127	Dengkol - Tirtomoyo	2	2	Prioritas Tinggi
128	Putatlor - Bulupitu	0	0	Prioritas Rendah
129	Kesamben - Balesari	0	0	Prioritas Rendah
130	Jln. Lingkar Wonosari	1	1	Prioritas Sedang
131	Sepanjang - Talangsuko	1	1	Prioritas Sedang
132	Jln. Lingkar Wagir	0	0	Prioritas Rendah
133	Jln. Tugu Mulyo	1	1	Prioritas Sedang
134	Landungsari - Tegalweru	2	2	Prioritas Tinggi
135	Jln. Karyawan	1	1	Prioritas Sedang
136	Jln. Pandesari - Songgoriti	1	1	Prioritas Sedang
137	Pakis - Sumberkradenan	0	0	Prioritas Rendah
138	Jln. Tirto Utomo	0	0	Prioritas Rendah
139	Wonorejo - Ketindan	1	1	Prioritas Sedang
140	Jln. Dampit - Pamotan	1	1	Prioritas Sedang
141	Jln. Krepyak Jaya	0	0	Prioritas Rendah
142	Talangsuko - Codo	1	1	Prioritas Sedang
143	Kambal - Selorejo	2	2	Prioritas Tinggi
144	Kebobang - Balesari	0	0	Prioritas Rendah
145	Jln. Diponegoro (Lawang)	1	1	Prioritas Sedang
146	Jln. Kertanegara (Tumpang)	1	1	Prioritas Sedang
147	Talangsuko - Kidangbang	1	1	Prioritas Sedang
148	Jln. Imam Bonjol - Hasyim Asyari	1	1	Prioritas Sedang
149	Pamotan - Garotan	1	1	Prioritas Sedang
150	Jln. Pendem - Batas Kota	1	1	Prioritas Sedang
151	Sekarpuro - Bunut	1	1	Prioritas Sedang
152	Jln. Simpang Ngurawan	1	1	Prioritas Sedang
153	Pagak - Sumbermanjing Kulon	1	1	Prioritas Sedang

154	Turen - Ketawang	0	0	Prioritas Rendah
155	Mulyoagung - Landungsari	1	1	Prioritas Sedang
156	Gedangan - Bantur	0	0	Prioritas Rendah
157	Jln. Anjasmara	1	1	Prioritas Sedang
158	Sidorejo - Sidomulyo	1	1	Prioritas Sedang
159	Sukopuro - Taji	0	0	Prioritas Rendah
160	Jln. Hayam Wuruk I (Gondanglegi)	0	0	Prioritas Rendah
161	Jln. Yos Sudarso	0	0	Prioritas Rendah
162	Jln. Jatirejoyoso - Gelanggang	2	2	Prioritas Tinggi
163	Jln. Pajang	1	1	Prioritas Sedang
164	Jln. Gunung Ceneng	1	1	Prioritas Sedang
165	Bantur - Sumbermanjing Kulon	0	0	Prioritas Rendah
166	Blayu - Patokpicis	1	1	Prioritas Sedang
167	Baturetno - Sidodadi	1	1	Prioritas Sedang
168	Tlogosari - Jogomulyan	0	0	Prioritas Rendah
169	Permanu - Babadan	1	1	Prioritas Sedang
170	Jln. Sumbersuko	0	0	Prioritas Rendah
171	Langlang - Klampok	0	0	Prioritas Rendah
172	Jln. Argo Puro	1	1	Prioritas Sedang
173	Jln. Abd. Rahman Saleh	1	1	Prioritas Sedang
174	Ngajum - Balesari	0	0	Prioritas Rendah
175	Jln. Pemuda	1	1	Prioritas Sedang
176	Jln. Salak (Turen)	2	2	Prioritas Tinggi
177	Lawang - Sumberporong	1	1	Prioritas Sedang
178	Jln. Pasar - Tawang Sari	1	1	Prioritas Sedang
179	Tulusbesar - Gubugklakah	1	1	Prioritas Sedang
180	Jln. Ringinsari - Sidodadi	1	1	Prioritas Sedang
181	Jln. Tawang Sari - Sumberwangi Kidul	1	1	Prioritas Sedang
182	Sumbertangkil - Lebakharjo	1	1	Prioritas Sedang
183	Jln. Gajahmada (Gondanglegi)	1	1	Prioritas Sedang
184	Jln. Kartini (Gedangan)	0	0	Prioritas Rendah
185	Jln. Sekarpuro - Wendit	1	1	Prioritas Sedang
186	Karangates - Kalipare	2	2	Prioritas Tinggi
187	Gondanglegi - Clumprit	0	0	Prioritas Rendah
188	Beringin - Bambang	0	0	Prioritas Rendah
189	Jln. Stadion Utara	1	1	Prioritas Sedang
190	Jln. Golek	1	1	Prioritas Sedang
191	Jln. KH. Syihabuddin	0	0	Prioritas Rendah
192	Banjararum - Mangliawan	0	0	Prioritas Rendah
193	Sananrejo - Sanankerto	1	1	Prioritas Sedang

194	Jln. Masjid (Singosari)	1	1	Prioritas Sedang
195	Bedali - Srigading	1	1	Prioritas Sedang
196	Jln. Mantung - Ngabab	2	2	Prioritas Tinggi
197	Jln. Brawijaya Gagak Asinan	0	0	Prioritas Rendah
198	Jln. Hasanudin (Tumpang)	0	0	Prioritas Rendah
199	Jln. Terusan Wisnuwardhana - Mangliawan	2	2	Prioritas Tinggi
200	Jln. Doho (Dampit)	0	0	Prioritas Rendah
201	Rejoyoso - Bantur	1	1	Prioritas Sedang
202	Jln. Ken Arok (Tumpang)	1	1	Prioritas Sedang
203	Jln. Sarangan	1	1	Prioritas Sedang
204	Jln. Melati Mangunrejo (Kepanjen)	1	1	Prioritas Sedang
205	Sumbermanjing Wetan - Gedangan	0	0	Prioritas Rendah
206	Jalan Lintas Selatan - Pantai Modangan	1	1	Prioritas Sedang
207	Jln. Pandesari (Pujon)	2	2	Prioritas Tinggi
208	Krebet Senggrong - Pringu	1	1	Prioritas Sedang
209	Candirenggo - Toyomarto	1	1	Prioritas Sedang
210	Jln. Banurejo	0	0	Prioritas Rendah
211	Jln. Jenggolo - Segaluh	2	2	Prioritas Tinggi
212	Jln. Sidomulyo (Kepanjen)	0	0	Prioritas Rendah
213	Jln. Pasar Bakroto	2	2	Prioritas Tinggi
214	Jln. Sidomulyo I (Bululawang)	1	1	Prioritas Sedang
215	Ngajum - Gunung Kawi	1	1	Prioritas Sedang
216	Jln. Penanggungan	1	1	Prioritas Sedang
217	Krebet - Gondanglegi	0	0	Prioritas Rendah
218	Pandesari - Sukomulyo	0	0	Prioritas Rendah
219	Babadan - Sumbersuko	1	1	Prioritas Sedang
220	Maguan - Ngasem	1	1	Prioritas Sedang
221	Jeru - Benjor	1	1	Prioritas Sedang
222	Jln. Hasanudin - Kedungpendaringan	1	1	Prioritas Sedang
223	Pandanlandung (Bts. Kota Malang) - Kalisongo	1	1	Prioritas Sedang
224	Segaran - Gedangan	1	1	Prioritas Sedang
225	Jln. Saptorenggo - Ampeldento	1	1	Prioritas Sedang
226	Jln. Welirang	1	1	Prioritas Sedang
227	Jln. Teuku Umar (Kepanjen)	1	1	Prioritas Sedang
228	Jln. Mentaraman	0	0	Prioritas Rendah
229	Amadanom - Gadungsari	1	1	Prioritas Sedang
230	Jln. Kyai Mojo (Gondanglegi)	0	0	Prioritas Rendah
231	Kepanjen - Kendalpayak	1	1	Prioritas Sedang

232	Jln. Sidomulyo IV (Bululawang)	0	0	Prioritas Rendah
233	Tlogosari - Mulyoasri	1	1	Prioritas Sedang
234	Kedungrejo (Bts. Kota Malang) - Tumpang	0	0	Prioritas Rendah
235	Kepanjen - Pagak	1	1	Prioritas Sedang
236	Jln. KH. Gozali	1	1	Prioritas Sedang
237	Jln. Adidarma	1	1	Prioritas Sedang
238	Jln. Kenanga	2	2	Prioritas Tinggi
239	Jln. Sido Utomo - Jatiorejoso	1	1	Prioritas Sedang
240	Jln. Tegal Mapan	1	1	Prioritas Sedang
241	Sumberoto - Jalan Lintas Selatan	1	1	Prioritas Sedang
242	Jln. Sumber Ayu	2	2	Prioritas Tinggi
243	Bumirejo - Sukodono	0	0	Prioritas Rendah
244	Donomulyo - Wates (Bts. Kab. Blitar)	0	0	Prioritas Rendah
245	Sumbermanjing Kulon - Pantai Kondang Iwak	0	0	Prioritas Rendah
246	Sengkaling - Junrejo (Bts. Kota Batu)	1	1	Prioritas Sedang
247	Jln. Nusantara	1	1	Prioritas Sedang
248	Ngenep - Sumbul	0	0	Prioritas Rendah
249	Jln. KH. Agus Salim	1	1	Prioritas Sedang
250	Tirtomoyo - Mulyoasri	1	1	Prioritas Sedang
251	Sumberejo (Bts. Kota Malang) - Petungsewu	2	2	Prioritas Tinggi
252	Jln. Cokroaminoto	1	1	Prioritas Sedang
253	Jatikerto - Plandi	2	2	Prioritas Tinggi
254	Jln. Pudak - Talangagung	1	1	Prioritas Sedang
255	Gunungsari - Argosuko	0	0	Prioritas Rendah
256	Dampit - Gadungsari	2	2	Prioritas Tinggi
257	Jln. Argo Tunggal	1	1	Prioritas Sedang
258	Ngadilangkung - Ngajum	1	1	Prioritas Sedang
259	Jln. Cempaka	1	1	Prioritas Sedang
260	Tirtomarto - Sidorenggo	0	0	Prioritas Rendah
261	Jln. Sultan Hasanuddin	0	0	Prioritas Rendah
262	Jln. Sunan Kalijogo	1	1	Prioritas Sedang
263	Jln. Mantung - Tawang Sari	1	1	Prioritas Sedang
264	Jln. R. A. Kartini (Lawang)	1	1	Prioritas Sedang
265	Jln. Diponegoro III (Bululawang)	1	1	Prioritas Sedang
266	Bandungrejo - Srigonco	1	1	Prioritas Sedang
267	Jalan Lingkar Barat Kepanjen	0	0	Prioritas Rendah
268	Sempalwadak - Pringu	1	1	Prioritas Sedang

269	Jln. Ds. Tempur - Kedungpendaringan	2	2	Prioritas Tinggi
270	Jln. Sido Makmur	0	0	Prioritas Rendah
271	Jln. Sido Topo	0	0	Prioritas Rendah
272	Kalipare - Donomulyo	1	1	Prioritas Sedang
273	Jln. Wijaya	0	0	Prioritas Rendah
274	Jln. Ronggowuni (Singosari)	0	0	Prioritas Rendah
275	Tamanasri - Pantai Licin	2	2	Prioritas Tinggi
276	Bantur -Jalan Lintas Selatan	0	0	Prioritas Rendah
277	Kluwut - Karangrejo	2	2	Prioritas Tinggi
278	Jalan Lintas Selatan-Ngliyep	0	0	Prioritas Rendah
279	Jln. Kromengan - Sumberpucung	1	1	Prioritas Sedang
280	Jln. Lingkar Bantur	2	2	Prioritas Tinggi
281	Klepu - Tambakasri	2	2	Prioritas Tinggi
282	Jln. Sidomukti II (Bululawang)	1	1	Prioritas Sedang
283	Jln. Lucari	1	1	Prioritas Sedang
284	Jln. Adi Santoso	1	1	Prioritas Sedang
285	Jln. Malang	1	1	Prioritas Sedang
286	Wonomulyo - Poncokusumo	0	0	Prioritas Rendah
287	Jln. Masjid (Dampit)	1	1	Prioritas Sedang
288	Jln. Hamid Rusdi	0	0	Prioritas Rendah
289	Sumbersuko (Bts. Kota Malang) - Tumpang	1	1	Prioritas Sedang
290	Jln. Sebaluh - Cobanrondo	0	0	Prioritas Rendah
291	Jln. Belakang Pasar	0	0	Prioritas Rendah
292	Tambakrejo - Jalan Lingkar Selatan	1	1	Prioritas Sedang
293	Jln. Keramat (Kepanjen)	0	0	Prioritas Rendah
294	Jln. Semeru (Kepanjen)	0	0	Prioritas Rendah
295	Jln. TGP (Sumberpucung)	0	0	Prioritas Rendah
296	Jln. Sultan Agung (Pagelaran)	0	0	Prioritas Rendah
297	Jln. Tawangsari 1 dan 2 (Lawang)	0	0	Prioritas Rendah
298	Jln. TMP Utara	1	1	Prioritas Sedang
299	Pagelaran - Kademangan	0	0	Prioritas Rendah
300	Kalipare - Sukowilangun	2	2	Prioritas Tinggi
301	Wonomulyo - Ngingit	1	1	Prioritas Sedang
302	Jln. Adi Mulyo	1	1	Prioritas Sedang
303	Kalisongo - Kucur	1	1	Prioritas Sedang
304	Jln. Melaten	0	0	Prioritas Rendah
305	Sukosari - Putukrejo	0	0	Prioritas Rendah
306	Ngadirejo - Plaosan	2	2	Prioritas Tinggi
307	Karangploso - Klampok	1	1	Prioritas Sedang

308	Jln. Sentong Lama	0	0	Prioritas Rendah
309	Jln. Kauman III	0	0	Prioritas Rendah
310	Gubugklakah- Ngadas (Bts. Kab. Lumajang)	1	1	Prioritas Sedang
311	Jln. Flamboyan	0	0	Prioritas Rendah
312	Jln. Sapta Marga - Jenggolo	1	1	Prioritas Sedang
313	Jln. Sultan Agung (Tumpang)	1	1	Prioritas Sedang
314	Jln. Aris Sujono	0	0	Prioritas Rendah
315	Jln. Bintoro (Dampit)	1	1	Prioritas Sedang
316	Purwodadi - Wediawu	0	0	Prioritas Rendah
317	Jln. Sidorejo I	1	1	Prioritas Sedang
318	Tumpakrejo - Sumberoto	1	1	Prioritas Sedang
319	Wonomulyo - Sumberejo	0	0	Prioritas Rendah
320	Tawangrenjeni - Sidorejo	1	1	Prioritas Sedang
321	Lawang - Gunung Jati	1	1	Prioritas Sedang
322	Pujon - Mantung	2	2	Prioritas Tinggi
323	Gadungsari - Sukorejo	1	1	Prioritas Sedang
324	Bedali - Sidodadi	1	1	Prioritas Sedang
325	Jln. Lawu (Kepanjen)	2	2	Prioritas Tinggi
326	Jln. Sidoagung	0	0	Prioritas Rendah
327	Jln. Tirto Taruno	0	0	Prioritas Rendah
328	Pagak - Kalipare	0	0	Prioritas Rendah
329	Jln. Pungkur Argo	0	0	Prioritas Rendah
330	Jln. Kramat	1	1	Prioritas Sedang
331	Pakis - Kemiri (Bts. Kab. Pasuruan)	1	1	Prioritas Sedang
332	Gondanglegi Wetan - Sepanjang	1	1	Prioritas Sedang
333	Kluwut - Gunung Kawi	1	1	Prioritas Sedang
334	Arjowilangun - Arjosari (Bts. Kab. Blitar)	0	0	Prioritas Rendah
335	Tanggung - Sidorejo	1	1	Prioritas Sedang
336	Jln. Suropati (Pagelaran)	1	1	Prioritas Sedang
337	Lawang - Wonosari	0	0	Prioritas Rendah
338	Jln. Suropati II (Bululawang)	1	1	Prioritas Sedang
339	Jln. Bengawan Solo (Wajak)	1	1	Prioritas Sedang
340	Asrikaton - Bandara Abdul Rahman Saleh	1	1	Prioritas Sedang
341	Jln. Stadion	0	0	Prioritas Rendah
342	Donomulyo - Sumbermanjing Kulon	0	0	Prioritas Rendah
343	Jln. Sidodadi (Kepanjen)	0	0	Prioritas Rendah
344	Gedangan - Tumpakrejo	1	1	Prioritas Sedang
345	Jln. Pahlawan - Summersuko	1	1	Prioritas Sedang

346	Jln. Krapyak	0	0	Prioritas Rendah
347	Sumbertempur - Kalilekso	1	1	Prioritas Sedang
348	Wringinanom - Sumberejo	1	1	Prioritas Sedang
349	Jln. Petung	0	0	Prioritas Rendah
350	Singosari - Toyomarto	1	1	Prioritas Sedang
351	Lawang - Wonorejo	2	2	Prioritas Tinggi
352	Sempol - Putukrejo	0	0	Prioritas Rendah
353	Jln. Sidoluhur	0	0	Prioritas Rendah
354	Jln. Argo Moyo	1	1	Prioritas Sedang
355	Jln. Sumber Kembar	1	1	Prioritas Sedang
356	Singosari - Jabung	2	2	Prioritas Tinggi
357	Karangsuko - Brongkal	1	1	Prioritas Sedang
358	Wirotaman - Jogomulyan	1	1	Prioritas Sedang
359	Jln. Ken Arok (Singosari)	1	1	Prioritas Sedang
360	Dengkol - Sukolilo	0	0	Prioritas Rendah
361	Jln. Indrakilo	1	1	Prioritas Sedang
362	Jln. Kertarajasa (Singosari)	0	0	Prioritas Rendah
363	Lebakharjo-Krajan Dua	1	1	Prioritas Sedang
364	Talangagung - Ngajum	2	2	Prioritas Tinggi
365	Malangsuko - Slamet	0	0	Prioritas Rendah
366	Jln. Hayam Wuruk III (Gondanglegi)	1	1	Prioritas Sedang
367	Jln. Sembojo	0	0	Prioritas Rendah
368	Putukrejo - Sukoraharjo	0	0	Prioritas Rendah
369	Jln. Suruji	1	1	Prioritas Sedang
370	Jln. Mandukoro	0	0	Prioritas Rendah
371	Jln. Kertanegara - Kendedes	1	1	Prioritas Sedang
372	Jln. Setiawan	1	1	Prioritas Sedang
373	Jln. Lingkar Stasiun Singosari	2	2	Prioritas Tinggi
374	Kedok - Sananrejo	0	0	Prioritas Rendah
375	Jln. Salak I (Turen)	0	0	Prioritas Rendah
376	Jedong (Bts. Kota Malang) - Pandanrejo	1	1	Prioritas Sedang
377	Jln. Sidodadi II (Bululawang)	0	0	Prioritas Rendah
378	Jln. Lingkar Pagak	0	0	Prioritas Rendah
379	Bululawang - Telogowaru (Bts. Kota Malang)	1	1	Prioritas Sedang
380	Pagelaran - Sidorejo	1	1	Prioritas Sedang
381	Gampingan - Dempok	1	1	Prioritas Sedang
382	Pandanlandung - Gondowangi	1	1	Prioritas Sedang
383	Mendalanwangi - Summersuko	2	2	Prioritas Tinggi
384	Jln. Wahidin Sudiro Husodo	1	1	Prioritas Sedang

385	Jln. Letjen Panjaitan (Gondanglegi)	2	2	Prioritas Tinggi
386	Jalan Lintas Selatan - Pantai Jonggring	0	0	Prioritas Rendah
387	Jln. A. Yani Turen	2	2	Prioritas Tinggi
388	Jln. Bromo	1	1	Prioritas Sedang
389	Jln. KH. Mustofa	1	1	Prioritas Sedang
390	Wonokerto - Sumberejo	0	0	Prioritas Rendah
391	Jln. Argo Bimo	1	1	Prioritas Sedang
392	Kedungbanteng - Tambakasri	2	2	Prioritas Tinggi
393	Gondanglegi - Wonokerto	0	0	Prioritas Rendah
394	Lumbangsari - Sukorejo	1	1	Prioritas Sedang
395	Tumpang - Duwet	0	0	Prioritas Rendah
396	Jln. Tawang Argo	0	0	Prioritas Rendah
397	Wagir - Gunung Kawi	1	1	Prioritas Sedang
398	Jalan Lintas Selatan- Pantai Balekambang	1	1	Prioritas Sedang
399	Jln. Kamboja	1	1	Prioritas Sedang