

MODEL OTOMATISASI CERDAS BERBASIS DATA SPASIAL UNTUK PRIORITAS PERBAIKAN JALAN RUSAK

Usman Nurhasan¹, Hendra Pradibta²

^{1,2} Jurusan, Teknologi Informasi, Politeknik Negeri Malang
usmannurhasan@polinema.ac.id, ² hendrapradibta@polinema.ac.id

Abstrak

Kerusakan jalan berdampak signifikan terhadap keselamatan dan efisiensi sosial-ekonomi. Pendekatan manual dalam menentukan prioritas perbaikan seringkali tidak akurat, terutama di wilayah dengan karakteristik spasial kompleks seperti Kabupaten Pasuruan. Penelitian ini mengusulkan sistem rekomendasi cerdas berbasis data spasial menggunakan model *hybrid*, yang mengintegrasikan metode *Simple Additive Weighting* (SAW) untuk perankingan dan *Random Forest* (RF) untuk validasi klasifikasi. Dataset mencakup 100 segmen jalan dengan atribut spasial dan non-spasial. SAW berhasil mengklasifikasikan segmen ke dalam tiga kelas prioritas dengan tingkat konsistensi 92%, yang didukung oleh hasil RF. Model RF juga menunjukkan kinerja baik dengan *precision*, *recall*, dan *F1-score* rata-rata sebesar 0,87, serta mampu mengidentifikasi fitur penting. Integrasi SAW-RF dalam kerangka GIS terbukti efektif meningkatkan efisiensi dan transparansi. Prototipe ini berpotensi untuk direplikasi di daerah lain menggunakan platform *open source*.

Kata kunci: Kerusakan Jalan; GIS, SAW, *Random Forest*, Sistem Rekomendasi

1. Pendahuluan

Kerusakan infrastruktur jalan merupakan permasalahan penting dalam pembangunan wilayah karena berdampak langsung pada aspek keselamatan transportasi, efisiensi distribusi logistik, dan kelancaran aktivitas sosial-ekonomi masyarakat. Tantangan utama dalam konteks ini tidak hanya terletak pada identifikasi segmen jalan yang mengalami kerusakan, tetapi juga pada penetapan skala prioritas perbaikan secara tepat, objektif, dan efisien (Khan et al., 2024). Kondisi ini menjadi semakin kompleks di wilayah dengan cakupan geografis luas dan karakteristik yang beragam, seperti Kabupaten Pasuruan, di mana proses perencanaan dan penanganan perbaikan jalan masih didominasi oleh pendekatan manual dan subjektivitas petugas lapangan (Firouraghi et al., 2022).

Akibat dari mekanisme perencanaan yang tidak berbasis data tersebut adalah potensi terjadinya ketidaktepatan dalam alokasi anggaran, keterlambatan penanganan kerusakan, serta ketimpangan dalam pelayanan infrastruktur antarwilayah. Dampak tersebut dirasakan secara langsung oleh masyarakat sebagai pengguna jalan, serta oleh pemerintah daerah yang memiliki tanggung jawab dalam pengambilan keputusan strategis terkait pengelolaan dan perbaikan infrastruktur jalan. Permasalahan ini semakin menonjol pada jaringan jalan sekunder dan tersier yang belum terintegrasi

dalam sistem pemantauan maupun dalam basis data informasi geospasial daerah (Adam et al., 2022).

Seiring dengan meningkatnya volume data baik yang bersumber dari laporan masyarakat, hasil survei lapangan, maupun data spasial teknis dari citra satelit dan peta jaringan jalan pendekatan analisis konvensional menjadi tidak lagi memadai. Oleh karena itu, dibutuhkan sistem pendukung keputusan yang adaptif, terotomatisasi, dan mampu mengolah data dalam skala besar secara efisien. Dalam konteks ini, pendekatan *Machine Learning* (ML) memiliki keunggulan dalam mengidentifikasi pola tersembunyi dari data multivariabel, serta memberikan hasil rekomendasi yang bersifat dinamis, akurat, dan berkelanjutan (Fang et al., 2023), (Jigar Shah, 2023). Kemampuan ML untuk memproses data numerik, kategorikal, dan spasial secara bersamaan menjadikannya sangat relevan dalam pengambilan keputusan berbasis data. Integrasi antara *Machine Learning* dan Sistem Informasi Geografis (GIS) memungkinkan pengembangan model rekomendasi yang lebih komprehensif (Singh, 2024). GIS berperan dalam mengelola, menyimpan, dan memvisualisasikan data spasial, sementara algoritma ML—seperti *Random Forest* dan metode pengambilan keputusan multikriteria seperti *Simple Additive Weighting* (SAW)—digunakan untuk menghitung dan mengurutkan prioritas perbaikan jalan berdasarkan kombinasi data spasial dan atribut kerusakan. Pendekatan ini memberikan peluang untuk menghasilkan sistem rekomendasi yang

bersifat transparan, objektif, dan dapat divalidasi secara visual oleh pengambil kebijakan.

Penelitian ini bertujuan untuk mengembangkan Model Otomatisasi Rekomendasi Cerdas Berbasis Analisis Data Spasial guna mendukung penentuan prioritas perbaikan jalan rusak di Kabupaten Pasuruan. Model ini dibangun dengan pendekatan *hybrid* yang menggabungkan kekuatan GIS dalam pemrosesan data spasial dan Machine Learning dalam otomatisasi klasifikasi dan perankingan. Data yang digunakan mencakup citra satelit, OpenStreetMap, laporan masyarakat, serta hasil survei lapangan, dan dianalisis menggunakan PostgreSQL + PostGIS serta divisualisasikan melalui Leaflet JS (Anugraha et al., 2022).

Kontribusi dari penelitian ini dapat dijabarkan dalam empat dimensi utama yang saling melengkapi. Pertama, dari sisi konseptual, penelitian ini menghadirkan suatu kerangka kerja integratif yang menggabungkan kekuatan *Geographic Information System* (GIS) dan algoritma *Machine Learning* (ML) sebagai pendekatan komprehensif dalam pengambilan keputusan terkait perbaikan infrastruktur jalan (Alkhaleel, 2024; Sui et al., 2023). Kedua, secara metodologis, penelitian ini memperkenalkan alur analisis spasial yang sistematis dengan memanfaatkan metode *Simple Additive Weighting* (SAW) sebagai teknik penentuan prioritas berbasis multikriteria. Ketiga, dalam ranah praktis, penelitian ini menghasilkan prototipe sistem rekomendasi yang dapat diimplementasikan oleh pemerintah daerah sebagai alat bantu dalam merencanakan perbaikan jalan secara lebih efisien, akurat, dan berbasis data. Keempat, dari aspek teknologis, model ini dirancang menggunakan arsitektur perangkat lunak berbasis sumber terbuka (*open source*), sehingga memberikan fleksibilitas untuk pengembangan lanjutan secara berkelanjutan serta meminimalkan ketergantungan terhadap sistem berlisensi.

Berdasarkan tujuan tersebut, penelitian ini merumuskan dua pertanyaan utama: RQ1: Bagaimana integrasi antara Sistem Informasi Geografis (GIS) dan algoritma *Machine Learning* dapat dimanfaatkan untuk melakukan klasifikasi dan pengurutan segmen jalan yang rusak secara otomatis berdasarkan kombinasi atribut spasial dan non-spasial? RQ2: Sejauh mana model *hybrid* yang dikembangkan dalam penelitian ini mampu meningkatkan efisiensi, objektivitas, dan transparansi dalam proses penentuan prioritas perbaikan jalan pada studi kasus wilayah Kabupaten Pasuruan?

2. Metode

Pada gambar 1 ditunjukkan model yang akan dikembangkan.



Gambar 1. Model Hybrid Untuk Sistem Rekomendasi Prioritas Perbaikan Jalan Rusak

Penelitian ini menggunakan metode pengembangan model *hybrid* berbasis *Simple Additive Weighting* (SAW) dan *Random Forest* untuk sistem rekomendasi prioritas perbaikan jalan rusak, dengan mengintegrasikan data spasial dan non-spasial. Metode ini termasuk dalam kategori *hybrid multi-criteria decision making* (MCDM) yang memadukan pendekatan pembobotan deterministik dan prediksi berbasis machine learning guna meningkatkan akurasi dan pemrosesan variabel kompleks dalam pengambilan keputusan infrastruktur.

2.1. Akuisisi dan Praproses Data

Tahap awal dalam pengembangan sistem rekomendasi dimulai dengan proses akuisisi dan integrasi data spasial serta non-spasial yang bersumber dari lembaga pemerintah dan hasil observasi lapangan. Dataset yang dikumpulkan mencakup lima kategori utama, yaitu: (1) data fisik kerusakan jalan, yang terdiri dari informasi panjang, lebar, jenis kerusakan, serta material permukaan jalan; (2) data spasial, berupa koordinat geografis, kedekatan ruas jalan terhadap fasilitas publik penting (seperti sekolah, rumah sakit, dan pasar), serta estimasi kepadatan permukiman; (3) data lalu lintas, yang mencakup volume kendaraan harian sebagai indikator tingkat penggunaan jalan; (4) data sosial-ekonomi, seperti jumlah keluhan masyarakat serta tingkat urgensi akses transportasi pada kawasan terkait; dan (5) data teknis dan administratif, seperti nilai usulan anggaran pemeliharaan serta frekuensi intervensi teknis sebelumnya (Akindele et al., 2023; Firouraghi et al., 2022).

Seluruh data yang terkumpul kemudian melalui tahap *preprocessing* untuk memastikan konsistensi dan kualitasnya. Proses ini meliputi normalisasi nilai numerik menggunakan metode *min-max scaling* agar seluruh fitur berada pada skala yang sebanding; imputasi terhadap nilai hilang dengan menggunakan teknik median atau mode; serta transformasi data spasial menjadi bentuk numerik, misalnya melalui perhitungan jarak Euclidean terhadap titik-titik strategis seperti fasilitas kesehatan

atau pusat pendidikan. Selain itu, atribut kategorikal seperti jenis permukaan jalan dikonversi ke dalam bentuk numerik melalui teknik encoding, seperti *one-hot encoding* atau label *encoding*, tergantung pada kompleksitas kategorinya (Putri & Nazhifah, 2022)..

2.2. Pemeringkatan Awal dengan Simple Additive Weighting (SAW)

Simple Additive Weighting (SAW) merupakan salah satu pendekatan paling umum dalam metode pengambilan keputusan multikriteria (*Multi-Criteria Decision Making/MCDM*). Keunggulan utama dari metode ini terletak pada kesederhanaannya dalam struktur perhitungan serta transparansinya dalam proses evaluasi alternatif (Nitayaprapha et al., 2025). Metode ini diterapkan sebagai pendekatan awal untuk menentukan tingkat prioritas perbaikan jalan berdasarkan kombinasi nilai multikriteria. Setiap segmen jalan dinilai berdasarkan enam kriteria utama yang mewakili dimensi fisik, spasial, dan sosial-ekonomi, dengan penilaian akhir dihitung dari hasil perkalian antara nilai yang telah dinormalisasi dan bobot masing-masing kriteria. Nilai dari masing-masing kriteria dinormalisasi untuk menyamakan skala, kemudian dikalikan dengan bobot kepentingan yang ditentukan, dan dijumlahkan untuk menghasilkan skor akhir. Skor inilah yang menjadi dasar peringkat awal prioritas perbaikan jalan. SAW bekerja dengan menjumlahkan seluruh nilai kriteria yang telah dinormalisasi, kemudian dikalikan dengan bobot preferensi yang sesuai untuk menghasilkan skor total dari masing-masing alternatif. Langkah-langkah dalam metode SAW dapat dijabarkan sebagai berikut:

1. Menentukan bobot preferensi untuk setiap kriteria w_j , dengan ketentuan bahwa total bobot seluruh kriteria harus bernilai satu:

$$\sum_{j=1}^n w_j = 1$$

2. Menyusun matriks keputusan $X = [x_{ij}]$, di mana x_{ij} merupakan nilai alternatif ke- i terhadap kriteria ke- j .
3. Melakukan normalisasi terhadap nilai x_{ij} menjadi r_{ij} , dengan mempertimbangkan jenis kriteria:

Untuk kriteria bertipe benefit:

$$r_{ij} = \frac{x_{ij}}{\max(x_j)}$$

Untuk kriteria bertipe cost:

$$r_{ij} = \frac{\min(x_j)}{x_{ij}}$$

4. Menghitung skor akhir dari masing-masing alternatif S_i berdasarkan kombinasi bobot dan nilai normalisasi:

$$S_i = \sum_{j=1}^n w_j \cdot r_{ij}$$

Mengurutkan nilai S_i untuk menentukan prioritas, di mana alternatif dengan skor tertinggi dipandang sebagai pilihan paling optimal. Hasil penelitian pendahulu menunjukkan bahwa metode SAW memiliki keunggulan dalam hal kejelasan logika perhitungan, serta fleksibel untuk diintegrasikan ke dalam sistem informasi geografis (GIS). Hal ini memungkinkan hasil pemeringkatan ditampilkan dalam bentuk peta digital yang interaktif, yang sangat bermanfaat dalam konteks pengambilan keputusan berbasis lokasi seperti dalam manajemen pemeliharaan infrastruktur jalan.

Pada Tabel 2 ditunjukkan kriteria yang digunakan.

Tabel 2. Kriteria SAW

No.	Kriteria	Jenis	Simbol	Deskripsi Singkat
1	Tingkat kerusakan	Benefit	C_1	Makin parah kerusakan, makin tinggi prioritas
2	Volume lalu lintas	Benefit	C_2	Mewakili intensitas penggunaan jalan
3	Usulan biaya perbaikan	Cost	C_3	Biaya rendah diprioritaskan sebagai bentuk efisiensi
4	Akses ke fasilitas publik	Benefit	C_4	Semakin dekat ke fasilitas vital, semakin prioritas
5	Kepadatan permukiman	Benefit	C_5	Area padat menunjukkan kebutuhan layanan infrastruktur lebih tinggi
6	Laporan masyarakat	Benefit	C_6	Indikator partisipasi dan persepsi kebutuhan masyarakat

2.3. Random Forest Classifier

Random Forest merupakan algoritma pembelajaran mesin berbasis *ensemble*. Algoritma ini membangun sejumlah model *decision tree* secara paralel, kemudian menggabungkan hasil prediksinya untuk menghasilkan keputusan akhir (Trianasari & Permadi, 2024). Dalam konteks klasifikasi, *Random Forest* menggunakan prinsip mayoritas suara (*majority voting*), sementara pada regresi digunakan nilai rata-rata (*mean aggregation*) dari seluruh prediksi pohon. Pendekatan Random Forest sangat sesuai untuk pemodelan sistem kompleks dengan jumlah variabel prediktor yang banyak, termasuk kombinasi fitur spasial dan non-spasial yang umum dijumpai dalam sistem informasi geografis (GIS). Selain itu, algoritma ini dikenal robust terhadap overfitting, mampu menangani data tidak linear, dan menyediakan *metrik feature importance* yang berguna dalam penentuan variabel paling berpengaruh terhadap hasil model (Choudhury et al., 2024).

2024). Proses kerja Random Forest secara umum terdiri atas beberapa tahapan utama:

1. Menentukan mayoritas voting

$$\hat{y} = \text{mode}(\{h_1(x), h_2(x), \dots, h_T(x)\})$$

Dimana:

\hat{y} = hasil prediksi akhir

$h_t(x)$ = prediksi dari pohon ke- t

T = jumlah total pohon

mode = nilai yang paling sering muncul (untuk klasifikasi)

2. Pembobotan Fitur

$$G_{ini} = 1 - \sum_{i=1}^C p_i^2$$

Dimana:

C = jumlah kelas

p_i = proporsi item dari kelas ke- i dalam node

3. Perhitungan Akurasi Random Forest

$$\text{Accuracy} = \frac{\text{Jumlah prediksi benar}}{\text{Total Prediksi}}$$

Dalam studi pendahulu, *Random Forest* telah berhasil digunakan dalam sistem peringkat dan klasifikasi prioritas perbaikan infrastruktur jalan. Data spasial seperti lokasi kerusakan, intensitas lalu lintas, serta jarak terhadap fasilitas publik diolah sebagai variabel masukan untuk memprediksi tingkat urgensi atau risiko kerusakan. Hasil prediksi tersebut kemudian divisualisasikan secara spasial dalam platform GIS untuk mendukung pengambilan keputusan yang bersifat lokasi spesifik dan data-driven.

Setelah pemeringkatan awal diperoleh melalui metode SAW, hasil klasifikasi (dalam bentuk label prioritas: Tinggi, Sedang, dan Rendah) digunakan sebagai target kelas (label ground truth) dalam proses pelatihan model *Random Forest Classifier* (Madhusudhan et al., 2021), (Raja et al., 2024). Pendekatan ini bertujuan untuk mengevaluasi ulang dan menyempurnakan pemeringkatan awal dengan mempertimbangkan kompleksitas hubungan antar variabel secara lebih adaptif. Model *Random Forest* dipilih karena kemampuannya dalam:

1. Mengakomodasi interaksi nonlinier antar fitur yang tidak dapat ditangkap secara optimal oleh metode deterministik seperti SAW,
2. Menyediakan analisis *feature importance*, sehingga dapat mengidentifikasi fitur mana yang paling berpengaruh terhadap hasil klasifikasi, dan
3. Meningkatkan akurasi melalui mekanisme ensemble learning yang menggabungkan prediksi dari banyak pohon keputusan secara agregatif.

Dataset pelatihan untuk *Random Forest* mencakup:

1. Seluruh variabel input yang digunakan dalam metode SAW,

2. Variabel turunan spasial, seperti kepadatan kerusakan per kilometer dan frekuensi kerusakan historis di wilayah tertentu,
3. Label klasifikasi hasil SAW yang digunakan sebagai acuan target awal (baseline ground truth).
4. Melalui pendekatan ini, sistem dapat mengidentifikasi inkonsistensi atau misalignment antara pemeringkatan awal SAW dan pola prediktif yang lebih kompleks, sekaligus menyediakan dasar evaluasi berbasis data historis dan hubungan spasial.

2.4. Integrasi Model Hybrid

(6)

Model *hybrid* yang diterapkan dalam penelitian ini terdiri atas dua fase integrative yaitu:

Fase 1 – Penilaian Awal dengan SAW (*Simple Additive Weighting*)

Pada tahap ini, metode SAW digunakan untuk menghitung skor prioritas awal berdasarkan pembobotan eksplisit terhadap setiap kriteria yang telah ditentukan. SAW berfungsi sebagai mekanisme preskriptif yang menyederhanakan proses pengambilan keputusan awal dengan memberikan nilai agregat terhadap alternatif berdasarkan nilai dan bobot kriteria.

Fase 2 – Evaluasi dan Koreksi dengan Random Forest

Hasil skor dari SAW kemudian dievaluasi menggunakan algoritma *Random Forest Classifier*. Tahap ini berfungsi untuk memverifikasi dan menyesuaikan klasifikasi prioritas berdasarkan pola empiris dari data historis serta mempertimbangkan karakteristik kontekstual, seperti kondisi spasial atau temporal.

3. Hasil dan Pembahasan

Penelitian ini menggunakan 100 data segmen jalan sebagai objek analisis, di mana masing-masing data merepresentasikan satu ruas jalan dengan atribut-atribut yang telah dijelaskan pada bab sebelumnya. Dalam pengumpulan data, penelitian ini menghadapi sejumlah tantangan yang berimplikasi pada kualitas hasil analisis. Pertama, ketersediaan data spasial masih terbatas dan tidak selalu konsisten, sehingga diperlukan validasi manual melalui survei lapangan. Kedua, perbedaan kualitas dan format data antar instansi menimbulkan potensi bias, khususnya pada pencatatan volume lalu lintas dan estimasi biaya perbaikan. Ketiga, pengolahan data yang bersifat heterogen membutuhkan proses pembersihan yang kompleks untuk mengatasi *missing values* dan *outliers*. Keempat, keterbatasan akses terhadap data *real-time* menyebabkan hasil analisis kurang adaptif terhadap perubahan kondisi lapangan.

Selain itu, metode *Simple Additive Weighting* (SAW) yang digunakan memiliki keterbatasan, yaitu asumsi independensi antar kriteria dan bobot yang

bersifat statis. Dalam konteks nyata, variabel dapat saling berinteraksi dan tingkat kepentingannya dapat berubah sesuai situasi, sehingga interpretasi hasil perlu dilakukan dengan mempertimbangkan keterbatasan tersebut. Untuk menggambarkan proses perhitungan, Tabel 3 menampilkan 10 sampel awal dari keseluruhan data. Tabel tersebut mencakup nilai asli masing-masing kriteria, hasil normalisasi, skor akhir dari metode Simple Additive Weighting (SAW), serta peringkat prioritas yang dihasilkan.

Tabel 3. Sample Data Awal dan Hasil SAW

Ind ex	C 1	C 2	C 3	C4	C 5	C 6	SAW_S core
0	4	1	2	32.021.971.406	1	4	0.3699
		2	0	.390.100	5		
		7	9		3		
1	5	2	1	17.612.459.761	5	1	0.5372
		3	4	.386.300	4	2	
		4	5		3		
2	3	3	2	18.111.269.156	7	3	0.3902
		0	8	.020.300	1		
		0	2		2		
3	5	9	2	36.571.828.264	3	1	0.7257
		3	2	.641.700	6	5	
		9	9		3		
4	5	8	1	44.958.402.737	1	1	0.7166
		7	6	.676.200	5	5	
		9	2		2		
5	2	1	3	4.446.723.478.	6	1	0.3661
		3	6	899.070	7		
		2	7		1		
6	3	1	4	3.921.390.174.	7	1	0.5469
		4	9	702.350	1	6	
		7	1		9		
7	3	6	1	32.459.550.661	1	1	0.5950
		0	0	.560.100	0	9	
		2	1		4		
8	3	5	3	0.51228582847	2	1	0.3933
		0	1	57393	0	1	
		6	7		2		
9	5	6	3	0.89198069906	2	1	0.5910
		7	4	36076	9	7	
		3	4		5		

Setelah dilakukan perhitungan skor menggunakan metode SAW, langkah berikutnya adalah melakukan klasifikasi skor ke dalam tiga kelas prioritas. Tujuan dari pengelompokan ini adalah untuk menyederhanakan interpretasi hasil serta mempermudah proses pengambilan keputusan strategis terkait urutan penanganan perbaikan. Klasifikasi dilakukan dengan pendekatan persentil, membagi 100 data ke dalam tiga kelas berdasarkan peringkat hasil SAW. Hasil klasifikasi ditunjukkan dalam Tabel 4.

Tabel 4. Ringkasan Klasifikasi Prioritas Berdasarkan Skor SAW

Kelas Prioritas	Kriteria Pemeringkatan	Jumlah Data	Persent ase
Prioritas Tinggi	Peringkat 1 – 30	30	30%
Prioritas Sedang	Peringkat 31 – 70	40	40%
Prioritas Rendah	Peringkat 71 – 100	30	30%

Sebagai bagian dari validasi hasil klasifikasi SAW, dilakukan pelatihan model Random Forest Classifier dengan label target berdasarkan hasil klasifikasi SAW. Tabel 5 berikut menyajikan contoh prediksi pada 10 data uji acak.

Tabel 5. Hasil Prediksi *Random Forest* terhadap Label SAW

I D	C 1	C 2	C 3	C 4	C 5	C 6	Label SAW (Aktual)	Predi ksi RF	St at us
D 0 1	4	8	3	0	9	1	Tinggi	Tingg i	Be na r
		0	0	.	0	2			
		0	0	5	0				
D 0 2	2	3	4	1	6	5	Sedang	Seda ng	Be na r
		0	0	.	0				
D 0 3	5	1	1	0	1	2	Tinggi	Tingg i	Be na r
		0	5	.	0	0			
		0	0	2	0				
		0	0	0	0				
D 0 4	3	6	2	2	7	1	Sedang	Rend ah	Sa lah
		0	0	.	0	0			
		0	0	0	0				
D 0 5	1	2	5	3	5	0	Rendah	Rend ah	Be na r
		0	0	.	0				
		0	0	0	0				
D 0 6	4	9	1	0	8	1	Tinggi	Tingg i	Be na r
		5	0	.	5	5			
		0	0	3	0				
D 0 7	2	5	2	1	7	8	Sedang	Seda ng	Be na r
		0	5	.	5				
		0	0	8	0				
D 0 8	3	4	3	2	6	6	Sedang	Seda ng	Be na r
		0	5	.	5				
		0	0	2	0				
D 0 9	1	1	4	3	4	2	Rendah	Seda ng	Sa lah
		5	5	.	0				
		0	0	5	0				
D 1 0	5	9	9	0	9	1	Tinggi	Tingg i	Be na r
		8	0	.	9	8			
		0	1	0					

Untuk mengevaluasi kinerja model Random Forest, digunakan confusion matrix dan classification report sebagaimana ditunjukkan pada Tabel 6 dan 7.

Tabel 6. Confusion Matrix Prediksi Random Forest dan Label SAW

Aktual \ Prediksi	Tinggi	Sedang	Rendah	Total
Tinggi	27	3	0	30
Sedang	2	35	3	40
Rendah	1	4	25	30
Total	30	42	28	100

Tabel 7. *Classification Report* – Evaluasi Kinerja Random Forest

Kelas Prioritas	Precisio n	Recal l	F1- Score	Suppor t
Tinggi	0.90	0.90	0.90	30
Sedang	0.83	0.88	0.85	40
Rendah	0.89	0.83	0.86	30
Rata-rata	0.87	0.87	0.87	100

Model *Random Forest* menunjukkan performa klasifikasi yang solid dengan tingkat kesalahan rendah. Nilai *precision* dan *recall* yang tinggi pada seluruh kelas menandakan bahwa model bekerja secara seimbang dalam mengenali semua kategori prioritas. Dengan *F1-score* rata-rata sebesar 0,87, model ini terbukti mampu mereplikasi hasil klasifikasi metode SAW secara konsisten dan dapat diandalkan untuk klasifikasi otomatis data baru.

Berdasarkan klasifikasi prioritas dari metode *Simple Additive Weighting* (SAW) yang telah tervalidasi melalui model *Random Forest*, disusun rekomendasi kebijakan untuk mempercepat dan mengoptimalkan penanganan perbaikan infrastruktur jalan. Pada Prioritas Tinggi, segmen ini memiliki kerusakan berat, lalu lintas padat, dekat fasilitas publik, dan berada di kawasan padat penduduk. Disarankan untuk segera diperbaiki dalam waktu maksimal tiga bulan, masuk dalam anggaran darurat atau perubahan, serta melibatkan koordinasi lintas instansi. Sedangkan pada prioritas sedang, ditandai oleh kerusakan menengah dan/atau biaya perbaikan tinggi. Penanganan dapat dijadwalkan dalam program tahunan (4–12 bulan), disertai evaluasi berkala dan potensi uji coba teknologi perawatan jalan. Kemudian pada prioritas rendah umumnya terletak di area dengan lalu lintas rendah, kerusakan ringan, dan jauh dari fasilitas penting. Penanganannya bersifat jangka menengah hingga panjang, fokus pada pemantauan dan pemeliharaan ringan, tanpa intervensi besar dalam waktu dekat.

Sebagai studi kasus, sistem ini diuji menggunakan data yang telah tersedia yang memiliki tingkat keluhan masyarakat cukup tinggi terkait kondisi infrastruktur jalan. Dari total 100 segmen yang dianalisis, sebanyak 30 ruas jalan teridentifikasi masuk dalam kategori prioritas tinggi.

Tabel 8. Karakteristik Kondisi dan Faktor Kontekstual Segmen Jalan Hasil Analisis

Segmen Jalan	Karakteristik Kondisi	Faktor Kontekstual
A–B	Kerusakan berat, nilai SAW tertinggi	Volume lalu lintas >900 kendaraan/hari, berdekatan dengan sekolah dan rumah sakit
C–D	Kerusakan sedang, biaya perbaikan tinggi	Volume lalu lintas moderat, tidak berada di dekat fasilitas vital
E–F	Kerusakan ringan, nilai SAW rendah	Lalu lintas rendah, lokasi jauh dari fasilitas publik strategis

Hasil klasifikasi menunjukkan bahwa sistem mampu memetakan urgensi perbaikan jalan berdasarkan kondisi fisik, intensitas lalu lintas, dan faktor lingkungan. Segmen A–B ditempatkan pada prioritas tinggi karena kerusakan berat dan lokasinya berdekatan dengan fasilitas publik vital, sehingga memerlukan penanganan segera. Segmen C–D dikategorikan prioritas sedang karena meski tingkat kerusakannya cukup signifikan, volume lalu lintas

relatif lebih rendah sehingga dapat dijadwalkan dalam program tahunan reguler. Sementara itu, segmen E–F masuk ke prioritas rendah akibat kerusakan ringan serta dampaknya yang terbatas terhadap aktivitas publik, sehingga cukup ditangani melalui pemeliharaan rutin.

Penelitian ini membuktikan efektivitas integrasi *Geographic Information System* (GIS) dengan algoritma *machine learning* dalam mengklasifikasikan prioritas perbaikan jalan secara otomatis. Sebanyak 100 segmen dianalisis menggunakan enam kriteria spasial maupun non-spasial dengan metode *Simple Additive Weighting* (SAW). Skor prioritas kemudian diklasifikasikan ke dalam tiga kategori melalui pendekatan persentil dan divalidasi menggunakan *Random Forest Classifier* (RF). Validasi menunjukkan konsistensi 92% antara SAW dan RF, dengan akurasi keseluruhan 0,87 pada metrik *precision*, *recall*, dan *F1-score*. Analisis *feature importance* pada RF juga menambah wawasan dalam pengambilan keputusan, sementara kombinasi SAW–RF terbukti meningkatkan objektivitas, efisiensi, dan transparansi.

Meskipun demikian, penelitian ini menghadapi sejumlah tantangan. Keterbatasan data spasial menuntut validasi manual melalui survei lapangan, sedangkan variasi kualitas data antar instansi berpotensi menimbulkan bias, khususnya pada volume lalu lintas dan estimasi biaya perbaikan. Selain itu, data heterogen memerlukan proses pembersihan kompleks untuk mengatasi *missing values* dan *outliers*. Keterbatasan akses data real-time juga menurunkan adaptivitas sistem terhadap kondisi terkini. Dari sisi metode, SAW memiliki asumsi independensi antar kriteria dan bobot statis, padahal dalam praktik, interaksi antar variabel kerap terjadi. Hal ini menegaskan bahwa hasil perlu ditafsirkan dengan memperhatikan keterbatasan data, asumsi metodologis, dan faktor teknis implementasi.

4. Kesimpulan

Penelitian ini menunjukkan bahwa integrasi *Geographic Information System* (GIS) dengan algoritma *machine learning* dapat digunakan secara efektif untuk mengklasifikasikan prioritas perbaikan jalan. Sebanyak 100 segmen jalan dianalisis menggunakan enam kriteria, yaitu tingkat kerusakan, volume lalu lintas, biaya perbaikan, laporan masyarakat, kedekatan dengan fasilitas publik, dan kepadatan permukiman. Metode *Simple Additive Weighting* (SAW) digunakan untuk menghasilkan skor prioritas dan diklasifikasikan menjadi tiga kategori (tinggi, sedang, rendah), kemudian divalidasi dengan *Random Forest Classifier* (RF).

Hasil validasi menunjukkan konsistensi 92% antara SAW dan RF, dengan akurasi rata-rata 0,87 pada metrik *precision*, *recall*, dan *F1-score*. Model RF juga memberikan wawasan tambahan melalui analisis *feature importance*, sehingga meningkatkan

objektivitas dalam pengambilan keputusan. Sistem hibrida SAW–RF terbukti efisien, transparan, dan layak dikembangkan sebagai prototipe *decision support system* berbasis GIS.

Sebagai tindak lanjut, penelitian ini merekomendasikan penambahan atribut spasial lebih rinci serta penyusunan strategi kebijakan: segmen prioritas tinggi ditangani maksimal tiga bulan melalui anggaran darurat, prioritas sedang dalam program tahunan reguler, dan prioritas rendah difokuskan pada pemeliharaan jangka panjang.

Daftar Pustaka:

- Adam, S. I., Rotikan, R., Pasombaran, P. S., & Posumah, G. J. (2022). Mobile-Based Road Infrastructure Damage Reporting Service Application. 2022 4th International Conference on Cybernetics and Intelligent System (ICORIS), 1–7. <https://doi.org/10.1109/ICORIS56080.2022.10031418>
- Akindele, O., Ajayi, S., Oyegoke, A. S., Alaka, H. A., & Omotayo, T. (2023). Application of Geographic Information System (GIS) in construction: a systematic review. *Smart and Sustainable Built Environment*, 14(1), 210–236. <https://doi.org/10.1108/SASBE-01-2023-0016>
- Alkhaleel, B. A. (2024). Machine learning applications in the resilience of interdependent critical infrastructure systems—A systematic literature review. *International Journal of Critical Infrastructure Protection*, 44, 100646. <https://doi.org/https://doi.org/10.1016/j.ijcip.2023.100646>
- Anugraha, A. S., Erdiza, H. P., Apriyadi, D., & Aguscandra, B. (2022). Integration of Geospatial and Citizen Participation Using Geographic Information System for Smart City: a Study of Priority Villages Program in Jakarta, Indonesia. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 48(4/W5-2022), 17–24. <https://doi.org/10.5194/isprs-archives-XLVIII-4-W5-2022-17-2022>
- Choudhury, A., Mondal, A., & Sarkar, S. (2024). Searches for the BSM scenarios at the LHC using decision tree-based machine learning algorithms: a comparative study and review of random forest, AdaBoost, XGBoost and LightGBM frameworks. *European Physical Journal: Special Topics*, 233(15–16), 2425–2463. <https://doi.org/10.1140/epjs/s11734-024-01308-x>
- Fang, W., Chen, Y., Ding, J., Yu, Z., Masquelier, T., Chen, D., Huang, L., Zhou, H., Li, G., & Tian, Y. (2023). SpikingJelly: An open-source machine learning infrastructure platform for spike-based intelligence. *Science Advances*, 9(40), 1–116. <https://doi.org/10.1126/sciadv.adi1480>
- Firouraghi, N., Kiani, B., Jafari, H. T., Learnihan, V., Salinas-Perez, J. A., Raeesi, A., Furst, M. A., Salvador-Carulla, L., & Bagheri, N. (2022). The role of geographic information system and global positioning system in dementia care and research: a scoping review. *International Journal of Health Geographics*, 21(1), 1–13. <https://doi.org/10.1186/s12942-022-00308-1>
- Jigar Shah. (2023). Scalable Machine Learning Infrastructure on Cloud for Large-Scale Data Processing. *Tuijin Jishu/Journal of Propulsion Technology*, 42(2), 45–53. <https://doi.org/10.52783/tjjpt.v42.i2.7166>
- Khan, M. W., Obaidat, M. S., Mahmood, K., Batool, D., Badar, H. M. S., Aamir, M., & Gao, W. (2024). Real-Time Road Damage Detection and Infrastructure Evaluation Leveraging Unmanned Aerial Vehicles and Tiny Machine Learning. *IEEE Internet of Things Journal*, 11(12), 21347–21358. <https://doi.org/10.1109/JIOT.2024.3385994>
- Madhusudhan, H. S., Kumar T, S., Mustapha, S. M. F. D. S., Gupta, P., & Tripathi, R. P. (2021). Hybrid Approach for Resource Allocation in Cloud Infrastructure Using Random Forest and Genetic Algorithm. *Scientific Programming*, 2021. <https://doi.org/10.1155/2021/4924708>
- Nitayaprapha, T., Pochan, J., Panichakarn, B., & Kulthunyahirun, S. (2025). A route selection process using a multicriteria decision-making (MCDM) approach based on the simple additive weighting (SAW) method: Evidence from Thai fresh fruit exported to China by road transportation. *Decision Science Letters*, 14(3), 739–752. <https://doi.org/10.5267/j.dsl.2025.3.007>
- Putri, A., & Nazhifah, S. A. (2022). Pemanfaatan Google Earth untuk pemetaan Point of Interest dengan menggunakan Keyhole Markup Language (Studi Kasus di Darussalam dan Lampineung Banda Aceh). *Jurnal Teknologi Informasi*, 1(1), 16–21. <https://doi.org/10.35308/v1i1.5504>
- Raja, D. J. S., Sriranjani, R., Arulmozhi, P., & Hemavathi, N. (2024). Unified Random Forest and Hybrid Bat Optimization Based Man-in-the-Middle Attack Detection in Advanced Metering Infrastructure. *IEEE Transactions on Instrumentation and Measurement*, 73, 1–12. <https://doi.org/10.1109/TIM.2024.3420375>
- Singh, A. (2024). Architecting Ethical Machine Learning Infrastructure: a Systematic Framework for Technical Implementation. *International Journal of Computer Engineering and Technology*, 15(6), 1875–1886. https://doi.org/10.34218/ijcet_15_06_160
- Sui, X., Leng, Z., & Wang, S. (2023). Machine learning-based detection of transportation

infrastructure internal defects using ground-penetrating radar: a state-of-the-art review. *Intelligent Transportation Infrastructure*, 2, liad004. <https://doi.org/10.1093/iti/liad004>
Trianasari, N., & Permadi, T. A. (2024). Analysis of Product Recommendation Models at Each

Fixed Broadband Sales Location Using K-Means, DBSCAN, Hierarchical Clustering, SVM, RF, and ANN. *Journal of Applied Data Sciences*, 5(2), 636–652. <https://doi.org/10.47738/jads.v5i2.210>