



Implementasi Data Mining untuk Penentuan Prioritas Perbaikan Jalan Menggunakan Metode Klasifikasi Algoritma *K-Nearest Neighbors*

Aurita Nur Hamidah¹, Fajar Nugraha², Syafi'ul Muzid³

^{1,2,3} Program Studi Sistem Informasi, Fakultas Teknik, Universitas Muria Kudus

Article Info:

Dikirim: 7 Mei 2025

Direvisi: 18 Mei 2025

Diterima: 6 Juni 2025

Tersedia Online: 30 Juni 2025

Penulis Korespondensi:

Aurita Nur Hamidah

Program Studi Sistem Informasi,

Fakultas Teknik,

Universitas Muria Kudus

Email: 202153039@std.umk.ac.id

Abstrak: *Infrastruktur jalan yang baik merupakan komponen vital dalam mendukung mobilitas masyarakat dan mempercepat pertumbuhan ekonomi suatu wilayah. Di Indonesia, kerusakan jalan yang cukup tinggi, termasuk di Kabupaten Pati, menimbulkan tantangan tersendiri dalam penentuan prioritas perbaikan yang tepat dan efisien. Penelitian ini bertujuan untuk mengembangkan sistem klasifikasi tingkat kerusakan jalan menggunakan algoritma K-Nearest Neighbors (K-NN), guna memberikan rekomendasi prioritas perbaikan secara objektif. Data yang digunakan berasal dari Dinas Pekerjaan Umum dan Tata Ruang (DPUTR) Kabupaten Pati, mencakup atribut fisik seperti panjang, lebar jalan, serta persentase kerusakan ringan dan berat. Pendekatan data mining diterapkan untuk menemukan pola tersembunyi dari data historis yang dapat dimanfaatkan dalam pengambilan keputusan berbasis bukti. Proses pengembangan sistem mengikuti kerangka kerja CRISP-DM, dimulai dari pemahaman bisnis, eksplorasi data, praproses, pemodelan, hingga evaluasi. Model K-NN yang dibangun dievaluasi menggunakan metrik accuracy, precision, recall, dan F1-score. Hasil evaluasi menunjukkan performa sangat baik dengan akurasi mencapai 98%, serta nilai precision dan recall tinggi di seluruh kelas. Sistem yang dikembangkan berhasil mengklasifikasikan jalan ke dalam tiga tingkat prioritas: rendah, sedang, dan tinggi. Dengan hasil tersebut, sistem ini mampu membantu pemerintah daerah dalam menetapkan prioritas perbaikan jalan secara lebih cepat, efisien, dan berbasis data.*

Kata kunci: *K-Nearest Neighbors, Data Mining, Prioritas Perbaikan Jalan*

Abstract: *Good road infrastructure is a vital component in supporting community mobility and accelerating economic growth in a region. In Indonesia, the high level of road damage, including in Pati Regency, poses its own challenges in determining the right and efficient repair priorities. This study aims to develop a road damage classification system using the K-Nearest Neighbors (K-NN) algorithm, in order to provide objective repair priority recommendations. The data used comes from the Public Works and Spatial Planning Agency (DPUTR) of Pati Regency, including physical attributes such as length, road width, and percentage of light and heavy damage. A data mining approach is applied to find hidden patterns from historical data that can be utilized in evidence-based decision making. The system development process follows the CRISP-DM framework, starting from business understanding, data exploration, preprocessing, modeling, to evaluation. The K-NN model built is evaluated using accuracy, precision, recall, and F1-score metrics. The evaluation results show very good performance with an accuracy of 98%, as well as high precision and recall values across all classes. The developed system successfully classifies roads into three priority levels: low, medium, and high. With these results, this system is able to help local governments in setting road repair priorities more quickly, efficiently, and based on data.*

Keywords: *K-Nearest Neighbors, Data Mining, Road Repair Priority*

1. PENDAHULUAN

Infrastruktur jalan memegang peranan penting dalam mendukung mobilitas masyarakat serta memperlancar aktivitas ekonomi di berbagai wilayah [1]. Di Indonesia, kondisi jalan yang tidak selalu optimal menjadi salah satu kendala dalam mendukung distribusi barang dan jasa secara efisien. Kerusakan yang terjadi pada jaringan jalan cukup signifikan dan membutuhkan perhatian serius, khususnya dalam hal pengelolaan serta pemeliharaan yang tepat waktu dan tepat sasaran [2]. Hal ini menuntut adanya sistem yang dapat membantu mengoptimalkan proses pengambilan keputusan agar perbaikan dapat dilakukan secara efisien dan berdampak maksimal.

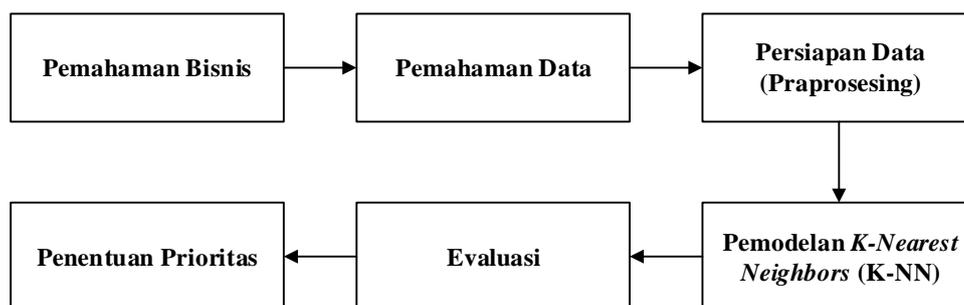
Di Kabupaten Pati, upaya pemerintah daerah dalam mengelola infrastruktur jalan mendapat tantangan tersendiri mengingat luasnya wilayah dan jumlah ruas jalan yang harus dikelola. Ketersediaan data mengenai kondisi jalan yang relatif lengkap menjadi modal penting dalam mendukung perencanaan perbaikan. Namun, keterbatasan anggaran serta sumber daya membuat pengelolaan tersebut tidak selalu berjalan sesuai harapan, sehingga diperlukan metode yang mampu memaksimalkan penggunaan data yang ada untuk menentukan prioritas perbaikan secara objektif dan efektif [3].

Perkembangan teknologi informasi, khususnya pada bidang data mining dan pembelajaran mesin, menawarkan peluang baru dalam meningkatkan kualitas pengelolaan infrastruktur jalan [4]. Data mining sendiri merupakan proses menemukan pola tersembunyi, informasi penting, atau pengetahuan yang bernilai dari kumpulan data dalam jumlah besar [5]. Tujuan dari data mining adalah mengidentifikasi pola-pola tersembunyi yang sebelumnya tidak dikenali, yang kemudian dapat dimanfaatkan untuk menyelesaikan berbagai permasalahan secara lebih efisien [6]. Dengan memanfaatkan algoritma yang tepat, seperti *K-Nearest Neighbors*, data historis dan kondisi fisik jalan dapat diolah untuk mengidentifikasi tingkat kerusakan secara otomatis dan mengklasifikasikannya ke dalam kategori prioritas. Pendekatan ini memungkinkan pengambilan keputusan yang lebih cepat dan berbasis bukti, sehingga diharapkan dapat mengatasi permasalahan keterbatasan sumber daya yang selama ini menjadi kendala utama [7].

Meski demikian, penerapan teknologi tersebut tidak terlepas dari berbagai tantangan, terutama terkait dengan kualitas dan kelengkapan data yang digunakan serta kemampuan sumber daya manusia dalam mengoperasikan sistem tersebut. Kesalahan input data atau kurangnya pemahaman teknis dapat berpengaruh pada akurasi hasil klasifikasi [8]. Oleh karena itu, penelitian ini diarahkan untuk mengkaji bagaimana algoritma *K-Nearest Neighbors* dapat diimplementasikan secara efektif untuk membantu pihak pengelola jalan dalam menentukan prioritas perbaikan yang lebih tepat dan efisien di Kabupaten Pati.

2. METODOLOGI PENELITIAN

Penelitian ini menggunakan pendekatan data mining dengan kerangka kerja CRISP-DM (*Cross Industry Standard Process for Data Mining*) sebagai pedoman utama dalam proses pengembangan model klasifikasi tingkat kerusakan jalan. Pendekatan ini dipilih karena bersifat sistematis dan fleksibel, sehingga dapat menyesuaikan dengan kebutuhan analisis data di lapangan [9]. Proses klasifikasi menggunakan algoritma *K-Nearest Neighbors* (K-NN), yang bertujuan untuk memberikan prediksi prioritas perbaikan jalan berdasarkan variabel kondisi fisik jalan. Alur pelaksanaan proses data mining tersebut mengikuti tahapan CRISP-DM, seperti yang ditunjukkan pada gambar 1.



Gambar 1. Alur Diagram CRISP-DM (*Cross Industry Standard Process for Data Mining*)

2.1. Pemahaman Bisnis

Tahap ini bertujuan untuk memahami permasalahan dan kebutuhan organisasi, yaitu Dinas Pekerjaan Umum dan Tata Ruang (DPUTR) Kabupaten Pati, dalam memantau dan menangani kerusakan jalan [10]. Permasalahan utama yang dihadapi adalah keterbatasan sumber daya dalam memantau kondisi jalan secara manual di seluruh wilayah.

Untuk mengatasi hal tersebut, dikembangkan sebuah sistem klasifikasi otomatis menggunakan algoritma K-NN yang dapat mengidentifikasi tingkat kerusakan jalan berdasarkan data yang tersedia, seperti volume kendaraan, jenis kerusakan, dan kondisi jalan. Hasil klasifikasi ini akan menjadi dasar dalam penentuan prioritas perbaikan jalan, sehingga proses pengambilan keputusan menjadi lebih cepat, efisien, dan tepat sasaran.

2.2. Pemahaman Data

Tahap ini difokuskan pada proses eksplorasi dan pemahaman terhadap data yang digunakan dalam penelitian [11]. Data dikumpulkan dari tiga sumber utama, yaitu:

1. Data sekunder diperoleh dari laporan resmi DPUTR Kabupaten Pati yang mencakup informasi kondisi fisik jalan, volume lalu lintas, jenis dan tingkat kerusakan.
2. Observasi lapangan, dilakukan untuk memverifikasi keakuratan data terhadap kondisi jalan sebenarnya di beberapa ruas jalan sebagai sampel.
3. Studi dokumentasi, berupa kajian literatur, standar teknis klasifikasi jalan, dan penelitian terdahulu yang berkaitan dengan penerapan algoritma K-NN dalam klasifikasi.

Tahapan ini penting untuk memastikan kualitas dan kelengkapan data, serta mengidentifikasi atribut-atribut yang paling relevan untuk proses klasifikasi.

2.3. Persiapan Data (Praprosesing)

Data preparation dilakukan untuk membersihkan dan mengubah data mentah menjadi format yang siap digunakan dalam pelatihan model [12]. Beberapa proses utama yang dilakukan adalah:

1. *Data Cleaning*, merupakan penghapusan data yang tidak konsisten, duplikat, atau memiliki nilai hilang. Teknik imputasi digunakan untuk menangani missing values [13].
2. *Data Transformation*, digunakan untuk memastikan format sesuai kebutuhan model K-NN. Langkah pentingnya adalah normalisasi, yaitu menyamakan rentang nilai setiap fitur [14]. Rumus normalisasi yang umum digunakan adalah:

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (1)$$

Keterangan:

- a. x' : Nilai yang telah dinormalisasi
 - b. x : Nilai asli
 - c. x_{min} : Nilai minimum dalam fitur
 - d. x_{max} : Nilai maksimum dalam fitur
3. *Feature Selection* dalam pemilihan fitur dilakukan untuk menentukan atribut yang paling berpengaruh dalam klasifikasi tingkat kerusakan [15]. Teknik seperti *Mutual Information* dan *Recursive Feature Elimination* (RFE) digunakan dalam proses ini.

2.4. Pemodelan K-Nearest Neighbors (K-NN)

Pada tahap ini dilakukan pembangunan model klasifikasi menggunakan algoritma *K-Nearest Neighbors* (K-NN). K-NN adalah metode *supervised learning* yang menentukan kelas dari data baru berdasarkan jarak terdekat dengan data-data dalam dataset [16]. Jarak antara dua titik dihitung menggunakan *Euclidean Distance*, dengan rumus:

$$d(i, j) = \sqrt{\sum_{k=1}^n (x_{ik} - x_{jk})^2} \quad (2)$$

Setelah menghitung jarak, algoritma akan memilih sejumlah k tetangga terdekat, kemudian menentukan kelas berdasarkan mayoritas kelas tetangga tersebut. Pemilihan nilai k yang optimal dilakukan melalui validasi model untuk menghindari *overfitting* maupun *underfitting*.

2.5. Evaluasi

Confusion Matrix adalah alat yang digunakan untuk mengevaluasi *accuracy*, *precision*, *recall* dan *F1-Score* dalam model, memastikan bahwa model bekerja sebagaimana yang diharapkan [17]. Presisi menunjukkan seberapa tepat hasil yang diperoleh terhadap kebutuhan data, sementara recall menunjukkan keberhasilan sistem dalam menemukan data yang benar. Hubungan antara nilai-nilai ini dapat dianalisis melalui *confusion matrix*, seperti yang ditunjukkan pada tabel 1.

Tabel 1. Confusion Matrix

	True	False
True (Positif)	TP (True Positif) <i>Correct result</i>	FP (False Positif) <i>Unexpected result</i>
False (Negative)	FN (False Negative) <i>Missing result</i>	TN (True Negative) <i>Correct absence of result</i>

Berikut adalah rumus dalam menghitung *precision*, *recall*, *accuracy*, dan *F1-Score*

$$a. \textit{Precision} : \frac{TP}{TP + FP} \quad (3)$$

$$b. \textit{Recall} : \frac{TP}{TP + FN} \quad (4)$$

$$c. \textit{Accuracy} : \frac{TP + TN}{TP + TN + FP + FN} \quad (5)$$

$$d. \textit{F1-Score} : \frac{2 \times \textit{Precision} \times \textit{Recall}}{\textit{Recall} + \textit{Precision}} \quad (6)$$

Di sini, **TP** adalah jumlah data dengan nilai positif yang benar, **FP** adalah data negatif dengan nilai prediksi positif, **FN** adalah data positif dengan prediksi negatif, dan **TN** adalah data negatif yang diprediksi negatif. Model dikatakan baik jika memiliki nilai *precision*, *recall*, dan *accuracy* yang tinggi, mengindikasikan kemampuan model dalam memprediksi dengan akurat.

2.6. Penentuan Prioritas

Tahapan akhir dari proses CRISP-DM yaitu implementasi model klasifikasi dalam sistem yang dapat digunakan oleh DPUTR. Hasil klasifikasi digunakan untuk menentukan prioritas perbaikan jalan, yang dibagi menjadi tiga kategori:

1. Prioritas Tinggi: kerusakan parah, volume kendaraan tinggi, dan terletak di jalur utama atau kawasan padat penduduk.
2. Prioritas Sedang: kerusakan sedang, lalu lintas moderat, berada di wilayah penunjang ekonomi.
3. Prioritas Rendah: kerusakan ringan, lalu lintas rendah, berada di daerah dengan aktivitas lalu lintas minim.

Hasil klasifikasi dan penentuan prioritas ini kemudian disajikan dalam bentuk laporan visualisasi yang dapat digunakan sebagai dasar pengambilan keputusan dalam perencanaan, penganggaran, dan pelaksanaan perbaikan jalan secara lebih strategis dan efisien.

3. HASIL DAN PEMBAHASAN

Berdasarkan permasalahan yang telah diuraikan sebelumnya, dibutuhkan sebuah sistem klasifikasi yang mampu memprediksi tingkat kerusakan jalan di Kabupaten Pati guna membantu Dinas Pekerjaan Umum dan Tata Ruang (DPUTR) dalam menentukan prioritas perbaikan secara lebih cepat, efisien, dan objektif. Proses klasifikasi ini dilakukan menggunakan algoritma *K-Nearest Neighbors* (K-NN) yang memanfaatkan pola dari data historis seperti jenis kerusakan, dan kondisi fisik jalan. Dengan menerapkan pendekatan data mining, sistem ini dirancang untuk mengekstraksi informasi yang relevan dari kumpulan data besar guna menghasilkan prediksi yang akurat. Hasil dari penelitian ini mencakup tahapan preprocessing data, pembangunan model klasifikasi, evaluasi performa model, serta implementasi sistem dalam bentuk rekomendasi prioritas perbaikan jalan yang dapat digunakan sebagai dasar pengambilan keputusan oleh pihak terkait.

3.1. Pengujian Model

a. Dataset Kerusakan Jalan

Dataset yang digunakan dalam penelitian ini merupakan data kondisi jalan di Kabupaten Pati yang diperoleh dari Dinas Pekerjaan Umum dan Tata Ruang (DPUTR). Dataset ini mencakup berbagai atribut penting terkait kondisi fisik dan klasifikasi kerusakan jalan, antara lain:

1. Nama Ruas Jalan: Identitas ruas jalan yang diamati.
2. Kecamatan Yang Dilalui: Wilayah administratif yang dilewati oleh ruas jalan tersebut.
3. Panjang dan Lebar Ruas: Dimensi fisik jalan yang mempengaruhi daya tahan terhadap beban lalu lintas.

4. Kondisi Jalan: Persentase kerusakan dari rusak ringan dan rusak berat.
 5. Total Kerusakan (%): Penjumlahan dari persentase kerusakan ringan dan berat.
- Data contoh dataset kerusakan jalan penelitian Untuk mendukung proses klasifikasi, digunakan data kondisi fisik jalan yang telah dikumpulkan sebagai bahan analisis. Data contoh dataset kerusakan jalan yang digunakan dalam penelitian ini dapat dilihat pada tabel 2.

Tabel 2. Contoh Dataset Kerusakan Jalan

No	Nama Ruas Jalan	Kecamatan	Panjang Ruas	Lebar Ruas	Rusak Ringan (%)	Rusak Berat (%)	Total Kerusakan (%)
1	JL. PEMUDA	PATI	1.24	11.0	0,00%	0,00%	0.0
2	JL. P. SUDIRMAN	PATI	1.45	11.0	0,00%	0,00%	0.0
3	JL. Dr. SUSANTO	PATI	1.02	11.0	0,00%	0,00%	0.0
4	JL. Dr. WAHIDIN	PATI	0.42	11.0	0,00%	0,00%	0.0
5	JL. ALUN-ALUN PATI	PATI	0.43	14.0	0,00%	0,00%	0.0

b. Hasil Normalisasi Data

Proses normalisasi dilakukan untuk menyetarakan skala nilai dari setiap fitur numerik, agar perhitungan jarak dalam algoritma *K-Nearest Neighbors* (K-NN) menjadi lebih akurat dan tidak berat sebelah terhadap fitur dengan skala besar. Normalisasi yang digunakan adalah metode *min-max normalization*, dengan rumus sebagai berikut:

Contoh Normalisasi (5 Data Pertama)

Fitur yang dinormalisasi:

1. Panjang Ruas
2. Lebar Ruas
3. Total Kerusakan (%)

Dari contoh data, nilai ekstrim untuk fitur:

1. Panjang Ruas: min = 0.42, max = 1.45
2. Lebar Ruas: min = 11.0, max = 14.0
3. Total Kerusakan (%): min = 0.0, max = 0.0 (karena semua 0.0 → tidak bisa dinormalisasi, maka bisa dihapus/dianggap tidak variatif dalam kasus ini)

Mari kita normalisasi Panjang Ruas dan Lebar Ruas dari data pertama:

Data ke-1

1. Panjang Ruas = 1.24
2. Lebar Ruas = 11.0

Normalisasi Panjang Ruas:

$$x' = \frac{1.45-0.42}{1.24-0.42} = \frac{1.03}{0.82} = 0.796$$

Normalisasi Lebar Ruas:

$$x' = \frac{11.0-11.0}{14.0-11.0} = \frac{0}{3} = 0.000$$

Data ke-2 (Panjang Ruas = 1.45, Lebar Ruas = 11.0):

Panjang Ruas:

$$x' = \frac{1.45-0.42}{1.45-0.42} = \frac{1.03}{1.03} = 1.000$$

Lebar Ruas:

$$x' = \frac{14.0-11.0}{11.0-11.0} = 0.000$$

Tabel berikut menunjukkan hasil normalisasi fitur pada lima data pertama, seperti yang ditampilkan pada tabel 3.

Tabel 3. Perhitungan Normalisasi Data

No	Panjang Ruas (x')	Lebar Ruas (x')
1	0.796	0.000
2	1.000	0.000
3	0.582	0.000
4	0.000	0.000
5	0.010	1.000

c. Perhitungan Jarak dan Penentuan Tetangga Terdekat

Proses klasifikasi menggunakan algoritma *K-Nearest Neighbors (K-NN)* dilakukan dengan menghitung jarak antar data untuk menemukan tetangga terdekat.

Contoh Perhitungan Jarak:

Berikut adalah contoh perhitungan jarak antara data ke-1 dan data ke-2 menggunakan dua fitur yang telah dinormalisasi, yaitu Panjang Ruas (x_1) dan Lebar Ruas (x_2), seperti ditunjukkan pada tabel 4.

Tabel 4. Perhitungan Jarak dan Penentuan Tetangga Terdekat

Fitur	Data ke-1 (x1)	Data ke-2 (x2)	$(x1 - x2)^2$
Panjang Ruas	0.796	1.000	$(0.796 - 1.000)^2 = 0.0416$
Lebar Ruas	0.000	0.000	$(0.000 - 0.000)^2 = 0.0000$
Total			0.0416

$$d(1,2) = \sqrt{0.0416} = 0.204$$

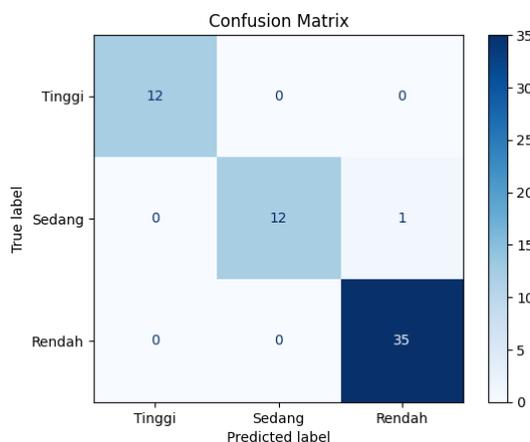
Perhitungan dilakukan untuk setiap data terhadap seluruh data lainnya dalam dataset. Setelah seluruh jarak dihitung, model K-NN memilih k data dengan jarak terpendek sebagai tetangga terdekat.

d. Evaluasi Model

Tahap evaluasi merupakan bagian penting dalam proses pengembangan model klasifikasi, karena memberikan gambaran mengenai seberapa baik model dalam memprediksi kelas target secara akurat. Dalam penelitian ini, evaluasi dilakukan terhadap model *K-Nearest Neighbors (K-NN)* yang digunakan untuk mengklasifikasikan tingkat kerusakan jalan berdasarkan fitur-fitur seperti panjang ruas, lebar ruas, dan total kerusakan.

Model diuji menggunakan 30% dari total data yang telah dipisahkan sebagai data uji (sebanyak 60 data). Hasil prediksi dari model dibandingkan dengan label aktual pada data uji untuk menghasilkan metrik evaluasi, yaitu:

1. *Confusion Matrix*



Gambar 2. Hasil Confusion Matrix

Berdasarkan *confusion matrix* yang ditunjukkan pada gambar 2, dapat dilihat bahwa semua data dengan label aktual Tinggi dan Rendah berhasil diklasifikasikan dengan benar oleh model, sementara hanya terdapat satu data dengan label aktual Sedang yang diklasifikasikan sebagai Tinggi. Hal ini menunjukkan bahwa

meskipun terdapat sedikit kesalahan dalam mengidentifikasi kelas Sedang, secara keseluruhan model mampu membedakan ketiga kelas dengan sangat baik dan tingkat kesalahan klasifikasinya tergolong sangat rendah.

2. Classification Report

```

== Classification Report ==
              precision    recall  f1-score   support

   Rendah      0.97      1.00      0.99         35
   Sedang      1.00      0.92      0.96         13
   Tinggi      1.00      1.00      1.00         12

 accuracy              0.98         60
 macro avg              0.99      0.97      0.98         60
 weighted avg          0.98      0.98      0.98         60

```

Untuk mendapatkan pemahaman yang lebih mendalam tentang performa model, digunakan metrik evaluasi lanjutan seperti *precision*, *recall*, dan *f1-score*. Berdasarkan hasil *classification report*, terlihat bahwa model memiliki kinerja yang sangat baik dalam mengklasifikasikan data ke dalam tiga kelas, Rendah, Sedang, dan Tinggi. Kelas Tinggi menunjukkan hasil sempurna dengan nilai *precision*, *recall*, dan *f1-score* sebesar 1.00, yang berarti seluruh data dalam kelas ini berhasil diidentifikasi tanpa kesalahan. Kelas Rendah juga mencatat performa yang sangat baik dengan *recall* 1.00 dan *precision* 0.97, menunjukkan bahwa hampir seluruh prediksi untuk kelas ini tepat sasaran. Sementara itu, kelas Sedang memiliki *precision* 1.00 namun *recall* sedikit lebih rendah, yaitu 0.92, yang mengindikasikan bahwa ada sebagian kecil data Sedang yang tidak berhasil dikenali dengan benar. Secara keseluruhan, model mencapai tingkat akurasi sebesar 98%, dan nilai rata-rata makro maupun tertimbang untuk semua metrik evaluasi berada di atas 0.97, mencerminkan stabilitas dan keandalan model dalam membedakan ketiga kelas dengan sangat baik.

e. Penentuan Prioritas Perbaikan Jalan

Penentuan prioritas perbaikan jalan merupakan langkah akhir yang penting, karena berperan dalam membantu instansi terkait, seperti DPUTR Kabupaten Pati, dalam merancang strategi perencanaan perbaikan infrastruktur jalan secara lebih efisien dan objektif. Prioritas ditentukan berdasarkan hasil klasifikasi tingkat kerusakan, yaitu:

1. Prioritas 1 (Tinggi): Jalan dalam kondisi rusak parah.
2. Prioritas 2 (Sedang): Jalan dengan kerusakan sedang.
3. Prioritas 3 (Rendah): Jalan yang mengalami kerusakan ringan.

Berikut merupakan daftar 10 jalan dengan prioritas tertinggi berdasarkan hasil klasifikasi menggunakan metode K-NN, seperti ditampilkan pada tabel 5.

Tabel 5. Daftar 10 Jalan Dengan Prioritas Tertinggi

No.	Nama Ruas Jalan	Total Kerusakan (%)	Prediksi Kelas	Prioritas
1	TPA	100.00	Tinggi	1
2	BUMIREJO	88.33	Tinggi	1
3	MUKTIHARJO	60.00	Tinggi	1
4	BADEGAN	66.23	Tinggi	1
5	H. MUNADI	77.61	Tinggi	1
6	PURI	73.33	Tinggi	1
7	JL. HANG TUAH	71.83	Tinggi	1
8	JL. KRAPYAK	64.29	Tinggi	1
9	TAYU	51.81	Tinggi	1
10	TLOGOWUNGU	94.27	Tinggi	1

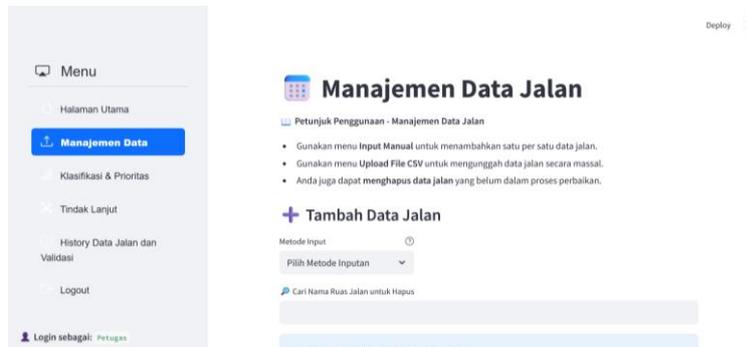
Ruas-ruas jalan tersebut tergolong dalam kelas kerusakan tinggi, sehingga sebaiknya menjadi fokus utama dalam perencanaan program perbaikan.

3.2. Pengembangan Sistem

Untuk mendukung proses pengumpulan, analisis, dan pengambilan keputusan terkait prioritas perbaikan jalan, sistem ini dikembangkan dengan berbagai halaman utama yang saling terintegrasi. Setiap halaman dirancang untuk memberikan kemudahan bagi pengguna sesuai dengan peran dan fungsinya. Berikut penjelasan dari setiap halaman yang terdapat dalam sistem:

a. Manajemen Data Jalan

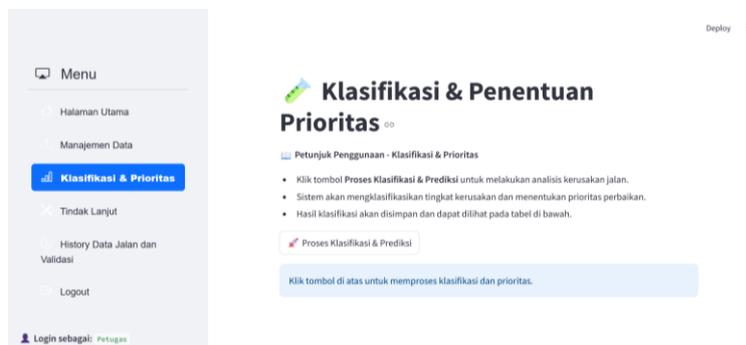
Halaman Manajemen Data Jalan berfungsi sebagai tempat input dan pengelolaan awal data jalan. Pada halaman ini, pengguna dapat melakukan input data jalan secara manual satu per satu atau melalui unggahan file CSV untuk mempermudah input data dalam jumlah besar. Selain itu, pengguna juga dapat melihat daftar data jalan yang telah dimasukkan dan menghapus data jalan yang belum masuk ke dalam proses tindak lanjut perbaikan. Halaman ini menjadi langkah awal yang krusial sebelum data dianalisis lebih lanjut melalui proses klasifikasi dan penentuan prioritas. Halaman berikut menunjukkan tampilan antarmuka untuk manajemen data jalan, seperti yang ditampilkan pada gambar 3.



Gambar 3. Halaman Manajemen Data Jalan

b. Klasifikasi dan Penentuan Prioritas

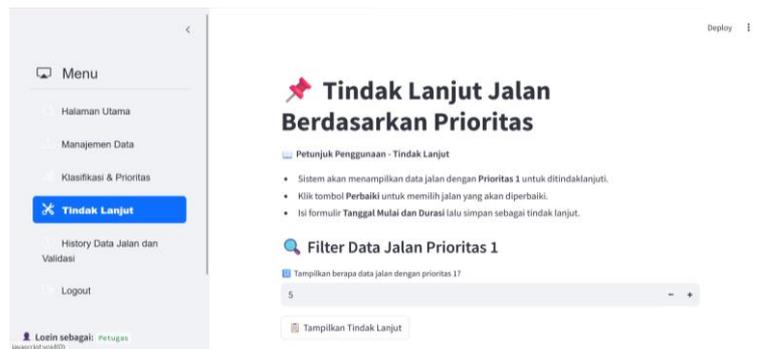
Halaman Klasifikasi dan Penentuan Prioritas digunakan untuk melakukan proses analisis data jalan berdasarkan tingkat kerusakan menggunakan algoritma klasifikasi, seperti *K-Nearest Neighbor* (KNN). Sistem akan mengelompokkan kondisi jalan ke dalam kategori seperti rusak parah, rusak sedang, atau rusak ringan. Berdasarkan hasil klasifikasi tersebut, sistem akan secara otomatis menentukan prioritas perbaikan, yaitu Prioritas 1, 2, atau 3. Hasil klasifikasi ini akan ditampilkan secara visual dan dapat diunduh dalam format PDF. Semua data hasil klasifikasi juga akan tersimpan ke dalam basis data untuk proses tindak lanjut berikutnya. Halaman berikut menunjukkan tampilan antarmuka untuk proses klasifikasi dan penentuan prioritas, seperti yang ditampilkan pada gambar 4.



Gambar 4. Halaman Klasifikasi dan Penentuan Prioritas

c. Tindak Lanjut

Halaman Tindak Lanjut disediakan khusus untuk petugas lapangan yang bertugas melaksanakan proses perbaikan jalan. Di halaman ini, ditampilkan daftar jalan yang termasuk dalam Prioritas 1 yang harus segera diperbaiki. Petugas dapat memilih jalan yang akan ditindaklanjuti, kemudian mengisi detail pelaksanaan seperti tanggal mulai dan estimasi durasi perbaikan. Informasi tersebut akan disimpan ke dalam sistem sebagai bukti progres pekerjaan di lapangan. Halaman berikut menunjukkan tampilan antarmuka untuk proses tindak lanjut, seperti yang ditampilkan pada gambar 5.



Gambar 5. Halaman Tindak Lanjut

d. **Histori Data Jalan dan Validasi**

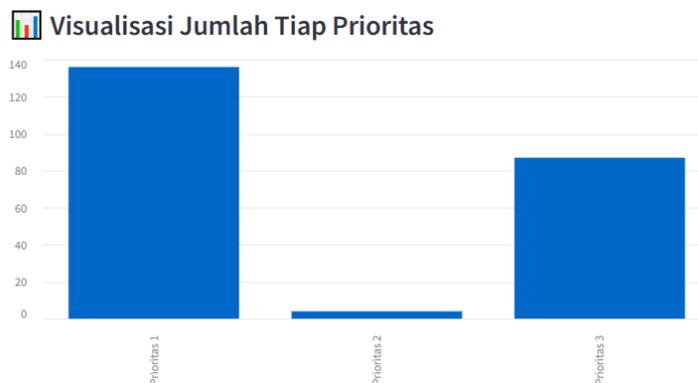
Halaman Histori Data Jalan dan Validasi menampilkan riwayat jalan yang telah diperbaiki serta digunakan untuk proses validasi oleh petugas. Setelah proses perbaikan selesai, jalan akan divalidasi ulang untuk memastikan bahwa kondisinya telah membaik dan layak dinyatakan sebagai “Jalan Baik”. Jika ditemukan kesalahan, petugas juga memiliki opsi untuk membatalkan validasi. Semua aktivitas validasi akan tercatat dalam sistem sebagai bagian dari histori untuk keperluan monitoring dan evaluasi. Halaman berikut menunjukkan tampilan untuk melihat histori data jalan dan proses validasi, seperti yang ditampilkan pada gambar 6.



Gambar 6. Halaman Histori Data Jalan dan Validasi

e. **Visualisasi Tingkat Prioritas**

Setelah model *K-Nearest Neighbors* (K-NN) berhasil dilatih dan diuji, langkah selanjutnya adalah melakukan klasifikasi terhadap seluruh data jalan dalam dataset. Setiap ruas jalan dipetakan ke dalam salah satu dari tiga kelas tingkat kerusakan, yaitu Rendah, Sedang, dan Tinggi, berdasarkan prediksi model terhadap fitur yang diberikan. Grafik berikut menunjukkan tingkat jumlah data berdasarkan kategori prioritas, seperti yang ditampilkan pada gambar 7.



Gambar 7. Grafik Jumlah Per Prioritas

4. **KESIMPULAN**

Berdasarkan hasil penelitian, dapat disimpulkan bahwa pengembangan sistem klasifikasi tingkat kerusakan dan penentuan prioritas perbaikan jalan menggunakan algoritma *K-Nearest Neighbors* (K-NN) mampu memberikan solusi yang efektif dan akurat bagi Dinas Pekerjaan Umum dan Tata Ruang (DPUTR) Kabupaten Pati dalam mengelola

infrastruktur jalan. Sistem ini dirancang berdasarkan tahapan CRISP-DM secara sistematis, mulai dari pemahaman permasalahan, eksplorasi data, *preprocessing*, pemodelan, evaluasi, hingga implementasi dalam bentuk aplikasi berbasis web. Hasil evaluasi menunjukkan bahwa model memiliki performa yang sangat baik dengan akurasi sebesar 98%, serta *precision*, *recall*, dan *f1-score* tinggi pada setiap kelas kerusakan. Model ini berhasil memetakan ruas jalan ke dalam tiga tingkat prioritas (tinggi, sedang, rendah) secara objektif, mendukung proses pengambilan keputusan yang lebih cepat, efisien, dan tepat sasaran. Selain itu, sistem ini juga berkontribusi pada peningkatan transparansi dan akuntabilitas, serta menjadi langkah awal menuju digitalisasi pengelolaan jalan di lingkungan pemerintah daerah.

UCAPAN TERIMA KASIH

Peneliti mengucapkan terima kasih yang sebesar-besarnya kepada semua pihak yang telah memberikan dukungan selama proses penelitian ini. Peneliti juga menyampaikan apresiasi dan penghargaan kepada Dinas Pekerjaan Umum dan Tata Ruang (DPUTR) Kabupaten Pati yang telah memberikan akses data dan informasi penting terkait kondisi infrastruktur jalan, sehingga memungkinkan dilakukannya analisis yang mendalam. Semoga hasil dari penelitian ini dapat memberikan kontribusi nyata dalam mendukung pengambilan keputusan perbaikan jalan yang lebih efisien, akurat, dan berbasis data di Kabupaten Pati, serta bermanfaat bagi pengembangan sistem pengelolaan infrastruktur jalan di masa mendatang.

DAFTAR PUSTAKA

- [1] A. R. Usman, S. E. Saleh, and S. Indriyani, "Dampak Perbaikan Infrastruktur Jalan Terhadap Pertumbuhan UMKM di Kecamatan Kota Selatan Kota Gorontalo," *Econ. Rev. J.*, vol. 4, no. 1, pp. 300–312, 2025.
- [2] R. S. Nuhun, L. Welendo, M. F. Almaliki, I. Ismayana, H. Herianto, and H. Harmianto, "Evaluasi Panjang dan Kondisi Jalan di Kabupaten/Kota Provinsi Sulawesi Tenggara: Implikasi Terhadap Pengelolaan Infrastruktur," *Innov. J. Soc. Sci. Res.*, vol. 4, no. 5, pp. 8824–8835, 2024.
- [3] A. G. Silalahi and others, "Sistem Informasi Geografis Prioritas Penanganan Kerusakan Infrastruktur Menggunakan Metode MOORA," *METIK J. (AKREDITASI SINTA 3)*, vol. 9, no. 1, pp. 170–180, 2025.
- [4] M. Indarto and B. Santoso, "Efektivitas Pemanfaatan Big Data dalam Pengambilan Keputusan Strategis di Industri Perbankan," *J. Bisnis dan Manaj.*, vol. 2, no. 4, pp. 1159–1182, 2024.
- [5] Z. F. Ahmed Arifi Hilman Rahman, "Implementasi Data Mining Menggunakan Algoritma Apriori," *J. SITECH Sist. Inf. dan Teknol.*, vol. 7, no. 2, 2024.
- [6] Z. F. Ahsin Ilallah, "Penerapan Data Mining Menggunakan Metode Naïve Bayes untuk Memprediksi Tingkat Perceraian Pasangan Muda di Kota Banyuwangi," *J. SITECH Sist. Inf. dan Teknol.*, vol. 7, 2024.
- [7] U. Labuhanbatu and M. Masrizal, *Data Mining Analisis Kepuasan Peserta Didik Berbasis Machine Learning*, no. June. 2025.
- [8] A. Ernawati, Z. Sitorus, M. Iqbal, D. Nasution, and others, "Penerapan Data Mining Untuk Klasifikasi Penduduk Miskin Di Kabupaten Labuhanbatu Menggunakan Random Forest Dan K-Nearest Neighbors," *Bull. Inf. Technol.*, vol. 6, no. 2, pp. 23–35, 2025.
- [9] P. T. I. U. P. Y. Nurirwan Saputra, "Pengenalan Data Mining," 2023.
- [10] R. Rianti and R. Andarsyah, "Memprediksi Tingkat Atrisi Karyawan Menggunakan Machine Learning," *J. Tekno Insentif*, vol. 18, no. 1, pp. 39–52, 2024.
- [11] S. Sutisna and M. N. Yuniar, "Klasifikasi kualitas air bersih menggunakan METODE Naive baiyes," *J. Sains dan Teknol.*, vol. 5, no. 1, pp. 243–246, 2023.
- [12] N. Wahyuningsih and H. Hendry, "Perbandingan metode klasifikasi dalam analisis sentimen masyarakat terhadap identitas kependudukan digital (ikd)," *JUPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.*, vol. 8, no. 4, pp. 1218–1227, 2023.
- [13] S. Pramuwidya, R. Y. Fa'rifah, and O. N. Pratiwi, "Klasifikasi Review Customer E-Commerce Menggunakan Metode K-Nearest Neighbor (Studi Kasus: Bukalapak)," *eProceedings Eng.*, vol. 10, no. 3, 2023.
- [14] M. A. Zidane *et al.*, "PENILAIAN KOMPARATIF METODE KLASIFIKASI NEURAL NETWORK DAN RANDOM FOREST UNTUK KNOWLEDGE DISCOVERY PADA PENYAKIT DIABETES," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 9, no. 3, pp. 4999–5006, 2025.
- [15] T. D. Arista, Y. Yusra, M. Fikry, and L. Oktavia, "Klasifikasi Sentimen Masyarakat di Twitter terhadap Kenaikan Harga BBM dengan Metode K-NN," *JUKI J. Komput. dan Inform.*, vol. 5, no. 1, pp. 140–150, 2023.
- [16] P. Putra, A. M. H. Pardede, and S. Syahputra, "Analisis Metode K-Nearest Neighbour (Knn) Dalam Klasifikasi Data Iris Bunga," *JTIK (Jurnal Tek. Inform. Kaputama)*, vol. 6, no. 1, pp. 297–305, 2022.
- [17] S. Lonang, A. Yudhana, and M. K. Biddinika, "Analisis Komparatif Kinerja Algoritma Machine Learning untuk Deteksi Stunting," *J. Media Inform. Budidarma*, vol. 7, no. 4, pp. 2109–2117, 2023.