

Sistem Pendukung Keputusan Prioritas Perbaikan Drainase Kota Manado Berbasis Web Menggunakan Algoritma C4.5 Decision Tree

Richard Gilardino Gabriel¹, Nelson Nainggolan², Dodisutarma Lapihu³

^{1,2,3}Sistem Informasi, Universitas Sam Ratulangi, Manado, Indonesia

Email: ¹richardgabriel106@student.unsrat.ac.id, ²n-nelson@unsrat.ac.id, ³dlapihu@unsrat.ac.id

Email Penulis Korespondensi: richardgabriel106@student.unsrat.ac.id

Abstrak– Sistem drainase merupakan infrastruktur penting dalam mengendalikan limpasan air permukaan untuk mengurangi risiko genangan dan banjir di kawasan perkotaan. Di Kota Manado, penentuan prioritas perbaikan drainase perlu dilakukan secara sistematis karena terdapat berbagai parameter penilaian dan keterbatasan sumber daya dalam penanganannya. Penelitian ini bertujuan mengembangkan Sistem Pendukung Keputusan (SPK) berbasis web menggunakan algoritma C4.5 Decision Tree untuk mengklasifikasikan prioritas perbaikan drainase ke dalam tiga kategori, yaitu rendah, sedang, dan tinggi. Data awal penelitian diperoleh melalui diskusi dan studi dokumentasi dengan Dinas Pekerjaan Umum dan Penataan Ruang (PUPR) Kota Manado yang mencakup 10 blok layanan drainase. Karena keterbatasan data historis, sebanyak 805 data dummy disusun berdasarkan sembilan parameter penilaian dan mekanisme pembobotan yang digunakan oleh PUPR Kota Manado. Dataset terdiri atas tinggi genangan, luas genangan, lama genangan, frekuensi genangan, kerugian ekonomi, gangguan sosial, gangguan transportasi, gangguan perumahan, dan kerugian harta benda. Data diproses melalui tahap preprocessing dan dibagi menjadi data pelatihan dan pengujian dengan rasio 80:20, masing-masing sebanyak 644 dan 161 data. Model klasifikasi dibangun menggunakan algoritma C4.5 dan dievaluasi menggunakan confusion matrix dengan metrik accuracy, precision, recall, dan F1-score. Hasil penelitian menunjukkan bahwa atribut gangguan transportasi memiliki nilai Gain Ratio tertinggi sebesar 0,179 sehingga dipilih sebagai root node pada pohon keputusan. Evaluasi model menghasilkan accuracy sebesar 76,40%, dengan 123 dari 161 data uji berhasil diklasifikasikan dengan benar. Sistem yang dikembangkan mampu menghasilkan rekomendasi prioritas perbaikan berdasarkan pola klasifikasi yang dipelajari dari dataset, sehingga dapat digunakan sebagai alat bantu dalam mendukung proses pengambilan keputusan perbaikan drainase secara lebih sistematis dan berbasis data.

Kata Kunci: Sistem Pendukung Keputusan, C4.5 Decision Tree, Drainase, Klasifikasi

Abstract– Drainage systems are important infrastructure for managing surface water runoff and reducing the risk of inundation and flooding in urban areas. In Manado City, determining drainage repair priorities requires a systematic approach because multiple assessment parameters must be considered while available resources for infrastructure intervention are limited. This study aims to develop a web-based Decision Support System (DSS) using the C4.5 Decision Tree algorithm to classify drainage repair priorities into three categories: low, medium, and high. Initial data were obtained through discussions and document studies with the Department of Public Works and Spatial Planning (PUPR) of Manado City, covering 10 drainage service blocks. Due to the limited availability of historical data, 805 dummy data instances were generated based on nine assessment parameters and the weighting mechanism used by PUPR Manado City. The dataset consisted of inundation depth, inundation area, inundation duration, inundation frequency, economic loss, social disruption, transportation disruption, housing disruption, and property damage. The data were preprocessed and divided into training and testing sets at an 80:20 ratio, consisting of 644 and 161 instances, respectively. The classification model was developed using the C4.5 algorithm and evaluated using a confusion matrix with accuracy, precision, recall, and F1-score metrics. The results showed that transportation disruption had the highest Gain Ratio of 0.179 and was therefore selected as the root node of the decision tree. Model evaluation achieved an accuracy of 76.40%, with 123 out of 161 testing instances correctly classified. The developed system is capable of generating drainage repair priority recommendations based on classification patterns learned from the dataset and can serve as a data-driven decision support tool for supporting drainage repair prioritization.

Keywords: Decision Support System, C4.5 Decision Tree, Classification

1. PENDAHULUAN

Sistem drainase merupakan salah satu infrastruktur penting di kawasan perkotaan yang berfungsi mengalirkan limpasan air permukaan untuk mengurangi risiko genangan dan banjir. Kondisi sistem drainase yang tidak mampu menampung atau mengalirkan limpasan secara optimal dapat mengganggu aktivitas masyarakat, memengaruhi fungsi infrastruktur, serta menimbulkan kerugian sosial dan ekonomi. Permasalahan tersebut dapat diperburuk oleh perubahan tata guna lahan, curah hujan ekstrem, sedimentasi saluran, dan penumpukan sampah yang berpotensi mengurangi kapasitas sistem drainase dalam

mengalirkan limpasan air. Kondisi tersebut dapat meningkatkan risiko terjadinya genangan dan banjir yang berdampak pada infrastruktur, aktivitas transportasi, kondisi sosial masyarakat, serta kerugian ekonomi. Oleh karena itu, pengelolaan dan perbaikan sistem drainase perlu dilakukan berdasarkan tingkat permasalahan dan urgensi penanganannya [1], [2].

Kota Manado merupakan salah satu wilayah yang menghadapi permasalahan genangan dan banjir. Evaria dan Istiqomah melaporkan bahwa kejadian banjir pada 27–28 Januari 2023 menyebabkan sekitar 400 unit rumah terendam. Pada kejadian tersebut, curah hujan terakumulasi sebesar 237,8 mm pada 27 Januari 2023 dan 216,7 mm pada 28 Januari 2023. Nilai tersebut termasuk dalam kategori hujan ekstrem dengan intensitas lebih dari 150 mm/hari berdasarkan kriteria BMKG [3]. Kondisi tersebut menunjukkan bahwa kejadian banjir di Kota Manado dapat menimbulkan dampak yang signifikan terhadap kawasan permukiman sehingga penanganan sistem drainase perlu dilakukan secara terarah dan sistematis.

Penentuan prioritas perbaikan drainase menjadi kompleks karena setiap blok layanan memiliki karakteristik genangan dan dampak yang berbeda. Berdasarkan Masterplan Drainase Kota Manado Tahun 2024, penilaian prioritas perbaikan mempertimbangkan sembilan parameter, yaitu tinggi genangan, luas genangan, lama genangan, frekuensi genangan, kerugian ekonomi, gangguan sosial, gangguan transportasi, gangguan perumahan, dan kerugian harta benda [4]. Parameter tersebut mencakup kondisi genangan serta dampak yang ditimbulkan terhadap aktivitas masyarakat dan lingkungan sekitar. Banyaknya parameter yang harus dipertimbangkan secara bersamaan menyebabkan proses penentuan prioritas memerlukan pengolahan data yang terstruktur agar rekomendasi yang dihasilkan dapat digunakan sebagai dasar pendukung pengambilan keputusan.

Salah satu pendekatan yang dapat digunakan adalah Sistem Pendukung Keputusan (SPK). SPK memanfaatkan data dan model analisis untuk membantu pengambil keputusan dalam menangani permasalahan yang kompleks sehingga proses penilaian dapat dilakukan secara lebih terstruktur. Dalam penelitian ini, SPK dikembangkan berbasis web untuk mengintegrasikan pengelolaan data, proses klasifikasi, evaluasi model, serta penyajian hasil dalam satu sistem. Pendekatan berbasis web juga memungkinkan hasil klasifikasi dan aturan keputusan dapat diakses serta ditelusuri oleh pengguna sesuai kebutuhan pengambilan keputusan [5], [6].

Algoritma C4.5 digunakan sebagai metode klasifikasi karena mampu menghasilkan pohon keputusan dan aturan IF–THEN yang dapat digunakan untuk menjelaskan dasar klasifikasi. Algoritma ini menggunakan Gain Ratio dalam menentukan atribut pemisah pada proses pembentukan pohon keputusan [7]. Karakteristik tersebut sesuai dengan penelitian yang melibatkan sembilan parameter penilaian dan membutuhkan hasil klasifikasi yang dapat diinterpretasikan. Selain itu, penggunaan algoritma C4.5 pada permasalahan klasifikasi telah diterapkan pada berbagai bidang, termasuk pemetaan kerusakan infrastruktur dan identifikasi faktor yang berkaitan dengan kejadian banjir [6], [8].

Sejumlah penelitian terdahulu telah menerapkan pendekatan berbeda dalam mendukung penentuan prioritas dan analisis permasalahan pada bidang infrastruktur. Imrona et al. menggunakan metode AHP dan COPRAS-G untuk menentukan prioritas perbaikan drainase pada jalan nasional di Kota Bandung [9], sedangkan Justam et al. menerapkan algoritma C4.5 dan Random Forest untuk pemetaan kerusakan jalan berbasis WebGIS [6]. Gama et al. menerapkan algoritma C4.5 untuk mengidentifikasi faktor-faktor yang berkontribusi terhadap kerentanan banjir di Kota Denpasar dengan mempertimbangkan curah hujan, jenis tanah, elevasi, penggunaan lahan, dan jarak dari sungai [10]. Penelitian Evaria dan Istiqomah menggunakan pendekatan analisis cuaca ekstrem untuk mengidentifikasi faktor penyebab banjir di Kota Manado [3].

Berdasarkan penelitian terdahulu, algoritma C4.5 telah diterapkan untuk mengidentifikasi faktor kerentanan banjir [10] dan pemetaan kerusakan infrastruktur [6], sedangkan metode AHP–COPRAS-G telah digunakan untuk menentukan prioritas perbaikan drainase [9]. Namun, penelitian tersebut memiliki fokus yang berbeda dengan penelitian ini. Gama et al. berfokus pada identifikasi faktor yang memengaruhi kerentanan banjir di Kota Denpasar, sedangkan penelitian ini berfokus pada klasifikasi prioritas perbaikan drainase berdasarkan sembilan parameter penilaian yang digunakan oleh PUPR Kota Manado [10]. Penelitian terkait banjir di Kota Manado juga berfokus pada analisis penyebab cuaca ekstrem dan belum membahas klasifikasi prioritas perbaikan drainase [3]. Dengan demikian, masih terdapat kebutuhan untuk mengintegrasikan algoritma C4.5 dengan parameter penilaian drainase PUPR Kota Manado ke dalam Sistem Pendukung Keputusan berbasis web untuk menghasilkan klasifikasi prioritas perbaikan drainase.

Penelitian ini mengatasi keterbatasan tersebut dengan mengembangkan SPK berbasis web yang mengintegrasikan algoritma C4.5 untuk mengklasifikasikan prioritas perbaikan drainase ke dalam tiga kategori, yaitu rendah, sedang, dan tinggi, berdasarkan sembilan parameter penilaian PUPR Kota Manado. Model yang dihasilkan tidak hanya memberikan kategori prioritas, tetapi juga membentuk pohon keputusan dan aturan IF–THEN sehingga dasar klasifikasi dapat ditelusuri. Kontribusi penelitian terletak pada

penerapan algoritma C4.5 untuk klasifikasi prioritas perbaikan drainase di Kota Manado dan integrasinya ke dalam SPK berbasis web sebagai alat bantu pengambilan keputusan. Berdasarkan permasalahan dan kesenjangan penelitian tersebut, penelitian ini bertujuan mengembangkan SPK prioritas perbaikan drainase Kota Manado berbasis web menggunakan algoritma C4.5 Decision Tree serta mengevaluasi kinerja model menggunakan accuracy, precision, recall, dan F1-score.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan tahapan yang meliputi identifikasi permasalahan, studi literatur, pengumpulan data dan dokumen, penyusunan dataset dummy, preprocessing data, pembentukan model klasifikasi menggunakan algoritma C4.5, pengembangan sistem pendukung keputusan berbasis web, dan evaluasi model. Tahapan tersebut dilakukan secara berurutan untuk menghasilkan model klasifikasi yang dapat diintegrasikan ke dalam sistem sebagai alat bantu penentuan prioritas perbaikan drainase. Tahapan tersebut dirancang agar proses penelitian berlangsung secara sistematis sehingga mampu menghasilkan sistem pendukung keputusan yang dapat membantu menentukan prioritas perbaikan drainase di Kota Manado.

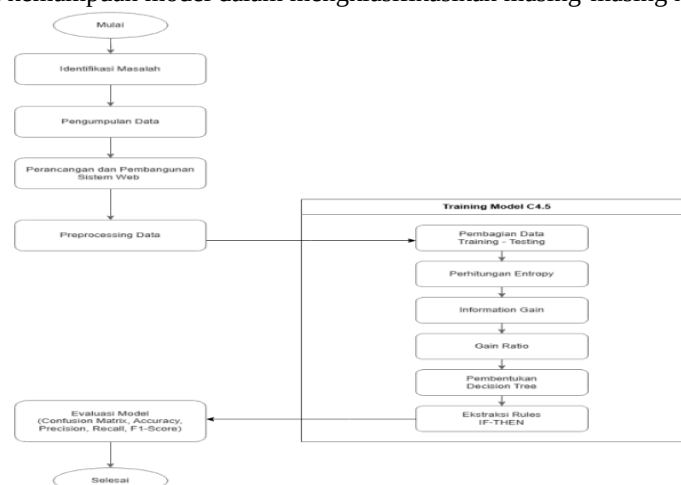
Tahap identifikasi permasalahan dilakukan melalui studi literatur dan analisis terhadap proses penentuan prioritas perbaikan drainase yang diterapkan oleh Dinas Pekerjaan Umum dan Penataan Ruang (PUPR) Kota Manado. Hasil analisis menunjukkan bahwa banyaknya lokasi drainase yang memerlukan penanganan serta keterbatasan sumber daya menyebabkan proses penentuan prioritas memerlukan pendekatan yang lebih objektif dan berbasis data.

Tahap berikutnya adalah pengumpulan data dilakukan melalui diskusi dengan Dinas Pekerjaan Umum dan Penataan Ruang (PUPR) Kota Manado serta studi dokumentasi terhadap dokumen terkait penilaian prioritas perbaikan drainase. Data yang diperoleh digunakan sebagai dasar untuk mengidentifikasi parameter, kriteria, dan mekanisme pembobotan dalam penentuan prioritas perbaikan drainase. Karena keterbatasan data historis yang tersedia, dataset pelatihan disusun dalam bentuk data dummy berdasarkan indikator, kriteria, dan sistem pembobotan yang digunakan oleh PUPR Kota Manado. Dataset yang dihasilkan terdiri atas 805 data dengan sembilan atribut input, yaitu tinggi genangan, luas genangan, lama genangan, frekuensi genangan, kerugian ekonomi, gangguan sosial, gangguan transportasi, gangguan perumahan, dan kerugian harta benda. Variabel target terdiri atas tiga kelas prioritas, yaitu Rendah, Sedang, dan Tinggi. Kesembilan atribut tersebut dipilih karena merepresentasikan kondisi genangan dan dampak yang ditimbulkan terhadap masyarakat serta lingkungan sekitar.

Selanjutnya dilakukan *preprocessing* untuk memastikan kualitas dan konsistensi data sebelum digunakan dalam pembentukan model klasifikasi. Tahap *preprocessing* meliputi pembersihan data kategorikal, pemeriksaan data duplikat, validasi dan normalisasi kategori, pemeriksaan nilai kosong (*missing value*), serta deteksi dan penanganan outlier. Pembersihan data kategorikal dilakukan dengan menyeragamkan format penulisan, termasuk perbedaan huruf, spasi, dan variasi penulisan. Data duplikat diperiksa berdasarkan identitas blok layanan untuk memastikan tidak terdapat data yang direpresentasikan lebih dari satu kali. Nilai pada atribut kategorikal kemudian divalidasi dan disesuaikan dengan kategori yang telah ditetapkan. Model klasifikasi kemudian dibangun menggunakan algoritma C4.5. Model yang dihasilkan diintegrasikan ke dalam sistem berbasis web sehingga pengguna dapat melakukan proses pelatihan model, prediksi data baru, visualisasi pohon keputusan, serta melihat hasil klasifikasi melalui satu aplikasi yang terintegrasi.

Setelah *preprocessing*, dataset dibagi menjadi data pelatihan dan data pengujian dengan rasio 80:20 menggunakan random seed 42. Dari 805 data, sebanyak 644 data digunakan sebagai data pelatihan dan 161 data digunakan sebagai data pengujian. Pembagian Dataset dengan rasio 80:20 untuk memberikan proporsi data yang lebih besar pada proses pembentukan model, sekaligus mempertahankan 20% data sebagai data pengujian yang tidak digunakan selama proses pelatihan. Pembagian ini memungkinkan kinerja model dievaluasi menggunakan data yang terpisah dari data pembentukan model. Data pelatihan digunakan untuk membentuk model klasifikasi menggunakan algoritma C4.5, sedangkan data pengujian digunakan untuk mengevaluasi model terhadap data yang tidak digunakan selama proses pembentukan model. Algoritma C4.5 membentuk pohon keputusan dengan menghitung entropy, Information Gain, Split Information, dan Gain Ratio pada setiap atribut. Atribut dengan nilai Gain Ratio tertinggi dipilih sebagai node pada pohon keputusan, kemudian proses pemisahan dilakukan secara berulang hingga terbentuk struktur pohon keputusan. Model yang terbentuk menghasilkan aturan klasifikasi dalam bentuk IF-THEN yang digunakan sebagai dasar prediksi terhadap data baru. Evaluasi model dilakukan menggunakan confusion matrix

dengan metrik accuracy, precision, recall, dan F1-score. Accuracy digunakan untuk mengukur proporsi keseluruhan data yang diklasifikasikan dengan benar, sedangkan precision, recall, dan F1-score digunakan untuk mengevaluasi kemampuan model dalam mengklasifikasikan masing-masing kelas prioritas.



Gambar 1. Tahapan Penelitian

2.2 Algoritma C4.5

Algoritma C4.5 digunakan sebagai metode klasifikasi untuk menentukan tingkat prioritas perbaikan drainase berdasarkan karakteristik setiap lokasi. Proses diawali dengan menghitung nilai *Entropy* untuk mengukur tingkat ketidakpastian data, kemudian dilanjutkan dengan perhitungan *Information Gain* dan *Gain Ratio* untuk menentukan atribut terbaik pada setiap proses pembentukan *Decision Tree*. Persamaan yang digunakan pada algoritma C4.5 mengacu pada metode yang diperkenalkan oleh Quinlan dalam pengembangan algoritma C4.5 Decision Tree [8].

Atribut dengan nilai *Gain Ratio* tertinggi dipilih sebagai simpul (*node*) pada pohon keputusan. Proses tersebut dilakukan secara berulang hingga seluruh data berada pada kelas yang sama atau tidak terdapat lagi atribut yang dapat digunakan sebagai pemisah. Hasil akhir dari proses ini berupa pohon keputusan (*Decision Tree*) beserta aturan klasifikasi (*IF-THEN Rules*) yang digunakan dalam proses prediksi.

Perhitungan entropi dilakukan menggunakan Persamaan.

$$Entropy(S) = -\sum_{i=1}^n p_i \log_2(p_i) \quad (1)$$

Selanjutnya nilai *Gain Ratio* dihitung menggunakan Persamaan.

$$GainRatio(S, A) = \frac{Gain(S, A)}{SplitInfo(S, A)} \quad (2)$$

Model yang telah terbentuk kemudian disimpan ke dalam sistem sehingga dapat digunakan pada proses prediksi data baru tanpa perlu membangun ulang model setiap kali sistem dijalankan.

2.3 Evaluasi Model

Evaluasi dilakukan setelah model klasifikasi berhasil dibangun menggunakan algoritma C4.5. Tujuan evaluasi adalah mengukur kemampuan model dalam mengklasifikasikan data prioritas perbaikan drainase yang tidak digunakan pada proses pelatihan. Pengujian dilakukan menggunakan data *testing* sebanyak 161 data yang diperoleh dari pembagian dataset dengan rasio 80:20.

Kinerja model dievaluasi menggunakan *confusion matrix* yang membandingkan hasil prediksi dengan kelas sebenarnya. Berdasarkan *confusion matrix* dihitung empat metrik evaluasi, yaitu *accuracy*, *precision*, *recall*, dan *F1-score*. *Accuracy* digunakan untuk mengukur persentase prediksi yang benar terhadap seluruh data uji, sedangkan *precision*, *recall*, dan *F1-score* digunakan untuk memberikan gambaran performa model pada setiap kelas prioritas. Pendekatan ini dipilih karena mampu memberikan evaluasi yang lebih komprehensif terhadap kemampuan model klasifikasi dibandingkan hanya menggunakan nilai akurasi.

Nilai *accuracy* dihitung menggunakan Persamaan:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

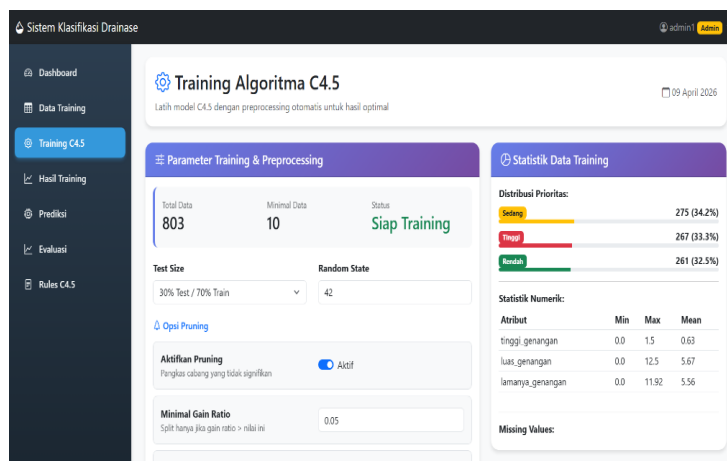
3. HASIL DAN PEMBAHASAN

3.1 Implementasi Sistem

Sistem Pendukung Keputusan (SPK) prioritas perbaikan drainase dikembangkan dalam bentuk aplikasi berbasis web dengan mengintegrasikan algoritma C4.5 sebagai metode klasifikasi. Sistem dirancang untuk membantu proses penentuan prioritas perbaikan drainase berdasarkan parameter penilaian yang digunakan oleh Dinas Pekerjaan Umum dan Penataan Ruang (PUPR) Kota Manado. Implementasi berbasis web memungkinkan seluruh proses, mulai dari pengelolaan data hingga prediksi, dilakukan melalui satu aplikasi yang terintegrasi.

Sistem dikembangkan menggunakan bahasa pemrograman Python dengan framework Flask, sedangkan MySQL digunakan sebagai basis data untuk menyimpan data pelatihan, hasil pelatihan model, serta riwayat prediksi. Antarmuka sistem dibangun menggunakan HTML, CSS, dan Bootstrap sehingga dapat diakses melalui peramban web tanpa memerlukan instalasi perangkat lunak tambahan.

Implementasi sistem dilakukan dengan mengintegrasikan algoritma C4.5 ke dalam Sistem Pendukung Keputusan berbasis web untuk mendukung proses klasifikasi prioritas perbaikan drainase. Fitur yang tersedia pada sistem meliputi pengelolaan data *training*, proses *preprocessing*, pelatihan model menggunakan algoritma C4.5, visualisasi *Decision Tree*, ekstraksi aturan keputusan (*IF-THEN Rules*), evaluasi model menggunakan *confusion matrix*, serta prediksi terhadap data baru. Integrasi tersebut memungkinkan proses pengolahan data dan penerapan model klasifikasi dilakukan dalam satu sistem sehingga hasil klasifikasi dapat digunakan sebagai rekomendasi dalam penentuan prioritas perbaikan drainase.



Gambar 2. Halaman Training C4.5

Gambar 2 menunjukkan halaman pelatihan model yang digunakan untuk membangun model klasifikasi menggunakan dataset pelatihan. Pada halaman ini pengguna dapat menjalankan proses pelatihan sehingga sistem menghitung nilai *Entropy*, *Information Gain*, dan *Gain Ratio* secara otomatis hingga terbentuk model *Decision Tree*. Model yang telah berhasil dibangun selanjutnya disimpan ke dalam sistem dan digunakan pada proses prediksi tanpa perlu dilakukan pelatihan ulang.

3.2 Pembentukan Model C4.5

Model klasifikasi pada penelitian ini dibangun menggunakan algoritma C4.5 berdasarkan dataset yang telah melalui tahap preprocessing. Dataset terdiri atas sembilan atribut yang merepresentasikan kondisi genangan dan dampaknya, yaitu tinggi genangan, luas genangan, lama genangan, frekuensi genangan, kerugian ekonomi, gangguan sosial, gangguan transportasi, gangguan perumahan, dan kerugian harta benda. Setelah proses preprocessing selesai, dataset dibagi menjadi data training dan testing dengan proporsi 80:20. Sebanyak 644 data digunakan sebagai data pelatihan untuk membangun model klasifikasi, sedangkan 161 data digunakan sebagai data pengujian.

Pembentukan model dilakukan melalui perhitungan nilai entropy dan Gain Ratio pada setiap atribut. Nilai Gain Ratio digunakan sebagai dasar pemilihan atribut terbaik dalam setiap proses pemisahan data (splitting). Atribut yang memiliki nilai Gain Ratio tertinggi dipilih sebagai simpul (node) pada pohon keputusan, kemudian proses yang sama dilakukan secara berulang pada setiap cabang hingga seluruh data

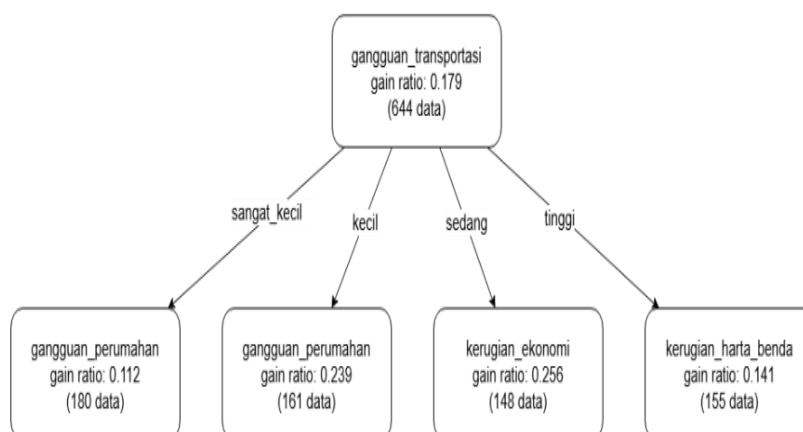
berhasil dikelompokkan ke dalam kelas prioritas yang sesuai. Tahapan ini menghasilkan struktur Decision Tree yang menjadi dasar proses klasifikasi pada sistem.

Tabel 1. Hasil Perhitungan Gain ratio per atribut untuk *root node*

Atribut	Information Gain	Split Info	Gain Ratio
Gangguan Transportasi	0.3574	1.9961	0.1790
Kerugian Ekonomi	0.3549	1.9981	0.1776
Gangguan Perumahan	0.3140	1.9995	0.1570
Gangguan Sosial	0.2791	1.9890	0.1403
Kerugian Harta Benda	0.2763	1.9965	0.1384
Frekuensi Genangan	0.0883	2.3156	0.0381
Tinggi Genangan	0.0122	0.4719	0.0258
Luas Genangan	0.0043	0.3538	0.0121
Lamanya Genangan	0.0052	0.7348	0.0071

Berdasarkan hasil perhitungan pada Tabel 1, atribut gangguan transportasi memperoleh nilai *Gain Ratio* tertinggi, yaitu 0,1790, sehingga dipilih sebagai simpul akar (*root node*) pada *Decision Tree*. Hal tersebut menunjukkan bahwa atribut tersebut memiliki kemampuan paling baik dalam membedakan kelas prioritas dibandingkan atribut lainnya pada dataset. Dengan kata lain, tingkat gangguan terhadap aktivitas transportasi menjadi faktor awal yang paling menentukan dalam proses klasifikasi prioritas perbaikan drainase.

Pemilihan atribut tersebut juga menunjukkan bahwa dampak genangan terhadap mobilitas masyarakat memiliki pengaruh yang besar dalam menentukan tingkat prioritas penanganan. Setelah simpul akar ditentukan, algoritma kembali menghitung nilai *Gain Ratio* pada setiap cabang sehingga terbentuk struktur pohon keputusan secara bertahap hingga seluruh data berhasil diklasifikasikan.



Gambar 3 Decision Tree Root Awal

Gambar 3 memperlihatkan struktur awal Decision Tree yang dihasilkan oleh algoritma C4.5. Struktur tersebut menggambarkan urutan atribut yang digunakan dalam proses klasifikasi berdasarkan nilai *Gain Ratio* pada setiap proses pemisahan data. Semakin ke bawah, pohon keputusan akan membentuk aturan klasifikasi yang semakin spesifik hingga menghasilkan kelas prioritas Rendah, Sedang, atau Tinggi. Model yang terbentuk kemudian menghasilkan sejumlah aturan keputusan (IF-THEN Rules) yang digunakan pada proses prediksi. Aturan tersebut merupakan representasi dari struktur Decision Tree sehingga setiap hasil klasifikasi dapat dijelaskan berdasarkan kombinasi atribut yang dimiliki suatu lokasi drainase. Dengan demikian, model tidak hanya menghasilkan kategori prioritas, tetapi juga memberikan penjelasan mengenai dasar pengambilan keputusan.

Berdasarkan hasil model yang telah dibangun, diperoleh sejumlah aturan keputusan yang menggambarkan hubungan antara atribut kondisi drainase dengan kategori prioritas perbaikan. Contoh aturan yang dihasilkan adalah sebagai berikut:

IF gangguan_transportasi = 'sedang' AND kerugian_ekonomi = 'sedang' AND kerugian_harta_benda = 'sedang' AND gangguan_sosial = 'kecil' THEN Prioritas = sedang

IF gangguan_transportasi = 'sedang' AND kerugian_ekonomi = 'sedang' AND kerugian_harta_benda = 'sedang' AND gangguan_sosial = 'tinggi' THEN Prioritas = tinggi

IF gangguan_transportasi = 'sedang' AND kerugian_ekonomi = 'sedang' AND kerugian_harta_benda = 'sangat_kecil' AND luas_genangan $\leq$ 5.00 THEN Prioritas = sedang

3.3 Evaluasi Model C4.5

Setelah proses pembentukan model selesai, dilakukan evaluasi untuk mengukur kemampuan algoritma C4.5 dalam mengklasifikasikan prioritas perbaikan drainase. Pengujian menggunakan 161 data uji yang tidak dilibatkan pada proses pelatihan sehingga hasil evaluasi dapat menggambarkan kemampuan model dalam melakukan klasifikasi terhadap data baru. Evaluasi dilakukan menggunakan *confusion matrix* yang membandingkan hasil prediksi model dengan kelas aktual. Berdasarkan matriks tersebut dihitung nilai *accuracy*, *precision*, *recall*, dan *F1-score* sehingga performa model dapat dianalisis baik secara keseluruhan maupun pada masing-masing kelas prioritas.

Aktual	Prediksi		
	Rendah	Sedang	Tinggi
Rendah	45	6	1
Sedang	9	29	6
Tinggi	0	16	49

Gambar 4. *Confusion Matrix*

Gambar 4 menunjukkan *Confusion matrix* menunjukkan bahwa sebanyak 45 data kelas Rendah, 29 data kelas Sedang, dan 49 data kelas Tinggi berhasil diklasifikasikan dengan benar. Kesalahan klasifikasi paling banyak terjadi pada data kelas Tinggi yang diprediksi sebagai kelas Sedang, yaitu sebanyak 16 data. Pada kelas Sedang, sebanyak 9 data diprediksi sebagai Rendah dan 6 data diprediksi sebagai Tinggi. Pola tersebut menunjukkan bahwa kelas Sedang memiliki kecenderungan lebih besar untuk beririsan dengan karakteristik kelas Rendah dan Tinggi, sehingga menjadi kelas yang lebih sulit dibedakan oleh model.

Berdasarkan hasil evaluasi, model memperoleh nilai *accuracy* sebesar 76,40%, dengan 123 dari 161 data pengujian berhasil diklasifikasikan sesuai dengan kelas aktual, sedangkan 38 data lainnya mengalami kesalahan klasifikasi. Nilai tersebut menunjukkan bahwa model mampu menghasilkan prediksi yang benar pada sebagian besar data pengujian, tetapi masih terdapat sejumlah data yang memiliki karakteristik yang menyebabkan model menghasilkan kelas yang berbeda dari kelas aktual.

Untuk memberikan gambaran yang lebih rinci mengenai performa model pada setiap kelas, nilai *precision*, *recall*, dan *F1-score* disajikan pada Tabel 2.

Tabel 2. Hasil Evaluasi

Kelas Prioritas	Precision	Recall	F1-Score
Rendah	83.33%	86.54%	84.91%
Sedang	56.86%	65.91%	61.05%
Tinggi	87.50%	75.38%	80.99
Accuracy			76.40%

Berdasarkan Tabel 2, kelas Rendah memperoleh nilai *precision* sebesar 83,33%, *recall* sebesar 86,54%, dan *F1-score* sebesar 84,91%, yang menunjukkan bahwa model mampu mengenali kelas tersebut dengan tingkat konsistensi yang tinggi. Pada kelas Sedang, nilai *precision* sebesar 56,86%, *recall* sebesar 65,91%, dan *F1-score* sebesar 61,05%, sehingga menjadi kelas dengan performa terendah dibandingkan dua kelas lainnya. Sementara itu, kelas Tinggi memperoleh nilai *precision* sebesar 87,50%, *recall* sebesar 75,38%, dan *F1-score* sebesar 80,99%, yang menunjukkan bahwa model memiliki kemampuan yang baik dalam mengidentifikasi lokasi yang termasuk prioritas tinggi. Hasil tersebut menunjukkan bahwa kelas Sedang memiliki nilai *precision*, *recall*, dan *F1-score* paling rendah dibandingkan dengan kelas Rendah dan Tinggi. Performa kelas Sedang yang lebih rendah menunjukkan bahwa model lebih sulit membedakan data yang termasuk dalam kategori Sedang dibandingkan dengan kategori Rendah dan Tinggi. Kondisi tersebut dapat terjadi karena karakteristik data pada kelas Sedang berada di antara dua kategori lainnya, sehingga beberapa kombinasi atribut memiliki pola yang mendekati kelas Rendah maupun Tinggi.

3.4 Hasil Prediksi

Setelah model klasifikasi berhasil dibentuk, sistem digunakan untuk melakukan prediksi terhadap data drainase yang belum pernah diproses sebelumnya. Proses prediksi memanfaatkan model C4.5 yang telah disimpan pada tahap pelatihan sehingga setiap data baru dapat langsung diklasifikasikan tanpa perlu melakukan proses pelatihan ulang.

ID Blok	Tinggi (m)	Luas (ha)	Lama (jam)	Frekuensi	Ekonomi	Sosial	Transport	Perumahan	Harta	Hasil	Tanggal	Aksi
Blok 1	0.30	12.50	4.00	KS	tinggi	sedang	tinggi	tinggi	kecil	Tinggi	04/07/2026 12:08	
Blok 2	0.10	8.00	4.00	KS	tinggi	sedang	sedang	tinggi	kecil	Tinggi	04/07/2026 12:08	
Blok 3	0.20	2.00	2.00	KS	tinggi	sedang	tinggi	tinggi	kecil	Tinggi	04/07/2026 12:08	
Blok 4	0.20	4.00	4.00	KS	tinggi	sedang	tinggi	tinggi	kecil	Tinggi	04/07/2026 12:08	
Blok 5	0.50	10.00	4.00	KS	tinggi	sedang	tinggi	tinggi	kecil	Tinggi	04/07/2026 12:08	
Blok 6	1.00	4.50	4.00	KS	sedang	sedang	kecil	sedang	kecil	Sedang	04/07/2026 12:08	
Blok 7	0.20	2.00	2.00	JR	sedang	sedang	sedang	tinggi	kecil	Sedang	04/07/2026 12:08	
Blok 8	0.50	6.00	4.00	KS	sedang	sedang	sedang	sedang	kecil	Sedang	04/07/2026 12:08	
Blok 9B	0.50	8.00	4.00	KS	kecil	sedang	kecil	sedang	kecil	Sedang	04/07/2026 12:08	
Blok 9A	1.00	10.50	8.00	KS	kecil	sedang	sedang	tinggi	kecil	Sedang	04/07/2026 12:08	

Gambar 5. Hasil Prediksi

Berdasarkan Gambar 5, pengguna memasukkan data kondisi drainase yang terdiri atas atribut tinggi genangan, luas genangan, lama genangan, frekuensi genangan, tingkat kerugian ekonomi, gangguan sosial, gangguan transportasi, gangguan perumahan, dan kerugian harta benda. Selanjutnya sistem memproses data tersebut menggunakan aturan (*IF-THEN Rules*) yang dihasilkan dari algoritma C4.5 dan menampilkan hasil klasifikasi dalam tiga kategori prioritas, yaitu Rendah, Sedang, dan Tinggi. Pada hasil prediksi terlihat bahwa beberapa lokasi, seperti Blok 1 hingga Blok 5, diklasifikasikan sebagai prioritas tinggi, sedangkan Blok 6 hingga Blok 9A diklasifikasikan sebagai prioritas sedang. Perbedaan hasil tersebut menunjukkan bahwa setiap lokasi memiliki kombinasi atribut yang berbeda sehingga menghasilkan tingkat prioritas yang berbeda pula sesuai dengan pola yang dipelajari oleh model. Hasil prediksi ini menunjukkan bahwa model C4.5 yang telah dibangun tidak hanya mampu menghasilkan struktur *Decision Tree* selama proses pelatihan, tetapi juga dapat diimplementasikan untuk memberikan rekomendasi prioritas terhadap data baru. Dengan demikian, sistem dapat dimanfaatkan sebagai alat bantu dalam proses penentuan prioritas perbaikan drainase secara lebih objektif dan berbasis data.

4. KESIMPULAN

Berdasarkan hasil penelitian, Sistem Pendukung Keputusan prioritas perbaikan drainase Kota Manado berbasis web berhasil dikembangkan dengan mengintegrasikan algoritma C4.5 sebagai model klasifikasi. Model dibangun menggunakan sembilan parameter penilaian dan menghasilkan tiga kategori prioritas, yaitu Rendah, Sedang, dan Tinggi. Hasil pembentukan *Decision Tree* menunjukkan bahwa atribut gangguan transportasi memiliki nilai Gain Ratio tertinggi sebesar 0,1790 sehingga dipilih sebagai root node. Model menghasilkan lebih dari 50 aturan *IF-THEN* yang digunakan dalam proses klasifikasi, sehingga hasil prediksi dapat ditelusuri berdasarkan jalur keputusan yang terbentuk. Evaluasi menggunakan 161 data pengujian menghasilkan accuracy sebesar 76,40%, dengan 123 data berhasil diklasifikasikan dengan benar. Kelas Rendah memperoleh F1-score sebesar 84,91%, kelas Sedang sebesar 61,05%, dan kelas Tinggi sebesar 80,99%. Perbedaan kinerja antar kelas menunjukkan bahwa model masih mengalami kesulitan dalam membedakan sebagian data pada kategori Sedang yang memiliki karakteristik yang berdekatan dengan kategori Rendah dan Tinggi. Secara praktis, sistem yang dikembangkan dapat digunakan sebagai alat bantu untuk mengolah parameter penilaian, menghasilkan klasifikasi prioritas, dan menyajikan aturan keputusan secara terstruktur sehingga dapat mendukung proses penentuan prioritas perbaikan drainase. Kontribusi penelitian ini terletak pada integrasi algoritma C4.5 dengan Sistem Pendukung Keputusan berbasis web untuk klasifikasi prioritas perbaikan drainase berdasarkan parameter penilaian yang digunakan PUPR Kota Manado. Namun, evaluasi model masih menggunakan metode hold-out dengan rasio 80:20 pada satu pembagian dataset. Penelitian selanjutnya dapat menerapkan k-fold cross-validation serta membandingkan C4.5 dengan algoritma klasifikasi lainnya untuk memperoleh evaluasi yang lebih komprehensif terhadap kestabilan dan kinerja model.

UCAPAN TERIMA KASIH

Terima kasih penulis sampaikan kepada para dosen pembimbing yang selalu memberikan arahan, masukan dan bantuan selama penelitian ini berlangsung. Ucapan terimakasih juga penulis sampaikan kepada semua pihak yang mendukung penelitian ini hingga bisa selesai.

REFERENCES

- [1] K. P. U. dan P. R. Indonesia, *Peraturan Menteri Pekerjaan Umum dan Perumahan Rakyat Nomor 12/PRT/M/2014 Tahun 2014 tentang Penyelenggaraan Sistem Drainase Perkotaan*. Indonesia: BN. 2014/NO.1451, Jdih.pu.go.id: 18 hlm., 2014.
- [2] A. Notanubun, R. J. Betaubun, and S. Johanes, "EVALUASI SISTEM DRAINASE RUAS JALAN LOKKI - IHA KULUR DESA LUHU KECAMATAN HUAMUAL KABUPATEN SERAM BAGIAN BARAT," *Journal Agregate*, vol. 2, no. 2, pp. 160–167, Sep. 2022, doi: 10.31959/ja.v2i2.1374.
- [3] Ni Luh Made Anik Evaria and Rafa Zafirah Istiqomah, "Analysis of Extreme Weather Causes of Flooding in Manado North Sulawesi," *Journal Of Global Research Publications*, vol. 1, no. 1, pp. 12–17, Jan. 2024, doi: 10.59435/jgrp.v1i1.2024.3.
- [4] Dinas Pekerjaan Umum dan Penataan Ruang Kota Manado, "Laporan Akhir Survey Kondisi dan Penyusunan Masterplan Drainase Kota Manado Tahun 2024," Manado, 2024.
- [5] N. Jajac, "Decision Support for Infrastructure Management of Public Institutions," *Sustainability (Switzerland)*, vol. 17, no. 5, Mar. 2025, doi: 10.3390/su17052096.
- [6] J. Justam, N. Jamilah, S. M. Umar, E. Erlita, and J. Ramba, "Penerapan Algoritma C4.5 dan Random Forest untuk Pemetaan Kerusakan Jalan dengan WebGIS," *Jurnal Ilmiah Sistem Informasi dan Teknik Informatika (JISTI)*, vol. 7, no. 2, pp. 326–339, Oct. 2024, doi: 10.57093/jisti.v7i2.270.
- [7] E. Vinck *et al.*, "A new model to evaluate the impact of SUDS on water quality," 2023, [Online]. Available: <https://hal.science/hal-04184255v1>
- [8] J. R. . Quinlan, *C4.5 : programs for machine learning*. Morgan Kaufmann Publishers, 1993.
- [9] M. Imrona, A. A. Budiutama, E. Darwiyanto, and D. Handayani, "Penerapan Metode AHP dan COPRAS-G untuk Menentukan Prioritas Perbaikan Drainase Pada Jalan Nasional Di Kota Bandung," *Indonesian Journal on Computing (Indo-JC)*, vol. 4, no. 1, p. 65, Mar. 2019, doi: 10.21108/INDOJC.2019.4.1.261.
- [10] A. W. O. Gama, M. Dennatan, I. G. N. P. Dharmayasa, M. M. Maw, I. P. Sugiana, and I. Suryanti, "Identifying Key Factors Causing Flooding Using Machine Learning," *Journal of Applied Data Sciences*, vol. 6, no. 1, pp. 115–130, Jan. 2025, doi: 10.47738/jads.v6i1.463.