

ANALISIS BENCANA BANJIR KAB. BOYOLALI MENGGUNAKAN METODE DECISION TREE PADA RAPID MINER

Muh Rizky Anggara Yunan Putra¹, Dwiningsih², Okta Viona Cahyanti³, Dewi Oktafiani⁴

Prodi Informatika, STMIK Amikom Surakarta¹⁻⁴

Email: muh.10372@mhs.amikomsolo.ac.id¹, dwiningsih.10353@mhs.amikomsolo.ac.id²,
okta.10383@mhs.amikomsolo.ac.id³, dewioktafiani@dosen.amikomsolo.ac.id⁴

Informasi	Abstract
Volume : 2 Nomor : 12 Bulan : Desember Tahun : 2025 E-ISSN : 3062-9624	<p><i>Flood disasters are a natural phenomenon that often occur in various regions in Indonesia, including cities such as Banjarnegara, Boyolali and Karanganyar. Floods have a significant impact on the economy, social and environment. Therefore, proper understanding and analysis of the factors that cause flooding is very important for effective mitigation efforts. This research uses a decision tree algorithm to analyze meteorological data including temperature, wind speed, humidity and rainfall from various cities in Indonesia. This data includes normal to hot temperature conditions, slightly calm wind speed to light gusts, high to moderate humidity, and extreme to very high rainfall. The results of the analysis show patterns and relationships between these variables and flood events. The decision tree algorithm is used to build a prediction model in the form of a decision tree, which makes interpretation and decision making easier. This research aims to identify the main factors that contribute to flooding and develop a prediction model that can be used to improve preparedness and response to flood disasters. By understanding the patterns and factors that cause flooding, it is hoped that more effective mitigation measures can be implemented to reduce the risk and impact of this disaster.</i></p>

Keyword: Flood, Decision Tree, Prediction, Accuracy

Abstrak

Bencana banjir merupakan fenomena alam yang sering terjadi di berbagai wilayah di Indonesia, termasuk kota-kota seperti Banjarnegara, Boyolali, dan Karanganyar. Banjir memiliki dampak yang signifikan terhadap ekonomi, sosial, dan lingkungan. Oleh karena itu, pemahaman dan analisis yang tepat mengenai faktor-faktor penyebab banjir sangat penting untuk upaya mitigasi yang efektif. Penelitian ini menggunakan algoritma decision tree untuk menganalisis data meteorologi yang mencakup suhu, kecepatan angin, kelembapan, dan curah hujan dari berbagai kota di Indonesia. Data tersebut mencakup kondisi temperatur normal hingga panas, kecepatan angin sedikit tenang hingga sedikit hembusan angin, kelembapan tinggi hingga sedang, dan curah hujan ekstrem hingga sangat tinggi. Hasil analisis menunjukkan pola dan hubungan antara variabel-variabel tersebut dengan kejadian banjir. Algoritma decision tree digunakan untuk membangun model prediksi dalam bentuk pohon keputusan, yang memudahkan interpretasi dan pengambilan keputusan. Penelitian ini bertujuan untuk mengidentifikasi faktor-faktor utama yang berkontribusi terhadap banjir dan mengembangkan model prediksi yang dapat digunakan untuk meningkatkan kesiapsiagaan dan respons terhadap bencana banjir. Dengan memahami pola dan faktor penyebab banjir, diharapkan langkah-langkah mitigasi yang lebih efektif dapat diterapkan untuk mengurangi risiko dan dampak bencana ini.

Kata Kunci: Banjir, Pohon Keputusan, Prediksi, Akurasi

A. PENDAHULUAN

Iklim tropis yang ada di Indonesia mengakibatkan banyak sekali bencana alam yang harus diterima oleh Indonesia[1], banjir adalah salah satu bencana alam yang sering terjadi di berbagai daerah di Indonesia, termasuk kota-kota seperti Banjarnegara, Boyolali, Karanganyar, dan lainnya. Banjir merupakan bencana yang sering dijumpai. Penyebab terjadinya banjir umumnya terdapat peningkatan volume air di bantaran sungai. Faktor-faktor yang mempengaruhi banjir di antaranya faktor alam dan faktor manusia. Hal yang dipengaruhi oleh faktor manusia yaitu seperti membuang sampah sembarangan, baik itu di jalan maupun di sungai, mendirikan bangunan di lingkungan hijau dan lain-lain. Pengaruh banjir dari faktor alam salah satunya ialah intensitas curah hujan yang tinggi[2], Dampak bencana ini sangat signifikan terhadap aspek ekonomi, sosial, dan lingkungan. Oleh karena itu, pemahaman dan analisis yang akurat mengenai faktor-faktor penyebab banjir sangat penting untuk upaya mitigasi dan penanggulangan yang lebih efektif.

Iklim merupakan hal sangat penting dalam kehidupan kita sehari-hari[3]. Perubahan iklim yang terjadi menyebabkan bencana hidrometeorologi banjir meningkat di Indonesia. Banjir merupakan bencana yang dipengaruhi oleh parameter cuaca dan iklim, seperti curah hujan, temperatur, angin, dan kelembapan[4], melakukan analisis tentang pengelompokan antara data dan atribut untuk mendukung pengambilan keputusan serta pembuatan kebijakan dalam memberikan informasi curah hujan[5].

Data Mining adalah suatu bidang interdisipliner yang berpatokan pada ilmu komputer (data base, kecerdasan buatan, pembelajaran mesin, grafis dan visualisasi model), statistik dan teknik (pengenalan pola, jaringan saraf)[6]. Anomali iklim adalah perubahan musiman dari rata-rata normal yang terjadi akibat pemanasan global, yakni peningkatan suhu bumi yang menyebabkan penguapan dan kenaikan curah hujan[7].

Dalam hal ini, data meteorologi seperti suhu, kecepatan angin, kelembapan, dan curah hujan menjadi variabel penting yang perlu diperhatikan. Dataset yang digunakan dalam penelitian ini mencakup berbagai kota di Indonesia dengan informasi mengenai kondisi temperatur, kecepatan angin, kelembapan, dan curah hujan. Sebagai contoh, kota Banjarnegara memiliki kondisi temperatur normal, kecepatan angin sedikit hembusan angin, kelembapan tinggi, dan curah hujan ekstrem. Informasi serupa juga diperoleh untuk kota-kota lain yang mengalami bencana banjir.

Untuk menganalisis data ini, digunakan algoritma decision tree, sebuah metode pembelajaran mesin yang efektif dalam mengidentifikasi pola dan hubungan antara variabel

input dengan kejadian banjir. Algoritma ini membangun model prediksi dalam bentuk pohon keputusan, yang memudahkan interpretasi hasil analisis dan pengambilan keputusan, bedasarkan penelitian [8] diketahui bahwa adanya perbedaan yang signifikan antara akurasi yang diperoleh dengan menggunakan library Gausian Naïve Bayes dari scikitlearn.org dan RapidMiner menjadikan sulit untuk disimpulkan apakah Naïve Bayes efektif untuk diterapkan pada dataset banjir.

Penelitian ini bertujuan untuk mengidentifikasi faktor-faktor utama yang berkontribusi terhadap terjadinya banjir di berbagai kota, serta untuk mengembangkan model prediksi yang dapat digunakan oleh pihak berwenang untuk meningkatkan kesiapsiagaan dan respons terhadap bencana banjir. Dengan memahami pola dan faktor penyebab banjir, diharapkan langkah-langkah mitigasi yang lebih efektif dapat diterapkan untuk mengurangi risiko dan dampak yang ditimbulkan oleh bencana ini.

B. METODE PENELITIAN

Metode penelitian yang digunakan adalah metode pengumpulan data, dimana kami mengumpulkan data dari 10 Jurnal kemudian kami analisis bagaimana pohon keputusannya dan akurasi dari penelitian jurnal-jurnal tersebut menggunakan aplikasi *RapidMiner*.



Gambar 1. Alur Metode Penelitian

1. Pengumpulan Data

Tahap pengumpulan data dilakukan dengan menganalisis tiap jurnal, kemudian ditemukan pada jurnal [9] ditemukan sebuah dataset curah hujan pada hasil penelitian yang dilakukan pada tahun 2023, dengan variable kota, temperatur, kecepatan angin, kelembapan, dan curah hujan

2. Analisis Pohon Keputusan dengan RapidMiner

Tahap Analisis Pohon Keputusan ini dilakukan dengan aplikasi Rapid Miner.

3. Analisis akurasi dengan RapidMiner

Tahap ini bertujuan untuk melihat apakah dataset dari tiap jurnal yang di analisis sesuai atau tidak, serta mencari cara untuk meningkatkan jumlah akurasi dari dataset tiap jurnal.

4. Menemukan Hasil

Tahap ini dalam merupakan tahapan akhir dimana hasil dibuat berdasarkan data hasil analisis yang sudah dilakukan. Dalam tahap ini peneliti mengambil kesimpulan dari hasil analisis keseluruhan penelitian untuk memecahkan masalah.

C. HASIL DAN PEMBAHASAN

Penelitian ini, peneliti menggunakan aplikasi RapidMiner pada prakteknya yang bertujuan untuk memudahkan peneliti untuk menentukan decision tree (pohon keputusan) dan akurasi dalam dataset yang sedang di teliti.

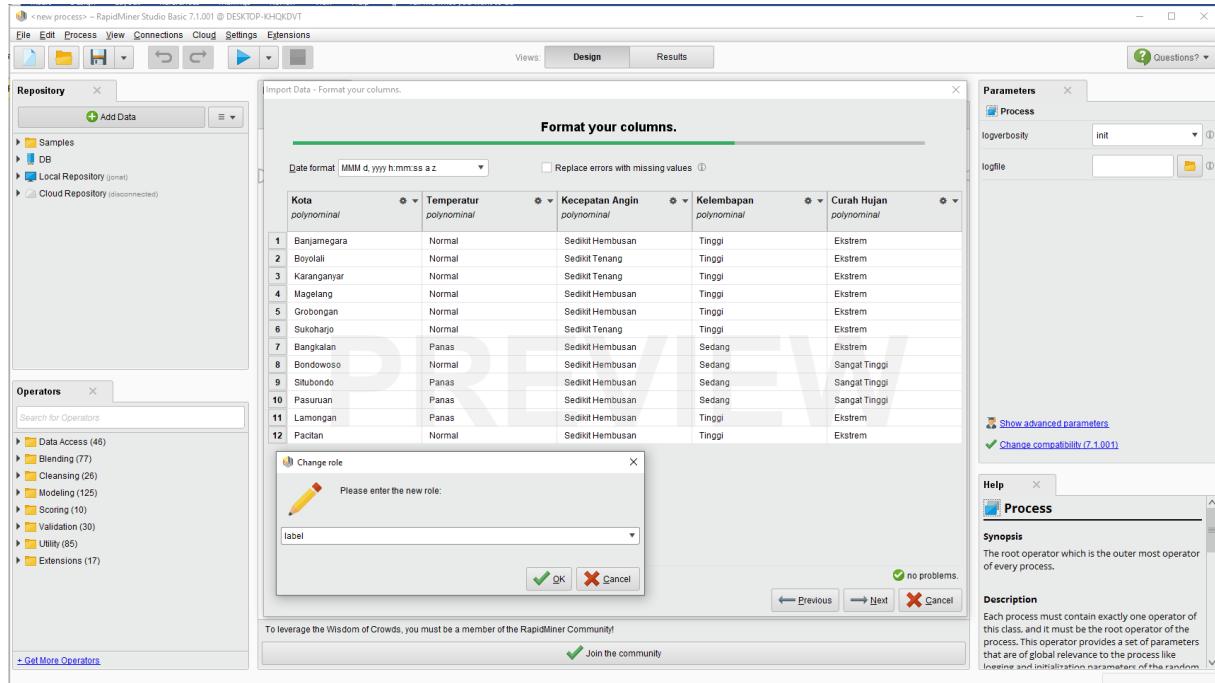
Pengumpulan Data

Kota	Temperatur	Kecepatan Angin	Kelembapan	Curah Hujan
Banjarnegara	Normal	Sedikit Hembusan	Tinggi	Ekstrem
Boyolali	Normal	Sedikit Tenang	Tinggi	Ekstrem
Karanganyar	Normal	Sedikit Tenang	Tinggi	Ekstrem
Magelang	Normal	Sedikit Hembusan	Tinggi	Ekstrem
Grobongan	Normal	Sedikit Hembusan	Tinggi	Ekstrem
Sukoharjo	Normal	Sedikit Tenang	Tinggi	Ekstrem
Bangkalan	Panas	Sedikit Hembusan	Sedang	Ekstrem
Bondowoso	Normal	Sedikit Hembusan	Sedang	Sangat Tinggi
Situbondo	Panas	Sedikit Hembusan	Sedang	Sangat Tinggi
Pasuruan	Panas	Sedikit Hembusan	Sedang	Sangat Tinggi
Lamongan	Panas	Sedikit Hembusan	Tinggi	Ekstrem
Pacitan	Normal	Sedikit Hembusan	Tinggi	Ekstrem

Tabel 1. Data Penelitian

Susun dataset tersebut dalam bentuk Microsoft Excel agar dapat dianalisis menggunakan aplikasi RapidMiner.

Menginput dataset

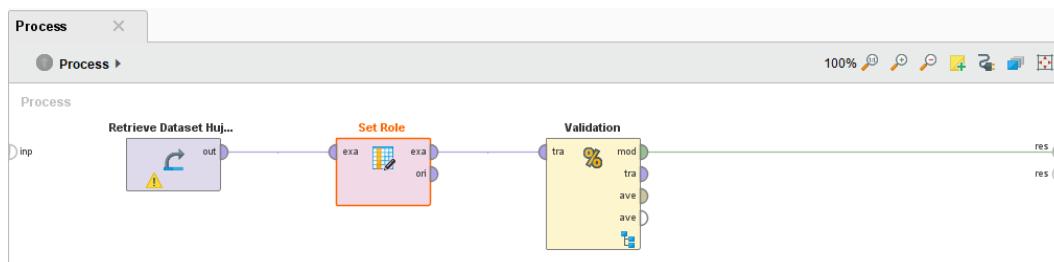


Gambar 2. Dataset pada RapidMiner

Saat menambahkan dataset kedalam RapidMiner jangan lupa jadikan atribut “Curah Hujan” menjadi sebuah label.

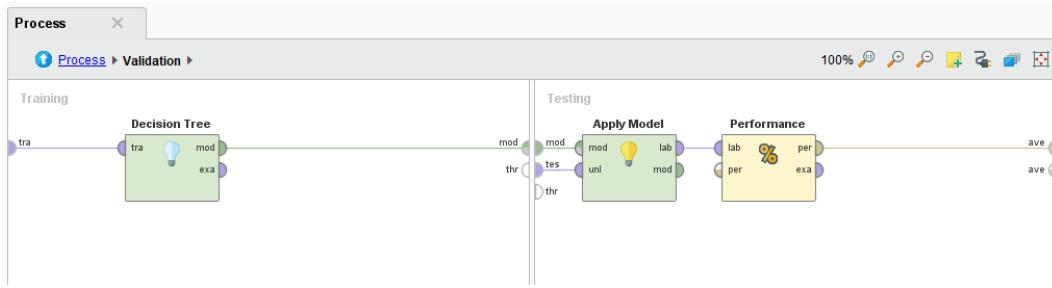
Analisis Pohon Keputusan

Susun dataset, Set Role, dan Validation seperti berikut :



Gambar 3. Design Rapid Miner

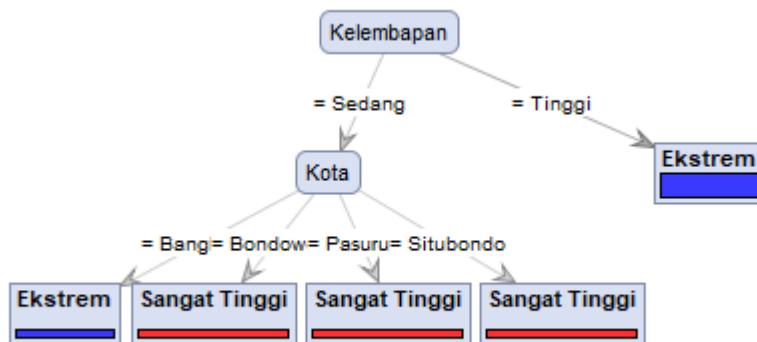
Cara membuat seperti gambar tersebut adalah masukkan dataset pada Local Repository kedalam kotak Process, langkah selanjutnya masukkan operator Set Role kedalam kotak process dan jangan lupa pada operator set role jadikan Curah Hujan menjadi label dan masukkan atrribut lain pada additional roles, kemudian masukkan juga operator Validation kedalam kotak process. Lalu sambungkan ke 3 hal tersebut dengan cara menghubungkan “out” pada dataset ke “exa” pada set role, dan “exa” kedua pada set role hubungkan ke “tra” pada validation sisanya tinggal menghubungkan “mod” pada validation ke “res” pada sisi kanan.



Gambar 4. Training dan Testing Validation

Dalam operator Validation masukkan operator Decision Tree pada kotak Training, dan operator Apply Model serta Performance pada kotak Testing. Jika sudah jangan lupa untuk menghubungkan semuanya dengan cara “tra” pada sisi kiri kotak training hubungkan pada “tra” yang terdapat pada operator decision tree kemudian “mod” pada decision tree hubungkan pada “mod” yang berada di sisi kanan kotak training. Kemudian pada kotak testing hubungkan “mod” dan “tes” pada sisi kiri kotak testing ke “mod” dan “tes” pada operator apply model, jika sudah dihubungkan jangan lupa juga hubungkan “lab” pada operator apply model ke “lab” yang berada di operator performance, terakhir hubungkan “per” pada operator performance ke “ave” yang berada di sisi kanan kotak testing.

Hasil Decision Tree

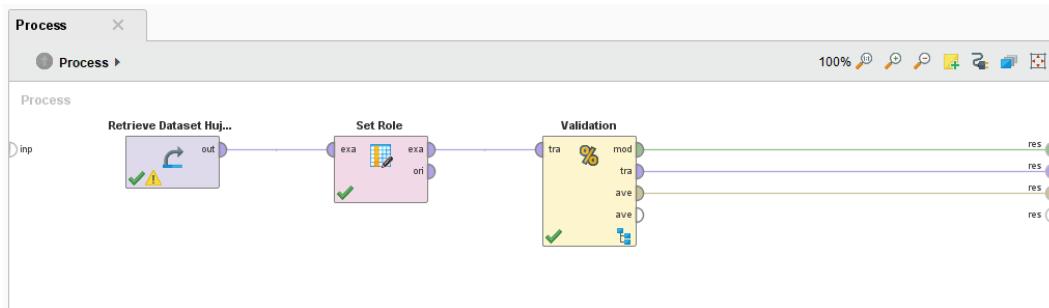


Gambar 5. Hasil Decision Tree

Hasil Pohon Keputusan tersebut menunjukkan alur Curah Hujan berdasarkan 2 faktor yaitu Kelembapan dan Kota.

- Kelembapan tinggi menunjukkan hasil curah hujan ekstrem
- Jika Kelembapan sedang pada kota Bangkalan hasil curah hujan menunjukkan ekstrem
- Jika kelembapan sedang pada kota Bondowoso, Pasuruan, dan Situbondo menunjukkan hasil curah hujan menunjukkan sangat tinggi

Analisis Akurasi



Gambar 6. Design RapidMiner Akurasi

Tidak jauh berbeda dengan Gambar 3. Yang membedakan terletak pada validation dimana kini “tra” dan “ave” validation dihubungkan ke “res” yang terletak di sisi kanan kotak Process.

Hasil Akurasi

accuracy: 95.00% +/- 15.00% (mikro: 91.67%)			
	true Ekstrem	true Sangat Tinggi	class precision
pred. Ekstrem	8	0	100.00%
pred. Sangat Tinggi	1	3	75.00%
class recall	88.89%	100.00%	

Gambar 7. Hasil Akurasi

Pada Gambar 7. terlihat bahwa akurasi yang didapatkan yaitu sebesar 91,67%.

D. KESIMPULAN

Setelah dilakukan analisis pohon keputusan dan akurasi pada tiap jurnal dengan menggunakan aplikasi RapidMiner dengan metode Decision Tree dapat ditarik kesimpulan sebagai berikut :

- Setelah peneliti menginputkan kota, temperature, kecepatan angin, dan kelembapan maka diperoleh hasil nilai akurasi yang tinggi sebesar 91,67%, akurasi ini menunjukkan bahwa Decision Tree memiliki kinerja yang baik dalam melakukan analisis akurasi curah hujan.
- Pengujian Akurasi menunjukkan bahwa semakin banyak data maka tingkat akurasi akan semakin tinggi.

SARAN

- Menambah atribut lain agar mendapatkan akurasi yang lebih tepat dan sempurna agar kedepannya bisa memprediksi bencana banjir yang akan datang bedasarkan curah hujan.

- Melakukan studi kasus ke tiap-tiap lokasi agar lebih real data yang digunakan tanpa adanya penipuan atau pemalsuan data.
- Mengembangkan penelitian dengan menggunakan Algoritma yang lain seperti Naïve Bayes, C5.0, SVM, ataupun Clustering agar lebih baik lagi.

E. DAFTAR PUSTAKA

- [1] D. Fitrianah, W. Gunawan, and A. Puspita Sari, "Studi Komparasi Algoritma Klasifikasi C5.0, SVM dan Naive Bayes dengan Studi Kasus Prediksi Banjir Comparative Study of Classification Algorithm between C5.0, SVM and Naive Bayes with Case Study of Flood Prediction," Februari, vol. 21, no. 1, pp. 1–11, 2022.
- [2] M. B. Arya Darmawan, F. Dewanta, and S. Astuti, "Analisis Perbandingan Algoritma Decision Tree, Random Forest, dan Naïve Bayes untuk Prediksi Banjir di Desa Dayeuhkolot," TELKA - Telekomunikasi Elektronika Komputasi dan Kontrol, vol. 9, no. 1, pp. 52–61, 2023, doi: 10.15575/telka.v9n1.52-61.
- [3] A. Febriansyah Istianto, A. Id Hadiana, and F. Rakhmat Umbara, "Prediksi Curah Hujan Menggunakan Metode Categorical Boosting (Catboost)," JATI (Jurnal Mahasiswa Teknik Informatika), vol. 7, no. 4, pp. 2930–2937, 2024, doi: 10.36040/jati.v7i4.7304.
- [4] H. Sharfina, P. Y. Utami, and I. Fakhruzi, "Prediksi Bencana Banjir Menggunakan Algoritma Deep Learning H2O Berdasarkan Data Curah Hujan," Jurnal Teknik Informatika dan Sistem Informasi, vol. 10, no. 4, pp. 2407–4322, 2023, [Online]. Available: <http://jurnal.mdp.ac.id>
- [5] M. A. Hasanah, S. Soim, and A. S. Handayani, "Implementasi CRISP-DM Model Menggunakan Metode Decision Tree dengan Algoritma CART untuk Prediksi Curah Hujan Berpotensi Banjir," Journal of Applied Informatics and Computing, vol. 5, no. 2, pp. 103–108, 2021, doi: 10.30871/jaic.v5i2.3200.
- [6] S. Dwiasnati and Y. Devianto, "Optimasi Prediksi Bencana Banjir menggunakan Algoritma SVM untuk penentuan Daerah Rawan Bencana Banjir," Prosiding SISFOTEK, pp. 202–207, 2021, [Online]. Available: <http://seminar.iaii.or.id/index.php/SISFOTEK/article/view/283>
- [7] M. Makmun Effendi and A. Siswandi, "Analysis Prediksi Wilayah Rawan Banjir dengan Algoritma K-Means," Journal of Information System Research (JOSH), vol. 5, no. 2, pp. 697–703, 2024, doi: 10.47065/josh.v5i2.4770.
- [8] S. Triyanto, A. Sunyoto, and M. R. Arief, "Analisis Klasifikasi Bencana Banjir Berdasarkan

Curah Hujan Menggunakan Algoritma Naïve Bayes,” JOISIE (Journal Of Information Systems And Informatics Engineering), vol. 5, no. 2, pp. 109–117, 2021, doi: 10.35145/joisie.v5i2.1785.

- [9] I. Risnawati et al., “Klasifikasi Data Mining Untuk Mengestimasi Potensi Curah Hujan Berdampak Banjir Daerah Menggunakan Algoritma C4.5,” Jurnal INSAN Journal of Information System Management Innovation, vol. 3, no. 2, pp. 78–84, 2023, doi: 10.31294/jinsan.v3i2.3050.
- [10] Tasya, Y., & Putri, R. A. (2023). Klasifikasi Tingkat Kerawanan Banjir Wilayah Medan Menggunakan Metode Naive Bayes Dan Algoritma J48. INTECOMS: Journal of Information Technology and Computer Science, 6(2), 657-665.