

Klasifikasi Data Mining Untuk Mengestimasi Potensi Curah Hujan Berdampak Banjir Daerah Menggunakan Algoritma C4.5

Indah Risnawati¹, Syifa Rahma², Fransisca Kusuma³, Najwa Salsabila⁴,
Arjun Nourmansyah⁵, Aria Farhan⁶, Eni Irfiani⁷

^{1,2,3,4,5,6,7} Sistem Informasi, Teknik dan Informatika, Universitas Bina Sarana Informatika
Jl. Kramat Raya No.98 - Jakarta 10450, Indonesia

e-mail: ¹indahrisnawati1998@gmail.com, ²syifa20rahma@gmail.com, ³siscakusuma11@gmail.com,
⁴njwsalsabila06@gmail.com, ⁵arjunnourmansyah12@gmail.com,
⁶quraniaria7@gmail.com, ⁷eni.enf@bsi.ac.id

Abstrak - Pada tahun 2020, di daerah Jawa Tengah dan Jawa Timur mengalami curah hujan yang tinggi. Hal ini menyebabkan seringnya terjadi bencana alam, terutama banjir. Akibat bencana banjir tersebut banyak unit rumah warga yang terendam setinggi 20-40 cm yang menyebabkan kerugian finansial bagi para warganya. Hal ini juga disebabkan oleh kerusakan lingkungan dan meluapnya tiga sungai besar karena ulah manusia itu sendiri. Dalam penelitian ini dibuat untuk menyadarkan masyarakat bahwa bukan hanya curah hujan tinggi yang menyebabkan bencana banjir, tapi juga ulah manusia yang merusak lingkungan. Terjadinya bencana banjir dikarenakan curah hujan tinggi di analisis menggunakan metode klasifikasi pada data mining yaitu algoritma C4.5. Pertama melakukan pengumpulan dataset lalu membuat model pohon keputusan untuk mendapatkan hasil dari klasifikasi curah hujan yang menyebabkan bencana banjir. Hasil akhir prediksi dengan menggunakan model pohon keputusan diperoleh tingkat akurasi sebesar 83,33% berdasarkan perhitungan estimasi menggunakan 60 data.

Kata Kunci: Curah Hujan, Dampak Banjir, Estimasi, Algoritma C4.5

Abstract - In 2020, the Central Java and East Java regions experienced high rainfall. This has led to frequent natural disasters, especially floods. Unfortunately, the flood disaster caused many residential units to be submerged at a height of 20-40 Cm, leading to financial losses for residents. The root cause of the flooding was not only due to high rainfall, but also due to environmental damage and the overflowing of three major rivers caused by human actions. To raise awareness about the issue, a study was conducted to analyze the occurrence of flood disasters using data mining's classification method, specifically the C4.5 algorithm. The study involved collecting datasets and creating a decision tree model to classify rainfall that causes flood disasters. The final result of the decision tree model showed an accuracy rate of 83,33% based on estimation calculations using 60 data.

Keywords: Rainfall, Flood Imct, Estimation, C4.5 Algorithm

PENDAHULUAN

Berdasarkan keterangan BMKG memasukinya tahun 2020, ketinggian curah hujan bisa dibilang memasuki kategori tinggi dan sangat tinggi dikarenakan pada saat itu iklim sedang tidak beraturan dan tidak dapat ditebak juga tidak sesuai prediksi lembaga manapun. Curah hujan adalah tingkat ketinggian air hujan yang jatuh pada daerah yang datar dengan asumsi air hujan tersebut tidak menguap, tidak meresap dan tidak mengalir dalam jangka waktu tertentu di suatu wilayah. Informasi curah hujan berguna dalam banyak bidang terutama informasi yang cepat, akurat dan rinci sangat penting (Muzani et al., 2021). Karena itu penting untuk membagi tingkatan curah hujan. Ada beberapa tingkatan curah hujan, antara lain curah hujan rendah, sedang, tinggi, dan ekstrem. Untuk bisa mengetahui masing-masing kategori hujan tersebut harus dihitung dari luas dan ketinggian milimeter hujan yang terjadi (Yani et al., 2022).

Kondisi cuaca mempunyai dampak yang signifikan terhadap bencana alam seperti banjir termasuk curah hujan yang menjadi salah satu faktornya. Curah hujan ekstrem adalah curah hujan yang bisa menyebabkan



berbagai bencana alam. Meski tidak selalu kejadian kondisi ini bisa dibilang kondisi yang berbahaya untuk wilayah-wilayah yang rawan terkena bencana ketika curah hujan sedang ekstrem (Sitepu et al., 2017).

Banjir merupakan bencana yang sering terjadi di daerah Jawa Tengah dan Jawa Timur. Penyebab banjir biasanya disebabkan oleh meningkatnya debit air di bantaran sungai. Faktor yang mempengaruhi terjadinya banjir adalah faktor alam seperti curah hujan yang ekstrem dan faktor manusia seperti masyarakat yang membuang sampah sembarangan (Arya Darmawan et al., 2023).

Risiko terjadinya banjir bandang akibat curah hujan yang tinggi akibat kondisi cuaca ekstrem dan daya serap air lahan di Indonesia menurun akibat alih fungsi hutan menjadi lahan perkebunan, pertanian, dan perikanan. Bencana banjir selalu menimbulkan kerugian, baik materil maupun manusia. Dampak perubahan ekosistem, baik yang bersifat sementara maupun permanen, dapat terjadi karena pengelolaan banjir yang tidak dilakukan dengan baik (Yuliandoko et al., 2017).

Berdasarkan pada tabel 1 penelitian diketahui penelitian yang kami lakukan untuk mengetahui tingginya curah hujan menggunakan data mining *rapid miner* di beberapa daerah Jawa Tengah dan Jawa Timur yang telah kita ketahui datanya melalui *website* NASA dan BMKG. Dalam penelitian ini menggunakan data mining dengan algoritma C4.5 untuk mengklasifikasi data curah hujan pada daerah Jawa Tengah dan Jawa Timur. Melalui metode ini klasifikasi data untuk menemukan keterangan empat kategori curah hujan, serta temperatur suhu, lalu kecepatan angin dan juga kelembapan. Solusi yang dihasilkan dari metode ini berupa *decision tree* yang bisa menunjukkan rata-rata tingginya curah hujan dalam perbulan dan juga pertahun pada daerah Jawa Tengah dan Jawa Timur.

Penelitian yang dibuat bertujuan agar mengetahui empat keterangan kategori curah hujan dalam perbulan juga pertahun pada daerah Jawa tengah dan Jawa Timur yang memiliki banyak hasil pertanian serta daerah yang sering mengalami bencana karena datangnya hujan dan badai agar para masyarakat juga paham apakah daerah mereka menjadi daerah yang memiliki rata-rata tingginya curah hujan dalam pertahun atau rendahnya curah hujan dalam pertahun. Kebutuhan akan prakiraan cuaca yang akurat juga diharapkan dapat benar-benar membantu masyarakat dalam beraktifitas sehari-hari yang membutuhkan informasi yang lebih akurat dalam hal prakiraan cuaca (Cahyanto et al., 2019).

Tabel 1. Literatur Penelitian

<i>Research Problem (RP)</i>	<i>Literatur Support</i>
RP1 Mengestimasi Curah Hujan di daerah Jawa Tengah dan Jawa Timur	<p>Banjir merupakan bencana yang sering terjadi di daerah Jawa Tengah dan Jawa Timur. Penyebab banjir biasanya disebabkan oleh meningkatnya debit air di bantaran sungai. Faktor yang mempengaruhi terjadinya banjir adalah faktor manusia dan faktor alam.</p> <p>Karena dampak curah hujan terhadap banjir, maka penting untuk memiliki teknologi dan juga informasi untuk mengelola datanya. Data tersebut akan diolah menjadi pengetahuan yang berfungsi sebagai referensi untuk mengidentifikasi pola dan metode tersembunyi dalam pengumpulan data, serta melakukan analisis kelompok untuk mendukung pengambilan keputusan dan pengembangan kebijakan dengan memberikan informasi kompreherensif tentang bagaimana curah hujan berdampak pada semua jenis aktivitas. Metode yang digunakan untuk mengestimasi curah hujan yaitu Data Mining dengan Algoritma C4.5 (Hasanah et al., 2021).</p>
RP2 Menentukan Klasifikasi Jenis Curah Hujan	<p>Untuk mengklasifikasi data curah hujan pada daerah Jawa Tengah dan Jawa Timur. Melalui metode ini klasifikasi data untuk menemukan keterangan empat kategori curah hujan, serta temperatur suhu. Lalu kecepatan angin dan juga kelembapan. Solusi yang dihasilkan dari metode ini berupa <i>decision tree</i> yang bisa menunjukkan rata-rata tingginya curah hujan dalam perbulan dan juga pertahun pada daerah Jawa Tengah dan Jawa Timur. Untuk dapat menentukan klasifikasi curah hujan menggunakan Data Mining dengan Algoritma C4.5.</p> <p>Banjir merupakan salah satu bencana paling berbahaya di dunia. Selain kejadian yang sering</p>

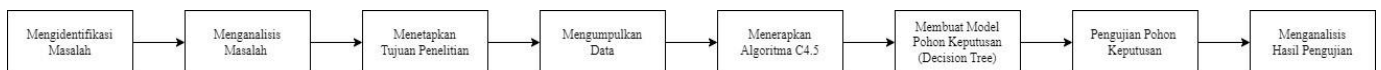
RP3 Pemanfaatan Data Dalam Curah Hujan Yang Tinggi

terjadi dan menimbulkan kerusakan cukup besar seperti di Indonesia, laporan Indeks Risiko Bencana Indonesia (IRBI) tahun 2020 menjelaskan bahwa bencana alam yang paling banyak terjadi adalah banjir yang berjumlah 1.070 kejadian (Triyanto et al., 2021). Data untuk menemukan keterangan empat kategori curah hujan, serta temperatur suhu, lalu kecepatan angin dan juga kelembapan. Pemanfaatan data ini juga bertujuan untuk mengetahui bencana dan bagus atau tidaknya hasil pertanian saat curah hujan tinggi. Banjir telah menyebabkan penurunannya sektor industri dan pertanian yang mempunyai peranan sangat penting dalam meningkatkan kualitas perekonomian masyarakat daerah Jawa Tengah dan Jawa Timur. Sedangkan sektor yang penting bagi peningkatan kualitas perekonomian masyarakat daerah Jawa Tengah dan Jawa Timur adalah sektor industri dan pertanian. Industri akan terhambat akibat banjir karena bergantung pada infrastruktur yang tidak terganggu. Sementara itu, sektor pertanian akan terkena dampak banjir di daerah Jawa Tengah dan Jawa Timur karena kelancaran sektor pertanian bergantung pada kondisi cuaca dan kondisi setempat (Khusaeri et al., 2017).

Sumber: Hasil Penelitian (2023)

METODE PENELITIAN

Dalam penelitian yang kami lakukan, sumber data diperoleh melalui data yang tersedia di *website* NASA dan BMKG di Provinsi Jawa Tengah dan Jawa Timur dengan jumlah 60 dataset dan metode yang digunakan adalah teknik data mining dengan menggunakan algoritma C4.5. Langkah-langkah penyelesaian masalah dapat dilihat pada Gambar 1. Gambar 1 menunjukkan langkah-langkah penyelesaian masalah meliputi mengidentifikasi masalah, menganalisis masalah, menetapkan tujuan penelitian, mengumpulkan data, menerapkan algoritma C4.5, membuat model pohon keputusan, pengujian pohon keputusan, dan menganalisis hasil pengujian.



Sumber: Hasil Penelitian (2023)

Gambar 1. Langkah-Langkah Penyelesaian Masalah

1. Mengidentifikasi Masalah

Pada tahap awal diperlukan untuk mengetahui permasalahan yang dihadapi subjek penelitian dan menentukan ruang lingkup penelitian.

2. Menganalisis Masalah

Pada penelitian yang kami lakukan ditemukan dampak masalah bencana banjir besar dari tingginya curah hujan di daerah Jawa Tengah dan Jawa Timur.

3. Menetapkan Tujuan Penelitian

Menurut dari permasalahan yang ada, diperlukan mengidentifikasi tujuan untuk dapat menyelesaikan masalah tersebut yaitu mengklasifikasi data mining untuk mengestimasi potensi curah hujan yang berdampak banjir di daerah Jawa Tengah dan Jawa Timur sesuai dengan metode yang akan digunakan.

4. Mengumpulkan Data

Pengumpulan data yang dikumpulkan secara observasi dari *website* NASA dan BMKG untuk mengetahui permasalahan yang diidentifikasi sebagai objek penelitian.

5. Menerapkan Algoritma C4.5

Algoritma C4.5 merupakan salah satu teknik klasifikasi pada machine learning yang digunakan pada proses data mining dengan membentuk sebuah pohon keputusan (*decision tree*). Tahapan awal dalam mengimplementasikan algoritma C4.5 dalam mengolah data mining yaitu:

1. Data Selection: Melakukan Pemisahan data yang sudah didapatkan dari sumber terpercaya, kemudian data dikelompokkan berdasarkan atribut beserta nilai didalamnya.
2. Pree-processing: dilakukan nya proses pembersihan data, bertujuan untuk mencegah duplikasi data dan memeriksa data yang tidak konsisten.
3. Transformation: melakukan beberapa perubahan data sesuai dengan format yang terdapat pada data mining.
4. Menghitung jumlah data dan menentukan nilai data disetiap nilai atribut yang ada.
5. Mengklasifikasikan data yang sudah dihitung menjadi beberapa kategori yaitu berpotensi hujan ekstrem atau tidaknya.

6. Membuat Model Pohon Keputusan (*Decision Tree*)

Tahapan pemodelan menggunakan algoritma C4.5 meliputi sebagai berikut (Sekarwati, 2020) :

1. Menyiapkan data training. Pada tahapan ini dataset digunakan sebagai dataset pada proses sebelumnya telah dikelompokkan dengan nilai kelas kelas tertentu.
2. Langkah kedua adalah mengidentifikasi akar pohon. Untuk menentukan akar pohon digunakan nilai informasi gain terbesar dari suatu atribut yang ada. Sebelum mendapatkan nilai informasi gain, terlebih dahulu kita mencari nilai entropi. Entropi adalah parameter untuk mengukur tingkat keragaman dalam dataset dan digunakan untuk menentukan seberapa informatif atribut input untuk menghasilkan atribut.

Rumus dari Entropi ditunjukkan dibawah ini:

$$Entropi(S) = \sum_{i=0}^n p_i * \log_2 p_i \quad (1)$$

Keterangan: S: himpunan kasus

n: jumlah partisi

pi: jumlah proporsi sampel untuk kelas *I*

3. Menentukan nilai informasi gain. Pada tahapan ini menghitung bagaimana informasi gain dari salah satu atribut selection measure yang digunakan untuk memilih test attribute tiap node pada *tree*. Berikut perhitungan informasi gain menggunakan rumus:

$$Gain(S, A) = Entropi(S) - \sum_{i=0}^n |S_i| / |S| * Entropi(S_i)$$

Keterangan: Entropi(*S_i*): entropi untuk kasus yang memiliki nilai *i*

S: himpunan kasus

A: atribut penentu

N: jumlah partisi atribut A

S_i: jumlah kasus nilai *i*

S: jumlah seluruh kasus dalam S

4. Mengulangi langkah kedua.

Fase ini mengulangi langkah kedua hingga semua *record* terpartisi. Fase ini berakhir ketika semua *record* di node N mempunyai kelas yang sama, tidak ada atribut di dalam record yang dapat dibagi lagi, dan tidak ada *record* kosong di cabang pohon.

7. Pengujian Pohon Keputusan

Berdasarkan hasil pemodelan dari algoritma C4.5 diperoleh model berupa pohon keputusan berdasarkan atribut dengan entropi maksimum yang terletak pada posisi paling atas dan nilai yang paling tinggi yaitu curah hujan yang ekstrem.

8. Menganalisis Hasil Pengujian

Pada tahap ini kita membuat bagaimana data training dan pengujian dari seluruh dataset. Menggunakan Validasi silang K-fold Cross diulang sebanyak kali, sehingga kumpulan record dapat dibagi secara acak menjadi k subset independen. Setelah setiap iterasi, ada subset yang tersisa untuk pelatihan. Kami merekomendasikan nilai K dari 10 untuk estimasi pengukuran terbaik. Hal ini menghasilkan 10 iterasi yang dihasilkan oleh pohon keputusan.

HASIL DAN PEMBAHASAN

3.1 Pengumpulan Data dan Penentuan Atribut

Pada tahap ini dibutuhkan data dengan parameter yang sesuai dengan apa yang ingin dihasilkan. Untuk dapat menunjang penelitian ini maka seluruh data yang dibutuhkan pada penelitian ini bersumber dari *website* resmi seperti BPS Jawa Tengah dan Jawa Timur, NASA dan BMKG daerah Jawa Tengah dan Jawa Timur. Data yang telah dikumpulkan menghasilkan lima atribut seperti pada Tabel 2 berikut yaitu dengan atribut keterangan curah hujan sebagai atribut target.

Tabel 2. Dataset Curah Hujan

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

Sumber: Hasil Penelitian (2023)

3.2 Penggunaan Algoritma C4.5 Dalam Pengolahan Data

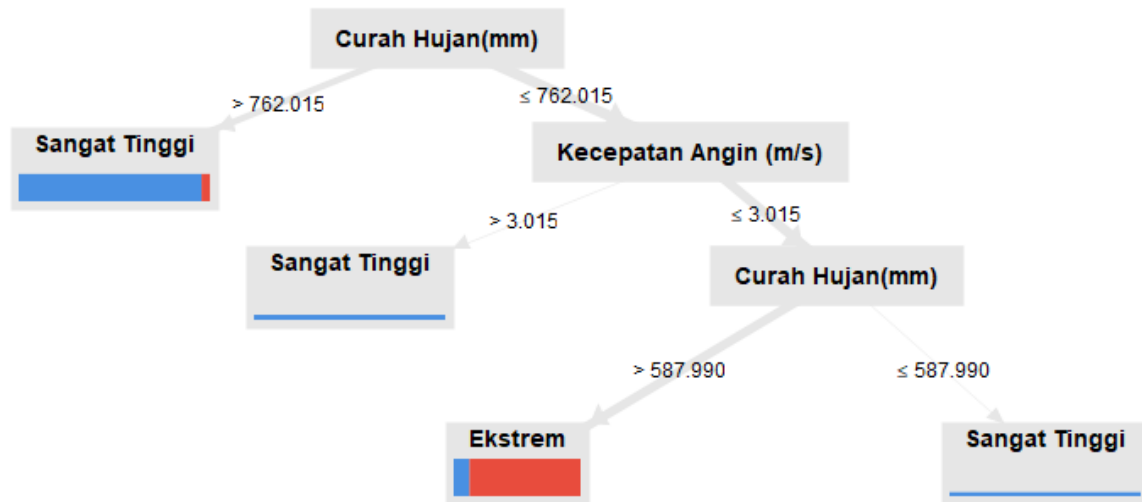
Pada tahap ini, data cuaca yang diperoleh memuat informasi rinci mengenai atribut-atribut yang akan menjadi fokus analisis, yakni Temperatur, Kecepatan Angin, Kelembapan, dan Curah Hujan. Proses pengolahan data dilakukan dengan cermat untuk mengidentifikasi pola dan hubungan antar atribut yang relevan dengan kondisi cuaca.

Tabel 3. Entropy Akar Cuaca

Atribut		Jumlah	Ya	Tidak	Entropy	Gain
Temperatur	Normal	60	41	19	0,90071968	0,376367177
	Panas	49	41	8	0,642064289	
Kecepatan Angin	Hembusan Angin	11	0	11	0	0,079069526
	Sedikit Tenang	38	30	8	0,74248757	
	Hembusan Angin Pelan	18	10	8	0,99107606	
Kelembapan	Tinggi	4	1	3	0,811278124	0,077541279
	Sedang	47	36	11	0,784992089	
Curah Hujan	Ekstrem	13	5	8	0,961236605	0,184136466
	Sangat Tinggi	29	13	16	0,992266639	
		31	28	3	0,458685816	

Sumber: Hasil Penelitian (2023)

3.3 Pembuatan Model Pohon Keputusan (Decision Tree)



Sumber: Hasil Penelitian (2023)

Gambar 2. Hasil *Decision Tree* Bulan Desember 2020

Berdasarkan gambar *decision tree* pada gambar diketahui bahwa curah hujan merupakan atribut yang menempati urutan paling awal yang dimana juga sebagai akar pohon. Dapat dilihat juga bahwa curah hujan yang bernilai lebih dari 762.015 merupakan nilai intensitas curah hujan yang sangat tinggi.

Sedangkan untuk yang bernilai kurang dari sama dengan 762.015 dan memiliki nilai kecepatan angin lebih dari 3.015 merupakan nilai intensitas curah hujan yang sangat tinggi. Sedangkan jika nilai kecepatannya kurang dari sama dengan 3.015, dengan nilai curah hujan sebesar 587.990 termasuk intensitas curah hujan yang ekstrem, sedangkan yang memiliki nilai curah hujan kurang dari sama dengan 587.990 termasuk intensitas curah hujan yang sangat tinggi.

Untuk dapat memahami lebih jelas berikut ringkasan dari penjelasan diatas:

1. Curah hujan (mm) > 762.015 : Sangat tinggi
2. Curah hujan (mm) ≤ 762.015
 - a. Kecepatan angin (m/s) > 3.015 : Sangat tinggi
 - b. Kecepatan angin (m/s) ≤ 3.015
 - Curah hujan (mm) > 587.990 : Ekstrem
 - Curah hujan (mm) ≤ 587.990 : Sangat tinggi

3.4 Hasil Keputusan

Informasi mengenai hasil evaluasi dataset dengan menerapkan metode *cross-validation* menggunakan *10-fold validation*, khususnya pada algoritma C4.5 yaitu dapat dilihat pada Gambar 3 dibawah ini.

Pada Gambar 3, disini kami menggunakan dataset pada bulan Desember di daerah Jawa Tengah dan Jawa Timur. Dalam hal ini, memiliki akurasi data yaitu 83.33%. (-/+) 17.57% dan rata-rata akurasi *micro* yaitu 83.33%.

Dimana untuk prediksi kenyataan sangat tinggi yaitu 25 dan untuk prediksi *true Ekstrem* yaitu 4, kemudian *class precision* memiliki akurasi 86.21% dari data prediksi curah hujan yang sangat tinggi.

Lalu ada prediksi kenyataan sangat tinggi yaitu 6 dan untuk prediksi *true ekstrem* yaitu 25, memiliki *class precision* dari prediksi curah hujan yang eskترم akurasi 80.65%.

Selanjutnya ada, *class recall* memiliki prediksi sangat tinggi ada 80.65% dan yang terakhir *class recall* prediksi *true ekstrem* ada 86.21%.

accuracy: 83.33% +/- 17.57% (micro average: 83.33%)

	true Sangat Tinggi	true Ekstrem	class precision
pred. Sangat Tinggi	25	4	86.21%
pred. Ekstrem	6	25	80.65%
class recall	80.65%	86.21%	

Sumber: Hasil Penelitian (2023)

Gambar 3 Hasil Keputusan Bulan Desember 2020

KESIMPULAN

Penggunaan teknik klasifikasi data mining, khususnya metode algoritma C4.5 membuktikan keberhasilannya dalam mengidentifikasi pola dan hubungan kompleks antara variabel curah hujan dengan potensi banjir di wilayah Surakarta, Jawa Tengah dengan hasil data curah hujan yang tinggi dari 60 dataset di daerah Jawa Tengah dan Jawa Timur memiliki hasil akhir prediksi dengan menggunakan model pohon keputusan diperoleh tingkat akurasi sebesar 83,33%. Keefektifan algoritma C4.5 sudah terbukti efektif dalam memprediksi potensi banjir. Penelitian ini memiliki dampak positif terhadap pengelolaan risiko bencana di wilayah tersebut. Informasi yang dihasilkan dari penelitian ini dapat digunakan oleh otoritas setempat untuk merancang kebijakan dan strategi dalam penanggulangan banjir dimusim curah hujan yang akan datang.

Dengan data diatas didapatkan bahwa curah hujan sangat berperan penting dalam penyebab banjir. Hal ini tidak berarti banjir hanya disebabkan dari curah hujan yang ekstrem. Faktor lain yang dapat memungkinkan menyebabkan banjir dalam penelitian ini adalah temperatur, kecepatan angin, serta kelembapan.

Pada penelitian ini diketahui bahwa bulan Desember di daerah Jawa Tengah dan Jawa Timur yang berpotensi banjir adalah Surakarta, Jawa Tengah. Meskipun penelitian ini memberikan kontribusi signifikan, masih terdapat ruang untuk penelitian lanjutan. Peningkatan dalam pengumpulan data yang lebih banyak dan akurat, peningkatan metode dan pemodelan yang lebih canggih dapat menjadi fokus untuk penelitian masa depan guna meningkatkan akurasi dan ketepatan prediksi curah hujan.

DAFTAR PUSTAKA

- Arya Darmawan, M. B., Dewanta, F., & Astuti, S. (2023). Analisis Perbandingan Algoritma Decision Tree, Random Forest, dan Naïve Bayes untuk Prediksi Banjir di Desa Dayeuhkolot. *TELKA - Telekomunikasi Elektronika Komputasi Dan Kontrol*, 9(1), 52–61. <https://doi.org/10.15575/telka.v9n1.52-61>
- Cahyanto, K. A., Muhamad, F. P. B., & Mulyani, E. (2019). Penerapan Dizcretize By Frequency Dalam Meningkatkan Akurasi Algoritma C4.5 Dalam Memprediksi Cuaca Pada Jalur Pantura Tegal-Pekalongan-Semarang. *JTT (Jurnal Teknologi Terapan)*, 5(2), 78. <https://doi.org/10.31884/jtt.v5i2.195>
- Hasanah, M. A., Soim, S., & Handayani, A. S. (2021). Implementasi CRISP-DM Model Menggunakan Metode Decision Tree dengan Algoritma CART untuk Prediksi Curah Hujan Berpotensi Banjir. *Journal of Applied Informatics and Computing*, 5(2), 103–108. <https://doi.org/10.30871/jaic.v5i2.3200>
- Khusaeri, A., Ilham, S., Nurhasanah, D., Delpidat, D., Anggri, A., Primajaya, A., & Sari, B. N. (2017). Algoritma C4.5 Untuk Pemodelan Daerah Rawan Banjir Studi Kasus Kabupaten Karawang Jawa Barat. *ILKOM Jurnal Ilmiah*, 9(2), 132–136. <https://doi.org/10.33096/ilkom.v9i2.128.132-136>
- Muzani, M. A., Sukri, M. I. A., Fauziah, S. N., Pradnya, W. M., & Suyonto, A. (2021). Algoritma Adaptive Neuro Fuzzy Inference System Untuk Perkiraan Intensitas Curah Hujan. *SISFOTEK: Sistem Informasi Dan Teknologi*, 5(1), 102–106.
- Sekarwati, A. C. R. dan K. A. (2020). Penerapan Data Mining Dengan Menggunakan Algoritma C4.5 Pada Klasifikasi Fasilitas Kesehatan Provinsi Di Indonesia. *Jurnal Ilmiah Komputasi*, 19(1), 27–38. <https://doi.org/10.32409/jikstik.19.1.153>
- Sitepu, F., Selintung, M., & Harianto, T. (2017). Pengaruh Intensitas Curah Hujan dan Kemiringan Lereng Terhadap Erosi Yang Berpotensi Longsor. *Jurnal Penelitian Enjiniring*, 21(1), 23–27. <https://doi.org/10.25042/jpe.052017.03>
- Triyanto, S., Sunyoto, A., & Arief, M. R. (2021). Analisis Klasifikasi Bencana Banjir Berdasarkan Curah Hujan Menggunakan Algoritma Naïve Bayes. *JOISIE (Journal Of Information Systems And Informatics Engineering)*, 5(2), 109–117. <https://doi.org/10.35145/joisie.v5i2.1785>
- Yani, V. I., Aradea, A., & Mubarak, H. (2022). Optimasi Prakiraan Cuaca Menggunakan Metode Ensemble pada Naïve Bayes dan C4.5. *Jurnal Teknik Informatika Dan Sistem Informasi*, 8(3), 607–619. <https://doi.org/10.28932/jutisi.v8i3.5455>
- Yuliandoko, H., Wardhany, V. A., & Pramono, S. H. (2017). Penggunaan Algoritma Decision Making Rules Pada Swod (Safety Water Overflow Detection) Untuk Sistem Pengambilan Keputusan Peringatan. *Prosiding Sentrinov*, 3(ISSN), 2477 – 2097.