

Prediksi Banjir Di Kota Pontianak Menggunakan Metode *Decision Tree C4.5*

¹Mifthahul Fitrah, ^{2*}Asrul Abdullah, ³Istikoma

^{1,2}Teknik Informatika, Universitas Muhammadiyah Pontianak, Pontianak, Indonesia

³Sistem Informasi, Universitas Muhammadiyah Pontianak, Pontianak, Indonesia

Email: mifthahulf17@gmail.com, asrul.abdullah@unmuhpnk.ac.id, istikoma@unmuhpnk.ac.id

ARTICLE INFO

Article History:

Diterima : 20-01-2025
Disetujui : 01-03-2025

Keywords:

Confusion Matrix
Data Mining
Decision Tree C4.5
Flood
Prediction



ABSTRACT

Abstract: – The city of Pontianak is prone to flooding during the rainy season due to its low land surface elevation of 0.1–1.5 meters above sea level and its location along the Kapuas River. Flood prediction in Pontianak needs to be conducted using predictive techniques from data mining. One of the predictive analysis methods is the C4.5 decision tree. This study aims to design a flood prediction model for Pontianak using the C4.5 decision tree method and to determine the obtained accuracy results. The research methodology follows several stages, including problem identification, literature review, data collection, data processing, testing, and conclusion. In this study, three experiments were conducted with different treatments. The model performance evaluation for these three experiments was carried out using a confusion matrix. Based on the evaluation results, the best model for flood prediction was obtained from the second experiment, achieving an accuracy, precision, recall, and F1-score of 98%, 72%, 76%, and 74%, respectively, on the test data

Abstrak: Kota Pontianak merupakan daerah rentan terjadinya bencana banjir saat musim hujan tiba karena memiliki ketinggian permukaan tanah 0,1-1,5 meter diatas permukaan laut dan dilalui oleh aliran Sungai Kapuas. Prediksi banjir di Kota Pontianak perlu dilakukan dengan memanfaatkan teknik prediksi dari data mining. Salah satu metode analisis prediktif adalah decision tree C4.5. Tujuan Penelitian ini adalah merancang model prediksi banjir di Kota Pontianak menggunakan metode *decision tree C4.5* serta mengetahui hasil akurasi yang diperoleh. Metode penelitian pada penelitian ini melalui beberapa tahapan yaitu identifikasi masalah, studi Pustaka, pengumpulan data, pengolahan data, pengujian, dan Kesimpulan. Pada penelitian ini dilakukan tiga percobaan dengan perlakuan yang berbeda. Pengecekan kinerja model dari tiga percobaan tersebut menggunakan confusion matrix. Dari pengecekan tiga percobaan tersebut didapatkan model terbaik untuk memprediksi banjir yaitu pada percobaan kedua dengan hasil akurasi, presisi, recall, dan F1-score sebesar 98%, 72%, 76% dan 74% pada data uji.



<https://doi.org/10.31764/justek.vXIY.ZZZ>



This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license

A. LATAR BELAKANG

Sebagai sumber daya yang sangat penting, data merupakan aset yang sangat berharga bagi berbagai kepentingan terutama di era transformasi dan digitalisasi informasi yang masif saat ini (Suryawijaya, 2023). Data adalah kumpulan objek dengan atribut-atribut tertentu, di mana objek tersebut adalah individu berupa data yang mana setiap data memiliki sejumlah atribut (Kurnia & Kusuma, 2020). Pembahasan dan penelitian mengenai data, salah satunya *data mining* semakin gencar dilakukan oleh semua pihak, dari lingkup kecil individu untuk penelitian, sektor industri, sektor bisnis, hingga sektor pemerintahan, salah satunya bencana alam di Kota Pontianak. Berdasarkan observasi yang dilakukan penulis, kejadian bencana di kota Pontianak, seperti banjir dan kebakaran hutan merupakan bencana alam yang kerap terjadi setiap tahunnya sehingga pemerintah perlu melakukan antisipasi sehingga tidak terulang lagi dimasa yang akan datang (Sarvina, 2018).

Secara umum, bencana alam dibedakan menjadi tiga, yaitu bencana geologi, hidrometeorologi dan biologi. Bencana hidrometeorologi adalah bencana yang disebabkan oleh kondisi hidrologi, meteorologi dan klimatologi seperti banjir, kekeringan, tanah longsor, badai, topan, puting beliung dan kebakaran hutan (Sarvina, 2018). Banjir didefinisikan sebagai massa air yang menggenangi di permukaan tanah yang relatif tinggi dan tidak dapat ditampung sehingga meluap secara alami serta menimbulkan genangan. Salah satu faktor meteorologis yang menyebabkan banjir yaitu hujan (Ka'u, 2021).

Hujan dengan intensitas tinggi dan dipengaruhi dengan keadaan geografis serta dikelilingi oleh daerah yang dilalui oleh aliran sungai seperti di kota Pontianak meningkatkan potensi terjadinya banjir (Alexsandrana et al., 2019). Ditinjau dari aspek geografis, Kota Pontianak hanya memiliki ketinggian permukaan tanah berkisar 0,1-1,5 meter di atas permukaan laut. Kontur topografis yang relatif datar dan dipengaruhi oleh pasang surut Sungai Kapuas menyebabkan kota Pontianak mudah untuk tergenang (Nurhidayati & Hadari Nawawi, 2022). Pencatatan untuk memprediksi terjadinya banjir perlu dilakukan dengan memanfaatkan teknik prediksi dari *data mining* (Nozomi, 2023). Pencatatan ini dilakukan untuk memudahkan pengolahan data dan analisis lebih lanjut mengenai banjir. Pencatatan yang dilakukan harus dilakukan dengan tepat, sehingga data yang dihasilkan dapat digunakan untuk memprediksi terjadinya banjir (Hasanah et al., 2021).

Prediksi merupakan metode untuk memperkirakan peristiwa di masa yang akan datang (Kafil, 2019). Prediksi dilakukan dengan mengumpulkan data-data lampau dengan variabel-variabel yang berlaku untuk memperkirakan masa depan dan digunakan dengan menggunakan

pendekatan tertentu untuk diolah, yang hasilnya dapat digunakan sebagai prediksi kejadian-kejadian di masa depan (Hidayat Nasution et al., 2024). Salah satu metode yang digunakan dalam analisis prediktif adalah *Decision Tree C4.5* (Lestari, 2020). Metode ini merupakan salah satu metode yang digunakan untuk melakukan klasifikasi atau pengelompokkan dan bersifat prediktif. Dasar metode *Decision Tree C4.5* adalah pembentukan pohon keputusan (*decision tree*) (Hidayah & Rozi, 2021). Metode *Decision Tree C4.5* digunakan pada penelitian ini karena metode ini dikenal sangat baik menangani dataset awal yang digunakan pada penelitian ini dengan melakukan transformasi data-data banjir dari tahun 2018-2022 kedalam dataset yang memiliki bentuk parameter kategorik. *Decision Tree C4.5* dikenal memiliki kemampuan dan akurasi yang baik dalam memodelkan data kategorik, baik data biner maupun non-biner (Iman et al., 2024). Pada proses transformasi penulis menambahkan parameter atau atribut yaitu curah hujan dan tinggi muka air sungai Kapuas berdasarkan observasi dan dokumentasi dalam proses pengumpulan dan transformasi data di objek penelitian. Berdasarkan hasil preprocessing dataset awal oleh penulis dengan metode *Decision Tree C4.5* dan pustaka *Scikit-Learn (sklearn-tree)* dari bahasa pemrograman *Python* dengan *DecisionTreeClassifier* dan kriteria entropi, diperoleh hasil perhitungan skor akurasi prediksi di angka 98%. Hasil perhitungan tersebut menunjukkan skor prediksi memiliki akurasi sangat baik dengan parameter yang digunakan.

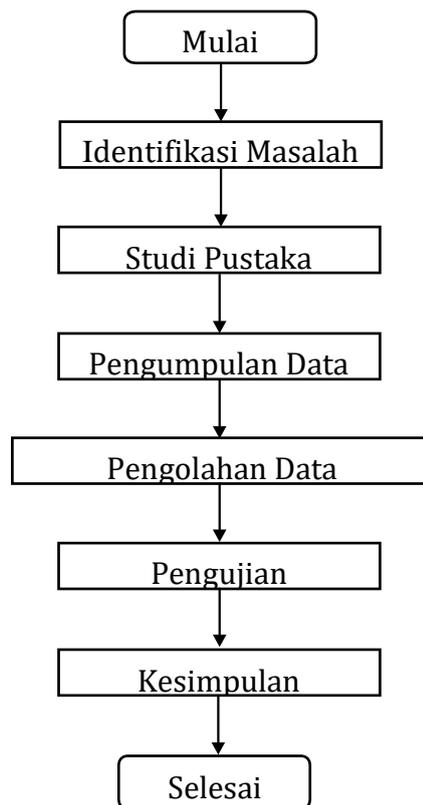
Beberapa studi telah menunjukkan bahwa *Decision Tree C4.5* memiliki akurasi yang baik dalam melakukan prediksi. Misalnya studi (Iman et al., 2024) menunjukkan penerapan *Decision Tree C4.5* untuk memprediksi potensi kunjungan wisata di Kabupaten Pati, Indonesia diperoleh akurasi sebesar 96,42%. Studi lain (Program et al., 2018) memprediksi penyakit diabetes menggunakan *Decision Tree C4.5* diperoleh akurasi sebesar 76,32%. Studi lain (Petra Valentino & Siska Narulita, 2023) menunjukkan bahwa *Decision Tree C4.5* memiliki akurasi yang sangat baik sebesar 88,35% dalam memprediksi penyakit jantung. Studi (Aji Primajaya et al., n.d.) menunjukkan *Decision Tree C4.5* memiliki akurasi sebesar 89,7859% dengan *percentage split 80% data training* dan *20% data testing* dalam memprediksi potensi kebakaran hutan. Studi (Anwar et al., 2022) menunjukkan bahwa *Decision Tree C4.5* memiliki akurasi yang baik sebesar 92% dalam memprediksi kelulusan mahasiswa. Studi (Harman, 2018) menunjukkan bahwa *Decision Tree C4.5* dapat menginformasikan indikator *website* yang paling baik sehingga dapat memberikan manfaat dalam pengembangan *website*. Hal ini jelas memberi manfaat dalam meningkatkan kualitas *website*. Beberapa penelitian terkait tersebut dengan berbagai objek prediksi yang berbeda tidak menggunakan metrik yang sama dengan jumlah data bervariasi dan

cenderung terlalu sedikit sehingga data awal penulis pada penelitian memiliki originalitasnya sendiri dan dapat diharapkan memberikan hasil prediksi yang lebih baik dari penelitian-penelitian sebelumnya.

Gap analisis penelitian ini dengan penelitian metode *Decision Tree C4.5* lainnya terletak pada penerapan metode *Decision Tree C4.5* dalam konteks prediksi banjir di Kota Pontianak yang belum banyak diteliti sebelumnya. Penelitian ini bertujuan untuk merancang model prediksi banjir di Kota Pontianak menggunakan metode *decision tree C4.5* dan mengetahui hasil akurasi dari model yang dibangun. Hasil penelitian ini diharapkan dapat menjadi pengetahuan dalam melihat pola-pola penyebab terjadinya banjir sehingga dapat membantu merumuskan strategi dan kebijakan yang lebih efektif untuk melakukan antisipasi saat akan terjadinya banjir di Kota Pontianak.

B. METODE PENELITIAN

Metode Penelitian yang digunakan dalam penelitian ini dapat dilihat pada diagram alir pada gambar 1 (Abrar, 2023).



Gambar 1 Diagram Alir Metode Penelitian

1. Identifikasi Masalah

Langkah yang dilakukan adalah mengidentifikasi permasalahan Kota Pontianak. Masalah di Kota Pontianak yakni tingginya frekuensinya terjadinya

hujan dan dilalui sungai kapuas menjadi permasalahan di kota pontianak yang rentan terjadinya banjir sehingga perlunya sistem agar dapat memprediksi terjadinya banjir.

2. Studi Pustaka

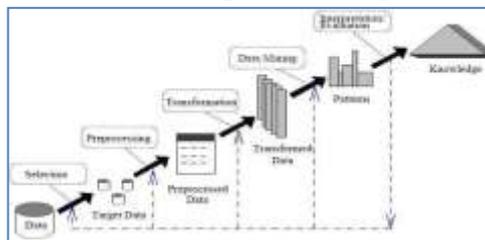
Pada tahap ini akan dilakukan pencarian mengenai penelitian sebelumnya. Studi pustaka dilakukan untuk mengumpulkan bahan referensi. Studi pustaka yang digunakan dapat berupa jurnal ilmiah terdahulu, buku, dan bahan-bahan lain yang dapat digunakan untuk mendukung penyelesaian penelitian.

3. Pengumpulan Data

Pada tahapan ini, tujuannya adalah untuk mendapatkan *dataset* yang sesuai dengan penelitian. Penelitian ini akan menggunakan *dataset* cuaca berbasis dampak dari tahun 2018-2022 yang bersumber dari BMKG Maritim Kota Pontianak

4. Pengolahan Data

Pengolahan data merupakan proses *data mining* yang bertujuan untuk mengeksplorasi data dan menemukan pola atau informasi yang berguna dari data tersebut. Pada penelitian ini, pengolahan data dilakukan dengan menggunakan metode *decision tree C4.5*. Tahapan pengolahan data dengan *Knowledge Discovery in Databases* dapat dilihat pada diagram gambar 2.



Gambar 2 Proses KDD (Karsito & Monika Sari, 2018)

Knowledge Discovery in Database (KDD) merupakan sekumpulan proses untuk menemukan pengetahuan atau informasi yang bermanfaat dari data. Proses *Knowledge Discovery in Databases* (KDD) secara umum sebagai berikut.

1) *Data Selection*

Pemilihan data dari banyaknya data yang tersedia perlu dilakukan sebelum tahap penggalian informasi dalam *knowledge data discovery* dimulai. Data tersebut akan digunakan untuk proses *data mining*, disimpan dalam suatu berkas, terpisah dari basis data. Sebelum proses *data mining* dapat dilaksanakan, perlu dilakukan proses pembersihan data. Proses ini antara lain membuang duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data, seperti kesalahan ketik (tipografi).

2) *Preprocessing*

Pada tahap ini data yang akan diolah bertujuan untuk menghindarkan dari data yang mengganggu (*noise*) atau data yang tidak konsisten.

3) *Transformation*

Coding merupakan proses transformasi pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses data mining. Proses *coding* dalam KDD sangat tergantung pada jenis atau pola informasi yang akan dicari dalam basis data.

4) *Data Mining*

Data mining adalah tahapan mencari pola atau informasi dalam data terpilih dengan menggunakan metode tertentu. Algoritma dalam *data mining* sangat bervariasi. Pemilihan algoritma yang tepat sangat bergantung pada tujuan dan proses KDD secara keseluruhan.

5) *Interpretation/Evaluation*

Pola informasi yang dihasilkan dari proses *data mining* perlu ditampilkan dalam bentuk yang mudah dimengerti.

5. Pengujian

Pengujian pada penelitian ini menggunakan *confusion matrix* yang dilakukan pada beberapa percobaan sebagai berikut:

- 1) Percobaan 1: Pada percobaan pertama ini hanya dilakukan *preprocessing data, splitting data* dan pemodelan *decision tree C4.5* lalu pengecekan akurasi
- 2) Percobaan 2: Pada percobaan ini dilakukan *preprocessing data, splitting data, balancing data* dan pemodelan *decision tree C4.5* lalu pengecekan akurasi
- 3) Percobaan 3: Pada percobaan kedua ini dilakukan *preprocessing, penanganan outlier, feature scaling, splitting data* dan pemodelan *decision tree C4.5* lalu pengecekan akurasi

6. Kesimpulan

Kesimpulan didapat berdasarkan dari analisis hasil pengujian yang telah dilakukan dan juga dari hasil analisis selama pembangunan aplikasi. Kesimpulan berupa hasil dari penelitian sesuai yang tergambar pada diagram alir penelitian (Karsito & Monika Sari, 2018).

C. HASIL DAN PEMBAHASAN

Penelitian ini dilakukan melalui serangkaian tahapan yang terstruktur, mencakup pengumpulan data, pemodelan menggunakan Decision Tree C4.5, evaluasi model, hingga interpretasi hasil.

1. Pengumpulan Data

Data yang digunakan pada penelitian ini merupakan data yang bersumber dari BMKG Maritim Kota Pontianak tahun 2018-2022. Tabel 1 menyajikan beberapa data pada dataset yang terbentuk dan digunakan dalam penelitian ini.

Tabel 1. Dataset

No	Tinggi Air Sungai Kapuas (cm)	Curah Hujan	Label
1	120	Hujan Ringan	Tidak Berpotensi banjir
2	160	Hujan Lebat	Berpotensi Banjir
3	110	Hujan Ringan	Tidak Berpotensi Banjir
4	150	Hujan Sedang	Berpotensi Banjir

2. Pemodelan Decision Tree C4.5

Setelah dataset siap, pemodelan dilakukan menggunakan algoritma Decision Tree C4.5. Algoritma ini dipilih karena kemampuannya dalam menangani fitur kategori dan kontinu, serta kemampuannya dalam menangani data yang tidak seimbang. Selain itu, Decision Tree C4.5 juga dapat memberikan pemahaman yang jelas dan interpretable mengenai bagaimana prediksi dibuat, yang sangat berguna dalam konteks ini. Proses pembuatan model Decision Tree C4.5 melibatkan pembuatan pohon keputusan berdasarkan data. Setiap node dalam pohon mewakili fitur dalam data, dan setiap cabang mewakili keputusan yang dibuat berdasarkan fitur tersebut. Algoritma ini akan terus membagi data berdasarkan fitur hingga mencapai kondisi berhenti, yang biasanya adalah ketika semua data dalam node berada dalam kelas yang sama atau ketika mencapai kedalaman maksimum. Untuk memahami lebih baik bagaimana algoritma Decision Tree C4.5 bekerja, kita dapat merujuk ke rumus entropi yang digunakan dalam algoritma ini:

Menghitung entropy dengan rumus sebagai berikut:

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

Keterangan :

- S = Data sample yang digunakan untuk training
- n = Jumlah partisi S
- pi = Proporsi dari S_i terhadap S
- A = Atribut

Menghitung gain dengan rumus sebagai berikut:

$$Gain(S, A) = Entropy(s) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i) \quad (2)$$

Keterangan :

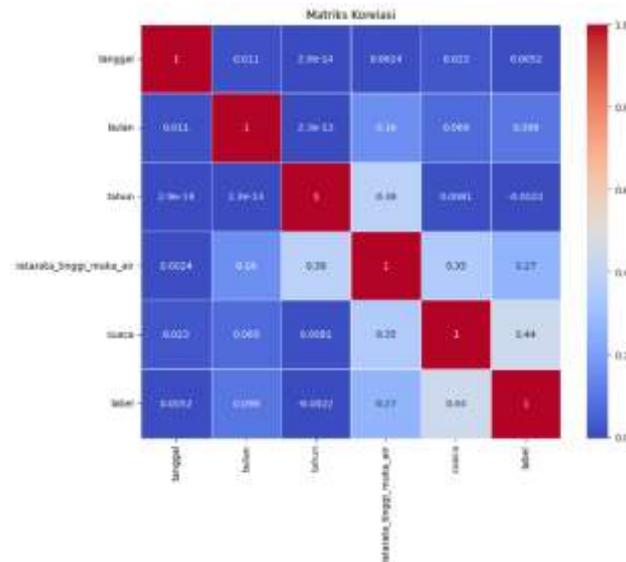
- S = Data sample yang digunakan untuk training
- n = Jumlah atribut A
- $|S_i|$ = Jumlah kasus dalam partisi ke i

$|S|$ = Jumlah kasus dalam S

A = Atribut

3. Korelasi Antar Fitur

Correlation matrix adalah metrik digunakan untuk mengecek korelasi antara setiap fitur pada dataset. Gambar korelasi antar fitur disajikan pada gambar 2.



Gambar 3 Mariks Korelasi

4. Hasil Pengujian

Pengujian dilakukan sebanyak tiga percobaan dengan perlakuan yang berbeda-beda untuk mendapatkan hasil akurasi yang terbaik.

- Percobaan 1: Pada Percobaan pertama hanya dilakukan preprocessing data, splitting data, pemodelan decision tree C4.5 lalu pengecekan akurasi. Hasil akurasi pada data latih dan data uji sangat baik yaitu 99%. Namun recall dan F1-score yang rendah pada data uji yaitu 53% dan 69%.
- Percobaan 2: Pada Percobaan kedua dilakukan preprocessing data, splitting data, balancing data dengan teknik SMOTE dan pemodelan decision tree C4.5 lalu pengecekan akurasi. Hasil akurasi pada data latih dan data uji sangat baik yaitu 97% dan 98%. Hasil presisi, recall dan F1-score pada data uji juga baik yaitu 72%, 76% dan 74%.
- Percobaan 3: Pada Percobaan ketiga ini dilakukan preprocessing data, splitting data, menghilangkan data outlier, standarisasi, balancing data dengan teknik SMOTE dan pemodelan decision tree C4.5 lalu pengecekan akurasi. Hasil akurasi pada data latih dan data uji sangat baik yaitu 99% dan 98%. Hasil presisi dan F1-score pada data uji kurang baik yaitu 50% dan 65%.

Dari hasil beberapa percobaan diatas, dapat dilihat pada percobaan 1, 2 dan 3 hasil akurasi data latih dan data uji yang diperoleh tidak jauh berbeda. Dapat diambil kesimpulan bahwa pada percobaan 2 model dapat memprediksi banjir lebih baik dibanding percobaan 1 dan 3 dengan perolehan hasil yang baik pada akurasi, presisi, recall dan F1-score. Hasil pengujian setiap percobaan dapat dilihat pada Tabel 2.

Tabel 1. Hasil Evaluasi Model

No	Pengujian	Data Latih (%)				Data Uji (%)			
		Akura si	Presi si	Reca ll	F1- Scor e	Akura si	Presi si	Reca ll	F1- Scor e
1	Percobaan 1	99	94	73	82	99	100	53	69
2	Percobaan 2	97	98	97	97	98	72	76	74
3	Percobaan 3	99	97	100	99	98	50	92	65

D. SIMPULAN DAN SARAN

Penelitian yang dilakukan dengan melalui beberapa tahapan mulai dari identifikasi masalah, studi pustaka, pengumpulan data, pengolahan data, pengujian, dan kesimpulan. Hasil pengujian dari tiga percobaan yang telah dilakukan, Terdapat hasil terbaik yaitu pada percobaan kedua dengan menangani data imbalance dengan teknik SMOTE dengan akurasi pada data latih 97% dan data uji 98%. Hasil confusion matrix pada recall 76%, presisi 72% dan F1-score 74%.

Penelitian selanjutnya dapat menggunakan *tunning hyperparameter* untuk meningkatkan kinerja model secara optimal serta dapat menerapkan teknik evaluasi model menggunakan *cross validation*.

UCAPAN TERIMA KASIH

Terima kasih kepada semua pihak yang terkait sehingga penelitian ini bisa berjalan dengan lancar terutama Badan Meteorologi Klimatologi dan Geofisika Maritim Kota Pontianak.

REFERENSI

- Abrar, I. N., Abdullah, A., & Sucipto, S. (2023). Liver Disease Classification Using the Elbow Method to Determine Optimal K in the K-Nearest Neighbor (K-NN) Algorithm. *Jurnal Sisfokom (Sistem Informasi Dan Komputer)*, 12(2), 218–228. <https://doi.org/10.32736/sisfokom.v12i2.1643>
- Aji Primajaya, B., Khusaeri, A., & Khusaeri. (n.d.). *JEPIN (Jurnal Edukasi dan Penelitian Informatika)*.
- Alexsandrana, Sitorus, S. H., & Midyanti, D. M. (2019). Menggunakan Logika Fuzzy Tsukamoto Berbasis Website (Studi Kasus Kota Pontianak). *Jurnal Komputer Dan Aplikasi*, 07(02), 61–70.
- Anwar, F. F., Jaya, A. I., & Abu, M. (2022). Prediksi Kelulusan Mahasiswa Tepat Waktu Menggunakan Metode Decision Tree dengan Penerapan Algoritma C4.5. *JURNAL ILMIAH MATEMATIKA DAN TERAPAN*, 19(1), 19–28.

- <https://doi.org/10.22487/2540766x.2022.v19.i1.15880>
- Harman, R. (2018). Penerapan Algoritma C.45 Untuk Memprediksi Indikator Website Yang Baik. *Jurnal Ilmiah Informatika*, 06(02), 54–59. <http://ejournal.upbatam.ac.id/index.php/jif> Penerapan
- 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>
- Hidayah, A. N. Z., & Rozi, A. F. (2021). Penerapan Data Mining Dalam Menentukan Kinerja Karyawan Terbaik Dengan Menggunakan Metode Algoritma C4.5 (Studi Kasus: Universitas Mercu Buana Yogyakarta). *Jurnal Information System & Artificial Intelligence*, 1(2), 117–127.
- Hidayat Nasution, S., Irsa Syahputri, N., & Aprilia, R. (2024). Penerapan Metode Least Square Dalam Prediksi Jumlah Produksi Padi Di Kabupaten Padang Lawas. *Jurnal Sains Dan Teknologi*, 7(2), 128–137. <http://journal.ummat.ac.id/index.php/justek>
- Iman, H. K. N., Latifah, N., Supriyono, S., & Nugraha, F. (2024). Pemanfaatan Metode Decision Tree dengan Algoritma C4.5 Untuk Prediksi Potensi Kunjungan Wisatawan. *Jurnal Sistem Komputer Dan Informatika (JSON)*, 5(3), 611. <https://doi.org/10.30865/json.v5i3.6684>
- Ka'u, A. A. (2021). ANALISIS TINGKAT KERAWANAN BANJIR DI KECAMATAN SANGTOMBOLANG KABUPATEN BOLAANG MONGONDOW. *Jurnal Spasial*, Volume 8(3), 102–114.
- Kafil, M. (2019). Penerapan Metode K-Nearest Neighbors Untuk Prediksi Penjualan Berbasis Web Pada Boutiq Dealove Bondowoso. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 3(2), 59–66. <https://doi.org/10.36040/jati.v3i2.860>
- Karsito, & Monika Sari, W. (2018). Prediksi Potensi Penjualan Produk Delifrance Dengan Metode Naive Bayes Di Pt. Pangan Lestari. *Jurnal Teknologi Pelita Bangsa*, 9(1), 67–78. <https://jurnal.pelitabangsa.ac.id/index.php/sigma/article/view/465>
- Kurnia, Y., & Kusuma, E. D. (2020). Pendugaan Kredit Macet Pada Koperasi Simpan Pinjam Flamboyan Binaan PPSW Jakarta Dengan Menggunakan Komparasi Algoritma Naïve Bayes dan C4.5. *Jurnal Algor*, 1, 66–74.
- Lestari, N. A. (2020). Penerapan data mining menggunakan metode decision tree C4.5 untuk prediksi tingkat kelulusan mahasiswa (Studi Kasus : STMIK WIT). *Jurnal Web Informatika Teknologi*, 5(2). <https://ejournal-wit.ac.id/index.php/J-WIT/article/view/50>
- Nozomi, I. (2023). Penerapan Data Mining Untuk Peringatan Dini Banjir Menggunakan Metode Klastering K-Means (Studi Kasus Kota Padang). *Jurnal Sains Informatika Terapan*, 2(2), 39–44. <https://doi.org/10.62357/jsit.v2i2.165>
- Nurhidayati, E., & Hadari Nawawi, J. H. (2022). *Konsep Blue-Green Infrastructure (Bgi) Melalui Permeable Pavements Pedestrian Dan Kolam Retensi Untuk Mitigasi Genangan Banjir Di Kota Pontianak*. 19(1), 1829–9172.
- Petra Valentino, & Siska Narulita. (2023). Performansi Algoritma Decision Tree (C4.5) untuk Prediksi Penyakit Jantung. *Jurnal Cakrawala Informasi*, 3(2), 18–24. <https://doi.org/10.54066/jci.v3i2.349>
- Program, N., Medis, S. R., Kesehatan, I., & Kesehatan, I.-I. (2018). *Implementasi Algoritma Decision Tree C4.5 untuk Prediksi Penyakit Diabetes* *Jurnal INOHIM* (Vol. 6, Issue 1). www.kaggle.com/uciml/pima-indians-diabetes-database
- Sarvina, Y. (2018). Aspek Hidrometeorologi dalam Menumbuhkan Budaya Sadar Bencana di Indonesia [Hydrometeorological Aspect in Raising Disaster Awareness

in Indonesia]. *Jurnal Dialog Penanggulangan Bencana*, 9(1), 9–17.

Suryawijaya, T. W. E. (2023). Memperkuat Keamanan Data melalui Teknologi Blockchain: Mengeksplorasi Implementasi Sukses dalam Transformasi Digital di Indonesia. *Jurnal Studi Kebijakan Publik*, 2(1), 55–68.
<https://doi.org/10.21787/jskp.2.2023.55-68>