

IMPLEMENTASI ALGORITMA RANDOM FOREST UNTUK MEMPREDIKSI KETINGGIAN AIR SEBAGAI SISTEM PERINGATAN DINI BANJIR TERINTEGRASI WEB

Ali Nurohman Fadilah¹, Fitri Yanti²

Teknik Informatika, Fakultas Ilmu Komputer, Universitas Pamulang

¹alinf3011@gmail.com, ²dosen00848@unpam.ac.id

ABSTRACT

Floods are among the most frequent disasters in regions with high rainfall, including Bogor Regency, creating a need for a system capable of predicting rising water levels early enough for preventive measures to be taken. This study applies the Random Forest Regression algorithm to predict the water level (TMA) one hour ahead using historical hydrological data from the Batu Beulah Water Level Monitoring Station, Bogor, comprising water level, rainfall, and discharge derived from a rating curve, and integrates the resulting model into a Next.js-based web application. Unlike locally run machine learning programs accessible only to technical users, the developed system separates the prediction service (FastAPI) from the web interface so that the output can be used by two actors, namely the station administrator and the general public. The dataset consists of 7,817 hourly records collected throughout 2024 with 31 input features, split chronologically into training, validation, and testing sets with a 70:15:15 ratio to avoid data leakage. The model was trained with 300 decision trees and evaluated using Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and the coefficient of determination (R^2), and compared against a persistence baseline. Testing results show an MAE of 0.0957 m, an RMSE of 0.2137 m, and an R^2 of 0.7918, with RMSE and R^2 outperforming the baseline. Current water level, calculated discharge, and the three-hour rolling mean of water level were the most influential features. Black Box and White Box testing confirmed that the system's core functions, including login, data management, prediction execution, and the display of Safe/Alert/Danger warning status, operate as intended.

Keywords: *Random Forest, Water Level Prediction, Flood Early Warning, Web Integration, Next.js*

ABSTRAK

Banjir merupakan salah satu bencana yang sering terjadi di wilayah dengan curah hujan tinggi, termasuk di Kabupaten Bogor, sehingga dibutuhkan sistem yang mampu memprediksi kenaikan tinggi muka air (TMA) secara lebih awal agar tindakan antisipasi dapat dilakukan lebih cepat. Penelitian ini bertujuan menerapkan algoritma Random Forest Regression untuk memprediksi TMA satu jam ke depan

menggunakan data hidrologi historis dari Pos Duga Air Batu Beulah, Bogor, yang mencakup data tinggi muka air, curah hujan, dan debit air hasil perhitungan rating curve, kemudian mengintegrasikan model tersebut ke dalam aplikasi web berbasis Next.js. Berbeda dari implementasi machine learning lokal yang hanya dapat diakses oleh pengguna teknis, sistem yang dibangun memisahkan layanan prediksi (FastAPI) dari antarmuka web sehingga hasil prediksi dapat digunakan oleh dua aktor, yaitu admin/pengelola pos duga air dan pengguna umum. Dataset yang digunakan terdiri atas 7.817 baris data per jam sepanjang tahun 2024 dengan 31 fitur masukan, dibagi secara kronologis menjadi data latih, validasi, dan uji dengan proporsi 70:15:15 untuk mencegah kebocoran data (data leakage). Model dilatih dengan 300 pohon keputusan dan dievaluasi menggunakan Mean Absolute Error (MAE), Root Mean Square Error (RMSE), dan koefisien determinasi (R^2), serta dibandingkan dengan baseline persistence. Hasil pengujian menunjukkan model memperoleh MAE sebesar 0,0957 m, RMSE sebesar 0,2137 m, dan R^2 sebesar 0,7918 pada data uji, dengan RMSE dan R^2 yang lebih baik dibandingkan baseline. Fitur TMA saat ini, debit terhitung, dan rata-rata bergerak TMA tiga jam menjadi fitur paling berpengaruh terhadap hasil prediksi. Pengujian Black Box dan White Box menunjukkan seluruh fungsi utama sistem, termasuk proses login, pengelolaan data, prediksi, dan penyajian status peringatan Aman/Siaga/Bahaya, berjalan sesuai kebutuhan.

Kata kunci: Random Forest, Prediksi Ketinggian Air, Peringatan Dini Banjir, Integrasi Web, Next.js

A. Pendahuluan

Banjir masih menjadi salah satu tantangan terbesar di Indonesia, terutama di wilayah dengan curah hujan tinggi dan karakteristik geografis yang kompleks. Kabupaten Bogor kerap mengalami peningkatan tinggi muka air (TMA) pada musim hujan sehingga memicu banjir lokal di beberapa kecamatan. Banjir tidak hanya berdampak pada kerusakan fisik, tetapi juga mengganggu aktivitas ekonomi, mobilitas, hingga kesehatan masyarakat, sebagaimana ditemukan

pada kajian banjir di Samarinda Utara yang menunjukkan bahwa genangan lebih dari 50 cm dapat melumpuhkan aktivitas ekonomi dan merusak aset rumah tangga secara signifikan (Anwar et al., 2022). Kompleksitas faktor penyebab banjir, mulai dari curah hujan ekstrem hingga perubahan penggunaan lahan, menuntut hadirnya pendekatan prediktif yang lebih akurat agar potensi banjir dapat diantisipasi sejak dini.

Perkembangan machine learning dalam dekade terakhir memberikan peluang besar untuk membangun sistem prediksi yang lebih presisi dalam konteks mitigasi banjir. Model berbasis algoritma ensemble, seperti Random Forest, terbukti unggul karena mampu menangkap pola non-linier, mengatasi noise, dan bekerja baik pada data multivariabel seperti curah hujan, TMA, debit air, maupun variabel temporal. Beberapa penelitian terdahulu memperkuat pemilihan Random Forest untuk kasus prediksi hidrometeorologi. Sitepu et al. (2025) memanfaatkan delapan variabel cuaca historis BMKG dengan 300 pohon keputusan dan memperoleh akurasi 97,97% dalam memprediksi banjir di Kota Manado. Sasoko et al. (2024) mengombinasikan K-Means dan Random Forest pada data TMA pintu air Jakarta dan memperoleh akurasi 95% serta F1-score 90%, dengan variabel waktu dan histori TMA terbukti berpengaruh besar. Aufa et al. (2025) mengintegrasikan Random Forest dan XGBoost melalui soft voting untuk memetakan potensi banjir di Pangkalpinang dan memperoleh akurasi 0,98. Nurkhaliza dan Wijayanto (2022) menunjukkan

Random Forest lebih unggul dibandingkan Support Vector Machine dalam memprediksi indeks mitigasi bencana, sedangkan Indaryono (2024) membuktikan Random Forest unggul atas Naïve Bayes dalam klasifikasi curah hujan. Natzir (2023) bahkan melaporkan performa sempurna Random Forest pada seluruh metrik evaluasi untuk data curah hujan historis 1901–2018.

Meskipun demikian, penelitian-penelitian tersebut umumnya berfokus pada pemetaan risiko atau analisis dampak pasca-banjir menggunakan data cuaca regional, bukan pada prediksi langsung kenaikan TMA dari data pengamatan harian pos duga air lokal. Sebagian besar implementasi model prediksi juga masih berhenti pada program machine learning yang dijalankan secara lokal sehingga hanya dapat digunakan oleh pengguna teknis, padahal hasil prediksi idealnya dapat diakses oleh lebih dari satu pihak, yaitu pengelola pos duga air yang mengelola data dan masyarakat umum yang membutuhkan informasi peringatan dini. Celah inilah yang mendasari perlunya integrasi model machine learning dengan aplikasi web agar proses prediksi dapat disajikan secara

interaktif dan digunakan oleh lebih banyak pelaku sistem.

Berdasarkan kondisi tersebut, penelitian ini bertujuan (1) mengolah data debit air, curah hujan, dan TMA dari Pos Duga Air Batu Beulah Bogor menjadi data masukan yang siap digunakan dalam pelatihan model prediksi; (2) menerapkan algoritma Random Forest untuk memprediksi ketinggian air satu jam ke depan dan menentukan status peringatan dini Aman, Siaga, atau Bahaya; dan (3) merancang integrasi model tersebut dengan aplikasi web berbasis Next.js melalui layanan prediksi (API) agar hasil prediksi dapat dimanfaatkan oleh admin/pengelola pos duga air maupun masyarakat umum. Kebaruan penelitian ini terletak pada penggunaan data operasional harian satu pos duga air lokal secara lengkap (TMA, curah hujan, debit hasil rating curve, dan informasi waktu) yang dipadukan dengan arsitektur layanan machine learning terpisah (FastAPI) dan antarmuka web modern (Next.js), sehingga hasil prediksi tidak berhenti sebagai eksperimen lokal, melainkan dapat langsung dioperasikan oleh dua aktor sistem.

B. Metode Penelitian

Penelitian ini menggunakan pendekatan pengembangan sistem prediksi terintegrasi web yang menggabungkan alur kerja machine learning dengan pengembangan aplikasi web, meliputi pengumpulan data, praproses data, rekayasa fitur, pelatihan dan evaluasi model Random Forest, penyimpanan model, pembuatan layanan prediksi, serta pengembangan antarmuka web.

Pengumpulan dan Sumber Data

Data yang digunakan merupakan data sekunder berupa TMA per jam dan curah hujan harian dari Pos Duga Air Batu Beulah, Bogor, sepanjang tahun 2024. Data TMA diekstraksi dari dokumen PDF pencatatan menggunakan pustaka pdfplumber, dengan deteksi nama bulan, label jam, serta pemetaan posisi nilai terhadap kolom tanggal; nilai berformat koma desimal dikonversi ke format titik. Data curah hujan diproses dengan pendekatan serupa, dengan tanda hubung pada tabel sumber diperlakukan sebagai curah hujan 0 mm setelah dilakukan pemeriksaan struktur dokumen. Setiap hasil ekstraksi divalidasi melalui pemeriksaan jumlah data per

bulan, distribusi 24 jam, nilai di luar rentang wajar, duplikasi timestamp, dan konsistensi total curah hujan bulanan, karena kesalahan posisi kolom pada PDF sumber dapat menghasilkan data yang valid secara sintaksis tetapi salah secara temporal.

Rekayasa Fitur

Debit air dihitung dari nilai TMA menggunakan persamaan rating curve pada Persamaan (1), karena data debit hasil pengukuran langsung tidak tersedia secara kontinu.

$$Q = 20,268 \times (H + 0,15)^{2,157} \quad (1)$$

dengan Q adalah debit dalam m³/s dan H adalah TMA dalam meter. Debit hasil perhitungan digunakan sebagai fitur turunan (engineered feature) dan tidak diklaim sebagai debit hasil pengukuran lapangan.

Data TMA per jam dan curah hujan harian digabungkan berdasarkan waktu pengamatan, disusun secara kronologis, dibersihkan dari duplikasi timestamp, kemudian dilengkapi dengan fitur turunan berupa komponen waktu (jam, hari, bulan, akhir pekan), komponen siklis (sin/cos jam dan bulan), nilai lag TMA (1–24 jam), perubahan dan rata-rata bergerak TMA, lag dan rata-rata bergerak curah

hujan, serta lag dan rata-rata bergerak debit, sebagaimana dirangkum pada Tabel 1. Baris dengan histori tidak lengkap pada 168 baris awal/akhir dihapus sebelum pembagian dataset.

Tabel 1. Kelompok Fitur pada Dataset

Kelompok	Fitur
Fitur dasar	tma, rainfall_mm, discharge_m3s
Fitur waktu	hour, day_of_week, day_of_month, month, is_weekend
Fitur siklis	hour_sin, hour_cos, month_sin, month_cos
Lag TMA	tma_lag_1h, 2h, 3h, 6h, 12h, 24h
Perubahan & rolling TMA	tma_diff_1h, tma_diff_3h, rolling mean 3h/6h/12h, rolling std 6h
Curah hujan	rainfall_lag_1d, rainfall_lag_2d, rolling sum 3d dan 7d
Debit	discharge_lag_1h, discharge_lag_3h, discharge_rolling_mean_6h

Dataset akhir terdiri atas 7.817 baris dengan 31 fitur masukan dan satu target, yaitu target_tma_next_1h (TMA satu jam berikutnya). Pembagian data dilakukan secara kronologis tanpa pengacakan dengan proporsi 70:15:15 (Tabel 2) untuk mencegah data leakage, yaitu kondisi ketika informasi masa depan ikut memengaruhi proses pelatihan sehingga evaluasi menjadi terlalu optimistis.

Tabel 2. Pembagian Dataset

Bagian	Jumlah	Persentase	Rentang Waktu
Training	5.471	70%	8 Jan 2024 20.00 – 15 Sep 2024 23.00
Validation	1.173	15%	16 Sep 2024 01.00 – 8 Nov 2024 01.00
Testing	1.173	15%	8 Nov 2024 02.00 – 31 Des 2024 22.00
Total	7.817	100%	Setelah 168 baris awal/akhir tidak lengkap dihapus

Algoritma Random Forest Regression

Random Forest membangun sejumlah B pohon keputusan dari sampel bootstrap dan subset fitur acak, kemudian menggabungkan hasil seluruh pohon untuk regresi dengan cara merata-ratakan

keluarannya, sebagaimana
 Persamaan (2).

$$\hat{y}_{RF}(x) = (1/B) \sum_{b=1}^B T_b(x) \quad (2)$$

Pemilihan pemisahan pada setiap pohon regresi dilakukan dengan memilih kandidat split yang menghasilkan penurunan Mean Squared Error (MSE) terbesar, dihitung menggunakan Persamaan (3)–(5), dengan $\bar{y}_S$ sebagai rata-rata target pada simpul S, S_L dan S_R sebagai simpul kiri/kanan hasil pemisahan, serta n_L dan n_R sebagai jumlah data pada masing-masing simpul turunan.

$$MSE(S) = (1/n) \sum_{i=1}^n (y_i - \bar{y}_S)^2 \quad (3)$$

$$MSE_{split} = (n_L/n) \cdot MSE(S_L) + (n_R/n) \cdot MSE(S_R) \quad (4)$$

$$Gain = MSE(S) - MSE_{split} \quad (5)$$

Model diimplementasikan menggunakan RandomForestRegressor dari scikit-learn dengan konfigurasi hyperparameter pada Tabel 3, dipilih melalui evaluasi pada data validasi untuk menyeimbangkan akurasi dan risiko overfitting.

Tabel 3. Konfigurasi Hyperparameter Model

Parameter	Nilai	Keterangan
<code>n_estimators</code>	300	Jumlah pohon keputusan
<code>max_depth</code>	25	Kedalaman maksimum tiap pohon
<code>min_samples_split</code>	5	Minimum sampel untuk membagi simpul
<code>min_samples_leaf</code>	2	Minimum sampel pada daun
<code>max_features</code>	sqrt	Jumlah fitur kandidat tiap split
<code>random_state</code>	42	Menjaga reproduktibilitas
<code>n_jobs</code>	-1	Menggunakan seluruh inti prosesor

Performa model dievaluasi menggunakan Mean Absolute Error (MAE), Root Mean Square Error (RMSE), dan koefisien determinasi (R^2) pada Persamaan (6)–(8), serta dibandingkan dengan baseline persistence (asumsi TMA tidak berubah dari jam sebelumnya) untuk memastikan peningkatan performa bukan semata-mata berasal dari sifat autokorelasi TMA.

$$MAE = (1/n) \sum_{i=1}^n |y_i - \hat{y}_i| \quad (6)$$

$$RMSE = \sqrt{[(1/n) \sum_{i=1}^n (y_i - \hat{y}_i)^2]} \quad (7)$$

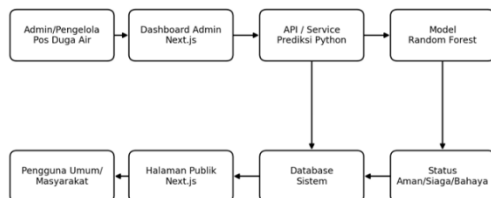
$$R^2 = 1 - [\sum_{i=1}^n (y_i - \hat{y}_i)^2 / \sum_{i=1}^n (y_i - \bar{y})^2] \quad (8)$$

Arsitektur dan Integrasi Sistem

Sistem dirancang dengan arsitektur berlapis untuk memisahkan tanggung jawab antarmuka, proses bisnis, layanan machine learning, dan penyimpanan data, sebagaimana ditunjukkan pada Gambar 1. Lapisan presentasi dibangun menggunakan Next.js 15, React 19, TypeScript, Tailwind CSS, dan Recharts untuk visualisasi; lapisan backend web

menggunakan Next.js API Routes dengan validasi Zod dan autentikasi NextAuth.js; layanan prediksi machine learning dibangun terpisah menggunakan FastAPI, scikit-learn, pandas, dan joblib; sedangkan penyimpanan data menggunakan Neon PostgreSQL melalui Drizzle ORM. Pemisahan layanan prediksi dari aplikasi web memungkinkan proses pemuatan model dan prediksi dikelola secara independen tanpa memicu pelatihan ulang pada setiap permintaan.

Arsitektur Sistem Prediksi Ketinggian Air



Gambar 1. Arsitektur Sistem Prediksi Ketinggian Air

Model yang telah dinyatakan layak disimpan dalam berkas joblib beserta metadata (daftar fitur, hyperparameter, rentang data pelatihan, dan hasil evaluasi) dalam format JSON agar versi model dapat ditelusuri. Layanan FastAPI menyediakan endpoint `/predict` yang menerima data hidrologi terbaru, membentuk fitur,

menjalankan model, menghitung status peringatan berdasarkan nilai ambang (threshold) Aman/Siaga/Bahaya, dan mengembalikan response JSON kepada backend Next.js. Sistem melibatkan dua aktor: admin/pengelola pos duga air yang mengelola data hidrologi, menjalankan prediksi, dan memantau dashboard; serta pengguna umum yang mengakses halaman publik untuk melihat status peringatan dan grafik kondisi sungai tanpa perlu melakukan login.

Skema Pengujian Sistem

Pengujian sistem dilakukan dengan dua pendekatan. Black Box Testing digunakan untuk menguji fungsi sistem dari sisi masukan dan keluaran, mencakup autentikasi, pengelolaan data hidrologi (CRUD), impor data, eksekusi prediksi, pengelolaan informasi peringatan, pengaturan threshold, serta halaman publik. White Box Testing digunakan untuk menguji struktur logika pada fungsi kritis, yaitu perhitungan rating curve, alur endpoint prediksi, dan skema validasi input, dengan jalur logika direpresentasikan melalui control

flow graph dan dihitung kompleksitas siklomatisnya menggunakan Persamaan (9), dengan E sebagai jumlah edge, N sebagai jumlah node, dan P sebagai jumlah komponen graf yang terhubung.

$$V(G) = E - N + 2P \quad (9)$$

C.Hasil Penelitian dan Pembahasan

Hasil Ekstraksi dan Karakteristik Data

Proses ekstraksi menghasilkan lebih dari 7.985 rekaman TMA per jam sepanjang Januari–Desember 2024 (cakupan sekitar 92% dari total jam dalam setahun) dan 366 rekaman curah hujan harian dengan total curah hujan 3.617,5 mm pada sekitar 155 hari hujan. Data yang tidak tersedia tidak diisi (imputasi) secara acak; fitur lag dan rolling hanya dibentuk pada baris dengan histori yang cukup, sedangkan baris tidak lengkap dihapus sebelum pembagian dataset. Rentang nilai pada dataset akhir adalah TMA 0,18–6,30 m, curah hujan 0–112,5 mm, dan debit terhitung 1,85–1.129,88 m³/s.

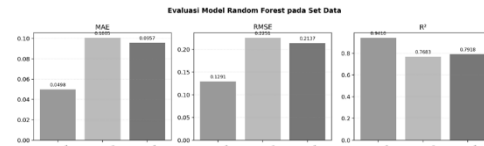
Evaluasi Kinerja Model

Model Random Forest Regression yang dilatih dengan konfigurasi pada Tabel 3 menghasilkan nilai

MAE, RMSE, dan R² pada data training, validation, dan testing sebagaimana disajikan pada Tabel 4 dan Gambar 2.

Tabel 4. Hasil Evaluasi Random Forest Regressor

Metrik	Training	Validation	Testing
MAE (m)	0,0498	0,1005	0,0957
RMSE (m)	0,1291	0,2251	0,2137
R ²	0,9410	0,7683	0,7918



Gambar 2. Grafik Evaluasi Model pada Data Training, Validation, dan Testing

Nilai MAE testing sebesar 0,0957 m menunjukkan rata-rata selisih absolut sekitar 9,57 cm antara TMA hasil prediksi dan TMA aktual satu jam berikutnya. RMSE testing sebesar 0,2137 m lebih besar daripada MAE karena metrik ini memberikan penalti lebih tinggi terhadap kesalahan berukuran besar, sedangkan R² sebesar 0,7918 menunjukkan model mampu menjelaskan sekitar 79,18% variasi TMA pada data testing. Selisih nilai metrik antara data training dan validation/testing mengindikasikan adanya penurunan generalisasi yang wajar pada data baru, namun masih berada pada rentang yang dapat diterima untuk kebutuhan peringatan dini operasional.

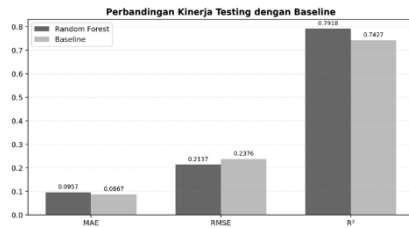
Perbandingan dengan Baseline Persistence

Untuk memastikan peningkatan performa model tidak semata-mata

berasal dari sifat autokorelasi TMA, hasil Random Forest dibandingkan dengan baseline persistence pada data testing (Tabel 5 dan Gambar 3).

Tabel 5. Perbandingan Model dengan Baseline pada Data Testing

Metrik	Random Forest	Baseline Persistence
MAE (m)	0,0957	0,0867
RMSE (m)	0,2137	0,2376
R ²	0,7918	0,7427



Gambar 3. Perbandingan Kinerja Random Forest dan Baseline Persistence

Baseline persistence menghasilkan MAE yang sedikit lebih rendah karena pada periode TMA relatif stabil, menyalin nilai TMA terkini sudah cukup akurat. Namun, Random Forest memiliki RMSE yang lebih rendah dan R² yang lebih tinggi, yang menunjukkan model lebih unggul dalam mengurangi kesalahan besar dan menjelaskan variasi data ketika terjadi perubahan TMA yang signifikan, misalnya pada periode hujan deras. Kondisi ini relevan dengan tujuan peringatan dini, karena kesalahan besar pada momen kenaikan TMA yang cepat berisiko lebih tinggi dibandingkan kesalahan kecil pada kondisi stabil.

Evaluasi time series cross-validation pada lima fold data historis menghasilkan rata-rata MAE sebesar $0,1658 \pm 0,0406$ m, dengan

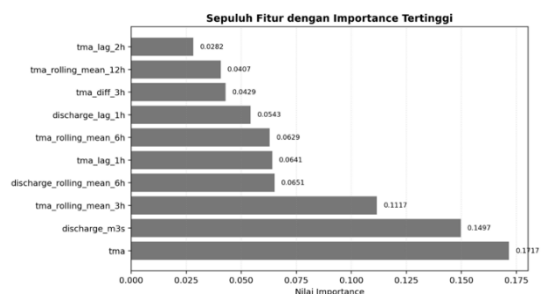
peningkatan error pada beberapa fold yang mengindikasikan adanya pergeseran distribusi data terkait variasi musiman dan kejadian hujan ekstrem. Temuan ini menegaskan pentingnya pemantauan performa model secara berkala ketika data operasional tahun berikutnya tersedia.

Feature Importance

Analisis feature importance pada Tabel 6 dan Gambar 4 menunjukkan bahwa TMA saat ini (tma), debit terhitung (discharge_m3s), dan rata-rata bergerak TMA tiga jam (tma_rolling_mean_3h) merupakan tiga fitur paling berpengaruh terhadap hasil prediksi model.

Tabel 6. Sepuluh Fitur dengan Importance Tertinggi

Peringkat	Fitur	Importance
1	tma	0,1717
2	discharge_m3s	0,1497
3	tma_rolling_mean_3h	0,1117
4	discharge_rolling_mean_6h	0,0651
5	tma_lag_1h	0,0641
6	tma_rolling_mean_6h	0,0629
7	discharge_lag_1h	0,0543
8	tma_diff_3h	0,0429
9	tma_rolling_mean_12h	0,0407
10	tma_lag_2h	0,0282



Gambar 4. Feature Importance Model Random Forest

Dominasi fitur TMA, debit, dan turunannya menunjukkan bahwa model mengandalkan kondisi hidrologi terkini dan tren jangka pendek untuk memprediksi TMA satu jam

berikutnya. Curah hujan memiliki kontribusi relatif lebih kecil karena pengaruhnya terhadap TMA dapat mengalami jeda waktu (time lag) dan tidak selalu terjadi pada hari pengamatan yang sama. Hasil feature importance ini bersifat deskriptif terhadap pemanfaatan fitur oleh pohon-pohon dalam model dan tidak dapat diinterpretasikan sebagai hubungan sebab-akibat.

Integrasi dan Pengujian Sistem Web

Model yang telah dievaluasi diintegrasikan dengan aplikasi web melalui layanan FastAPI dan disajikan melalui dua kelompok antarmuka: halaman publik (beranda, status, grafik, dan informasi) yang dapat diakses tanpa login, serta halaman admin (dashboard, kelola data hidrologi, prediksi, kelola informasi peringatan, pengaturan threshold, dan grafik) yang dilindungi autentikasi NextAuth.js. Pengujian Black Box terhadap tujuh kelompok fungsi utama, meliputi login, CRUD data hidrologi, impor data, eksekusi prediksi, kelola informasi peringatan, pengaturan threshold, dan halaman publik, menunjukkan seluruh 33 skenario uji berstatus “Sesuai” dengan hasil yang diharapkan, termasuk

penanganan kondisi gagal seperti data kosong, layanan prediksi tidak aktif, dan input tidak valid.

Pengujian White Box pada fungsi rating curve menunjukkan seluruh delapan kasus uji, mencakup nilai normal, nilai batas bawah, dan sifat monotonik fungsi, dinyatakan lulus. Pengujian jalur logika pada endpoint POST /api/predictions menghasilkan control flow graph dengan 11 node dan 12 edge pada satu komponen terhubung, sehingga kompleksitas siklomatis $V(G) = 12 - 11 + 2(1) = 3$, yang berarti terdapat tiga jalur independen minimum, yaitu jalur data tidak tersedia, jalur layanan prediksi gagal, dan jalur prediksi berhasil, yang seluruhnya telah diuji dan disimpan pada basis path pengujian. Validasi input pada skema pendaftaran data, autentikasi, informasi peringatan, dan threshold turut diuji pada dua lapisan, yaitu validasi sisi klien dan validasi pada API, sehingga permintaan yang dimodifikasi secara langsung tetap tertolak. Secara keseluruhan, hasil pengujian fungsional dan struktural menunjukkan bahwa model prediksi dan aplikasi web yang dikembangkan telah berjalan sesuai kebutuhan sistem yang dirancang.

D. Kesimpulan

Penelitian ini berhasil mengolah data hidrologi Pos Duga Air Batu Beulah Bogor tahun 2024 menjadi dataset prediksi sebanyak 7.817 baris dengan 31 fitur masukan, serta menerapkan algoritma Random Forest Regression untuk memprediksi TMA satu jam ke depan dengan MAE sebesar 0,0957 m, RMSE sebesar 0,2137 m, dan R^2 sebesar 0,7918 pada data testing. Perbandingan dengan baseline persistence menunjukkan bahwa Random Forest memiliki RMSE yang lebih rendah dan R^2 yang lebih tinggi, sehingga lebih unggul dalam menjelaskan variasi data serta mengurangi kesalahan besar ketika terjadi perubahan TMA yang signifikan. Fitur TMA saat ini, debit terhitung, dan rata-rata bergerak TMA tiga jam terbukti menjadi fitur paling berpengaruh terhadap hasil prediksi.

Model yang dihasilkan berhasil diintegrasikan dengan aplikasi web berbasis Next.js melalui layanan FastAPI dan basis data Neon PostgreSQL, sehingga hasil prediksi tidak lagi berhenti sebagai program machine learning lokal, melainkan dapat digunakan oleh admin/pengelola pos duga air untuk

mengelola data dan menjalankan prediksi, serta oleh pengguna umum untuk melihat status peringatan dini Aman, Siaga, atau Bahaya melalui halaman publik. Pengujian Black Box dan White Box menunjukkan seluruh fungsi utama sistem serta jalur logika kritis pada endpoint prediksi telah berjalan sesuai kebutuhan, dengan kompleksitas siklomatis sebesar tiga jalur independen.

Penelitian selanjutnya disarankan untuk mengintegrasikan sensor atau perangkat Internet of Things agar pembaruan data TMA dan curah hujan dapat berlangsung secara real-time, menambahkan variabel hidrometeorologi lain seperti kelembapan tanah atau intensitas hujan per jam, membandingkan Random Forest dengan algoritma lain seperti XGBoost atau model deret waktu, serta menerapkan pemantauan dan pelatihan ulang model secara berkala untuk menjaga performa prediksi ketika pola musim atau kondisi sungai berubah. Validasi nilai ambang status peringatan bersama pihak berwenang dan pengelola pos duga air juga diperlukan sebelum sistem diterapkan secara operasional di lapangan.

Kesimpulan akhir yang diperoleh dalam penelitian dan saran perbaikan yang dianggap perlu ataupun penelitian lanjutan yang relevan.

DAFTAR PUSTAKA

- Aji Bayu Prasetyo, D., & Alfa Susetyo, Y. (2022). Implementasi Information Schema Database pada PostgreSQL untuk Pembuatan Tabel Informasi dengan Menggunakan Python di PT XYZ. *Jurnal Ilmiah Matrik*, 9(3). <http://jurnal.mdp.ac.id>
- Aldi Ramadani. (2025). Sistem Informasi Cuti Kepegawaian pada Rumah Sakit Umum Daerah Kabupaten Batu Bara. *Modem: Jurnal Informatika dan Sains Teknologi*, 3(1), 67–75. <https://doi.org/10.62951/modem.v3i1.350>
- Annisa Tri Hidayati, Widyantoro, A. E., & Ramadhani, H. J. (2023). Perancangan Sistem Informasi Wirausaha Mahasiswa (SIWIRMA) Berbasis Web dengan Unified Modelling Language (UML). *Jurnal Penelitian Rumpun Ilmu Teknik*, 2(4), 86–107. <https://doi.org/10.55606/juprit.v2i4.2906>
- Anwar, Y., Ningrum, M. V. R., & Setyasih, I. (2022). Dampak Bencana Banjir terhadap Ekonomi Masyarakat di Kecamatan Samarinda Utara, Kota Samarinda. *JPG (Jurnal Pendidikan Geografi)*, 9(1). <https://doi.org/10.20527/jpg.v9i1.12457>
- Aufa, R. A., Altiarika, E., Pramudyantoro, A., Pratama, Y. B., & Wahyuzi, Z. (2025). Pengenalan Pola Kasus Potensi Banjir di Pangkalpinang dengan Algoritma Random Forest dan XGBoost Menggunakan Google Earth Engine. *Journal of Information Technology and Society (JITS)*, 3(1), 1–6.
- Baehaqi, A., Basit, M. S., Indrajit, R. E., & Kurniawan, R. D. (2023). Front End Learning Management System Development Using the Next.js Framework. *Jurnal Teknik Informatika (Jutif)*, 4(4), 899–911. <https://doi.org/10.52436/1.jutif.2023.4.4.1273>
- Dahri, N., & Setya Hadi, H. (n.d.). Perancangan Sistem Informasi Arsip Digital Berbasis Web pada Kejaksaan Negeri Padang. *Jurnal Manajemen Teknologi Informatika*, 1(3), 125.
- Efendi, I., & Annisa, S. (2022). Perancangan Aplikasi Penjualan Baju Berbasis Desktop Menggunakan Java NetBeans. *Jurnal Ilmiah Amik Mitra Gama*, 6(1). <http://ojsamik.amikmitragama.ac.id>
- Fahmi Idris, J., Prastyo, Y. D., Hardiansyah, R., Magnus, A., Noventona, F., Haikal, F., Gusdevi, H., & Naseer, M. (2025). Pengujian Fungsional dan Struktural Aplikasi Pengajuan Cuti dengan Metode Black Box dan White Box. *Jurnal Teknik Informatika*, 7.
- Fatman, Y., Khoirun Nafisah, N., & Bendoro Jembar Pambudi, P. (2023). Implementasi Payment Gateway dengan Menggunakan Midtrans pada Website UMKM Geberco. *Jurnal KomtekInfo*, 10(2), 64–

72.
<https://doi.org/10.35134/komtekinfo.v10i2.364>
- Gusdevi, H., Kuswayati, S., Iqbal, M., Fikri, M., Bakar, A., Novianti, N., & Ramadan, R. (2022). Pengujian White-Box pada Aplikasi Debt Manager Berbasis Android. *Jurnal Teknik Informatika STT Bandung*, 4.
- Intan Mayla Faiza, Gunawan, & Wresti Andriani. (2022). Machine Learning: Prosperity of Rainfall, Water Discharge, and Flood with Web Application in Deli Serdang. *Jurnal Penelitian Komunikasi dan Opini Publik*, 22(2).
<https://doi.org/10.33299/jpkop.22.2.1752>
- Jurnal, H., Fendi Prasetyo, S., & Feri Efendi, T. (2024). Implementasi Sistem Prediksi Curah Hujan dengan Penerapan Jaringan Syaraf Tiruan Berbasis Website. *Jurnal Riset Teknik Komputer (JURTIKOM)*, 1(2).
<https://doi.org/10.69714/gn6nym06>
- Lina Halifah, J. S. (2025). Perancangan Sistem Informasi Peminjaman dan Pengembalian Buku Berbasis Website. Universitas Pamulang.
- Mamok Andri Saubekti, G. L. (2024). Pembuatan Desain UI/UX dengan Metode Prototyping pada Aplikasi Layanan Pengadilan Negeri Bale Bandung Menggunakan Figma. Universitas Pamulang.
- Muna, F., Khotimah, T., & Jazuli, A. (2023). Sistem Administrasi Perpustakaan Desa Kaliputu Berbasis Web. *Jurnal Mahasiswa Teknik Informatika*, 7(2).
- Natzir, S. M. (2023). Perbandingan Kinerja Model Pembelajaran Mesin dalam Prediksi Banjir Menggunakan KNN, Naive Bayes, dan Random Forest. *Jurnal Ilmiah Hoaq*, 14(1), 59–64.
<https://doi.org/10.52972/hoaq.vol14no1.p59-64>
- Nurkhaliza, A. A., & Wijayanto, A. W. (2022). Perbandingan Algoritma Klasifikasi Support Vector Machine dan Random Forest pada Prediksi Status Indeks Mitigasi dan Kesiapsiagaan Bencana (IMKB) Satuan Kerja BPS di Indonesia Tahun 2020. *Jurnal Informatika Universitas Pamulang*, 7(1), 2622–4615.
<https://doi.org/10.32493/informatika.v7i1.16117>
- Parameswari, D. C., & Warsono, H. (2023). Efektivitas Aplikasi Sistem Peringatan Dini dalam Upaya Mitigasi Bencana Banjir di Kota Jakarta. *Jurnal Ilmu Administrasi Publik, Universitas Diponegoro. Penelitian Ilmu Komputer*, J., & Wisnu Widagdo, B. (2025). Penerapan Algoritma Genetika untuk Penjadwalan Mengajar Guru (Studi Kasus: SD Negeri Kunciran 7). *Jurnal Ilmu Komputer*, 3(3).
<https://mypublikasi.com/>
- Prakoso Indaryono, N. A. (2024). Analisa Perbandingan Algoritma Random Forest dan Naïve Bayes untuk Klasifikasi Curah Hujan Berdasarkan Iklim di Indonesia. *JUPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, 9(1), 158–167.
<https://doi.org/10.29100/jupi.v9i1.4421>
- Puan Maharani. (2025). Pengembangan Website PT.

- Rantangin Digital Indonesia Menggunakan Framework Next.js dan Tailwind CSS. Repeater: Publikasi Teknik Informatika dan Jaringan, 3(1), 129–137.
<https://doi.org/10.62951/repeater.v3i1.355>
- Rahman Akbar, M., Zurfadly, A., & Apriani, M. (2025). Perancangan Database Elite Hotel Tembilahan Menggunakan ERD (Entity Relationship Diagram). Jurnal Sistem Informasi (TEKNOFILE), 3, 105–117.
- Salman, H. A., Kalakech, A., & Steiti, A. (2024). Random Forest Algorithm Overview. Babylonian Journal of Machine Learning, 2024, 69–79.
<https://doi.org/10.58496/bjml/2024/007>
- Sasoko, W. H., Pujiharto, E. W., Haris, R., & Kania, A. Y. (2024). Prediksi Banjir di DKI Jakarta dengan Menggunakan Algoritma K-Means dan Random Forest. Journal of Informatics and Computer Media (JICOM).
<https://ejurnalunsam.id/index.php/jicom/>
- Sitepu, J. P., Mongi, C. E., Widian, S. A., & Ketaren, E. (2025). Prediksi Banjir di Kota Manado Menggunakan Metode Random Forest. Jurnal Ilmiah STMIK Time.
<http://ejournal.stmik-time.ac.id>
- Timuhingide, A. A., Rianto, I., & Komansilan, T. (2025). Digitalisasi Desa Tandengan Berbasis Website. Journal of Education Method and Technology, 3.
- Wayahdi, M. R., & Ruziq, F. (2023). Pemodelan Sistem Penerimaan Anggota Baru dengan Unified Modeling Language (UML) (Studi Kasus: Programmer Association of Battuta). Jurnal Minfo Polgan, 12(1), 1514–1521.
<https://doi.org/10.33395/jmp.v12i1.12870>
- Zulfa, A. A., & Arifudin, O. (2025). Peran Sistem Informasi Akademik Berbasis Web dalam Upaya Meningkatkan Efektivitas dan Efisiensi Pengelolaan Akademik di Perguruan Tinggi. Jurnal Tahsinia, 6(1).