

LAPORAN PENELITIAN

■ Floodzy: Prediksi Risiko Banjir Berbasis Machine Learning

Analisis *Feature Importance* Menggunakan XGBoost dan
SHAP

Oleh:

MATT YUDHA
NIM: 123456789

PROGRAM STUDI TEKNIK INFORMATIKA

UNIVERSITAS TEKNOLOGI KREATIF

2025

Daftar Isi

Daftar Isi	2
Abstrak	3
BAB I: PENDAHULUAN	3
BAB II: METODOLOGI PENELITIAN	3
BAB III: HASIL DAN ANALISIS	4
BAB IV: KESIMPULAN	5

Abstrak

Sistem peringatan dini banjir merupakan komponen krusial dalam mitigasi bencana. Proyek Floodzy mengembangkan model **machine learning** menggunakan algoritma XGBoost untuk memprediksi ketinggian air sungai berdasarkan data sensorik historis. Laporan ini berfokus pada analisis **feature importance** untuk mengidentifikasi variabel paling berpengaruh. Dengan menggunakan metode **built-in** dari XGBoost dan SHAP (**SHapley Additive exPlanations**), penelitian ini menunjukkan bahwa fitur seperti *riverlevelcm*, *rainmm*, dan *soilmoisture* secara konsisten menjadi prediktor utama. Hasil analisis ini tidak hanya memvalidasi pilihan model tetapi juga memberikan wawasan mendalam tentang faktor-faktor hidrologis yang mendorong terjadinya banjir, yang dapat digunakan untuk menyempurnakan strategi penempatan sensor dan kebijakan penanggulangan bencana.

BAB I: PENDAHULUAN

Banjir adalah salah satu bencana alam yang paling sering terjadi dan merusak di Indonesia. Kerugian yang ditimbulkan tidak hanya bersifat material, tetapi juga dapat mengancam keselamatan jiwa. Oleh karena itu, pengembangan sistem peringatan dini (EWS - **Early Warning System**) yang andal menjadi sebuah keharusan.

Proyek Floodzy bertujuan untuk membangun EWS berbasis data dengan memanfaatkan kemajuan dalam bidang **machine learning**. Dengan menganalisis data historis dari berbagai sensor (curah hujan, ketinggian air, kelembapan tanah), model diharapkan dapat memberikan prediksi akurat beberapa jam sebelum kejadian. Keberhasilan model ini sangat bergantung pada pemilihan fitur yang relevan. Laporan ini akan membahas secara mendalam proses analisis dan interpretasi fitur-fitur yang digunakan dalam model prediksi Floodzy.

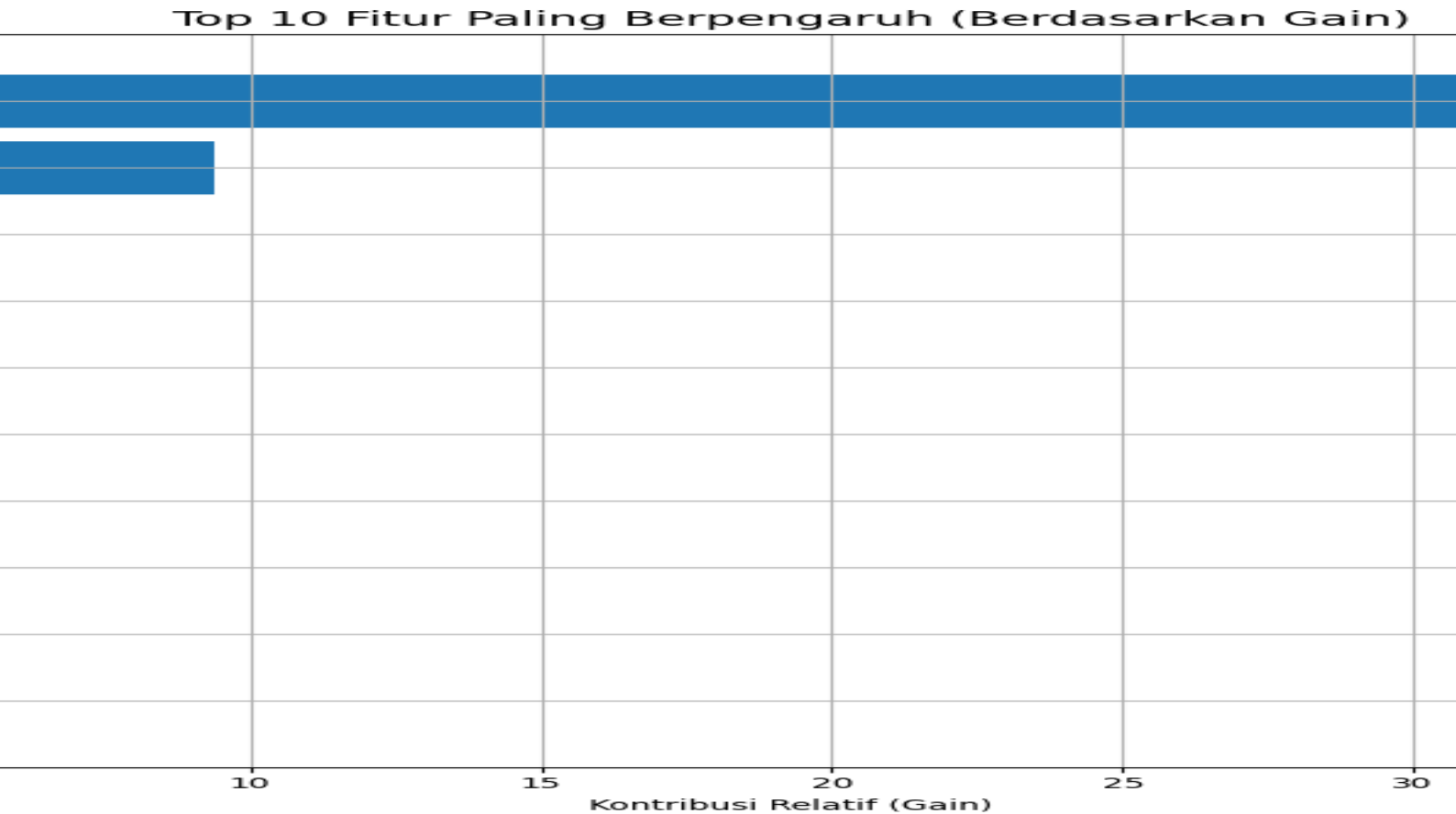
BAB II: METODOLOGI PENELITIAN

Metodologi penelitian dibagi menjadi beberapa tahapan utama:

- 1 **Pengumpulan Data:** Data dikumpulkan dari sensor IoT yang tersebar di beberapa titik DAS (Daerah Aliran Sungai), mencakup variabel seperti curah hujan (mm), ketinggian air sungai (cm), kelembapan tanah (%), suhu (°C), dan kecepatan angin (km/j) selama periode 24 bulan.
- 2 **Model Machine Learning:** Algoritma yang digunakan adalah XGBoost (eXtreme Gradient Boosting), yang dikenal karena performa tinggi dan kemampuannya menangani data tabular kompleks.
- 3 **Analisis Feature Importance:** Dua metode digunakan untuk mengukur pentingnya fitur:
 1. *Importance Type 'weight'* dari XGBoost.
 2. *SHAP Summary Plot* untuk dampak global tiap fitur terhadap prediksi.

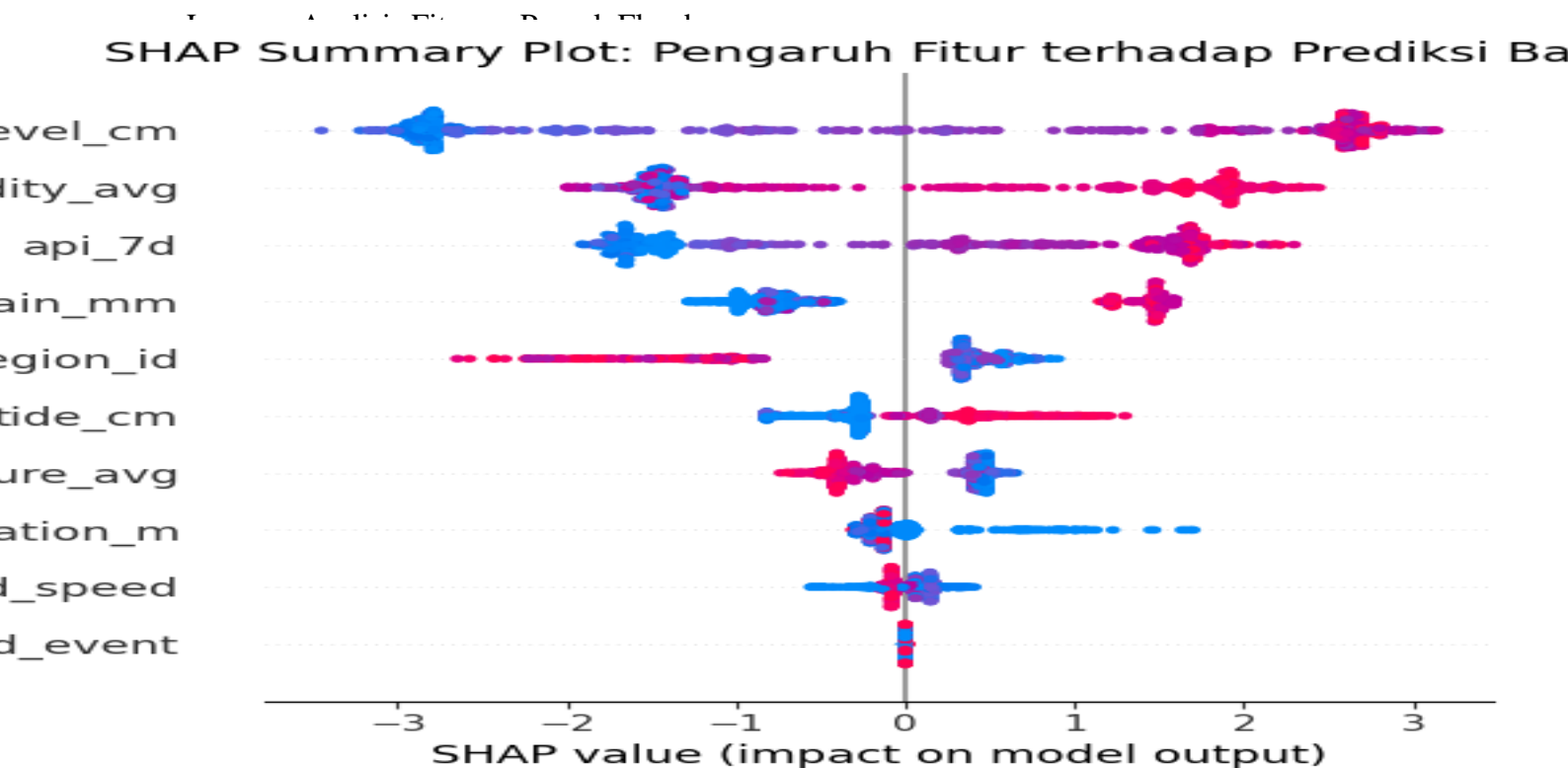
BAB III: HASIL DAN ANALISIS

Setelah model XGBoost dilatih, analisis **feature importance** dilakukan untuk menginterpretasi faktor-faktor yang paling mempengaruhi prediksi ketinggian air. Hasilnya disajikan dalam dua visualisasi utama.



Gambar 1: Peringkat 10 Fitur Teratas Berdasarkan XGBoost ('weight')

Gambar 1 menunjukkan frekuensi penggunaan setiap fitur untuk membuat keputusan di dalam model. Fitur *river_level_cm* menjadi yang paling sering digunakan, mengindikasikan bahwa kondisi sungai saat ini adalah prediktor terkuat untuk kondisi di masa depan.



Gambar 2: SHAP Summary Plot untuk Menganalisis Dampak Fitur

Gambar 2 memberikan detail lebih kaya. Titik-titik merah menunjukkan nilai fitur yang tinggi, sedangkan biru menunjukkan nilai rendah. Terlihat jelas bahwa nilai *river_level_cm* yang tinggi (merah) memiliki dampak SHAP positif yang besar, artinya sangat mendorong prediksi banjir. Hal yang sama berlaku untuk *rain_mm*.

BAB IV: KESIMPULAN

Berdasarkan analisis yang telah dilakukan, dapat disimpulkan bahwa fitur-fitur hidrologis utama seperti ketinggian air sebelumnya (*river_level_cm*), curah hujan (*rain_mm*), dan kelembapan tanah (*soil_moisture*) merupakan prediktor paling signifikan dalam model Floodzy. Visualisasi dari XGBoost dan SHAP secara konsisten menempatkan fitur-fitur ini di peringkat teratas.

Temuan ini mengonfirmasi bahwa model XGBoost tidak hanya berkinerja baik secara prediktif, tetapi juga dapat diinterpretasikan dan sejalan dengan pemahaman domain hidrologi. Hasil ini memberikan dasar yang kuat untuk implementasi model Floodzy dalam sistem peringatan dini banjir yang andal dan dapat dipertanggungjawabkan.