

INTISARI

Kejadian banjir di Indonesia semakin kompleks seiring dengan perubahan penggunaan dan tutupan lahan (*Land Use and Land Cover/LULC*) yang cepat serta dinamika kondisi hidrometeorologi. Model prediksi kejadian banjir berbasis data (*data-driven model*) umumnya masih menggunakan kumpulan data (*data set*) statis dan pendekatan pembelajaran batch (*batch learning*), sehingga sering kali kurang mampu beradaptasi terhadap perubahan kondisi lingkungan dari waktu ke waktu. Penelitian ini bertujuan untuk menganalisis pengaruh dinamika LULC terhadap prediksi kejadian banjir di Pulau Jawa dengan membandingkan dua paradigma pembelajaran mesin (*machine learning*), yaitu pembelajaran batch (*batch learning*) dan pembelajaran bertahap (*incremental learning*). *Data set* yang digunakan mencakup periode 2019-2024 dan berasal dari berbagai sumber, meliputi indeks LULC berbasis citra satelit untuk merepresentasikan vegetasi, kawasan terbangun, dan badan air permukaan; variabel hidrometeorologi seperti akumulasi curah hujan dan fitur jeda waktu (*lag*), serta variabel topografi dan jenis tanah. Wilayah studi dianalisis berdasarkan batas daerah aliran sungai (DAS) dan batas administrasi untuk menangkap heterogenitas spasial kejadian banjir.

Model prediksi berbasis data dikembangkan untuk memprediksi kejadian banjir dengan tiga pendekatan representasi LULC, yaitu tanpa LULC, dengan LULC statis, dan dengan LULC dinamis yang diperbarui secara berkala dari observasi satelit. Proses pelatihan model dilakukan menggunakan strategi pembelajaran batch dan pembelajaran bertahap untuk mengevaluasi kemampuan masing-masing pendekatan dalam menghadapi perubahan kondisi lingkungan. Untuk meningkatkan interpretabilitas model, digunakan pendekatan kecerdasan buatan yang dapat dijelaskan (*Explainable Artificial Intelligence/XAI*), khususnya *Shapley Additive Explanations* (SHAP), yaitu metode untuk mengukur kontribusi setiap variabel terhadap hasil prediksi, serta *partial dependence plot* (PDP) untuk menganalisis hubungan antara nilai variabel dan probabilitas kejadian banjir.

Hasil penelitian menunjukkan bahwa perlakuan LULC sebagai variabel dinamis memberikan performa prediksi yang lebih baik dibandingkan tanpa representasi LULC maupun dengan LULC statis, karena model mampu menangkap perubahan karakteristik permukaan lahan yang memengaruhi proses limpasan dan kerentanan banjir. Pendekatan *incremental learning* juga menunjukkan kemampuan adaptasi yang lebih baik terhadap perubahan distribusi data dibandingkan *batch learning*. Analisis interpretabilitas menunjukkan bahwa curah hujan jangka pendek merupakan faktor dominan dalam memicu banjir, namun pengaruhnya dimoderasi oleh kondisi LULC dan karakteristik fisik wilayah. Analisis PDP mengidentifikasi nilai ambang curah hujan sebesar 30,99 mm dan nilai ambang indeks vegetasi sebesar 0,31 yang memicu peningkatan probabilitas kejadian banjir secara signifikan hingga 0,6. Secara keseluruhan, penelitian ini menegaskan pentingnya representasi LULC dinamis, pembelajaran adaptif, dan interpretabilitas model dalam pengembangan sistem prediksi kejadian banjir berbasis data yang lebih responsif terhadap perubahan lingkungan.



UNIVERSITAS
GADJAH MADA

Dinamika LULC dalam Prediksi Kejadian Banjir di Pulau Jawa: Perbandingan Batch dan Incremental Learning berbasis Explainable Artificial Intelligence (XAI)

Aryanta Adri Herlangga, Vempi Satriya Adi Hendrawan, S.T., M.Env., Ph.D.

Universitas Gadjah Mada, 2026 | Diunduh dari <http://etd.repository.ugm.ac.id/>

Kata kunci: prediksi kejadian banjir, representasi LULC dinamis, *incremental learning*, *batch learning*, *machine learning*, *Explainable Artificial Intelligence (XAI)*, Pulau Jawa.

ABSTRACT

Flood events in Indonesia have become increasingly complex as a result of rapid Land Use and Land Cover (LULC) changes and evolving hydrological conditions. Conventional flood prediction models typically depend on static data sets and batch learning approaches, which often fail to adapt to temporal variations in environmental variables. This study examines the impact of LULC dynamics on flood prediction over Java Island by comparing batch learning and incremental learning machine learning paradigms. A multi-source data set covering the 2019–2024 period was assembled, incorporating satellite-derived LULC indices, hydrometeorological variables, topographical variables, soil type, and historical flood records. The analysis was done using both watershed and administrative boundaries to account for spatial variation in flood behavior.

A data-driven prediction model was developed to predict flood occurrence using three LULC representation scenarios: without LULC, with static LULC, and with dynamic LULC periodically updated from satellite observations. Model training was conducted using both batch learning and incremental learning strategies to evaluate the ability of each approach to respond to changing environmental conditions. To improve model interpretability, an Explainable Artificial Intelligence (XAI) approach was applied, particularly Shapley Additive Explanations (SHAP), a method used to quantify the contribution of each variable to model predictions, and partial dependence plot (PDP), which analyzes the relationship between feature values and flood occurrence probability.

The results show that treating LULC as a dynamic variable improves predictive performance compared with models without LULC representation or with static LULC representation, as the model can better capture changes in land surface characteristics that influence runoff processes and flood susceptibility. Incremental learning also demonstrates better adaptability to changes in data distribution than batch learning. The interpretability analysis indicates that short-term rainfall is the dominant factor triggering flood occurrence, although its effect is moderated by LULC conditions and the physical characteristics of the region. The PDP analysis identifies a rainfall threshold of 30.99 mm and a vegetation index threshold of 0.31, beyond which the probability of flood occurrence increases significantly up to 0.6. Overall, this study emphasizes the importance of dynamic LULC representation, adaptive learning, and model interpretability in developing a more responsive data-driven flood prediction system under changing environmental conditions.

Keywords: Flood prediction, Land Use Land Cover (LULC), Dynamic LULC representation, Incremental learning, Batch learning, Machine learning, Explainable Artificial Intelligence (XAI), Java Island.