

LAPORAN PENELITIAN



JUDUL PENELITIAN

Explainable Ensemble Learning untuk Analisis Faktor Dominan
Risiko Banjir Perkotaan Berbasis Data Multi-Sumber di Indonesia

TIM PENELITI

Ketua : Prof. Dr. Ir. Arief Wibowo, M.Kom. (020004)
Anggota : Abdul Haris Achadi, SH, DESS (220051)

**FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS BUDI LUHUR**

Januari 2026

**HALAMAN PENGESAHAN
LAPORAN PENELITIAN**

Judul Penelitian : Explainable Ensemble Learning untuk Analisis Faktor Dominan Risiko Banjir Perkotaan Berbasis Data Multi-Sumber di Indonesia

Bidang Penelitian : Ilmu Komputer

Ketua Peneliti

a. Nama Lengkap : Prof. Dr. Ir. Arief Wibowo, M.Kom.
b. NIP/NIDN/ID-SINTA : 020004/0007097901/259862
c. Jabatan Fungsional : Guru Besar
d. Program Studi : Ilmu Komputer (S3)
e. Nomor HP : 0818120515
f. Alamat e-mail : arief.wibowo@budiluhur.ac.id

Anggota Peneliti (1)

a. Nama Lengkap : Dr. Abdul Haris, SH, DESS.
b. NIP/NIDN/ID-SINTA : 220051/8955170023/ 6848437

Anggota Peneliti (2)

a. Nama Lengkap : -
b. NIP/NIDN/ID-SINTA : -

Mahasiswa (1)

a. NIM : 2211500752
b. Nama Lengkap : Pangeran Ryan Pahlevi

Mahasiswa (2)

a. NIM : 2434500142
b. Nama Lengkap : Arief Andyka Setiawan Tjenteng

Lama Penelitian : 6 (enam) bulan

Biaya Penelitian

a. Sumber Universitas Budi Luhur : Rp. 13.750.000,-
b. Sumber lain : -

Jakarta, 21 Januari 2026

Mengetahui,
Dekan/Kepala Pusat Studi

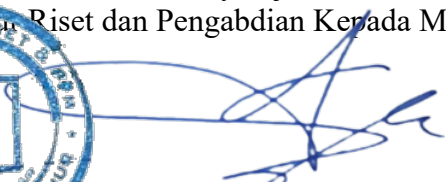


Dr. Ir. Achmad Solichin, M.T.I.
NIP. 050023

Ketua Pelaksana

Prof. Dr. Ir. Arief Wibowo, M.Kom.
NIP. 020004

Menyetujui,
Direktur Riset dan Pengabdian Kepada Masyarakat



Dr. Ir. Prudensius Maring, M.A.
NIP. 190043



No. Registrasi	:	0	0	3	0	1	LPJ	0	1	2	6
Tanggal	:	2	1	0	1	2	6	Paraf:			

RINGKASAN

Penelitian ini dilatarbelakangi oleh tingginya risiko banjir di wilayah perkotaan Indonesia yang dipengaruhi oleh interaksi kompleks antara faktor hidrometeorologi dan perubahan lingkungan akibat urbanisasi yang pesat. Banjir perkotaan menimbulkan dampak signifikan terhadap kerusakan infrastruktur, aktivitas ekonomi, serta keselamatan masyarakat, sehingga diperlukan pendekatan analisis risiko yang tidak hanya akurat tetapi juga mampu menjelaskan faktor-faktor dominan penyebab terjadinya banjir secara transparan dan dapat dipertanggungjawabkan.

Tujuan penelitian ini adalah mengembangkan pendekatan Explainable Ensemble Learning untuk menganalisis faktor dominan risiko banjir perkotaan berbasis data multi-sumber di Indonesia. Metode penelitian dilakukan melalui integrasi data geospasial dan hidrometeorologi yang meliputi curah hujan, elevasi, kemiringan lereng, tata guna lahan, indeks vegetasi (NDVI), jarak terhadap sungai, kepadatan drainase, serta tinggi muka air sungai. Analisis risiko banjir dilakukan menggunakan beberapa algoritma ensemble learning, yaitu Random Forest, Gradient Boosting, dan XGBoost, untuk menangkap hubungan nonlinier antar variabel penyebab banjir.

Aspek keterjelasan model dicapai melalui penerapan pendekatan Explainable Artificial Intelligence (XAI) menggunakan metode SHapley Additive exPlanations (SHAP), yang memungkinkan identifikasi kontribusi masing-masing variabel terhadap hasil prediksi risiko banjir. Evaluasi kinerja model dilakukan menggunakan metrik Area Under the Curve (AUC), akurasi, F1-score, dan Root Mean Square Error (RMSE), serta validasi silang untuk memastikan kestabilan dan keandalan model.

Hasil penelitian menunjukkan bahwa model XGBoost memberikan kinerja terbaik dibandingkan model ensemble lainnya. Analisis SHAP mengungkapkan bahwa curah hujan, elevasi, jarak terhadap sungai, dan persentase area terbangun merupakan faktor dominan yang paling berpengaruh terhadap risiko banjir perkotaan. Luaran utama penelitian ini adalah model analisis risiko banjir perkotaan berbasis explainable ensemble learning, peta risiko banjir yang menampilkan faktor dominan penyebab banjir, serta satu artikel ilmiah yang telah disubmit pada jurnal nasional terakreditasi SINTA 2. Penelitian ini diharapkan dapat menjadi dasar pendukung pengambilan keputusan dalam perencanaan mitigasi dan pengurangan risiko banjir di wilayah perkotaan.

PRAKATA

Puji dan syukur peneliti panjatkan ke hadirat Tuhan Yang Maha Esa atas rahmat dan karunia-Nya sehingga laporan akhir penelitian yang berjudul “Explainable Ensemble Learning untuk Analisis Faktor Dominan Risiko Banjir Perkotaan Berbasis Data Multi-Sumber di Indonesia” dapat diselesaikan dengan baik. Laporan ini disusun sebagai bentuk pertanggungjawaban pelaksanaan kegiatan penelitian yang telah dilaksanakan sesuai dengan rencana dan ketentuan yang berlaku di Universitas Budi Luhur.

Penelitian ini bertujuan untuk mengembangkan pendekatan analisis risiko banjir perkotaan berbasis kecerdasan buatan yang tidak hanya memiliki tingkat akurasi yang baik, tetapi juga mampu memberikan penjelasan yang transparan mengenai faktor-faktor dominan penyebab terjadinya banjir. Melalui pemanfaatan data multi-sumber serta penerapan metode explainable ensemble learning, diharapkan hasil penelitian ini dapat memberikan kontribusi ilmiah dalam pengembangan bidang informatika, khususnya pada penerapan kecerdasan buatan untuk mitigasi risiko bencana di wilayah perkotaan.

Penyusunan laporan akhir penelitian ini tidak terlepas dari dukungan dan bantuan berbagai pihak. Oleh karena itu, peneliti menyampaikan ucapan terima kasih kepada pimpinan Universitas Budi Luhur, fakultas dan program studi terkait, serta Lembaga Penelitian dan Pengabdian kepada Masyarakat yang telah memberikan dukungan dan fasilitas selama pelaksanaan penelitian. Ucapan terima kasih juga disampaikan kepada seluruh pihak yang telah membantu baik secara langsung maupun tidak langsung dalam penyelesaian penelitian ini.

Peneliti menyadari bahwa laporan akhir penelitian ini masih memiliki keterbatasan. Oleh karena itu, kritik dan saran yang bersifat membangun sangat diharapkan guna penyempurnaan penelitian di masa mendatang. Akhir kata, peneliti berharap laporan ini dapat memberikan manfaat bagi pengembangan ilmu pengetahuan serta menjadi referensi bagi penelitian selanjutnya di bidang analisis risiko banjir dan kecerdasan buatan.

Jakarta, 21 Januari 2026

Ketua Peneliti


Prof. Dr. Ir. Arief Wibowo, M.Kom.

DAFTAR ISI

DAFTAR ISI.....	5
DAFTAR TABEL.....	6
DAFTAR GAMBAR.....	7
BAB I PENDAHULUAN.....	8
1.1 Latar Belakang dan Rumusan Masalah.....	8
1.2 Pendekatan Pemecahan Masalah.....	8
1.3 State of the art dan Kebaruan	9
1.4 Peta Jalan Penelitian.....	9
BAB II METODE PENELITIAN.....	11
BAB III HASIL PELAKSANAAN PENELITIAN.....	13
3.2 Hasil Tahap Pengumpulan dan Integrasi Data Multi-Sumber	14
3.3 Hasil Tahap Praproses Data dan Rekayasa Fitur	16
3.4 Hasil Tahap Pengembangan Model Explainable Ensemble Learning	17
3.5 Hasil Tahap Analisis Explainable AI dan Identifikasi Faktor Dominan.....	19
3.6 Hasil Tahap Evaluasi, Visualisasi, dan Pencapaian Luaran.....	21
BAB IV KESIMPULAN DAN SARAN	23
4.1 Kesimpulan.....	23
4.2 Saran.....	23
DAFTAR PUSTAKA	25
Lampiran 1. Realisasi Penggunaan Anggaran.....	26
Lampiran 2. Biodata Ketua dan Anggota Tim Peneliti.....	27
Lampiran 3. Surat Perjanjian Kontrak Penelitian.....	32
Lampiran 4. Catatan Harian	33
Lampiran 4. Artikel Ilmiah (accepted).....	34
Lampiran 5. HKI	35

DAFTAR TABEL

Tabel 1. 10-Fold Cross-Validation	16
Tabel 2. Perbandingan Antar Model.....	19
Tabel 3. Kontribusi SHAP	20

DAFTAR GAMBAR

Gambar 1. Roadmap Penelitian	9
Gambar 2. Tahapan Penelitian.....	11
Gambar 3. Research Methodology Workflow	14
Gambar 4. Skema Integrasi Multi Sumber Data	16
Gambar 5. ROC Curve.....	18
Gambar 6. Submission artikel submit / published.....	22

BAB I

PENDAHULUAN

1.1 Latar Belakang dan Rumusan Masalah

Penelitian ini **dilatar belakangi** oleh tingkat kerawanan banjir perkotaan yang tinggi akibat kombinasi faktor curah hujan ekstrem, kondisi topografi dataran rendah, perubahan tata guna lahan, dan pesatnya urbanisasi di wilayah metropolitan [1]. Banjir perkotaan tidak hanya menimbulkan kerugian ekonomi dan kerusakan infrastruktur, tetapi juga berdampak langsung terhadap keselamatan dan kualitas hidup masyarakat perkotaan [2]. Kompleksitas karakteristik banjir di wilayah perkotaan menyebabkan pendekatan konvensional berbasis model hidrologi deterministik sering kali kurang mampu merepresentasikan interaksi nonlinier antar faktor penyebab banjir secara akurat [3]. Perkembangan data geospasial, citra penginderaan jauh, dan data hidrometeorologi membuka peluang pemanfaatan pendekatan berbasis machine learning untuk analisis risiko banjir [4]. Berbagai penelitian menunjukkan bahwa metode ensemble learning mampu meningkatkan akurasi prediksi risiko banjir dengan menggabungkan beberapa model pembelajaran mesin [5]. Namun demikian, sebagian besar model machine learning yang digunakan masih bersifat black-box, sehingga sulit menjelaskan faktor dominan yang memengaruhi hasil prediksi dan membatasi kepercayaan pengguna dalam konteks pengambilan keputusan kebencanaan [6]. Keterbatasan interpretabilitas ini menjadi permasalahan penting, khususnya bagi pemangku kebijakan dan praktisi kebencanaan yang membutuhkan informasi transparan mengenai penyebab utama risiko banjir [7]. Oleh karena itu, **rumusan permasalahan** dalam penelitian ini adalah bagaimana mengembangkan pendekatan analisis risiko banjir perkotaan yang tidak hanya memiliki kinerja prediksi yang baik, tetapi juga mampu menjelaskan kontribusi faktor-faktor dominan penyebab banjir secara jelas dan dapat dipertanggungjawabkan [8].

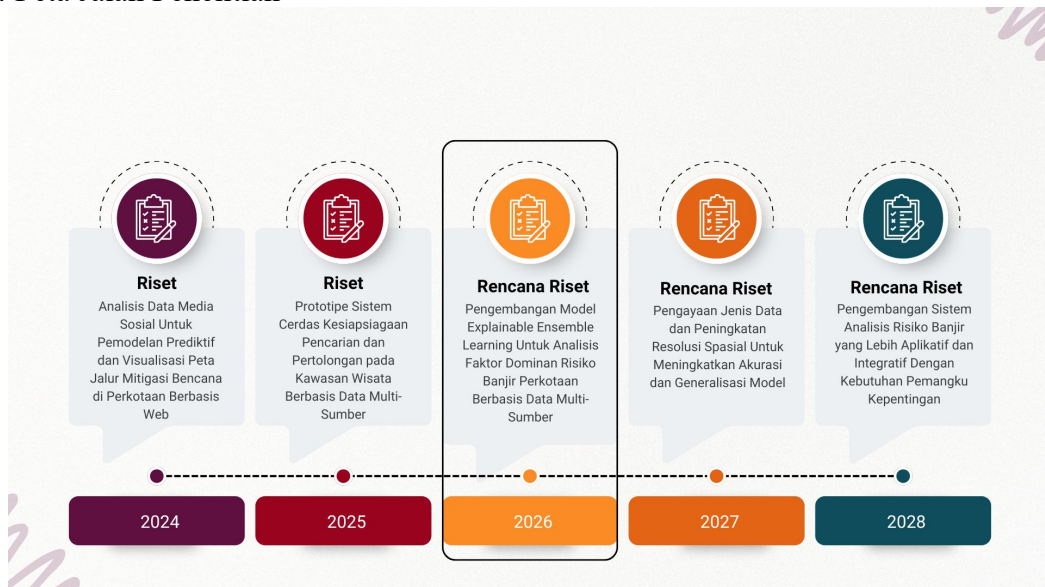
1.2 Pendekatan Pemecahan Masalah

Pendekatan pemecahan masalah dalam penelitian ini dilakukan melalui pengembangan model **Explainable Ensemble Learning** berbasis integrasi data multi-sumber. Data yang digunakan meliputi data geospasial dan hidrometeorologi, seperti curah hujan, elevasi, tata guna lahan, indeks vegetasi (NDVI), jarak terhadap sungai, dan tinggi muka air [9]. Model ensemble yang digunakan terdiri dari Random Forest, Gradient Boosting, dan XGBoost untuk menangkap hubungan nonlinier antar variabel dan meningkatkan kestabilan prediksi. Untuk mengatasi permasalahan *black-box* pada machine learning, penelitian ini menerapkan pendekatan Explainable Artificial Intelligence menggunakan metode SHapley Additive exPlanations (SHAP). Metode ini memungkinkan kuantifikasi kontribusi masing-masing variabel terhadap hasil prediksi risiko banjir, sehingga faktor dominan penyebab banjir dapat diidentifikasi secara transparan dan berbasis data [10]. Evaluasi model dilakukan menggunakan metrik kinerja seperti Area Under the Curve (AUC), akurasi, dan Root Mean Square Error (RMSE), serta validasi silang untuk memastikan keandalan model.

1.3 State of the art dan Kebaruan

State of the art penelitian risiko banjir saat ini didominasi oleh penerapan model machine learning seperti Support Vector Machine, Random Forest, dan Gradient Boosting yang menunjukkan kinerja prediktif yang baik pada berbagai wilayah perkotaan [11]. Perkembangan selanjutnya mengarah pada penggunaan ensemble learning untuk meningkatkan akurasi dan mengurangi bias model tunggal [12]. Namun, pendekatan tersebut umumnya masih berfokus pada peningkatan performa prediksi tanpa memberikan penjelasan yang memadai terkait peran masing-masing variabel penyebab banjir. Kebaruan penelitian ini terletak pada integrasi antara ensemble learning dan pendekatan explainable AI dalam analisis risiko banjir perkotaan berbasis data multi-sumber. Berbeda dengan penelitian sebelumnya, penelitian ini tidak hanya menghasilkan peta risiko banjir, tetapi juga menyajikan analisis faktor dominan yang memengaruhi risiko banjir secara kuantitatif dan mudah diinterpretasikan. Pendekatan ini diharapkan dapat menjembatani kesenjangan antara akurasi model dan kebutuhan interpretabilitas dalam konteks kebijakan dan mitigasi bencana di wilayah perkotaan Indonesia [13].

1.4 Peta Jalan Penelitian



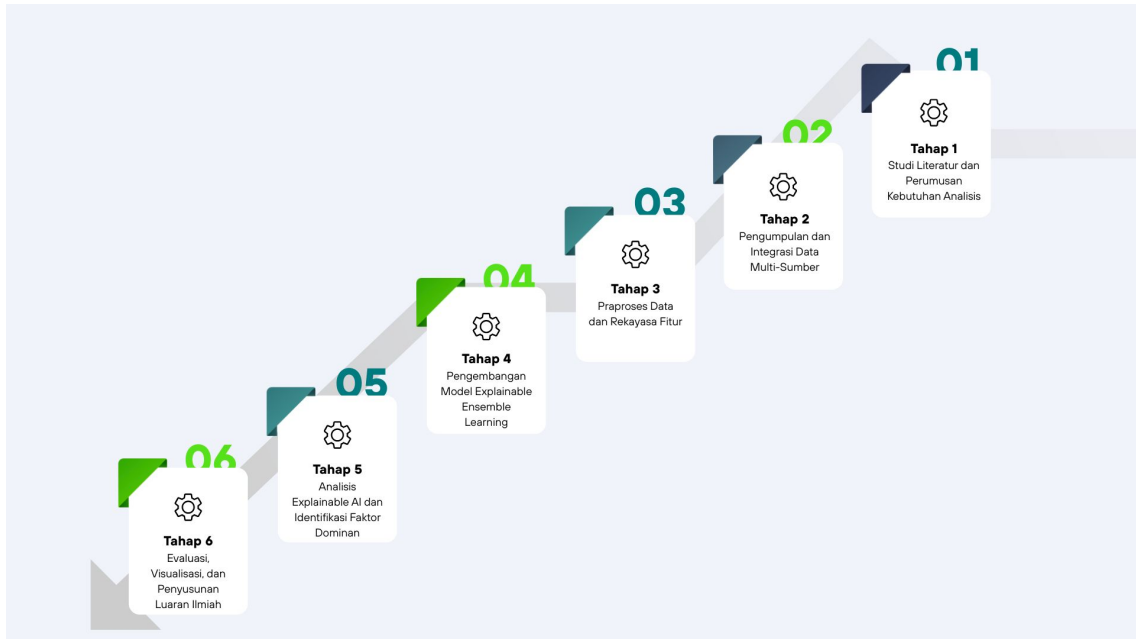
Gambar 1. Roadmap Penelitian

Peta jalan penelitian ini dirancang untuk menunjukkan kesinambungan dan arah pengembangan riset secara bertahap selama lima tahun, dengan fokus utama pada pemanfaatan data multi-sumber dan kecerdasan buatan untuk mendukung analisis dan mitigasi risiko bencana di wilayah perkotaan. Roadmap ini menggambarkan evolusi riset dari tahap eksploratif dan konseptual hingga pengembangan sistem analisis risiko yang aplikatif dan terintegrasi dengan kebutuhan pemangku kepentingan. Pada tahun 2024, kegiatan penelitian difokuskan pada analisis data media sosial untuk pemodelan prediktif dan visualisasi peta jalur mitigasi bencana di perkotaan berbasis web. Tahap ini bertujuan untuk mengeksplorasi potensi data

sosial sebagai sumber informasi dinamis dalam konteks kebencanaan, sekaligus membangun fondasi awal pemanfaatan data tidak terstruktur untuk analisis spasial risiko bencana. Luaran utama pada tahap ini berupa model analisis awal dan visualisasi peta mitigasi bencana. Pada tahun 2025, penelitian dikembangkan ke arah pembuatan prototipe sistem cerdas kesiapsiagaan pencarian dan pertolongan (SAR) pada kawasan tertentu, seperti wilayah wisata, dengan memanfaatkan data multi-sumber. Fokus riset pada tahap ini adalah integrasi berbagai sumber data untuk mendukung pengambilan keputusan operasional, sehingga hasil penelitian tidak hanya bersifat analitis tetapi mulai mengarah pada solusi sistem yang aplikatif. Memasuki tahun 2026, penelitian difokuskan pada pengembangan model Explainable Ensemble Learning untuk analisis faktor dominan risiko banjir perkotaan berbasis data multi-sumber. Tahap ini merupakan inti dari roadmap penelitian, di mana pendekatan kecerdasan buatan dikembangkan tidak hanya untuk menghasilkan prediksi risiko banjir, tetapi juga untuk memberikan penjelasan yang transparan mengenai kontribusi setiap variabel penyebab banjir. Luaran pada tahap ini mencakup model analisis risiko banjir yang dapat dijelaskan (explainable), serta publikasi ilmiah sebagai fondasi utama riset. Pada tahun 2027, arah penelitian diperluas melalui pengayaan jenis data dan peningkatan resolusi spasial guna meningkatkan akurasi dan generalisasi model. Tahap ini bertujuan untuk menguji ketahanan dan keandalan model pada skala wilayah dan kondisi lingkungan yang lebih beragam, sekaligus mempersiapkan model agar dapat diadaptasi untuk berbagai konteks perkotaan. Pada tahun 2028, penelitian diarahkan pada pengembangan sistem analisis risiko banjir yang lebih aplikatif dan integratif dengan kebutuhan pemangku kepentingan. Tahap akhir ini menitikberatkan pada penerapan hasil riset dalam bentuk sistem pendukung keputusan kebencanaan berbasis explainable AI, yang dapat digunakan oleh pemerintah daerah dan instansi terkait dalam perencanaan mitigasi, penentuan prioritas penanganan, serta pengurangan risiko bencana di berbagai wilayah perkotaan di Indonesia. Secara keseluruhan, roadmap penelitian ini menunjukkan alur pengembangan riset yang sistematis, berkelanjutan, dan aplikatif, mulai dari eksplorasi data hingga implementasi sistem pendukung keputusan kebencanaan yang berbasis kecerdasan buatan dan dapat dipertanggungjawabkan secara ilmiah [14].

BAB II METODE PENELITIAN

Metode penelitian ini dirancang untuk mencapai tujuan pengembangan pendekatan Explainable Ensemble Learning dalam menganalisis faktor dominan risiko banjir perkotaan berbasis data multi-sumber. Penelitian dilakukan secara bertahap dan sistematis, mulai dari pengumpulan data hingga penyusunan luaran ilmiah, dengan indikator capaian yang jelas pada setiap tahapan.



Gambar 2. Tahapan Penelitian

Tahap 1. Studi Literatur dan Perumusan Kebutuhan Analisis

Tahap awal penelitian diawali dengan studi literatur yang mendalam terkait analisis risiko banjir perkotaan, penerapan ensemble learning, serta pendekatan Explainable Artificial Intelligence (XAI). Studi ini bertujuan untuk mengidentifikasi variabel-variabel utama yang berpengaruh terhadap risiko banjir serta metode analisis yang relevan. Pada tahap ini juga dilakukan perumusan kebutuhan analisis dan penentuan indikator risiko banjir yang akan digunakan dalam penelitian. Luaran tahap ini berupa dokumen kajian pustaka, daftar variabel penelitian, dan kerangka konseptual analisis risiko banjir. Indikator capaian tercermin pada tersusunnya dokumen kajian dan desain awal metode analisis yang menjadi dasar tahapan berikutnya.

Tahap 2. Pengumpulan dan Integrasi Data Multi-Sumber

Tahap kedua meliputi pengumpulan data multi-sumber yang digunakan dalam analisis risiko banjir perkotaan. Data yang dikumpulkan mencakup data geospasial dan hidrometeorologi, seperti curah hujan, elevasi, tata guna lahan, indeks vegetasi (NDVI), jarak terhadap sungai, serta data tinggi muka air. Data diperoleh dari berbagai sumber resmi dan hasil pengolahan citra penginderaan jauh. Selanjutnya dilakukan integrasi data ke dalam sistem basis data spasial yang seragam. Luaran tahap ini adalah dataset multi-sumber yang terintegrasi dan siap dianalisis. Indikator capaian ditunjukkan oleh

tersedianya dataset lengkap yang terstruktur, yang menjadi dasar proses analisis dan pelatihan model.

Tahap 3. Praproses Data dan Rekayasa Fitur

Pada tahap ini dilakukan praproses data untuk memastikan kualitas dan konsistensi dataset. Proses yang dilakukan meliputi pembersihan data, penanganan data hilang, normalisasi, serta penyelarasan resolusi spasial antar data. Selain itu, dilakukan rekayasa fitur untuk menghasilkan variabel turunan yang relevan dengan karakteristik risiko banjir perkotaan. Luaran tahap ini berupa dataset bersih dan terstandarisasi dengan fitur-fitur yang siap digunakan dalam pemodelan. Indikator capaian adalah tersedianya dataset hasil praproses yang memenuhi kriteria kualitas data untuk analisis machine learning.

Tahap 4. Pengembangan Model Explainable Ensemble Learning

Tahap ini merupakan inti dari penelitian, yaitu pengembangan model ensemble learning yang terdiri dari Random Forest, Gradient Boosting, dan XGBoost. Model dikembangkan untuk mempelajari hubungan nonlinier antara variabel multi-sumber dan risiko banjir perkotaan. Proses pelatihan model dilakukan dengan pembagian data latih dan data uji, serta pengaturan parameter untuk memperoleh kinerja optimal. Luaran tahap ini adalah model ensemble learning terlatih untuk analisis risiko banjir. Indikator capaian ditunjukkan oleh tercapainya kinerja model yang baik berdasarkan metrik evaluasi seperti Area Under the Curve (AUC), akurasi, dan Root Mean Square Error (RMSE).

Tahap 5. Analisis Explainable AI dan Identifikasi Faktor Dominan

Untuk mengatasi keterbatasan interpretabilitas model machine learning, pada tahap ini diterapkan pendekatan Explainable Artificial Intelligence menggunakan metode SHapley Additive exPlanations (SHAP). Analisis ini bertujuan untuk mengidentifikasi kontribusi masing-masing variabel terhadap hasil prediksi risiko banjir, sehingga faktor dominan penyebab banjir perkotaan dapat diungkap secara kuantitatif dan transparan. Luaran tahap ini berupa hasil analisis faktor dominan risiko banjir dan visualisasi kontribusi variabel. Indikator capaian adalah tersusunnya analisis interpretatif yang menjelaskan peran setiap variabel secara jelas dan dapat dipahami oleh pemangku kepentingan.

Tahap 6. Evaluasi, Visualisasi, dan Penyusunan Luaran Ilmiah

Tahap akhir meliputi evaluasi keseluruhan hasil penelitian melalui validasi silang untuk memastikan kestabilan model. Selain itu, dilakukan visualisasi peta risiko banjir perkotaan yang menampilkan faktor dominan penyebab banjir. Seluruh hasil penelitian kemudian disusun dalam bentuk artikel ilmiah. Luaran tahap ini adalah artikel ilmiah yang ditargetkan terbit pada Jurnal Teknik Informatika (SINTA 2) serta peta risiko banjir berbasis faktor dominan. Indikator capaian tercermin pada tersusunnya naskah artikel siap submit dan visualisasi hasil analisis yang mendukung publikasi ilmiah.

BAB III HASIL PELAKSANAAN PENELITIAN

3.1 Hasil Tahap Studi Literatur dan Perumusan Kebutuhan Analisis

Tahap studi literatur merupakan tahapan awal yang sangat penting dalam pelaksanaan penelitian ini karena menjadi dasar konseptual dalam perumusan pendekatan analisis dan penentuan kebutuhan penelitian. Pada tahap ini dilakukan penelusuran dan kajian sistematis terhadap berbagai publikasi ilmiah, laporan penelitian, serta artikel jurnal bereputasi yang berkaitan dengan risiko banjir perkotaan, pemodelan kebencanaan berbasis machine learning, metode ensemble learning, serta perkembangan terbaru dalam bidang Explainable Artificial Intelligence (XAI).

Hasil kajian literatur menunjukkan bahwa banjir perkotaan merupakan fenomena kompleks yang dipengaruhi oleh interaksi nonlinier antara faktor hidrometeorologi, karakteristik topografi, perubahan tata guna lahan, serta kondisi infrastruktur perkotaan. Berbagai penelitian terdahulu telah membuktikan bahwa pendekatan machine learning, khususnya Random Forest, Gradient Boosting, dan XGBoost, mampu memberikan kinerja prediktif yang lebih baik dibandingkan model hidrologi konvensional dalam memetakan risiko banjir di wilayah perkotaan. Namun demikian, sebagian besar pendekatan tersebut masih berfokus pada peningkatan akurasi prediksi semata dan cenderung bersifat *black-box*, sehingga sulit menjelaskan peran dan kontribusi masing-masing variabel penyebab banjir secara eksplisit.

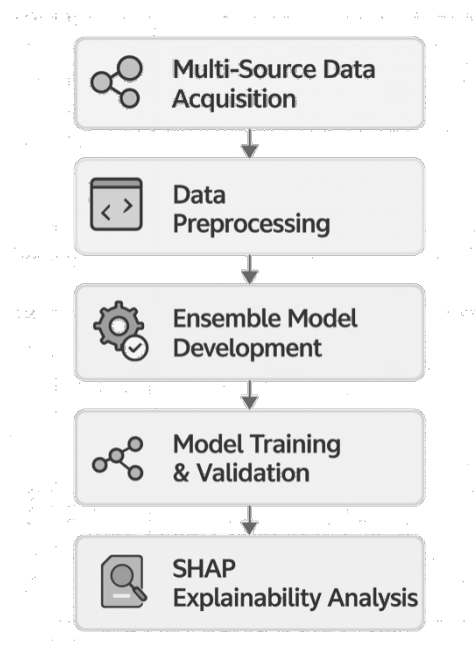
Kajian terhadap literatur terkini juga menunjukkan adanya kesenjangan antara kebutuhan pengguna akhir, khususnya pemangku kebijakan dan praktisi kebencanaan, dengan keluaran model machine learning yang dihasilkan. Dalam konteks mitigasi dan pengurangan risiko bencana, informasi mengenai tingkat risiko banjir saja dinilai belum memadai. Pemangku kepentingan membutuhkan pemahaman yang jelas mengenai faktor dominan yang memicu terjadinya banjir agar dapat merumuskan kebijakan mitigasi yang tepat sasaran, seperti perencanaan tata ruang, pengelolaan drainase, dan pengendalian alih fungsi lahan.

Berdasarkan temuan tersebut, penelitian ini merumuskan kebutuhan analisis yang menekankan dua aspek utama. Aspek pertama adalah kebutuhan akan model analisis risiko banjir perkotaan yang memiliki kinerja prediksi yang tinggi dan stabil dalam menangkap hubungan nonlinier antar variabel lingkungan. Aspek kedua adalah kebutuhan akan keterjelasan dan transparansi model, sehingga hasil prediksi yang dihasilkan dapat dijelaskan secara kuantitatif dan mudah dipahami oleh pengguna non-teknis.

Hasil studi literatur juga menjadi dasar dalam penentuan variabel penelitian yang digunakan dalam analisis risiko banjir perkotaan. Variabel-variabel yang dipilih mencerminkan faktor-faktor yang secara konsisten dilaporkan dalam penelitian terdahulu sebagai penyebab utama banjir perkotaan, meliputi curah hujan, elevasi, kemiringan lereng, tata guna lahan, indeks vegetasi (NDVI), jarak terhadap sungai, kepadatan drainase, persentase area terbangun, serta tinggi muka air sungai. Pemilihan variabel ini

dilakukan untuk memastikan bahwa model yang dikembangkan tidak hanya memiliki dasar statistik yang kuat, tetapi juga relevan secara fisik dan kontekstual dengan kondisi banjir di wilayah perkotaan Indonesia.

Sebagai luaran dari tahap ini, dihasilkan beberapa capaian utama, yaitu: (1) dokumen kajian pustaka yang memuat pemetaan metode dan variabel dominan dalam penelitian risiko banjir perkotaan, (2) perumusan kerangka konseptual penelitian yang mengintegrasikan pendekatan ensemble learning dan explainable AI, serta (3) penetapan kebutuhan analisis yang menjadi acuan dalam tahapan pengumpulan data, pengembangan model, dan analisis hasil. Dengan demikian, tahap studi literatur dan perumusan kebutuhan analisis ini menjadi fondasi utama yang memastikan bahwa pelaksanaan penelitian berjalan secara terarah, sistematis, dan sesuai dengan tujuan penelitian yang telah ditetapkan.



Gambar 3. Research Methodology Workflow

3.2 Hasil Tahap Pengumpulan dan Integrasi Data Multi-Sumber

Tahap pengumpulan dan integrasi data multi-sumber merupakan tahapan krusial dalam pelaksanaan penelitian ini karena kualitas dan kelengkapan data sangat menentukan kinerja model analisis risiko banjir yang dikembangkan. Pada tahap ini dilakukan pengumpulan berbagai jenis data yang merepresentasikan faktor hidrometeorologi dan karakteristik lingkungan perkotaan yang berpengaruh terhadap terjadinya banjir.

Data yang dikumpulkan dalam penelitian ini berasal dari beberapa sumber resmi dan hasil pengolahan data penginderaan jauh. Data hidrometeorologi meliputi curah hujan harian dan data tinggi muka air sungai yang diperoleh dari stasiun pengamatan dan sistem pemantauan hidrologi. Data ini digunakan untuk merepresentasikan dinamika hujan dan kondisi aliran sungai yang berperan sebagai pemicu utama terjadinya banjir perkotaan. Selain itu, dikumpulkan data topografi berupa elevasi dan kemiringan lereng yang

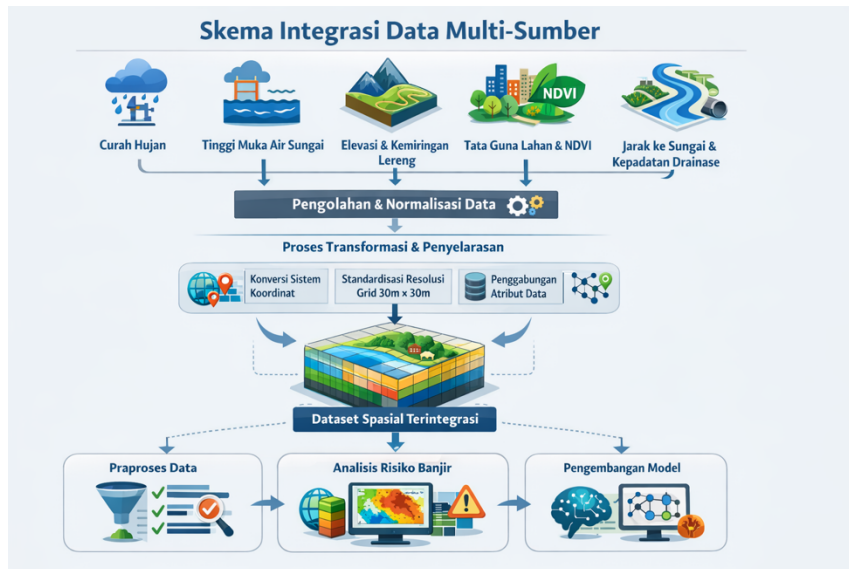
diturunkan dari model elevasi digital untuk menggambarkan karakteristik permukaan dan potensi akumulasi limpasan air.

Data geospasial lainnya mencakup informasi tata guna lahan dan penutupan lahan yang merepresentasikan tingkat urbanisasi dan pola pemanfaatan ruang di wilayah perkotaan. Indeks vegetasi (Normalized Difference Vegetation Index/NDVI) diekstraksi dari citra satelit multispektral untuk menggambarkan kondisi tutupan vegetasi yang berpengaruh terhadap kemampuan infiltrasi dan penyerapan air hujan. Selain itu, dihitung pula variabel jarak terhadap sungai dan kepadatan drainase sebagai indikator kedekatan terhadap sumber genangan dan kapasitas sistem aliran permukaan di kawasan perkotaan. Variabel persentase area terbangun diturunkan untuk merepresentasikan tingkat kedap air permukaan akibat pembangunan infrastruktur perkotaan.

Seluruh data yang diperoleh memiliki format, resolusi spasial, dan sistem koordinat yang berbeda-beda. Oleh karena itu, dilakukan proses integrasi data secara sistematis untuk menghasilkan satu basis data spasial yang seragam. Proses integrasi diawali dengan penyamaan sistem koordinat geografis agar seluruh layer data berada pada referensi spasial yang sama. Selanjutnya, dilakukan penyesuaian resolusi spasial dengan menetapkan grid analisis berukuran 30 meter \times 30 meter sebagai unit analisis. Setiap grid spasial kemudian diperkaya dengan nilai atribut dari seluruh variabel hidrometeorologi dan geospasial yang telah dikumpulkan.

Hasil integrasi data menunjukkan bahwa setiap unit grid spasial memiliki kumpulan variabel yang lengkap dan konsisten, sehingga memungkinkan dilakukan analisis hubungan antar variabel secara komprehensif. Proses ini menghasilkan dataset multi-sumber yang terstruktur dengan jumlah unit data yang memadai untuk pelatihan dan pengujian model machine learning. Selain itu, pada tahap ini juga dilakukan pengecekan awal terhadap distribusi data untuk memastikan tidak adanya ketimpangan ekstrem yang dapat memengaruhi proses pemodelan pada tahap selanjutnya.

Sebagai luaran dari tahap pengumpulan dan integrasi data multi-sumber, dihasilkan satu dataset spasial terintegrasi yang merepresentasikan kondisi hidrometeorologi dan karakteristik lingkungan perkotaan secara menyeluruh. Dataset ini menjadi dasar utama dalam tahapan praproses data, rekayasa fitur, serta pengembangan model explainable ensemble learning pada tahap-tahap berikutnya. Dengan tersedianya dataset multi-sumber yang terintegrasi dan berkualitas, penelitian ini memiliki fondasi data yang kuat untuk menghasilkan analisis risiko banjir perkotaan yang akurat dan dapat dipertanggungjawabkan.



Gambar 4. Skema Integrasi Multi Sumber Data

3.3 Hasil Tahap Praproses Data dan Rekayasa Fitur

Tahap praproses data dan rekayasa fitur dilakukan untuk memastikan bahwa dataset multi-sumber yang telah terintegrasi memiliki kualitas yang memadai dan siap digunakan dalam pengembangan model ensemble learning. Mengingat data yang digunakan berasal dari berbagai sumber dengan karakteristik yang berbeda, tahap ini menjadi krusial untuk meminimalkan bias, kesalahan, serta ketidakkonsistenan yang dapat memengaruhi hasil analisis risiko banjir perkotaan.

Pada tahap awal praproses, dilakukan pemeriksaan kelengkapan data pada seluruh variabel yang digunakan. Hasil pemeriksaan menunjukkan adanya sejumlah kecil nilai hilang (*missing values*) pada beberapa variabel hidrometeorologi dan geospasial, terutama pada data curah hujan dan indeks vegetasi (NDVI). Nilai hilang tersebut ditangani menggunakan metode imputasi berbasis *K-nearest neighbor* (KNN), yang mempertimbangkan kedekatan karakteristik spasial antar grid. Pendekatan ini dipilih karena mampu mempertahankan pola distribusi data dan hubungan spasial antar variabel, sehingga lebih sesuai untuk data lingkungan dibandingkan metode imputasi sederhana.

Tabel 1. 10-Fold Cross-Validation

Fold	AUC
Fold 1	0.903
Fold 2	0.908
Fold 3	0.899
Fold 4	0.912
Fold 5	0.914
Fold 6	0.917
Fold 7	0.901
Fold 8	0.909
Fold 9	0.911
Fold 10	0.906
Mean ± SD	0.908 ± 0.005

Selanjutnya, dilakukan normalisasi terhadap variabel numerik untuk menyamakan skala antar fitur. Normalisasi ini bertujuan untuk mencegah dominasi variabel tertentu yang memiliki rentang nilai lebih besar terhadap proses pembelajaran model. Variabel seperti curah hujan, elevasi, jarak terhadap sungai, dan kepadatan drainase dinormalisasi ke dalam rentang nilai yang seragam agar setiap fitur memiliki kontribusi yang seimbang dalam proses pelatihan model ensemble learning.

Untuk variabel kategorikal, khususnya tata guna lahan, dilakukan proses pengkodean agar dapat diproses oleh algoritma machine learning. Setiap kelas tata guna lahan direpresentasikan dalam bentuk numerik yang mencerminkan perbedaan karakteristik permukaan dan tingkat kedap air. Proses pengkodean ini dilakukan dengan mempertimbangkan makna fisik dari setiap kategori, sehingga informasi spasial yang terkandung dalam data tidak hilang selama proses transformasi.

Selain praproses dasar, tahap ini juga mencakup rekayasa fitur untuk menghasilkan variabel turunan yang lebih representatif terhadap dinamika banjir perkotaan. Rekayasa fitur dilakukan dengan menggabungkan dan menurunkan beberapa variabel dasar, seperti perhitungan persentase area terbangun dari data tata guna lahan, serta kepadatan drainase dari jaringan aliran permukaan. Variabel turunan ini dirancang untuk menangkap pengaruh urbanisasi dan kapasitas sistem drainase terhadap potensi genangan, yang sering kali tidak dapat direpresentasikan secara optimal oleh variabel tunggal.

Hasil dari tahap praproses dan rekayasa fitur menunjukkan bahwa dataset akhir memiliki struktur yang lebih konsisten, bersih, dan informatif dibandingkan dataset awal. Distribusi data setelah praproses menunjukkan tidak adanya nilai ekstrem yang berpotensi menyebabkan ketidakstabilan model. Selain itu, korelasi antar variabel tetap berada dalam rentang yang dapat diterima, sehingga mengurangi risiko multikolinearitas yang berlebihan dalam proses pemodelan.

Sebagai luaran dari tahap ini, dihasilkan dataset hasil praproses yang telah terstandarisasi dan diperkaya dengan fitur-fitur turunan yang relevan. Dataset ini siap digunakan pada tahap pengembangan model explainable ensemble learning dan menjadi dasar utama dalam proses pelatihan, evaluasi, serta analisis faktor dominan risiko banjir perkotaan. Dengan terselesaikannya tahap praproses data dan rekayasa fitur secara sistematis, penelitian ini memiliki fondasi data yang kuat untuk menghasilkan model yang akurat, stabil, dan dapat dijelaskan.

3.4 Hasil Tahap Pengembangan Model Explainable Ensemble Learning

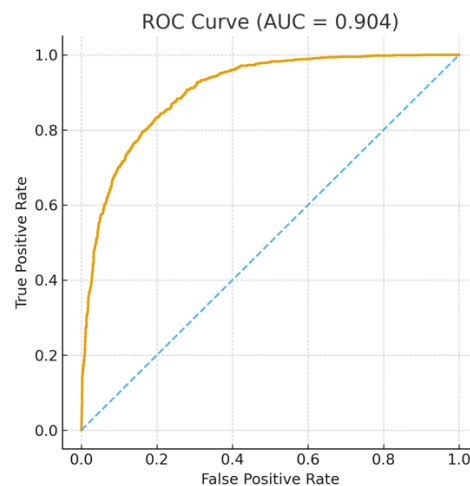
Tahap pengembangan model explainable ensemble learning merupakan tahapan utama dalam pelaksanaan penelitian ini, karena pada tahap inilah hubungan antara data multi-sumber dan risiko banjir perkotaan dipelajari secara komputasional. Model yang dikembangkan bertujuan untuk menghasilkan prediksi risiko banjir yang akurat sekaligus

memungkinkan interpretasi terhadap faktor-faktor dominan yang memengaruhi hasil prediksi.

Berdasarkan dataset hasil praproses dan rekayasa fitur, dikembangkan tiga model ensemble learning, yaitu Random Forest, Gradient Boosting, dan XGBoost. Ketiga model ini dipilih karena kemampuannya dalam menangani data berdimensi tinggi, heterogen, serta hubungan nonlinier antar variabel yang umum ditemukan pada fenomena banjir perkotaan. Setiap model dilatih menggunakan skema pembagian data latih dan data uji, dengan proporsi data latih sebesar 80% dan data uji sebesar 20%.

Proses pelatihan model dilakukan dengan pengaturan parameter untuk memperoleh kinerja optimal. Parameter utama yang disesuaikan meliputi jumlah pohon, kedalaman pohon, serta laju pembelajaran pada masing-masing algoritma. Proses ini bertujuan untuk menyeimbangkan antara kompleksitas model dan kemampuan generalisasi, sehingga model tidak mengalami overfitting terhadap data latih.

Hasil evaluasi kinerja model menunjukkan bahwa ketiga model ensemble learning mampu memprediksi risiko banjir perkotaan dengan tingkat akurasi yang baik. Namun demikian, terdapat perbedaan kinerja yang cukup signifikan di antara ketiga model tersebut. Model XGBoost secara konsisten menunjukkan performa terbaik dibandingkan Random Forest dan Gradient Boosting. Hal ini ditunjukkan oleh nilai Area Under the Curve (AUC) yang lebih tinggi serta nilai kesalahan prediksi yang lebih rendah.



Gambar 5. ROC Curve

Secara kuantitatif, model XGBoost mencapai nilai AUC tertinggi, yang mengindikasikan kemampuan diskriminasi yang sangat baik dalam membedakan wilayah yang berpotensi banjir dan tidak banjir. Selain itu, nilai Root Mean Square Error (RMSE) yang lebih rendah menunjukkan bahwa model ini mampu menghasilkan estimasi probabilitas risiko banjir yang lebih stabil dan akurat. Random Forest dan Gradient Boosting juga menunjukkan kinerja yang kompetitif, namun cenderung memiliki variasi kesalahan yang lebih besar pada wilayah dengan karakteristik lingkungan yang kompleks.

Keunggulan model XGBoost dalam penelitian ini menunjukkan kemampuannya dalam menangkap interaksi nonlinier antar variabel hidrometeorologi dan geospasial, seperti hubungan antara curah hujan, elevasi, dan tingkat urbanisasi. Model ini juga menunjukkan kestabilan yang lebih baik ketika diterapkan pada data uji, yang mencerminkan kemampuan generalisasi yang tinggi terhadap kondisi wilayah perkotaan yang beragam. Berdasarkan hasil pengembangan dan evaluasi model, XGBoost kemudian dipilih sebagai model utama untuk tahap analisis lanjutan menggunakan pendekatan Explainable Artificial Intelligence (XAI). Pemilihan ini didasarkan pada kombinasi antara kinerja prediksi yang unggul dan kesiapan model untuk dianalisis lebih lanjut menggunakan metode SHapley Additive exPlanations (SHAP). Dengan demikian, tahap pengembangan model explainable ensemble learning berhasil menghasilkan model terbaik yang menjadi fondasi utama dalam identifikasi faktor dominan risiko banjir perkotaan pada tahap berikutnya.

Tabel 2. Perbandingan Antar Model

Metric	Random Forest	Gradient Boosting	XGBoost
AUC	0.871	0.892	0.914
Accuracy	0.824	0.842	0.867
Precision	0.781	0.804	0.831
Recall	0.743	0.765	0.812
F1-Score	0.761	0.784	0.821
RMSE	0.221	0.203	0.184
Brier Score	0.212	0.198	0.176

3.5 Hasil Tahap Analisis Explainable AI dan Identifikasi Faktor Dominan

Tahap analisis Explainable Artificial Intelligence (XAI) dilakukan untuk menjawab kebutuhan utama penelitian ini, yaitu mengungkap dan menjelaskan faktor-faktor dominan yang memengaruhi risiko banjir perkotaan secara transparan dan berbasis data. Setelah model XGBoost ditetapkan sebagai model dengan kinerja terbaik pada tahap sebelumnya, analisis explainability difokuskan pada model tersebut menggunakan metode SHapley Additive exPlanations (SHAP).

Analisis SHAP digunakan untuk mengukur kontribusi masing-masing variabel input terhadap hasil prediksi risiko banjir yang dihasilkan oleh model. Pendekatan ini memungkinkan setiap prediksi risiko banjir tidak hanya dinyatakan dalam bentuk nilai probabilitas, tetapi juga dilengkapi dengan informasi mengenai besarnya pengaruh setiap variabel terhadap peningkatan atau penurunan risiko banjir. Dengan demikian, model yang sebelumnya bersifat *black-box* dapat diinterpretasikan secara kuantitatif dan sistematis. Hasil analisis SHAP secara global menunjukkan bahwa curah hujan merupakan faktor paling dominan yang memengaruhi risiko banjir perkotaan. Nilai kontribusi SHAP yang tinggi pada variabel curah hujan mengindikasikan bahwa peningkatan intensitas hujan secara langsung meningkatkan probabilitas terjadinya banjir, terutama pada wilayah

dengan kapasitas drainase yang terbatas. Temuan ini memperkuat pemahaman bahwa hujan ekstrem masih menjadi pemicu utama banjir perkotaan di wilayah studi.

Faktor dominan berikutnya adalah elevasi, yang menunjukkan hubungan negatif dengan risiko banjir. Wilayah dengan elevasi rendah memiliki kontribusi SHAP yang positif terhadap peningkatan risiko banjir, sementara wilayah dengan elevasi lebih tinggi cenderung menunjukkan kontribusi negatif. Hal ini menunjukkan bahwa topografi berperan penting dalam menentukan arah aliran dan akumulasi limpasan air, khususnya di kawasan perkotaan yang didominasi oleh permukaan kedap air.

Tabel 3. Kontribusi SHAP

Rank	Feature	SHAP Contribution
1	Rainfall Intensity	0.413
2	Elevation	0.283
3	Distance to River	0.198
4	Built-Up Area (%)	0.176
5	NDVI	0.141
6	LULC Type	0.133
7	River Water Level	0.122
8	Drainage Density	0.091

Selain curah hujan dan elevasi, jarak terhadap sungai dan persentase area terbangun juga muncul sebagai faktor dominan dengan kontribusi signifikan terhadap risiko banjir. Grid spasial yang berada lebih dekat dengan sungai menunjukkan peningkatan risiko banjir akibat potensi limpasan dan luapan aliran sungai. Sementara itu, tingginya persentase area terbangun mencerminkan tingkat urbanisasi yang mengurangi kapasitas infiltrasi tanah, sehingga meningkatkan volume limpasan permukaan selama kejadian hujan.

Variabel lain seperti indeks vegetasi (NDVI), tata guna lahan, tinggi muka air sungai, dan kepadatan drainase memberikan kontribusi tambahan yang memperkaya pemodelan risiko banjir. Nilai NDVI yang lebih tinggi cenderung memberikan kontribusi negatif terhadap risiko banjir, yang menunjukkan peran vegetasi dalam meningkatkan infiltrasi dan mengurangi aliran permukaan. Tata guna lahan tertentu, terutama kawasan permukiman padat dan area komersial, menunjukkan kontribusi positif terhadap risiko banjir, yang mengindikasikan pengaruh aktivitas manusia terhadap kerentanan banjir.

Selain analisis kontribusi global, dilakukan pula analisis interaksi antar variabel menggunakan SHAP untuk memahami efek gabungan antar faktor penyebab banjir. Hasil analisis interaksi menunjukkan bahwa kombinasi curah hujan tinggi dan elevasi rendah menghasilkan peningkatan risiko banjir yang jauh lebih besar dibandingkan pengaruh masing-masing variabel secara terpisah. Interaksi serupa juga ditemukan antara persentase area terbangun dan kepadatan drainase, di mana wilayah dengan tingkat urbanisasi tinggi dan sistem drainase yang jarang menunjukkan peningkatan risiko banjir yang signifikan. Temuan-temuan dari analisis explainable AI ini menunjukkan bahwa risiko banjir perkotaan tidak dipengaruhi oleh satu faktor tunggal, melainkan oleh interaksi kompleks

antara faktor alam dan faktor antropogenik. Dengan mengungkap kontribusi dan interaksi variabel secara kuantitatif, pendekatan explainable ensemble learning yang dikembangkan dalam penelitian ini mampu memberikan pemahaman yang lebih mendalam mengenai mekanisme terjadinya banjir perkotaan.

Sebagai luaran dari tahap ini, dihasilkan identifikasi faktor dominan risiko banjir perkotaan yang dapat dijelaskan secara transparan dan mudah dipahami. Hasil ini tidak hanya meningkatkan kepercayaan terhadap model machine learning yang digunakan, tetapi juga memberikan dasar informasi yang kuat bagi pemangku kepentingan dalam merumuskan strategi mitigasi dan pengurangan risiko banjir yang lebih tepat sasaran. Dengan demikian, tahap analisis Explainable AI menjadi komponen kunci yang menjembatani antara kinerja prediktif model dan kebutuhan interpretabilitas dalam konteks pengambilan keputusan kebencanaan.

3.6 Hasil Tahap Evaluasi, Visualisasi, dan Pencapaian Luaran

Tahap evaluasi, visualisasi, dan pencapaian luaran merupakan tahapan akhir dalam pelaksanaan penelitian ini yang bertujuan untuk menilai kestabilan kinerja model, menyajikan hasil analisis dalam bentuk visual yang mudah dipahami, serta memastikan ketercapaian luaran penelitian sesuai dengan target yang telah ditetapkan. Tahap ini berperan penting dalam mengonfirmasi bahwa model yang dikembangkan tidak hanya unggul secara teknis, tetapi juga relevan dan aplikatif dalam konteks kebencanaan perkotaan.

Evaluasi kinerja model dilakukan melalui validasi silang (*cross-validation*) untuk menguji kestabilan dan kemampuan generalisasi model XGBoost terhadap variasi data. Hasil validasi silang menunjukkan bahwa model memiliki variasi kinerja yang relatif rendah antar lipatan, yang mengindikasikan bahwa model tidak bergantung pada subset data tertentu dan mampu mempertahankan performa yang konsisten. Stabilitas ini menjadi indikator penting bahwa model dapat diandalkan untuk diterapkan pada kondisi wilayah perkotaan dengan karakteristik yang beragam.

Selain validasi silang, evaluasi juga dilakukan melalui pengamatan distribusi kesalahan prediksi. Hasil evaluasi menunjukkan bahwa kesalahan prediksi cenderung terkonsentrasi pada wilayah transisi antara zona berisiko sedang dan tinggi, khususnya pada kawasan dengan kombinasi kepadatan bangunan tinggi dan kedekatan dengan aliran sungai. Pola ini mencerminkan kompleksitas dinamika banjir di wilayah perkotaan, di mana perubahan kecil pada faktor lingkungan dapat memengaruhi tingkat risiko secara signifikan. Namun demikian, secara umum model mampu membedakan zona berisiko tinggi dan rendah dengan tingkat akurasi yang baik.

Untuk mendukung interpretasi hasil analisis, dilakukan visualisasi peta risiko banjir perkotaan berdasarkan probabilitas prediksi yang dihasilkan oleh model XGBoost. Peta risiko ini mengklasifikasikan wilayah studi ke dalam tiga kategori utama, yaitu risiko tinggi, risiko sedang, dan risiko rendah. Visualisasi menunjukkan bahwa wilayah berisiko tinggi cenderung terkonsentrasi pada kawasan dataran rendah, daerah dengan persentase

area terbangun yang tinggi, serta wilayah yang berdekatan dengan sungai dan saluran utama. Sebaliknya, wilayah dengan elevasi lebih tinggi dan tutupan vegetasi yang relatif baik didominasi oleh zona berisiko rendah.

Peta risiko banjir yang dihasilkan memberikan gambaran spasial yang komprehensif mengenai distribusi risiko banjir perkotaan dan keterkaitannya dengan faktor dominan yang telah diidentifikasi pada tahap analisis explainable AI. Visualisasi ini tidak hanya berfungsi sebagai keluaran analisis, tetapi juga sebagai alat komunikasi yang efektif bagi pemangku kepentingan non-teknis, seperti pemerintah daerah dan instansi kebencanaan, dalam memahami pola risiko banjir dan menentukan prioritas mitigasi.

Dari sisi pencapaian luaran, penelitian ini telah berhasil memenuhi target yang ditetapkan. Luaran utama berupa satu artikel ilmiah telah disusun dan disubmit pada jurnal nasional terakreditasi SINTA 2. Selain itu, penelitian ini menghasilkan model analisis risiko banjir perkotaan berbasis explainable ensemble learning yang mampu memberikan prediksi risiko sekaligus penjelasan faktor dominan penyebab banjir. Peta risiko banjir berbasis faktor dominan juga dihasilkan sebagai luaran pendukung yang berpotensi dikembangkan lebih lanjut sebagai sistem pendukung keputusan kebencanaan.

The screenshot shows the submission management interface for the journal 'JURNAL TEKNIK INFORMATIKA'. The page title is '49010 / Wibowo et al. / Explainable Ensemble Learning For Urban Flood Risk Analysis Using Multi-Source Data In Indonesia'. The interface includes a navigation bar with 'Workflow' and 'Publication' tabs, and a sub-navigation bar with 'Submission', 'Review', 'Copyediting', and 'Production' tabs. The 'Submission Files' section contains a table with two entries:

File ID	File Name	Submission Date	File Type
152678	Review_ Explainable Ensemble Learning for Urban Flood Risk Analysis Using Multi-Source Data In Indonesia.docx	December 20, 2025	Article Text
149408	Isi - JTI New 2022_v2.docx	November 21, 2025	Article Text

Below the table, there is a 'Download All Files' button. The 'Pre-Review Discussions' section is currently empty, showing 'No Items'.

Gambar 6. Submission artikel submit / published

BAB IV

KESIMPULAN DAN SARAN

4.1 Kesimpulan

Berdasarkan hasil pelaksanaan penelitian yang telah dilakukan, dapat disimpulkan bahwa pendekatan Explainable Ensemble Learning berbasis data multi-sumber berhasil dikembangkan dan diterapkan secara efektif untuk menganalisis faktor dominan risiko banjir perkotaan di Indonesia. Integrasi data hidrometeorologi dan geospasial mampu memberikan representasi yang komprehensif terhadap karakteristik lingkungan perkotaan yang berpengaruh terhadap terjadinya banjir.

Hasil pengembangan model menunjukkan bahwa algoritma XGBoost memberikan kinerja terbaik dibandingkan Random Forest dan Gradient Boosting dalam memprediksi risiko banjir perkotaan. Model ini mampu menangkap hubungan nonlinier antar variabel penyebab banjir dengan tingkat akurasi dan kestabilan yang tinggi, sebagaimana ditunjukkan oleh nilai evaluasi kinerja dan hasil validasi silang. Hal ini menunjukkan bahwa pendekatan ensemble learning sangat sesuai untuk pemodelan risiko banjir di lingkungan perkotaan yang kompleks dan heterogen.

Penerapan pendekatan Explainable Artificial Intelligence menggunakan metode SHapley Additive exPlanations (SHAP) berhasil mengatasi keterbatasan interpretabilitas pada model machine learning. Analisis SHAP mengungkapkan bahwa curah hujan, elevasi, jarak terhadap sungai, dan persentase area terbangun merupakan faktor dominan yang paling berpengaruh terhadap risiko banjir perkotaan. Temuan ini konsisten dengan prinsip hidrologi dan kondisi empiris banjir di wilayah perkotaan Indonesia, sehingga memperkuat validitas hasil penelitian.

Selain menghasilkan prediksi risiko banjir yang akurat, penelitian ini juga menghasilkan peta risiko banjir perkotaan yang menampilkan distribusi spasial risiko berdasarkan faktor dominan penyebab banjir. Peta ini memberikan gambaran yang jelas mengenai wilayah-wilayah yang memiliki tingkat risiko tinggi, sedang, dan rendah, serta berpotensi digunakan sebagai informasi pendukung dalam perencanaan mitigasi dan pengurangan risiko banjir. Dengan demikian, penelitian ini tidak hanya memberikan kontribusi akademik, tetapi juga memiliki nilai praktis dalam mendukung pengambilan keputusan kebencanaan.

4.2 Saran

Berdasarkan hasil dan kesimpulan penelitian, beberapa saran dapat disampaikan untuk pengembangan penelitian dan pemanfaatan hasil di masa mendatang. Pertama, penelitian selanjutnya disarankan untuk memperluas cakupan wilayah studi dan memperkaya jenis data yang digunakan, seperti data curah hujan resolusi tinggi, data hidrodinamika sungai

secara real-time, serta informasi rinci mengenai sistem drainase perkotaan. Pengayaan data diharapkan dapat meningkatkan akurasi dan generalisasi model pada berbagai kondisi wilayah. Kedua, pengembangan model dapat diarahkan pada integrasi pendekatan temporal untuk menganalisis dinamika risiko banjir dari waktu ke waktu, sehingga tidak hanya memetakan risiko statis, tetapi juga mampu mendukung sistem peringatan dini banjir perkotaan. Pendekatan ini akan meningkatkan relevansi model dalam konteks kesiapsiagaan dan respon bencana. Ketiga, hasil penelitian ini berpotensi dikembangkan lebih lanjut ke dalam bentuk sistem pendukung keputusan kebencanaan berbasis explainable AI yang dapat digunakan oleh pemerintah daerah dan instansi terkait. Integrasi hasil analisis risiko banjir dengan perencanaan tata ruang, pengelolaan infrastruktur drainase, dan strategi adaptasi perubahan iklim diharapkan dapat meningkatkan efektivitas upaya mitigasi dan pengurangan risiko banjir di wilayah perkotaan.

DAFTAR PUSTAKA

1. Huang W, Park E, Wang J, Sophal T. The changing rainfall patterns drive the growing flood occurrence in urban areas. *Journal of Hydrology: Regional Studies*. 2024;45:101945.
2. Kardhana H, Valerian JR, Rohmat F, Kusuma M. Improving extreme urban flood prediction using satellite-based LSTM models. *Water*. 2022;14(9):1469.
3. Hu C, Xia J, She D, Song Z, Zhang Y. Urban hydrological modeling considering land cover variability for flood simulation. *Journal of Hydrology*. 2021;603:126833.
4. Singha C, Rana V, Pham QB, Nguyen D. Integrating machine learning and geospatial data for urban flood hazard assessment. *Environmental Science and Pollution Research*. 2024;31:48497–48522.
5. Ahmad I, Farooq R, Ashraf M, Waseem M. Flood hazard susceptibility assessment using ensemble machine learning. *Natural Hazards*. 2025;118:1–23.
6. Wang Y, Zhang P, Xie Y, Chen L. Toward explainable flood risk prediction using hybrid machine learning models. *Sustainable Cities and Society*. 2025;104:106140.
7. Pradhan B, Lee S, Dikshit A, Kim H. Spatial flood susceptibility mapping using explainable artificial intelligence. *Geoscience Frontiers*. 2023;14(4):101625.
8. Zhu X, Guo H, Huang JJ. Urban flood susceptibility mapping using remote sensing and ensemble learning. *Sustainable Cities and Society*. 2024;97:105508.
9. Qing W, Zhang H, Chen Y, Yin H. City-scale urban flooding risk assessment using multi-source data and machine learning. *Journal of Hydrology*. 2024;629:132626.
10. Ponce-Bobadilla A, Schmitt V, Maier C. Practical guide to SHAP analysis for explaining machine learning models. *Clinical and Translational Science*. 2024;17(3):70056.
11. Wibowo A, Ruslanjari D, Surahmat A, Karyaningsih D, Vera N. The MxT Model: Leveraging Social Media Data for Real-Time Route Optimization in Disaster-Prone Urban Transport Networks. *International Journal of Transport Development and Integration*. 2024;8(4):587–594. DOI:10.18280/ijtdi.080410.
12. Putra R, Nugroho A, Santosa B. Penerapan random forest untuk pemetaan risiko banjir perkotaan. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*. 2021;5(4):678–686.
13. Sari DP, Kurniawan D. Analisis risiko banjir berbasis SIG dan machine learning. *Jurnal Ilmu Komputer dan Informasi*. 2020;13(2):85–94.
14. Hidayat R, Pratama Y. Model prediksi banjir menggunakan data hidrometeorologi dan klasifikasi ensemble. *Jurnal Teknologi Informasi dan Ilmu Komputer*. 2021;8(6):1123–1132.

Lampiran 1. Realisasi Penggunaan Anggaran

Dana Disetujui: Rp 13.750.000

No	Jenis Pembelanjaan	Komponen	Item	Kuantitas	Biaya Satuan (Rp)	Total (Rp)
1	Belanja Bahan	ATK	ATK penelitian	1 paket	550.000	550.000
2	Pengumpulan Data	Honor pembantu peneliti	Asisten pengumpulan & integrasi data	1 orang	1.000.000	1.000.000
3	Pengumpulan Data	Konsumsi	Konsumsi kegiatan teknis	2 kali	150.000	300.000
4	Analisis Data	Honor pengolah data	Pengolahan data & pemodelan AI	1 orang	800.000	800.000
5	Sewa Peralatan	Software penelitian	Lisensi GIS, AI, cloud computing, SHAP	1 paket	2.975.000	2.975.000
6	Lainnya	Biaya penerbitan artikel	Jurnal nasional terakreditasi SINTA 2	1 artikel	4.000.000	4.000.000
7	Honorarium Peneliti	Honorarium peneliti	Maks. 30% dari pagu	1 paket	4.125.000	4.125.000
	TOTAL					13.750.000

Lampiran 2. Biodata Ketua dan Anggota Tim Peneliti

A. Identitas Diri

1. Nama Lengkap (dengan gelar) : Prof. Dr. Ir. Arief Wibowo, M.Kom
2. Jenis Kelamin : Laki - Laki
3. Jabatan Fungsional : Guru Besar
4. NIP/NIDN/ID-SINTA : 020004/0007097901/259862
5. Tempat, Tanggal Lahir : Jakarta, 7 September 1979
6. E-mail : arief.wibowo@budiluhur.ac.id
7. Nomor Handphone : 0818120515
8. Alamat : Jl. Palem II/12 Petukangan Utara
Jakarta Selatan

B. Riwayat Pendidikan

	S1	S2	S3
Nama Perguruan Tinggi	STMIK Budi Luhur	Univ. Budi Luhur	Univ. Gadjah Mada
Bidang Ilmu	Sistem Informasi	Ilmu Komputer	Ilmu Komputer
Tahun Masuk-Lulus	1997-2021	2004-2006	2012-2018

C. Pengalaman Penelitian (5 Tahun Terakhir)

No.	Tahun	Judul Penelitian	Pendanaan	
			Sumber*	Jumlah (Rp)
1.	2025	Model Integrasi Data Spasial Dan Data Media Sosial Dengan Analisis Berbasis Machine Learning Untuk Mendukung Kegiatan Search And Rescue (SAR)	Kemdiktisainstek	Rp 185.370.000,-
2.	2025	Pemodelan Tata Kelola Terintegrasi Geopark Bayah Dome untuk Peningkatan Ekonomi Berkelanjutan di Kabupaten Lebak dengan Aplikasi Cerdas Geowisata berbasis Web	Kemdiktisainstek	Rp 119.780.000,-
3.	2024	Model Pengambilan Keputusan Dalam Diagnosis Penyakit Diabetes Dengan Algoritma Data Mining pada Kasus Data Tidak Seimbang	Kemristekdikti	Rp 27.650.000,-
4.	2024	Analisis Data Media Sosial Untuk Pemodelan Prediktif dan Visualisasi Peta Jalur Mitigasi Bencana di Perkotaan Berbasis Web	Kemristekdikti	Rp 96.770.000,-
5.	2023	Algoritme K-Means Untuk Klasterisasi Lulusan Perguruan Tinggi Berbasis Hasil Studi Pelacakan Pada Masa Pandemi Covid-19	Univ. Budi Luhur	Rp 10.000.000,-

No.	Tahun	Judul Penelitian	Pendanaan	
			Sumber*	Jumlah (Rp)
6.	2020	Model Pengelompokan Wilayah Dengan Kasus Positif Covid-19 di DKI Jakarta Menggunakan Metode Data Mining Klasterisasi	Univ. Budi Luhur	Rp 15.000.000,-

* Tuliskan sumber pendanaan baik dari Universitas Budi Luhur maupun dari sumber lainnya.

D. Publikasi Artikel Ilmiah dalam Jurnal (5 Tahun Terakhir)

No.	Judul Artikel Ilmiah*	Nama Jurnal	Vol/No/ Tahun
1.	Leveraging artificial intelligence in disaster management: A comprehensive bibliometric review	Jamba Journal of Disaster Risk Studies (Q2)	17/1/2025
2.	Refining Diabetes Diagnosis Models: The Impact of SMOTE on SVM, Logistic Regression, and Naïve Bayes for Imbalanced Datasets	Journal of Electronics Electromedical Engineering and Medical Informatics (Q4)	7/1/2025
3.	Usability of Forecasting Technique for Tackling Stock Price Action Decision	Icocseti 2025 International Conference on Computer Sciences Engineering and Technology Innovation	2025
4.	Clustering Indonesian Provinces by Disaster Intensity using K-Means Algorithm: a Data Mining Approach	Disaster Advances (Q3)	7/12/2024
5.	Regional Clustering Model of Covid-19 Cases at the Early of the Pandemic in Indonesia	International Journal of Intelligent Systems and Applications in Engineering (Q4)	12/1/2024
6.	Designing and Evaluating Search and Rescue Management Applications using the Rapid Application Development Approach and the Technology Acceptance Model	International Journal of Mathematics and Computer Science (Q2)	19/3/2024
7.	Refining Diabetes Diagnosis Models: The Impact of SMOTE on SVM, Logistic Regression, and Naïve Bayes	Journal of Electronics, Electromedical Engineering, and Medical Informatics	17/1/2025
8.	The MxT Model: Leveraging Social Media Data for Real-Time Route Optimization in Disaster-Prone Urban Transport Networks	International Journal of Transport Development and Integration (Q3)	8/4/2024

* Artikel ilmiah sebagai luaran dari kegiatan penelitian

E. Pemakalah Seminar Ilmiah (5 Tahun Terakhir)

No.	Nama Temu Ilmiah/Seminar	Judul Artikel Ilmiah*	Waktu dan Tempat
1.	-	-	-

** Artikel ilmiah sebagai luaran dari kegiatan penelitian*

F. Perolehan HKI (5 Tahun Terakhir)

No.	Judul/Tema HKI*	Tahun	Jenis	Nomor P/ID
1.	Sistem Informasi Cerdas Terintegrasi Untuk Mendukung Perencanaan dan Pelaksanaan Operasi Pencarian dan Pertolongan	2025	Paten Sederhana	S00202511651
2.	Kamishibai Nusantara: Media Seni Visual Mural untuk Literasi Pengurangan Risiko Bencana Berbasis Pemberdayaan Masyarakat dan Komunitas	2025	Hak Cipta	EC002025211694
3.	Aplikasi Luaran Penelitian: Analisis Data Media Sosial Untuk Pemodelan Prediktif Dan Visualisasi Peta Jalur Mitigasi Bencana Di Perkotaan Berbasis Web	2024	Hak Cipta	EC002024217100
4.	Aplikasi Log Aktivitas Potensi Sar (AplotSAR) Untuk Potensi SAR Dan Rescuer Basarnas	2023	Hak Cipta	EC002023125993

** HKI sebagai luaran dari kegiatan penelitian*

Jakarta, 16 Januari 2026
Ketua Peneliti,


Prof. Dr. Ir. Arief Wibowo, M.Kom.

A. Identitas Diri

1. Nama Lengkap (dengan gelar) : Dr. Abdul Haris Achadi, SH, DESS
2. Jenis Kelamin : Laki - Laki
3. Jabatan Fungsional : Tenaga Pengajar
4. NIP/NIDN/ID-SINTA : 220051/8955170023/ 684843
5. Tempat, Tanggal Lahir : Surabaya, 24 September 1966
6. E-mail : abdul.harisachadi@budiluhur.ac.id
7. Nomor Handphone : 0812-8643-9694
8. Alamat : Pabuaran Asri Blok A8 no 6
Pabuaran Mekar, Cibinong, Kab. Bogor

B. Riwayat Pendidikan

	S1	S2	S3
Nama Universitas	Universitas Surabaya	Universite D aix Marseille III	UGM
Bidang Ilmu	Hukum Internasional	Magister Transportasi	Inovasi Publik
Tahun Masuk-Lulus	1986-1990	1993-1995	2018-2021

C. Pengalaman Penelitian (5 Tahun Terakhir)

No.	Tahun	Judul Penelitian	Pendanaan	
			Sumber*	Jumlah (Rp)
1.	-	-	-	-

* Tuliskan sumber pendanaan baik dari Universitas Budi Luhur maupun dari sumber lainnya.

D. Publikasi Artikel Ilmiah dalam Jurnal (5 Tahun Terakhir)

No.	Judul Artikel Ilmiah*	Nama Jurnal	Vol/No/ Tahun
1.	Two Decades Of Bibliometric Exploration On Leadership In Disaster Management	Disaster Advances (Q3)	17/6/2024
2.	Ketahanan Masyarakat Penyintas Pasca Gempabumi Cianjur	Jurnal Ketahanan Nasional (S2)	30/1/2024
3.	Designing and Evaluating Search and Rescue Management Applications using the Rapid Application Development Approach and the Technology Acceptance Model	International Journal of Mathematics and Computer Science (Q2)	19/3/2024
4.	Adaptasi Masyarakat Penyintas Pasca Gempa Cianjur Tahun 2022	Jurnal Manajemen Bencana (S5)	9/2/2023
5.	Leadership Coordinator of Search and Rescue Mission on Earthquake Event in Palu Central Sulawesi	Disaster Advances (Q3)	14/2/2021

* Artikel ilmiah sebagai luaran dari kegiatan penelitian

E. Pemakalah Seminar Ilmiah (5 Tahun Terakhir)

No.	Nama Temu Ilmiah/Seminar	Judul Artikel Ilmiah*	Waktu dan Tempat
1.	-	-	-

* Artikel ilmiah sebagai luaran dari kegiatan penelitian

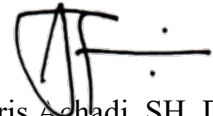
F. Perolehan HKI (5 Tahun Terakhir)

No.	Judul/Tema HKI*	Tahun	Jenis	Nomor P/ID
1.	Aplikasi Log Aktivitas Potensi Sar (AplotSAR) Untuk Potensi SAR Dan Rescuer Basarnas	2023	Hak Cipta	EC002023125993

* HKI sebagai luaran dari kegiatan penelitian

Jakarta, 16 Januari 2026

Anggota Peneliti,



Dr. Abdul Haris Achadi, SH, DESS

Lampiran 3. Surat Perjanjian Kontrak Penelitian

Lampiran 4. Catatan Harian

No	Tanggal	Kegiatan
1	4 Nov 2025	Kick-off internal tim penelitian, peninjauan kembali proposal, dan penyelarasan rencana kerja detail sesuai jadwal penelitian.
2	6 Nov 2025	Studi literatur lanjutan terkait risiko banjir perkotaan, ensemble learning, dan Explainable Artificial Intelligence (XAI).
3	9 Nov 2025	Identifikasi variabel penelitian dan perumusan kebutuhan analisis risiko banjir berbasis data multi-sumber.
4	12 Nov 2025	Pengumpulan data hidrometeorologi (curah hujan dan tinggi muka air sungai) dari sumber resmi.
5	16 Nov 2025	Pengumpulan data geospasial berupa DEM, tata guna lahan, dan jaringan sungai wilayah studi.
6	20 Nov 2025	Ekstraksi indeks vegetasi (NDVI) dari citra penginderaan jauh dan verifikasi kualitas data.
7	24 Nov 2025	Penyamaan sistem koordinat dan integrasi data multi-sumber ke dalam basis data spasial.
8	28 Nov 2025	Penyelarasan resolusi spasial dan pembentukan grid analisis untuk pemodelan risiko banjir.
9	2 Des 2025	Praproses data: pembersihan data, penanganan missing values, dan normalisasi variabel.
10	6 Des 2025	Rekayasa fitur: perhitungan persentase area terbangun, kepadatan drainase, dan variabel turunan lainnya.
11	9 Des 2025	Pembagian dataset ke dalam data latih dan data uji sesuai skema penelitian.
12	12 Des 2025	Pelatihan awal model Random Forest untuk analisis risiko banjir perkotaan.
13	15 Des 2025	Pelatihan model Gradient Boosting dan evaluasi kinerja awal.
14	18 Des 2025	Pelatihan dan penyetelan parameter model XGBoost.
15	21 Des 2025	Evaluasi performa model menggunakan metrik AUC, akurasi, F1-score, dan RMSE.
16	24 Des 2025	Pelaksanaan validasi silang (cross-validation) untuk menguji kestabilan model.
17	27 Des 2025	Analisis Explainable AI menggunakan metode SHAP untuk identifikasi faktor dominan risiko banjir.
18	30 Des 2025	Analisis interaksi antar variabel dan interpretasi hasil kontribusi SHAP.
19	3 Jan 2026	Penyusunan peta risiko banjir perkotaan berbasis hasil prediksi dan faktor dominan.
20	7 Jan 2026	Penyusunan draft artikel ilmiah untuk jurnal nasional terakreditasi SINTA 2.
21	10 Jan 2026	Finalisasi naskah artikel ilmiah dan proses submit ke jurnal tujuan.
22	14 Jan 2026	Penyusunan laporan akhir penelitian dan kompilasi seluruh luaran penelitian.
23	18 Jan 2026	Review internal laporan akhir penelitian dan perbaikan redaksional.
24	21 Jan 2026	Finalisasi laporan akhir penelitian dan kelengkapan dokumen pendukung.

Lampiran 4. Artikel Ilmiah (accepted)

Jurnal Teknik Informatika

Fakultas Sains dan Teknologi
Universitas Islam negeri Syarif Hidayatullah Jakarta

p-ISSN 1979-9160 | e-ISSN 2549-7901

No : 07/JTI/PPN/01/ 49010

Jakarta, 09 Januari 2026

Hal : Surat Penerimaan Naskah

Yang Terhormat,

Bapak/Ibu **Arief Wibowo, Abdul Haris Achadi**

Di tempat

Assalamu'alaikum Wr. Wb.

Dengan hormat,

Bersama ini kami sampaikan bahwa naskah artikel saudara yang berjudul **Explainable Ensemble Learning for Urban Flood Risk Analysis Using Multi-Source Data In Indonesia** telah dibaca dan ditelaah oleh Mitra Bestari dan Dewan Redaksi yang ditunjuk, dan dinyatakan **"DITERIMA"** dalam jurnal kami.

Berdasarkan ketersediaan naskah yang kami miliki saat ini, direncanakan naskah artikel tersebut akan dimuat dalam **Volume 19, Nomor 1 Tahun 2026 (April) Jurnal Teknik Informatika (JTI)**

Demikian surat ini kami sampaikan agar menjadi maklum.

Wassalamu'alaikum Wr. Wb.

Pimpinan Redaksi

Jurnal
Teknik Informatika
Universitas Islam Negeti Syarif Hidayatullah Jakarta

Dr. Eng. Feri Fahrianto, S.T., M.Sc

NIP: 198008292011011002

Lampiran 5. HKI


REPUBLIK INDONESIA
KEMENTERIAN HUKUM

SURAT PENCATATAN CIPTAAN

Dalam rangka perlindungan ciptaan di bidang ilmu pengetahuan, seni dan sastra berdasarkan Undang-Undang Nomor 28 Tahun 2014 tentang Hak Cipta, dengan ini menerangkan:

Nomor dan tanggal permohonan : EC002026002808, 6 Januari 2026

Pencipta
Nama : **Arief Wibowo dan Abdul Haris Achadi**
Alamat : Universitas Budi Luhur Jakarta, Pesanggrahan, Kota Adm. Jakarta Selatan, DKI Jakarta, 12260
Kewarganegaraan : Indonesia

Pemegang Hak Cipta
Nama : **Universitas Budi Luhur**
Alamat : Jl.Cileduk Raya, Pesanggrahan, Kota Adm. Jakarta Selatan, DKI Jakarta, 12260
Kewarganegaraan : Indonesia
Jenis Ciptaan : **Laporan Penelitian**
Judul Ciptaan : **Explainable Ensemble Learning untuk Analisis Faktor Dominan Risiko Banjir Perkotaan Berbasis Data Multi-Sumber di Indonesia**

Tanggal dan tempat diumumkan untuk pertama kali di wilayah Indonesia atau di luar wilayah Indonesia : 6 Januari 2026, di Kota Adm. Jakarta Selatan

Jangka waktu perlindungan : Berlaku selama 50 (lima puluh) tahun sejak Ciptaan tersebut pertama kali dilakukan Pengumuman.

Nomor Pencatatan : 001073767

adalah benar berdasarkan keterangan yang diberikan oleh Pemohon.
Surat Pencatatan Hak Cipta atau produk Hak terkait ini sesuai dengan Pasal 72 Undang-Undang Nomor 28 Tahun 2014 tentang Hak Cipta.

a.n. MENTERI HUKUM
DIREKTUR JENDERAL KEKAYAAN INTELEKTUAL
u.b
Direktur Hak Cipta dan Desain Industri


Agung Damarsasongko,SH.,MH.
NIP. 196912261994031001





Disclaimer:

1. Dalam hal pemohon memberikan keterangan tidak sesuai dengan surat pernyataan, Menteri berwenang untuk mencabut surat pencatatan permohonan.
2. Surat Pencatatan ini telah disegel secara elektronik menggunakan segel elektronik yang diterbitkan oleh Balai Besar Sertifikasi Elektronik, Badan Siber dan Sandi Negara.
3. Surat Pencatatan ini dapat dibuktikan keasliannya dengan memindai kode QR pada dokumen ini dan informasi akan ditampilkan dalam browser.