



## ISIAN SUBSTANSI PROPOSAL PENELITIAN

*Petunjuk: Pengusul hanya diperkenankan mengisi di tempat yang telah disediakan sesuai dengan petunjuk pengisian dan tidak diperkenankan melakukan modifikasi template atau penghapusan di setiap bagian.*

### JUDUL

Tuliskan judul penelitian.

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

### RINGKASAN

Isian ringkasan penelitian tidak lebih dari 300 kata yang berisi urgensi, tujuan, metode, dan luaran yang ditargetkan.

**Urgensi** penelitian ini didorong oleh tingginya frekuensi dan dampak banjir perkotaan di Indonesia yang dipengaruhi oleh interaksi kompleks antara faktor hidrometeorologi dan perubahan lingkungan perkotaan. Upaya mitigasi banjir membutuhkan pemahaman yang komprehensif tidak hanya terhadap tingkat risiko, tetapi juga terhadap faktor-faktor dominan yang memengaruhi terjadinya banjir. Pendekatan analisis risiko banjir berbasis machine learning yang ada umumnya masih bersifat black-box, sehingga sulit digunakan sebagai dasar pengambilan keputusan oleh pemangku kebijakan. **Tujuan** penelitian ini adalah mengembangkan pendekatan Explainable Ensemble Learning untuk menganalisis faktor dominan risiko banjir perkotaan berbasis data multi-sumber di Indonesia, serta menghasilkan model yang tidak hanya akurat tetapi juga mampu memberikan penjelasan yang transparan mengenai kontribusi masing-masing variabel penyebab banjir. **Metode** yang digunakan dalam penelitian ini meliputi integrasi data multi-sumber berupa data geospasial dan hidrometeorologi, antara lain curah hujan, elevasi, tata guna lahan, indeks vegetasi (NDVI), jarak terhadap sungai, dan tinggi muka air. Analisis risiko banjir dilakukan menggunakan model ensemble learning yang terdiri dari Random Forest, Gradient Boosting, dan XGBoost untuk menangkap hubungan nonlinier antar variabel. Aspek keterjelasan model dicapai melalui penerapan Explainable Artificial Intelligence menggunakan metode SHapley Additive exPlanations (SHAP) untuk mengidentifikasi faktor-faktor dominan yang memengaruhi risiko banjir. Evaluasi model dilakukan menggunakan metrik kinerja seperti Area Under the Curve (AUC), akurasi, dan Root Mean Square Error (RMSE), serta validasi silang untuk memastikan kestabilan model. **Luaran** yang ditargetkan dari 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 ditargetkan terbit pada jurnal nasional terakreditasi SINTA 2, yaitu Jurnal Teknik Informatika.

### KATA KUNCI

Isian kata kunci maksimal 5 kata yang dipisahkan dengan tanda titik koma (;)

Banjir Perkotaan; Explainable AI; Ensemble Learning; Data Multi-Sumber; Risiko Bencana.

## PENDAHULUAN

Pendahuluan penelitian tidak lebih dari 1000 kata yang terdiri dari:

1. Latar belakang dan rumusan permasalahan yang akan diteliti
2. Pendekatan pemecahan masalah
3. *State of the art* dan kebaruan
4. Peta jalan (*road map*) penelitian 5 tahun

*Sitasi disusun dan ditulis berdasarkan sistem nomor sesuai dengan urutan pengutipan, mengikuti format Vancouver*

### 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].

### 2. PENDEKATAN PEMECAHAN MASALAH

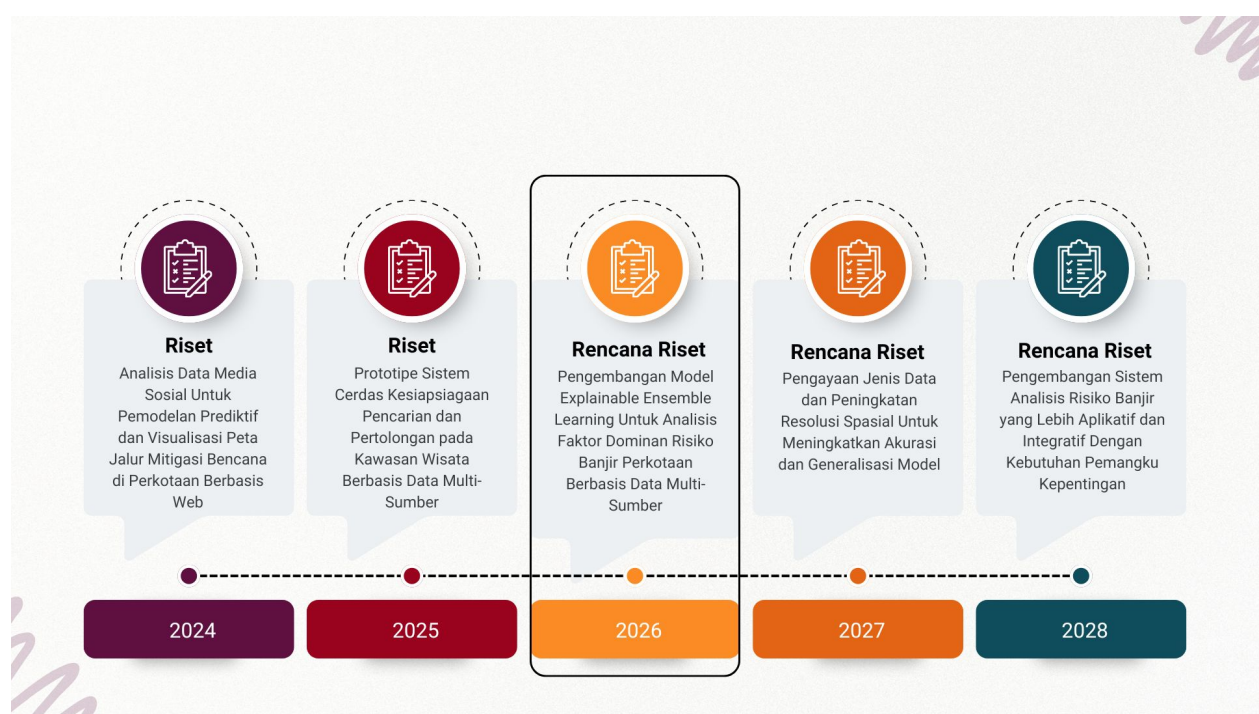
**Pendekatan pemecahan masalah** dalam penelitian ini dilakukan melalui pengembangan model Explainable Ensemble Learning berbasis integrasi data multi-sumber. Data yang digunakan meliputi data geospasial dan hidrometeorologi, seperti curah hujan, elevasi, tata guna lahan, indeks vegetasi (NDVI), jarak 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.

### 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].

### 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].

## **METODE**

Isian metode atau cara untuk mencapai tujuan yang telah ditetapkan ditulis tidak melebihi 1000 kata. Bagian ini dapat dilengkapi dengan diagram alir penelitian yang menggambarkan apa yang sudah dilaksanakan dan yang akan dikerjakan selama waktu yang diusulkan. Metode penelitian harus dibuat secara utuh dengan penahapan yang jelas, mulai dari awal bagaimana proses dan luarannya, dan indikator capaian yang ditargetkan yang tercermin dalam Rencana Anggaran Biaya (RAB).
--

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, sebagaimana terlihat pada Gambar 2.



Gambar 2. Tahapan Penelitian

Berdasarkan Gambar 2, maka penelitian ini terdiri dari 6 tahap, yaitu:

### **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.

## **JADWAL PENELITIAN**

Jadwal penelitian disusun berdasarkan pelaksanaan penelitian, ditulis dengan mengisi langsung tabel berikut dan diperbolehkan menambahkan baris sesuai banyaknya jenis kegiatan.

No.	Nama Kegiatan	Tahun 2025-2026					
		Okt	Nov	Des	Jan	Feb	Mar
1.	Studi Literatur dan Perumusan Kebutuhan Analisis	√					
2.	Pengumpulan dan Integrasi Data Multi-Sumber		√				
3.	Praproses Data dan Rekayasa Fitur		√				
4.	Pengembangan Model Explainable Ensemble Learning		√	√			
5.	Analisis Explainable AI dan Identifikasi Faktor Dominan			√			
6.	Evaluasi, Visualisasi Hasil, dan Penyusunan Artikel Ilmiah			√			
7.	Submission artikel Ilmiah				√	√	
8.	Pelaporan Penelitian						√

#### LUARAN DAN TARGET CAPAIAN

Tuliskan target luaran wajib dan tambahan (jika ada) yang akan dihasilkan.

No.	Kategori Luaran	Jenis Luaran	Target Capaian
1.	Artikel Ilmiah	Artikel bereputasi nasional terindeks SINTA 2 Jurnal Teknik Informatika <a href="https://journal.uinjkt.ac.id/ti/index">https://journal.uinjkt.ac.id/ti/index</a>	published

#### RENCANA ANGGARAN BIAYA

**Total RAB: Rp 13.750.000,-**

Jenis Pembelanjaan	Komponen	Item	Kuantitas	Biaya Satuan (Rp)	Total (Rp)
Belanja Bahan	ATK	ATK penelitian	1 paket	550.000	550.000
Pengumpulan Data	Honor pembantu peneliti	Asisten pengumpulan & integrasi data	1 orang	1.000.000	1.000.000
Pengumpulan Data	Konsumsi	Konsumsi kegiatan teknis	2 kali	150.000	300.000
Analisis Data	Honor pengolah data	Pengolahan data & pemodelan AI	1 orang	800.000	800.000

Sewa Peralatan	Software penelitian	Lisensi GIS, AI, cloud computing, SHAP	1 paket	2.975.000	2.975.000
Lainnya	Biaya proofreading dan penerbitan artikel	Jurnal nasional SINTA 2 (JTI-UIN Syarif Hidayatullah Jakarta)	1 artikel	4.000.000	4.000.000
Honorarium Peneliti	Honorarium peneliti	30% dari pagu	1 paket	4.125.000	4.125.000
<b>Jumlah Anggaran yang diusulkan</b>					<b>13.750.000</b>

## DAFTAR PUSTAKA

Sitasi disusun dan ditulis berdasarkan sistem nomor sesuai dengan urutan pengutipan, mengikuti format Vancouver. Sumber pustaka mengutamakan hasil penelitian pada jurnal ilmiah yang terkini (maksimal 5 tahun terakhir). Hanya pustaka yang disitasi pada usulan penelitian yang dicantumkan dalam Daftar Pustaka.

- 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.
- Kardhana H, Valerian JR, Rohmat F, Kusuma M. Improving extreme urban flood prediction using satellite-based LSTM models. *Water*. 2022;14(9):1469.
- 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.
- 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.
- Ahmad I, Farooq R, Ashraf M, Waseem M. Flood hazard susceptibility assessment using ensemble machine learning. *Natural Hazards*. 2025;118:1–23.
- 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.
- Pradhan B, Lee S, Dikshit A, Kim H. Spatial flood susceptibility mapping using explainable artificial intelligence. *Geoscience Frontiers*. 2023;14(4):101625.
- Zhu X, Guo H, Huang JJ. Urban flood susceptibility mapping using remote sensing and ensemble learning. *Sustainable Cities and Society*. 2024;97:105508.
- 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.
- 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.
- 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.