

**PENGEMBANGAN SISTEM PREDIKSI POTENSI  
KEBAKARAN HUTAN DAN LAHAN DI  
INDONESIA BERBASIS *MACHINE LEARNING*  
DAN XAI**

**Proposal Tugas Akhir**

Oleh

**Moh Afnan Fawaz  
18222111**



**PROGRAM STUDI SISTEM DAN TEKNOLOGI INFORMASI  
SEKOLAH TEKNIK ELEKTRO DAN INFORMATIKA  
INSTITUT TEKNOLOGI BANDUNG  
Desember 2025**

## **LEMBAR PENGESAHAN**

# **PENGEMBANGAN SISTEM PREDIKSI POTENSI KEBAKARAN HUTAN DAN LAHAN DI INDONESIA BERBASIS *MACHINE LEARNING* DAN XAI**

## **Proposal Tugas Akhir**

Oleh

**Moh Afnan Fawaz  
18222111**

Program Studi Sistem dan Teknologi Informasi  
Sekolah Teknik Elektro dan Informatika  
Institut Teknologi Bandung

Proposal Tugas Akhir ini telah disetujui dan disahkan  
di Bandung, pada tanggal 5 Desember 2025

Pembimbing

Dr. Ir. Dimitri Mahayana, M.Eng.  
NIP. 19680809 199102 1 001

## DAFTAR ISI

<b>DAFTAR GAMBAR . . . . .</b>	<b>iv</b>
<b>DAFTAR TABEL . . . . .</b>	<b>v</b>
<b>I PENDAHULUAN . . . . .</b>	<b>1</b>
I.1 Latar Belakang . . . . .	1
I.2 Rumusan Masalah . . . . .	3
I.3 Tujuan . . . . .	4
I.4 Batasan Masalah . . . . .	5
I.5 Metodologi . . . . .	5
<b>II STUDI LITERATUR . . . . .</b>	<b>8</b>
II.1 Kebakaran Hutan dan Lahan (Karhutla) di Indonesia . . . . .	8
II.2 Faktor Pemicu Karhutla . . . . .	9
II.2.1 Faktor Meteorologis . . . . .	9
II.2.2 Faktor Antropogenik . . . . .	9
II.2.3 Faktor Biofisik . . . . .	10
II.3 Indeks Cuaca Kebakaran . . . . .	10
II.4 <i>Artificial Intelligence</i> . . . . .	11
II.5 <i>Machine Learning</i> . . . . .	13
II.5.1 Konsep <i>Supervised Learning</i> . . . . .	14
II.5.2 <i>Linear Regression</i> . . . . .	15
II.5.3 <i>Random Forest</i> . . . . .	15
II.5.4 <i>Extreme Gradient Boosting</i> . . . . .	16
II.6 Metrik Evaluasi . . . . .	17
II.6.1 <i>Confusion Matrix</i> . . . . .	18
II.6.2 Kurva ROC dan Area AUC . . . . .	19
II.6.3 Metrik Regresi . . . . .	20
II.7 <i>Explainable Artificial Intelligence (XAI)</i> . . . . .	21
II.7.1 SHAP . . . . .	22
II.7.2 <i>Permutation Feature Importance</i> . . . . .	23
II.8 Sistem Informasi Geospasial . . . . .	24
II.9 Penelitian Terkait . . . . .	25
<b>III ANALISIS MASALAH . . . . .</b>	<b>27</b>
III.1 Analisis Kondisi . . . . .	27

III.1.1	<i>Business Understanding</i>	27
III.1.1.1	Kondisi Saat Ini	27
III.1.2	Kebutuhan Fungsional	30
III.1.3	Kebutuhan Non-fungsional	31
III.2	Analisis Pemilihan Solusi	32
III.2.1	Alternatif Solusi	32
III.2.2	Analisis Penentuan Solusi	34
III.3	<i>Data Understanding</i>	36
III.3.1	Deskripsi Umum <i>Dataset</i>	36
III.3.2	Kelengkapan dan Distribusi Data	37
<b>IV</b>	<b>DESAIN KONSEP SOLUSI</b>	<b>40</b>
<b>V</b>	<b>RENCANA SELANJUTNYA</b>	<b>44</b>
V.1	Rencana implementasi	44
V.1.1	Alur Implementasi Teknis	44
V.1.2	Lingkungan Pengembangan dan Daftar Alat	46
V.1.3	Linimasa Penggerjaan	48
V.2	Rencana Verifikasi dan Evaluasi	48
V.2.1	Verifikasi Model Prediksi dan XAI	49
V.2.1.1	Verifikasi Model Prediksi	49
V.2.1.2	Verifikasi XAI	49
V.2.2	Validasi Model Prediksi dan XAI	50
V.2.2.1	Validasi Model Prediksi	50
V.2.2.2	Validasi hasil XAI	50
V.3	Analisis Risiko dan Mitigasi	51

## DAFTAR GAMBAR

I.1	Karhutla di Indonesia pada tahun 2015-2024 (Kementerian Lingkungan Hidup dan Kehutanan (KLHK) 2025) . . . . .	1
I.2	Metodologi CRISP-DM . . . . .	6
II.1	Ilustrasi posisi hierarki ML dan DL pada area AI (Sarker 2022) . . . . .	12
II.2	Struktur umum ML pada fase pelatihan dan pengujian (Sarker 2022)	13
II.3	Stuktur model (Random Forest) (Sarker 2021) . . . . .	16
II.4	Contoh penerapan <i>confusion matrix</i> pada kasus label data karhutla .	18
II.5	Kurva ROC dan area ROC . . . . .	19
II.6	Konsep skema XAI . . . . .	22
II.7	Contoh <i>layer GIS</i> (GAO 2003) . . . . .	25
III.1	Konsep sistem penanganan karhutla di Indonesia saat ini . . . . .	29
III.2	Jumlah <i>hotspot</i> per tahun (VIIRS NOAA-20)	38
III.3	Data persebaran fitur <i>confidence</i> . . . . .	39
IV.1	Model konseptual <i>to-be</i> . . . . .	40
IV.2	Sistem prediktif dengan ML, XAI, dan penyajian informasi visual .	41
V.1	Diagram alur rencana implementasi . . . . .	44

## DAFTAR TABEL

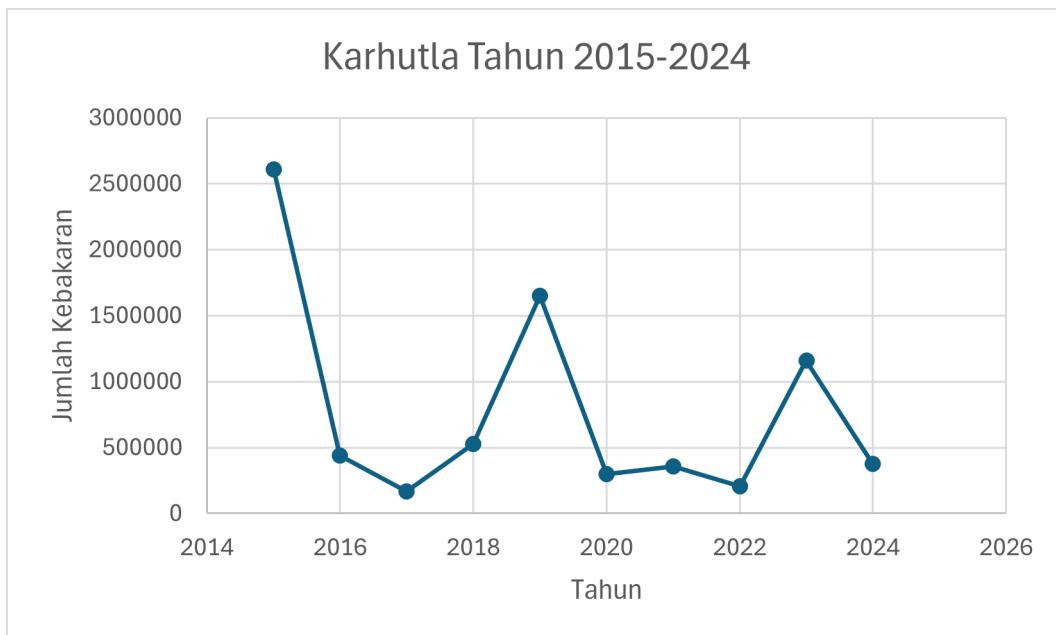
II.1	Kategori <i>Machine Learning</i> (Sarker 2022) . . . . .	14
III.1	Kebutuhan Fungsional Sistem (versi ringkas dan problem-driven) . .	31
III.2	Kebutuhan non-fungsional sistem . . . . .	32
III.3	Matriks penilaian alternatif solusi menggunakan WSM . . . . .	35
III.4	Deskripsi atribut pada <i>dataset VIIRS 375 m</i> . . . . .	37
IV.1	Perbandingan Ringkas Sistem <i>As-Is</i> dan <i>To-Be</i> . . . . .	43
V.1	Lingkungan pengembangan dan daftar alat yang digunakan . . . . .	47
V.2	Rencana pemanfaatan layanan berbayar dan estimasi biaya . . . . .	48
V.3	Lini masa penyelesaian tugas akhir . . . . .	48
V.4	Identifikasi risiko dan rencana mitigasi . . . . .	51

# BAB I

## PENDAHULUAN

### I.1 Latar Belakang

Indonesia merupakan negara dengan kawasan hutan tropis terbesar ketiga di dunia dengan keanekaragaman hayatinya yang tergolong tinggi. Namun, ancaman terhadap kebakaran hutan kerap terjadi di Indonesia dan memberikan dampak terhadap ekologi, ekonomi, lingkungan, kesehatan, dan sosial (Karurung, Lee, dan Lee 2025). Meskipun tren karhutla telah mengalami penurunan sesuai yang ditunjukkan pada Gambar I.1, kebakaran hutan dan lahan (karhutla) tetap menjadi salah satu sumber utama emisi karbon. Selain itu, kejadian deforestasi hutan sering kali berhubungan erat dengan bencana karhutla (Vetrita dkk. 2025).



Gambar I.1 Karhutla di Indonesia pada tahun 2015-2024 (Kementerian Lingkungan Hidup dan Kehutanan (KLHK) 2025)

Kebakaran hebat terjadi terutama saat musim kemarau dan tahun-tahun kering. Berdasarkan data dari Kementerian Lingkungan Hidup dan Kehutanan (KLHK), Indonesia mengalami kebakaran hebat yang membakar 2,6 juta hektare hutan dan lahan yang dipicu oleh kekeringan akibat El Niño pada tahun 2015. Peristiwa tersebut menghasilkan emisi karbon dalam jumlah besar disertai kabut berkepanjangan di seluruh Indonesia dan negara tetangga dengan total kerugian ekonomi yang berkisar antara 1,62 hingga 2,7 miliar USD (United States Dollars) (Prasetyo dkk. 2022). Selain kerugian ekonomi, kebakaran hutan juga memberikan dampak yang besar bagi kesehatan dan lingkungan.

Sebagaimana yang dijelaskan pada Hein dkk. (2022), kebakaran hutan pada tahun 2015 menghasilkan partikel PM<sub>2,5</sub> (partikel halus berdiameter 2,5 mikrometer yang biasanya merupakan asap berbahaya) selama 24 jam dengan jumlah di atas 2.000 µg/m<sup>3</sup>. Partikel ini memberikan dampak buruk pada kesehatan seperti memperburuk asma, memicu gejala pernapasan, meningkatkan angka rawat inap, hingga berkontribusi terhadap kanker paru-paru. Pada jangka panjang, kebakaran hutan dapat merusak lahan serta menjadi bencana yang menimbulkan hilangnya biodiversitas dan menyebabkan pemanasan global (Thoha dan Triani 2021).

Menurut studi dari Karurung, Lee, dan Lee (2025), penyebab kebakaran hutan dapat beragam mulai dari fase ENSO (*El Niño-Southern Oscillation*), curah hujan di sekitar Juli hingga September, faktor meteorologi dan topografi, hingga aktivitas manusia. Meninjau dari beberapa penyebab tersebut, pemerintah dan beberapa penelitian telah melakukan usaha untuk melakukan penanganan dan pencegahan terhadap karhutla. Penelitian yang dilakukan oleh Qamariyanti, Usman, dan Rahmawati (2023) menunjukkan bahwa di Indonesia sudah terdapat beberapa upaya pencegahan yang dilakukan oleh pemerintah pada tahun 2020 seperti membuat peraturan dan pasal terhadap tindakan pembukaan lahan dengan kebakaran hutan. Selain itu, Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) juga telah membuat pendekatan identifikasi karhutla melalui perhitungan Fire Weather Index yang memanfaatkan faktor meteorologis dalam melakukan identifikasi. Walaupun begitu, Fire Weather Index yang dipandang sebagai perhitungan indeks kebakaran yang telah teruji secara internasional belum mengintegrasikan faktor-faktor lainnya yang menurut beberapa studi menjadi penyebab terjadinya bencana karhutla di Indonesia.

Di sisi lain, pemantauan karhutla di Indonesia sudah mulai bergantung pada data satelit *hotspot* harian dari produk Visible Infrared Imaging Radiometer Suite (VIIRS) 375 m. Produk ini kemudian diintegrasikan dengan sistem informasi kebakaran na-

sional. Namun, studi oleh Indradjad dkk. (January 2024) menemukan bahwa data VIIRS hanya menemukan sekitar 52-53% kejadian lapangan yang memiliki *hotspot*, sedangkan untuk kebakaran yang lebih besar dari 14 hektare tingkat deteksinya meningkat hingga sekitar 83%. Hal ini menunjukkan masih adanya beberapa kebakaran kecil dan menengah yang tidak terdeteksi oleh satelit.

Dalam beberapa tahun terakhir, teknologi *remote sensing* dan *geospatial* telah muncul sebagai alat untuk mendeteksi dan memetakan kebakaran menggunakan faktor alam. Bahkan, pengembangan teknologi telah mencapai pemanfaatan integrasi *Machine Learning* (ML) dalam melakukan pemantauan dengan lebih efektif. Akan tetapi, integrasi data satelit dengan ML atau *Deep Learning* (DL) untuk prediksi dan pemodelan kebakaran masih merupakan bidang yang relatif jarang dieksplorasi di wilayah Asia Tenggara dan memiliki potensi besar untuk dikembangkan (Eaturu dan Vadrevu 2025).

Selain itu, pemanfaatan model ML dalam tindakan prediktif kebakaran hutan sering kali hanya berfokus kepada hasil atau akurasi dari model. Padahal dalam memberikan informasi, alasan dalam penentuan hasil prediktif sangat diperlukan untuk memberikan wawasan yang signifikan. Kurangnya transparansi saat menggunakan model ML merupakan hambatan bagi para pengelola kebakaran hutan karena algoritma ini dianggap sebagai "*black box*" sehingga sulit memahami hasil yang diberikan oleh model (Abdollahi dan Pradhan 2023). Maka dari itu, pendekatan *Explainable Artificial Intelligence* (XAI) perlu diterapkan dalam konteks kebakaran hutan untuk menginterpretasikan kontribusi dari faktor-faktor terhadap hasil model.

## I.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan dapat diketahui bahwa saat ini langkah prediktif terhadap kebakaran hutan masih dalam tahap penelitian dan kurang dieksplorasi khususnya di wilayah Asia Tenggara. Selain itu, sistem yang sudah ada saat ini terkadang menghambat pengambilan keputusan pemangku kepentingan, khususnya pada langkah mitigasi. Model prediktif dapat dimanfaatkan untuk memberikan wawasan dan acuan dasar dalam menyusun langkah mitigasi seperti penentuan wilayah patroli. Meskipun BMKG saat ini menyediakan indeks bahaya kebakaran berbasis Fire Weather Index yang valid secara meteorologis untuk pemantauan harian, indeks yang tersedia masih terbatas pada variabel meteorologis dan belum mengintegrasikan komponen biofisik maupun antropogenik. Walaupun sistem tersebut sudah menerapkan pendekatan *forecasting*, faktor yang digunakan

masih berfokus ke faktor meteorologis.

Oleh karena itu, langkah prediktif diperlukan untuk memberikan informasi terkait indikator karhutla dengan menggunakan faktor yang lebih luas dan relevan menurut banyak penelitian sekaligus memberikan acuan informasi yang lebih cepat sebagai bahan perancangan mitigasi. Namun, pendekatan prediktif seperti memanfaatkan ML sering kali menimbulkan keraguan disebabkan hasil prediksi yang diberikan bersifat *black box* sehingga hasil yang diberikan masih diragukan oleh pemangku kepentingan. Berdasarkan penjelasan sebelumnya, didefinisikan beberapa rumusan masalah sebagai berikut.

1. Bagaimana mengembangkan model prediksi risiko kebakaran hutan dan lahan berbasis ML dengan memanfaatkan faktor-faktor penyebab karhutla?
2. Bagaimana cara menerapkan pendekatan XAI untuk menjelaskan peran faktor-faktor utama yang mempengaruhi hasil prediksi risiko terjadinya karhutla?
3. Bagaimana bentuk penerjemahan hasil prediksi sehingga dapat memberikan wawasan yang relevan untuk mendukung keputusan mitigasi?

### I.3 Tujuan

Berdasarkan rumusan masalah yang telah dituliskan pada bagian sebelumnya, berikut adalah tujuan utama dari penggerjaan tugas akhir ini.

1. Membangun model prediksi risiko kebakaran hutan dan lahan berbasis ML dengan memanfaatkan data hotspot satelit, informasi meteorologi, dan faktor lingkungan pada suatu wilayah studi di Indonesia.
2. Menerapkan pendekatan XAI untuk menginterpretasikan faktor-faktor utama yang memengaruhi hasil prediksi risiko karhutla pada model yang dikembangkan.
3. Merancang prototipe visualisasi spasial yang menyajikan hasil prediksi risiko karhutla secara informatif dan mudah dipahami sebagai bahan pertimbangan dalam upaya pencegahan dan mitigasi.

Keberhasilan tugas akhir ini dievaluasi melalui pencapaian akurasi model yang me-lampaui 0,80 dalam memprediksi titik api dan nilai F1-score mencapai 0,75 untuk kelas risiko tinggi serta kemampuan sistem dalam memberikan alasan di balik hasil prediksi model dan penyajian informasi secara informatif dan jelas yang dibungkus dalam bentuk prototipe visual. Selain itu, nilai *fidelity* yang ditargetkan untuk XAI adalah sekitar 0,80 sehingga penjelasan yang dihasilkan diharapkan cukup konsisten dengan keputusan model. Selain itu, tugas akhir ini tidak ditujukan sebagai langkah untuk menggantikan sistem Fire Weather Index yang telah digunakan saat ini, tu-

gas akhir ini akan tetap memanfaatkan indeks tersebut sebagai salah satu masukan dalam model prediksi risiko terjadinya karhutla.

#### I.4 Batasan Masalah

Adapun untuk batasan masalah dalam penggerjaan tugas akhir ini adalah sebagai berikut.

1. Untuk mengurangi masalah keterbatasan data dan beban infrastruktur pengembangan, tugas akhir difokuskan pada wilayah Kalimantan dan menggunakan model ML.
2. Jenis data yang digunakan dibatasi pada data satelit hotspot, data meteorologi, dan data biofisik yang tersedia secara terbuka.
3. Prototipe penyajian hasil prediksi bersifat demonstratif dan berbasis data historis yang telah diproses secara *offline*.
4. Cakupan waktu yang diprediksi adalah skala bulanan untuk mengurangi kompleksitas data dan mendukung perencanaan strategis.

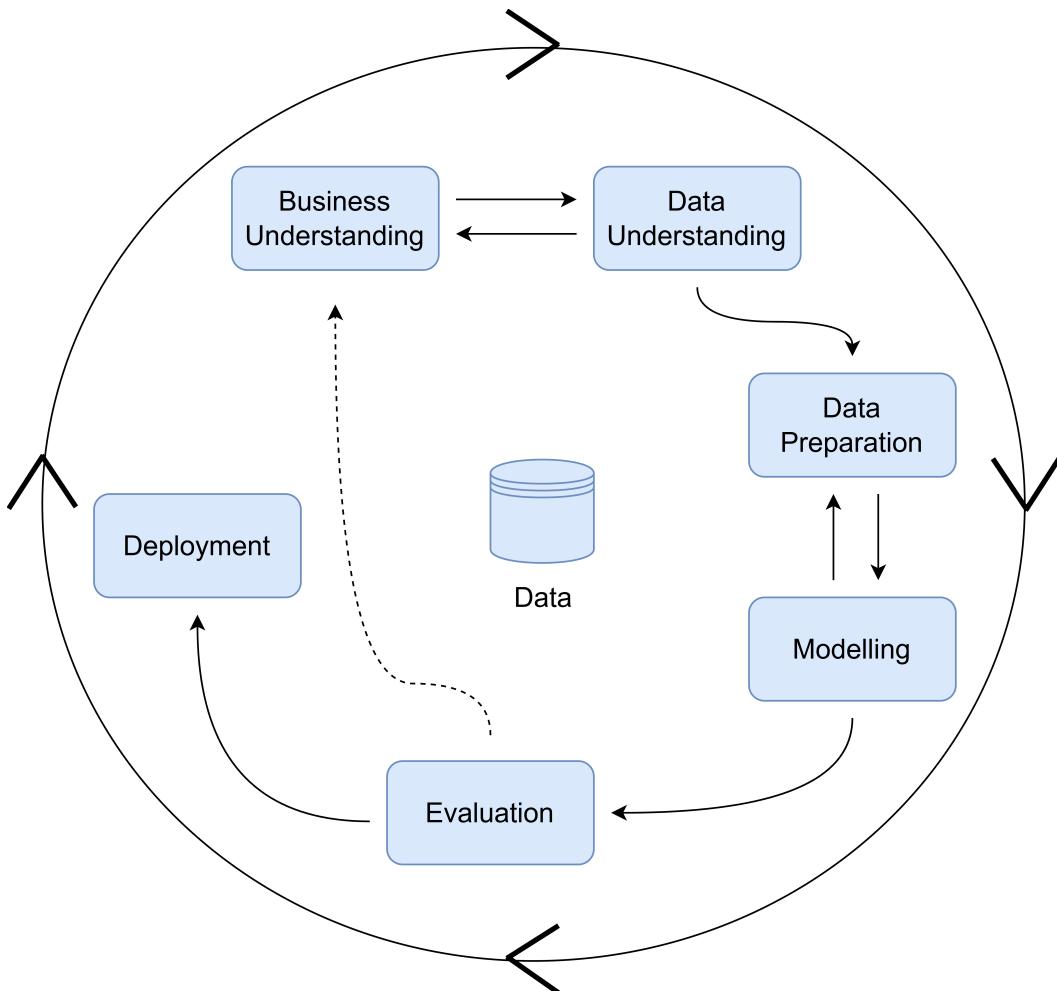
#### I.5 Metodologi

Metodologi yang digunakan dalam tugas akhir ini mengacu pada kerangka kerja CRISP-DM (*Cross-Industry Standard Process for Data Mining*). CRISP-DM dipilih karena menyediakan tahapan yang sistematis untuk mengembangkan solusi berbasis data yang dimulai dari pemahaman masalah hingga pembangunan model. Metodologi ini diadaptasi agar sesuai dengan konteks pemodelan risiko karhutla dan pengembangan prototipe visual yang disertai penerapan XAI.

Menurut Gambar I.2, tahapan metodologi yang akan dilaksanakan adalah sebagai berikut.

1. *Business Understanding*

Pada tahap ini dilakukan pemahaman terhadap konteks kebakaran hutan dan lahan di wilayah studi beserta dampak yang ditimbulkan. Kegiatan yang dilakukan meliputi studi literatur mengenai karhutla di Indonesia, penyebab dan dampak dari karhutla, serta identifikasi kebutuhan akan informasi risiko sebagai langkah pemenuhan tindakan mitigasi. Hasil dari tahap ini adalah pemahaman konteks, perumusan masalah dan tujuan, pemodelan, ruang lingkup wilayah studi, serta kriteria keberhasilan pengembangan model dan prototipe. Selain itu, langkah pemahaman diawali dengan melakukan riset dan kajian untuk memahami permasalahan dan pendekatan solusi yang dilakukan dengan



Gambar I.2 Metodologi CRISP-DM

memanfaatkan artikel-artikel terbitan tahun 2021 ke atas agar memiliki relevansi yang signifikan. Artikel yang digunakan diusahakan memiliki reputasi yang baik dan berasal dari jurnal yang terindeks Scopus atau Sinta. Namun, beberapa artikel yang terbit di bawah tahun 2021 terkadang digunakan sebagai dasar atau dukungan studi.

## 2. *Data Understanding*

Tahap ini berfokus pada pengenalan dan eksplorasi data yang akan digunakan. Data yang dimaksud mencakup data *hotspot* atau kebakaran dari satelit, data meteorologi, serta data lingkungan/biofisik wilayah studi. Kegiatan meliputi identifikasi sumber data, pengumpulan data historis, analisis struktur dan format data, serta eksplorasi awal pola spasial dan temporal kejadian karhutla. Selain itu, dilakukan pula penilaian awal terhadap kualitas data, seperti keberadaan nilai hilang, inkonsistensi, atau ketidakseimbangan kelas antara kejadian kebakaran dan non-kebakaran.

### *3. Data Preparation*

Pada tahap ini dilakukan pengolahan data agar siap digunakan dalam pemodelan. Kegiatan yang dilakukan antara lain pemilihan periode waktu yang akan dianalisis, pembatasan data pada wilayah studi, pembentukan representasi spasial (pembentukan *grid*), penggabungan data hotspot dengan data meteorologi dan data lingkungan, serta pembersihan data (penanganan nilai hilang, duplikasi, dan nilai yang tidak wajar). Pada tahap ini juga dilakukan analisis terhadap fitur yang relevan dengan risiko karhutla serta pembagian data menjadi pelatihan dan pengujian.

### *4. Modelling*

Tahap pemodelan berfokus pada pembangunan model *Supervised Machine Learning* untuk memprediksi risiko kebakaran hutan dan lahan. Kegiatan meliputi pemilihan model yang sesuai, skema pelatihan dan validasi, serta penyesuaian parameter model. Pada tahap ini juga dilakukan penerapan teknik XAI sebagai langkah untuk menjelaskan kontribusi masing-masing fitur terhadap hasil prediksi.

### *5. Evaluation*

Evaluasi model dilakukan dengan meninjau nilai kuantitatif yang dihasilkan dengan menggunakan beberapa metrik utama yang umum digunakan pada evaluasi model, termasuk Accuracy, Precision, Recall, F1-Score, serta Area Under ROC Curve (AUC-ROC). Selain itu evaluasi kualitatif juga dilakukan berdasarkan hasil interpretasi dari XAI untuk menilai apakah faktor-faktor yang muncul selaras dengan domain terkait karhutla di wilayah studi.

### *6. Deployment*

Pada tahap terakhir, hasil pemodelan dan interpretasi XAI disusun dalam bentuk prototipe penyajian hasil prediksi berupa visualisasi spasial yang menampilkan peta risiko karhutla pada wilayah studi. Prototipe ini dirancang agar informasi risiko dan faktor-faktor penjelas dapat disajikan secara sederhana dan objektif sehingga mudah dipahami oleh pengguna.

## **BAB II**

### **STUDI LITERATUR**

Bab studi literatur menjelaskan kajian-kajian yang relevan dengan permasalahan yang diangkat pada tugas akhir ini. Penjelasan dalam bab ini akan mencakup tinjauan terhadap karhutla dan penerapan model *Machine Learning* (ML) serta XAI, dan literatur terkait seputar solusi prediktif terhadap bencana karhutla. Kajian yang dilakukan berguna untuk memberikan pemahaman dan landasan dalam memahami domain dan menerapkan solusi yang relevan.

#### **II.1 Kebakaran Hutan dan Lahan (Karhutla) di Indonesia**

Kebakaran hutan dan lahan (karhutla) merupakan kejadian berkobarnya api yang meliputi hutan serta lahan gambut/mineral yang sering dipicu oleh kombinasi faktor antropogenik dan kondisi lingkungan serta dapat terjadi dengan frekuensi tahunan (Nurhayati dkk. 2021). Peristiwa ini menjadi ancaman berulang di Indonesia, salah satunya di wilayah Kalimantan. Dampak yang ditimbulkan oleh karhutla mengakibatkan kerusakan hingga ke berbagai aspek:

1. Kerusakan Ekologis: Karhutla menyebabkan kerusakan lahan serta menjadi bencana yang menimbulkan hilangnya biodiversitas dan menyebabkan pemanasan global (Thoha dan Triani 2021).
2. Dampak Kesehatan: Asap lintas batas (*transboundary haze*) yang dihasilkan membawa partikel aerosol berbahaya (PM<sub>2.5</sub>) yang terbukti secara klinis menyebabkan peningkatan signifikan pada kasus infeksi saluran pernapasan akut (ISPA), asma, dan bahkan kematian prematur pada populasi yang terpapar (Syahid dan Hidayat 2022).
3. Kerugian Ekonomi: Bank Dunia memperkirakan kerugian ekonomi akibat karhutla pada tahun 2019 saja mencapai 5,2 miliar USD di Indonesia, setara dengan 0,5% dari PDB negara. Kerugian ini mencakup kerusakan pada sektor pertanian, kehutanan, transportasi, dan pariwisata, serta biaya pemadaman

yang sangat besar (Jain dkk. 2021).

## II.2 Faktor Pemicu Karhutla

Penyebab karhutla di Indonesia umumnya disebabkan hasil interaksi antara kondisi lingkungan yang membuat hutan dan lahan menjadi rentan dan aktivitas manusia yang menyulut api. Sejumlah studi mengatakan bahwa variabilitas dari iklim (seperti ENSO dan kekeringan), karakteristik biofisik (keberadaan lahan gambut), serta faktor antropogenik (aktivitas manusia) saling memperkuat satu sama lain dalam membentuk pola karhutla di Indonesia.

### II.2.1 Faktor Meteorologis

Faktor meteorologis berperan sebagai pemicu yang menentukan seberapa mudah suatu daerah terbakar. Cuaca dan musim merupakan faktor yang dapat mempengaruhi kondisi lingkungan. Sejumlah penelitian juga menunjukkan bahwa anomali iklim berkaitan kuat dengan intensitas kebakaran di Indonesia. Studi oleh Fitriany dkk. (2021) menemukan bahwa kejadian kebakaran hutan di Indonesia berkorelasi dengan indeks iklim jangka panjang seperti El Niño-Southern Oscillation (ENSO) dan Indian Ocean Dipole (IOD). ENSO sendiri memiliki dua fase yang mempengaruhi kondisi iklim lingkungan.

1. El Niño: Fenomena ini secara signifikan mengurangi curah hujan dan memperpanjang durasi musim kemarau di sebagian besar wilayah Indonesia (Syahid dan Hidayat 2022). Analisis data historis menunjukkan korelasi positif yang kuat antara indeks ENSO dan jumlah *hotspot* (titik deteksi kebakaran) yang terdeteksi oleh satelit. Selama episode El Niño yang kuat, penurunan kelembaban tanah dan vegetasi menciptakan kondisi lingkungan yang sangat kondusif bagi penyebaran api yang cepat dan luas.
2. La Niña: Sebaliknya, La Niña membawa kondisi yang lebih basah ke Indonesia, yang secara efektif menekan aktivitas kebakaran dan mengurangi jumlah *hotspot* tahunan (Syahid dan Hidayat 2022). Siklus ENSO ini menjelaskan mengapa intensitas Karhutla di Indonesia sangat berfluktuasi dari tahun ke tahun.

### II.2.2 Faktor Antropogenik

Walaupun iklim menciptakan kondisi yang rentan, sebagian besar api awal (*ignition*) di Indonesia justru berasal dari aktivitas manusia, bukan penyebab alami. Faktor antropogenik merupakan faktor yang berhubungan dengan aktivitas manusia. Se-

jumlah kajian menunjukkan bahwa pemicu utama karhutla di Indonesia didominasi oleh aktivitas manusia. Nurhayati dkk. (2021) menemukan bahwa pola kebakaran di lahan gambut Sumatra banyak terjadi di kawasan dengan tekanan pemanfaatan lahan tinggi. Sementara itu, menurut Rossita dkk. (2023) kebakaran berulang yang terjadi di Riau berhubungan dengan praktik pengelolaan lahan dan drainase. Temuan-temuan ini menegaskan bahwa faktor sosial dan tata kelola lahan berkontribusi besar terhadap peningkatan frekuensi bencana karhutla.

### **II.2.3 Faktor Biofisik**

Faktor biofisik berkaitan dengan karakteristik fisik dan wilayah ekologis yang memengaruhi kerentanan suatu area terhadap kebakaran. Beberapa faktor utama yang banyak dibahas dalam literatur antara lain keberadaan lahan gambut, jenis tutupan lahan (vegetasi), dan topografi. Rossita dkk. (2023) menunjukkan bahwa gambut mencakup sekitar 8% dari luas daratan Indonesia, dan analisis mereka di satu kesatuan hidrologis gambut di Riau menemukan tren peningkatan luas gambut terbakar antara 2001-2020, dengan sekitar 33% area yang terbakar mengalami kebakaran berulang. Hal ini mengindikasikan perangkap degradasi, di mana lahan yang pernah terbakar menjadi semakin rentan untuk terbakar kembali pada tahun-tahun berikutnya.

## **II.3 Indeks Cuaca Kebakaran**

Salah satu pendekatan yang banyak digunakan untuk menilai tingkat bahaya kebakaran karhutla adalah Fire Weather Index (FWI). Studi oleh Fitriani dkk. (2023) menjelaskan bahwa FWI menjadi dasar dari Fire Danger Rating System (FDRS) yang dioperasikan oleh BMKG untuk memantau potensi bahaya karhutla secara harian. Sistem ini menggabungkan informasi cuaca harian bahaya karhutla seperti suhu, kelembapan, kecepatan angin, dan curah hujan untuk menghasilkan indeks yang menggambarkan tingkat risiko kebakaran di suatu wilayah.

Fitriani dkk. (2023) bahwa FDRS Indonesia telah disesuaikan dengan sistem FWI yang awalnya dikembangkan di Kanada dan kemudian diadaptasi untuk kondisi iklim di Indonesia. Dalam sistem ini, indeks FWI terdiri dari enam komponen yang dibagi menjadi dua kelompok utama, yaitu kode kelembapan bahan bakar (*fuel moisture codes*) dan indeks perilaku kebakaran (*fire behaviour indices*).

1. Fine Fuel Moisture Code (FFMC)

Menggambarkan kelembapan lapisan permukaan atas tanah dan bahan bakar

yang sangat halus seperti ranting kecil. Nilai FFMC yang tinggi menunjukkan kondisi bahan bakar permukaan sangat kering sehingga mudah terbakar.

2. Duff Moisture Code (DMC)

Menggambarkan kelembapan lapisan bahan organik pada kedalaman menengah di bawah permukaan tanah. Kode ini peka terhadap kekeringan pada skala mingguan

3. Drought Code (DC)

Menggambarkan kelembapan lapisan bahan organik yang lebih dalam dan berhubungan dengan kekeringan jangka panjang. Nilai DC yang tinggi menandakan kondisi sangat kering dan potensi kebakaran bawah permukaan yang sulit dipadamkan

4. Initial Spread Index (ISI)

Merupakan indeks yang menggabungkan FFMC dan kecepatan angin untuk memperkirakan laju awal penjalaran api jika terjadi penyalaan. Semakin kering bahan bakar dan semakin kencang angin, semakin besar nilai ISI.

5. Build-Up Index (BUI)

Menggabungkan DMC dan DC untuk merepresentasikan akumulasi bahan bakar kering yang tersedia. BUI yang besar menunjukkan banyaknya bahan bakar yang siap dikonsumsi api.

6. Fire Weather Index (FWI)

Merupakan indeks utama yang diperoleh dari kombinasi ISI dan BUI dan digunakan sebagai indikator intensitas potensial kebakaran hutan dan lahan.

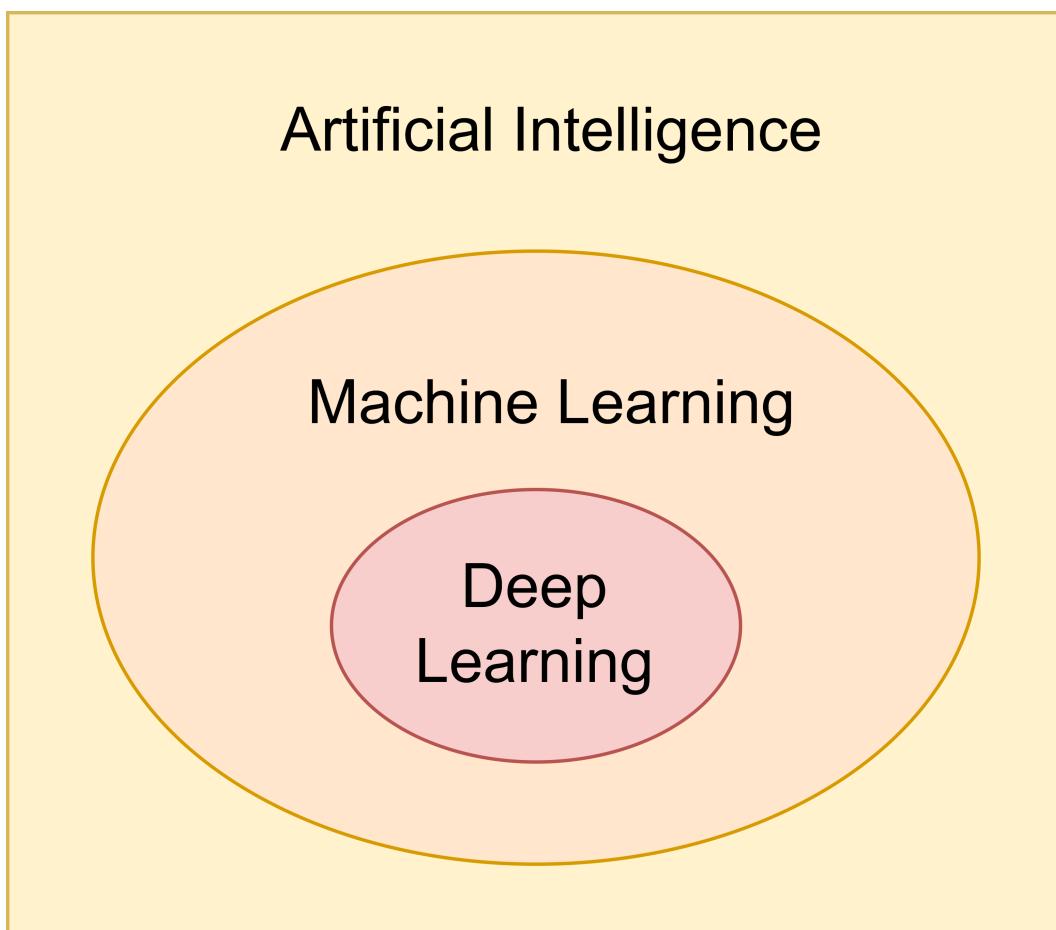
Menurut penjelasan pada laman resmi BMKG, FWI menunjukkan besarnya intensitas api jika terjadi kebakaran dan sangat dipengaruhi nilai ISI dan BUI.

Laman resmi BMKG menyajikan peta FWI dan indeks terkait dalam bentuk peta spasial untuk melihat beberapa kategori tingkat bahaya untuk membantu kepentingan dalam memantau potensi kebakaran pada skala nasional. FWI dipandang sebagai kerangka indeks bahaya berbasis cuaca yang sudah teruji dan digunakan secara operasional di Indonesia. Namun, pendekatan ini berfokus ke variabel meteorologis, sedangkan banyak penelitian yang menunjukkan bahwa risiko karhutla ternyata dipengaruhi oleh banyak faktor selain meteorologis.

## II.4 *Artificial Intelligence*

*Artificial Intelligence* (AI) adalah bidang ilmu komputer yang luas dan berkaitan dengan pembangunan mesin pintar yang mampu melakukan tugas-tugas yang biasanya membutuhkan kecerdasan manusia (Sarker 2022). Penelitian ini juga menjelaskan bahwa AI berhasil mentransformasikan sebagian besar industri, mulai dari keuangan

an dan manufaktur hingga transportasi. AI membantu manusia dalam mengambil keputusan dengan berperan sebagai sistem yang dapat mengolah data dalam jumlah besar, mengekstraksi pola, dan menghasilkan prediksi atau rekomendasi untuk mendukung pengambilan keputusan. Gupta dkk. (2022) menjelaskan bahwa integrasi AI dan *big data* telah menjadi fondasi utama sistem pengambilan keputusan berbasis data di berbagai sektor untuk saat ini. Pada tugas akhir, peran AI relevan dalam tindakan prediksi risiko kebakaran hutan serta penyajian informasi untuk membantu pengelola karhutla dalam menetapkan langkah mitigasi.

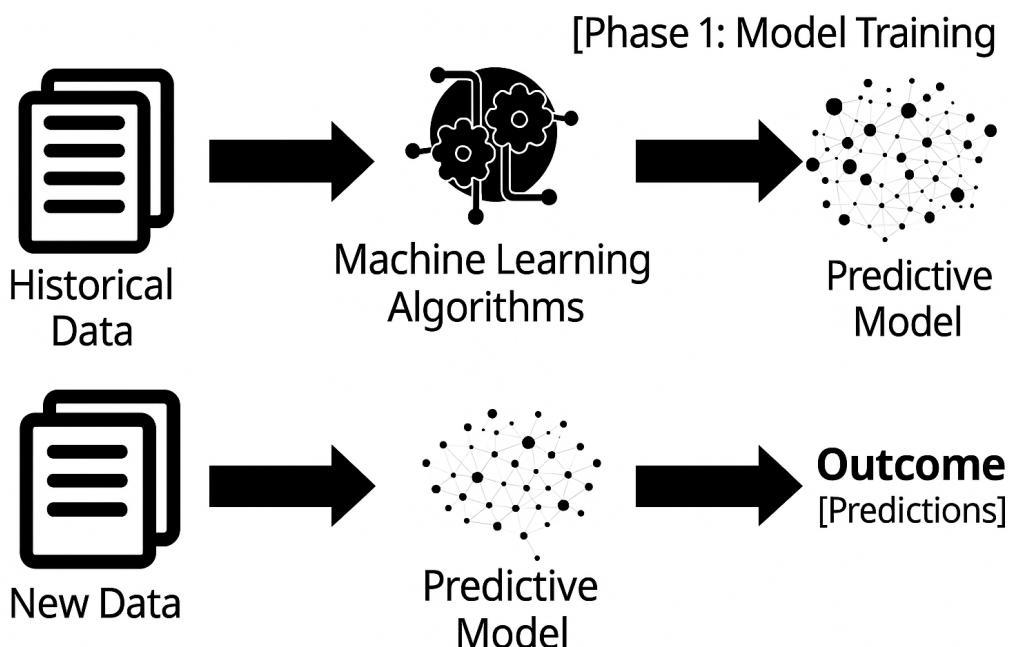


Gambar II.1 Ilustrasi posisi hierarki ML dan DL pada area AI (Sarker 2022)

*Artificial Intelligence*, *Machine Learning* (ML), dan *Deep Learning* (DL) saling berkaitan secara hierarkis seperti yang ditunjukkan pada Gambar II.1 dan berinteraksi dalam membangun sistem perangkat lunak yang cerdas. Menurut Sarker (2022), ML dan DL merupakan bagian dari teknologi AI yang berfokus pada pembelajaran berbasis data untuk mengembangkan sistem cerdas dan mengotomatisasi proses. DL sendiri merupakan subbagian ML yang memanfaatkan jaringan saraf berlapis untuk mempelajari representasi yang kompleks, sehingga berperan penting dalam analitik tingkat lanjut dan pengambilan keputusan cerdas.

## II.5 Machine Learning

*Machine Learning* (ML) merupakan cabang utama dari AI yang berfokus pada bagaimana komputer dapat belajar dari data untuk meningkatkan kinerja tugas tertentu tanpa diprogram secara eksplisit untuk setiap aturan. Menurut Sarker (2021), ML telah berkembang pesat dalam konteks analisis data dan komputasi serta meningkatkan pengalaman otomatisasi dan mendorong revolusi industri keempat (Industri 4.0). Dalam AI, ML telah muncul sebagai metode pilihan untuk mengembangkan perangkat lunak praktis untuk visi komputer, pengenalan suara, pemrosesan bahasa alami, kendali robot, dan lainnya.



Gambar II.2 Struktur umum ML pada fase pelatihan dan pengujian (Sarker 2022)

Dalam konteks organisasi, Gupta dkk. (2022) menunjukkan bahwa pemanfaatan model ML yang didasari oleh infrastruktur *big data* mampu meningkatkan kualitas pengambilan keputusan melalui hasil prediksi, klasifikasi, atau deteksi pola anomali yang lebih akurat dibandingkan pendekatan konvensional. Pada Gambar II.2, kerangka umum model prediktif berbasis ML ditandai dengan dua fase sebelum menjadi sebuah model prediktif. Kerangka model ML diawali dengan fase 1, yaitu pelatihan dari data historis dan menghasilkan prediksi untuk dijadikan sebagai data uji pada fase 2. Secara garis besar, Sarker (2022) membagi ML ke dalam tiga kategori utama, yaitu *supervised learning*, *unsupervised learning*, dan *reinforcement learning*.

Tabel II.1 Kategori *Machine Learning* (Sarker 2022)

Tipe pembelajaran	Proses pembentukan model	Tugas umum
<i>Supervised learning</i>	Algoritma atau model belajar dari data berlabel (pendekatan berbasis tugas)	Klasifikasi, Regresi
<i>Unsupervised learning</i>	Algoritma atau model belajar dari data tidak berlabel (pendekatan berbasis data)	Klastering, Asosiasi, Reduksi dimensi
<i>Semi-supervised learning</i>	Model dibangun menggunakan kombinasi data berlabel dan tidak berlabel	Klasifikasi, Klastering
<i>Reinforcement learning</i>	Model belajar berdasarkan <i>reward</i> atau penalti (pendekatan berbasis lingkungan)	Klasifikasi, Kontrol

Pada *supervised learning*, model dilatih dengan fitur masukan dan label sehingga dapat melakukan prediksi pada data baru yang belum pernah dilihat. Pendekatan ini sangat relevan dengan prediksi risiko karhutla karena data historis sudah mencatat faktor-faktor yang menyebabkan suatu lokasi kebakaran atau tidak. *Unsupervised learning* digunakan ketika label tidak tersedia dan model berfokus dalam menemukan pola atau struktur tersembunyi. Disebabkan data tidak memiliki label, model harus bisa memprediksi pola menggunakan metode seperti *clustering*. Terdapat model yang menggabungkan antara *supervised* dan *unsupervised* bernama *semi-supervised learning* yang biasa digunakan ketika jumlah data berlabel tidak cukup banyak sehingga memanfaatkan data tidak berlabel untuk memperkaya data. *Reinforcement learning* mempelajari strategi atau tindakan terbaik melalui interaksi dengan lingkungan dan umpan balik berupa *reward* atau *penalty*. Model ini lebih relevan untuk digunakan pada sistem kontrol atau penjadwalan dinamis pada pemodelan risiko statis.

### II.5.1 Konsep *Supervised Learning*

Sarker (2021) mendefinisikan *supervised learning* sebagai pendekatan ML sebagai model yang belajar dari data yang sudah memiliki label. Setiap contoh data terdiri atas pasangan masukan dan keluaran yang benar sehingga model dapat mempelajari pemetaan dari fitur-fitur pada data hingga ke label berdasarkan seluruh contoh data tersebut. Pendekatan ini banyak digunakan pada tugas klasifikasi dan regresi di berbagai domain karena secara langsung memanfaatkan informasi target yang diketahui saat pelatihan.

Secara matematis, data latih pada *supervised learning* dapat dituliskan sebagai

$$D = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\} \quad (\text{II.1})$$

Pada Persamaan II.1,  $x_i \in \mathbb{R}^d$  merupakan vektor fitur untuk contoh ke  $i$ , dan  $y_i$  yang merupakan label atau nilai target yang bersesuaian. Pada *supervised learning*, kesalahan yang dimiliki model akan menjadi hukuman dan semakin besar kesalahannya akan semakin besar hukumannya. Jadi, tugas utama *supervised learning* adalah memilih suatu model ( $f(x; \theta)$ ) dan kemudian mencari parameter  $\theta$  yang membuat hasil model dapat sedekat mungkin dengan label yang benar. Untuk menilai seberapa bagus model tersebut, terdapat fungsi kerugian yang dapat mengukur seberapa jauh prediksi dari label sebenarnya.

$$R^{(\theta)} = \frac{1}{n} \sum_{i=1}^n L(y_i, f(x_i; \theta)) \quad (\text{II.2})$$

Rata-rata kerugian pada seluruh data latih (*empirical risk*) dinyatakan pada Persamaan II.2. Secara konsep, proses pelatihan model *supervised learning* dapat dipahami sebagai upaya pencarian nilai  $\theta$  yang dapat meminimalkan kerugian rata-rata.

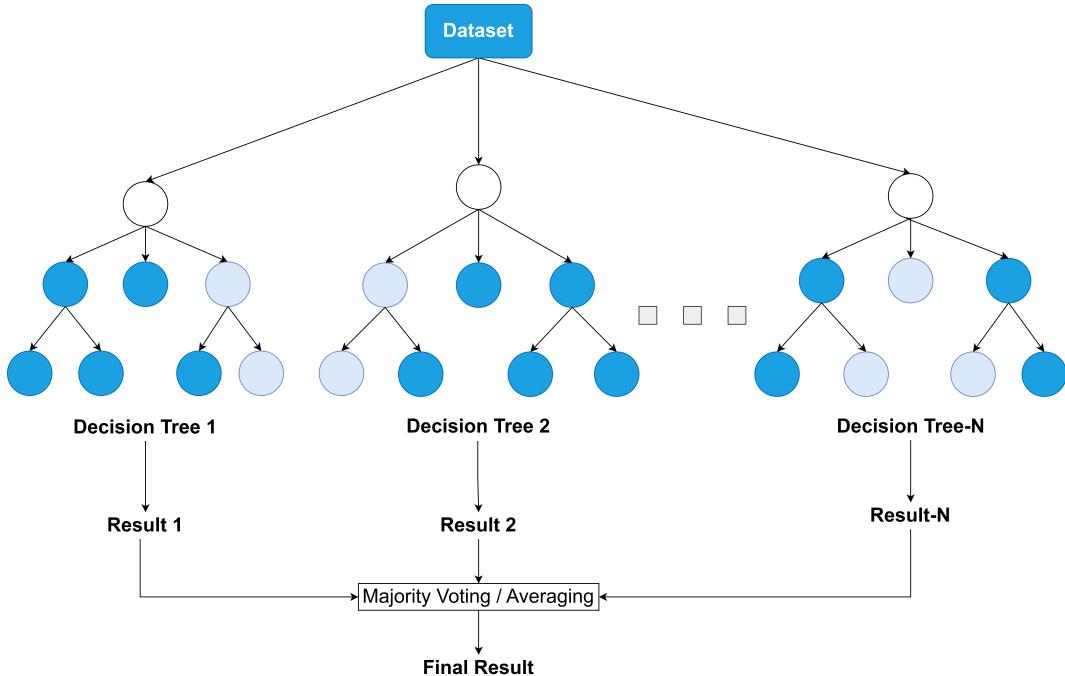
### II.5.2 Linear Regression

*Linear regression* (LR) adalah contoh paling sederhana dan klasik dari model *supervised learning*. Model ini menggambarkan hubungan antara sekumpulan fitur dan suatu nilai keluaran secara kontinu dengan menggambar garis lurus (untuk satu fitur) atau bidang (untuk banyak fitur). Pada Persamaan II.3,  $\beta_0$  merupakan titik potong dengan sumbu vertikal dan setiap  $\beta_j$  menunjukkan seberapa besar kontribusi suatu fitur  $x_j$  terhadap perubahan nilai *output*  $y$ . Istilah  $\varepsilon$  digunakan sebagai galat karena terdapat variasi yang bisa saja tidak bisa dijelaskan oleh kombinasi linier dari fitur-fitur tersebut.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon. \quad (\text{II.3})$$

### II.5.3 Random Forest

*Random Forest* (RF) adalah pendekatan ML yang menggabungkan prediksi beberapa pohon keputusan untuk menciptakan model yang lebih tangguh dan akurat. Dengan melakukan akumulasi keluaran melalui teknik seperti pemungutan suara (untuk klasifikasi) atau *averaging* (untuk regresi), RF dapat meningkatkan kemam-



Gambar II.3 Stuktur model (Random Forest) (Sarker 2021)

puan prediksinya dan meminimalkan terjadinya *overfitting* (Karurung, Lee, dan Lee 2025). Selain itu, menurut Sarker (2021) setiap pohon pada RF dilatih dengan sampel data yang berbeda dan hanya melihat subset acak dari fitur pada setiap pemisahan. Hal ini mengakibatkan masing-masing pohon dapat belajar pola yang berbeda sehingga kesalahan yang dibuat satu pohon dapat dikoreksi ketika digabungkan dengan pohon lain.

$$\hat{y} = \text{mode}(h_1(x), h_2(x), \dots, h_K(x)) \quad (\text{II.4})$$

Pada langkah prediksi untuk tugas klasifikasi, RF melakukan langkah *majority vote* dalam memilih keluaran. Pada Persamaan II.4, metode *majority vote* ditentukan dengan memilih kelas mana yang paling sering dipilih dari seluruh pohon. Pendekatan ini membantu mengurangi varians yang biasanya muncul pada satu pohon keputusan saja sekaligus meningkatkan akurasi model. Selain menghasilkan prediksi kelas risiko, RF juga menyediakan ukuran pentingnya suatu fitur *feature importance* yang dapat dimanfaatkan sebagai awal bagi analisis *explainable AI* yang lebih mendalam.

#### II.5.4 *Extreme Gradient Boosting*

XGBoost (*Extreme Gradient Boosting*) merupakan algoritma *gradient boosting* berbasis pohon keputusan yang banyak digunakan dalam studi prediksi karena menggabungkan kinerja prediksi yang tinggi dengan efisiensi komputasi serta mampu

menangani data dalam skala besar. Berbeda dengan RF, XGBoost menyusun pohon secara bertahap untuk memperbaiki kesalahan model sebelumnya. Wiens dkk. (2025) menyebutkan persamaan prediksi XGBoost untuk contoh ke- $i$  sebagai penjumlahan keluaran dari  $big(K)$  pohon sebagai berikut.

$$F_k(x) = F_0(x) + \varepsilon T_1(x) + \varepsilon T_2(x) + \cdots + \varepsilon T_k(x) \quad (\text{II.5})$$

Pada Persamaan II.5,  $F_0(x)$  biasanya berupa model awal sederhana (misalnya nilai rata-rata),  $T_m(x)$  merepresentasikan pohon regresi pada iterasi ke- $m$ , dan  $\varepsilon$  adalah *learning rate* yang mengendalikan besar langkah pembaruan model. Hal ini sesuai dengan pendapat Sarker (2021) bahwa gradien yang digunakan pada XGBoost berguna untuk meminimalkan fungsi kerugian yang serupa dengan cara kerja jaringan saraf tiruan (*Artificial Neural Network*).

## II.6 Metrik Evaluasi

Dalam pengembangan model *Machine Learning* untuk melakukan prediksi, langkah evaluasi sangat diperlukan untuk menilai apakah model memiliki kemampuan yang cukup untuk memprediksi label. Langkah untuk melakukan evaluasi tersebut dapat menggunakan metrik evaluasi untuk melihat kinerja dan performa model. Jenis metrik evaluasi yang dapat digunakan sebagai bahan evaluatif berbeda-beda setiap modelnya dengan mempertimbangkan sifat data, jenis tugas, dan kebutuhan interpretasi. Berdasarkan penjelasan dari Miller dkk. (2024), metrik evaluasi dikelompokkan berdasarkan informasi yang berasal dari *confusion matrix*, seperti *accuracy*, *precision*, *recall*, *F1-score* serta metrik berbasis kurva seperti *Receiver Operating Characteristic* (ROC) dan *Area Under the Curve* (AUC).

Pada konteks karhutla sendiri, berbagai studi menunjukkan penggunaan metrik yang berbeda-beda tergantung cara mereka melakukan prediksi dan jenis model prediksi yang digunakan. Walaupun begitu, setiap studi menggunakan kombinasi metrik untuk memberikan evaluasi model secara jelas. Penelitian yang dilakukan Abdollahi dan Pradhan (2023) menggunakan metrik akurasi, *precision*, *recall*, *F1-Score*, dan AUC-ROC dalam melakukan evaluasi model *wildfire susceptibility* dan menilai kemampuan model dari berbagai sudut pandang. Di sisi lain, studi dari Karurung, Lee, dan Lee (2025) memanfaatkan metrik yang sama seperti studi sebelumnya untuk melakukan perbandingan antara model RF dan XGBoost. Sementara itu, studi dari Eaturu dan Vadrevu (2025) yang membandingkan model ML dan DL dalam melakukan prediksi memanfaatkan *Root Mean Square Error* (RMSE), *Mean Abs-*

*lute Error (MAE), dan  $R^2$  (coefficient of determination) untuk mengevaluasi model prediksi kebakaran di Asia Tenggara.*

### II.6.1 Confusion Matrix

Miller dkk. (2024) dan Sokolova dan Lapalme (2009) menjelaskan bahwa *confusion matrix* merupakan representasi dasar untuk menilai performa model yang melakukan klasifikasi berdasarkan rangkuman prediksi benar dan salah pada tiap kelas. Mereka juga menambahkan bahwa metode ini berfungsi untuk menjadi basis dalam perhitungan berbagai metrik seperti *accuracy*, *precision*, dan *recall*. Untuk contoh klasifikasi biner seperti ini, *confusion matrix* memuat empat kategori, yaitu *true positive* (TP), *true negative* (TN), *false positive* (FP), dan *false negative* (FN). Empat kategori tersebut, dapat memberikan wawasan dan diturunkan menjadi beberapa perhitungan evaluatif.

Model memprediksi "terjadi kebakaran" dan kebakaran memang terjadi.	True Positive      False Negative	Model memprediksi "tidak terjadi kebakaran", tetapi sebenarnya terjadi kebakaran (missed detection)
Model memprediksi "terjadi kebakaran", tetapi sebenarnya tidak terjadi kebakaran ( <i>false alarm</i> ).	False Positive      True Negative	Model memprediksi "tidak terjadi kebakaran" dan memang tidak terjadi kebakaran.

Gambar II.4 Contoh penerapan *confusion matrix* pada kasus label data karhutla

1. Akurasi: efektivitas keseluruhan sebuah model *classifier*

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (\text{II.6})$$

2. Precision: proporsi prediksi positif yang benar-benar positif

$$\text{Precision} = \frac{TP}{TP + FP} \quad (\text{II.7})$$

3. Recall (Sensitivity): kemampuan model menemukan seluruh kelas positif

$$\text{Recall} = \frac{TP}{TP + FN} \quad (\text{II.8})$$

4. F1-Score: rata-rata harmonik dari precision dan recall

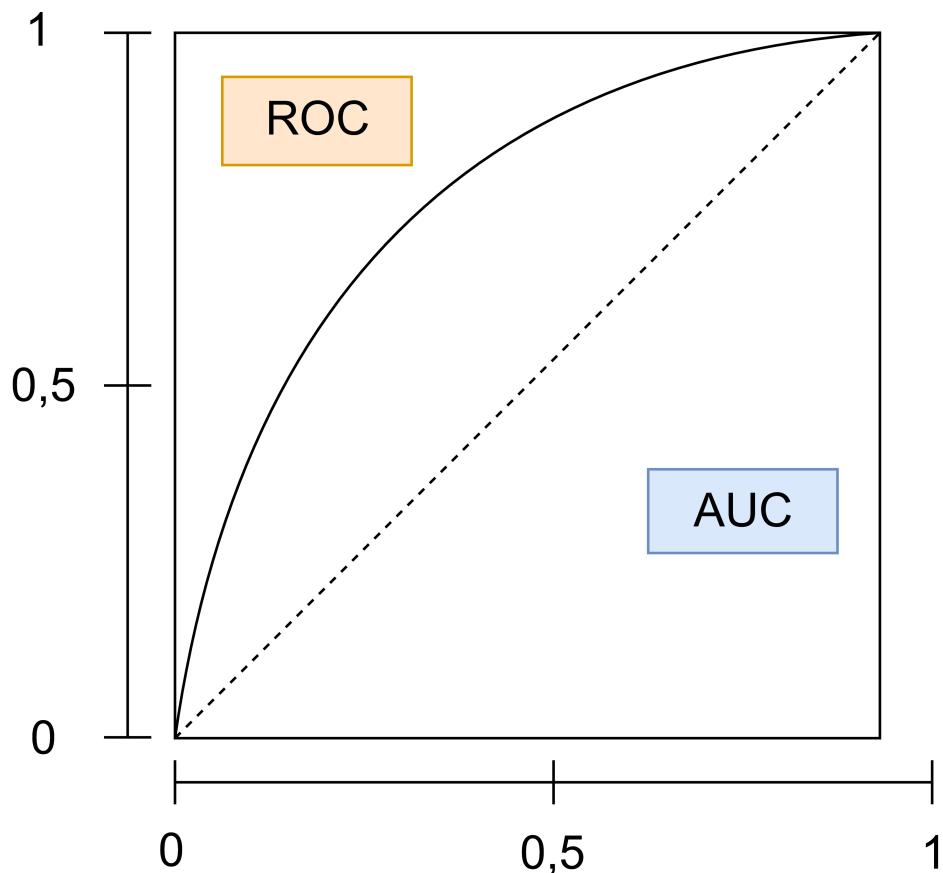
$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (\text{II.9})$$

5. Specificity: kemampuan model mengidentifikasi kelas negatif dengan benar

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (\text{II.10})$$

Akurasi saja sering kali tidak cukup untuk menilai sebuah model, terutama ketika jumlah kelas tidak seimbang. Dalam situasi tersebut, kombinasi metrik seperti *F1-Score* atau *precision-recall* lebih informatif untuk melihat kemampuan model dalam menangkap kelas minoritas.

## II.6.2 Kurva ROC dan Area AUC



Gambar II.5 Kurva ROC dan area ROC

Untuk model yang menghasilkan skor probabilitas atau skor risiko, penggunaan kurva *Receiver Operating Characteristic* ROC dan *Area Under the Curve* (AUC) sebagai metrik evaluasi yang tidak bergantung pada satu nilai batas saja. ROC meru-

pakan evaluasi yang bergantung dari nilai *True Positive Rate* (TPR) terhadap *False Positive Rate* (FPR) pada nilai ambang keputusan dengan persamaan sebagai berikut.

$$\text{TPR} = \frac{TP}{TP + FN}, \quad (\text{II.11})$$

$$\text{FPR} = \frac{FP}{FP + TN}. \quad (\text{II.12})$$

AUC dijelaskan sebagai luas area bawah kurva ROC yang apabila nilainya mendekati 1 model menunjukkan kemampuan pemisahan kelas dengan sangat baik. Namun, menurut Lavazza, Morasca, dan Rotoloni (2025), sebenarnya baik ROC maupun AUC secara konseptual dikembangkan untuk klasifikasi biner karena keduanya bergantung langsung pada definisi dua kelas (positif dan negatif) melalui TPR dan FPR. Meskipun dapat diperluas menjadi multikelas, interpretasinya kurang representif dibandingkan situasi biner. Selain itu AUC sering kali terlalu optimis pada dataset yang tidak seimbang sehingga diperlukan untuk melakukan kombinasi metrik dalam melihat performa model. Hal ini selaras dengan penelitian yang dilakukan oleh Karurung, Lee, dan Lee (2025) maupun Abdollahi dan Pradhan (2023) karena keduanya mengombinasikan AUC dengan metrik evaluasi lainnya.

### II.6.3 Metrik Regresi

Saat label berupa nilai kontinu, metrik yang digunakan adalah metrik regresi. Dalam studi Eaturu dan Vadrevu (2025), langkah evaluasi yang dilakukan adalah dengan menggunakan kombinasi RMSE, MAE, dan koefisien determinasi untuk menilai kualitas prediksi terhadap data observasi. Secara umum, RMSE didefinisikan sebagai

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2}, \quad (\text{II.13})$$

dengan  $y_i$  nilai aktual,  $\hat{y}_i$  nilai prediksi, dan  $n$  jumlah pengamatan. RMSE memberikan penalti lebih besar pada kesalahan yang besar karena selisih dipangkatkan terlebih dahulu sebelum dirata-ratakan, sehingga sensitif terhadap *outlier*. Nilai RMSE yang lebih kecil menunjukkan bahwa rata-rata selisih kuadrat antara prediksi dan observasi lebih rendah. Salah satu metrik lainnya yang digunakan untuk mengevaluasi

label kontinu adalah MAE.

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i|. \quad (\text{II.14})$$

MAE mengukur kesalahan rata-rata dalam satuan yang sama dengan target. Berbeda dengan RMSE, MAE memberikan bobot yang sama pada setiap kesalahan sehingga interpretasinya lebih langsung. Intinya, semakin kecil nilai MAE berarti prediksi model lebih bagus atau lebih dekat dengan nilai aktual.

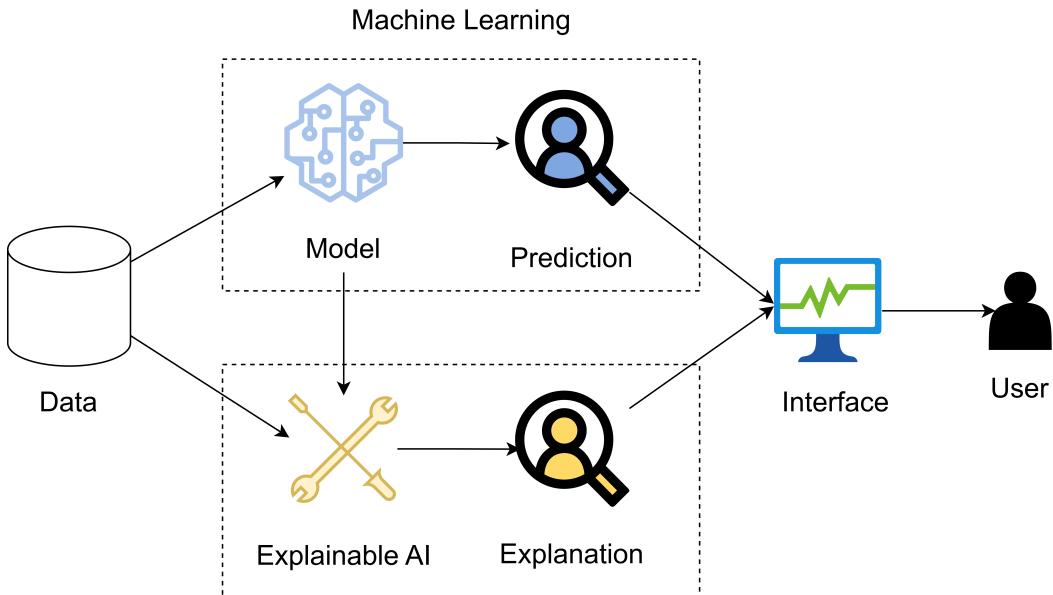
## II.7 Explainable Artificial Intelligence (XAI)

Model ML seperti RF dan XGBoost sering kali mencapai akurasi tinggi dalam melakukan prediksi. Namun, seluruh model prediksi sering dianggap sebagai "*black box*" atau kotak hitam karena tidak adanya penjelasan dasar di balik prediksi yang dihasilkan. Kondisi ini sering kali berujung pada keraguan karena pengguna tidak memahami alasan dibalik hasil prediksi tersebut. Dampak yang dihasilkan bisa berujung ke kesalahan interpretasi yang menganggap model prediksi selalu benar. Padahal model prediksi bisa saja memaksakan kemampuannya agar terlihat bagus pada metrik evaluasi walaupun hasil prediksinya tidak relevan dengan konteks data yang diberikan.

Bidang *Explainable Artificial Intelligence* (XAI) hadir untuk mengatasi permasalahan tersebut. Ali dkk. (2023) menekankan bahwa XAI berperan penting untuk membangun kepercayaan terhadap sistem AI, mendukung proses *debugging* model, dan menyediakan penjelasan yang dapat diakses oleh pemangku kepentingan non-teknis sebagai upaya untuk mewujudkan *trustworthy* AI. Pada Gambar II.6, XAI akan membantu menjelaskan hasil prediksi model dengan memanfaatkan informasi yang tersedia kepada user.

Secara umum, interpretabilitas terhadap model otomatis mendukung dua variasi, yaitu memahami keseluruhan perilaku model atau memahami satu prediksi. Literatur Ali dkk. (2023) dan Adadi dan Berrada (2018) menjelaskan bahwa pemahaman di balik keputusan atau prediksi model dibedakan menjadi dua subkelas.

1. Interpretabilitas Global: Penjelasan global merangkup perilaku model di seluruh data. Kelas ini memudahkan pemahaman logika suatu model serta menjadi aspek yang krusial dalam mengambil keputusan tingkat populasi (keseluruhan), seperti tren konsumsi obat-obatan atau perubahan iklim. Kelas ini juga dapat memahami apakah model yang dikembangkan berperilaku secara



Gambar II.6 Konsep skema XAI

wajar atau overfit ke pola yang tidak wajar.

2. Interpretabilitas Lokal: Penjelasan lokal berfokus pada menjelaskan alasan di balik keputusan spesifik atau prediksi tunggal (lokal). Kelas ini bermanfaat untuk memberikan dasar penjelasan mengapa suatu model memberikan keputusan spesifik tersebut.

Selain itu, XAI juga bersifat *post-hoc* yang artinya penjelasan dihasilkan setelah model selesai dilatih dan *model-agnostic* yang artinya XAI dapat diterapkan pada berbagai jenis model tanpa mengubah strukturnya.

### II.7.1 SHAP

SHAP (*SHapley Additive Explanations*) merupakan salah satu metode XAI yang paling populer untuk menjelaskan prediksi model ML. Sebagai contoh, pada penelitian Abdollahi dan Pradhan (2023), penggunaan SHAP yang dikombinasikan dengan ML dapat menghasilkan sebuah pemetaan kerentanan kebakaran hutan. Penelitian tersebut menunjukkan adanya faktor-faktor seperti kelembapan, kemiringan lereng, dan kondisi vegetasi yang disebutkan sebagai kontributor terbesar berdasarkan skor kerentanan kebakaran. Lundberg dan Lee (2017) memperkenalkan SHAP sebagai kerangka terstruktur untuk mengetahui *feature importance* yang memanfaatkan konsep nilai Shapley. Dalam kerangka ini, setiap fitur dianggap sebagai "emain" yang berkontribusi terhadap hasil nilai prediksi model. Pendapat tersebut diperkuat oleh Bifarin (2023) yang menyatakan SHAP berfungsi sebagai interpretabilitas model ML seperti RF dan XGBoost yang berfokus pada interpretabilitas lokal. Penjelas-

an tersebut selaras dengan hasil pada Abdollahi dan Pradhan (2023) yang memanfaatkan SHAP untuk menunjukkan kontribusi faktor dengan menghitung rata-rata kontribusi suatu fitur terhadap prediksi dengan mempertimbangkan kemungkinan kombinasi fitur lain.

$$f(x) \approx \phi_0 + \sum_{j=1}^M \phi_j, \quad (\text{II.15})$$

Menurut Persamaan ..,  $\phi_0$  adalah nilai dasar dan  $\phi_j$  adalah nilai SHAP untuk fitur ke- $j$ . Nilai  $\phi_j$  positif berarti fitur tersebut berkontribusi untuk meningkatkan prediksi menuju aktual label, sedangkan nilai negatif menurunkannya. Nilai  $\phi_j$  didefinisikan sebagai nilai berikut:

$$\phi_{ij} = \sum_{S \subseteq F \setminus \{j\}} \frac{|S|!(M - |S| - 1)!}{M!} [f_{S \cup \{j\}}(x) - f_S(x)], \quad (\text{II.16})$$

dengan  $F$  merupakan himpunan dari semua fitur dan  $S$  merupakan subset fitur yang tidak mengandung fitur itu sendiri ( $j$ ). Istilah  $f_S(x)$  menyatakan prediksi model ketika hanya terdapat fitur  $S$ . Rumus tersebut menyatakan bahwa kontribusi suatu fitur dihitung sebagai kontribusi marjinal fitur  $j$ , yaitu selisih prediksi model ketika fitur  $j$  disertakan dan ketika fitur tersebut tidak disertakan dalam suatu subset fitur  $S$ . Nilai selisih ini kemudian dirata-ratakan pada seluruh kombinasi subset  $S$  untuk mendapatkan nilai kontribusi akhir  $\phi_j$ .

SHAP merupakan metode yang dapat digunakan baik untuk penjelasan lokal maupun global. Dalam penjelasan lokal, SHAP memanfaatkan persamaannya untuk menghitung kontribusi masing-masing fitur terhadap suatu hasil prediksi. Pada penjelasan global, SHAP merata-ratakan nilai  $|\phi_j|$  di seluruh data sehingga mendapatkan sebuah peringkat pentingnya fitur yang menggambarkan bagaimana fitur tersebut berkontribusi dalam prediksi yang dilakukan model.

### II.7.2 Permutation Feature Importance

*Permutation Feature Importance* (PFI) merupakan metode XAI yang berfokus pada kelas interpretabilitas global. PFI mengukur seberapa penting sebuah fitur bagi model dengan melihat penurunan performa model ketika nilai fitur tersebut diacak sehingga hubungan antara fitur dan label menjadi rusak. Fumagalli dkk. (2023) menyebut PFI sebagai ukuran *model-agnostic* yang telah mapan untuk memperoleh *global feature importance* pada berbagai jenis model ML.

Langkah yang dilakukan oleh PFI untuk satu fitur adalah:

1. Hitung kinerja awal model pada data uji, seperti akurasi, presisi, atau AUC ( $M_{awal}$ ).
2. Lakukan pengacakan nilai kolom fitur di seluruh data uji sehingga hubungan antara fitur dengan label menjadi rusak.
3. Hitung ulang kinerja model pada data yang sudah diacak ( $M_{acak,j}$ ).
4. Definisikan skor pentingnya suatu fitur dengan persamaan berikut.

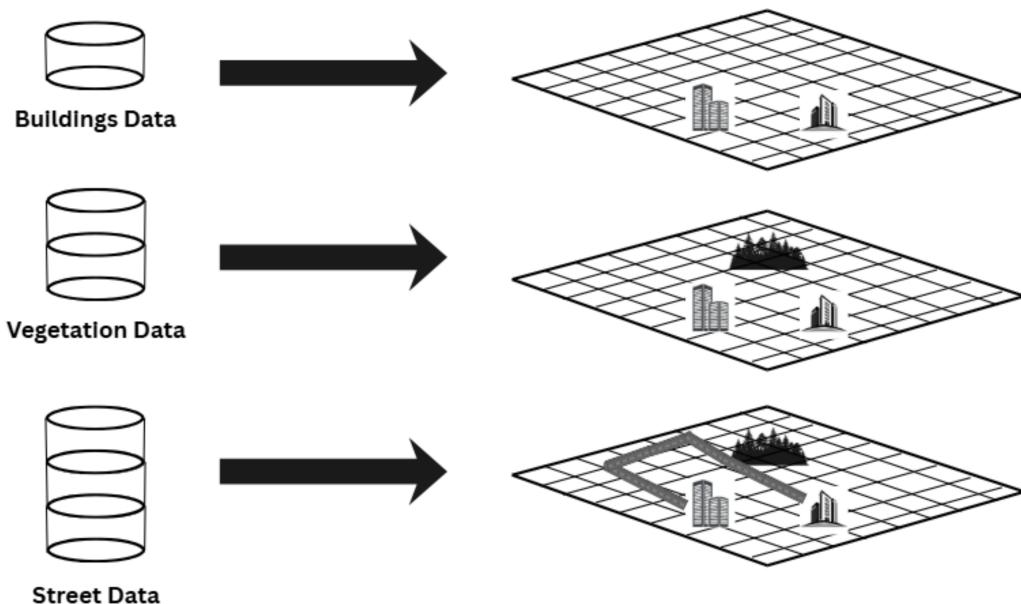
$$Imp_j = M_{awal} - M_{acak,j}. \quad (\text{II.17})$$

Jika nilai  $Imp_j$  besar, artinya ketika fitur itu dihilangkan, performa model turun jauh yang berarti fitur tersebut memiliki kontribusi yang besar dalam langkah prediksi model. Jika hampir tidak ada perubahan, fitur tersebut tidak berpengaruh. Pengulangan permutasi disarankan untuk memperoleh distribusi *importance* dan menguji apakah nilai fitur tersebut benar-benar signifikan.

## II.8 Sistem Informasi Geospasial

Sistem informasi geospasial atau *Geographic Information System* (GIS) merupakan sistem yang dirancang untuk mengelola, mengorganisasi, menganalisis, dan memvisualisasikan data spasial dan alfanumerik yang berguna untuk membantu sistem pendukung keputusan (Teixeira dkk. 2021). Penelitian tersebut menekankan bahwa perkembangan GIS telah memperluas ketersediaan dan pemanfaatan *geospatial tools* untuk menganalisis data ruang di berbagai skala. Dengan kata lain, GIS dapat menghubungkan lokasi dengan lebih spesifik dan kompleks dan digunakan sebagai analisis pengambilan keputusan.

Secara umum, data geospasial pada GIS dipresentasikan dalam dua bentuk utama, yaitu vektor dan raster. Data vektor menyimpan geometri sebagai titik, garis, atau poligon untuk merepresentasikan objek dalam bentuk diskrit, seperti jalan, gedung, vegetasi, dan lainnya. Sementara itu, data raster memodelkan permukaan bumi sebagai *grid* sel dengan suatu resolusi tertentu. Pada Gambar II.7, bentuk *layer-layer* digabungkan sehingga memungkinkan operasi analisis seperti *overlay*, perhitungan jarak, pemotongan wilayah *clip*, dan *resampling* solusi. Pada salah satu contoh studi untuk menganalisis karhutla, Vetrata dkk. (2025) memanfaatkan GIS untuk menyajikan pemetaan bulanan area terbakar di Indonesia dengan memanfaatkan citra satelit dan serangkaian data geospasial sehingga analisis dapat dilakukan untuk melihat deret waktu area terbakar yang konsisten.



Gambar II.7 Contoh *layer* GIS (GAO 2003)

## II.9 Penelitian Terkait

Untuk memahami lebih dalam terkait permasalahan, studi dilakukan untuk mencari penelitian yang membahas permasalahan dan mencoba memahami langkah solutif yang telah di uji coba sebagai percobaan penyelesaian masalah. Penelitian yang dilakukan oleh Karurung, Lee, dan Lee (2025) menjelaskan tentang kebakaran hutan yang berlokasi di Indonesia dengan fokus pada variasi musiman. Wilayah studi yang dilakukan meliputi provinsi Kalimantan, Sumatra, dan Papua. Mereka mengembangkan sebuah algoritma ML termasuk RF dan XGBoost dari lima belas faktor kerentanan, seperti curah hujan, suhu, kelembapan, indeks vegetasi, topografi, jarak permukiman, hingga keberadaan lahan gambut. Prediksi yang dilakukan berfokus pada pemetaan tingkat kerentanan kebakaran pada tiga kondisi musim, yaitu hujan, kering, dan gabungan. Hasilnya menunjukkan bahwa model berbasis *ensemble tree* mampu mencapai akurasi tinggi dengan nilai AUC di atas 0,9 dan mengungkap bahwa pengaruh faktor lingkungan seperti curah hujan dan kelembapan menjadi dominan.

Studi selanjutnya dilakukan dengan melihat penelitian yang dilakukan Chen dkk. (2024) yang memanfaatkan model *Support Vector Machine* (SVM), RF, dan *Neural Network* dalam melakukan klasifikasi area *burning* dan *burnt* pada tiga periode waktu (sebelum kebakaran, saat kebakaran, dan setelah kebakaran). Mereka menggunakan citra Sentinel-1B dan Sentinel-2A untuk mendapatkan citra di lokasi

Xintian County, Tiongkok. Fitur yang digunakan merupakan kombinasi dari index vegetasi, parameter radar, dan kanal optik. Secara umum, model SVM menunjukkan kinerja paling baik pada fase pra-kebakaran, sedangkan RF dan *Neural Network* memberikan hasil yang lebih baik untuk pemetaan area yang sudah terbakar. Studi ini menegaskan bahwa pemilihan algoritma dan fitur citra sangat mempengaruhi kualitas pemetaan kebakaran berbasis penginderaan jauh (*remote sensing*).

Studi dari Abdollahi dan Pradhan (2023) memberikan pandangan lain tentang prediksi kerentanan kebakaran hutan dengan memberikan pendekatan XAI untuk memberikan penjelasan dari hasil prediksi. Wilayah studi yang dilakukan berada di Gippsland, Victoria (Australia) dengan menggunakan serangkaian faktor pemicu kebakaran seperti curah hujan, suhu, dan topografi. Mereka menggunakan model DL untuk memprediksi kerentanan kebakaran hutan lalu menerapkan metode SHAP untuk menafsirkan kontribusi masing-masing faktor. Hasil yang diberikan SHAP mampu menunjukkan fitur-fitur kunci yang mendorong peningkatan kerentanan kebakaran seperti kelembapan, kemiringan lereng, dan jarak ke jalan. Pendekatan ini memperlihatkan bagaimana XAI dapat mengatasi kasus "*black box*" dalam menjelaskan hasil yang diberikan oleh model prediksi.

Secara keseluruhan, ketiga penelitian tersebut memberikan gambaran penerapan ML dan XAI pada kasus kebakaran hutan dan lahan. Namun, eksplorasi pemanfaatan ML dan XAI untuk kebakaran hutan di Asia Tenggara, khususnya Indonesia, masih relatif terbatas. Integrasi berbagai langkah solutif seperti pemodelan kerentanan berbasis *ensemble tree*, pemanfaatan berbagai *raster layer* lingkungan dalam kerangka SIG, serta penerapan XAI yang terhubung dengan dashboard interaktif belum banyak dikaji secara terpadu. Ruang inilah yang akan diisi oleh tugas akhir ini.

## **BAB III**

### **ANALISIS MASALAH**

#### **III.1 Analisis Kondisi**

Bab ini membahas kondisi aktual dari karhutla di Indonesia, khususnya di wilayah Kalimantan serta bagaimana kondisi mempengaruhi kebutuhan akan sistem prediksi risiko karhutla berbasis data. Analisis yang dilakukan mengikuti alur metodologi dari CRISP-DM, yaitu dimulai dari pemahaman terhadap konteks masalah, proses bisnis, dan pemangku kepentingan yang terlibat dalam pengendalian bencana karhutla (*business understanding*). Selanjutnya, pemahaman terhadap data yang digunakan untuk memahami karakteristik seputar karhutla dilakukan untuk memahami bagaimana solusi-solusi yang telah diterapkan saat ini dalam menghadapi langkah mitigasi karhutla (*data understanding*).

Melalui analisis kondisi ini, diharapkan kondisi kesenjangan (*gap*) antara kebutuhan lapangan dengan kemampuan sistem dapat terlihat dengan jelas. Kesenjangan inilah yang kemudian akan dimanfaatkan untuk menjadi dasar dalam perumusan kebutuhan sistem, perancangan alternatif solusi, serta pemilihan pendekatan yang paling relevan untuk dikembangkan pada tugas akhir ini.

##### **III.1.1 *Business Understanding***

###### **III.1.1.1 Kondisi Saat Ini**

Walaupun Indonesia merupakan salah satu negara tropis terbesar di dunia, ancaman terhadap karhutla masih terjadi secara berulang dan menimbulkan dampak yang luas terhadap ekologi, ekonomi, kesehatan, dan sosial. Karhutla di Indonesia tidak hanya dipengaruhi oleh satu faktor, melainkan merupakan gabungan atau hasil interkasi antara variabilitas iklim, kondisi biofisik, dan aktivitas manusia sebagaimana telah dibahas pada Bab II. Hal ini membuat pola terjadinya karhutla menjadi kompleks, berbeda antar wilayah, dan berubah dari waktu ke waktu.

Saat ini, beberapa lembaga di Indonesia telah membuat sistem untuk mengendalikan dan mengelola bencana karhutla sebagai langkah pemantauan dan peringatan dini. Kementerian Lingkungan Hidup Dan Kehutanan (KLHK) mengembangkan sebuah sistem melalui SiPongi+ dengan menyajikan informasi *hotspot* saat ini dan luas kebakaran melalui sebuah situs web interaktif yang dapat diakses secara terbuka. Sementara itu, BMKG menyediakan indeks bahaya kebakaran melalui perhitungan matematis, seperti *Fire Weather Index* (FWI), *Initial Spread Index* (ISI), dan lainnya. Sistem-sistem ini sangat penting untuk memantau titik api dan kondisi bahaya kebakaran secara harian untuk mendukung respon reaktif yang cepat di lapangan. Namun dari sudut pandang pengelolaan risiko kebakaran, masih terdapat sejumlah tantangan yang dihadapi pemangku kepentingan, khususnya ketika mereka ingin menyusun langkah mitigasi yang lebih proaktif dan terarah karena kurangnya informasi dengan jangka waktu yang lebih panjang.

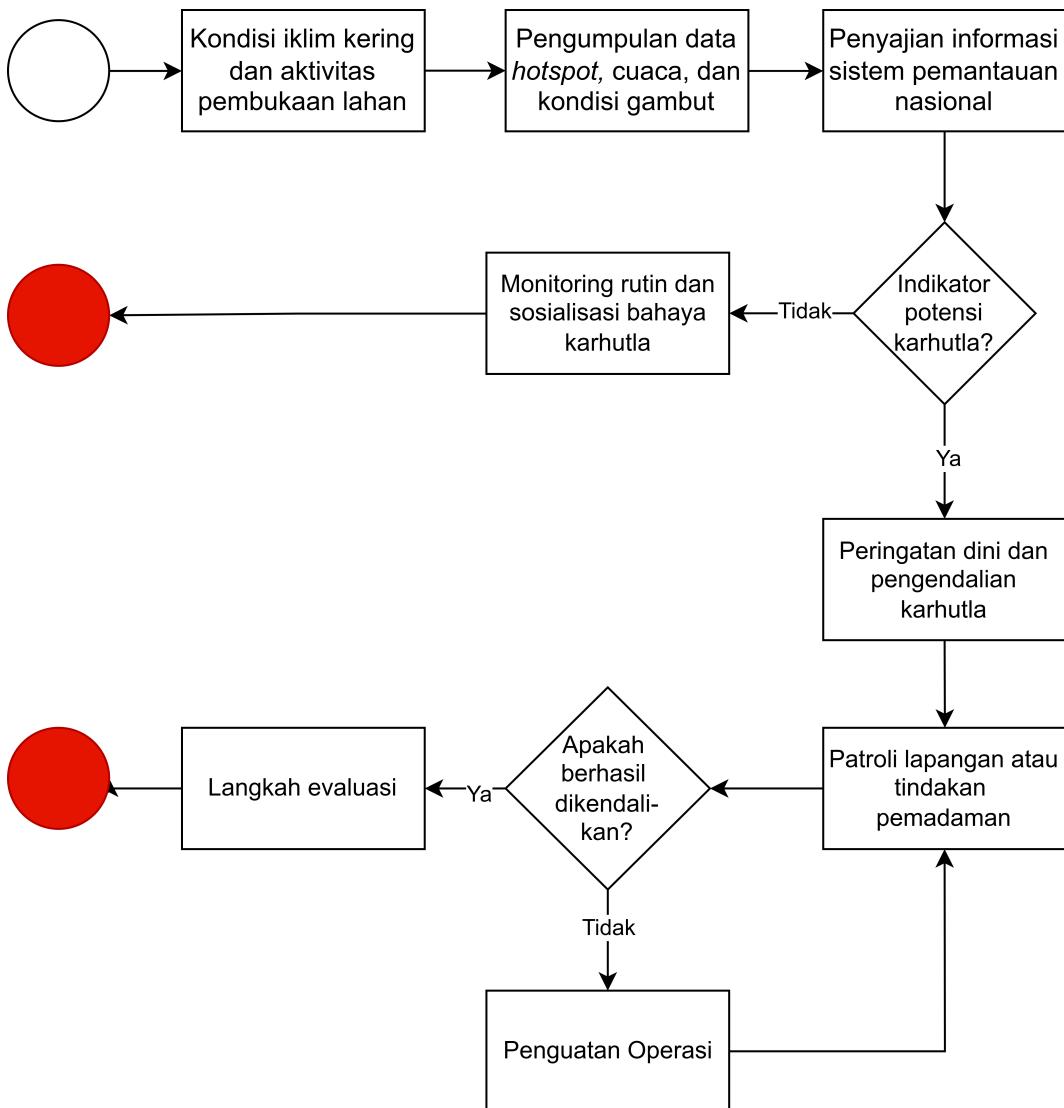
Secara ringkas, alur pengendalian karhutla yang berjalan di Indonesia saat ini digambarkan pada model konseptual pada Gambar III.1. Pembuatan model konseptual tersebut dilakukan dengan memanfaatkan informasi yang didapatkan pada situs resmi KLHK dan berdasarkan studi dari Prayoga dan Koestoer (December 2021) sehingga memiliki validasi terhadap kondisi saat ini.

Proses tersebut diawali oleh kondisi iklim dan aktivitas pembukaan lahan yang kemudian akan meningkatkan potensi munculnya titik api. Langkah selanjutnya adalah pengumpulan data titik api, cuaca, dan kondisi lahan gambut yang didapatkan dari penyajian informasi lewat sistem pemantauan nasional seperti SiPongi+. Informasi tersebut digunakan untuk menilai apakah indikator menunjukkan potensi karhutla sehingga terdapat langkah tindak lanjut yang akan dilakukan. Apabila kebakaran belum terkendali, terdapat langkah penguatan operasi. Namun bila sudah terkendali, tahap evaluasi akan dilakukan. Langkah ini memiliki satu kelemahan efektivitas, yaitu keputusan penanggulangan baru direncanakan apabila titik api telah terdeteksi sehingga bisa menghambat efektivitas penanggulangan. Prayoga dan Koestoer (December 2021) juga menilai bahwa alur ini cenderung berorientasi pada respon sehingga diperlukan penguatan tahap kesiapsiagaan melalui pemanfaatan data yang lebih ilmiah dan prediktif.

Beberapa tantangan utama yang muncul antara lain:

1. Lingkungan risiko yang kompleks dan dinamis.

Risiko karhutla ditentukan oleh kombinasi faktor meteorologik (cuaca, suhu, angin, ENSO), faktor biofisik (jenis tutupan lahan, lahan gambut, topografi),



Gambar III.1 Konsep sistem penanganan karhutla di Indonesia saat ini

dan faktor antropogenik (aktivitas manusia). Kombinasi ini membuat langkah rekognisi yang telah dilakukan saat ini menggunakan pendekatan WFI menjadi kurang relevan disebabkan hanya bergantung pada satu faktor, yaitu faktor meteorologik.

2. Pola kejadian yang tidak merata secara ruang dan waktu.  
Data historis menunjukkan bahwa kejadian karhutla cenderung terjadi pada wilayah dan periode tertentu, misalnya pada tipe lahan gambut dan saat musim kemarau yang diperkuat dengan fenomena ENSO. Beberapa kawasan mengalami kebakaran berulang, Sementara area lain relatif jarang terbakar. Tanpa alat bantu analisis yang memanfaatkan pola historis ini secara sistematis, pengelola akan kesulitan mengidentifikasi wilayah mana yang rentan sehingga perlu menjadi prioritas mitigasi.

3. Fokus informasi pada kejadian dan bahaya jangka pendek.

Informasi yang disediakan oleh pemantauan yang ada umumnya berfokus pada langkah reaktif dengan memanfaatkan indeks bahaya kebakaran berbasis cuaca. Informasi tersebut sangat berguna untuk operasi tanggap darurat, namun belum sepenuhnya menjawab kebutuhan perencanaan mitigasi. Patroli yang dilakukan berdasarkan data saat ini atau historis yang mungkin saja menyebabkan keterlambatan tindakan. Hal ini disebabkan belum adanya acuan prediksi yang bisa dimanfaatkan untuk menentukan wilayah yang menjadi prioritas sehingga bisa dilakukan operasi tanggap darurat dengan lebih cepat.

4. Terbatasnya pemahaman terstruktur tentang faktor penyebab.

Bagi pemangku kepentingan, pertanyaan yang diajukan sebelum mengambil keputusan bukan hanya tentang "di mana risiko tertinggi saat ini", tetapi juga "mengapa wilayah tersebut memiliki risiko yang tinggi". Saat ini, hubungan antara faktor-faktor penyebab karhutla sering kali dianalisis secara terpisah. Hal tersebut akan menghambat pemangku kepentingan dalam melakukan penyusunan keputusan strategi mitigasi karhutla.

5. Langkah penguatan operasi yang belum optimal.

Untuk meningkatkan langkah operasional dalam mengendalikan bencana karhutla, pemenuhan sumber daya sering kali menjadi tantangan dalam hal ini. Dalam meningkatkan operasi, mobilisasi sumber daya dan rencana pengendalian membutuhkan waktu perencanaan administrasi dan eksekusi fisik yang memakan waktu mingguan. Maka dari itu, diperlukan sebuah informasi dari perencanaan dari jauh-jauh hari untuk membuat keputusan strategis yang tidak mendadak.

Rangkaian tantangan di atas menunjukkan bahwa meskipun Indonesia saat ini telah memiliki beberapa sistem pemantauan dan peringatan dini, masih terdapat kesenjangan dalam hal ketersediaan informasi risiko untuk membantu melakukan langkah mitigasi dengan mengetahui wilayah mana yang paling rentan diserta faktor-faktor apa yang paling berkontribusi. Analisis inilah yang akan menjadi fokus pada tahap selanjutnya dalam tugas akhir untuk menentukan langkah yang tepat untuk membantu mengatasi permasalahan ini.

### **III.1.2 Kebutuhan Fungsional**

Berdasarkan pemahaman terhadap masalah, langkah selanjutnya adalah merumuskan kebutuhan sistem secara konseptual untuk mengetahui bagaimana sistem yang relevan dengan permasalahan. Kebutuhan fungsional menggambarkan fungsi inti apa saja yang harus disediakan oleh suatu sistem informasi pendukung risiko kar-

Tabel III.1 Kebutuhan Fungsional Sistem (versi ringkas dan problem-driven)

ID	Kebutuhan Fungsional	Penjelasan singkat (dikaitkan dengan masalah)
F1	Integrasi data karhutla dan faktor lingkungan	Sistem dapat mengintegrasikan data historikal karhutla dengan data faktor lingkungan yang relevan agar analisis hasil memiliki dasar yang jelas.
F2	Analisis pola kebakaran	Sistem mampu mengidentifikasi pola kebakaran secara spasial dan temporal berdasarkan rangkaian data historis.
F3	Analisis keterkaitan faktor risiko dengan kejadian	Sistem perlu menyediakan hasil analisis pola dan penyajian informasi analisis untuk memperbaiki hasil transparansi pola.
F4	Penyajian informasi visual	Sistem mampu menyajikan informasi pola dan faktor penyebab untuk memberikan wawasan kepada pemangku kepentingan.

hutla. Kebutuhan fungsional berguna untuk menyusun dan memahami bagaimana sistem harus mampu untuk memberikan fitur yang dapat membantu mengatasi permasalahan. Kebutuhan fungsional yang diidentifikasi dirangkum pada Tabel III.1.

### III.1.3 Kebutuhan Non-fungsional

Kebutuhan non-fungsional tidak berbicara tentang "apa yang dilakukan" sistem, melainkan bagaimana karakteristik sistem tersebut agar relevan dengan konteks pengendalian karhutla di Indonesia. Dengan menganalisis kebutuhan non-fungsional dari sistem, pengetahuan terhadap kelayakan teknis dapat dipahami untuk pengembangan yang relevan. Berikut hasil analisis kebutuhan non-fungsional pada permasalahan terkait.

Selain itu, pada tugas akhir ini kebutuhan non-fungsional difokuskan pada tiga aspek utama, yaitu akurasi model, ketersediaan dan waktu respon sistem, serta kemudahan penggunaan. Akurasi model diperlukan agar informasi risiko yang dihasilkan tidak menyesatkan pengambil keputusan, ketersediaan dan waktu respon menentukan sejauh mana sistem dapat mendukung analisis secara interaktif, sedangkan kemudahan penggunaan memastikan bahwa sistem dapat dimanfaatkan oleh pemangku kepentingan yang tidak selalu memiliki latar belakang teknis. Rangkuman kebutuhan non-fungsional tersebut disajikan pada Tabel III.2.

Tabel III.2 Kebutuhan non-fungsional sistem

ID	Kebutuhan Non-fungsional	Indikator / Target
NF1	Akurasi Model	Model analisis/prediksi risiko karhutla yang digunakan dalam sistem harus mencapai tingkat kinerja yang dapat diterima, dengan nilai <i>accuracy</i> mencapai 0,80 dan <i>F1-score</i> pada kelas risiko tinggi adalah 0,75.
NF2	Keandalan	Sistem mampu memberikan hasil deteksi atau analisis risiko karhutla yang konsisten dan dapat dipercaya, disertai bentuk keluaran yang jelas (misalnya kelas risiko atau skor risiko) sehingga tidak menimbulkan ambiguitas.
NF3	Kemudahan penggunaan	Sistem harus disajikan dalam bentuk yang mudah dipahami dan cukup responsif, sehingga pengguna non-teknis dapat mengoperasikan fitur utama.

### III.2 Analisis Pemilihan Solusi

Berdasarkan hasil identifikasi kebutuhan fungsional dan non-fungsional, langkah penentuan beberapa alternatif solusi dilakukan untuk mencari solusi yang secara konseptual dapat membantu memenuhi kebutuhan tersebut. Alternatif-alternatif ini disusun dengan merangkup pendekatan yang digunakan pada penelitian terdahulu dan membandingkannya dengan rancangan sistem yang diusulkan pada tugas akhir. Agar pemilihan tidak hanya berdasarkan preferensi subjektif, proses evaluasi alternatif dilakukan dengan menggunakan pendekatan *Multi-Criteria Decision Analysis* (MCDA) dengan metode pembobotan sederhana yang diadaptasi dari studi O'Shea dkk. (2026).

#### III.2.1 Alternatif Solusi

Untuk langkah prediktif karhutla, terdapat beberapa pendekatan yang digunakan, terutama dalam aspek penyampaian informasi dan akurasi model. Alternatif solusi yang dipertimbangkan mencakup tiga pendekatan dari literatur dan satu pendekatan yang diusulkan dalam tugas akhir ini.

##### 1. Alternatif Solusi 1 (A1): Model Kerentanan Karhutla Berbasis Random

**Forest dan XGBoost** (Karurung, Lee, dan Lee 2025).

Karurung, Lee, dan Lee (2025) mengembangkan model kerentanan kebakaran hutan di provinsi Sumatra, Kalimantan, dan Papua menggunakan algoritma RF dan XGBoost dengan berbagai faktor kondisi (meteorologis, biofisik, dan antropogenik). Penelitian ini menghasilkan peta kerentanan dan *grid* risiko serta menunjukkan bahwa pemisahan data berdasarkan musim dapat meningkatkan performa model. Pendekatan ini efektif dalam pengintegrasian data dan analisis pola kebakaran karena mendapatkan akurasi model yang tinggi. Namun, penyajian yang diberikan belum disajikan dalam bentuk visual sederhana sehingga sulit dipahami bagi pengguna.

2. **Alternatif Solusi 2 (A2): Sistem Pemantauan dan Informasi *Hotspot Karhutla SiPongi+*** (Kementerian Lingkungan Hidup dan Kehutanan (KLHK) 2025).

KLHK melalui SiPongi+ merepresentasikan sistem pemantauan karhutla secara *near real-time* dilengkapi riwayat kejadian dan level risiko kebakaran. Sistem ini disajikan melalui perhitungan indikator *hotspot* lewat FWI serta antarmuka peta dan grafik di web serta memiliki respon cepat untuk langkah karhutla secara reaktif. Walaupun begitu, sistem ini masih lebih bersifat *early warning* sehingga belum mampu melakukan analisis pola historis dan integrasi faktor untuk melakukan langkah prediktif menggunakan *machine learning* sebagai acuan langkah mitigasi untuk pemangku kepentingan.

3. **Alternatif Solusi 3 (A3): Model Kerentanan Karhutla Berbasis Explainable AI (XAI)** (Abdollahi dan Pradhan 2023).

Abdollahi dan Pradhan (2023) mengembangkan model kerentanan kebakaran dengan memanfaatkan teknik XAI untuk menjelaskan kontribusi masing-masing faktor risiko terhadap kerentanan kebakaran. Studi ini menghasilkan grafik kontribusi dari masing-masing hasil yang diprediksi model *machine learning* melalui penjelasan kontribusi faktor. Alternatif ini relevan dengan kebutuhan sistem karena menunjukkan integrasi data karhutla disertai penjelasan pola yang dapat memberikan penjelasan terhadap hasil model yang dapat membantu pengambilan keputusan pemangku kepentingan. Namun penyajian yang dilakukan masih bersifat statis sehingga belum dapat memberikan informasi visual yang mudah digunakan oleh pengguna.

4. **Alternatif Solusi 4 (A4): Sistem Informasi Prediksi Karhutla Berbasis *Machine Learning*, XAI, dan Penyajian *Dashboard* Interaktif.**

Alternatif ini merupakan rancangan yang diusulkan dalam tugas akhir ini. Sistem ini dirancang sebagai prototipe sistem analitik berbasis data historis. Sis-

tem ini mengintegrasikan data historis *hotspot* dengan faktor lingkungan yang relevan untuk dijadikan analisis prediktif. Hasil prediksi yang diberikan akan dijelaskan melalui XAI untuk menghilangkan aspek "black box" dan disajikan dalam bentuk *dashboard* interaktif sehingga dapat membantu pengambilan keputusan mitigasi pengelolaan karhutla oleh pemangku kepentingan.

### III.2.2 Analisis Penentuan Solusi

Setelah empat alternatif solusi diidentifikasi, langkah berikutnya adalah menentukan solusi mana yang paling sesuai untuk dikembangkan pada tugas akhir ini. Penentuan solusi dilakukan berdasarkan pendekatan MCDA agar menghindari aspek subjektif. Secara umum, MCDA digunakan ketika terdapat beberapa alternatif dan kriteria yang harus dipertimbangkan. Menurut O'Shea dkk. (2026), konsep penentuan solusi dilakukan dengan mendefinisikan kriteria, menentukan bobot pada setiap kriteria, memberikan penilaian terhadap setiap alternatif, dan menggabungan penilaian sehingga menjadi satu skor total untuk mengambil keputusan solusi. Dalam penelitian ini, metode MCDA yang digunakan adalah *Weighted Sum Model* (WSM) dengan alasan kesederhanaan konsep dan transparansi serta cocok untuk jumlah kriteria yang tidak terlalu banyak. Konsep WSM dijelaskan pada

$$S_i = \sum_{j=1}^n w_j x_{ij} \quad (\text{III.1})$$

dengan :

- $S_i$  = skor total alternatif ke- $i$ ,
- $w_j$  = bobot kriteria ke- $j$  (dengan  $\sum_{j=1}^m w_j = 1$ ),
- $x_{ij}$  = skor alternatif ke- $i$  pada kriteria ke- $j$ .

Kriteria diambil dari kebutuhan fungsional dan non-fungsional dan disederhanakan menjadi empat kriteria:

1. Kesesuaian dengan kebutuhan fungsional (C1): Sejauh mana alternatif mampu memenuhi empat kebutuhan fungsional utama.
2. Akurasi dan kualitas informasi risiko (C2): Seberapa besar potensi pendekatan tersebut untuk menghasilkan model risiko yang memadai sehingga tidak menyesatkan pemangku. Kriteria ini terkait langsung dengan kebutuhan non-fungsional NF1.
3. Kinerja dan kelayakan implementasi (C3): Seberapa realistik suatu pendekatan diwujudkan sebagai prototipe tugas akhir, dilihat dari ketersediaan data, kompleksitas model, dan kebutuhan komputasi berdasarkan NF2.

4. Kemudahan penggunaan (C4): Sejauh mana pendekatan tersebut memberikan penyajian yang mudah dipahami oleh pemangku kepentingan dan membantu mengambil keputusan yang sesuai dengan NF3.

Bobot masing-masing kriteria ditentukan dengan mempertimbangkan tujuan tugas akhir dengan berdasarkan pertimbangan kebutuhan sistem. Berdasarkan pertimbangan tersebut, bobot yang digunakan adalah sebagai berikut.

1. C1 (kesesuaian dengan kebutuhan fungsional): 0,35
2. C2 (akurasi dan kualitas informasi risiko): 0,25
3. C3 (kinerja dan kelayakan implementasi): 0,20
4. C4 (kemudahan penggunaan): 0,20

Bobot tersebut memiliki total 1,00 dan menyesuaikan dengan kebutuhan fungsional dan non-fungsional. Dari keempat alternatif yang diajukan, penilaian dilakukan dengan menggunakan nilai kriteria yang telah ditentukan sebelumnya. Perhitungan yang dihasilkan akan menentukan hasil akhir yang bertujuan sebagai penentuan solusi. Penilaian masing-masing alternatif dilakukan dengan melihat kriteria pada skala 1-5 dengan penjelasan angka 1 berarti sangat kurang dalam memenuhi kriteria dan angka 5 berarti sangat memenuhi kriteria. Skor ditentukan secara kualitatif berdasarkan kelebihan dan kekurangan masing-masing alternatif yang telah dijelaskan sebelumnya pada Subbab III.2.1 Alternatif Solusi.

Secara ringkas, A1 dinilai kuat pada pemodelan risiko namun belum menyediakan sistem informasi yang siap pakai. A2 dinilai kuat pada penyajian pemantauan operasional, namun sangat lemah pada analisis historis dan pemodelan prediktif. A3 cenderung kuat pada akurasi dan memiliki interpretabilitas pada hasil prediksi, tetapi visualisasi yang diberikan masih statis dan fokusnya cenderung pada hasil analisis. A4 dirancang khusus untuk menyeimbangkan kekurangan dari keempat kriteria dalam konteks provinsi Kalimantan. Matriks penilaian alternatif dan skor total dihitung menggunakan rumus WSM sesuai pada persamaan III.1 dan dirangkum pada Tabel III.3

Tabel III.3 Matriks penilaian alternatif solusi menggunakan WSM

Alternatif	C1	C2	C3	C4	Skor total $S_i$
Bobot $w_j$	0,35	0,25	0,20	0,20	—
A1	4	4	3	3	3,60
A2	3	2	2	4	2,75
A3	4	4	3	4	3,80
A4	5	4	4	4	4,35

Berdasarkan Tabel III.3, Alternatif Solusi 4 (A4) memperoleh skor total tertinggi ( $S_{A4} = 4,35$ ) sehingga dipilih sebagai solusi yang akan dikembangkan pada tugas akhir ini. A4 dinilai paling seimbang dalam memenuhi kebutuhan fungsional dan non-fungsional, serta masih realistik untuk diwujudkan sebagai prototipe sistem analitik berbasis data historis untuk Kalimantan.

Alternatif lain tetap dimanfaatkan sebagai rujukan: A1 dan A3 sebagai acuan pemodelan dan pendekatan XAI, sedangkan A2 sebagai gambaran kondisi sistem pemantauan karhutla yang sudah ada (SiPongi+) dan dasar untuk mengidentifikasi kebutuhan analisis yang belum terpenuhi.

### **III.3 Data Understanding**

Dalam mengembangkan solusi, langkah awal yang dilakukan adalah mencari dan memahami data-data yang tersedia untuk memahami konteks permasalahan dengan lebih dalam. Selain itu, data juga bermanfaat untuk melatih model prediktif sehingga dapat dimanfaatkan dalam sistem. Pencarian data bersifat wajib untuk langkah awal, namun terkadang data yang digunakan belum cukup sehingga pencarian data selanjutnya bersifat iteratif. Pada subbab ini, data yang akan dipahami adalah data utama yang akan digunakan pada tugas akhir ini, yaitu data *hotspot* karhutla dari *National Aeronautics and Space Administration* (NASA) bernama *Fire Information for Resource Management System* (FIRMS). Data tersebut merupakan data yang akan menjadi label atau target karena berisikan titik letak terjadinya *hotspot*.

#### **III.3.1 Deskripsi Umum Dataset**

Data FIRMS diperoleh dalam format *shapefile* (shp) yang dikonversi ke file *Comma Separated Values* (csv) dengan memanfaatkan *Quantum Geographic Information System* (QGIS). QGIS merupakan perangkat lunak GIS yang gratis, *open-source*, dan multifungsi yang memungkinkan pengguna untuk membuat, mengedit, memvisualisasikan, menganalisis, dan mempublikasikan data spasial atau geografis. Data FIRMS yang diperoleh berupa produk *Visible Infrared Imaging Radiometer Suite* (VIIRS) 375 m *Active Fire* untuk satelit *National Oceanic and Atmospheric Administration* (NOAA-20). Satelit ini memiliki keunggulan berupa resolusi spasial yang lebih tinggi dibandingkan *Moderate Resolution Imaging Spectroradiometer* (MODIS) 1 km sehingga mampu mendeteksi panas berskala kecil seperti kebakaran lahan gambut atau pembukaan lahan dengan api.

*Dataset* ini diambil pada situs resmi NASA FIRMS. *Dataset* ini berisi 846.831 baris

data *hotspot* dengan 15 fitur utama yang mencakup periode observasi dari 1 April 2018 sampai 1 Januari 2023. Pemilihan lokasi data befokus pada wilayah Indonesia. Untuk memberikan gambaran awal tentang data, Tabel III.4 menunjukkan deskripsi singkat tentang seluruh fitur beserta contoh nilai dari masing-masing fitur.

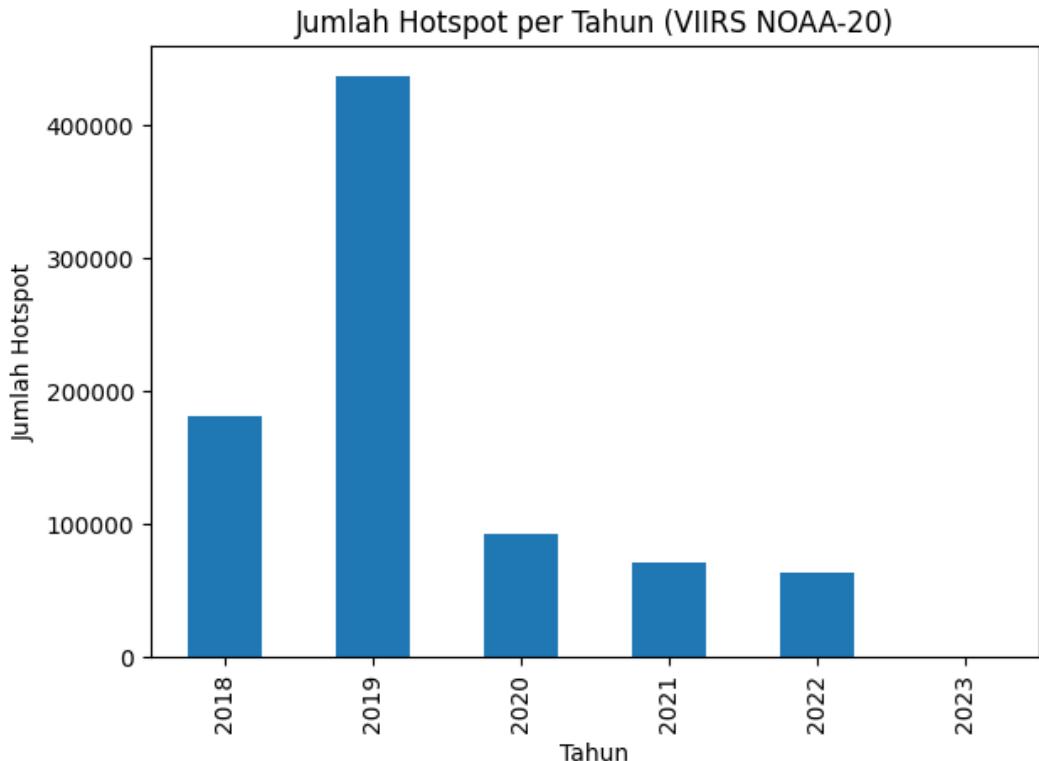
Tabel III.4 Deskripsi atribut pada *dataset VIIRS 375 m*

Nama Fitur	Tipe Data	Deskripsi Singkat	Contoh Nilai
LATITUDE	<i>float</i>	Lintang geografis titik <i>hotspot</i> (derajat, WGS84)	-3.85
LONGITUDE	<i>float</i>	Bujur geografis titik <i>hotspot</i> (derajat, WGS84)	136.36
BRIGHTNESS	<i>float</i>	<i>Brightness temperature</i> kanal termal utama VIIRS (Kelvin)	330.34
SCAN	<i>float</i>	Ukuran piksel pada arah <i>scan</i> satelit (derajat)	0.40
TRACK	<i>float</i>	Ukuran piksel pada arah <i>track</i> satelit (derajat)	0.37
ACQ_DATE	<i>string</i>	Tanggal akuisisi dalam UTC (format YYYY/MM/DD)	2018/04/01
ACQ_TIME	<i>int</i>	Waktu akuisisi dalam UTC (HHMM)	416
SATELLITE	<i>string</i>	Kode satelit penginderaan; seluruh record berisi N20 (NOAA-20)	N20
INSTRUMENT	<i>string</i>	Nama sensor; seluruh record berisi VIIRS	VIIRS
CONFIDENCE	<i>string</i>	Tingkat kepercayaan deteksi <i>hotspot</i> (l=low, n=nominal, h=high)	n
VERSION	<i>int</i>	Versi produk FIRMS; seluruh record berisi 2 (Collection 2)	2
BRIGHT_T31	<i>float</i>	<i>Brightness temperature</i> kanal referensi T31 (Kelvin)	297.54
FRP	<i>float</i>	<i>Fire Radiative Power</i> (MW), indikator intensitas kebakaran	3.93
DAYNIGHT	<i>string</i>	Penanda kondisi siang/malam saat deteksi (D = day, N = night)	D
TYPE	<i>int</i>	Kode jenis deteksi sesuai definisi FIRMS	0

### III.3.2 Kelengkapan dan Distribusi Data

Pemeriksaan awal terhadap struktur data adalah dengan menunjukkan bahwa seluruh fitur tidak memiliki nilai kosong (*missing values*) dan data duplikat. Hasil analisis menemukan bahwa data tidak memiliki nilai kosong dan tidak ditemukan

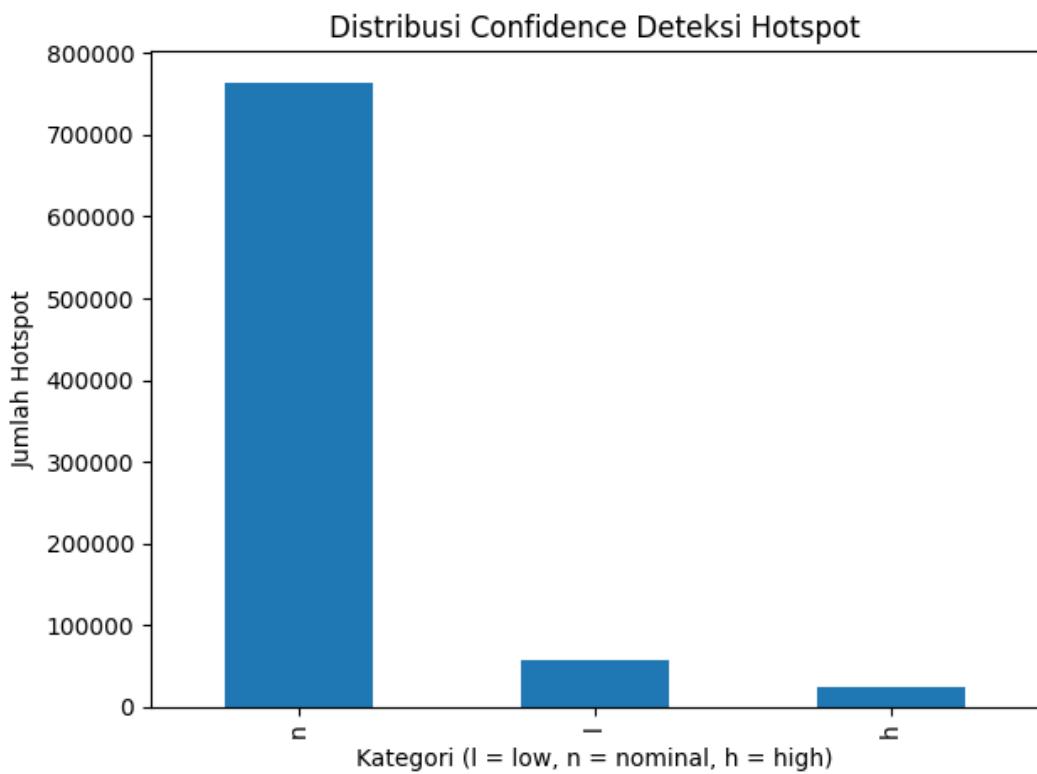
baris duplikat. Selanjutnya, analisis distribusi data dilakukan dengan melihat jumlah *hotspot* berdasarkan data tahunan. Distribusi jumlah *hotspot* per tahun ditunjukkan pada 2.



Gambar III.2 Jumlah *hotspot* per tahun (VIIRS NOAA-20)

Pada Gambar III.1, jumlah data *hotspot* yang terjadi pada tahun 2023 sangat kecil. Hal ini tidak sejalan dengan pola tahunan sebelumnya maupun statistik kebakaran dari sumber resmi yang ditunjukkan pada Bab 1. Hal ini bisa jadi mengindikasikan bahwa data tahun 2023 pada *dataset* ini kemungkinan tidak lengkap disebabkan hanya mengambil bulan Januari pada tanggal 1. Oleh karena itu, pada tahap pembedahan selanjutnya, data tahun 2023 direncanakan untuk tidak digunakan. Namun, data ini akan tetap dilaporkan sebagai bagian dari kondisi *dataset*.

Analisis distribusi selanjutnya yang dilakukan adalah dengan menilai distribusi dari fitur tingkat kepercayaan deteksi (*confidence*). Hal ini bertujuan untuk memahami bagaimana *dataset* ini merepresentasikan titik api karena seluruh titik api yang muncul dan terdeteksi oleh satelit akan dimasukkan sebagai data. Ini bisa menjadi kesalahpahaman karena bisa saja titik api yang ditunjukkan bukanlah sebuah kebakaran, namun hanya percikan api yang mungkin disebabkan oleh aktivitas manusia. Distribusi intensitas api ditunjukkan pada Gambar III.2.



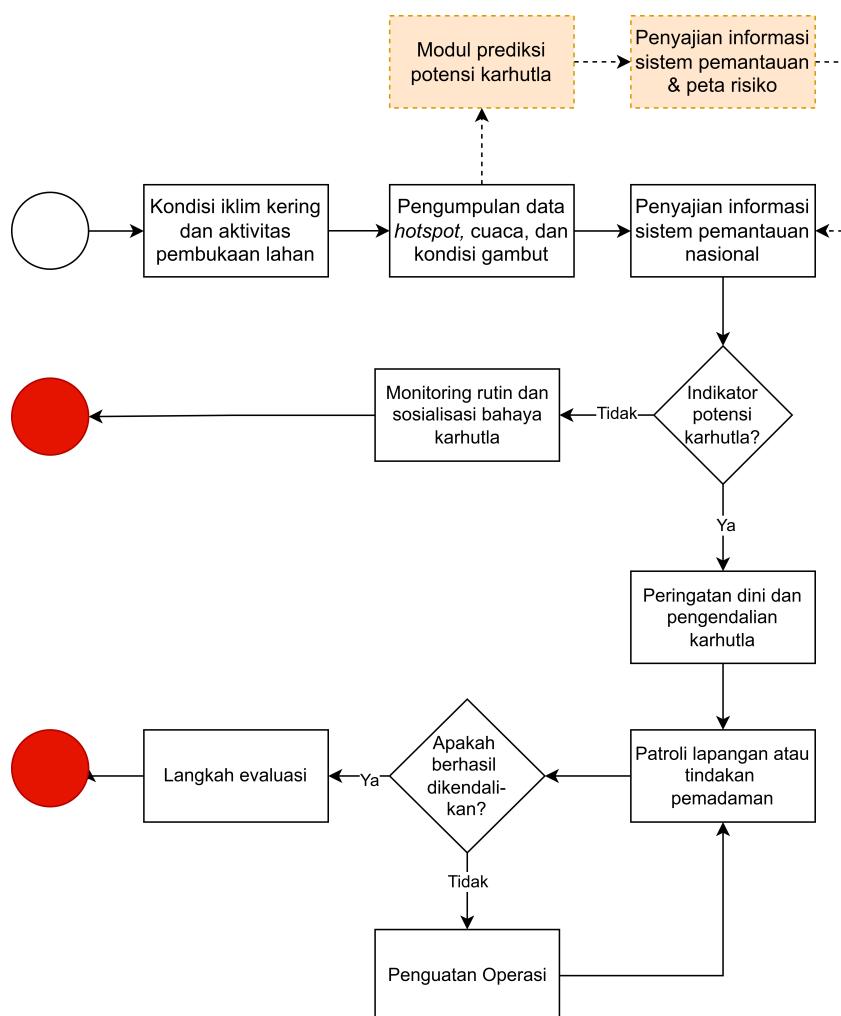
Gambar III.3 Data persebaran fitur *confidence*

Pada Gambar III.3, dominasi kategori nominal menunjukkan bahwa sebagian besar deteksi berada pada tingkat kepercayaan standar yang disarankan oleh FIRMS. Data dengan label *low* menunjukkan bahwa titik tersebut bisa ditandai sebagai kandidat *noise* disebabkan titik api yang dimunculkan bukan merupakan kebakaran. Maka dari itu, pada pengembangan selanjutnya persiapkan data akan memanfaatkan batas kepercayaan yang sesuai agar dapat memberikan informasi kebakaran yang relevan.

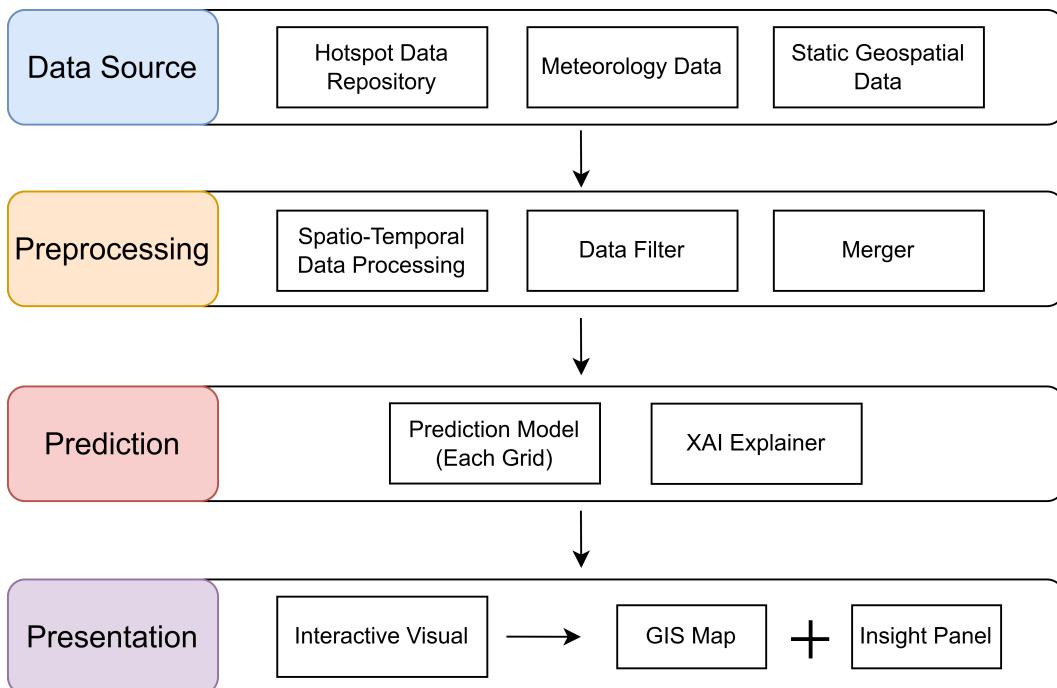
## BAB IV

### DESAIN KONSEP SOLUSI

Bab ini menjelaskan tentang desain yang diajukan berdasarkan hasil penentuan solusi alternatif pada Bab III. Penjelasan yang dilakukan berupa ilustrasi diagram sehingga terdapat konsep sistem yang menggambarkan bentuk sistem solusi.



Gambar IV.1 Model konseptual *to-be*



Gambar IV.2 Sistem prediktif dengan ML, XAI, dan penyajian informasi visual

Berdasarkan analisis masalah dan pemilihan alternatif pada Bab III, solusi yang dipilih adalah pengembangan sistem prediksi potensi karhutla berbasis ML dan XAI yang menyajikan informasi risiko dalam bentuk peta dan ringkasan faktor utama. Sistem ini ditujukan untuk melengkapi sistem pemantauan yang sudah ada sehingga pemangku kepentingan dapat memperoleh informasi risiko yang lebih proaktif untuk mendukung penentuan prioritas mitigasi.

Gambar IV.1 menunjukkan model konseptual setelah penambahan solusi usulan (*to-be*). Meninjau dari kondisi saat ini (*as-is*) yang ditunjukkan pada Gambar III.1, langkah yang dilakukan masih bersifat reaktif setelah mendapatkan acuan informasi titik api berdasarkan data dari satelit untuk merencanakan tindakan pengelolaan seperti patroli. Namun, kondisi *to-be* yang ditandai dengan kotak berwarna oranye memberikan pendekatan yang berbeda dari kondisi *as-is*, yaitu menggunakan model prediktif dalam memberikan informasi yang lebih cepat. Modul prediksi menerima data yang sudah dikumpulkan, menghasilkan level risiko karhutla per wilayah analisis, dan mengirimkannya ke sistem pemantauan sehingga indikator potensi karhutla kini mempertimbangkan informasi risiko prediktif di samping kondisi yang sedang berlangsung.

Gambar IV.2 menunjukkan konsep desain sistem prediksi potensi karhutla yang diusulkan. Secara ringkas, tahapan utama sistem dijelaskan sebagai berikut.

## 1. Pengumpulan data historis bulanan

Sistem mengumpulkan data historis titik *hotspot*, faktor antropogenik, faktor biofisik, dan faktor meteorologis untuk setiap *grid* pada wilayah Provinsi Kalimantan. Seluruh data dikumpulkan dalam satuan waktu bulanan sehingga setiap rekaman mempresentasikan kombinasi *grid*-bulan. Skala bulanan dipilih karena relevan dengan perencanaan operasional mitigasi sekaligus lebih efisien dari sisi volume dan pengolahan data

## 2. Pemrosesan dan integrasi data

Data mentah dari berbagai sumber akan melalui tahap pemrosesan seperti pembersihan nilai dan penyesuaian format. Selain itu, langkah filter dan integrasi data juga dilakukan untuk mengetahui penyebab suatu *grid* memiliki titik api atau tidak. Hasil tahap ini adalah *dataset* bulanan per *grid* yang siap digunakan untuk pelatihan dan evaluasi model

## 3. Pemodelan prediktif berbasis ML

*Dataset* terproses akan dijadikan masukan bagi model ML agar dapat menjadi model prediktif karhutla. Model ML yang digunakan adalah RF dan XGBoost agar memiliki *benchmark* dan dapat menentukan model yang terbaik. Model berbasis *ensemble tree* sangat relevan untuk memetakan hubungan kombinasi dari faktor meteorologis, biofisik, dan antropogenik. Output utama dari model ini adalah skor atau kelas (level) risiko karhutla bulanan untuk setiap *grid* wilayah. Resolusi waktu prediksi bulanan ini dipilih untuk mendukung perencanaan strategis dan manajemen logistik yang secara administratif tidak bisa dilakukan dalam siklus harian. Jadi, prediksi ini melengkapi sistem harian BMKG yang bersifat taktis dan operasional.

## 4. Analisis faktor menggunakan XAI

Untuk memberikan penjelasan dari hasil model, sistem menerapkan pendekatan XAI dengan memanfaatkan SHAP sebagai interpretabilitas lokal dan PFI untuk interpretabilitas global. Metode ini menghasilkan informasi kontribusi masing-masing faktor terhadap prediksi risiko sehingga memberikan penjelasan dari hasil model.

## 5. Penyajian informasi kepada pemangku kepentingan

Hasil dari model memberikan kerentanan risiko bulanan dengan level rendah, sedang, dan tinggi yang disertai dengan hasil analisis XAI dan disajikan dalam bentuk peta grafik risiko. Penyajian ini bertujuan agar pemangku kepentingan seperti instansi pengelola karhutla dapat memahami hasil prediksi dan menyusun rencana mitigasi dengan lebih cepat karena dapat memanfaatkan hasil prediksi sebagai acuan dasar. Walaupun prediksi bukanlah hasil

yang pasti, hasil yang diberikan model tidak sekedar sebuah hasil. Namun, hasil yang diberikan disertai penjelasan sehingga analisis lanjutan mungkin tetap perlu dilakukan sebelum menyusun rencana mitigasi.

Berdasarkan penjelasan kondisi sistem *as-is* dan *to-be* sebelumnya, analisis penjelasan perbandingan dilakukan sesuai pada Tabel IV.1. Tabel IV.1 merangkum perbedaan utama antara sistem *as-is* dan sistem *to-be*. Secara singkat, sistem usulan tidak menggantikan mekanisme pemantauan yang sudah ada, tetapi menambahkan modul prediktif bulanan dan penjelasan faktor yang diharapkan dapat menjawab kebutuhan informasi yang diidentifikasi pada Bab III.

Tabel IV.1 Perbandingan Ringkas Sistem *As-Is* dan *To-Be*

Aspek	<i>As-Is</i> (saat ini)	<i>To-Be</i> (solusi usulan)
Jenis informasi utama	<i>Hotspot</i> dan indeks bahaya kebakaran yang menggambarkan kondisi saat ini.	Level risiko karhutla bulanan per <i>grid</i> yang dibangun dari data historis multi-faktor.
Fungsi utama	Mendukung monitoring dan respon reaktif ketika bahaya sudah tinggi.	Menambah dukungan untuk perencanaan mitigasi proaktif melalui peta risiko dan informasi level potensi kebakaran.
Pemahaman faktor penyebab	Analisis faktor meteorologis, biofisik, dan antropogenik dilakukan terpisah dan bersifat kualitatif. Fokus saat ini hanya menggunakan FWI sebagai identifikasi faktor meteorologis.	<i>Explainable AI</i> memberikan kontribusi relatif tiap faktor terhadap risiko, baik pada tingkat lokal maupun global.
Bentuk penyajian	Portal pemantauan dan laporan operasional.	Visualisasi dalam bentuk peta risiko bulanan.

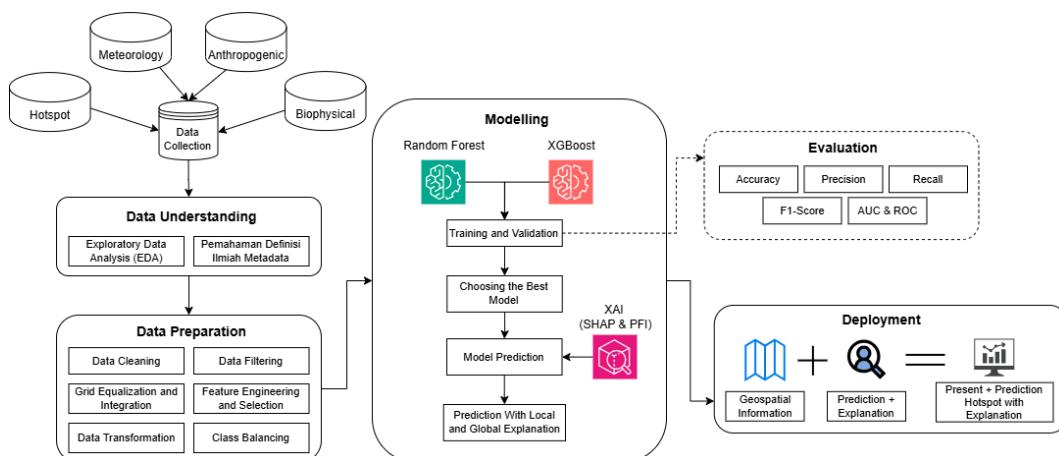
# BAB V

## RENCANA SELANJUTNYA

### V.1 Rencana implementasi

Mengacu pada metodologi CRISP-DM yang telah dijelaskan pada Bab 1, rencana implementasi akan dilakukan dengan mengikuti langkah-langkah yang terdapat pada metodologi ini. Selain itu, penjelasan seperti lingkungan pengembangan disertai alat-alat yang digunakan juga dibahas pada bagian ini. Untuk memberikan bayangan implementasi, diagram alur pengembangan dan linimasa disusun pada bagian ini.

#### V.1.1 Alur Implementasi Teknis



Gambar V.1 Diagram alur rencana implementasi

Alur keseluruhan implementasi digambarkan pada Gambar V.1 yang memperlihatkan bagaimana proses implementasi dirancang dengan mengikuti metodologi yang telah dibahas sebelumnya. Setiap blok merepresentasikan satu kelompok kegiatan utama yang akan dikerjakan selama tugas akhir.

##### 1. *Data Collection*

Tahap ini sebenarnya sempat dilakukan pada pengembangan proposal tugas

akhir dan mendapatkan data *hotspot* dari VIIRS. Namun, disebabkan hasil studi literatur sebelumnya menemukan bahwa bencana karhutla dapat diidentifikasi lewat berbagai faktor, tahap ini dilakukan kembali sebagai lanjutan dari langkah *data understanding* yang telah dilakukan pada Bab III. Data *hotspot* akan dilengkapi dengan tiga kelompok data penjelas, yaitu faktor meteorologis (curah hujan, suhu, kelembapan, dan indeks iklim), faktor biofisik (tutupan lahan, lahan gambut, indeks vegetasi, topografi), dan faktor antropogenik (kedekatan dengan pemukiman).

## 2. *Data Understanding*

Setelah data terkumpul, langkah *Exploratory Data Analysis* (EDA) dan peneilaahan metadata dilakukan untuk memahami tujuan dan maksud data, distribusi data, pola dasar, dan kualitas data. Pada tahap ini, definisi ilmiah yang terdapat pada data akan dikaji untuk memahami maksud fitur tersebut dan korelasinya dengan faktor penyebab karhutla berdasarkan studi literatur sebelumnya. Kegiatan ini memastikan bahwa data yang masuk ke tahap berikutnya memang relevan dengan tujuan pemodelan.

## 3. *Data Preparation*

Pada tahap ini, data akan disiapkan agar dapat menjadi data yang bersih untuk dijadikan bahan *training* oleh model prediktif. Output pada tahap ini adalah *dataset* tabular yang terstruktur per kombinasi *grid* dalam kurun waktu bulan dan siap untuk digunakan pada tahap pemodelan. Kegiatan *data preparation* meliputi:

- (a) *Data cleaning* dan *filtering* merupakan langkah untuk menangani *missing value*, kesalahan pencatatan, data duplikasi, dan *outlier* pada data.
- (b) *Grid equalization and integration* merupakan langkah untuk melakukan penyamaan resolusi spasial ke *grid* analisis (dalam bentuk 1 km x 1 km) dan melakukan integrasi seluruh layer spasial ke dalam kerangka *grid* yang sama. Hal ini dilakukan untuk mendapatkan informasi faktor penyebab suatu titik api muncul pada *grid* tertentu.
- (c) *Data transformation* merupakan langkah untuk melakukan transformasi data yang diperlukan (normalisasi, pengkodean variabel kategorik, dan lainnya).
- (d) *Feature engineering and selection* merupakan langkah untuk melakukan pemrosesan fitur seperti pemanfaatan fitur FWI yang dikonversi menjadi bulanan dan memilih fitur mana yang paling berkorelasi dengan label agar dapat menghilangkan fitur yang menjadi *noise*. Pada langkah ini juga dilakukan pemrosesan terhadap label data yaitu data *hotspot* sehingga

ga memiliki tiga kelas risiko (rendah, sendang, dan tinggi) berdasarkan *threshold* yang akan ditentukan.

- (e) *Class balancing* merupakan langkah untuk melakukan penyeimbangan data agar distribusi kelas tidak timpang ke salah satu kelas. Hal ini dilakukan agar model tidak terlalu bias terhadap kelas mayoritas.

#### 4. *Modelling*

Model yang dikembangkan adalah Random Forest dan XGBoost, sesuai dengan rancangan pada Bab IV. *Dataset* yang telah siap digunakan akan dibagi menjadi data latih dan validasi. Kedua model dilatih dan dibandingkan berdasarkan metrik kinerja yang ditetapkan (akurasi, *precision*, *recall*, *F1-score*, serta nilai AUC dan kurva ROC). Proses ini dapat dilakukan secara iteratif dengan penyesuaian *hyperparameter* hingga memperoleh performa yang memenuhi kriteria. Model dengan kinerja terbaik dipilih sebagai model utama yang akan digunakan untuk menghasilkan prediksi risiko karhutla. Model akan dikembangkan mengikuti data label sehingga akan menghasilkan tiga tingkat risiko atau kerentanan karhutla, yaitu rendah, sedang, dan tinggi.

#### 5. *Explainable AI* dan *Evaluation*

Metode XAI berupa SHAP dan PFI diterapkan ke model terbaik berdasarkan hasil seleksi sebelumnya. PFI digunakan untuk memperoleh gambaran penting fitur secara global, sedangkan SHAP digunakan untuk menjelaskan kontribusi fitur secara lokal pada masing-masing *grid*. Hasil prediksi dan penjelasan ini kemudian dievaluasi menggunakan metrik klasifikasi seperti akurasi, *precision*, *recall*, *F1-score*, serta nilai AUC dan kurva ROC.

#### 6. *Deployment*

Tahap ini merupakan tahap terakhir yang masuk dalam skala *prototype*. Model prediksi risiko dan hasil penjelasan model digabungkan serta disajikan dalam bentuk geospasial yang memberikan informasi risiko terjadinya karhutla bulanan serta ringkasan faktor utama yang memengaruhi munculnya risiko tersebut. Prototipe penyajian ini diharapkan mampu menampilkan kondisi saat ini, prediksi *hotspot*, serta penjelasan faktor sehingga dapat dijadikan ilustrasi bagi pemangku kepentingan dalam mengambil keputusan mitigasi.

### V.1.2 Lingkungan Pengembangan dan Daftar Alat

Lingkungan pengembangan dan alat yang digunakan dalam tugas akhir dirangkum pada Tabel V.1. Pemilihan alat berfokus pada ketersediaan, dukungan terhadap analisis spasial, keandalan sebagai bahan latihan prediktif, serta efisiensi biaya. Sebagian besar alat yang digunakan merupakan perangkat lunak yang bertujuan untuk

membantu langkah pencarian data, pemrosesan data, pengembangan model, hingga pengembangan visual untuk penyajian informasi. Namun, beberapa perangkat lunak tidak bisa digunakan secara terbuka, seperti perangkat untuk melakukan pelatihan model. Hal tersebut disebabkan karena infrastruktur yang dimiliki tidak cukup untuk melakukan pelatihan model. Informasi detail terkait alat yang akan digunakan tercantum pada Tabel V.1.

Tabel V.1 Lingkungan pengembangan dan daftar alat yang digunakan

<b>Kategori</b>	<b>Alat / Platform</b>	<b>Fungsi Utama</b>
Hardware lokal	Laptop dengan prosesor Intel Core i5 Gen-11, RAM 16 GB, GPU NVIDIA RTX 3050 4 GB	Komputasi untuk mengembangkan sistem prediksi karhutla.
<i>Cloud computing</i>	Google Cloud (mesin virtual dengan GPU)	Pelatihan model Random Forest dan XGBoost apabila beban komputasi melebihi kapasitas laptop lokal.
Software SIG	QGIS	Pengolahan data spasial, filter peta, <i>overlay</i> , pemotongan wilayah studi, dan pembentukan grid 1 km × 1 km.
Software analisis	Python dan Jupyter Notebook	Lingkungan utama untuk membuat skrip pemrosesan data, pelatihan model, evaluasi, dan penerapan <i>explainable AI</i> .
Pustaka Python	pandas, NumPy, scikit-learn, xgboost, shap, GeoPandas, dan pustaka terkait lainnya	Pustaka untuk membantu pengolahan data tabular dan numerik, implementasi model Random Forest dan XGBoost, perhitungan nilai SHAP.
Layanan data	NASA FIRMS, BMKG, KLHK, Climate Engine, Google Earth Engine (GEE)	Sumber data hotspot, variabel meteoroologis, biofisik, dan antropogenik yang digunakan sebagai masukan pemodelan risiko.
Prototipe visual	QGIS layout, pustaka Python (Folium)	Penyusunan peta risiko karhutla dan prototipe tampilan geospasial sederhana untuk mendemonstrasikan hasil model.

Tabel V.2 Rencana pemanfaatan layanan berbayar dan estimasi biaya

Aspek	Layanan	Perkiraan biaya per bulan (Rp)
Komputasi untuk pelatihan model	Google Colab Pro	170.000
Penyimpanan dan pengelolaan data	Google One (Google Drive, paket Google AI Pro)	158.000
	<b>Total</b>	<b>328.000</b>

Alat yang digunakan tentunya tidak gratis dan memiliki biaya sehingga pencatatan terkait biaya dicantumkan pada Tabel V.2. Biaya yang dicantumkan disusun dalam kurun waktu bulan. Pengembangan sistem dilakukan dalam waktu kurang lebih 4 bulan sehingga total biaya yang akan dikeluarkan adalah Rp 1.312.000.

### V.1.3 Linimasa Pengerjaan

Linimasa pelaksanaan tugas akhir dirangkum pada Tabel V.3. Setiap periode menggambarkan fokus kegiatan utama yang akan dikerjakan hingga tahap implementasi dan evaluasi prototipe selesai.

Tabel V.3 Linimasa penyelesaian tugas akhir

Waktu	Kegiatan
September 2024 – Oktober 2024	Pencarian topik tugas akhir.
Oktober 2024 – Desember 2024	<i>Business Understanding</i> dan penyusunan proposal tugas akhir.
Desember 2024	<i>Data collecting</i> dan <i>Exploratory Data Analysis</i> (tahap <i>data understanding</i> ).
Januari 2025 – Juni 2025	<i>Data collecting, data understanding, data preparation, modelling, evaluation, dan deployment</i> secara iteratif hingga prototipe sistem prediksi risiko karhutla selesai.

## V.2 Rencana Verifikasi dan Evaluasi

Subbab ini membahas tentang verifikasi dan evaluasi dari sistem model prediktif. Pengujian ini dilakukan agar sistem dapat bekerja sesuai dengan tujuan yang telah ditentukan sebelumnya. Langkah validasi dilakukan dengan melakukan verifikasi terhadap modul prediksi dan hasil dari XAI dalam memberikan penjelasan hasil prediksi.

## V.2.1 Verifikasi Model Prediksi dan XAI

Langkah verifikasi dilakukan untuk memastikan model prediksi dan XAI yang dikembangkan dapat berjalan dengan baik dan relevan dengan bahasan.

### V.2.1.1 Verifikasi Model Prediksi

Verifikasi dilakukan untuk memastikan bahwa seluruh komponen modul prediksi telah terimplementasi sesuai desain pada Subbab V.1.1 dan tidak menghasilkan kesalahan logika. Metode pengujian yang direncanakan meliputi:

1. Verifikasi struktur dan konsistensi data

Memeriksa kembali jumlah rekaman dan distribusi kelas sebelum dan sesudah tahap *data preparation*, memastikan setiap kombinasi grid–bulan memiliki label yang benar dan tidak terjadi kehilangan data yang tidak diinginkan.

2. Uji alur pemodelan

Menjalankan alur pelatihan pada sampel data untuk memastikan seluruh fungsi berjalan, tidak terjadi error, dan menghasilkan keluaran (prediksi dan metrik) dengan format yang sesuai.

3. Reproduksibilitas hasil

Mengatur *random seed* dan mendokumentasikan konfigurasi eksperimen untuk menilai konsistensi dari hasil metrik.

Modul prediksi dianggap lolos verifikasi apabila:

1. Tidak terdapat error pada saat alur pelatihan dan prediksi dijalankan.
2. Jumlah dan struktur data pada setiap tahap konsisten dengan rancangan.
3. Prediksi yang dihasilkan dapat dipetakan kembali.

### V.2.1.2 Verifikasi XAI

Verifikasi yang dilakukan untuk XAI adalah dengan menyusun beberapa pertanyaan analitis terhadap hasil prediksi. Pertanyaan tersebut di antaranya:

1. Mengapa *grid A* pada bulan Juli diklasifikasikan dengan risiko tinggi?
2. Faktor apa yang paling memengaruhi peningkatan risiko di Kalimantan Tengah pada beberapa bulan ini?

Untuk memberikan verifikasi, sebenarnya pertanyaan yang diberikan dapat melebihi jumlah kedua pertanyaan di atas. Jadi kedua pertanyaan tersebut merupakan contoh gambaran tentang bagaimana melakukan verifikasi terhadap XAI. Dengan memberikan pertanyaan tersebut, apabila XAI dapat memberikan penjelasan yang dapat menjawab pertanyaan tersebut, verifikasi terhadap XAI dapat dinyatakan berhasil. Hal ini karena XAI berhasil menjadi penjelas dari kondisi *blackbox* yang dialami

oleh model dan dapat membantu pemangku kepentingan dalam memahami kondisi dan mengambil keputusan.

### V.2.2 Validasi Model Prediksi dan XAI

Validasi bertujuan menilai apakah model prediksi yang dikembangkan memiliki performa yang memadai dan apakah hasil *explainable AI* selaras dengan pengetahuan domain. Metode pengujian dan kriteria keberhasilan dirangkum sebagai berikut.

#### V.2.2.1 Validasi Model Prediksi

Validasi kuantitatif model prediksi dilakukan menggunakan data uji yang tidak digunakan pada saat pelatihan. Dua metrik utama yang digunakan adalah:

1. Akurasi rata-rata dari model adalah 0,80 sesuai yang telah ditetapkan pada tujuan sebelumnya.
2. *F1-score* untuk kelas risiko tinggi adalah 0,75.

#### V.2.2.2 Validasi hasil XAI

Untuk komponen XAI, validasi kuantitatif dilakukan dengan mengukur *fidelity*, yaitu seberapa baik penjelasan mengikuti perilaku model asli. Kabir, Hossain, dan Andersson (September 2025) menempatkan *fidelity* sebagai salah satu dimensi utama dalam evaluasi XAI, berdampingan dengan aspek lain seperti komprehensibilitas dan stabilitas. Dalam tugas akhir ini, *fidelity* lokal didefinisikan sebagai tingkat kesesuaian antara prediksi model asli  $f(x)$  dan model penjelas yang lebih sederhana  $g(x)$  (misalnya model yang dibangun dari atribut penting hasil SHAP) pada sejumlah sampel uji di sekitar suatu wilayah. Untuk kasus klasifikasi, salah satu bentuk perhitungannya dapat dituliskan sebagai:

$$\text{Fidelity} = \frac{1}{N} \sum_{i=1}^N \mathbb{I}(f(x_i) = g(x_i)),$$

Target yang ditetapkan dalam tugas akhir ini adalah nilai *fidelity* rata-rata sekitar 0,80 pada beberapa sampel wilayah uji. Nilai ini dianggap cukup untuk menunjukkan bahwa pola kontribusi fitur yang ditampilkan XAI tidak menyimpang jauh dari keputusan yang diambil model, sekaligus realistik untuk prototipe awal.

### V.3 Analisis Risiko dan Mitigasi

Pelaksanaan tugas akhir yang melibatkan pengolahan data spasial berskala besar dan pengembangan model *machine learning* memiliki sejumlah risiko teknis maupun nonteknis. Pada subbab ini diidentifikasi beberapa risiko utama yang dapat memengaruhi kualitas hasil maupun kelancaran jadwal, beserta strategi mitigasi yang direncanakan. Analisis dilakukan secara kualitatif dengan menilai tingkat kemungkinan (*likelihood*) dan dampak (*impact*) setiap risiko pada skala rendah, sedang, dan tinggi.

Tabel V.4 Identifikasi risiko dan rencana mitigasi

Risiko	Kemungkinan	Dampak	Strategi mitigasi
Ketersediaan dan kelengkapan data faktor penjelas	Sedang	Tinggi	Mengutamakan pengumpulan variabel yang paling kritis (misalnya hotspot, curah hujan, lahan gambut, dan tutupan lahan) lebih dahulu. Mencari data berdasarkan artikel sehingga memiliki banyak alternatif sumber yang bisa digunakan atau memanfaatkan data historikal saja sebagai alat pelatihan model.
Kualitas data yang rendah (missing value, inkonsistensi, noise)	Rendah	Sedang	Melakukan <i>exploratory data analysis</i> untuk mengidentifikasi masalah kualitas data, menerapkan prosedur pembersihan, mencatat variabel yang kualitasnya sangat rendah dan, bila perlu, mengevaluasi ulang kelayakan variabel tersebut untuk dimasukkan ke model.
Keterbatasan sumber daya komputasi	Tinggi	Rendah	Memanfaatkan layanan <i>cloud computing</i> seperti Google Colab sebagai alat untuk melakukan pelatihan data dan pengembangan prototipe.

*Bersambung ke halaman berikutnya*

Tabel V.4 Identifikasi risiko dan rencana mitigasi (lanjutan)

Risiko	Kemungkinan	Dampak	Strategi mitigasi
Performa model tidak mencapai kriteria keberhasilan	Sedang	Tinggi	Melakukan langkah iterasi dan <i>hyperparameter</i> untuk mencari konfigurasi terbaik dari model. Teknik <i>feature engineering</i> juga dilakukan untuk menyesuaikan data agar hasil model lebih baik.
Biaya layanan komputasi awan melebihi estimasi	Sedang	Sedang	Mengatur penggunaan pada layanan (hanya digunakan saat benar-benar melakukan pengembangan model dan prototipe) sehingga dapat menekan anggaran.

Analisis pada Tabel V.4 menunjukkan bahwa risiko utama berasal dari aspek ketersediaan data, keterbatasan komputasi, dan kemungkinan performa model yang belum memenuhi target awal. Dengan adanya strategi mitigasi yang direncanakan sejak awal, diharapkan dampak risiko-risiko tersebut dapat dikendalikan sehingga tujuan utama tugas akhir tetap dapat dicapai dalam batas waktu dan sumber daya yang tersedia.

## DAFTAR PUSTAKA

- Abdollahi, Arnick, dan Biswajeet Pradhan. 2023. “Explainable artificial intelligence (XAI) for interpreting the contributing factors feed into the wildfire susceptibility prediction model”. *Science of the Total Environment* 879. ISSN: 18791026. <https://doi.org/10.1016/j.scitotenv.2023.163004>.
- Adadi, Amina, dan Mohammed Berrada. 2018. “Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI)”. *IEEE Access* 6:52138–52160. ISSN: 21693536. <https://doi.org/10.1109/ACCESS.2018.2870052>.
- Ali, Sajid, Tamer Abuhmed, Shaker El-Sappagh, Khan Muhammad, Jose M. Alonso-Moral, Roberto Confalonieri, Riccardo Guidotti, Javier Del Ser, Natalia Díaz-Rodríguez, dan Francisco Herrera. 2023. “Explainable Artificial Intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence”. *Information Fusion* 99. ISSN: 15662535. <https://doi.org/10.1016/j.inffus.2023.101805>.
- Bifarin, Olatomiwa O. 2023. “Interpretable machine learning with treebased shapley additive explanations: Application to metabolomics datasets for binary classification”. *PLoS ONE* 18 (5 May). ISSN: 19326203. <https://doi.org/10.1371/journal.pone.0284315>.
- Chen, Xinbao, Yaohui Zhang, Shan Wang, Zecheng Zhao, Chang Liu, dan Junjun Wen. 2024. “Comparative study of machine learning methods for mapping forest fire areas using Sentinel-1B and 2A imagery”. *Frontiers in Remote Sensing* 5. ISSN: 26736187. <https://doi.org/10.3389/frsen.2024.1446641>.
- Eaturu, Aditya, dan Krishna Prasad Vadrevu. 2025. “Evaluation of machine learning and deep learning algorithms for fire prediction in Southeast Asia”. *Scientific Reports* 15 (1). ISSN: 20452322. <https://doi.org/10.1038/s41598-025-00628-9>.

Fitriani, Rizqi Nur, Bambang Joko Priatmadi, Akhmad Rizalli Saidy, Muhammad Syahdan, dan Program Studi Pengelolaan Sumberdaya Alam dan Lingkungan Program Pascasarjana. 2023. “PENENTUAN TINGKAT RESIKO KEBAKARAN HUTAN DAN LAHAN MENGGUNAKAN METODE INDEKS CUACA KEBAKARAN (FIRE WEATHER INDEX) DAN JUMLAH TITIK PANAS (HOTSPOT) DI KABUPATEN BANJAR PROVINSI KALIMANTAN SELATAN Forest and Land Fire Risk Level Determination Using The Fire Weather Index Method and The Number Of Hotspots in Banjar District, South Kalimantan Province”. 19 (2). ISSN: 2302-3708. [https://cds.climate.copernicus.eu/..](https://cds.climate.copernicus.eu/)

Fitriany, Anni Arumsari, Piotr J. Flatau, Khoirunurrofik Khoirunurrofik, dan Nelly Florida Riama. 2021. “Assessment on the use of meteorological and social media information for forest fire detection and prediction in riau, indonesia”. *Sustainability (Switzerland)* 13 (20). ISSN: 20711050. <https://doi.org/10.3390/su132011188>.

Fumagalli, Fabian, Maximilian Muschalik, Eyke Hüllermeier, dan Barbara Hammer. 2023. “Incremental permutation feature importance (iPFI): towards online explanations on data streams”. *Machine Learning* 112 (12): 4863–4903. ISSN: 15730565. <https://doi.org/10.1007/s10994-023-06385-y>.

GAO, United States General Accounting Office. 2003. *Challenges to Effective Data Sharing*. Technical report. <https://www.gao.gov/cgi-bin/getrpt?GAO-03-874T>.

Gupta, Shivam, Sachin Modgil, Samadrita Bhattacharyya, dan Indranil Bose. 2022. “Artificial intelligence for decision support systems in the field of operations research: review and future scope of research”. *Annals of Operations Research* 308 (1-2): 215–274. ISSN: 15729338. <https://doi.org/10.1007/s10479-020-03856-6>.

Hein, Lars, Joseph V. Spadaro, Bart Ostro, Melanie Hammer, Elham Sumarga, Resti Salmayenti, Rizaldi Boer, Hesti Tata, Dwi Atmoko, dan Juan Pablo Castañeda. 2022. “The health impacts of Indonesian peatland fires”. *Environmental Health: A Global Access Science Source* 21 (1). ISSN: 1476069X. <https://doi.org/10.1186/s12940-022-00872-w>.

- Indradjad, Andy, Muhammad Dimyati, Yenni Vetrita, Erna Sri Adiningsih, dan Rokhmatuloh Rokhmatuloh. January 2024. “Enhancing Fire Monitoring Method over Peatlands and Non-Peatlands in Indonesia Using Visible Infrared Imaging Radiometer Suite Data”. *Fire* 7 (1). ISSN: 25716255. <https://doi.org/10.3390/fire7010009>.
- Jain, P., S. C. P. Coogan, S. G. Subramanian, M. Crowley, S. Taylor, dan M. D. Flannigan. 2021. “A review of machine learning applications in wildfire science and management”. *Environmental Reviews* 29 (4): 438–461. <https://doi.org/10.1139/er-2020-0132>.
- Kabir, Sami, Mohammad Shahadat Hossain, dan Karl Andersson. September 2025. *A Review of Explainable Artificial Intelligence from the Perspectives of Challenges and Opportunities*. <https://doi.org/10.3390/a18090556>.
- Karurung, Wulan Salle, Kangjae Lee, dan Wonhee Lee. 2025. “Assessment of forest fire vulnerability prediction in Indonesia: Seasonal variability analysis using machine learning techniques”. *International Journal of Applied Earth Observation and Geoinformation* 138. ISSN: 1872826X. <https://doi.org/10.1016/j.jag.2025.104435>.
- Kementerian Lingkungan Hidup dan Kehutanan (KLHK). 2025. “SIPONGI: Sistem Informasi Pengendalian Kebakaran Hutan dan Lahan”. Diakses 28 Oktober 2025. Diakses pada October 28, 2025. <https://sipongi.menlhk.go.id/>.
- Lavazza, Luigi, Sandro Morasca, dan Gabriele Rotoloni. 2025. “Software Defect Prediction evaluation: New metrics based on the ROC curve”. *Information and Software Technology* 187. ISSN: 09505849. <https://doi.org/10.1016/j.infsof.2025.107865>.
- Lundberg, Scott M., dan Su-In Lee. 2017. “A Unified Approach to Interpreting Model Predictions”. Dalam *Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS 2017)*, 4765–4774. Curran Associates, Inc. ISBN: 9781510860964.
- Miller, Catriona, Theo Portlock, Denis M. Nyaga, dan Justin M. O’Sullivan. 2024. *A review of model evaluation metrics for machine learning in genetics and genomics*. <https://doi.org/10.3389/fbinf.2024.1457619>.

- Nurhayati, Ati Dwi, Bambang Hero Saharjo, Leti Sundawati, Syartinilia Syartinilia, dan Mark A. Cochrane. 2021. “Forest and peatland fire dynamics in South Sumatra Province”. *Forest and Society* 5 (2): 591–603. ISSN: 25494333. <https://doi.org/10.24259/fs.v5i2.14435>.
- O’Shea, Richard, Peter Deeney, Evangelos Triantaphyllou, Luis Diaz-Balteiro, dan S. Armagan Tarim. 2026. “Weight stability intervals for multi-criteria decision analysis using the weighted sum model”. *Expert Systems with Applications* 296. ISSN: 09574174. <https://doi.org/10.1016/j.eswa.2025.128460>.
- Prasetyo, Lilik Budi, Yudi Setiawan, Aryo Adhi Condro, Kustiyo Kustiyo, Erianto Indra Putra, Nur Hayati, Arif Kurnia Wijayanto, Almi Ramadhi, dan Daniel Murdiyarso. 2022. “Assessing Sumatran Peat Vulnerability to Fire under Various Condition of ENSO Phases Using Machine Learning Approaches”. *Forests* 13 (6). ISSN: 19994907. <https://doi.org/10.3390/f13060828>.
- Prayoga, M. Bayu Rizky, dan Raldi Hendro Koestoer. December 2021. “Improving Forest Fire Mitigation in Indonesia: A Lesson from Canada”. *Jurnal Wilayah dan Lingkungan* 9 (3): 293–305. ISSN: 2338-1604. <https://doi.org/10.14710/jwl.9.3.293-305>.
- Qamariyanti, Yulia, Rachmadi Usman, dan Diana Rahmawati. 2023. “Pencegahan dan Penanggulangan Kebakaran Lahan Gambut dan Hutan”. *Jurnal Ilmu Lingkungan* 21 (1): 132–142. ISSN: 1829-8907. <https://doi.org/10.14710/jil.21.1.132-142>.
- Rossita, Annuri, Rizaldi Boer, Lars Hein, Dodik Ridho Nurrochmat, dan Ahmad Riqqi. 2023. “Peatland fire regime across Riau peat hydrological unit, Indonesia”. *Forest and Society* 7 (1): 76–94. ISSN: 25494333. <https://doi.org/10.24259/FS.V7I1.21996>.
- Sarker, Iqbal H. 2021. *Machine Learning: Algorithms, Real-World Applications and Research Directions*. <https://doi.org/10.1007/s42979-021-00592-x>.
- . 2022. *AI-Based Modeling: Techniques, Applications and Research Issues Towards Automation, Intelligent and Smart Systems*. <https://doi.org/10.1007/s42979-022-01043-x>.
- Sokolova, Marina, dan Guy Lapalme. 2009. “A systematic analysis of performance measures for classification tasks”. *Information Processing and Management* 45 (4): 427–437. ISSN: 03064573. <https://doi.org/10.1016/j.ipm.2009.03.002>.

- Syahid, L. N., dan R. Hidayat. 2022. “Analysis of the Relationship between El Niño-Southern Oscillation and Forest/Land Fires in Sumatra and Kalimantan”. *Jurnal Pengelolaan Sumberdaya Alam dan Lingkungan* 12 (1): 127–137. <https://doi.org/10.29244/jpsl.12.1.127-137>.
- Teixeira, Hugo, Alexandre Magalhães, André Ramalho, Maria de Fátima Pina, dan Hernâni Gonçalves. 2021. *Indoor Environments and Geographical Information Systems: A Systematic Literature Review*. <https://doi.org/10.1177/21582440211050379>.
- Thoha, Achmad Siddik, dan Hesty Triani. 2021. “A spatial model of forest and land fire vulnerability level in the Dairi district, North Sumatra, Indonesia”. *Biodiversitas* 22 (8): 3319–3326. ISSN: 20854722. <https://doi.org/10.13057/biodiv/d220827>.
- Vetrita, Yenni, Israr Albar, Imam Santoso, Indah Prasasti, Tatik Kartika, Ahmad Basyiruddin Usman, Anna Tosiani, dkk. 2025. “Monthly mapping of Indonesia’s burned areas: implementation, history, techniques, and future directions”. *International Journal of Remote Sensing* 46 (2): 636–660. ISSN: 13665901. <https://doi.org/10.1080/01431161.2024.2421942>.
- Wiens, Matthew, Alissa Verone-Boyle, Nick Henscheid, Jagdeep T. Podichetty, dan Jackson Burton. 2025. “A Tutorial and Use Case Example of the eXtreme Gradient Boosting (XGBoost) Artificial Intelligence Algorithm for Drug Development Applications”. *Clinical and Translational Science* 18 (3). ISSN: 17528062. <https://doi.org/10.1111/cts.70172>.