

Model Pembelajaran Mesin Yang Dapat Dijelaskan Untuk Menjelaskan Parameter Iklim Yang Berpengaruh Berdasarkan Prediksi Curah Hujan

Hamdan Gani¹, Annisa Dwi Damayanti², Muhammad Muflih Mubarak³, Helmy Gani⁴, Muslimin⁵
^{1,3,5}Politeknik ATI Makassar, ²Universitas Hasanuddin, ³Sekolah Tinggi Ilmu Kesehatan Makassar
¹hamdangani@atim.ac.id, ²annisa.dd@unhas.ac.id, ³21OSP611@atim.ac.id, dan ⁴ganihelmy@gmail.com

ABSTRAK

Prediksi curah hujan yang akurat dan dapat ditafsirkan sangat penting untuk pengambilan keputusan dalam bidang pertanian, manajemen bencana, dan ketahanan iklim. Studi ini mengusulkan kerangka pembelajaran mesin (ML) yang komprehensif dan dapat ditafsirkan, yang tidak hanya meningkatkan akurasi prediksi tetapi juga menjelaskan faktor-faktor utama yang mempengaruhi kejadian curah hujan. Memanfaatkan data iklim selama sembilan tahun dari Climate Data Services - Bureau of Meteorology of Australia, enam model ML—XGBoost, CatBoost, LightGBM, Random Forest, Decision Tree, dan Logistic Regression—dikembangkan dan dievaluasi. Alur kerja praproses data mencakup pembersihan, rekayasa fitur, dan penanganan ketidakseimbangan kelas melalui Synthetic Minority Oversampling (SMOTE). Untuk memastikan transparansi, teknik Kecerdasan Buatan yang Dapat Dijelaskan (XAI) seperti SHAP dan LIME diintegrasikan, menyediakan interpretasi global dan lokal dari prediksi model. Hasil penelitian menunjukkan bahwa XGBoost mencapai akurasi tertinggi sebesar 95,35%, mengungguli model lainnya. Analisis SHAP mengidentifikasi Humidity3pm dan Pressure3pm sebagai prediktor paling berpengaruh, sementara LIME memberikan wawasan terperinci tentang variasi prediksi lokal. Temuan ini menyoroti kehandalan framework yang diusulkan dalam kinerja prediktif dan interpretabilitas yang ditingkatkan, memungkinkan para pakar memahami dinamika iklim yang mendasarinya. Studi ini memberikan pendekatan ML yang baru untuk prediksi curah hujan, menjembatani kesenjangan antara kompleksitas model dan penerapan praktis.

Kata kunci: Prediksi Curah Hujan, Pembelajaran Mesin, Kecerdasan Buatan yang Dapat Dijelaskan (XAI), Analisis Data Iklim

1. PENDAHULUAN

Bergantung pada kondisi iklim, hujan turun di lingkungan dalam berbagai bentuk dan fitur. Secara umum, ada tiga jenis curah hujan yang berbeda berdasarkan tempat asalnya [1], [2]. Curah hujan relief atau orografis muncul ketika udara dipaksa naik ke wilayah dataran tinggi setelah dipaksa melewati laut. Akibatnya, udara mendingin, mengembunkan uap air udara untuk menghasilkan hujan. Sementara itu, curah hujan frontal atau siklon terjadi ketika udara hangat dipaksa naik di atas udara dingin. Uap air mengembun saat udara hangat mendingin, menghasilkan awan dan hujan. Jenis curah hujan terakhir adalah curah hujan konveksi. Jenis curah hujan tersebut paling banyak terjadi di daerah tropis yang panas dan udara panas naik menjadi dingin dan mengembun untuk menghasilkan hujan saat mendingin dan badai petir dapat terbentuk saat udara naik dengan cepat karena panas. Gambar 1 mengilustrasikan ketiga jenis curah hujan tersebut.

Dalam banyak aspek kehidupan kita sehari-hari, curah hujan tetap menjadi salah satu variabel iklim yang paling penting. Curah hujan sangat penting bagi semua spesies hidup di dunia. Curah hujan memiliki banyak dampak, baik positif maupun negatif. Dampak positifnya meliputi air yang dibutuhkan untuk kehidupan dan pertumbuhan tanaman [3], air untuk pembangkit listrik tenaga air [4], irigasi tanaman [5], dan peningkatan produksi ekosistem bruto (GEP) [6]. Di sisi negatifnya, dampak curah hujan meliputi kerusakan tanaman [7], erosi tanah [8], kerusakan infrastruktur [9], dan gangguan jaringan transportasi akibat hujan lebat [10]. Selain itu, dalam



dampak global, curah hujan ekstrem dapat memperlambat ekonomi global [11]. Oleh karena itu, memahami fenomena curah hujan sangat penting untuk membuat tindakan yang tepat guna mengelola sumber daya air, produktivitas tanaman, dan berbagai konteks kehidupan secara efisien, termasuk transportasi, ekonomi, industri, perjalanan, pertanian, olahraga, dan bidang lainnya [12], [13].

Saat ini, prediksi curah hujan menjadi hal yang krusial karena curah hujan yang ekstrem dan tidak teratur dapat menimbulkan beberapa dampak, termasuk kerusakan bangunan [14] dan kerusakan pertanian dan tanaman [15], sehingga diperlukan model peramalan yang lebih baik untuk sistem peringatan dini yang dapat mengurangi risiko curah hujan terhadap kehidupan kita. Secara umum, peramalan ini bermanfaat bagi petani untuk penggunaan sumber daya air yang efektif dan meningkatkan produktivitas pertanian [16]. Selain itu, penting bagi pemerintah untuk memastikan akses penduduk suatu negara terhadap makanan dan air bersih serta mengelola masalah kualitas air dan sistem peringatan banjir [17]. Oleh karena itu, banyak penelitian telah menyarankan sistem peramalan curah hujan yang akurat sebagai strategi mitigasi agar siap menghadapi situasi terburuk apa pun.

Peramalan curah hujan menggunakan parameter cuaca historis seperti suhu, tekanan udara, kelembaban, sinar matahari, tekanan udara, dan lainnya sebagai kumpulan data. Selama bertahun-tahun, penelitian tentang prakiraan curah hujan telah mengeksplorasi hubungan antara curah hujan dan parameter cuaca seperti koordinat geografis (yaitu, lintang dan bujur), dan karakteristik meteorologi lainnya (yaitu, tekanan, suhu, kecepatan angin, dan kelembaban). Para peneliti memanfaatkan metode konvensional seperti teknik statistik untuk memprediksi curah hujan [18]. Namun, prakiraan curah hujan masih sulit karena kompleksitasnya dan fenomena non-linearitas [19]. Akibatnya, upaya telah dilakukan untuk mengurangi non-linearitas ini dengan memanfaatkan beberapa teknik [20] seperti analisis wavelet, analisis spektrum singular, dan dekomposisi mode empiris tetapi metode tersebut memiliki keterbatasan pada daya pemrosesan dan memakan waktu [19]. Oleh karena itu, para peneliti saat ini telah menerapkan pendekatan ML secara signifikan untuk prakiraan curah hujan.

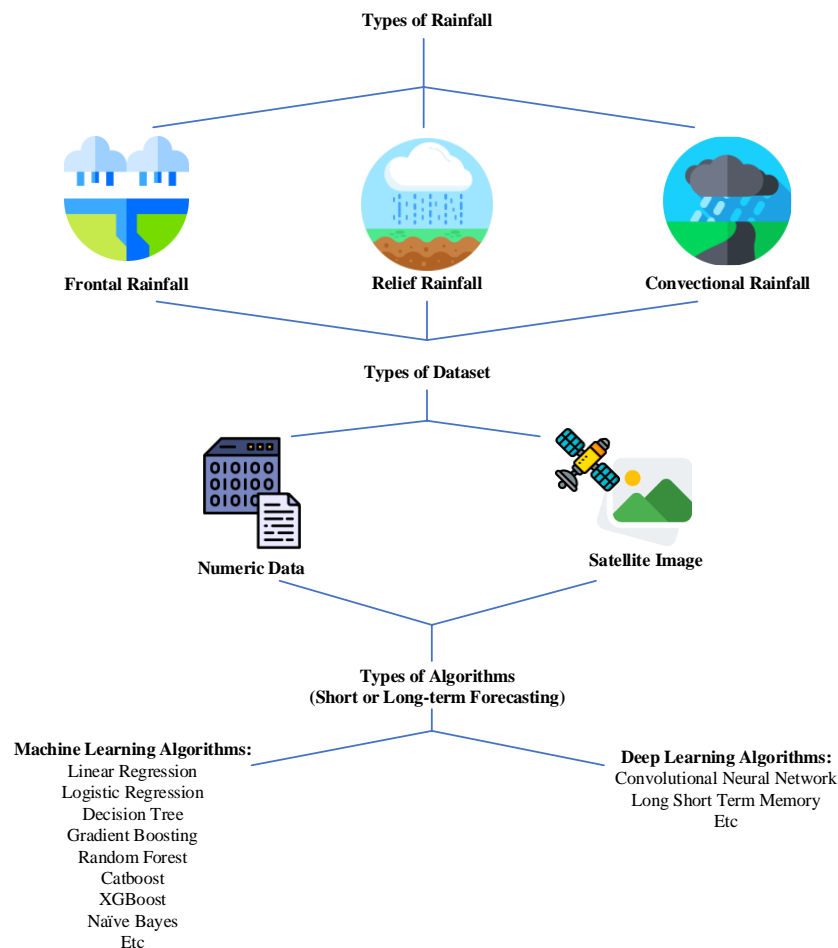
Pendekatan ML telah menjadi solusi untuk tantangan prediksi curah hujan karena dapat mengatasi non-linearitas kumpulan data curah hujan dan juga dapat muncul dengan teknologi lain seperti Internet of Things [21]. Misalnya, menggunakan kumpulan data numerik historis, Esteves et al. [22] menggunakan jaringan saraf tiruan (ANN) untuk prediksi deret waktu dari sepuluh wilayah pertanian di Brasil. Balamurugan dan Manojkumar [23] mempelajari prakiraan hujan jangka pendek menggunakan regresi logistik, pohon keputusan, hutan acak, dan pendekatan statistik. Yin et al. [24] mengusulkan model jaringan saraf konvolusional (CNN) yang disebut CNN curah hujan berbasis gambar (irCNN) menggunakan kumpulan data citra satelit. Model ini dikembangkan menggunakan citra curah hujan yang dikumpulkan dari sensor padat yang ada (yaitu, telepon pintar atau kamera transportasi). Boonyuen et al. [25] menggunakan jaringan saraf convolutional (CNN) dengan model inception-v3 untuk memperkirakan curah hujan harian. Input dari model ini adalah citra satelit dari wilayah-wilayah di Asia.

Meskipun penelitian ekstensif mengenai teknologi ML untuk perkiraan curah hujan telah dilakukan, tidak ada yang menjelaskan faktor-faktor kritis yang memengaruhi curah hujan. Sepengetahuan kami, masalah ini belum diteliti di tempat lain. Oleh karena itu, memahami variabel-variabel yang berpengaruh menyebabkan curah hujan sangat penting, karena (a) dapat berkontribusi untuk memprediksi curah hujan dengan akurasi prediksi yang lebih tinggi dan memberikan prediksi algoritma ML yang lebih baik [26], (b) memberikan informasi peringatan dini untuk mengevaluasi parameter meteorologi yang dapat memengaruhi curah hujan [27], (c) memungkinkan pemerintah untuk memahami fenomena curah hujan dan menggunakan pengetahuan tersebut untuk mencegah dampak curah hujan seperti banjir dan dampak sosial ekonomi lainnya [19].

Interpretable ML adalah metode yang menjelaskan cara kerja internal model ML sehingga orang dapat dengan mudah memahaminya [28]. Metode ini sekarang digunakan dalam banyak penelitian. Dalam studi eksperimental, Arteaga et al [29] memanfaatkan teknik Interpretable ML untuk menjelaskan terjadinya tingkat keparahan cedera dalam kecelakaan lalu lintas menggunakan penambahan teks dengan pendekatan pembelajaran mesin yang dapat ditafsirkan. Teknik Interpretable ML juga telah digunakan di bidang lain, seperti bidang medis dan permainan digital. Misalnya, teknik ini telah menggambarkan individu yang berisiko terkena hipertensi berdasarkan data kebugaran kardiorespirasi [30] dan menjelaskan konten berbahaya dalam permainan video [31]. Meskipun penelitian sebelumnya telah menunjukkan janji teknik Interpretable ML mengenai analisis interpretabilitas, tidak ada penelitian yang menggunakannya untuk memeriksa fenomena curah hujan. Studi penelitian terkini hanya berupaya memperoleh metrik dengan akurasi prediksi tertinggi [20] tetapi tidak memiliki analisis menyeluruh tentang faktor-faktor kritis yang memengaruhi curah hujan. Studi tinjauan sebelumnya juga telah mengemukakan bahwa tidak ada studi curah hujan yang berfokus pada keterjelasan [26]. Studi kami membahas kelalaian penelitian ini yang diidentifikasi dalam makalah eksperimental dan tinjauan sebelumnya.

Studi ini menyajikan sistem prakiraan yang dapat ditafsirkan untuk memahami fenomena curah hujan. Pertama, studi ini menyajikan pengembangan sistem prakiraan yang dapat ditafsirkan. Studi ini menggunakan enam model ML terbimbing untuk prediksi curah hujan, yaitu, eXtreme Gradient Boosting (XGBoost), CatBoost, Light Gradient Boosting Machine (LightGBM), Random Forest (RF), Decision Tree (DT), dan Logistic Regression (LR). Kemudian, studi ini menerapkan kerangka kerja ML yang dapat ditafsirkan untuk menjelaskan faktor-faktor kritis yang memengaruhi curah hujan berdasarkan prediksi curah hujan. Studi ini mengusulkan

beberapa penjelasan untuk meningkatkan keterjelasan model menggunakan analisis interpretabilitas global dan lokal. ML yang dapat diinterpretasikan juga mengungkap titik-titik perubahan variabel iklim di mana transisi dari tingkat curah hujan rendah ke tinggi terjadi.



Gambar 1. Berbagai jenis curah hujan dan algoritma ML

2. TINJAUAN PUSTAKA

Selama dua dekade terakhir, banyak peneliti telah berfokus pada peningkatan akurasi algoritma ML dalam peramalan curah hujan. Di sini, beberapa studi terkait dibahas. Untuk kumpulan data numerik, dalam Abebe dan Endalie [32], peneliti menerapkan Jaringan Syaraf Tiruan (JST) dan model Sistem Inferensi Neuro-Fuzzy Adaptif (ANFIS) dalam meramalkan curah hujan bulanan jangka panjang (dari 2011 hingga 2021) di 92 stasiun meteorologi Ethiopia menggunakan komponen geografis dan periodisitas (bujur, lintang, dan ketinggian). Menurut temuan mereka, model ANFIS berkinerja lebih baik daripada model JST di semua kriteria evaluasi. Hudnurkar dan Rayavarapu [33] menggunakan JST untuk memprediksi rata-rata bulanan variabel iklim, yaitu suhu udara (AT), suhu permukaan laut (SST), dan tekanan permukaan laut (SLP) di seluruh dunia, untuk meramalkan curah hujan monsun musim panas bulanan dan musiman di negara bagian Maharashtra, India. Temuan mereka menunjukkan bahwa model variabel iklim individual untuk AT mengungguli semua faktor evaluasi lainnya kecuali untuk kapasitas curah hujan maksimum. Lebih lanjut, Pirone et al. [34] menggunakan model peta pengorganisasian mandiri (SOM) untuk mengkategorikan sepuluh pola sirkulasi skala besar untuk setiap musim hujan. Kemudian, mereka membangun model jaringan saraf konvolusional (CNN) untuk mengekstrak informasi spasial (seperti ketinggian, kelembaban spesifik, dan tekanan permukaan laut rata-rata) dan model jaringan memori jangka pendek panjang (LSTM) untuk mengekstrak deret waktu. Cekungan Sungai Huaihe (antara 2007 dan 2021) digunakan untuk membandingkan pendekatan mereka (SOM-CNN-LSTM) dengan metode tolok ukur lainnya (yaitu, CNN, LSTM, dan CNN-LSTM). Hasilnya menunjukkan bahwa metode

yang mereka usulkan (SOM-CNN-LSTM) mengalahkan metode tolok ukur yang ada untuk semua waktu tunggu dan setiap musim.

Untuk citra satelit, Ouallouche et al. [35] menggunakan ML, seperti Jaringan Syaraf Tiruan (JST), Support Vector Machine (SVM), dan Random Forest (RF), untuk klasifikasi dan estimasi intensitas curah hujan. Hasilnya menunjukkan bahwa teknik JST mengungguli yang lain. Antal dkk. [36] meneliti kinerja Prediksi Regresi Kriging Bayesian Empiris (EBKRP) dan Regresi Tertimbang Geografis (GWR) untuk memprediksi distribusi spasial curah hujan tahunan dan musiman untuk citra satelit di Rumania. Temuan mereka mengonfirmasi bahwa sepuluh variabel relevan untuk memprediksi curah hujan tahunan dan sembilan untuk musiman. Prediksi curah hujan keseluruhan lebih dipengaruhi oleh topografi (ketinggian, lereng, kekasaran permukaan) dan jarak ke badan laut (Laut Hitam dan Laut Adriatik). Studi mereka mengungkapkan GWR sebagai metode terbaik untuk memprediksi curah hujan tahunan dan musiman di Rumania. Yang dkk. [37] meneliti teknik pembelajaran mendalam untuk menghitung reflektivitas radar komposit menggunakan pembacaan dari satelit cuaca geostasioner terbaru Tiongkok FY-4A dan informasi topografi. Temuan mereka menunjukkan bahwa model pembelajaran mendalam dapat menciptakan kembali posisi umum, bentuk, dan kekuatan gema radar. Selain itu, analisis observasi radar rekonstruksi menunjukkan bahwa model Attention U-Net yang dimodifikasi berkinerja lebih baik daripada model U-Net konvensional berdasarkan mekanisme perhatian.

Dalam paper sebelumnya, sebagian besar peneliti menggunakan ML yang diawasi untuk menemukan pola tersembunyi dalam data historis dan memprediksi curah hujan. Para peneliti sebagian besar menggunakan lebih dari satu metode dalam kerangka kerja yang diusulkan untuk prediksi. Namun, tidak ada studi curah hujan yang berfokus pada keterjelasan model ML mereka. Oleh karena itu, model ML terutama dianggap sebagai kotak hitam yang membuat praktisi enggan menggunakan model ML karena kurangnya keterjelasan dan kepercayaan. Untuk mengatasi hal ini, penelitian ini mengusulkan sistem peramalan yang dapat ditafsirkan menggunakan analisis global dan lokal yang dapat ditafsirkan untuk menjelaskan faktor-faktor penting yang memengaruhi curah hujan.

3. METODE PENELITIAN

Dataset diekstrak dari Climate Data Services - Bureau of Meteorology of Australia dari tahun 2008 hingga 2017 [15]. Dataset terdiri dari 145.461 instance dan 23 fitur, yang 22 di antaranya independen, dan satu dependen atau kelas target. Kelas target menunjukkan apakah akan ada hujan besok atau tidak. Jika fitur yang diprediksi memiliki nilai 1, maka akan turun hujan besok; jika nilainya 0, maka tidak akan turun hujan besok. Karakteristik akhir dataset dirangkum dalam Tabel 1.

Tahap pra-pemrosesan data terdiri dari beberapa aktivitas: pembersihan, penanganan nilai yang hilang, konversi variabel kategorikal dan kontinu, penghapusan outlier, dan penyeimbangan dataset menggunakan teknik SMOOTE untuk mendapatkan akurasi maksimum. Proses pembersihan data bertujuan untuk menghapus nilai yang hilang dalam dataset dengan menggunakan teknik imputasi rata-rata. Teknik konversi membawa nilai atribut ke dalam variabel tertentu. Pada aktivitas terakhir dari tahap pra-pemrosesan, kumpulan data dibagi menjadi dua subset: data latih dan data uji, dengan rasio aturan pemisahan kelas 80:20.

Eksperimen dilakukan pada platform Windows 10 dengan unit pemrosesan grafis (GPU) 16 GB, penyimpanan SSD 256 GB, prosesor Intel Core i7 1,80 GHz, dan RAM 8 GB. Lingkungan Python (versi 3.7.6), Scikit-learn, dan pustaka Keras digunakan untuk mengembangkan algoritme. Terakhir, kerangka kerja SHAP dan LIME digunakan untuk analisis interpretabilitas global dan lokal prediksi curah hujan.

Tabel 1. Atribut kumpulan data

Attribute Name	Attribute Type	Meaning	Measurement
Date	Categorical	Day of the month	Date
Location	Categorical	Location of city	-
Minimum Temperature	Continuous	The minimum temperature is 24 hours to 9 am. Sometimes only known to the nearest whole degree.	Degrees Celsius
Maximum Temperature	Continuous	Maximum temperature in the 24 hours to 9 am. Sometimes only known to the nearest whole degree.	Degrees Celsius
Rainfall	Continuous	Precipitation (rainfall) in the 24 hours to 9 am. Sometimes only known to the nearest whole millimeters.	Millimeters (mm)
Evaporation	Continuous	"Class A" pan evaporation in the 24 hours to 9 am	Millimeters (mm)
Sunshine	Continuous	Bright Sunshine in the 24 hours to midnight	Hours
Wind Gust Direction	Categorical	The direction of the strongest gust in the 24 hours to midnight	16 compass points
Wind Gust Speed	Continuous	Speed of strongest wind gust in the 24 hours to midnight	Kilometers per Hours
Wind Direction at 9 am	Categorical	Wind direction averaged over 10 minutes prior to 9 am	Compass points
Wind Direction at 3 pm	Categorical	Wind direction averaged over 10 minutes prior to 3 pm	Compass points
Wind Speed at 9 am	Continuous	Wind speed averaged over 10 minutes prior to 9 am	Kilometers per Hours
Wind Speed at 3 pm	Continuous	Wind speed averaged over 10 minutes prior to 3 pm	Kilometers per Hours
Humidity at 9 am	Continuous	Relative humidity at 9 am	Percent
Humidity at 3 pm	Continuous	Relative humidity at 3 pm	Percent
Pressure at 9 am	Continuous	Atmospheric pressure reduced to mean sea level at 9 am	Hectopascals (hPa)
Pressure at 3 pm	Continuous	Atmospheric pressure reduced to mean sea level at 3 pm	Hectopascals (hPa)
Cloud at 9 am	Continuous	Fraction of sky obscured by cloud at 9 am	Eighths



Cloud at 3 pm	Continuous	Fraction of sky obscured by clouds at 3 pm	Eighths
The temperature at 9 am	Continuous	The temperature at 9 am	Degrees Celsius
The temperature at 3 pm	Continuous	The temperature at 3 pm	Degrees Celsius
Rain Today	Categorical	This occurs of rain on this day	Yes/No

3.1 Pengembangan Sistem Peramalan yang Dapat Diinterpretasikan

Proses pengembangan sistem peramalan yang dapat ditafsirkan ditampilkan pada Gambar 1. Kerangka kerja yang diusulkan terutama terdiri dari dua lapisan: Lapisan pertama berkaitan dengan mengekstraksi atribut cuaca dari stasiun cuaca. Tahap pra-pemrosesan data terdiri dari beberapa aktivitas seperti yang dijelaskan pada bagian bahan dan metode. Setelah melakukan aktivitas pra-pemrosesan, kumpulan data siap untuk tahap klasifikasi, di mana kumpulan data pelatihan dan pengujian keduanya diberikan sebagai input ke enam ML yang diawasi (yaitu, XGBoost, CatBoost, LightGBM, RF, DT, dan LR). Semua algoritme ini dioptimalkan secara iteratif selama pelatihan dan pengujian untuk mencapai akurasi yang lebih tinggi. Setelah analisis kinerja, model yang paling terlatih dipilih untuk lapisan berikutnya. Pada lapisan kedua, semua model kereta digunakan sebagai input ke ML yang dapat ditafsirkan untuk menjelaskan fenomena curah hujan.

Proses pengembangan sistem peramalan yang dapat ditafsirkan ditampilkan pada Gambar 1. Langkah pertama adalah pemrosesan data, di mana semua fitur (variabel independen) dari kumpulan data cuaca dan hujan besok (variabel dependen) digabungkan. Set data lengkap dibagi menjadi set data pelatihan dan pengujian, yang digunakan untuk memprediksi kinerja model ML dalam memperkirakan prediksi curah hujan. Langkah selanjutnya adalah pembuatan model. Beberapa model ML dikembangkan dalam langkah ini, termasuk XGBoost, CatBoost, LightGBM, RF, DT, dan LR. Akurasi, presisi, recall, dan skor F1 adalah empat metrik yang digunakan. Persentase kasus yang dikategorikan dengan benar oleh model dikenal sebagai akurasi. Recall atau sensitivitas mengacu pada rasio prediksi positif aktual, sedangkan presisi mengacu pada jumlah kasus yang sesuai dengan kelas yang ditentukan dan contoh yang dikategorikan ke dalam kelas tersebut. Skor F1 atau ukuran F1 menggambarkan akurasi klasifikasi mengenai nilai presisi dan recall rata-rata. Akurasi klasifikasi yang lebih baik ditunjukkan oleh nilai skor F1 yang mendekati 1. Metrik evaluasi dihitung dalam (1), (2), (3), dan (4):

$$Accuracy = \frac{Correctly\ classified\ data}{Total\ data} \quad (1)$$

$$Recall = \frac{TP}{TP+FN} \quad (2)$$

$$Precision = \frac{TP}{TP+FP} \quad (3)$$

$$F1 = 2 * \frac{Precision*Recall}{Precision+Recall} \quad (4)$$

TP merupakan singkatan dari "true positive", FN untuk "false negative", dan FP untuk "false positive".

4. HASIL DAN PEMBAHASAN

Pada bagian ini, penelitian ini membandingkan kinerja enam model ML. Kemudian, penelitian ini menyajikan analisis interpretabilitas untuk menjelaskan faktor-faktor penting yang memengaruhi curah hujan.

3.1 Analisis Kinerja

Bagian ini menyajikan kinerja komparatif model ML. Konfigurasi model ML digambarkan dalam Gambar 2. Semua model dievaluasi menggunakan empat metrik pada data uji, seperti yang digambarkan dalam Tabel 2. Kinerja keseluruhan mengungkapkan bahwa model XGBoost mengungguli model lain dalam memprediksi curah hujan. Dalam hal metrik akurasi, Tabel 2 menunjukkan bahwa XGBoost berkinerja terbaik, dengan akurasi prediksi sebesar 95,35%, dan CatBoost dan RF mencapai kinerja yang sebanding sekitar 93-92%. Dalam kinerja keseluruhan, XGBoost menunjukkan kinerja terbaik dibandingkan dengan model lain, dengan 95,49%. Selain itu, untuk F1, Recall, Precision, dan Accuracy, XGBoost juga memiliki kinerja yang optimal. Sebagai perbandingan, LR memiliki kinerja terburuk dari kelima model ini. Hasil kami menunjukkan bahwa model-model, seperti XGBoost, CatBoost, dan RF, mencapai kinerja yang unggul. Meskipun temuan menunjukkan kinerja yang lebih baik, penulis mencatat bahwa kinerja algoritma dipengaruhi oleh berbagai faktor, seperti kompleksitas dan konfigurasi model serta kualitas data. Hasil eksperimen mengungkapkan bahwa kinerja model ensemble,



XGBoost, lebih baik daripada model lain dan konsisten dengan hasil sebelumnya [23]. Dengan demikian, penelitian kami menggunakan model XGBoost untuk analisis interpretabilitas global dan lokal berikutnya.

Tabel 2. Perbandingan Model ML

Model	Accuracy	Precision	Recall	F1 Score	Overall Performance
XGBoost	95.35%	92.98%	98.14%	95.49%	95.49%
CatBoost	93.33%	90.78%	96.50%	93.55%	93.54%
LightGBM	87.40%	84.76%	91.29%	87.91%	87.84%
RF	92.44%	90.25%	95.21%	92.66%	92.64%
DT	87.30%	85.01%	90.67%	87.75%	87.68%
LR	80.48%	81.09%	79.64%	80.36%	80.39%

3.2 Analisis Interpretabilitas

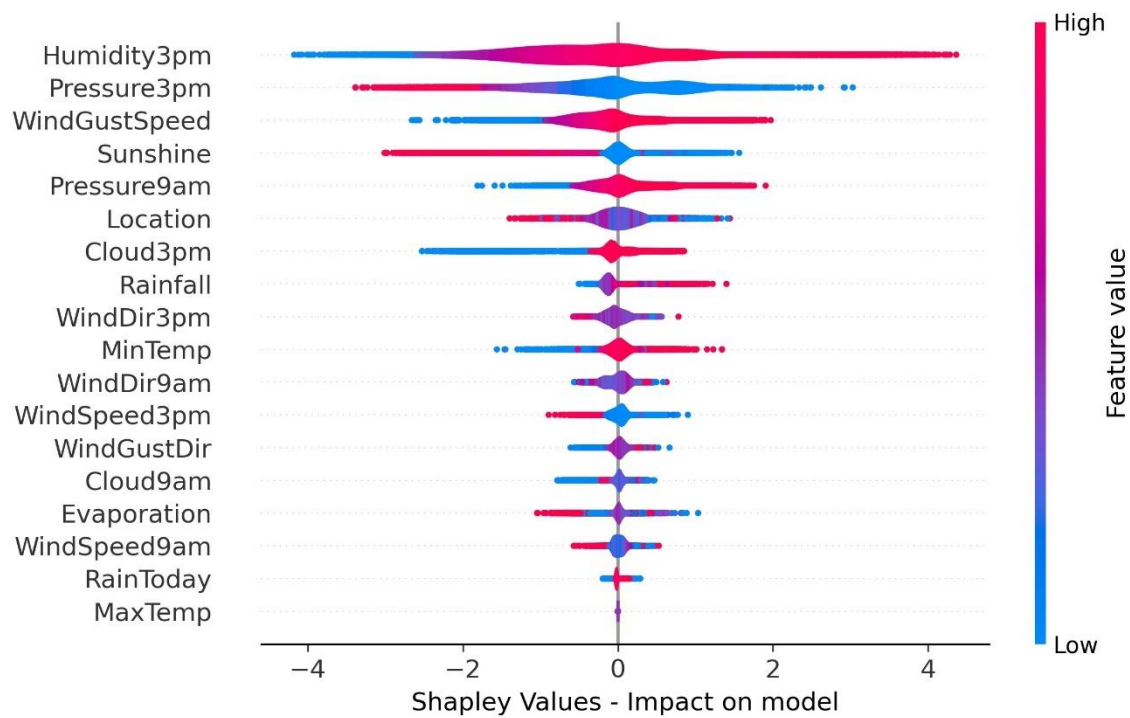
Pada bagian ini, studi ini menyediakan hasil ML yang dapat dijelaskan untuk penjelasan global dan penjelasan lokal menggunakan penjelas SHAP dan LIME.

3.2.1 Analisis Interpretabilitas Global

Tujuan analisis interpretabilitas global adalah untuk menentukan perilaku yang diprediksi dari model pembelajaran mesin dalam hal distribusi nilai lengkap untuk variabel inputnya. Nilai SHAP yang lebih tinggi menunjukkan bahwa model tersebut mencoba untuk memprediksi curah hujan yang lebih tinggi berdasarkan nilai fitur yang sesuai. Demikian pula, nilai SHAP negatif berarti bahwa model tersebut mencoba untuk mendorong prediksi curah hujan ke nilai yang lebih rendah. Perlu dicatat bahwa nilai SHAP tidak mewakili hubungan kausal tetapi menggambarkan perilaku model.

Bagian ini menyajikan hasil analisis interpretabilitas global menggunakan model XGBoost yang memanfaatkan kerangka kerja SHAP. Gambar 2 adalah plot ringkasan SHAP, dengan setiap titik mewakili satu titik data dalam kumpulan data. Lokasi titik-titik pada sumbu x mencerminkan pengaruh nilai karakteristik individual pada prediksi ROP. Nama variabel ditampilkan pada sumbu y dalam urutan kepentingan menaik dari atas ke bawah. Nilai SHAP yang lebih tinggi menunjukkan nilai ROP yang lebih tinggi, seperti halnya titik merah. Nilai SHAP yang lebih rendah sama dengan nilai ROP yang lebih rendah, seperti halnya titik biru. Nilai SHAP yang positif memengaruhi prediksi, menyebabkan model memperkirakan ROP yang lebih besar. Nilai SHAP yang negatif menunjukkan dampak negatif, menyebabkan model memperkirakan ROP yang lebih rendah.

Gambar 2 menunjukkan bahwa lima fitur teratas yang terutama memengaruhi prediksi curah hujan global, dari urutan tertinggi hingga terendah, adalah Humidity3pm, Pressure3pm, Sunshine, WindGustSpeed, dan Pressure9am. Fitur lainnya dapat diabaikan karena memiliki dampak minimal pada prediksi curah hujan, menurut "analisis interpretabilitas global". Seperti yang digambarkan dalam Gambar 2, SHAP untuk analisis interpretabilitas global mengungkap urutan relatif pentingnya fitur pada prediksi curah hujan dalam urutan Humidity3pm > Pressure3pm > WindGustSpeed > Sunshine > Pressure9am. Misalnya, model ML mendorong prediksi curah hujan lebih tinggi (yaitu, nilai Shapley yang lebih tinggi untuk output) saat Humidity3pm, WindGustSpeed, dan Pressure9am tinggi dengan titik merah. Saat variabel-variabel ini memiliki nilai Shapley yang lebih tinggi, model XGBoost lebih mungkin diprediksi akan mengalami hujan besok. Sebaliknya, Pressure3pm dan Sunshine rendah, diwakili oleh titik biru.

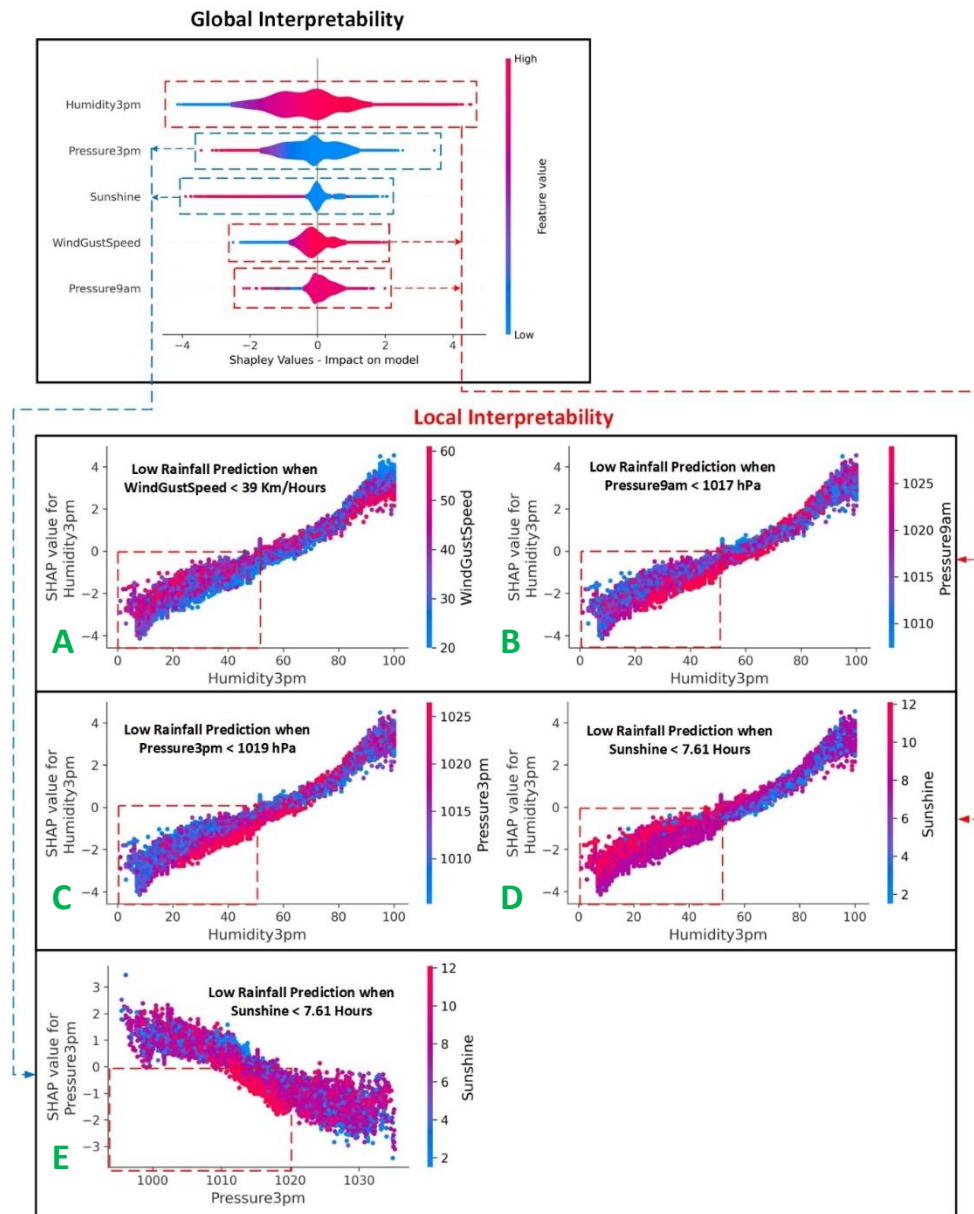


Gambar 2. Plot ringkasan SHAP untuk menjelaskan pengaruh fitur global pada prediksi curah hujan. Humidity3pm adalah fitur yang paling penting, sementara Pressure9am adalah fitur yang paling tidak berpengaruh untuk prediksi curah hujan. Nilai SHAP yang lebih tinggi menunjukkan bahwa model memprediksi nilai curah hujan yang lebih tinggi dan sebaliknya. Misalnya, nilai Humidity3pm Tinggi (merah) sesuai dengan nilai SHAP yang lebih tinggi, yang menunjukkan nilai Humidity3pm yang tinggi.

3.2.2 Analisis Interpretabilitas Lokal

Sementara “analisis interpretabilitas global” SHAP memberikan gambaran umum tentang pengaruh setiap fitur pada prediksi model, analisis interpretabilitas lokal menguraikan tentang bagaimana prediksi model bervariasi dengan nilai fitur dan interaksi. Setiap nilai SHAP parameter input dapat menggambarkan kontribusi setiap fitur. Seperti yang ditunjukkan pada Gambar 4, nilai fitur dan nilai SHAP yang sesuai diplot pada sumbu x (nilai fitur) dan sumbu y (nilai SHAP), masing-masing. Plot ketergantungan lebih lanjut disorot oleh interaksi fitur (melalui bilah warna), yang menunjukkan efek fitur gabungan pada prediksi. Perlu dicatat bahwa nilai SHAP tidak mewakili hubungan kausal tetapi lebih mengkarakterisasi perilaku model.

Analisis interpretabilitas lokal juga memberikan gambaran umum tentang pengaruh setiap fitur pada prediksi model, seperti yang digambarkan pada Gambar 3. Sumbu y kanan berarti nilai setiap fitur. Untuk sumbu y kiri, nilai SHAP positif berarti mendorong model untuk memprediksi curah hujan besok, dan nilai SHAP negatif berarti nilai fitur yang lebih rendah berkontribusi negatif terhadap prediksi curah hujan besok. Seperti yang ditunjukkan pada Gambar 3, nilai SHAP negatif dari Pressure3pm dan Sunshine berkontribusi negatif terhadap prediksi curah hujan besok. Sebaliknya, nilai SHAP positif dari Humidity3pm, WindGustSpeed, dan Pressure9am berarti fitur ini secara substansial memengaruhi prediksi model terhadap curah hujan besok.



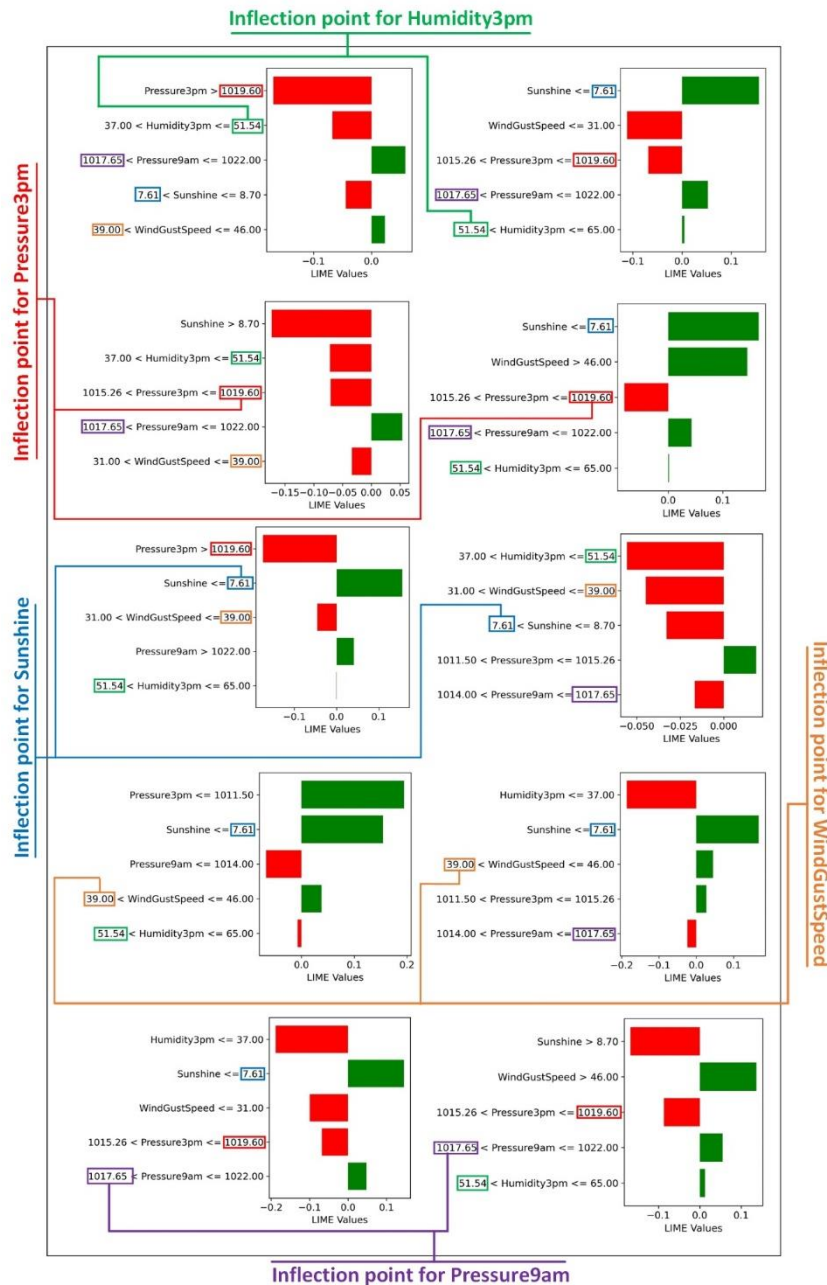
Gambar 3. Plot interpretasi lokal model XGBoost Rainfall antara fitur yang berpengaruh dan curah hujan. Nilai SHAP pada sumbu y menunjukkan pengaruh nilai fitur masing-masing terhadap curah hujan.

Pada Gambar 3(A), efek ketergantungan & interaksi mengindikasikan bahwa nilai curah hujan tinggi dapat dicapai ketika Humidity3pm $\leq 51.54\%$ dan WindGustSpeed ≤ 39 Km/jam. Nilai curah hujan tinggi juga dapat dicapai dengan WindGustSpeed > 39 Km/jam yang lebih tinggi jika Humidity3pm $> 85\%$. Gambar 3(B) memberi tahu kita bahwa nilai curah hujan tinggi dapat dicapai ketika Humidity3pm $\leq 51\%$ dan Pressure9am < 1017 hPa. Dari Gambar 3(C), kita juga dapat mengatakan bahwa curah hujan tinggi dapat diperoleh dengan menggunakan $\sim 20\% \leq$ Humidity3pm $\leq 60\%$ dengan Pressure3pm ≤ 1019 hPa. Seperti yang ditunjukkan pada Gambar 3(D), setiap titik mewakili kecelakaan lalu lintas, ketika nilai curah hujan yang lebih tinggi dapat dicapai ketika Humidity3pm $\leq 51\%$ dan Sunshine < 7.61 Jam. Perlu dicatat juga bahwa Humidity3pm $> 51\%$ dan Sunshine > 7.61 Jam merupakan titik balik di mana nilai SHAP meningkat secara substansial, yang berarti curah hujan jarang terjadi. Selain itu, hubungan yang lebih rumit antara Humidity3pm dan Sunshine dapat dieksplorasi berdasarkan Gambar 3(D). Misalnya, ketika Pressure3pm sekitar $-2 <$ Pressure3pm < 0 , semakin rendah Sunshine, semakin rendah nilai SHAP, sehingga sesuai dengan lebih banyak curah hujan. Gambar 3(E) menunjukkan hubungan terbalik antara Pressure3pm dan Sunshine. Nilai curah hujan yang lebih tinggi dapat diperoleh ketika Pressure3pm < 1019 hPa dan Sunshine < 7.61 Jam.

Untuk menyelidiki lebih lanjut titik infleksi iklim yang mendorong tingkat curah hujan, kami melakukan analisis "interpretabilitas lokal" terperinci dengan LIME yang ditunjukkan pada Gambar 4. Plot ini menjelaskan prediksi dari model curah hujan XGBoost untuk meningkatkan keterjelasannya. Hal ini menunjukkan bahwa titik infleksi Humidity3pm, WindGustSpeed, dan Pressure9am adalah $\sim 51,54\%$, ~ 39 Km/jam, dan $\sim 1017,65$ hPa. Berdasarkan interpretasi ini, kita dapat berhipotesis:

- H1: jika Humidity3pm $\leq \sim 51,54$, WindGustSpeed $\leq \sim 39$ Km/jam, dan Pressure9am $< \sim 1017,65$ hPa, nilai laju curah hujan akan tinggi.
- H2: jika Pressure3pm $< \sim 1019$ hPa, dan Sunshine $\leq \sim 7,61$ jam, nilai laju curah hujan akan tinggi.

Selain itu, hasil LIME konsisten dengan penjelasan SHAP yang diuraikan dalam bagian ini. Oleh karena itu, penjelasan global dan lokal SHAP dan LIME yang digabungkan memberikan representasi akurat dari proses iklim dunia nyata dan meningkatkan keterjelasan model ML dengan memberikan penjelasan sederhana untuk pemahaman manusia.



Gambar 4. Plot interpretasi lokal model curah hujan XGBoost



Bagian ini membahas analisis faktor pengaruh yang menyebabkan prediksi curah hujan besok dan membandingkannya dengan temuan penelitian sebelumnya. Hasil analisis interpretabilitas global menunjukkan bahwa fitur-fitur berikut (dalam urutan kepentingan) berkontribusi paling besar terhadap prediksi curah hujan: Humidity3pm > WindGustSpeed > Pressure9am. Ketika variabel-variabel ini memiliki nilai Shapley yang lebih tinggi, model tersebut lebih mungkin memprediksi curah hujan besok (kelas=ya).

Sedangkan untuk analisis interpretabilitas lokal, hasilnya menunjukkan bahwa titik infleksi Humidity3pm, WindGustSpeed, dan Pressure9am adalah ~51,54 %, ~39 Km/jam, dan ~1017,65 hPa dan Sinar Matahari <= ~7,61 jam.

Secara keseluruhan, ditemukan bahwa model ML yang dapat diinterpretasikan memberikan hasil yang menjanjikan mengenai variabel-variabel yang memengaruhi curah hujan. Metode ini secara efektif menangkap metode yang digunakan model ML untuk menghasilkan prediksi. Model ML yang dapat diinterpretasikan dianggap dapat memuaskan minat, tujuan, harapan, kebutuhan, dan tuntutan untuk sistem buatan (Langer et al. 2021). Studi ini menunjukkan bagaimana model ML yang dapat diinterpretasikan dapat digunakan untuk mengidentifikasi variabel yang memengaruhi curah hujan. Menggabungkan penjelasan global dan lokal memberikan gambaran tentang curah hujan yang sebenarnya dan penjelasan langsung yang mudah dipahami orang.

Keterbatasan studi kami adalah bahwa analisis interpretabilitas curah hujan didasarkan pada kumpulan data kami di lokasi tertentu (misalnya, Australia), sehingga wilayah lain tidak dieksplorasi. Studi mendatang harus menggunakan kumpulan data regional lain untuk menyelidiki fenomena curah hujan.

5. KESIMPULAN DAN SARAN

5.1 Kesimpulan

Studi ini menyajikan sistem prakiraan curah hujan yang dapat ditafsirkan untuk memahami fenomena curah hujan. Akurasi prediksi dari enam ML yang diawasi diintegrasikan ke dalam kerangka kerja yang diusulkan. Teknik ML yang digunakan adalah XGBoost, CatBoost, LightGBM, RF, DT, dan LR. Untuk percobaan tersebut, data cuaca historis selama sembilan tahun (dari 2008 hingga 2017) untuk kota Australia diekstraksi dari situs web prakiraan cuaca yang terdiri dari berbagai fitur terkait cuaca. Aktivitas pra-pemrosesan dilakukan pada kumpulan data yang diekstraksi untuk meningkatkan akurasi prediksi, termasuk pembersihan, penanganan nilai yang hilang, konversi variabel kategoris dan kontinu, penghapusan outlier, dan penyeimbangan kumpulan data menggunakan teknik SMOOTE. Kerangka kerja yang disajikan dalam penelitian ini akan diperluas di masa mendatang dengan mengeksplorasi penggabungan teknik ML ensemble lainnya pada kumpulan data yang lebih beragam.

Selain itu, teknik pemilihan fitur yang tepat juga akan menjadi tambahan praktis bagi sistem, yang memastikan prediksi yang hemat biaya. Selain prediksi curah hujan, sistem prakiraan curah hujan yang dapat ditafsirkan juga akan digunakan untuk area lain, seperti prediksi suhu, prediksi energi surya, tingkat keparahan kecelakaan, dll. Upaya akan dilakukan untuk menggabungkan berbagai bentuk Jaringan Syaraf Tiruan dalam proses peramalan, seperti Multi-Layer Perceptron (MLP) dan jaringan Long Short-Term Memory (LSTM). Studi ini menerapkan kerangka kerja ML yang dapat dijelaskan berbasis SHAP dan LIME untuk meningkatkan kemampuan menjelaskan model XGBoost. Komponen interpretabilitas global dari kerangka kerja SHAP mengungkapkan bahwa kepentingan relatif variabel pada prediksi curah hujan di seluruh area studi adalah dalam urutan Humidity3pm > Pressure3pm > WindGustSpeed > Sunshine > Pressure9am. Kerangka kerja tersebut secara akurat menangkap fisika dasar yang mendorong curah hujan sehingga model tersebut memperkirakan curah hujan yang lebih tinggi ketika Humidity3pm, WindGustSpeed, dan Pressure9am tinggi, dan Pressure3pm dan Sunshine rendah dan sebaliknya. Komponen interpretabilitas lokal dari kerangka kerja SHAP menunjukkan bagaimana prediksi model bervariasi dengan nilai fitur dan interaksi menggunakan plot ketergantungan dan interaksi SHAP yang digabungkan dengan model ML. Interpretabilitas lokal dari kerangka kerja LIME mengukur titik infleksi kritis pada setiap prediktor yang mengarah pada transisi dari tingkat curah hujan rendah ke tinggi. Analisis kami mengonfirmasi bahwa keterjelasan model XGBoost ditingkatkan ketika model tersebut digabungkan dengan kerangka kerja SHAP dan LIME XAI.

Sebagai kesimpulan, model XGBoost yang terintegrasi dengan kerangka kerja XAI menunjukkan kinerja yang luar biasa dalam (i) analisis kinerja mengungguli algoritme lain, (ii) mengidentifikasi dan mengukur urutan pentingnya prediktor yang memengaruhi curah hujan target besok sambil mempertahankan interpretabilitas fisik dari dinamika input-output, dan (iii) menentukan titik infleksi pada prediktor iklim di mana transisi dari curah hujan rendah ke tingkat curah hujan tinggi terjadi.

5.2 Saran

Mengembangkan teknik ML ensemble lebih lanjut dapat dilakukan dengan menggabungkan model ensemble lain seperti Stacking, Bagging, atau Boosting pada dataset yang lebih beragam. Selain itu, menggabungkan pendekatan ML dengan Jaringan Syaraf Tiruan seperti MLP dan LSTM membantu menangkap





pola yang lebih kompleks dalam data. Untuk meningkatkan akurasi prediksi, teknik pemilihan fitur canggih seperti Recursive Feature Elimination (RFE) atau Shapley Additive Explanations (SHAP) dapat digunakan untuk memastikan hanya fitur yang relevan. Sistem ini juga bisa dikembangkan untuk memprediksi fenomena cuaca lain seperti suhu, energi surya, dan tingkat keparahan kecelakaan. Pengujian dan validasi sistem dengan data terbaru dan dari berbagai lokasi penting untuk memastikan generalisasi dan robustnya model. Evaluasi dan penyesuaian model secara berkala berdasarkan hasil prediksi meningkatkan keandalannya. Menyelidiki teknik Explainable AI (XAI) lain seperti Local Interpretable Model-agnostic Explanations (LIME) atau Layer-wise Relevance Propagation (LRP) dapat memperjelas keputusan model.

DAFTAR PUSTAKA

- [1] S. Chapman *et al.*, “Climate Change Impacts on Extreme Rainfall in Eastern Africa in a Convection-Permitting Climate Model,” *Journal of Climate*, vol. 36, no. 1, pp. 93–109, Jan. 2023, doi: 10.1175/JCLI-D-21-0851.1.
- [2] A. D. Griffiths, P. C. Treble, P. Hope, and I. Rudeva, “Rainfall Stable Water Isotope Variability in Coastal Southwestern Western Australia and Its Relationship to Climate on Multiple Timescales,” *Journal of Geophysical Research: Atmospheres*, vol. 127, no. 1, p. e2021JD035433, Jan. 2022, doi: 10.1029/2021JD035433.
- [3] C. Kyei-Mensah, R. Kyerematen, and S. Adu-Acheampong, “Impact of Rainfall Variability on Crop Production within the Worobong Ecological Area of Fanteakwa District, Ghana,” *Advances in Agriculture*, vol. 2019, pp. 1–7, May 2019, doi: 10.1155/2019/7930127.
- [4] L. Wei, L. Jiheng, G. Junhong, B. Zhe, F. Lingbo, and H. Baodeng, “The Effect of Precipitation on Hydropower Generation Capacity: A Perspective of Climate Change,” *Frontiers in Earth Science*, vol. 8, no. September, pp. 1–13, Sep. 2020, doi: 10.3389/feart.2020.00268.
- [5] H. R. Bedane, K. T. Beketie, E. E. Fantahun, G. L. Feyisa, and F. A. Anose, “The impact of rainfall variability and crop production on vertisols in the central highlands of Ethiopia,” *Environmental Systems Research*, vol. 11, no. 1, p. 26, Dec. 2022, doi: 10.1186/s40068-022-00275-3.
- [6] C. Wu, J. M. Chen, A. R. Desai, P. M. Lafleur, and S. B. Verma, “Positive impacts of precipitation intensity on monthly CO₂ fluxes in North America,” *Global and Planetary Change*, vol. 100, pp. 204–214, Jan. 2013, doi: 10.1016/j.gloplacha.2012.10.019.
- [7] Z. Fatima *et al.*, “Field Crops and Climate Change,” in *Building Climate Resilience in Agriculture*, Cham: Springer International Publishing, 2022, pp. 83–94.
- [8] X. Peng and Q. Dai, “Drivers of soil erosion and subsurface loss by soil leakage during karst rocky desertification in SW China,” *International Soil and Water Conservation Research*, vol. 10, no. 2, pp. 217–227, Jun. 2022, doi: 10.1016/j.iswcr.2021.10.001.
- [9] W. Jiang and Y. Tan, “Overview on failures of urban underground infrastructures in complex geological conditions due to heavy rainfall in China during 1994–2018,” *Sustainable Cities and Society*, vol. 76, p. 103509, Jan. 2022, doi: 10.1016/j.scs.2021.103509.
- [10] V. H. S. de Abreu, A. S. Santos, and T. G. M. Monteiro, “Climate Change Impacts on the Road Transport Infrastructure: A Systematic Review on Adaptation Measures,” *Sustainability*, vol. 14, no. 14, p. 8864, Jul. 2022, doi: 10.3390/su14148864.
- [11] X.-Z. Liang, “Extreme rainfall slows the global economy,” *Nature*, vol. 601, no. 7892, pp. 193–194, Jan. 2022, doi: 10.1038/d41586-021-03783-x.
- [12] B. Bochenek and Z. Ustrnul, “Machine Learning in Weather Prediction and Climate Analyses—Applications and Perspectives,” *Atmosphere*, vol. 13, no. 2, p. 180, Jan. 2022, doi: 10.3390/atmos13020180.
- [13] C. Z. Basha, N. Bhavana, P. Bhavya, and V. Sowmya, “Rainfall Prediction using Machine Learning Deep Learning Techniques,” in *Proceedings of the International Conference on Electronics and Sustainable Communication Systems, ICESC 2020*, Jul. 2020, pp. 92–97, doi: 10.1109/ICESC48915.2020.9155896.
- [14] M. Dietze *et al.*, “More than heavy rain turning into fast-flowing water – a landscape perspective on the 2021 Eifel floods,” *Natural Hazards and Earth System Sciences*, vol. 22, no. 6, pp. 1845–1856, Jun. 2022, doi: 10.5194/nhess-22-1845-2022.
- [15] A. Sen, R. Roy, and S. R. Dash, “Smart Farming Using Machine Learning and IoT,” in *Agricultural Informatics*, vol. 3, Wiley, 2021, pp. 13–34.
- [16] H. T. Pham, J. Awange, M. Kuhn, B. Van Nguyen, and L. K. Bui, “Enhancing Crop Yield Prediction Utilizing Machine Learning on Satellite-Based Vegetation Health Indices,” *Sensors*, vol. 22, no. 3, p. 719, Jan. 2022, doi: 10.3390/s22030719.
- [17] W. Sanders, D. Li, W. Li, and Z. N. Fang, “Data-Driven Flood Alert System (FAS) Using Extreme Gradient Boosting (XGBoost) to Forecast Flood Stages,” *Water (Switzerland)*, vol. 14, no. 5, p. 747, 2022, doi: 10.3390/w14050747.





- [18] S. Fahad, F. Su, S. U. Khan, M. R. Naeem, and K. Wei, "Implementing a novel deep learning technique for rainfall forecasting via climatic variables: An approach via hierarchical clustering analysis," *Science of The Total Environment*, vol. 854, p. 158760, Jan. 2023, doi: 10.1016/j.scitotenv.2022.158760.
- [19] A. Y. Barrera-Animas, L. O. Oyedele, M. Bilal, T. D. Akinosho, J. M. D. Delgado, and L. A. Akanbi, "Rainfall prediction: A comparative analysis of modern machine learning algorithms for time-series forecasting," *Machine Learning with Applications*, vol. 7, no. October 2021, p. 100204, 2022, doi: 10.1016/j.mlwa.2021.100204.
- [20] R. Thamilselvan, M. Prabhu, K. T. Selvi, V. Karthik, and T. Bhuvaneshwaran, "An Extensive Review of Rainfall Prediction using Machine Learning and Deep Learning Techniques," *Proceedings of the 2nd International Conference on Artificial Intelligence and Smart Energy, ICAIS 2022*, pp. 198–205, 2022, doi: 10.1109/ICAIS53314.2022.9742746.
- [21] Q. Liu, Y. Zou, X. Liu, and N. Linge, "A survey on rainfall forecasting using artificial neural network," *International Journal of Embedded Systems*, vol. 11, no. 2, p. 240, 2019, doi: 10.1504/IJES.2019.098300.
- [22] J. T. Esteves, G. de Souza Rolim, and A. S. Ferraudo, "Rainfall prediction methodology with binary multilayer perceptron neural networks," *Climate Dynamics*, vol. 52, no. 3–4, pp. 2319–2331, 2019, doi: 10.1007/s00382-018-4252-x.
- [23] M. S. Balamurugan and R. Manojkumar, "Study of short term rain forecasting using machine learning based approach," *Wireless Networks*, vol. 27, no. 8, pp. 5429–5434, Nov. 2021, doi: 10.1007/s11276-019-02168-3.
- [24] H. Yin, F. Zheng, H. F. Duan, D. Savic, and Z. Kapelan, "Estimating Rainfall Intensity Using an Image-Based Deep Learning Model," *Engineering*, 2022, doi: 10.1016/j.eng.2021.11.021.
- [25] K. Boonyuen, P. Kaewprapha, and P. Srivihok, "Daily rainfall forecast model from satellite image using Convolution neural network," in *Proceeding of 2018 3rd International Conference on Information Technology, IncIT 2018*, 2018, pp. 1–7, doi: 10.23919/INCIT.2018.8584886.
- [26] A. R. Naik, A. V. Deorankar, and P. B. Ambhore, "Rainfall Prediction based on Deep Neural Network: A Review," *2nd International Conference on Innovative Mechanisms for Industry Applications, ICIMIA 2020 - Conference Proceedings*, no. Icimia, pp. 98–101, 2020, doi: 10.1109/ICIMIA48430.2020.9074892.
- [27] T. Kashiwao, K. Nakayama, S. Ando, K. Ikeda, M. Lee, and A. Bahadori, "A neural network-based local rainfall prediction system using meteorological data on the Internet: A case study using data from the Japan Meteorological Agency," *Applied Soft Computing Journal*, vol. 56, pp. 317–330, Jul. 2017, doi: 10.1016/j.asoc.2017.03.015.
- [28] A. Holzinger, G. Langs, H. Denk, K. Zatloukal, and H. Müller, "Causability and explainability of artificial intelligence in medicine," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 9, no. 4, pp. 1–13, 2019, doi: 10.1002/widm.1312.
- [29] C. Arteaga, A. Paz, and J. W. Park, "Injury severity on traffic crashes: A text mining with an interpretable machine-learning approach," *Safety Science*, vol. 132, no. August, p. 104988, 2020, doi: 10.1016/j.ssci.2020.104988.
- [30] R. Elshawi, M. H. Al-Mallah, and S. Sakr, "On the interpretability of machine learning-based model for predicting hypertension," *BMC Medical Informatics and Decision Making*, vol. 19, no. 1, p. 146, Dec. 2019, doi: 10.1186/s12911-019-0874-0.
- [31] F. Zhipeng and H. Gani, "Interpretable Models for the Potentially Harmful Content in Video Games Based on Game Rating Predictions," *Applied Artificial Intelligence*, vol. 36, no. 1, 2022, doi: 10.1080/08839514.2021.2008148.
- [32] W. T. Abebe and D. Endalie, "Artificial intelligence models for prediction of monthly rainfall without climatic data for meteorological stations in Ethiopia," *Journal of Big Data*, vol. 10, no. 1, p. 2, Jan. 2023, doi: 10.1186/s40537-022-00683-3.
- [33] S. Hudnurkar and N. Rayavarapu, "On the performance analysis of rainfall prediction using mutual information with artificial neural network," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 13, no. 2, p. 2101, Apr. 2023, doi: 10.11591/ijece.v13i2.pp2101-2113.
- [34] D. Pirone, L. Cimorelli, G. Del Giudice, and D. Pianese, "Short-term rainfall forecasting using cumulative precipitation fields from station data: a probabilistic machine learning approach," *Journal of Hydrology*, vol. 617, p. 128949, Feb. 2023, doi: 10.1016/j.jhydrol.2022.128949.
- [35] F. Ouallouche, K. Labadi, Y. Mohia, M. Lazri, and S. Ameer, "Artificial Intelligence for Satellite Image Processing: Application to Rainfall Estimation," in *Lecture Notes in Electrical Engineering*, vol. 959, Springer, 2023, pp. 165–174.
- [36] A. Antal, A. Dumitrescu, S. Cheval, and P. M. P. Guerreiro, "Enhanced precipitation prediction using DEM-based predictors and satellite imagery," *International Journal of Climatology*, Jan. 2023, doi: 10.1002/joc.7987.





- [37] L. Yang *et al.*, "Radar Composite Reflectivity Reconstruction Based on FY-4A Using Deep Learning," *Sensors*, vol. 23, no. 1, p. 81, Dec. 2022, doi: 10.3390/s23010081.

