

INTEGRASI LONG SHORT-TERM MEMORY DAN LARGE LANGUAGE MODEL UNTUK PREDIKSI CURAH HUJAN DAN MITIGASI KEKERINGAN

Heru Hidayatulloh^{1*}, Sri Mulyono^{2*}, Luluk Indah Rahmawati^{3*}

^{1,3}Universitas Islam Sultan Agung, Indonesia

heruhidayatulloh2@gmail.com, ^{1*}sri.m@unissula.ac.id, ^{2*}

lulukindah22@gmail.com, ^{3*}

Info Artikel	Abstrak
Pengajuan : 28-07-2026 Peninjauan : 10-08-2026 Revisi : 19-08-2026 Diterima : 25-08-2026 Diterbitkan : 31-08-2026 Kata Kunci: Long Short-Term Memory, Large Language Model, Prediksi Curah Hujan, Mitigasi Kekeringan, Prediksi Deret Waktu	<i>Kekeringan merupakan salah satu bencana hidrometeorologi yang berkaitan dengan rendahnya curah hujan dan dapat berdampak pada ketersediaan air, pertanian, serta aktivitas masyarakat. Penelitian ini bertujuan mengembangkan sistem prediksi curah hujan menggunakan Long Short-Term Memory (LSTM) yang diintegrasikan dengan Large Language Model (LLM) melalui Gemini AI untuk mendukung interpretasi kondisi curah hujan dan mitigasi risiko kekeringan di Kabupaten Kendal. Data yang digunakan berupa 447 data historis curah hujan dasarian dari BMKG dengan rentang nilai curah hujan historis sebesar 0–[nilai maksimum] mm. Tahapan penelitian meliputi pembagian data secara kronologis menjadi 80% training, 10% validation, dan 10% testing, normalisasi Min-Max Scaling, pembentukan sliding window 24 periode, pelatihan LSTM, serta iterative forecasting selama 36 periode dasarian. Hasil evaluasi menghasilkan MAE sebesar 20,8357 mm, MSE sebesar 736,7793 mm², RMSE sebesar 27,1437 mm, dan R² sebesar 0,7339. Penentuan potensi risiko kekeringan didasarkan pada kategori curah hujan rendah menggunakan threshold yang ditetapkan dalam penelitian dan tidak menggunakan indeks kekeringan meteorologis seperti SPI atau SPEI. Hasil prediksi kemudian diklasifikasikan berdasarkan threshold tersebut sebelum dikirimkan ke Gemini AI. Gemini AI berperan sebagai komponen interpretasi yang mengeksekusi prompt engineering berbasis hasil klasifikasi untuk menghasilkan informasi dan rekomendasi mitigasi dalam bahasa alami, bukan sebagai penentu utama risiko. Seluruh pipeline prediksi, klasifikasi, dan interpretasi berhasil diimplementasikan dalam aplikasi web berbasis Streamlit. Sistem memiliki keterbatasan berupa potensi akumulasi galat pada iterative forecasting jangka panjang dan belum mempertimbangkan indeks kekeringan meteorologis.</i>
How to Cite: Hidayatulloh, H., Mulyono, S., & Rahmawati, L. I. (2026). Integrasi long short-term memory dan large language model untuk prediksi curah hujan dan mitigasi kekeringan. <i>Jurnal Rekayasa Sistem Informasi dan Teknologi (JRSIT)</i>, 4(1), 38–52. DOI : 10.70248/jrsit.v4i1.4455 This is an open access article under https://creativecommons.org/licenses/by/4.0/ Copyright© 2026 by Author. Published by Yayasan Nuraini Ibrahim Mandiri	



PENDAHULUAN

Indonesia merupakan negara beriklim tropis yang memiliki dua musim, yaitu musim hujan dan musim kemarau. Musim kemarau yang berlangsung dalam waktu lama dengan intensitas curah hujan yang rendah dapat memicu terjadinya bencana hidrometeorologi, salah satunya kekeringan (Kratzert et al., 2018). Kekeringan terjadi akibat defisit curah hujan dalam periode tertentu yang menyebabkan berkurangnya ketersediaan air di suatu wilayah. Kondisi tersebut tidak hanya berdampak pada sektor pertanian, tetapi juga dapat memengaruhi ketersediaan air bersih, ketahanan pangan, serta stabilitas sosial dan ekonomi masyarakat (Supriyanto et al., 2024); (Juliansyah et al., 2025); (Zhang et al., 2022). Oleh karena itu, mitigasi kekeringan memerlukan informasi iklim yang dapat digunakan sebagai dasar untuk mengantisipasi kondisi

kekeringan dan mendukung pengambilan keputusan.

Curah hujan merupakan salah satu parameter klimatologi yang berperan penting dalam menentukan kondisi musim, khususnya di wilayah tropis seperti Indonesia (Wani et al., 2024). Ketepatan dalam memprediksi curah hujan menjadi aspek strategis untuk menghadapi ketidakpastian iklim serta mendukung pengelolaan sumber daya air, sektor pertanian, dan mitigasi bencana hidrometeorologi (Hawali, 2025);(Maulana et al., 2026);(Vellore et al., 2026);(Mceachran et al., 2022). Kabupaten Kendal memiliki wilayah yang membentang dari kawasan pesisir hingga dataran rendah dan wilayah perbukitan, sehingga kondisi geografis dan karakteristik wilayah dapat menyebabkan variasi kondisi curah hujan antarperiode. Variabilitas tersebut menjadikan informasi curah hujan yang lebih terperinci diperlukan untuk mendukung pemantauan kondisi hidrometeorologi di wilayah Kabupaten Kendal. Dalam penelitian ini digunakan data curah hujan dasarian karena resolusi tersebut dapat merepresentasikan perubahan curah hujan dalam periode yang lebih pendek dibandingkan data bulanan, tetapi tetap lebih teragregasi dibandingkan data harian sehingga sesuai dengan ketersediaan data penelitian dan kebutuhan prediksi kondisi curah hujan secara berkala. Oleh karena itu, pengembangan metode prediksi curah hujan dengan tingkat akurasi yang baik menjadi salah satu topik penelitian yang terus berkembang.

Long Short-Term Memory (LSTM) memiliki kemampuan dalam mempelajari dependensi temporal pada data deret waktu, tetapi penerapannya pada peramalan jangka panjang memiliki tantangan ketika model *single-step* digunakan secara rekursif (Hosseini et al., 2025);(Xiang et al., 2020);(Mao et al., 2021);(Kratzert et al., 2019). Pada skema tersebut, hasil prediksi pada suatu periode digunakan kembali sebagai masukan untuk periode berikutnya(Mishra et al., 2025). Kesalahan pada prediksi awal dapat ikut terbawa ke langkah berikutnya sehingga berpotensi menyebabkan akumulasi galat seiring bertambahnya horizon peramalan (Miri et al., 2026); (Kratzert et al., 2021);(Ibrahim et al., 2025). Kondisi tersebut menjadi tantangan teknis karena semakin panjang horizon prediksi, semakin banyak hasil prediksi model yang digunakan kembali sebagai masukan sehingga ketidakpastian prediksi dapat meningkat. Tantangan ini menjadi semakin relevan dalam penelitian ini karena model digunakan untuk menghasilkan prediksi selama 36 periode dasarian atau setara dengan satu tahun. Pencapaian prediksi hingga 36 dasarian diperlukan agar sistem tidak hanya memberikan informasi satu periode ke depan, tetapi juga memberikan gambaran kecenderungan curah hujan dalam jangka waktu satu tahun yang dapat digunakan sebagai informasi pendukung perencanaan mitigasi. Permasalahan tersebut menjadi semakin penting pada penelitian ini karena model digunakan untuk menghasilkan prediksi selama 36 periode dasarian atau setara dengan satu tahun. Strategi *iterative forecasting* dipilih karena model yang dikembangkan merupakan model *single-step*, tetapi hasil peramalan jangka panjang perlu dipahami sebagai estimasi kecenderungan, bukan sebagai prediksi dengan tingkat kepastian yang sama pada setiap horizon. Permasalahan akumulasi galat pada strategi rekursif telah dikenal dalam penelitian *multi-step time-series forecasting*. (Meteorologi & Inten, 2025);(Supriatna et al., 2025);(Hermawan et al., 2025); (Prasetya et al., 2026);(Farman et al., 2026). Dengan demikian, tantangan penelitian tidak hanya terletak pada kemampuan LSTM menghasilkan prediksi satu periode ke depan, tetapi juga pada kemampuan mempertahankan kestabilan pola ketika prediksi diperluas hingga 36 periode secara iteratif(Raval et al., 2021);(Hop et al., 2024). Selain itu, hasil prediksi numerik perlu diterjemahkan menjadi informasi yang dapat dipahami oleh pengguna nonteknis. Oleh karena itu, penelitian ini menggabungkan model LSTM dengan Gemini AI sebagai

komponen interpretasi berbasis *prompt engineering*, sehingga hasil numerik dapat disajikan bersama kategori curah hujan, informasi potensi risiko, dan rekomendasi mitigasi (Talan, 2026);(Wang, 2025).

Untuk mendukung pemanfaatan hasil prediksi oleh pengguna nonteknis, diperlukan mekanisme yang mampu menghubungkan keluaran numerik model dengan informasi yang mudah dipahami tanpa menghilangkan konteks hasil prediksi. Salah satu perkembangan kecerdasan buatan yang berpotensi mengatasi keterbatasan tersebut adalah *Large Language Model* (LLM) (Xu et al., 2025);(Onyutha, 2025). LLM merupakan model bahasa berbasis kecerdasan buatan yang mampu memahami konteks dan menghasilkan informasi dalam bahasa alami. Kemampuan tersebut memungkinkan LLM mengubah hasil analisis yang bersifat teknis menjadi informasi yang lebih komunikatif (Cohen et al., 2026);(Karimanzira et al., 2025). Dalam bidang kebencanaan, integrasi LLM telah dimanfaatkan untuk mendukung analisis informasi secara otomatis dan meningkatkan efektivitas penyampaian informasi kepada pengguna. (Rifqi et al., 2025); (Gemma, 2025); (Hendra et al., 2023) Dalam penelitian ini, aliran informasi dirancang secara terstruktur, yaitu hasil prediksi numerik LSTM terlebih dahulu diklasifikasikan berdasarkan threshold curah hujan yang digunakan dalam penelitian, kemudian nilai prediksi, periode, kategori curah hujan, dan indikator potensi risiko dimasukkan ke dalam *prompt* sebagai konteks terstruktur sebelum dikirimkan kepada Gemini AI. Dengan demikian, Gemini AI tidak menentukan kategori risiko secara bebas, tetapi berperan menginterpretasikan informasi yang telah dihasilkan melalui model dan aturan klasifikasi. Selain itu, (Ripai et al., 2026); (Bendi, 2024)(Wiranta & Situmorang, 2025); (Desifa Ramdani Minhar, 2021);(Saengtabtim et al., 2025) menjelaskan bahwa LLM dapat berperan sebagai antarmuka cerdas yang menerjemahkan hasil analisis teknis menjadi rekomendasi yang lebih mudah dipahami oleh pengambil keputusan nonteknis. Integrasi tersebut kemudian diwujudkan dalam bentuk *dashboard* interaktif berbasis web agar hasil prediksi, kategori kondisi curah hujan, dan rekomendasi mitigasi dapat disajikan secara lebih komunikatif serta mendukung pengambil keputusan nonteknis dalam melakukan pemantauan dan perencanaan mitigasi (Press, 2026). Temuan tersebut menunjukkan bahwa LLM memiliki potensi untuk menginterpretasikan hasil prediksi curah hujan menjadi informasi tingkat risiko kekeringan beserta rekomendasi mitigasinya.

Berdasarkan uraian tersebut, penelitian ini mengusulkan integrasi *Long Short-Term Memory* (LSTM) dan *Large Language Model* (LLM) untuk mendukung mitigasi risiko kekeringan di Kabupaten Kendal. LSTM digunakan untuk memprediksi curah hujan berdasarkan data historis, sedangkan LLM melalui Gemini AI digunakan untuk menginterpretasikan hasil prediksi menjadi informasi mengenai kondisi curah hujan, tingkat risiko kekeringan, serta rekomendasi mitigasi dalam bahasa alami. Integrasi tersebut dirancang melalui alur prediksi, klasifikasi berbasis threshold, penyusunan konteks dalam *prompt*, dan interpretasi menggunakan Gemini AI yang selanjutnya disajikan melalui *dashboard* interaktif. Integrasi kedua metode tersebut diharapkan menghasilkan sistem yang tidak hanya memiliki kemampuan prediksi yang baik, tetapi juga mampu menyajikan informasi yang lebih komunikatif sehingga dapat mendukung proses pengambilan keputusan dalam mitigasi risiko kekeringan.

METODE PENELITIAN

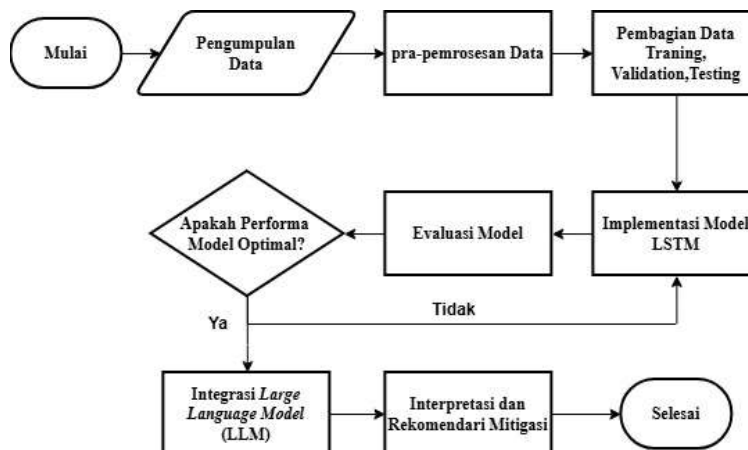
Penelitian ini menggunakan pendekatan *time series forecasting* untuk memprediksi curah hujan menggunakan metode *Long Short-Term Memory* (LSTM) yang diintegrasikan dengan *Large Language Model* (LLM) melalui Gemini AI. Tahapan

penelitian meliputi pengumpulan data, prapemrosesan data, pembangunan model LSTM, evaluasi model, *forecasting* curah hujan, dan interpretasi hasil prediksi menggunakan Gemini AI.

Persiapan Dataset

Tahap pertama penelitian adalah menyiapkan dataset yang akan digunakan sebagai data masukan pada proses pelatihan model. Dataset diperoleh dari Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) berupa data historis curah hujan dasarian Kabupaten Kendal. Data yang digunakan berjumlah 447 data dengan empat atribut, yaitu Tahun, Bulan, Dasarian, dan Curah Hujan (RR). Variabel curah hujan (RR) digunakan sebagai variabel target dalam proses prediksi, sedangkan atribut Tahun, Bulan, dan Dasarian digunakan sebagai penanda urutan waktu pengamatan.

Karena penelitian ini menggunakan pendekatan *time series forecasting*, seluruh data disusun secara kronologis berdasarkan waktu pengamatan. Penyusunan tersebut bertujuan agar model dapat mempelajari hubungan antarperiode sehingga mampu mengenali pola perubahan curah hujan dari waktu ke waktu. Tahap persiapan dataset juga meliputi identifikasi karakteristik data untuk memastikan bahwa dataset layak digunakan pada proses pemodelan.



Gambar 1 Flowchart Penelitian

Preprocessing Data

Tahap *preprocessing* dilakukan untuk meningkatkan kualitas data sebelum digunakan pada proses pelatihan model. Proses ini terdiri atas beberapa tahapan, yaitu data cleaning, penanganan missing value, normalisasi data, pembentukan data deret waktu (*sliding window*), dan pembagian dataset.

Pada tahap *data cleaning*, dilakukan pemeriksaan terhadap struktur dataset, tipe data, data duplikat, dan konsistensi atribut. Apabila ditemukan nilai yang hilang (*missing value*), dilakukan penanganan menggunakan metode interpolasi agar data tetap bersifat kontinu sesuai karakteristik data deret waktu.

Selanjutnya dilakukan normalisasi menggunakan metode Min-Max Scaling untuk mengubah rentang nilai data menjadi 0–1. Normalisasi bertujuan mengurangi perbedaan skala antar data sehingga proses pelatihan model menjadi lebih stabil dan mempercepat proses konvergensi.

Dataset yang telah dinormalisasi kemudian dibagi menjadi 80% data pelatihan, 10% data validasi, dan 10% data pengujian. Pembagian dilakukan tanpa proses *shuffle* agar urutan waktu tetap terjaga.

Tahap terakhir adalah pembentukan data menggunakan teknik *sliding window*.

Teknik ini membentuk sejumlah data historis sebagai masukan (*input sequence*) untuk memprediksi nilai curah hujan pada periode berikutnya. Dengan pendekatan tersebut, model LSTM dapat mempelajari pola temporal yang terdapat pada data historis.

Pelatihan Model LSTM

Model prediksi dibangun menggunakan arsitektur *Long Short-Term Memory* (LSTM) yang dirancang untuk mempelajari pola temporal pada data curah hujan. Arsitektur model terdiri atas dua lapisan LSTM, dua lapisan *Dropout*, dan satu lapisan *Dense* sebagai lapisan keluaran.

Lapisan LSTM pertama digunakan untuk menangkap pola jangka panjang pada data historis, sedangkan lapisan LSTM kedua berfungsi menyempurnakan representasi fitur sebelum menghasilkan prediksi. Lapisan *Dropout* diterapkan setelah setiap lapisan LSTM untuk mengurangi risiko *overfitting*, sementara lapisan *Dense* menghasilkan nilai prediksi curah hujan.

Model dikompilasi menggunakan *Adam Optimizer* dengan fungsi kehilangan (*loss function*) *Mean Squared Error* (MSE). Selama proses pelatihan diterapkan mekanisme *Early Stopping* berdasarkan nilai *validation loss* untuk menghentikan pelatihan secara otomatis apabila performa model tidak lagi mengalami peningkatan. Pendekatan ini bertujuan memperoleh model dengan performa terbaik tanpa mengalami *overfitting*.

Forecasting Curah Hujan

Model LSTM yang telah selesai dilatih digunakan untuk melakukan *forecasting* curah hujan selama 36 periode dasarian atau satu tahun. Proses *forecasting* dilakukan secara iteratif, yaitu setiap hasil prediksi digunakan kembali sebagai data masukan untuk menghasilkan prediksi pada periode berikutnya hingga seluruh periode prediksi terpenuhi.

Karena model LSTM yang digunakan merupakan model *single-step*, peramalan 36 periode menggunakan skema rekursif sehingga terdapat potensi akumulasi galat. Kesalahan pada satu periode dapat terbawa sebagai masukan pada periode berikutnya dan menyebabkan ketidakpastian meningkat seiring bertambahnya horizon prediksi. Dalam penelitian ini, akumulasi galat tidak dikoreksi menggunakan model tambahan atau *recalibration*, sehingga hasil 36 dasarian diposisikan sebagai estimasi kecenderungan curah hujan jangka panjang, bukan sebagai prediksi dengan tingkat kepastian yang sama pada seluruh horizon.

Hasil prediksi 36 dasarian kemudian diakumulasikan berdasarkan bulan untuk memperoleh total curah hujan bulanan selama satu tahun periode prediksi. Hasil *forecasting* berupa nilai curah hujan dalam satuan milimeter (mm). Nilai hasil akumulasi bulanan selanjutnya diklasifikasikan menggunakan aturan *threshold* yang digunakan dalam penelitian untuk memperoleh kategori curah hujan dan indikator potensi risiko kekeringan.

Klasifikasi Curah Hujan dan Potensi Risiko Kekeringan

Hasil akumulasi curah hujan bulanan kemudian diklasifikasikan berdasarkan nilai *threshold* yang digunakan dalam penelitian untuk menentukan kategori curah hujan dan indikator potensi risiko kekeringan. Dengan demikian, klasifikasi dilakukan pada nilai curah hujan bulanan hasil akumulasi, bukan pada nilai prediksi dasarian secara individual.

Tabel 1 Acuan Kategori Curah Hujan Bulanan

Curah Hujn	Kategori	Potensi Resiko
≤ 100	Rendah	Tinggi
$>101-300$	Menengah	Sedang

>301–500	Tinggi	Rendah
>500	Sangat Tinggi	Rendah

Kategori tersebut digunakan sebagai indikator potensi risiko berdasarkan curah hujan bulanan dan bukan sebagai indeks kekeringan meteorologis seperti SPI atau SPEI.

Interpretasi Menggunakan Gemini AI

Tahap interpretasi dilakukan dengan mengintegrasikan hasil *forecasting* curah hujan menggunakan Gemini AI sebagai implementasi *Large Language Model* (LLM). Sebelum dikirimkan ke Gemini AI, hasil *forecasting* diproses terlebih dahulu oleh sistem untuk menghasilkan informasi terstruktur yang terdiri atas periode prediksi, nilai curah hujan, kategori curah hujan, dan indikator potensi risiko berdasarkan aturan klasifikasi penelitian. Struktur tersebut digunakan sebagai konteks yang tetap dipertahankan dalam *prompt* sehingga Gemini AI tidak menentukan kategori risiko berdasarkan interpretasi bebas terhadap nilai numerik.



Gambar 2 Alur Integrasi System

Informasi tersebut diberikan kepada Gemini AI untuk menghasilkan interpretasi kondisi curah hujan, penjelasan tingkat risiko kekeringan, serta rekomendasi mitigasi dalam bentuk bahasa alami. Dengan demikian, Gemini AI berperan sebagai komponen interpretasi yang memproses konteks hasil prediksi dan klasifikasi melalui *prompt engineering*, sedangkan nilai prediksi dan indikator risiko tetap ditentukan oleh model LSTM dan aturan klasifikasi yang telah ditetapkan.

Integrasi tersebut memungkinkan hasil prediksi numerik yang bersifat teknis diterjemahkan menjadi informasi yang lebih komunikatif dan mudah dipahami oleh pengguna. Oleh karena itu, kombinasi LSTM dan LLM dalam penelitian ini tidak ditujukan untuk meningkatkan akurasi prediksi, tetapi untuk menghubungkan hasil prediksi dengan penyajian interpretasi dan rekomendasi mitigasi yang lebih kontekstual.

Evaluasi Model

Evaluasi dilakukan untuk mengukur performa model LSTM dalam memprediksi curah hujan dengan membandingkan hasil prediksi terhadap data aktual pada data pengujian. Pengukuran dilakukan menggunakan empat metrik evaluasi, yaitu *Mean Absolute Error* (MAE), *Mean Squared Error* (MSE), *Root Mean Squared Error* (RMSE), dan *Coefficient of Determination* (R^2).

Nilai MAE digunakan untuk mengukur rata-rata kesalahan absolut antara hasil prediksi dan data aktual. MSE digunakan untuk mengukur rata-rata kuadrat kesalahan prediksi sehingga lebih sensitif terhadap kesalahan yang besar. RMSE menunjukkan besarnya kesalahan prediksi dalam satuan yang sama dengan data asli sehingga lebih mudah diinterpretasikan. Sementara itu, R^2 digunakan untuk mengukur kemampuan model dalam menjelaskan variasi data aktual. Semakin kecil nilai MAE, MSE, dan RMSE

serta semakin mendekati satu nilai R^2 , maka semakin baik performa model yang dihasilkan.

Evaluasi dilakukan pada data testing yang tidak digunakan untuk proses *fit scaler* maupun pelatihan model. Evaluasi ini digunakan untuk mengukur kemampuan model dalam menghasilkan prediksi satu periode berdasarkan pola historis, sedangkan forecasting 36 dasarian dilakukan secara rekursif dan tidak dievaluasi terhadap data aktual karena periode tersebut merupakan periode mendatang. Oleh karena itu, hasil 36 dasarian diperlakukan sebagai estimasi kecenderungan dan perlu mempertimbangkan potensi akumulasi galat pada horizon yang lebih panjang.

Hasil evaluasi selanjutnya digunakan untuk menentukan kelayakan model sebelum diterapkan pada proses forecasting dan integrasi dengan Gemini AI.

HASIL PENELITIAN DAN PEMBAHASAN

Hasil Prapemrosesan Data

Data historis curah hujan dasarian Kabupaten Kendal yang digunakan dalam penelitian ini terdiri atas 447 data pengamatan dengan empat atribut, yaitu Tahun, Bulan, Dasarian, dan Curah Hujan (RR). Setelah tahap prapemrosesan dilakukan, dataset tidak memiliki data duplikat maupun *missing value*, sehingga seluruh data dapat digunakan pada proses pelatihan model.

Tabel 2 Pembagian Dataset

Dataset	Jumlah Data
Training	357
Validation	45
Testing	45
Total	447

Proses normalisasi menggunakan metode *Min-Max Scaling* berhasil mengubah rentang nilai curah hujan ke interval 0–1, sehingga data memiliki skala yang seragam dan sesuai sebagai masukan model LSTM. Selanjutnya, dataset dibagi secara berurutan menjadi 357 data pelatihan (80%), 45 data validasi (10%), dan 45 data pengujian (10%) untuk mempertahankan karakteristik data *time series*. Hasil pembagian dataset ditunjukkan pada Tabel 2.

Tabel 3 Hasil Pembentukan Sliding Window

Membentuk Sliding Window	
Training Shipe	333, 24, 1
Validation Shpe	22, 24, 1
Testing Shape	21, 24, 1

Pembentukan data deret waktu menggunakan teknik *Sliding Window* dengan $\text{TIME_STEP} = 24$ menghasilkan data masukan berukuran (333, 24, 1) untuk pelatihan, (21, 24, 1) untuk validasi, dan (21, 24, 1) untuk pengujian. Bentuk data tersebut menunjukkan bahwa setiap prediksi memanfaatkan 24 periode historis sebagai masukan untuk memperkirakan curah hujan pada periode berikutnya. Ilustrasi hasil pembentukan *Sliding Window* ditunjukkan pada Gambar 3.

Hasil prapemrosesan menunjukkan bahwa dataset telah memenuhi kebutuhan model LSTM, baik dari aspek kualitas data maupun struktur data deret waktu. Normalisasi dan pembentukan *Sliding Window* memungkinkan model mempelajari hubungan temporal antarperiode secara lebih efektif sehingga mendukung proses pelatihan dan prediksi curah hujan.

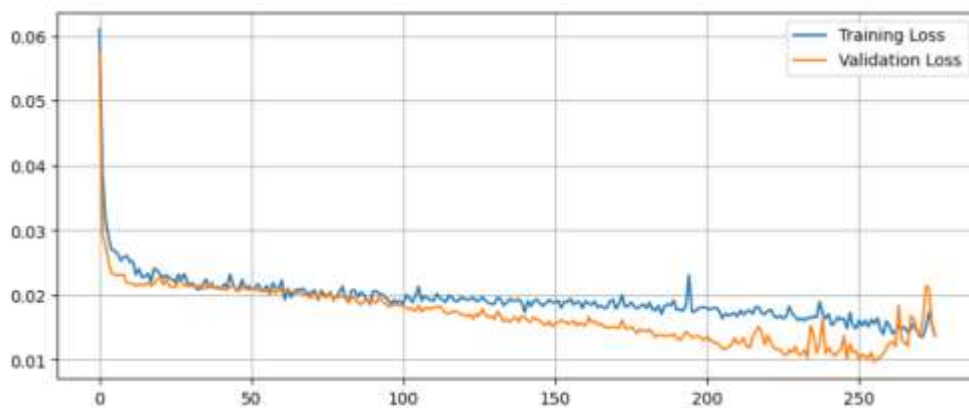
Hasil Pelatihan dan Evaluasi Model LSTM

Model prediksi curah hujan dibangun menggunakan arsitektur *Long Short-Term Memory* (LSTM) yang terdiri atas dua lapisan LSTM, dua lapisan *Dropout*, dan satu lapisan *Dense*. Lapisan LSTM pertama menggunakan 64 unit dengan parameter *return_sequences=True* untuk mempertahankan informasi sekuensial yang diteruskan ke lapisan berikutnya, sedangkan lapisan LSTM kedua menggunakan 32 unit dengan fungsi aktivasi *tanh*. Setiap lapisan LSTM diikuti *Dropout* sebesar 0,2 untuk mengurangi risiko *overfitting*. Lapisan *Dense* dengan satu unit digunakan untuk menghasilkan nilai prediksi curah hujan. Model menggunakan *Adam Optimizer* dengan fungsi kehilangan *Mean Squared Error* (MSE).

Tabel 4 Arsitektur Model LSTM

Lapisan	Konfigurasi
LSTM 1	64 unit, <i>return_sequences=True</i>
Dropout 1	0,2
LSTM 2	32 unit, aktivasi <i>tanh</i>
Dropout 2	0,2
Dense	1 unit
Optimizer	Adam
Loss Function	Mean Squared Error (MSE)

Proses pelatihan dilakukan menggunakan maksimum 300 epoch, *batch size* sebesar 8, serta mekanisme *Early Stopping* dengan nilai *patience* 20 dan *restore best weights* bernilai *True*. Grafik *training loss* dan *validation loss* menunjukkan bahwa kedua kurva mengalami penurunan secara konsisten hingga mencapai kondisi konvergen. Jarak kedua kurva relatif berdekatan sehingga menunjukkan bahwa model memiliki kemampuan generalisasi yang baik dan tidak mengalami *overfitting* secara signifikan.

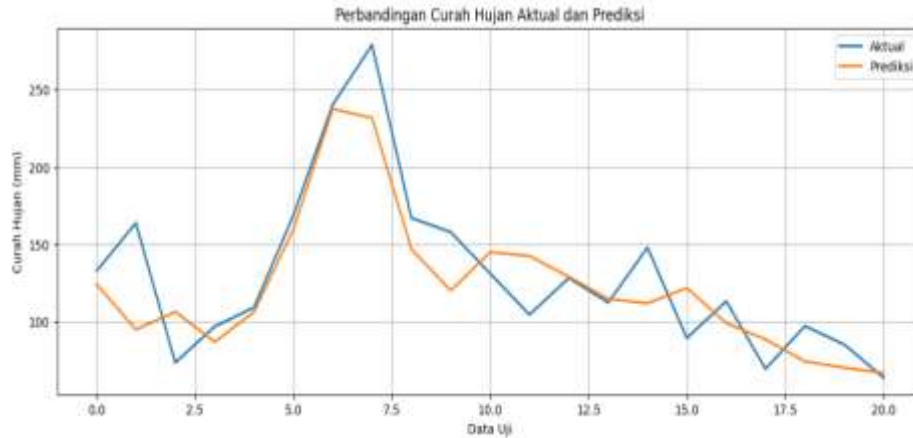


Gambar 3 Grafik *Training Loss* dan *Validation Loss*

Berdasarkan Gambar 3, nilai *training loss* dan *validation loss* mengalami penurunan yang cukup tajam pada tahap awal pelatihan, kemudian menurun secara bertahap hingga mendekati kondisi konvergen. Kedua kurva relatif berdekatan pada sebagian besar proses pelatihan, meskipun terdapat fluktuasi pada beberapa *epoch* di bagian akhir. Kondisi tersebut menunjukkan bahwa model dapat mempelajari pola data dengan baik tanpa menunjukkan indikasi *overfitting* yang dominan selama proses pelatihan.

Setelah proses pelatihan selesai, model diuji menggunakan data pengujian untuk mengetahui kemampuan prediksi terhadap data yang belum pernah dipelajari

sebelumnya. Perbandingan antara data aktual dan hasil prediksi menunjukkan bahwa model mampu mengikuti pola perubahan curah hujan meskipun masih terdapat selisih pada beberapa periode. Hal tersebut menunjukkan bahwa model berhasil menangkap pola temporal pada data historis dengan cukup baik.



Gambar 4 Perbandingan Curah Hujan Aktual dan Hasil Prediksi

Berdasarkan Gambar 4, hasil prediksi secara umum mampu mengikuti arah perubahan curah hujan aktual, terutama pada pola kenaikan dan penurunan secara umum. Namun, pada beberapa periode terdapat perbedaan antara nilai aktual dan prediksi, terutama ketika terjadi perubahan curah hujan yang cukup besar. Hal ini menunjukkan bahwa model mampu menangkap pola temporal data, tetapi belum sepenuhnya merepresentasikan seluruh variasi curah hujan aktual.

Tabel 5 Hasil Evaluasi Model

Metrik	Nilai
MAE	20.8357
MSE	736.7793
RMSE	27.1437
R^2	0.7339

Berdasarkan Tabel 6, model menghasilkan nilai MAE sebesar 20,8357 mm, MSE sebesar 736,7793 mm², dan RMSE sebesar 27,1437 mm. Nilai tersebut menunjukkan besarnya kesalahan antara hasil prediksi dan data aktual pada data pengujian. Sementara itu, nilai R^2 sebesar 0,7339 menunjukkan bahwa model mampu menjelaskan sebesar 73,39% variasi curah hujan pada data pengujian. Hasil evaluasi tersebut menunjukkan bahwa model memiliki kemampuan prediksi yang memadai untuk digunakan pada tahap forecasting.

Hasil *Forecasting* Curah Hujan

Forecasting dilakukan menggunakan model LSTM yang telah memenuhi kriteria evaluasi untuk memprediksi curah hujan selama 36 periode dasarian atau setara dengan satu tahun. Proses *forecasting* dilakukan secara iteratif, yaitu setiap hasil prediksi digunakan kembali sebagai data masukan untuk menghasilkan prediksi pada periode berikutnya hingga seluruh periode prediksi terpenuhi.

Hasil *forecasting* menunjukkan adanya perubahan nilai curah hujan antarperiode, dengan kecenderungan penurunan pada beberapa periode awal sebelum mengalami peningkatan kembali. Berdasarkan klasifikasi BMKG, hasil

prediksi dapat dikelompokkan ke dalam kategori curah hujan rendah, menengah, tinggi, dan sangat tinggi. Informasi kategori tersebut digunakan sebagai dasar dalam menentukan tingkat risiko kekeringan sebelum dilakukan interpretasi menggunakan Gemini AI. Dengan demikian, proses *forecasting* tidak hanya menghasilkan nilai numerik, tetapi juga menjadi dasar penyusunan informasi mitigasi.

Tabel 6 Hasil *Forecasting* Curah Hujan

Tahun	Bulan	Prediksi Bulanan (mm)
2026	Jun	173,24
2026	Jul	136,34
2026	Agu	123,54
2026	Sep	171,02
2026	Okt	227,08
2026	Nov	278,43
2026	Des	341,72
2027	Jan	447,25
2027	Feb	465,76
2027	Mar	441,85
2027	Apr	353,44
2027	Mei	279,87

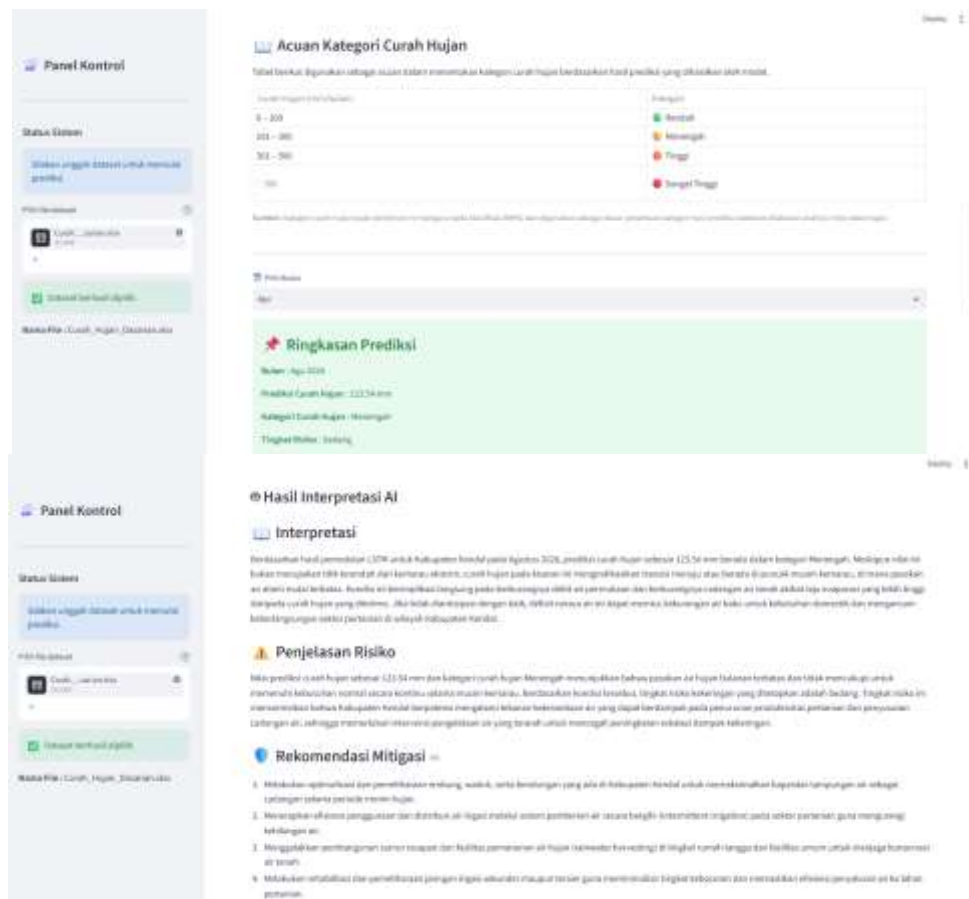
Berdasarkan Tabel 7, curah hujan diprediksi menurun hingga Agustus 2026, kemudian meningkat dan mencapai nilai tertinggi 465,76 mm pada Februari 2027 sebelum kembali menurun hingga Mei 2027. Pada *recursive forecasting*, penggunaan hasil prediksi sebagai input berikutnya berpotensi menyebabkan akumulasi galat dan *loss of variance*, sehingga hasil jangka panjang perlu dipahami sebagai kecenderungan curah hujan. Selanjutnya, 12 prediksi bulanan digunakan untuk klasifikasi berdasarkan *threshold* dan menjadi konteks bagi Gemini AI dalam menghasilkan interpretasi serta rekomendasi mitigasi.

Hasil Implementasi Sistem

Hasil implementasi menunjukkan bahwa sistem berhasil mengintegrasikan proses prediksi curah hujan menggunakan *Long Short-Term Memory* (LSTM) dengan interpretasi berbasis Gemini AI sebagai implementasi *Large Language Model* (LLM) dalam satu alur kerja. Sistem dikembangkan menggunakan *framework Streamlit* sehingga pengguna dapat menjalankan proses prediksi melalui antarmuka yang sederhana tanpa perlu berinteraksi langsung dengan model LSTM maupun Gemini AI.

Setelah proses prediksi dijalankan, sistem secara otomatis menampilkan hasil *forecasting* curah hujan dalam bentuk nilai prediksi dan kategori curah hujan berdasarkan klasifikasi BMKG. Selanjutnya, hasil tersebut dikirimkan ke Gemini AI untuk menghasilkan interpretasi kondisi curah hujan, analisis tingkat risiko kekeringan, serta rekomendasi mitigasi dalam bentuk bahasa alami

Implementasi ini menunjukkan bahwa integrasi antara LSTM dan Gemini AI mampu mengubah hasil prediksi yang semula berupa nilai numerik menjadi informasi yang lebih komunikatif dan mudah dipahami. Pengguna tidak hanya memperoleh informasi mengenai besarnya curah hujan yang diprediksi, tetapi juga mendapatkan penjelasan mengenai kondisi yang diperkirakan terjadi beserta rekomendasi tindakan mitigasi yang sesuai. Dengan demikian, sistem yang dikembangkan tidak hanya berfungsi sebagai alat prediksi curah hujan, tetapi juga sebagai pendukung pengambilan keputusan dalam mitigasi risiko kekeringan.



Gambar 5 Implementasi Sistem

Gambar 7 menunjukkan tampilan sistem yang terdiri atas hasil prediksi, klasifikasi risiko, serta interpretasi dan rekomendasi mitigasi berbasis Gemini AI. Pada bagian hasil prediksi, sistem menampilkan nilai curah hujan bulanan beserta kategori dan tingkat risiko kekeringan. Selanjutnya, hasil tersebut diinterpretasikan oleh Gemini AI menjadi penjelasan mengenai kondisi curah hujan, tingkat risiko, dan rekomendasi mitigasi dalam bentuk bahasa alami. Tampilan tersebut menunjukkan bahwa sistem mampu menyajikan hasil prediksi secara lebih informatif dan mudah dipahami oleh pengguna.

KESIMPULAN

Penelitian ini berhasil mengembangkan sistem prediksi curah hujan menggunakan metode *Long Short-Term Memory* (LSTM) yang diintegrasikan dengan *Large Language Model* (LLM) melalui Gemini AI untuk mendukung mitigasi risiko kekeringan. Model LSTM menghasilkan MAE sebesar 20,8357 mm, MSE sebesar

736,7793 mm², RMSE sebesar 27,1437 mm, dan R² sebesar 0,7339, yang menunjukkan kemampuan menjelaskan 73,39% variasi curah hujan. Model digunakan untuk melakukan *forecasting* selama 36 periode dasarian. Penerapan *iterative forecasting* berpotensi menyebabkan akumulasi galat dan *loss of variance* pada horizon yang semakin panjang, serta model memiliki keterbatasan dalam merepresentasikan fenomena iklim ekstrem yang dapat memengaruhi kondisi kekeringan.

Hasil *forecasting* kemudian diklasifikasikan menggunakan aturan *threshold* untuk memperoleh indikator risiko kekeringan dan diinterpretasikan oleh Gemini AI menjadi informasi serta rekomendasi mitigasi dalam bahasa alami. Integrasi tersebut menunjukkan bahwa Gemini AI dapat digunakan untuk mengonversi hasil prediksi dan klasifikasi berbasis aturan menjadi informasi yang lebih mudah dipahami, dengan risiko dan kategori tetap ditentukan oleh model LSTM dan aturan *threshold*. Sistem berhasil diimplementasikan secara interaktif berbasis web menggunakan Streamlit dan dapat digunakan sebagai informasi pendukung bagi BPBD dan Dinas Pertanian Kabupaten Kendal dalam pemantauan kondisi curah hujan dan perencanaan mitigasi kekeringan, bukan sebagai satu-satunya dasar pengambilan keputusan. Keandalan penggunaan pada skala luas masih perlu diuji, khususnya terkait *latency*, konsistensi keluaran Gemini AI, dan kestabilan koneksi API.

Pernyataan mengenai Penggunaan Kecerdasan Buatan (AI).

AI digunakan dalam penyusunan naskah ini. Penulis menggunakan perangkat kecerdasan buatan semata-mata untuk membantu perbaikan bahasa, pemeriksaan tata bahasa, serta penyuntingan demi kejelasan dan keterbacaan. Penulis tetap bertanggung jawab penuh atas akurasi, orisinalitas, integritas, dan konten akademis naskah tersebut. Seluruh konten yang dibantu oleh AI telah ditinjau dan diverifikasi oleh penulis sebelum naskah diserahkan.

Konflik kepentingan

Penulis menyatakan tidak ada konflik kepentingan.

Kontribusi penulis

Konseptualisasi, Heru Hidayatulloh; pengumpulan data, Heru Hidayatulloh; metodologi, Heru Hidayatulloh; analisis formal, Heru Hidayatulloh; pengembangan model, Heru Hidayatulloh; pengembangan sistem, Heru Hidayatulloh; visualisasi, Heru Hidayatulloh; penulisan—penyusunan draf awal, Luluk Indah Rahmawati; penulisan—peninjauan dan penyuntingan, Heru Hidayatulloh dan Luluk Indah Rahmawati. Penulis telah membaca dan menyetujui versi akhir naskah ini.

Ucapan terima kasih

Penelitian ini didukung oleh Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) melalui penyediaan data historis curah hujan yang digunakan dalam penelitian ini. Penulis mengucapkan terima kasih atas dukungan dan penyediaan data yang diberikan sehingga dapat mendukung pelaksanaan dan penyelesaian penelitian ini.

DAFTAR PUSTAKA

Bendi, M. I. (2024). *Informasi Peringatan Dini Potensi Kekeringan Meteorologis Provinsi*

- Nusa Tenggara Timur. 7(2), 46–54.
- Cohen, I., Littor, N., Elyashar, A., Cohen, O., & Puzis, R. (2026). Progress in Disaster Science Effects of uncertainty and affective content on large language models ' disaster assessment: A controlled comparison using synthetic disaster data ☆. *Progress in Disaster Science*, 30(December 2025), 100571. <https://doi.org/10.1016/j.pdisas.2026.100571>
- Desifa Ramdani Minhar, F. A. (2021). MITIGASI BENCANA DALAM MENGATASI KEKERINGAN DI KALURAHAN GAYAMHARJO KAPANEWON PRAMBANAN KABUPATEN SLEMAN DAERAH ISTIMEWA YOGYAKARTA. 5(1), 368–381.
- Farman, H., Hussain, M. A., Shaikh, S., & Hassan, S. (2026). Dual framework for rainfall prediction : a multi-seed machine and deep learning evaluation across Pakistan ' s climatic regimes. 1–48.
- Gemma, M. D. A. N. (2025). PREDIKSI CURAH HUJAN DI KABUPATEN BOGOR MENGGUNAKAN LONG SHORT-TERM. 9(1), 147–157.
- Hawali, M. L. (2025). Pemilihan Neuron LSTM dan LSTM Bayesian Optimization Untuk Prediksi Curah Hujan Bulanan Berbasis Iklim. 15(2), 376–382.
- Hendra, Y., Mukhtar, H., & Hafsari, R. (2023). Jurnal Software Engineering and Information System (SEIS) PREDIKSI CURAH HUJAN DI KOTA PEKANBARU MENGGUNAKAN LSTM. 3(2).
- Hermawan, T., & Zuliarso, E. (2025). Perbandingan Metode Recurrent Neural Network (RNN) dan Long Short-Term Memory (LSTM) untuk Prediksi Curah Hujan. 7(2), 1450–1463. <https://doi.org/10.47065/bits.v7i2.8099>
- Hop, F. J., Linneman, R., Schnitzler, B., Bomers, A., & Booi, M. J. (2024). Real time probabilistic inundation forecasts using a LSTM neural network. *Journal of Hydrology*, 635(June 2023), 131082. <https://doi.org/10.1016/j.jhydrol.2024.131082>
- Hosseini, F., Prieto, C., & Alvarez, C. (2025). Ensemble learning of catchment-wise optimized LSTMs enhances regional rainfall-runoff modelling – case Study : Basque Country, Spain. 646(August 2024). <https://doi.org/10.1016/j.jhydrol.2024.132269>
- Ibrahim, A., Osman, A., Aldahoul, N., Lun, K., Feng, Y., Lin, J., Elshafie, A., Sherif, M., & Najah, A. (2025). Climate Risk Management A review on machine learning models for drought monitoring and forecasting. *Climate Risk Management*, 50(August), 100758. <https://doi.org/10.1016/j.crm.2025.100758>
- Juliansyah, D., Amlelia, Q., Najiyah, K., Ramdani, R. N., & Nadira, N. (2025). Bencana Kekeringan dalam Perspektif Mitigasi , Rehabilitasi , dan Rekonstruksi serta Dampaknya : Studi Komparatif antara Nusa Tenggara Barat dan Jerman. 4(2), 414–420.
- Karimanzira, D., Rauschenbach, T., & Hellmund, T. (2025). Improved Flood Management and Risk Communication Through Large Language Models. 1–25.
- Kratzert, F., Gauch, M., Nearing, G., Hochreiter, S., & Klotz, D. (2021). Niederschlags-Abfluss-Modellierung mit Long Short-Term Memory (LSTM). 270–280. <https://doi.org/10.1007/s00506-021-00767-z>
- Kratzert, F., Klotz, D., Brenner, C., Schulz, K., & Herrnegger, M. (2018). Rainfall – runoff modelling using Long Short-Term Memory (LSTM) networks. 6005–6022.
- Kratzert, F., Klotz, D., Herrnegger, M., & Sampson, A. K. (2019). Toward Improved Predictions in Ungauged Basins: Exploiting the Power of Machine Learning. <https://doi.org/10.1029/2019WR026065>
- Mao, G., Wang, M., Liu, J., Wang, Z., Wang, K., Meng, Y., Zhong, R., Wang, H., & Li, Y. (2021). Comprehensive comparison of artificial neural networks and long short-

- term memory networks for rainfall-runoff simulation. *Physics and Chemistry of the Earth*, 123, 103026. <https://doi.org/10.1016/j.pce.2021.103026>
- Maulana, M. R., Fazilatunnisa, A., Febriansyah, M. Y., Muiz, A., & Fauzan, I. (2026). *Analisis dan Prediksi Curah Hujan Bulanan Kota Serang Berbasis Apache Spark Menggunakan Dataset BPS Provinsi Banten*. 2(1), 15–21.
- Mceachran, Z., Ghosh, R., Renganathan, A., Sharma, S., Lindsay, K., Steinbach, M., Nieber, J., Duffy, C., & Kumar, V. (2022). *Knowledge - Guided Machine Learning for Operational Flood Forecasting*. 1–18. <https://doi.org/10.1029/2024WR039064>
- Meteorologi, S., & Inten, R. (2025). *Prediksi Curah Hujan Harian Menggunakan Model ARIMA dan LSTM di Stasiun Meteorologi Radin Inten II Lampung*. 20(x), 35–44.
- Miri, S. M., Kavianpour, M. R., & Alizadeh, M. J. (2026). *Delta feature and random forest – enhanced LSTM – attention forecasts with probabilistic postprocessing for rainfall forecasting*. 1–17.
- Mishra, P., Ray, S., Lal, P., Nair, S. B., Matuka, A., Tashkandy, Y., & Emam, W. (2025). *Climate modeling for South Asia: statistical and deep learning for rainfall and temperature prediction*. 1–25.
- Onyutha, C. (2025). *Embracing large language model (LLM) technologies in hydrology research OPEN ACCESS Embracing large language model (LLM) technologies in hydrology. Llm*.
- Prasetya, M. R., Dewi, R. C., Asyhar, A. H., & Hamid, A. (2026). *Prediksi Curah Hujan Harian Menggunakan Metode Principal Component Analysis (PCA) dan Long Short-Term Memory (LSTM)*. 6(1), 1–13.
- Press, A. I. N. (2026). *npj Natural Hazards Article in Press Rescue plan intelligent generation for natural disasters : an integrated approach based on Large Language Models TIES IN ES*.
- Raval, M., Sivashanmugam, P., Pham, V., Gohel, H., & Kaushik, A. (2021). Automated predictive analytics tool for rainfall forecasting. *Scientific Reports*, 1–13. <https://doi.org/10.1038/s41598-021-95735-8>
- Rifqi, R. M., Hajar, M. R., & Haryanto, W. T. (2025). *DETEKSI BANJIR BERBASIS LLM (LARGE LANGUAGE MODELS) MENGGUNAKAN DATA TWITTER / X VIA CHATBOT WHATSAPP*. 2, 61–67.
- Ripai, I., Baswardono, W., Abdurahman, N., Rijanto, E., Afrianto, I., Sumitra, I. D., Informasi, D. S., Indonesia, U. K., Bandung, K., Buatan, K., Networks, G. N., Models, L. L., Banji, M., & Perkotaan, B. (2026). *DIGITAL TWINS UNTUK MANAJEMEN BANJIR PERKOTAAN : 10(2)*, 1972–1978.
- Saengtabtum, K., Leelawat, N., Aumnoysombat, R., Saengwongwattana, N., Suktavornprasit, G., Imamura, F., & Tang, J. (2025). Harnessing generative AI for enhanced disaster management : a systematic review. *Big Earth Data*, 9(4), 651–675. <https://doi.org/10.1080/20964471.2025.2521157>
- Supriatna, D., Anggai, S., Informatika, T., & Pamulang, U. (2025). *Analisis Prediksi Curah Hujan di Kota Tangerang Menggunakan Metode LSTM dan GRU*. 5(2), 119–131.
- Supriyanto, A., Zuliarso, E., Suharmanto, E. T., Amalina, H., & Damaryanti, F. (2024). *DROUGHT PREDICTION USING LSTM MODEL WITH STANDARDIZED PRECIPITATION INDEX ON THE NORTH COAST OF CENTRAL JAVA*. 5(6), 1873–1882.
- Talan, T. (2026). *Machine learning-based rainfall prediction across temporal scales : model benchmarking and explainability analysis*. 1–17.
- Vellore, R. K., Islam, S., Khare, M., & Kulkarni, S. (2026). *Scientific Reports Article in Press A deep learning framework for rainfall forecasting in mumbai metropolitan region IN A Deep Learning Framework for Rainfall Forecasting in Mumbai IN*. 0–37.

- Wang, F. (2025). *Deep Learning Model for Real- - Time Flood Forecasting in Fast- - Flowing Watershed. February 2023*, 18–22. <https://doi.org/10.1111/jfr3.70036>
- Wani, O. A., Mahdi, S. S., Kumar, S. S., Gagnon, A. S., Danish, F., Al-ansari, N., & Hendawy, S. El. (2024). *Predicting rainfall using machine learning , deep learning , and time series models across an altitudinal gradient in the North-Western Himalayas*.
- Wiranta, D. O., & Situmorang, A. (2025). *Kajian Risiko Bencana Kekeringan di Kabupaten Indramayu*. 7(2), 104–128.
- Xiang, Z., Yan, J., & Demir, I. (2020). *A Rainfall - Runoff Model With LSTM - Based Sequence - to - Sequence Learning Water Resources Research*. 1–17. <https://doi.org/10.1029/2019WR025326>
- Xu, F., Ma, J., Li, N., & Cheng, J. C. P. (2025). *International Journal of Disaster Risk Reduction Large language model applications in disaster management: An interdisciplinary review. International Journal of Disaster Risk Reduction*, 127(March), 105642. <https://doi.org/10.1016/j.ijdr.2025.105642>
- Zhang, C., Brodeur, Z. P., Steinschneider, S., & Herman, J. D. (2022). *Leveraging Spatial Patterns in Precipitation Forecasts Using Deep Learning to Support Regional Water Management Water Resources Research*. 1–18. <https://doi.org/10.1029/2021WR031910>