

## PREDIKSI CURAH HUJAN DI KABUPATEN BOGOR MENGUNAKAN *LONG SHORT-TERM MEMORY* DAN *GEMMA 2*

Indra Wijaya<sup>1</sup>, Herlawati Herlawati<sup>2\*</sup>, Rafika Sari<sup>3</sup>  
<sup>1,2,3</sup> Informatika / Ilmu Komputer, Universitas Bhayangkara Jakarta Raya  
\*email: [herlawati@ubharajaya.ac.id](mailto:herlawati@ubharajaya.ac.id)

DOI: <https://doi.org/10.31603/komtika.v9i1.13639>

Received: 23-05-2025, Revised: 28-05-2025, Accepted: 31-05-2025

### ABSTRACT

*Bogor Regency is an area that often experiences prolonged rainfall, especially during the rainy season. High rainfall causes problems such as floods and landslides. Therefore, accurate rainfall prediction is important for various needs, especially in disaster mitigation. This study aims to implement the Long Short-Term Memory (LSTM) algorithm as a model for prediction of historical rainfall data and use the Large Language Model (LLM) GEMMA 2 to provide interpretation of prediction results and recommendations based on the prediction results. The methods used include data collection from the BMKG online data website totaling 1804 data, data pre-processing, model building, model performance evaluation, and interpretation of results using LLM. The results of this study show that LSTM is able to produce the best performance by showing MSE 201.92 mm<sup>2</sup>, Root Mean Square Error (RMSE) of 14.21 mm. the RMSE value shows an average error of 14.21 mm. In addition, the interpretation provided by LLM GEEMA 2 to help understand the prediction and provide practical recommendations for disaster mitigation due to rainfall.*

**Keywords:** *Rainfall, Prediction, LSTM, GEMMA 2, Bogor.*

### ABSTRAK

Kabupaten Bogor adalah daerah sering mengalami curah hujan berkepanjangan, terlebih pada musim hujan. Curah hujan yang tinggi menimbulkan permasalahan seperti banjir dan tanah longsor. Oleh karena itu prediksi curah hujan yang akurat penting untuk berbagai kebutuhan terutama dalam mitigasi bencana. Studi ini bertujuan mengimplementasikan algoritma *Long Short-Term Memory* (LSTM) sebagai model untuk prediksi dari data historis curah hujan dan menggunakan *Large Language Model* (LLM) *GEMMA 2* untuk memberikan interpretasi hasil prediksi serta rekomendasi berdasarkan hasil prediksi. Metode yang digunakan meliputi pengumpulan data dari *website* data online BMKG yang berjumlah 1804 data, pra-pemrosesan data, pembuatan model, evaluasi performa model, serta interpretasi hasil menggunakan LLM. Hasil dari penelitian ini menunjukkan bahwa LSTM mampu menghasilkan performa terbaik dengan menunjukkan MSE 201,92 mm<sup>2</sup>, *Root Mean Square Error* (RMSE) sebesar 14,21 mm. nilai RMSE menunjukkan kesalahan rata – rata 14,21 mm. Selain itu interpretasi yang diberikan oleh LLM GEEMA 2 untuk membantu memahami prediksi dan memberikan rekomendasi praktis untuk mitigasi bencana akibat curah hujan.

**Keywords:** Curah hujan, Prediksi, LSTM, *GEMMA 2*, Bogor.

## PENDAHULUAN

Indonesia adalah negara yang beriklim tropis dan memiliki karakteristik curah hujan yang bervariasi sepanjang tahun dengan pola spasial dan temporal yang beragam. Hal ini memicu bencana hidrometeorologi seperti banjir dan tanah longsor [1]. Semakin tinggi suatu tempat maka semakin besar curah hujan yang diterima, hal ini dipengaruhi oleh topografis suatu wilayah. Salah satu wilayah dengan topografi berbukit dan pegunungan seperti Kabupaten Bogor lebih rentan terhadap bencana seperti tanah longsor [2][3]. Kondisi ini diperparah dengan degradasi lahan dan perubahan tata guna lahan yang menyebabkan peningkatan risiko bencana tanah longsor, sebagaimana yang terjadi di Kecamatan Sukajaya, Kabupaten Bogor [4].

Pada penelitian sebelumnya algoritma *Long Short-Term Memory* (LSTM) dimanfaatkan untuk memprediksi cuaca dengan akurasi yang cukup baik menunjukkan bahwa LSTM mampu memprediksi cuaca dengan akurasi cukup tinggi [5][6][7][8][9][10]. Namun, Sebagian besar penelitian tersebut hanya berfokus pada performa prediktif tanpa memperhastikan bagaimana hasil prediksi dapat dipahami secara luas, baik dalam Masyarakat umum atau pemangku kebijakan.

Seiring dengan perkembangan *Artificial Intelligence*, *Large Language Model* (LLM) seperti GEMMA 2 memberikan pendekatan baru untuk menjabatani hasil prediksi yang sering kali masih sulit dipahami oleh Masyarakat umum. Meski LSTM dan LLM sering digunakan secara terpisah dalam berbagai domain, kombinasi keduanya dalam prediksi curah hujan masih jarang ditemui. Penelitian ini mengisi kekosongan tersebut dengan mengintegrasikan model prediktif LSTM dan GEMMA 2 sebagai alat interpretasi hasil dengan bahasa alami. Implementasi LLM dalam prediksi telah dilakukan oleh Fadillah et al. pada domain analisis emosi di media sosial menggunakan integrasi LightGBM dan LLM Gemini, yang mencapai akurasi 99%. Pendekatan serupa dapat diadaptasi untuk prediksi curah hujan dengan modifikasi input data dan pelatihan model berbasis time-series [11].

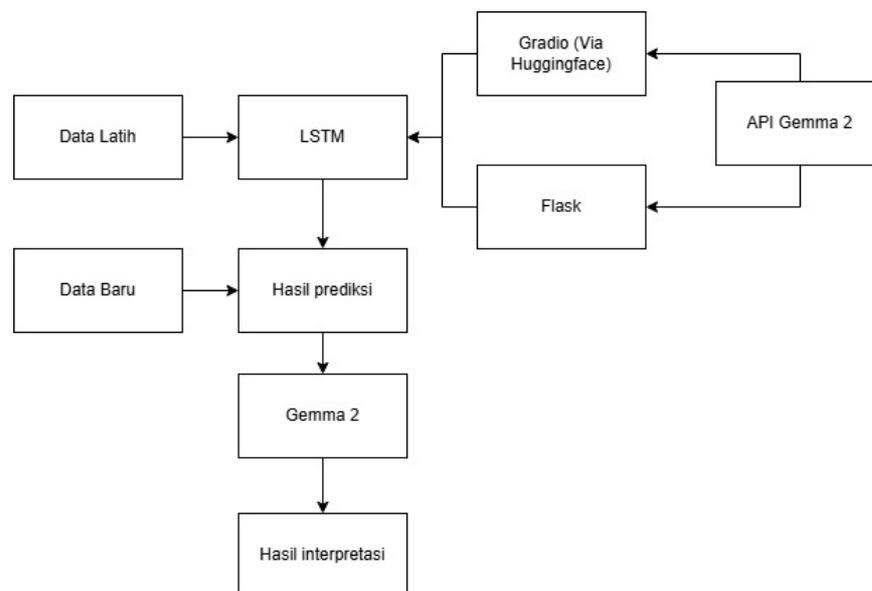
Penelitian ini menggunakan metode CRISP-DM yang mencakup tahap pemahaman bisnis dan data, persiapan data (termasuk penanganan missing value, deteksi outlier, dan normalisasi), pemodelan dengan algoritma LSTM untuk prediksi curah hujan, evaluasi menggunakan MSE dan RMSE, serta deployment melalui Gradio dan Flask. Selain itu, penelitian ini memanfaatkan Large Language Model (LLM) GEMMA 2 untuk memberikan interpretasi hasil prediksi dalam bentuk narasi bahasa alami dan rekomendasi praktis. Proses ini mengindikasikan penggunaan prompt engineering, yaitu teknik untuk menyusun input yang tepat agar GEMMA 2 dapat menghasilkan penjelasan yang relevan dan mudah dipahami oleh pengguna non-teknis.

Selain menggunakan metode tradisional seperti ARIMA atau moving average, prediksi curah hujan juga dapat dilakukan dengan pendekatan berbasis deep learning, salah satunya adalah LSTM, yaitu jenis jaringan saraf tiruan yang dirancang khusus untuk menangani data deret waktu dengan ketergantungan jangka panjang. Metode ini memiliki sejumlah keunggulan dibandingkan metode klasik, antara lain kemampuannya dalam mengingat informasi penting dalam jangka panjang sehingga mampu mengenali pola musiman dan tren jangka panjang pada data curah hujan.

Dalam perkembangannya, model LSTM mulai digantikan oleh arsitektur yang lebih modern seperti Transformer, yang mampu menangani dependensi jangka panjang dengan lebih efisien dan paralel. Model transformer seperti BERT, T5, dan GPT telah mendominasi berbagai tugas pemrosesan bahasa alami dan deret waktu. Saat ini, model *Large Language Model* (LLM) berbasis GPT semakin banyak digunakan karena kemampuannya dalam memahami konteks dan menghasilkan teks secara lebih alami. Beberapa contoh LLM yang berkembang pesat antara lain Gemma dari Google, LLaMA dari Meta, GPT-4 dari OpenAI, serta DeepSeek. Di antara model-model tersebut, Gemma memiliki keunggulan dari sisi efisiensi dan optimalisasi sumber daya, karena dirancang ringan dan terbuka, sehingga lebih mudah diintegrasikan ke dalam perangkat dengan keterbatasan komputasi, tanpa mengorbankan performa pada tugas-tugas inferensi.

## METODE

Penelitian ini bertujuan untuk membangun sistem prediksi curah hujan harian di Kabupaten Bogor dengan mengintegrasikan LSTM dan GEMMA 2.



Gambar 1. Alir Penelitian

Gambar 1 merupakan alur dari penelitian proses diawali dengan pengumpulan data cuaca dari BMKG, yang mencakup parameter seperti suhu rata-rata, kelembapan rata-rata, lamanya penyinaran matahari, dan curah hujan. Data tersebut digunakan sebagai data latih untuk melatih model LSTM karena kemampuannya dalam menangani data deret waktu. Setelah model berhasil dilatih, model siap menerima data cuaca baru untuk menghasilkan prediksi cuaca ke depan. Hasil dari prediksi ini bersifat numerik dan teknis, sehingga diperlukan interpretasi agar bisa lebih mudah dipahami oleh pengguna umum. Oleh karena itu, hasil prediksi dikirim ke model Gemma 2 melalui API, yang akan memberikan penjelasan naratif berdasarkan prompt engineering, seperti menjelaskan dampak dari suhu ekstrem atau curah hujan tinggi.

Untuk menguji sistem ini, digunakan antarmuka Gradio yang di-host melalui HuggingFace Space. Di tahap ini, pengguna dapat mencoba memberikan input dan melihat hasil prediksi serta interpretasi dari model secara langsung melalui web. Sementara itu, untuk

proses deployment akhir, digunakan framework Flask sebagai server backend. Flask menangani permintaan dari pengguna, memproses data menggunakan model LSTM, mengakses API Gemma 2 untuk interpretasi, dan mengembalikan hasil dalam bentuk yang mudah dipahami. Dengan integrasi ini, sistem tidak hanya mampu melakukan prediksi cuaca yang akurat, tetapi juga menjelaskan hasilnya dalam bahasa alami yang informatif dan kontekstual, meningkatkan kemanfaatan hasil prediksi dalam pengambilan keputusan.

#### 1. Data dan Sumber

Data yang digunakan diperoleh dari situs resmi BMKG (<https://dataonline.bmkg.go.id>) dengan rentang waktu Januari 2020 hingga Maret 2025. Data meliputi empat variabel utama: suhu rata-rata (TAVG), kelembaban rata-rata (RH\_AVG), lamanya penyinaran matahari (SS), dan curah hujan (RR) sebagai target prediksi.

#### 2. Persiapan data

Seluruh data yang diunduh dari <https://dataonline.bmkg.go.id> akan disatukan menjadi satu dataset untuk mempermudah dalam proses analisis dan pelatihan model. Tahap persiapan data meliputi beberapa Langkah yaitu

##### a. Pembersihan nilai hilang(*missing value*)

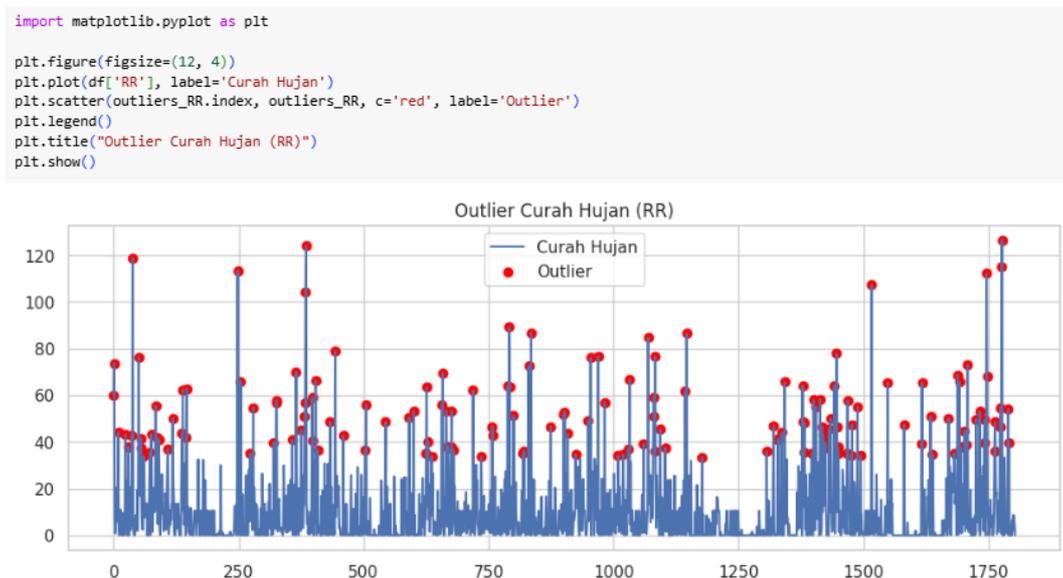
Pada Gambar 2 Menunjukkan proses penanganan *missing Value* pada data curah hujan menggunakan metode *imputasi* rata-rata (*mean*). Setiap kolom yang memiliki nilai hilang akan diisi dengan nilai rata-rata dari masing-masing kolom. Untuk kolom RH\_avg, nilai rata-rata akan dibulatkan terlebih dahulu sebelum digunakan sebagai pengganti.

```
df['Tavg'].fillna(df['Tavg'].mean(), inplace=True)
df['RH_avg'].fillna(int(df['RH_avg'].mean()), inplace=True)
df['RR'].fillna(df['RR'].mean(), inplace=True)
df['SS'].fillna(df['SS'].mean(), inplace=True)
```

Gambar 2. Penanganan *Missing Value*

##### b. Deteksi *outlier*

Gambar 3 menunjukkan deteksi *outlier* pada data yang digunakan dengan metode IQR. Proses ini dimulai dengan mengidentifikasi kuartil 1 (Q1), yang merupakan nilai 25% data terendah yaitu, dan kuartil 3 (Q3) yang merupakan nilai di mana 75% data terendah berada. IQR dihitung sebagai selisih antara Q3 dan Q1. Untuk menentukan batas *outlier*, batas dihitung dengan mengurangi 1.5 IQR dari Q1, dan batas atas dihitung dengan tambahan 1.5 kali IQR 3. Setiap titik data yang berada dibawah dan diatas akan dianggap sebagai *outlier*. Gambar 3 menunjukkan *outlier* pada kolom curah hujan memiliki banyak nilai ekstrem (*outlier*), terutama pada hari – hari dengan curah hujan ekstrem yang jarang terjadi. Untuk menangani *outlier* tersebut maka dilakukanlah normalisasi pada data dengan menggunakan *RocbusScaler*.



Gambar 3. Outlier Pada Kolom RR

c. Normalisasi data dengan *Robustscaler*

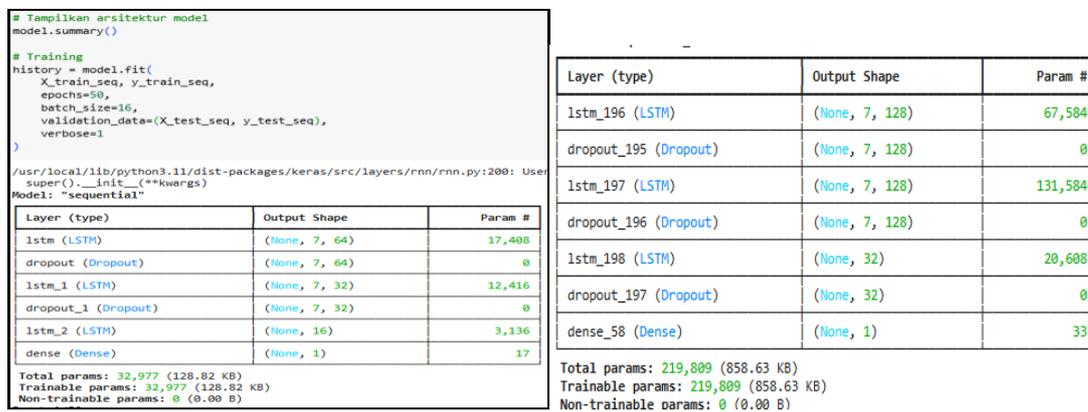
Pada data curah hujan yang digunakan nilai ekstrem yang valid, dinormalisasikan menggunakan *Robustscaler*, yang lebih tahan terhadap pengaruh outlier. Teknik ini mengskalakan data berdasarkan IQR sehingga distribusi data tetap representative meskipun ada anomali nilai yang tinggi [13].

Tabel 1. Hasil Normalisasi RobustScaler

Tavg	RH_avg	RR	SS
-2.0909	1.2222	4.2093	-0.7374
-1.3636	1.0000	5.2336	-0.7374
-0.4545	0.6666	0.2317	-0.7374
-0.5454	0.8888	-0.2168	-0.6927

### 3. Modeling

Model yang digunakan adalah LSTM, jenis jaringan saraf berulang (RNN) yang dirancang untuk mengatasi *vanishing gradient* dan mampu mempelajari pola jangka panjang dalam data deret waktu [14]. LSTM digunakan untuk menangkap pola curah hujan harian, dengan penyesuaian parameter seperti *epoch*, *batch size*, dan jumlah neuron untuk mengoptimalkan performa. Gambar 4 adalah arsitektur LSTM yang digunakan dalam penelitian ini, pada model ini memiliki beberapa lapisan layer. Lapisan layer bisa dianalogikan seperti tahapan proses berpikir. Dengan layer pertama yang *output/unit* 64 yang berupa *hiden state* dan *cell state*, *time stap* 7 yang berarti setiap langkah waktu 7 layer lstm menghasilkan output. Dan jumlah parameter pada layer pertama sebanyak 17.408.



(a) (b)  
Gambar 2. Arsitektur LSTM Non-Tuning (a) dan LSTM Tuning (b)

Pada Gambar 4 menunjukkan arsitektur LSTM yang setelah menggunakan parameter tuning dari optuna. Model terbagi menjadi empat layer, diselingi dengan dropout untuk mencegah overfitting.

#### 4. Evaluasi

Evaluasi model dilakukan dengan menggunakan dua metrik utama, yaitu *Mean Squared Error* (MSE) dan *Root Mean Squared Error* (RMSE). MSE mengukur rata-rata dari kuadrat selisih antara nilai prediksi dan aktual, memberikan penalti lebih besar terhadap kesalahan yang besar, serta sangat cocok digunakan sebagai fungsi kerugian dalam proses pelatihan model [15]. RMSE merupakan akar dari MSE dan memiliki satuan yang sama dengan target asli (dalam hal ini mm), sehingga lebih mudah diinterpretasikan secara praktis sebagai rata-rata kesalahan model. Kedua metrik ini secara bersama-sama memberikan gambaran yang komprehensif mengenai performa model baik selama pelatihan maupun pengujian.

a. *Root Mean Squared Error* (RMSE) adalah metrik evaluasi yang mengukur rata-rata kesalahan antara nilai aktual dengan nilai prediksi, sebagaimana ditunjukkan pada persamaan (1).

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

b. *Mean Squared Error* adalah mengukur rata-rata selisih kuadrat antara nilai prediksi dan nilai aktual. MSE memberikan penalti lebih besar untuk kesalahan yang besar karena selisihnya dikuadratkan, sebagaimana ditunjukkan pada persamaan (2).

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (2)$$

## HASIL DAN PEMBAHASAN

Hasil evaluasi pada Tabel 2 menunjukkan bahwa model LSTM dengan parameter tuning memiliki performa lebih baik dibandingkan model non-tuning. Nilai MSE dan RMSE lebih rendah pada semua skala, baik setelah normalisasi maupun pada skala asli.

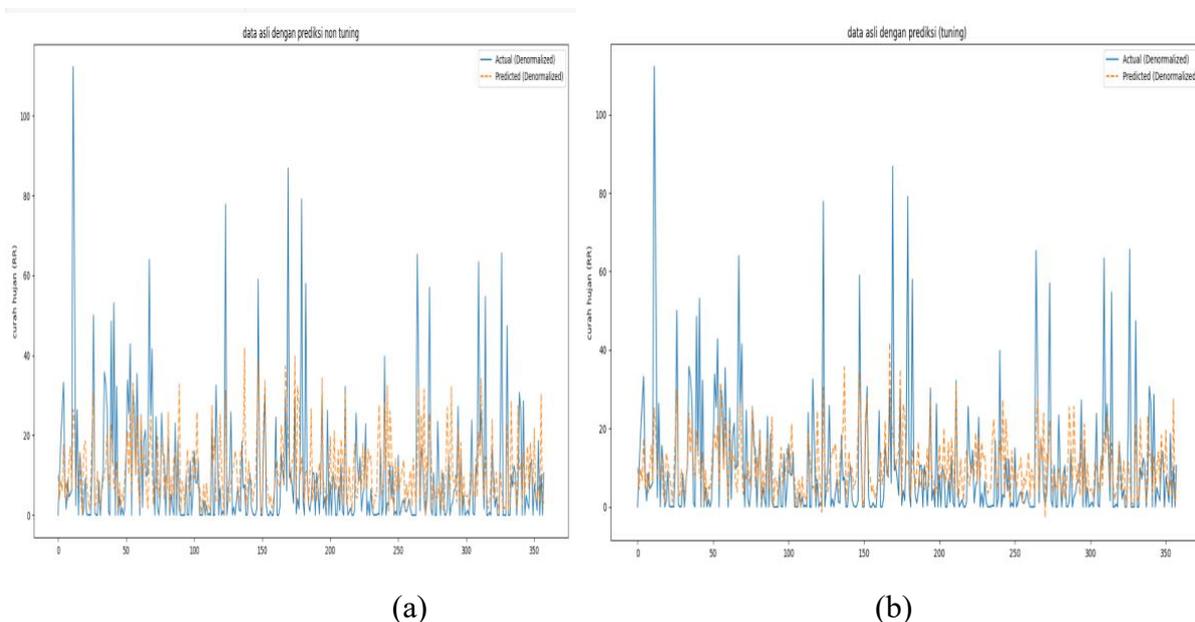
Penurunan nilai RMSE dari 14,2097 mm menjadi 13,7168 mm pada skala asli menunjukkan bahwa model hasil tuning lebih akurat dalam memprediksi curah hujan harian di Kabupaten Bogor.

Selain itu, nilai RMSE yang dicapai termasuk dalam kategori akurat dan dapat diterima menurut beberapa penelitian sebelumnya yang menggunakan metode serupa. Beberapa studi melaporkan bahwa nilai RMSE di bawah 15 mm sudah dianggap cukup baik dalam konteks prediksi curah hujan harian dengan algoritma LSTM [5][16][7]. Hal ini menandakan bahwa proses tuning hyperparameter menggunakan Optuna berhasil meningkatkan akurasi dan kemampuan generalisasi model terhadap data uji.

Tabel 2. Evaluasi Model LSTM

<b>Model</b>	<b>MSE (normalisasi)</b>	<b>MSE (skala asli)</b>	<b>RMSE (normalisasi)</b>	<b>RMSE (skala asli)</b>
LSTM non-tuning	1,1287	201,9152	1,0624	14,2097
LSTM parameter tuning	1,0518	188,1511	1,0256	13,7168

Gambar 5 menunjukkan perbandingan antara nilai aktual dan nilai prediksi curah hujan harian pada data uji. Terlihat bahwa model dengan parameter tuning (b) lebih mampu mengikuti pola fluktuasi curah hujan dibandingkan model tanpa tuning (a).



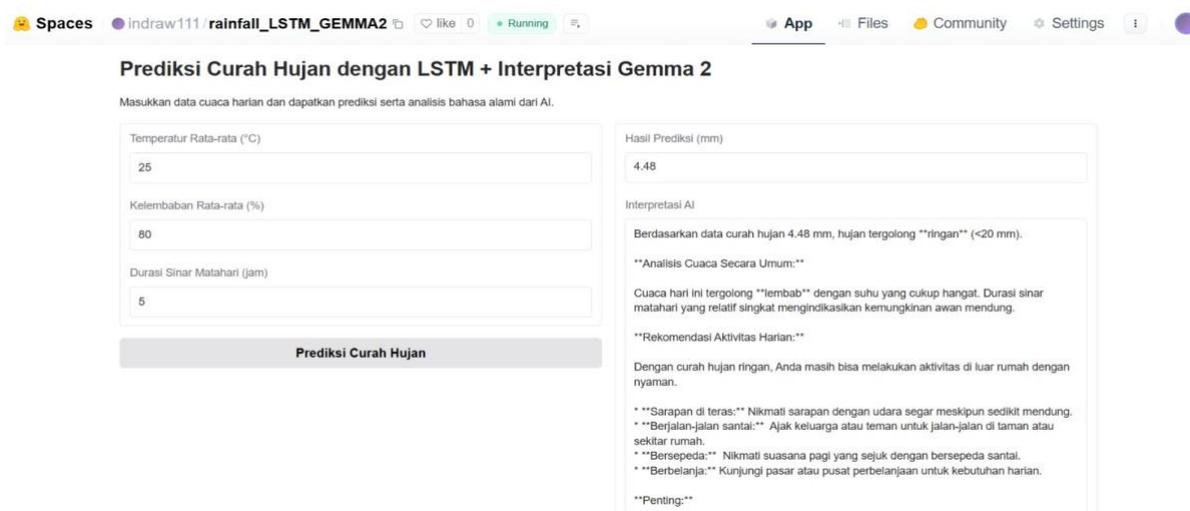
Gambar 5. Perbandingan Grafik Hasil Prediksi

Perbandingan nilai prediksi dari kedua model terhadap data aktual ditampilkan pada Tabel 3. Terlihat bahwa model dengan parameter tuning menghasilkan nilai prediksi yang lebih mendekati data aktual pada sebagian besar observasi.

Tabel 3. Perbandingan Hasil Prediksi

No	Data asli	LSTM non tuning	LSTM parameter tuning
1	0	10,035	10
2	10,1	4,159	4,534
3	18,5	8,761	9,721
4	26,4	6,140	6,781
5	6,2	7,148	6,252
6	33,2	17,566	17,014

Gambar 6 menunjukkan antarmuka hasil implementasi model prediksi curah hujan menggunakan dua pendekatan deployment berbeda namun sama-sama berbasis Hugging Face Inference API: Antarmuka awal menggunakan Gradio langsung yang dihosting di Hugging Face Spaces. Pendekatan ini digunakan untuk pengujian cepat dan interaktif, karena Gradio mendukung frontend yang otomatis dan responsif tanpa perlu membuat server backend manual.



Gambar 6. Hasil Deployment Gradio Dan Flask

Gambar 7 menampilkan antarmuka deployment akhir menggunakan Flask sebagai backend. Model tetap diakses melalui Hugging Face Inference API, tetapi komunikasi dilakukan melalui endpoint Flask, yang kemudian merender hasil ke halaman web yang dikembangkan sendiri. Pendekatan ini dipilih untuk kebutuhan deployment akhir yang lebih fleksibel dan dapat dikustomisasi, termasuk untuk integrasi interpretabilitas model (SHAP/interpretasi AI) ke dalam antarmuka pengguna.

**Prediksi Curah Hujan AI**

Temperatur Rata-rata (°C):  
25

Kelembaban Rata-rata (%):  
70

Durasi Sinar Matahari (jam):  
5

Mulai Prediksi

**Hasil Prediksi: 3.72 mm**

Hujan Ringan

**Interpretasi AI:**

Berdasarkan data cuaca, prediksi curah hujan 3.72 mm tergolong **hujan ringan** (1-5 mm).

**## Analisis Cuaca**

Cuaca hari ini akan cukup cerah dengan durasi sinar matahari 5 jam dan temperatur udara yang nyaman 25.0°C. Kelembaban udara sebesar 70% menandakan udara terasa lembab.

Karena hujan hanya ringan, kemungkinan besar tidak akan mengganggu aktivitas di luar ruangan.

**## Rekomendasi Aktivitas Harian**

Dengan cuaca seperti ini, cocok untuk melakukan aktivitas di luar ruangan seperti:

- Olahraga ringan:** jogging, sepeda, yoga di taman.
- Berjalan-jalan santai:** menikmati suasana taman, mengunjungi pasar tradisional, atau sekadar jalan-jalan sore.
- Piknik:** siapa yang tahu, mungkin akan ada pelangi cantik setelah hujan!

Namun tetap perlu membawa payung atau jas hujan tipis sebagai persiapan jika hujan tiba-tiba.

**Catatan:** Meskipun hujan tergolong ringan, tetap pantau kondisi cuaca terkini saat beraktivitas di luar ruangan untuk berjaga-jaga.

Gambar 7. Hasil Deployment Flask

## KESIMPULAN

Penelitian ini berhasil membangun model prediksi curah hujan harian di Kabupaten Bogor menggunakan arsitektur LSTM. Model yang dioptimasi dengan tuning hyperparameter menggunakan Optuna menunjukkan performa lebih baik dibandingkan model non-tuning, dengan penurunan nilai RMSE dari 14,21 mm menjadi 13,72 mm. Evaluasi dilakukan pada data yang telah dinormalisasi dan dikembalikan ke skala asli, dengan hasil yang konsisten menunjukkan keunggulan model tuning. Visualisasi prediksi juga memperlihatkan bahwa model tuning lebih akurat dalam mengikuti pola curah hujan aktual. Dua metode deployment digunakan: (1) Gradio via Hugging Face Spaces untuk pengujian cepat, dan (2) Flask dengan Hugging Face Inference API untuk kebutuhan deployment akhir. Keduanya berhasil menampilkan hasil prediksi secara real-time dan mendukung interpretasi model, menjadikan sistem ini siap diterapkan dalam praktik.

## UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih yang sebesar-besarnya kepada para reviewer atas masukan, koreksi, dan saran yang sangat berharga dalam penyempurnaan artikel ini. Ucapan terima kasih juga disampaikan kepada Universitas Bhayangkara Jakarta Raya yang telah memberikan dukungan penuh, baik dalam bentuk fasilitas, pendanaan, maupun motivasi

selama proses penelitian berlangsung. Dukungan tersebut sangat berarti dalam mewujudkan penelitian ini hingga tahap publikasi.

#### DAFTAR PUSTAKA

- [1] M. Ariska, S. Suhadi, S. Supari, M. Irfan, and I. Iskandar, “Annual and Interannual Rainfall Variability in Indonesia Using Empirical Orthogonal Function (EOF) Analysis and Its Response to Ocean-Atmosphere Dynamics,” *Jurnal Ilmu Fisika*, vol. 16, no. 2, pp. 151–165, 2024.
- [2] H. Setiawan, A. Wibowo, and S. Supriatna, “Pembuatan Peta Curah Hujan untuk Evaluasi Kesesuaian Rencana Tata Ruang Kawasan Hutan Kabupaten Bogor,” *Geomedia Majalah Ilmiah dan Informasi Kegeografian*, vol. 19, no. 2, pp. 113–121, Nov. 2021, doi: 10.21831/gm.v19i2.43227.
- [3] E. M. Lesik, H. L. Sianturi, A. S. Geru, and B. Bernandus, “Analisis Pola Hujan Dan Distribusi Hujan Berdasarkan Ketinggian Tempat Di Pulau Flores,” *Jurnal Fisika : Fisika Sains dan Aplikasinya*, vol. 5, no. 2, pp. 118–128, 2020, doi: 10.35508/fisa.v5i2.2451.
- [4] K. Sukajaya *et al.*, “Analisis Pasca Bencana Tanah Longsor 1 Januari 2020 Dan Evaluasi Penataan Kawasan Di,” *Jurnal Geografi Gea*, vol. 20, no. 2, pp. 197–213, 2020.
- [5] Y. Hendra, H. Mukhtar, B. Baidarus, and R. Hafsari, “Prediksi Curah hujan di Kota Pekanbaru Menggunakan LSTM (*Long Short Term Memory*),” *Journal of Software Engineering and Information Systems*, vol. 3, no. 2, pp. 74–81, 2021, doi: 10.37859/seis.v3i2.5606.
- [6] J. Badriyah, A. Fariza, and T. Harsono, “Prediksi Curah Hujan Menggunakan *Long Short Term Memory*,” *Jurnal Media Informatika Budidarma*, vol. 6, no. 3, p. 1297, 2022, doi: 10.30865/mib.v6i3.4008.
- [7] M. Rizki, S. Basuki, and Y. Azhar, “Implementasi Deep Learning Menggunakan Arsitektur *Long Short Term Memory* (LSTM) Untuk Prediksi Curah Hujan Kota Malang,” *Jurnal Repositor*, vol. 2, no. 3, pp. 331–338, Mar. 2020, doi: 10.22219/repositor.v2i3.470.
- [8] S. A. Jofipasi, Admi Salma, Dodi Vionanda, and Dina Fitria, “*Prediction Of Bogor City Rainfall Parameters Using Long Short Term Memory (LSTM)*,” *UNP Journal of Statistics and Data Science*, vol. 1, no. 5, pp. 434–440, Nov. 2023, doi: 10.24036/ujsds/vol1-iss5/110.
- [9] R. Farikhul Firdaus and I. V. Papatungan, “Prediksi Curah Hujan di Kota Bandung Menggunakan Metode *Long Short Term Memory*,” *Jurnal Penelitian Inovatif*, vol. 2, no. 3, pp. 453–460, 2022, doi: 10.54082/jupin.99.
- [10] A. Wijayanto, A. Sugiharto, and R. Santoso, “Identifikasi Dini Curah Hujan Berpotensi Banjir Menggunakan Algoritma *Long Short-Term Memory* (LSTM) Dan *Isolation Forest*,” *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 11, no. 3, pp. 637–646, 2024, doi: 10.25126/jtiik.938718.
- [11] D. H. Fadillah *et al.*, “Implementasi Lightgbm dan LLM Gemini pada Website Psychobot untuk Analisis Emosi Saat Bersosial Media,” *Jurnal Riset dan Aplikasi Mahasiswa Informatika (JRAMI)*, vol. 6, no. 01, pp. 224–233, Jan. 2025, doi:

10.30998/jrami.v6i01.13500.

- [12] F. Martinez-Plumed *et al.*, “CRISP-DM Twenty Years Later: From Data Mining Processes to Data Science Trajectories,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 33, no. 8, pp. 3048–3061, 2021, doi: 10.1109/TKDE.2019.2962680.
- [13] Scikit-learn developers, “*RobustScaler — scikit-learn 1.6.1 documentation.*” Accessed: May 06, 2025. [Online]. Available: <https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.RobustScaler.html>
- [14] S. M. Al-Selwi, M. F. Hassan, S. J. Abdulkadir, and A. Muneer, “*LSTM Inefficiency in Long-Term Dependencies Regression Problems,*” *Journal of Advanced Research in Applied Sciences and Engineering Technology*, vol. 30, no. 3, pp. 16–31, 2023, doi: 10.37934/araset.30.3.1631.
- [15] A. Geron, *Hands-on Machine Learning with Scikit-Learn, Keras & TensorFlow*. New York: O’Reilly Media, 2019.
- [16] W. Surta, K. T. Basuki, E. S. Negara, and Y. N. Kunang, “*Rainfall Prediction in Palembang City Using the GRU and LSTM methods,*” *Journal of Data Science*, vol. 4, no. 2, 2023, [Online]. Available: <http://eprints.intimal.edu.my/1730/>

