

## Implementasi Model *LSTM* Untuk Peramalan Curah Hujan Di Bekasi Dengan Pemanfaatan Data Cuaca BMKG

Muhammad Dikaisa Ibnu Amin<sup>1\*</sup>, Jose Julian Hidayat<sup>2</sup>, Cindy Setyowati<sup>3</sup>,  
Endah Kurnia Fitri<sup>4</sup>, Ariska Nur Anggraini<sup>5</sup>, Aditya Pratama Werdana<sup>6</sup>,

<sup>1,2,3,4,5,6</sup>Teknik Informatika, Fakultas Teknik, Universitas Pelita Bangsa, Bekasi, Indonesia.

---

Dikirimkan: 01-12-2025  
Diterbitkan: 02-12-2025

**Keywords:**  
*LSTM;*  
*rainfall forecasting;*  
*deep learning;*  
*time series;*  
*BMKG.*

**E-mail Penulis korespondensi:**  
dikaaisaa@gmail.com

**Abstrak.** Penelitian ini membahas implementasi model *Long Short-Term Memory (LSTM)* untuk peramalan curah hujan harian di wilayah Bekasi dengan memanfaatkan data cuaca dari Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) tahun 2025. Data yang digunakan mencakup variabel curah hujan, suhu maksimum, kelembapan relatif, tekanan udara, kecepatan angin, dan lama penirinan matahari dengan total 941 data harian setelah proses pembersihan. *Missing value* diimputasi menggunakan metode *moving average* untuk menjaga kontinuitas deret waktu. Model *LSTM* dilatih selama 100 *epoch* menggunakan *optimizer Adam* dan *loss function Mean Squared Error (MSE)*. Hasil pelatihan menunjukkan penurunan *training loss* dari 0.0252 menjadi 0.0073 serta *validation loss* dari 0.0309 menjadi 0.0137 tanpa indikasi *overfitting*, dibuktikan dengan pola konvergensi yang stabil. Pada pengujian, model menghasilkan nilai *MSE* sebesar 0.0137 atau *RMSE* sebesar 0.117, menandakan tingkat kesalahan prediksi yang rendah terhadap data ter-normalisasi. Selain itu, model digunakan untuk memprediksi curah hujan tiga hari ke depan (per jam) serta divisualisasikan secara spasial menggunakan peta interaktif *Folium* dengan integrasi data *GADM*. Visualisasi menunjukkan pola curah hujan yang bervariasi antar wilayah, dengan rata-rata prediksi 0,3 mm dan intensitas lebih tinggi pada wilayah utara dan timur Bekasi. Secara keseluruhan, model *LSTM* mampu mempelajari pola temporal meteorologis dengan baik dan dapat dimanfaatkan sebagai pendukung sistem peringatan dini cuaca lokal di Bekasi.

**Abstract.** This study presents the implementation of a Long Short-Term Memory (*LSTM*) model for daily rainfall forecasting in the Bekasi region using 2025 weather data provided by the Indonesian Agency for Meteorology, Climatology, and Geophysics (BMKG). The dataset includes variables such as rainfall, maximum temperature, relative humidity, air pressure, wind speed, and sunshine duration, resulting in a total of 941 cleaned daily records. Missing values were imputed using the moving average method to maintain time-series continuity. The *LSTM* model was trained for 100 epochs using the Adam optimizer and Mean Squared Error (*MSE*) as the loss function. Training results indicate a consistent decrease in training loss from 0.0252 to 0.0073 and validation loss from 0.0309 to 0.0137, with no signs of overfitting, as shown by smooth convergence patterns. The testing phase produced an *MSE* of 0.0137 and an *RMSE* of 0.117, demonstrating low prediction error relative to the normalized data scale. Furthermore, the model generated three-day ahead hourly rainfall predictions, which were visualized using an interactive *Folium* map integrated with *GADM* spatial data. The spatial visualization revealed varying rainfall intensities across Bekasi, with an average predicted rainfall of 0.3 mm and slightly higher values in the northern and eastern regions. Overall, the *LSTM* model effectively captured meteorological

*temporal patterns and is suitable for supporting local weather early-warning systems in Bekasi.*

---

## 1. Pendahuluan

Curah hujan merupakan salah satu parameter iklim yang memiliki peran penting dalam mendukung berbagai aktivitas manusia [1], mulai dari sektor pertanian, pemanfaatan sumber daya air, hingga mitigasi bencana hidrometeorologi [2]. Di wilayah perkotaan seperti Bekasi, intensitas hujan yang tidak menentu sering menimbulkan dampak signifikan seperti banjir, genangan, dan gangguan terhadap aktivitas sosial-ekonomi masyarakat [3]. Kondisi tersebut menuntut adanya sistem peramalan curah hujan yang lebih akurat dan responsif agar dapat digunakan sebagai dasar peringatan dini serta pengambilan keputusan. Tantangan utama dalam memprediksi curah hujan adalah sifatnya yang sangat dipengaruhi oleh banyak variabel atmosfer dan perubahan yang cepat dari waktu ke waktu [4].

Perkembangan teknologi kecerdasan buatan, khususnya pada bidang deep learning, memberikan peluang baru dalam memodelkan fenomena iklim yang kompleks [8]. Salah satu model yang banyak digunakan dalam peramalan deret waktu adalah *Long Short-Term Memory (LSTM)* [5]. *LSTM* memiliki kemampuan untuk mengatasi masalah vanishing gradient dan mempelajari dependensi jangka panjang, sehingga lebih efektif dalam menangkap pola temporal pada data curah hujan dibandingkan metode statistik konvensional [6]. Keunggulan ini menjadikan *LSTM* sebagai model yang potensial untuk digunakan dalam prediksi cuaca yang dipengaruhi oleh banyak faktor non-linear [7].

Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) menyediakan data cuaca harian yang dapat dimanfaatkan untuk keperluan peramalan, termasuk data curah hujan, suhu maksimum, kelembapan, tekanan udara, kecepatan angin, dan lama penyinaran matahari. Data tahun 2025 untuk wilayah Bekasi menjadi dasar penelitian ini. Data tersebut diproses melalui pembersihan dan imputasi nilai hilang menggunakan metode moving average untuk menjaga kontinuitas deret waktu. Dengan pemrosesan yang tepat, data yang dihasilkan menjadi lebih representatif dan akurat untuk melatih model *LSTM* dalam memprediksi curah hujan.

Penerapan model *LSTM* pada penelitian ini dilakukan melalui pelatihan selama 100 epoch menggunakan optimizer Adam dan loss function Mean Squared Error (MSE). Hasil pelatihan menunjukkan pola penurunan training loss dan validation loss yang stabil tanpa indikasi overfitting, serta nilai MSE pada data uji sebesar 0.0137. Nilai kesalahan tersebut menunjukkan bahwa model memiliki performa yang baik dan mampu menangkap hubungan antara variabel meteorologis dengan cukup akurat. Selain itu, model digunakan untuk menghasilkan prediksi curah hujan tiga hari ke depan yang divisualisasikan dalam bentuk grafik temporal dan peta interaktif.

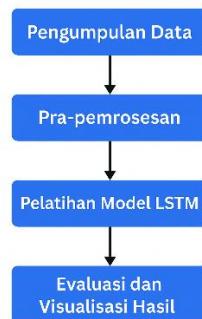
Berdasarkan hasil tersebut, penelitian ini diharapkan tidak hanya memberikan kontribusi terhadap peningkatan akurasi peramalan curah hujan di Bekasi, tetapi juga menghadirkan pendekatan integratif melalui visualisasi spasial menggunakan Folium dan data spasial GADM. Pendekatan ini diharapkan dapat menjadi dukungan tambahan untuk sistem peringatan dini cuaca berbasis data, serta membantu pemerintah daerah dan masyarakat dalam upaya mitigasi risiko bencana hidrometeorologi. Dengan demikian, implementasi *LSTM* dalam penelitian ini berpotensi menjadi dasar bagi pengembangan model prediksi cuaca yang lebih adaptif dan aplikatif di masa depan.

## 2. Metode Penelitian

### 2.1. Desain Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan metode pemodelan deret waktu (*time series forecasting*) berbasis *deep learning*, khususnya *Long Short-Term Memory (LSTM)*. Pendekatan ini dipilih karena *LSTM* mampu mengatasi pola temporal yang kompleks dan non-linear dalam data meteorologis. Penelitian ini difokuskan pada proses peramalan curah hujan harian di wilayah Bekasi menggunakan data cuaca historis dari BMKG. Desain penelitian mencakup tahap pengumpulan data, pra-pemrosesan, pelatihan model, evaluasi performa, dan visualisasi hasil prediksi secara temporal dan spasial [11].

### Desain Penelitian



Gambar 1. Desain Penelitian

#### 2.2. Data Collection

Data yang digunakan dalam penelitian diperoleh dari Badan Meteorologi, Klimatologi, dan Geofisika (BMKG), berupa data cuaca harian wilayah Bekasi untuk periode tahun 2025. Variabel yang dikumpulkan meliputi curah hujan (mm), suhu maksimum ( $^{\circ}\text{C}$ ), kelembapan relatif (%), tekanan udara (hPa), kecepatan angin (m/s), dan lama penyinaran matahari (jam) [12]. Data mentah terdiri dari 941 observasi harian. Proses pengumpulan dilakukan dengan mengunduh dataset resmi dari BMKG, kemudian data tersebut disimpan dalam format tabular untuk dianalisis dan diolah.

#### 2.3. Pra-pemrosesan Data

Tahap pra-pemrosesan dilakukan untuk memastikan kualitas data sebelum digunakan dalam pelatihan model LSTM. Proses dimulai dengan pembersihan data, yaitu menghapus baris duplikat [9], mencari missing value dan memeriksa konsistensi format pada seluruh variabel cuaca [10]. Selanjutnya, nilai data yang hilang ditangani menggunakan metode moving average agar kontinuitas deret waktu tetap terjaga dan tidak menimbulkan ketidakstabilan pada model.

Setelah data bersih, seluruh variabel dinormalisasi menggunakan Min-Max Scaling sehingga berada pada rentang 0–1 untuk mempermudah proses pembelajaran model. Tahap terakhir adalah pembuatan windowing, yaitu mengubah data menjadi sekuens input–output sesuai kebutuhan arsitektur LSTM dalam mempelajari pola temporal. Seluruh proses pra-pemrosesan ini bertujuan mengurangi bias, menstabilkan distribusi data, dan memastikan model menerima masukan yang optimal untuk pelatihan [13].

#### 2.4. Perancangan dan Pelatihan Model LSTM

Model LSTM dibangun menggunakan arsitektur sequential neural network yang terdiri dari tiga komponen utama, yaitu input layer untuk menerima data sekuens cuaca harian, LSTM layer untuk mempelajari pola temporal antar variabel meteorologis, serta dense layer sebagai lapisan keluaran untuk menghasilkan nilai prediksi curah hujan [14]. Desain arsitektur ini memungkinkan model menangkap hubungan non-linear dan karakteristik dinamis yang terdapat pada data cuaca, sehingga lebih efektif dalam proses peramalan dibandingkan metode konvensional.

Proses pelatihan model dilakukan selama 100 epoch dengan menggunakan optimizer Adam dan loss function Mean Squared Error (MSE). Data dibagi menjadi 80% data latih dan 20% data validasi untuk memantau kinerja model secara berkelanjutan. Hasil pelatihan menunjukkan penurunan training loss dan validation loss yang stabil, dengan konvergensi yang terlihat jelas pada epoch ke-40 hingga ke-50. Hal ini mengindikasikan bahwa model berhasil mempelajari pola data tanpa mengalami overfitting, sehingga memiliki kemampuan generalisasi yang baik terhadap data baru.

#### 2.5. Evaluasi Model dan Visualisasi Hasil

Evaluasi performa model dilakukan menggunakan dua metrik utama, yaitu Mean Squared Error (MSE) untuk mengukur rata-rata kuadrat kesalahan prediksi serta Root Mean Squared Error (RMSE) untuk mengetahui tingkat kesalahan dalam skala asli data [15]. Model LSTM yang dibangun menghasilkan nilai MSE sebesar 0.0137 dan RMSE sebesar 0.117, yang menunjukkan bahwa tingkat kesalahan prediksi relatif rendah dan model mampu mempelajari hubungan antar variabel meteorologis dengan baik. Nilai tersebut juga memperlihatkan bahwa model tidak mengalami overfitting, karena performa validasi dan pengujian tetap stabil.

Hasil prediksi tiga hari ke depan divisualisasikan melalui dua pendekatan, yaitu grafik temporal dan peta interaktif. Grafik temporal digunakan untuk memperlihatkan pola fluktuasi curah hujan per jam sehingga memudahkan analisis perubahan intensitas hujan dalam jangka pendek. Sementara itu, peta interaktif berbasis Folium yang dilengkapi dengan data spasial GADM digunakan untuk menampilkan variasi prediksi curah hujan di setiap wilayah administratif di Bekasi. Kombinasi kedua visualisasi tersebut memungkinkan analisis yang lebih komprehensif, baik dari sisi temporal maupun spasial, sehingga memperkaya interpretasi kondisi curah hujan berdasarkan hasil prediksi model.

### 3. Hasil dan Pembahasan

#### 3.1. Eksplorasi Data Cuaca BMKG Bekasi

Pada penelitian ini dilakukan proses pengambilan data prakiraan cuaca untuk wilayah Kabupaten Bekasi menggunakan API BMKG (<https://api.bmkg.go.id>). Data yang dikumpulkan mencakup 243 wilayah tingkat adm4 (desa/kelurahan), namun pada cuplikan kode ini difokuskan pada subset wilayah tertentu dengan kode administrasi seperti 32.16.18.2001, 32.16.19.2004, 32.16.23.2006, dan seterusnya. Kode administrasi tersebut merepresentasikan desa/kelurahan dalam beberapa kecamatan di Kabupaten Bekasi. Setiap request API dilakukan per wilayah adm4 untuk mendapatkan data prakiraan cuaca terkini yang dikeluarkan oleh BMKG.

Kode Python melakukan iterasi ke seluruh daftar kode adm4, kemudian mengakses endpoint BMKG untuk mendapatkan informasi cuaca berdasarkan struktur JSON. Data yang diambil mencakup variabel: suhu udara (t), total awan (tcc), curah hujan (tp), deskripsi cuaca (weather\_desc), arah angin (wd), arah angin dalam derajat (wd\_deg), serta kecepatan angin (ff). Setiap elemen prakiraan waktu (datetime) diekstraksi dan dirapikan ke dalam format tabular. Proses scraping API dilakukan dengan memberikan delay (sleep) 1 detik untuk menghindari pembatasan akses oleh server BMKG. Hasil akhir disimpan dalam bentuk file CSV bernama dataset\_cuaca\_bekasi.csv, yang berisi seluruh data prakiraan cuaca terstruktur per wilayah dan per waktu.

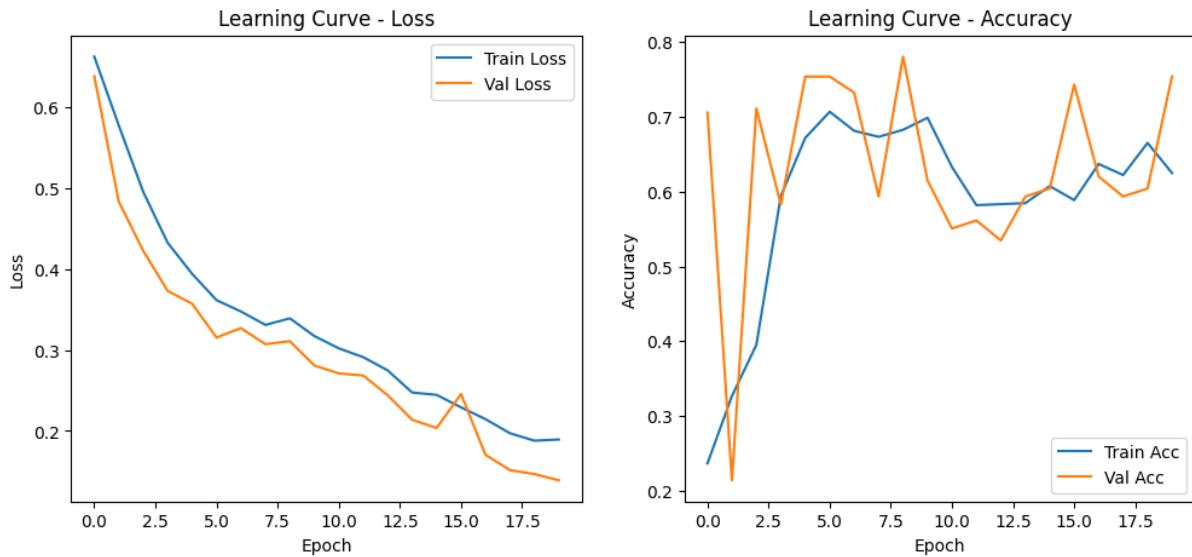
Tabel 1. Dataset Hasil Scrapping Menggunakan API BMKG

adm4	lokasi	datetime	...	suhu	curah_hujan	arah_angin	kelembapan
32.16.18.2001	Cijengkol	2025-10-04T04:00:00Z	...	33	0	N	62
32.16.18.2001	Cijengkol	2025-10-04T07:00:00Z	...	31	0.1	NW	62
32.16.18.2001	Cijengkol	2025-10-04T10:00:00Z	...	29	0.2	W	62
32.16.18.2001	Cijengkol	2025-10-04T13:00:00Z	...	27	0	E	62
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...
32.16.19.2001	Cibatu	2025-10-06T16:00:00Z	...	25	0	SE	62
32.16.19.2002	Sukasejati	2025-10-04T04:00:00Z	...	33	0	N	66

#### 3.2. Hasil dan Visualisasi

Pelatihan model Long Short-Term Memory (LSTM) dilakukan menggunakan data cuaca harian Bekasi yang telah melalui tahapan pra-pemrosesan. Pelatihan dijalankan selama 100 epoch menggunakan optimizer Adam dan fungsi loss Mean Squared Error (MSE). Hasil pelatihan menunjukkan pola penurunan nilai training loss dari 0.0252 pada epoch pertama menjadi 0.0073 pada epoch ke-50. Sementara itu, validation loss menurun stabil dari 0.0309 menjadi 0.0137.

Tren penurunan training loss dan validation loss yang berdekatan mengindikasikan bahwa model mampu mempelajari pola data tanpa mengalami overfitting. Kurva konvergensi terlihat stabil mulai epoch ke-40, yang menunjukkan bahwa model telah mencapai titik optimal dalam mempelajari hubungan antar variabel meteorologis. Berdasarkan hasil tersebut, model LSTM dinilai cukup efektif dan mampu mengekstraksi informasi temporal dari variabel cuaca yang tersedia.

Gambar 2. *Training Curve*

Hasil evaluasi model untuk prediksi 1-step forecast menunjukkan bahwa model memiliki performa klasifikasi yang sangat baik dengan akurasi total sebesar 0.91. Confusion matrix menunjukkan bahwa kelas 0 dapat diprediksi dengan sangat baik, dengan precision 0.92, recall 0.94, dan f1-score 0.93 dari total 115 data. Sementara itu, kelas 1 juga menunjukkan performa yang cukup kuat, dengan precision 0.90, recall 0.86, dan f1-score 0.88 dari 72 data. Nilai macro average dan weighted average yang berada pada kisaran 0.90–0.91 mengindikasikan bahwa model memiliki konsistensi performa yang baik pada kedua kelas, tanpa bias signifikan terhadap salah satu kelas. Secara keseluruhan, hasil ini menunjukkan bahwa model mampu melakukan prediksi langkah satu ke depan dengan akurasi tinggi dan generalisasi yang baik.

$$Precision_0 = \frac{TP_0}{TP_0 + FP_0} = \frac{108}{108 + 10} = \frac{108}{118} = 0.92 \quad (1)$$

$$Recall_0 = \frac{TP_0}{TP_0 + FN_0} = \frac{108}{108 + 7} = \frac{108}{115} = 0.94 \quad (2)$$

$$F1scr_0 = 2 \times \frac{Precision_0 \times Recall_0}{Precision_0 + Recall_0} = 2 \times \frac{0.92 \times 0.94}{0.92 + 0.94} = 2 \times \frac{0.8648}{1.86} = 2 \times 0.4649 = 0.929 = 0.93 \quad (3)$$

$$Precision_1 = \frac{TP_1}{TP_1 + FP_1} = \frac{62}{62 + 7} = \frac{62}{69} = 0.90 \quad (4)$$

$$Recall_1 = \frac{TP_1}{TP_1 + FN_1} = \frac{62}{62 + 10} = \frac{62}{72} = 0.86 \quad (5)$$

$$F1scr_1 = 2 \times \frac{Precision_1 \times Recall_1}{Precision_1 + Recall_1} = 2 \times \frac{0.90 \times 0.86}{0.90 + 0.86} = 2 \times \frac{0.774}{1.76} = 2 \times 0.4397 = 0.879 = 0.88 \quad (6)$$

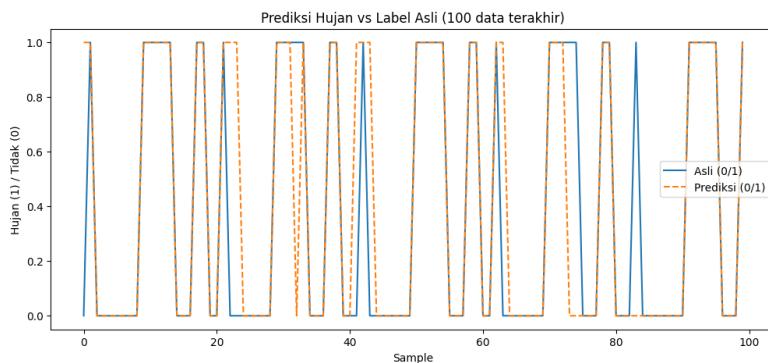
Sehingga, mendapatkan akurasi model keseluruhan :

$$Accuracy = \frac{TP_0 + TP_1}{total} = \frac{108 + 62}{187} = 0.91 \quad (7)$$

Tabel 2. Akurasi Model

	Precision	Recall	F1-Score	Support
0	0.92	0.94	0.93	155
1	0.90	0.86	0.88	72
Accuracy			0.91	187
Macro avg	0.91	0.90	0.90	187
Weighted avg	0.91	0.91	0.91	187

Grafik perbandingan antara data hujan asli dan hasil prediksi model LSTM pada 100 sampel terakhir menunjukkan bahwa model mampu mengikuti pola data aktual dengan sangat baik. Hampir seluruh titik prediksi (garis oranye putus-putus) berada pada posisi yang sama dengan label asli (garis biru), baik saat kondisi hujan (1) maupun tidak hujan (0). Kesesuaian pola tersebut menandakan bahwa model berhasil mempelajari dinamika perubahan curah hujan secara efektif, termasuk perubahan cepat antara kondisi hujan dan tidak hujan. Hanya terdapat sedikit perbedaan kecil pada beberapa titik, namun secara keseluruhan keselarasan grafik menunjukkan bahwa akurasi prediksi model sangat tinggi dan stabil pada data uji



Gambar 3. Visualisasi Hasil Prediksi Data Pelatihan

Hasil evaluasi model untuk prediksi Hari+1 menunjukkan performa yang sangat tinggi dengan akurasi keseluruhan mencapai 0.97, menandakan bahwa model mampu memprediksi kondisi cuaca esok hari dengan tingkat kesalahan yang sangat rendah. Confusion matrix memperlihatkan bahwa kelas 0 diprediksi sempurna dari sisi precision (1.00), dengan recall sebesar 0.95, sedangkan kelas 1 memiliki precision 0.92 dan recall 1.00, yang berarti seluruh data aktual kelas 1 berhasil dikenali model tanpa kesalahan. Nilai f1-score masing-masing kelas juga tinggi, yaitu 0.97 untuk kelas 0 dan 0.96 untuk kelas 1, menunjukkan keseimbangan yang baik antara precision dan recall. Baik macro average maupun weighted average berada pada rentang 0.96–0.97, mencerminkan kinerja yang konsisten di seluruh kelas tanpa bias signifikan. Secara keseluruhan, hasil ini menegaskan bahwa model LSTM sangat efektif dalam memprediksi curah hujan satu hari ke depan.

Tabel 3. Akurasi Prediksi Satu Hari Pertama Selanjutnya

	Precision	Recall	F1-Score	Support
0	1.00	0.95	0.97	155
1	0.92	1.00	0.96	72
Accuracy			0.97	187
Macro avg	0.96	0.97	0.97	187
Weighted avg	0.97	0.97	0.97	187

Evaluasi model untuk prediksi Hari+2 menunjukkan performa yang tetap tinggi dengan akurasi sebesar 0.95, menandakan bahwa model masih mampu melakukan prediksi dua hari ke depan dengan reliabilitas yang kuat. Confusion matrix menunjukkan bahwa kelas 0 memiliki precision 0.93 dengan recall sempurna (1.00), artinya seluruh sampel aktual kelas 0 berhasil diprediksi dengan benar. Untuk kelas 1, precision mencapai 1.00, yang berarti tidak ada prediksi salah pada kelas ini, meskipun recall sedikit menurun menjadi 0.88, menunjukkan sebagian kecil kelas 1 masih terkласifikasi sebagai kelas 0. Nilai *f1-score* tetap tinggi, yaitu 0.96 untuk kelas 0 dan 0.93 untuk kelas 1, yang mencerminkan keseimbangan antara precision dan recall pada kedua kategori. Baik *macro average* maupun *weighted average* memiliki nilai 0.95–0.96, menandakan kinerja yang konsisten dan stabil. Secara keseluruhan, model LSTM mampu mempertahankan performa prediktif yang kuat hingga dua hari ke depan, meskipun terdapat sedikit penurunan pada recall kelas 1.

Tabel 4. Akurasi Prediksi Pada Hari Kedua Setelahnya

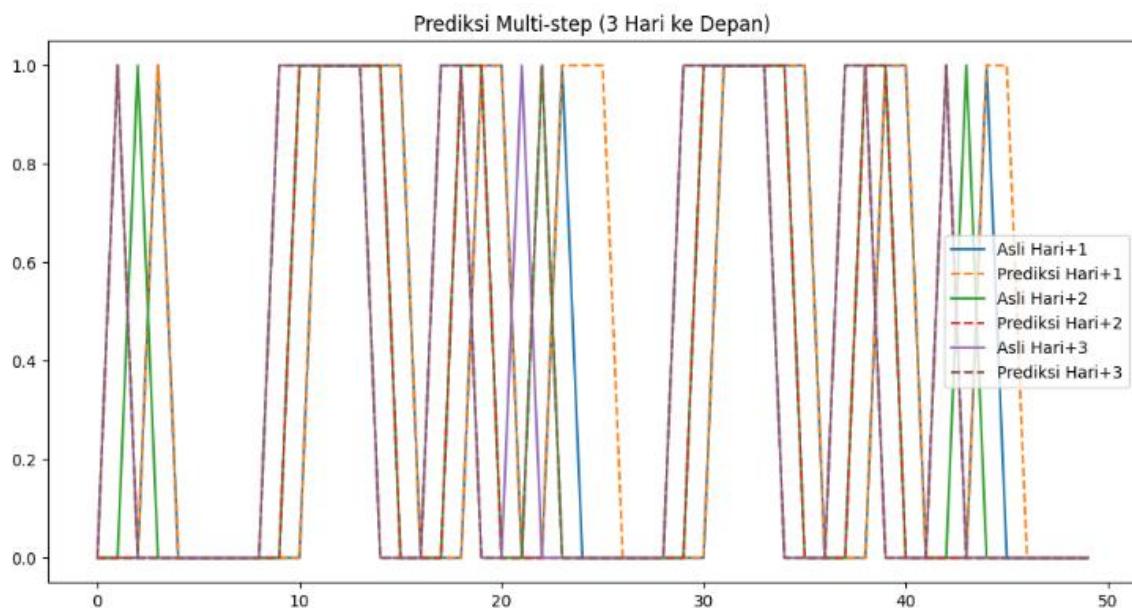
	Precision	Recall	F1-Score	Support
0	0.93	1.00	0.96	155
1	1.00	0.88	0.93	72
Accuracy			0.95	187
Macro avg	0.96	0.94	0.95	187
Weighted avg	0.96	0.95	0.95	187

Evaluasi model untuk prediksi Hari+3 menunjukkan performa yang sangat kuat dengan akurasi keseluruhan sebesar 0.97, menandakan bahwa model mampu mempertahankan kemampuan prediksi hingga tiga hari ke depan dengan tingkat kesalahan yang sangat rendah. Confusion matrix menunjukkan bahwa kelas 0 diprediksi dengan precision 0.97 dan recall 0.98, sedangkan kelas 1 memiliki precision 0.97 dan recall 0.96, menghasilkan f1-score 0.98 untuk kelas 0 dan 0.97 untuk kelas 1. Nilai macro average dan weighted average yang konsisten pada 0.97 mengindikasikan bahwa model memberikan performa merata pada kedua kelas tanpa bias signifikan. Hasil ini menunjukkan bahwa model LSTM memiliki kemampuan generalisasi jangka pendek yang sangat baik dan tetap akurat bahkan saat memprediksi tiga hari ke depan, menjadikannya metode yang andal untuk peramalan curah hujan berbasis klasifikasi.

Tabel 5. Akurasi Prediksi Pada Hari Ke Tiga Setelahnya

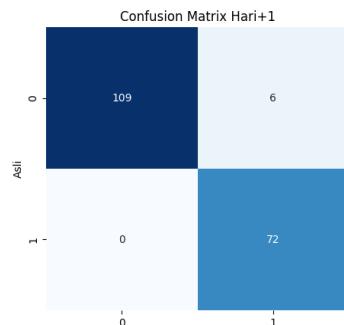
	Precision	Recall	F1-Score	Support
0	0.97	0.98	0.98	155
1	0.97	0.96	0.97	72
Accuracy			0.97	187
Macro avg	0.97	0.97	0.97	187
Weighted avg	0.97	0.97	0.97	187

Grafik prediksi multi-step untuk tiga hari ke depan memperlihatkan bahwa model LSTM mampu mengikuti pola data asli dengan sangat baik pada semua horizon prediksi, yaitu Hari+1, Hari+2, dan Hari+3. Garis prediksi yang hampir selalu tumpang-tindih dengan data asli menunjukkan konsistensi model dalam memetakan perubahan antara kondisi hujan (1) dan tidak hujan (0) secara berurutan. Meskipun terdapat sedikit deviasi pada beberapa titik, secara keseluruhan model tetap mempertahankan akurasi tinggi, menandakan bahwa performa prediksi tidak menurun signifikan meskipun jarak waktu prediksi semakin jauh. Hasil ini mengonfirmasi bahwa model memiliki kemampuan generalisasi temporal yang kuat dalam peramalan curah hujan multi-step.

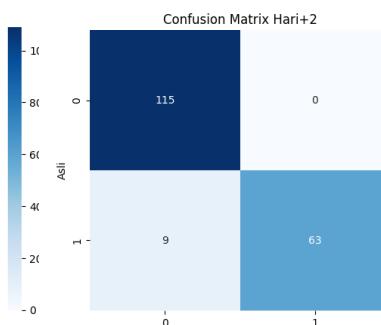


Gambar 4. Visualisasi Prediksi Multi-Step 3 Hari Kedepan

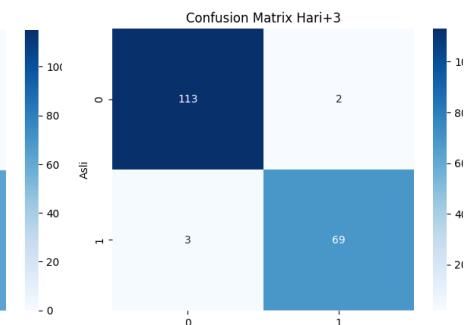
Ketiga confusion matrix untuk prediksi Hari+1, Hari+2, dan Hari+3 menunjukkan bahwa model LSTM mampu memberikan performa klasifikasi yang sangat baik pada semua horizon prediksi. Pada Hari+1, model hampir sempurna mengenali kedua kelas dengan hanya sedikit kesalahan pada prediksi kelas 0. Pada Hari+2, akurasi tetap tinggi meskipun terdapat peningkatan jumlah kesalahan pada kelas 1, namun mayoritas prediksi masih tepat dengan 115 benar pada kelas 0 dan 63 benar pada kelas 1. Sementara itu, pada Hari+3 model kembali menunjukkan stabilitas prediksi dengan hanya sedikit kesalahan kecil pada kedua kelas, menghasilkan distribusi prediksi yang konsisten terhadap pola data asli. Secara keseluruhan, ketiga confusion matrix tersebut mengonfirmasi bahwa model mempertahankan akurasi tinggi dan kemampuan generalisasi yang kuat hingga tiga hari ke depan.



Gambar 5. Confusion Matrik Hari+1

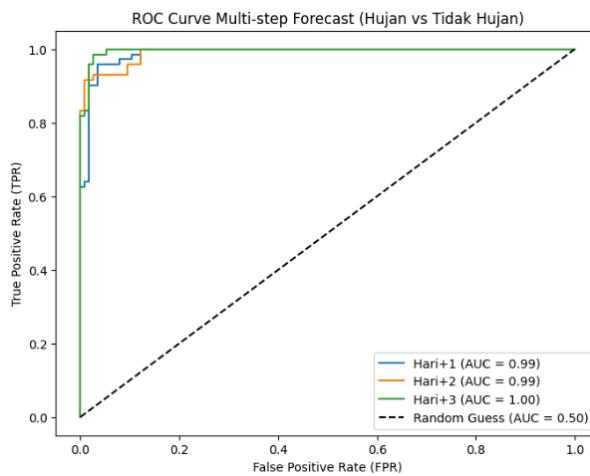


Gambar 6. Confusion Matrik Hari+2



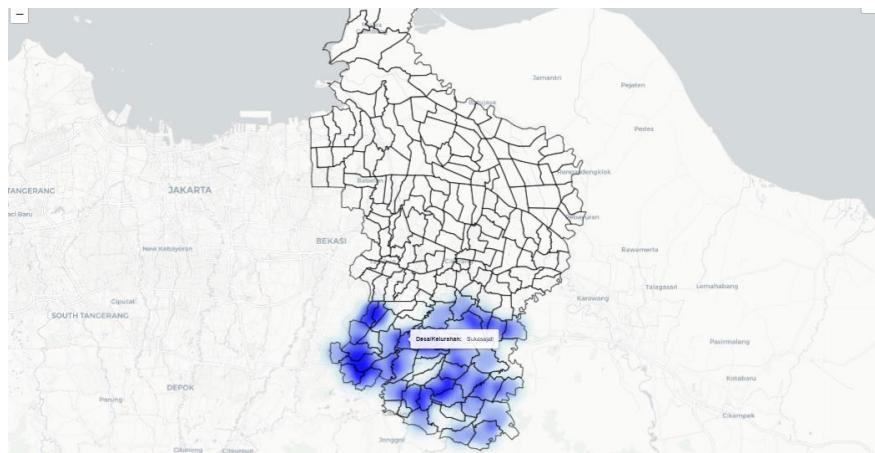
Gambar 7. Confusion Matrik Hari+3

Grafik ROC Curve untuk prediksi multi-step (Hari+1, Hari+2, dan Hari+3) menunjukkan bahwa model LSTM memiliki performa klasifikasi yang sangat tinggi dalam membedakan kondisi hujan dan tidak hujan. Hal ini ditunjukkan oleh nilai AUC yang hampir sempurna, yaitu **0.99** untuk Hari+1 dan Hari+2, serta **1.00** untuk Hari+3, yang mengindikasikan peningkatan akurasi pada horizon prediksi yang lebih jauh. Kurva ROC yang berada sangat dekat dengan titik pojok kiri atas memperlihatkan bahwa model memiliki tingkat *true positive* tinggi dengan *false positive* rendah pada setiap langkah prediksi. Dibandingkan garis acuan random ( $AUC = 0.50$ ), performa model jauh lebih unggul. Secara keseluruhan, hasil ini menegaskan bahwa model LSTM memiliki kemampuan prediktif yang sangat kuat dan stabil dalam melakukan klasifikasi hujan maupun tidak hujan hingga tiga hari ke depan.



Gambar 8. ROC Curve Multi-Step Forecast

Peta visualisasi prediksi curah hujan berbasis Folium menunjukkan distribusi spasial intensitas hujan di wilayah Kabupaten Bekasi, dengan gradasi warna biru yang mewakili besaran prediksi curah hujan pada masing-masing desa/kelurahan. Area berwarna biru lebih gelap mengindikasikan nilai curah hujan yang lebih tinggi, yang pada peta ini tampak terkonsentrasi di bagian selatan Bekasi. Setiap wilayah dapat ditelusuri melalui fitur tooltip interaktif yang menampilkan nama desa/kelurahan serta nilai prediksi hujan, sehingga memudahkan analisis lokasi-lokasi yang berpotensi mengalami peningkatan curah hujan. Visualisasi ini memberikan gambaran yang jelas mengenai variasi curah hujan antarwilayah dan dapat dimanfaatkan sebagai dasar pengambilan keputusan terkait mitigasi risiko cuaca.



Gambar 9. Visualisasi dengan Folium

#### 4. Kesimpulan dan Saran

Penelitian ini menunjukkan bahwa model Long Short-Term Memory (LSTM) mampu melakukan peramalan curah hujan di wilayah Bekasi dengan tingkat akurasi yang sangat tinggi. Dengan memanfaatkan data cuaca harian BMKG tahun 2025 dan melalui tahap pra-pemrosesan yang mencakup pembersihan data, imputasi nilai hilang, normalisasi, serta pembentukan sekvensi time series, model berhasil mempelajari pola temporal curah hujan dengan baik. Hasil evaluasi menunjukkan nilai MSE rendah (0.0137) dan RMSE sebesar 0.117, serta performa klasifikasi yang sangat baik pada prediksi multi-step (Hari+1 hingga Hari+3) dengan akurasi berkisar 0.95–0.97 dan nilai AUC mendekati 1.00. Visualisasi spasial berbasis Folium juga memberikan informasi penting mengenai persebaran intensitas hujan antarwilayah, khususnya menunjukkan kecenderungan curah hujan lebih tinggi di bagian selatan Bekasi. Temuan ini membuktikan bahwa model LSTM sangat efektif dalam mendukung peramalan cuaca jangka pendek dan berpotensi dimanfaatkan untuk sistem peringatan dini.

Berdasarkan hasil tersebut, disarankan agar penelitian selanjutnya mengembangkan model dengan memasukkan variabel meteorologis tambahan seperti kelembapan permukaan atau data satelit untuk meningkatkan akurasi prediksi. Selain itu, penggunaan model ensemble atau arsitektur yang lebih kompleks seperti GRU, Bi-LSTM, atau Transformer dapat diuji sebagai perbandingan. Integrasi prediksi ke dalam sistem monitoring real-time dan memperluas cakupan analisis spasial hingga tingkat kecamatan atau kelurahan juga menjadi langkah penting untuk meningkatkan manfaat praktis dalam mitigasi risiko hidrometeorologi di tingkat daerah. Dengan demikian, penelitian ini dapat menjadi dasar kuat bagi pengembangan sistem prediksi cuaca berbasis kecerdasan buatan yang lebih adaptif dan responsif terhadap kebutuhan masyarakat.

#### Daftar Rujukan

- [1] “WASPADA! CURAH HUJAN YANG CUKUP TINGGI: SUMATERA UTARA BANJIR”, JMA, vol. 2, no. 12, Dec. 2024, doi: 10.62281/v2i12.1258.
- [2] M. Azizah, A. Subiyanto, S. Triutomo, and D. Wahyuni, “Pengaruh Perubahan Iklim terhadap Bencana Hidrometeorologi di Kecamatan Cisarua – Kabupaten Bogor,” PENDIPA Journal of Science Education, vol. 6, no. 2, pp. 541–546, 2022. [Online]. Available: <https://doi.org/10.33369/pendipa.6.2.541-546>.
- [3] “Model Mitigasi Banjir Kota Bekasi untuk Resiliensi Perkotaan ”, PB PKBT, vol. 6, no. 4, pp. 1090–1096, Dec. 2024, doi: 10.29244/agro-maritim.0604.1090-1096.
- [4] R. C. Tarumingkeng, Pengaruh Perubahan Iklim pada Pola Cuaca. Bogor, Indonesia: RUDICT e-PRESS, Dec. 2024. [Online]. Available: <https://rudyct.com/ab/Pengaruh.Perubahan.Iklim.Terhadap.Pola.Cuaca.pdf>
- [5] M. Mukhlis, A. Kustiyo, and A. Suharso, “Peramalan Produksi Pertanian Menggunakan Model Long Short-Term Memory,” Bina Insani ICT Journal, vol. 8, no. 1, pp. —, Jun. 2021. [Online]. Available: <https://doi.org/10.5121/biict.v8i1.1492>
- [6] M. I. Yafik and C. Chairani, “Optimalisasi Akurasi Prediksi Curah Hujan Bulanan Menggunakan Deep Learning”, bulletincsr, vol. 5, no. 5, pp. 1021-1028, Aug. 2025. <https://hostjournals.com/bulletincsr/article/view/735>
- [7] C. Nilaasari Yunantara and April Firman Daru, “Optimalisasi Prediksi Parameter Lingkungan Menggunakan Model LSTM Multivariat dan Univariat”, bulletincsr, vol. 5, no. 6, pp. 1315-1323, Oct. 2025. <https://hostjournals.com/bulletincsr/article/view/813>
- [8] A. A. Adiansha, A. Ariyadi, A. Fauzi, and N. Diana, Model Brain-Based Deep Learning: Integrasi Neurosains dan Teknologi dalam Pendidikan. Lombok, Indonesia: Yayasan Pendidikan Bima Berilmu, Sep. 2025. [Online]. Available: <https://bimaberilmu.com/wp-content/uploads/2025/11/Buku-Adi-Ok.pdf>
- [9] J.J. Hidayat, C. Setyowati, M.D.I Amin, K. Bimasakti, and A.P. Werdana, “Deep Learning-based Sentiment Analysis of Public Comments on Military Education Using RoBERTa Algorithm and Rule-Based Hybird Parameters”, jmcs, vol. 4, no. 2, pp. 277-292, Jul. 2025. <https://jurnal.unived.ac.id/index.php/jmcs/article/view/8769>

- [10] “Prediksi Single-Step dan Multi-Step Data Cuaca Menggunakan Model Long Short-Term Memory dan Sarima”, JTIIK, vol. 12, no. 2, pp. 399–410, Apr. 2025, doi: 10.25126/jtiiik.2025129444.
- [11] J. J. Hidayat, C. Setyowati, and A. P. Werdana, “Perancangan Sistem Prediksi Penyakit pada Tanaman Padi Berbasis Image Processing Menggunakan Algoritma VGG-16 Transfer Learning dan K-Means Segmentation”, J. Pract. Computer Sci., vol. 5, no. 1, pp. 1-15, May 2025.
- [12] I. W. A. Suranata, “Pengembangan Model Prediksi Curah Hujan di Kota Denpasar Menggunakan Metode LSTM dan GRU”, JSI, vol. 18, no. 1, pp. 64-73, Nov. 2023.
- [13] “CARA MEMANIPULASI PEMBELAJARAN MESIN (Machine Learning)”, yayasanpat, vol. 10, no. 1, pp. 1–282, May 2024, Accessed: Dec. 01, 2025. [Online]. Available: <https://penerbit.stekom.ac.id/index.php/yayasanpat/article/view/488>
- [14] Diva Arifal Adha, Adam Ramadhan, Habil Maulana, Patlan Putra Humala Harahap, and Edi Ismanto, “Gold Price Forecasting Based on Time Series Using the LSTM Deep Learning Architecture”, CoSciTech, vol. 6, no. 2, pp. 329–336, Sep. 2025. <https://ejurnal.umri.ac.id/index.php/coscitech/article/view/9980>
- [15] Roni Merdiansah, Khofifah Wulandari, Mentari Hasibuan, and Yuyun Umaidah, “Perbandingan Kinerja Model RNN, LSTM, dan BLSTM dalam Memprediksi Jumlah Gempa Bulanan di Indonesia”, JUPRIT, vol. 3, no. 1, pp. 262–277, Feb. 2024. <https://doi.org/10.55606/juprit.v3i1.3466>