



# PREDIKSI CURAH HUJAN BULANAN DI MEDAN MENGGUNAKAN METODE LONG SHORT-TERM MEMORY (LSTM)

Dedek<sup>1)</sup>, Lailan Sofinah Harahap<sup>2)</sup>, Muhammad Rayhans Adrian<sup>3)</sup>

<sup>1)</sup>Ilmu Komputer, Fakultas Sains dan Teknologi, Universitas Islam Negeri Sumatera Utara, Medan, Indonesia  
Email: [dedek0701232058@uinsu.ac.id](mailto:dedek0701232058@uinsu.ac.id)

<sup>2)</sup>Ilmu Komputer, Fakultas Sains dan Teknologi, Universitas Islam Negeri Sumatera Utara, Medan, Indonesia  
Email: [lailansofinah@uinsu.ac.id](mailto:lailansofinah@uinsu.ac.id)

<sup>3)</sup>Ilmu Komputer, Fakultas Sains dan Teknologi, Universitas Islam Negeri Sumatera Utara, Medan, Indonesia  
Email: [rayhansadrian12@gmail.com](mailto:rayhansadrian12@gmail.com)

## Abstract

This study aims to predict monthly rainfall in Medan City using the Long Short-Term Memory (LSTM) method. The data utilized in this research comprises monthly rainfall figures and the number of rainy days for the 2015–2023 period, obtained from the Meteorology, Climatology, and Geophysics Agency (BMKG) Region I Medan via official publications of the Central Statistics Agency (BPS) of North Sumatra Province. The pre-processing stage involves data cleaning, normalization, and the construction of a time series dataset using a sliding window structure. The LSTM model was developed with two hidden layers and optimized using the Adam algorithm. Evaluation results indicate that the LSTM model effectively captures seasonal patterns and rainfall trends, as evidenced by a low Root Mean Square Error (RMSE) value. This study is expected to serve as a reference for hydrometeorological disaster mitigation in the Medan region.

**Keywords:** Rainfall, LSTM, Prediction, Time series, BMKG Region I Medan.

## Abstrak

Penelitian ini bertujuan untuk memprediksi curah hujan bulanan di Kota Medan menggunakan metode Long Short-Term Memory (LSTM). Data yang digunakan dalam penelitian ini merupakan data curah hujan bulanan dan jumlah hari hujan periode 2015–2023 yang diperoleh dari Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) Wilayah I Medan melalui publikasi resmi Badan Pusat Statistik (BPS) Provinsi Sumatera Utara. Proses pra-pemrosesan meliputi pembersihan data, normalisasi, dan pembentukan dataset time series dengan struktur sliding window. Model LSTM dibangun dengan dua lapisan tersembunyi (hidden layers) dan dioptimasi menggunakan algoritma Adam. Hasil evaluasi menunjukkan bahwa model LSTM mampu menangkap pola musiman dan tren curah hujan dengan baik, yang dibuktikan dengan nilai Root Mean Square Error (RMSE) yang rendah. Penelitian ini diharapkan dapat menjadi acuan dalam mitigasi bencana hidrometeorologi di wilayah Medan.

**Kata Kunci:** Curah hujan, LSTM, Prediksi, Deret waktu, BMKG Wilayah I Medan.



## PENDAHULUAN

Curah hujan merupakan parameter iklim yang krusial karena dampaknya yang luas terhadap berbagai sektor kehidupan, khususnya pertanian dan mitigasi bencana. Di wilayah Sumatera Utara, variabilitas curah hujan yang tinggi seringkali memicu risiko bencana hidrometeorologis seperti banjir dan tanah longsor. Selain itu, ketidakpastian pola hujan juga berdampak signifikan terhadap produktivitas komoditas perkebunan, seperti kopi, yang sangat bergantung pada kestabilan iklim untuk masa tanam dan panen yang optimal. Oleh karena itu, prediksi curah hujan yang akurat menjadi instrumen vital untuk sistem peringatan dini (early warning system) guna meminimalkan kerugian materiil dan korban jiwa.

Selama beberapa dekade, metode statistik konvensional seperti Autoregressive Integrated Moving Average (ARIMA) menjadi standar dalam peramalan cuaca. Namun, metode ini memiliki keterbatasan dalam menangkap pola non-linear yang kompleks dan dinamis dari data iklim jangka panjang. Seiring perkembangan teknologi, pendekatan Machine Learning, khususnya Deep Learning, mulai mendominasi penelitian meteorologi. Algoritma seperti Long Short-Term Memory (LSTM) dinilai lebih unggul dibandingkan metode statistik maupun model jaringan saraf tiruan biasa karena kemampuannya menyimpan informasi jangka panjang (long-term dependencies) dan mengatasi masalah vanishing gradient pada data deret waktu.

Berbagai penelitian terdahulu telah membuktikan efektivitas LSTM. Saputra et al. menunjukkan bahwa LSTM mampu memprediksi curah hujan di Kabupaten Badung dengan tingkat kesalahan yang sangat rendah. Hal serupa ditemukan oleh Nailah et al., di mana optimasi parameter pada model jaringan saraf tiruan mampu menghasilkan akurasi tinggi untuk peramalan cuaca di Denpasar. Penelitian lain oleh Suranata juga menegaskan bahwa LSTM memiliki performa yang lebih stabil dibandingkan Gated Recurrent Unit (GRU) dalam menangani fluktuasi data hujan.

Berdasarkan urgensi dan potensi tersebut, penelitian ini bertujuan untuk menerapkan metode LSTM dalam memprediksi curah hujan bulanan di Kota Medan menggunakan data historis 2015–2023.

## METODOLOGI PENELITIAN

### Data dan Sumber Data

Penelitian ini menggunakan pendekatan kuantitatif berbasis Deep Learning dengan algoritma Long Short-Term Memory (LSTM) varian multivariate untuk memprediksi curah hujan bulanan. Data utama berasal dari Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) Wilayah I Medan yang dipublikasikan melalui Badan Pusat Statistik (BPS) Provinsi Sumatera Utara. Dataset mencakup deret waktu selama sembilan tahun, mulai Januari 2015 hingga Desember 2023, dengan total 108 bulan observasi. Berbeda dengan penelitian sebelumnya yang hanya menggunakan satu variabel, penelitian ini mengintegrasikan dua variabel iklim, yaitu curah hujan bulanan (mm) dan jumlah hari hujan (hari), untuk meningkatkan sensitivitas model terhadap anomali cuaca ekstrem.

### Pra-pemrosesan Data dan Pembentukan Dataset

Sebelum diolah oleh model, serangkaian tahapan pra-pemrosesan dilakukan untuk menjamin integritas dan kualitas data. Proses dimulai dengan penggabungan sembilan berkas data tahunan menjadi satu kerangka data kronologis yang utuh. Tahap pembersihan data (data cleaning) dilakukan dengan mengeliminasi baris non-observasi dan menangani nilai hilang (missing values) agar kontinuitas deret waktu tetap terjaga.

Mengingat perbedaan skala antara curah hujan (ratusan mm) dan hari hujan (puluhan hari), teknik normalisasi Min-Max Scaling diterapkan untuk mentransformasi seluruh nilai ke rentang [0, 1]. Transformasi ini penting untuk mempercepat konvergensi gradien selama pelatihan model LSTM. Selanjutnya, data dibentuk ulang menjadi format supervised learning dengan struktur tiga dimensi [Sampel, Time Steps, Fitur]. Penelitian ini menggunakan jendela waktu (window size) 12 bulan ke

belakang sebagai input untuk memprediksi kondisi satu bulan ke depan. Dataset kemudian dipisahkan secara sekuensial menjadi 80% data latih (training set) untuk mempelajari pola dan 20% data uji (testing set) untuk validasi kinerja model pada data yang belum pernah dilihat.

### Arsitektur dan Pelatihan Model LSTM

Model dirancang untuk menangani input multivariate dengan lapisan input yang menerima dua fitur sekaligus (curah hujan dan hari hujan). Struktur jaringan saraf terdiri dari dua lapisan tersembunyi (hidden layers) LSTM dengan masing-masing 64 dan 32 unit neuron. Untuk mencegah model menghafal data latih (overfitting), lapisan Dropout dengan rasio 0,2 disisipkan di antara lapisan LSTM.

Lapisan output terdiri dari satu neuron Dense dengan fungsi aktivasi linear yang menghasilkan prediksi kontinu. Proses pelatihan menggunakan algoritma optimasi Adam dengan fungsi kerugian Mean Squared Error (MSE), dijalankan selama 150 epochs dengan ukuran batch 16, untuk memastikan model mampu mempelajari pola musiman secara mendalam.

### Evaluasi Kinerja Model

Kinerja model divalidasi menggunakan tiga metrik utama: Mean Absolute Error (MAE) untuk mengukur rata-rata penyimpangan absolut, Root Mean Square Error (RMSE) untuk mendeteksi sensitivitas terhadap kesalahan besar, dan Koefisien Determinasi ( $R^2$ ) untuk menilai seberapa baik model menjelaskan variabilitas data aktual. Penggunaan metrik ini sesuai standar evaluasi dalam penelitian klimatologi modern, sehingga memastikan objektivitas hasil prediksi.

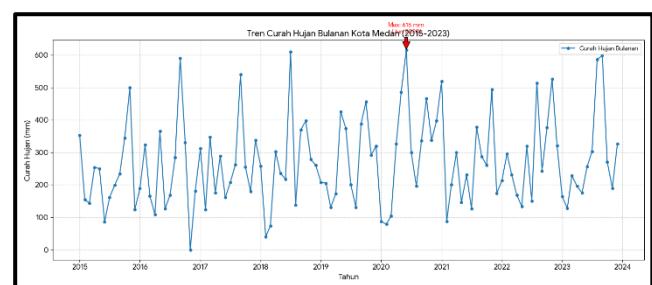
## HASIL DAN PEMBAHASAN

Analisis statistik dilakukan untuk memetakan karakteristik data klimatologis Kota Medan sebelum pemodelan. Data observasi mencakup sembilan tahun penuh, dari Januari 2015 hingga Desember 2023. Ringkasan statistik yang meliputi rata-rata curah hujan, nilai ekstrem, dan rata-rata jumlah hari hujan disajikan pada Tabel 3.1.

**Tabel 1.** Statistik Curah dan Hari Hujan Kota Medan (2015–2023)

Tahun	Curah Hujan		Hari Hujan	
	Rata-rata (mm)	Maksimum (mm)	Minimum (mm)	Rata-rata (hari)
2015	233.58	499	86	15.75
2016	235.83	590	0	16.75
2017	265.75	540	124	20.25
2018	265.08	610	40	34.00
2019	275.08	456	131	18.92
2020	310.75	615	79	19.42
2021	267.08	519	88	16.33
2022	291.08	526	134	19.17
2023	285.08	598	128	20.08

Berdasarkan Tabel 3.1, curah hujan di Kota Medan memiliki variabilitas yang tinggi, dengan rata-rata bulanan berkisar antara 233,58 mm hingga 310,75 mm. Tahun 2020 tercatat sebagai tahun terbaik dengan rata-rata curah hujan tertinggi (310,75 mm) dan nilai maksimum ekstrem mencapai 615 mm pada bulan Juni. Data juga menunjukkan korelasi antara intensitas hujan dan frekuensi kejadian hujan; tahun-tahun dengan curah hujan tinggi umumnya memiliki rata-rata hari hujan di atas 19 hari per bulan.



**Gambar 1.** Tren Curah Hujan Bulanan Kota Medan (2015-2023)

Untuk memahami dinamika musiman, tren curah hujan bulanan divisualisasikan pada Gambar 1, yang memperlihatkan pola iklim monsun ekuatorial dengan dua puncak musim hujan (bimodal). Lonjakan intensitas hujan secara konsisten terjadi pada periode akhir tahun (September hingga Desember), sementara anomali ekstrem terlihat pada tahun 2020 dan 2023, di mana curah hujan melampaui 500 mm, jauh di atas rata-rata historis.

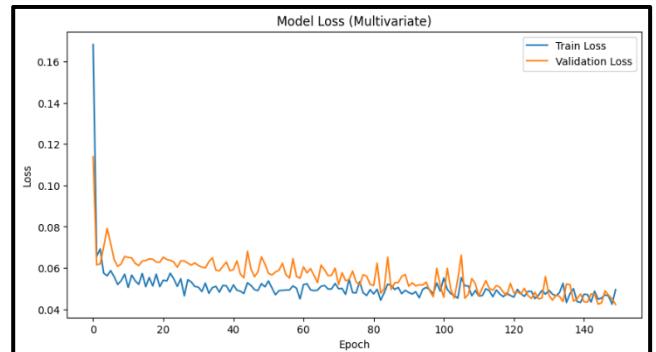


**Gambar 2.** Tren Curah Hujan Bulanan Kota Medan (2015-2023)

Selain pola bulanan, akumulasi curah hujan tahunan ditampilkan pada Gambar 2, yang memperlihatkan tren peningkatan total curah hujan dari tahun 2015 hingga puncaknya pada 2020, sebelum sedikit menurun dan stabil kembali pada 2022–2023. Tren ini mengindikasikan pengaruh fenomena iklim regional yang meningkatkan presipitasi di wilayah Indonesia bagian barat. Fluktuasi data historis ini menjadi landasan penting bagi model LSTM untuk mempelajari pola ketidakpastian cuaca.

### Hasil dan Evaluasi Model LSTM

Model LSTM dilatih menggunakan pendekatan multivariate dengan dua variabel input: curah hujan (mm) dan jumlah hari hujan (hari). Proses pelatihan berlangsung selama 150 epoch untuk memastikan model menangkap pola data secara optimal. Dinamika pembelajaran model divisualisasikan pada Gambar 3.



**Gambar 3.** Grafik Loss Pelatihan VS Validasi Model

Grafik pada Gambar 3. menunjukkan penurunan nilai loss yang tajam pada 10 epoch pertama, kemudian melandai dan stabil hingga akhir pelatihan. Celah yang sangat tipis antara Training Loss dan Validation Loss menunjukkan kemampuan generalisasi model yang baik, serta minim risiko overfitting maupun underfitting.

Evaluasi pada data uji menunjukkan bahwa model menghasilkan performa lebih baik dibandingkan pendekatan univariate. Nilai evaluasi yang diperoleh adalah RMSE sebesar 126,56 mm, MAE sebesar 92,92 mm, dan Koefisien Determinasi ( $R^2$ ) sebesar 0,26. Peningkatan nilai  $R^2$  menunjukkan bahwa penambahan variabel jumlah hari hujan memberikan konteks tambahan bagi model untuk mengenali intensitas hujan bulanan.

### Pembahasan

Hasil penelitian menegaskan bahwa penerapan metode Multivariate LSTM meningkatkan akurasi dibandingkan metode univariate. Dengan memasukkan variabel jumlah hari hujan, model tidak hanya belajar dari besaran curah hujan masa lalu, tetapi juga frekuensi kejadiannya, sehingga lebih sensitif terhadap pola bulan basah.

Nilai RMSE 126,56 mm dan  $R^2$  0,26 menunjukkan bahwa model mampu menjelaskan sekitar 26% variabilitas data curah hujan Kota Medan. Meskipun angka ini moderat, capaian ini signifikan mengingat sifat data curah hujan tropis yang stokastik dan penuh noise. Temuan ini sejalan dengan Tanjung et al., yang menyatakan bahwa pendekatan multivariat penting untuk meningkatkan performa model Deep Learning pada data klimatologi kompleks.



Keterbatasan model terlihat pada prediksi nilai ekstrem ( $>500$  mm), di mana model cenderung melakukan under-prediction. Hal ini umum terjadi pada fungsi kerugian Mean Squared Error (MSE) yang merata-ratakan kesalahan, sehingga "menghaluskan" lonjakan data yang tajam. Namun, pola prediksi yang mengikuti tren naik-turun data aktual menunjukkan bahwa LSTM efektif menangkap ketergantungan jangka panjang (long-term dependencies) sesuai karakteristik data deret waktu.

## KESIMPULAN

Penelitian ini berhasil mengembangkan model prediksi curah hujan bulanan di Kota Medan menggunakan metode Long Short-Term Memory (LSTM) yang dikembangkan oleh Hochreiter dan Schmidhuber. Dengan pendekatan multivariate, model ini terbukti mampu mempelajari pola ketergantungan jangka panjang (long-term dependencies) dari data deret waktu yang kompleks. Hasil analisis dan evaluasi menunjukkan bahwa penerapan arsitektur LSTM dengan dua variabel input, yaitu curah hujan dan jumlah hari hujan, menghasilkan kinerja yang cukup baik, dengan nilai RMSE sebesar 126,56 mm dan Koefisien Determinasi ( $R^2$ ) sebesar 0,26. Temuan ini mengonfirmasi keandalan metode Deep Learning dalam menangani volatilitas data cuaca tropis dibandingkan metode konvensional, sejalan dengan hasil penelitian Supriatna et al.

Analisis data historis 2015–2023 mengungkap bahwa Kota Medan memiliki karakteristik iklim yang sangat dinamis, dengan tren peningkatan curah hujan tahunan. Anomali ekstrem tercatat pada tahun 2020 dan 2023, ketika curah hujan bulanan melampaui 500 mm. Meskipun model mampu mengikuti tren musiman dengan baik, terdapat kecenderungan under-prediction pada kejadian curah hujan ekstrem, yang menunjukkan bahwa data historis curah hujan saja belum cukup untuk menangkap kompleksitas anomali cuaca ekstrem secara penuh.

Berdasarkan temuan ini, beberapa saran dapat diajukan untuk penelitian selanjutnya. Pertama, penambahan variabel atmosferik lain secara multivariate,

seperti suhu permukaan laut (Sea Surface Temperature), kelembaban udara, dan indeks ENSO (El Niño-Southern Oscillation), diharapkan dapat meningkatkan sensitivitas model terhadap anomali iklim ekstrem. Kedua, penerapan teknik optimasi hyperparameter otomatis, misalnya Bayesian Optimization, dapat membantu menemukan kombinasi lapisan dan neuron yang paling optimal tanpa uji coba manual. Ketiga, penggunaan resolusi data yang lebih detail, seperti data harian alih-alih bulanan, dapat memperbanyak jumlah sampel data latih, sehingga model dapat mempelajari pola mikro iklim dengan lebih presisi.

## DAFTAR PUSTAKA

- Arizka, A., & Lailan, S. H. (2024). Pemanfaatan jaringan saraf tiruan untuk prediksi curah hujan di Sumatera Utara. *MARS*, 2(6), 1–6.
- Dilla Sabine, F., Sitinjak, M. A., Ryacudu, J. T., Way Hui, D., Jatiagung, K., & Selatan, L. (2025). Analisis jumlah produksi kopi dan curah hujan di Provinsi Lampung menggunakan metode Seasonal Autoregressive Moving Average. *Jurnal Nasional Komputasi dan Teknologi Informasi (JNKTI)*, 8(2).
- Firdaus, R., & Mukhtar, H. (2023). Prediksi indeks harga produsen pertanian karet di Indonesia menggunakan metode LSTM. *Jurnal Fasilkom*, 13(01), 1–6.
- Hawali, M. L., & Walid, W. (2025). Pemilihan neuron LSTM dan LSTM Bayesian Optimization untuk prediksi curah hujan bulanan berbasis iklim. *Jurnal Fasilkom*, 15(2), 376–382.
- Hendra, Y., Mukhtar, H., & Hafsari, R. (2023). Prediksi curah hujan di Kota Pekanbaru menggunakan LSTM (Long Short Term Memory). *Journal of Software Engineering and Information System (SEIS)*, 74–81.
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- Lattifia, T., Buana, P. W., & Rusjanythi, N. K. D. (2022). Model prediksi cuaca menggunakan metode LSTM. *JITTER Jurnal Ilmiah Teknologi dan Komput*, 3(1), 994–1000.



- Nailah, F., Larasati, D. I., Siswanto, S., & Kalondeng, A. (2024). Optimasi metode jaringan saraf tiruan backpropagation untuk peramalan curah hujan bulanan di Kota Denpasar. *MATHunesa: Jurnal Ilmiah Matematika*, 12(1), 134–140.
- Saputra, B. D., Hiryanto, L., & Handhayani, T. (2023). Prediksi curah hujan di Kabupaten Badung, Bali menggunakan metode Long Short-Term Memory. *Jurnal Ilmu Komputer dan Sistem Informasi*, 11(2).
- Sunendar, N., Putro, H. P., & Hesananda, R. (2025). Prediksi penjualan aerosol menggunakan algoritma ARIMA, LSTM dan GRU. *INSOLOGI: Jurnal Sains dan Teknologi*, 4(1), 113–126.
- Supriatna, D., & Anggai, S. (2025). Analisis prediksi curah hujan di Kota Tangerang menggunakan metode LSTM dan GRU. *Jurnal Pustaka AI (Pusat Akses Kajian Teknologi Artificial Intelligence)*, 5(2), 119–131.
- Suranata, I. W. A. (2023). Pengembangan model prediksi curah hujan di Kota Denpasar menggunakan metode LSTM dan GRU. *Jurnal Sistem dan Informatika (JSI)*, 18(1), 64–73.
- Tanjung, R., Listiani, A., & Lestari, F. (2024). Prediksi multivariate time series parameter cuaca menggunakan Long Short-Term Memory (LSTM). In *Prosiding Seminar Nasional Sains Data* (pp. 445–456).
- Wayan, I., & Suranata, A. (n.d.). Pengembangan model prediksi curah hujan di Kota Denpasar menggunakan metode LSTM dan GRU.
- Wijayanto, A., Sugiharto, A., & Santoso, R. (2024). Identifikasi dini curah hujan berpotensi banjir menggunakan algoritma Long Short-Term Memory (LSTM) dan Isolation Forest. *Jurnal Teknologi Informasi dan Ilmu Komputer*, 11(3), 637–646.