

**Prediksi Curah Hujan Harian Menggunakan Long Short-Term Memory (LSTM)
pada Dataset *Rainfall in India* Kaggle**

Proposal ini disusun untuk memenuhi tugas mata kuliah Deep Learning



Nama Kelompok:

Fikri Armia Fahmi (2023071018)

Dava Ferdian Hadiputra (2023071021)

FAKULTAS TEKNOLOGI DAN DESAIN

UNIVERSITAS PEMBANGUNAN JAYA

SEPTEMBER, 2025

KATA PENGANTAR

Puji syukur kami panjatkan ke hadirat Tuhan Yang Maha Esa, karena atas rahmat dan karunia-Nya kami dapat menyusun proposal penelitian dengan judul *“Prediksi Curah Hujan Harian Menggunakan Long Short-Term Memory (LSTM) pada Dataset Rainfall in India Kaggle”* sebagai salah satu tugas pada mata kuliah **Deep Learning** yang diampu oleh **Dr. Ida Nurhaida, M.T.**

Proposal ini disusun dengan tujuan untuk memperdalam pemahaman kami mengenai penerapan metode deep learning dalam menyelesaikan permasalahan nyata, khususnya pada bidang meteorologi dan analisis deret waktu (*time series*). Pemilihan topik prediksi curah hujan didasarkan pada pentingnya informasi iklim yang akurat bagi sektor pertanian, pengelolaan sumber daya air, serta mitigasi bencana.

Kami menyadari bahwa proposal ini masih memiliki keterbatasan, baik dari segi isi maupun penyajiannya. Oleh karena itu, kami sangat mengharapkan kritik dan saran yang membangun dari dosen maupun rekan mahasiswa agar penelitian ini dapat disempurnakan di masa mendatang.

Akhir kata, kami mengucapkan terima kasih kepada **Dr. Ida Nurhaida, M.T.** selaku dosen pengampu, serta semua pihak yang telah memberikan dukungan dalam penyusunan proposal ini. Semoga proposal ini dapat memberikan manfaat bagi pengembangan ilmu pengetahuan, khususnya dalam penerapan deep learning untuk prediksi curah hujan.

Tangerang Selatan, 21 September 2025

Penyusun

DAFTAR ISI

KATA PENGANTAR	2
DAFTAR ISI	3
Table of Contents.....	3
BAB I	4
1.1 Latar Belakang	4
1.2 Rumusan Masalah	5
1.3 Tujuan Proyek	5
BAB II	6
2.1 Deep Learning & LSTM dalam Prediksi Curah Hujan.....	6
2.2 Variabel Input dalam Prediksi Curah Hujan	6
2.3 Arsitektur dan Varian LSTM	7
2.4 Pre-processing & Sliding Window	8
2.5 Evaluasi Model.....	8
BAB III	10
BAB IV	12
4.1 Akuisisi Data.....	12
4.2 Preprocessing dan Augmentasi Data	12
4.3 Pemodelan CNN	13
4.4 Optimisasi dan Regularisasi.....	13
4.5 Evaluasi Model.....	14
4.6 Analisis Hasil	14
BAB V	15
BAB VI	16
DAFTAR PUSTAKA	17

BAB I PENDAHULUAN

1.1 Latar Belakang

Curah hujan musiman di India memainkan peranan penting terhadap produktivitas sektor pertanian, manajemen air, serta ketahanan pangan dan ekonomi nasional. Variabilitas yang tinggi, terutama pada musim monsun, sering menyebabkan risiko tinggi seperti kekeringan hingga banjir, yang memerlukan prediksi curah hujan harian yang sangat akurat guna mendukung mitigasi bencana dan perencanaan sektor terkait.

Dalam beberapa tahun terakhir, metode tradisional seperti model statistik dan numerical weather prediction (NWP) terbukti memiliki keterbatasan dalam menangkap pola non-linear pada distribusi curah hujan yang kompleks. Studi perbandingan menunjukkan bahwa model deep learning, khususnya LSTM, jauh lebih efektif dalam mempelajari dependensi temporal jangka panjang (long-range dependencies) pada data meteorologi.

Misalnya, Patro dan Bartakke (2024) menerapkan algoritma LSTM untuk **prediksi curah hujan harian di wilayah** India dengan data stasiun otomatis IMD, serta variabel meteorologi seperti suhu, kelembapan, kecepatan angin, dan tekanan udara. Model ini berhasil menunjukkan performa tinggi dengan RMSE rendah dan Koefisien Determinasi (R^2) hingga mendekati 0.93. Selain itu, Zoremsanga dan Hussain (2024) memperkenalkan Bidirectional LSTM untuk prediksi curah hujan bulanan rata-rata seluruh India, yang menunjukkan peningkatan akurasi mulai dari 2% hingga 8% dibanding model LSTM dan RNN konvensional. Jurnal lain seperti Chen et al. (2022) juga membuktikan efektivitas LSTM berbasis fixed sliding window untuk peramalan distribusi curah hujan bulanan, menunjukkan bahwa LSTM dapat menangkap pola temporal yang stabil dari data yang panjang. Selain itu, penelitian oleh Sarkar (2023) di Barak River Basin, India menggunakan LSTM multivariate berhasil meramalkan curah hujan bulanan di tingkat regional dengan hasil yang menjanjikan. Dan kajian terbaru oleh Sham et al. (2025) menerapkan LSTM dengan indeks iklim (seperti ENSO dan IOD) untuk proyeksi curah hujan di daerah Barak, mendapat hasil akurasi tinggi.

Model deep learning baru seperti ARLSTM (Autoregressive LSTM) juga menunjukkan performa unggul di berbagai badan air India dengan median Nash–Sutcliffe Efficiency (NSE) sekitar 0.71, lebih baik dibanding model lain. Berdasarkan temuan-temuan ini, penggunaan metode LSTM baik unidirectional, bidirectional, maupun varian hybrid sangat potensial untuk diterapkan dalam prediksi curah hujan, termasuk pada skala harian, dengan tingkat akurasi yang lebih tinggi daripada pendekatan tradisional. Oleh karena itu, penelitian ini bertujuan membangun model LSTM yang mengolah data curah hujan (bisa interpolasi dari data bulanan seperti Rainfall in India di Kaggle) untuk menghasilkan prediksi harian yang akurat.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dipaparkan, maka rumusan masalah dalam penelitian/proyek ini adalah sebagai berikut:

1. Bagaimana cara mengolah dan memanfaatkan dataset curah hujan di India agar dapat digunakan untuk membangun model prediksi curah hujan harian?
2. Bagaimana penerapan metode *Long Short-Term Memory (LSTM)* dalam memprediksi curah hujan harian berdasarkan data historis curah hujan India?
3. Sejauh mana tingkat akurasi prediksi curah hujan harian yang dihasilkan oleh metode LSTM dibandingkan dengan pendekatan tradisional atau model *machine learning* lainnya?
4. Bagaimana hasil prediksi curah hujan harian tersebut dapat mendukung perencanaan sektor pertanian, pengelolaan sumber daya air, dan mitigasi bencana hidrometeorologi di India?

1.3 Tujuan Proyek

Tujuan Umum

Membangun model prediksi curah hujan harian dengan tingkat akurasi tinggi menggunakan metode Long Short-Term Memory (LSTM) berbasis data historis curah hujan India, sehingga dapat memberikan informasi yang lebih andal untuk mendukung pengambilan keputusan di sektor pertanian, sumber daya air, dan mitigasi bencana.

Tujuan Khusus

Secara lebih rinci, tujuan khusus proyek ini adalah:

1. Melakukan eksplorasi, pembersihan, dan transformasi dataset *Rainfall in India* agar dapat digunakan untuk prediksi curah hujan skala harian.
2. Mengembangkan model prediksi curah hujan berbasis LSTM yang mampu menangkap pola temporal dan musiman pada data historis.
3. Mengevaluasi performa model prediksi curah hujan LSTM dengan metrik evaluasi seperti RMSE, MAE, dan R^2 .
4. Membandingkan hasil prediksi LSTM dengan model tradisional atau metode *machine learning* lain (misalnya regresi, ARIMA, atau RNN standar).
5. Menyajikan hasil prediksi curah hujan harian yang dapat dimanfaatkan oleh pemerintah, lembaga meteorologi, dan masyarakat luas sebagai dasar perencanaan yang lebih tepat.

BAB II

LANDASAN TEORI

2.1 Deep Learning & LSTM dalam Prediksi Curah Hujan

Long Short-Term Memory (LSTM) merupakan salah satu varian dari *Recurrent Neural Network (RNN)* yang banyak digunakan pada data runtun waktu karena mampu mengatasi permasalahan *vanishing gradient* dan *exploding gradient*. Hal ini membuat LSTM lebih unggul dibandingkan RNN biasa dalam menyimpan informasi jangka panjang yang sering kali sangat penting dalam pola iklim maupun hidrologi (Chai et al., 2024).

Dalam konteks prediksi curah hujan, keunggulan LSTM sangat relevan karena pola hujan tidak hanya dipengaruhi oleh data harian, tetapi juga interaksi jangka panjang antar-musim, fenomena global seperti El Niño dan La Niña, serta faktor geografis. Model tradisional seperti ARIMA atau regresi linier sering gagal menangkap hubungan non-linear tersebut, sedangkan LSTM dapat mempelajarinya secara bertahap melalui mekanisme *gate* (Endalie et al., 2022).

Tabel 2.1. Contoh Studi Penggunaan LSTM dalam Prediksi Curah Hujan

Peneliti (Tahun)	Pendekatan/Model	Lokasi/Objek Data	Hasil Utama
Patro & Bartakke (2024)	LSTM (data harian)	India	$R^2 \approx 0,93$ - jauh lebih baik dibanding model klasik
Chai et al. (2024)	FA-LSTM (Firefly untuk optimasi)	Data prediksi hujan	$R^2 \approx 0,97$ - unggul atas GRU, RNN standar, Random Forest, dan SVM
Sarmah et al. (2023)	Stacked LSTM	Timur Laut India	MAE lebih rendah dibanding regresi linear berganda, Random Forest, XGBoost

2.2 Variabel Input dalam Prediksi Curah Hujan

Prediksi curah hujan umumnya tidak cukup hanya menggunakan data hujan sebelumnya. Faktor iklim lain seperti suhu, kelembapan, tekanan atmosfer, dan indeks iklim global sangat berpengaruh. Sebagai contoh, fenomena ENSO (El Niño–Southern Oscillation) diketahui dapat meningkatkan atau menurunkan intensitas hujan di wilayah tropis, sehingga penambahan variabel ENSO/IOD dapat membantu model LSTM mengenali pola yang lebih kompleks (Park et al., 2025).

Variabel umum yang digunakan dalam penelitian hidrologi modern:

1. Suhu udara (*temperature*) → mempengaruhi penguapan.
2. Kelembapan relatif (*humidity*) → indikator jumlah uap air di atmosfer.

3. Tekanan udara (*atmospheric pressure*) → berkaitan dengan sistem cuaca.
4. Kecepatan angin (*wind speed*) → berperan dalam distribusi awan hujan.
5. Indeks iklim global (ENSO, IOD, PDO) → memberi gambaran fenomena skala besar.

Tabel 2.2. Perbandingan model dengan dan tanpa variabel tambahan

Studi	Dataset	Model	Input	Hasil
Park et al. (2025)	Basin di Korea	LSTM	Curah hujan saja	RMSE lebih tinggi
Park et al. (2025)	Basin di Korea	LSTM + indeks iklim	Curah hujan + suhu + ENSO/IOD	RMSE lebih rendah, R^2 lebih tinggi

2.3 Arsitektur dan Varian LSTM

Arsitektur Long Short-Term Memory (LSTM) dibangun atas dasar *Recurrent Neural Network (RNN)*, namun memiliki struktur internal yang lebih kompleks untuk mengatasi kelemahan utama RNN, yaitu masalah *vanishing gradient*. LSTM menggunakan unit memori (*memory cell*) yang berfungsi menyimpan informasi dalam jangka panjang, dengan mekanisme pengendali berupa tiga *gate*, yaitu *input gate*, *forget gate*, dan *output gate* (Sarmah et al., 2023).

1. **Input Gate** berfungsi mengatur seberapa besar informasi baru dari *input* saat ini dimasukkan ke dalam *cell state*.
2. **Forget Gate** bertugas memutuskan informasi apa yang perlu dihapus dari memori, sehingga model tidak menyimpan data yang tidak relevan.
3. **Output Gate** menentukan informasi apa yang dikeluarkan dari *cell state* untuk diteruskan ke lapisan berikutnya.

Dengan mekanisme ini, LSTM mampu mempertahankan hubungan jangka panjang dalam data runtun waktu, sehingga relevan untuk prediksi curah hujan yang dipengaruhi oleh pola musiman maupun fenomena iklim global.

Selain LSTM standar, terdapat beberapa varian yang telah banyak digunakan pada penelitian meteorologi:

- **Stacked LSTM:** terdiri dari beberapa lapisan LSTM yang ditumpuk, dengan tujuan meningkatkan kapasitas representasi model. Varian ini terbukti efektif dalam mempelajari pola kompleks, namun memerlukan sumber daya komputasi yang lebih besar (Sarmah et al., 2023).

- **Bidirectional LSTM:** memproses data secara maju (*forward*) dan mundur (*backward*), sehingga model memperoleh konteks temporal yang lebih lengkap. Hal ini sangat berguna dalam prediksi curah hujan yang bergantung pada pola historis dan proyeksi ke depan (Park et al., 2025).
- **Hybrid LSTM:** menggabungkan LSTM dengan metode lain, misalnya algoritma optimisasi Firefly atau model CNN. Tujuannya adalah meningkatkan akurasi prediksi dengan mengoptimalkan bobot atau mengekstraksi fitur tambahan dari data input (Chai et al., 2024).

2.4 Pre-processing & Sliding Window

Tahapan *pre-processing* merupakan komponen penting sebelum data curah hujan digunakan untuk melatih model LSTM. Data meteorologi pada umumnya memiliki keterbatasan seperti nilai hilang (*missing values*), skala yang berbeda antar-variabel, serta keberadaan *outlier*. Jika tidak ditangani dengan benar, hal ini dapat menurunkan performa model secara signifikan (Chen et al., 2022).

Prosedur *pre-processing* yang umum dilakukan meliputi:

1. **Penanganan nilai hilang (*missing values*):** dilakukan dengan interpolasi linier, metode statistik, atau pendekatan berbasis model.
2. **Normalisasi atau standardisasi data:** bertujuan menyamakan skala antar-variabel agar proses pelatihan lebih stabil.
3. **Pemisahan data latih dan uji (*train-test split*):** untuk memastikan model dapat dievaluasi secara obyektif.
4. **Penerapan metode *sliding window*:** mengubah data runtun waktu menjadi pasangan input-output dengan jendela waktu tertentu.

Metode *sliding window* secara khusus digunakan untuk membantu LSTM menangkap dinamika temporal yang lebih panjang. Sebagai contoh, data curah hujan 30 hari terakhir dapat digunakan untuk memprediksi curah hujan pada hari ke-31. Dibandingkan pendekatan tanpa *sliding window*, metode ini terbukti lebih mampu mengidentifikasi pola berulang dalam data iklim, meskipun menambah kompleksitas perhitungan (Chen et al., 2022).

2.5 Evaluasi Model

Evaluasi model prediksi curah hujan harus dilakukan secara komprehensif dengan menggunakan metrik yang sesuai untuk data runtun waktu. Reliabilitas suatu model tidak dapat hanya ditentukan oleh akurasi sederhana, melainkan memerlukan analisis kuantitatif terhadap kesalahan prediksi dan kemampuan model menjelaskan variasi data (Park et al., 2025).

Metrik yang sering digunakan antara lain:

1. **Mean Absolute Error (MAE):** menghitung rata-rata kesalahan absolut antara nilai

prediksi dan observasi, sehingga mudah diinterpretasikan.

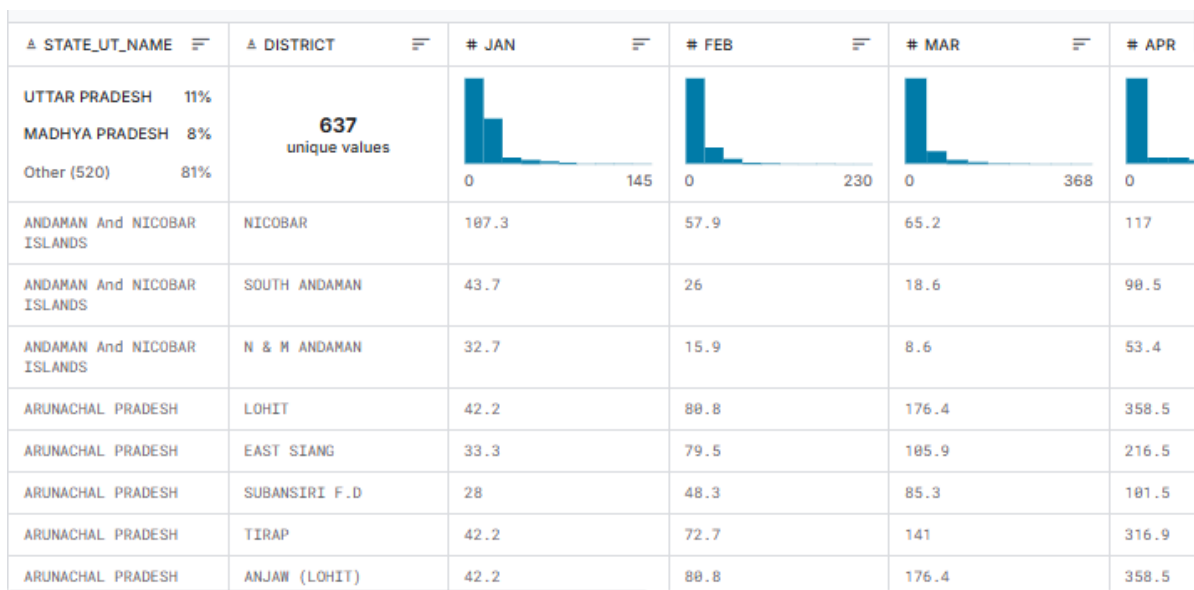
2. **Mean Squared Error (MSE)**: menghitung rata-rata kuadrat kesalahan, dengan sifat memperbesar pengaruh kesalahan besar.
3. **Root Mean Squared Error (RMSE)**: akar dari MSE, banyak digunakan karena memiliki satuan yang sama dengan variabel asli.
4. **Koefisien Determinasi (R^2)**: menunjukkan proporsi variasi data yang dapat dijelaskan oleh model. Semakin mendekati 1, semakin baik model tersebut.
5. **Nash–Sutcliffe Efficiency (NSE)**: metrik yang sering digunakan dalam hidrologi, menilai kemampuan model dalam memprediksi data dibandingkan dengan rata-rata historis.

Apabila prediksi dilakukan dalam bentuk klasifikasi, misalnya hujan atau tidak hujan, evaluasi dapat menggunakan **confusion matrix** dengan turunan metrik seperti *precision*, *recall*, dan *F1-score*. Dengan demikian, evaluasi model menjadi lebih menyeluruh dan dapat mencerminkan performa pada berbagai skenario penggunaan (Endaliev et al., 2022).

BAB III

DATASET

Dataset yang digunakan dalam penelitian ini adalah *Rainfall in India Dataset* yang tersedia di platform Kaggle dan dikompilasi oleh Anand (2019). Dataset ini berisi data curah hujan bulanan dari tahun 1901 hingga 2017, yang mencakup lebih dari 36 wilayah meteorologi di India. Secara keseluruhan, dataset terdiri dari sekitar 1.400 baris data dengan 124 kolom yang mewakili bulan serta rata-rata curah hujan untuk masing-masing wilayah. Data yang tersedia bersifat numerik dan direpresentasikan dalam satuan *milimeter (mm)*.



Gambar 3. Dataset Rainfall in India dari Kaggle

Distribusi data dalam dataset ini mencerminkan variasi curah hujan yang signifikan antarwilayah maupun antarmusim, mengingat India memiliki kondisi iklim yang beragam serta dipengaruhi oleh fenomena monsun. Hal ini menimbulkan tantangan dalam pemodelan, karena pola hujan yang tidak merata dapat memengaruhi kemampuan model dalam melakukan generalisasi. Selain itu, data yang bersifat bulanan perlu dilakukan transformasi ke dalam bentuk harian menggunakan pendekatan interpolasi atau metode statistik lainnya agar lebih sesuai dengan kebutuhan prediksi curah hujan harian.

Pemilihan dataset ini didasarkan pada beberapa pertimbangan. Pertama, dataset ini bersifat terbuka dan dapat diakses publik, sehingga mendukung prinsip keterbukaan ilmu pengetahuan dan memungkinkan replikasi penelitian oleh pihak lain. Kedua, dataset ini memiliki rentang waktu yang sangat panjang (lebih dari satu abad), sehingga dapat memberikan informasi historis yang kaya dan relevan untuk melatih model deep learning dalam menangkap pola jangka panjang. Ketiga, dataset ini telah banyak digunakan dalam penelitian terdahulu yang membahas prediksi iklim dan analisis deret waktu, sehingga hasil penelitian ini dapat dibandingkan dengan studi sebelumnya sebagai bentuk benchmarking.

Selain itu, relevansi dataset ini sangat tinggi karena curah hujan merupakan faktor kunci dalam sektor pertanian, pengelolaan sumber daya air, serta mitigasi bencana banjir dan kekeringan di India. Dengan demikian, pemilihan *Rainfall in India Dataset* tidak hanya memenuhi aspek teknis dalam pembelajaran mesin, tetapi juga memiliki relevansi sosial dan praktis yang kuat terhadap tantangan nyata dalam pengelolaan lingkungan dan pembangunan berkelanjutan di kawasan tersebut.

BAB IV

METODOLOGI

4.1 Akuisisi Data

Dataset utama penelitian ini adalah **Rainfall in India** yang tersedia di platform Kaggle dan dikompilasi oleh Anand (2019). Dataset berisi data curah hujan bulanan dari tahun 1901 hingga 2017 pada 36 wilayah meteorologi di India

- Dataset diunduh dalam format CSV.
- Data kemudian diekstraksi menjadi tabel numerik dengan kolom mewakili wilayah dan baris mewakili waktu.
- Dataset dibagi menjadi tiga subset: **data latih (70%)**, **data validasi (15%)**, dan **data uji (15%)**.

Tabel 4.1. Pembagian Dataset

Jenis Data	Persentase	Jumlah Periode (\pm)
Data Latih	70%	81 tahun
Data Validasi	15%	17 tahun
Data Uji	15%	19 tahun

4.2 Preprocessing dan Augmentasi Data

Agar data sesuai untuk pemodelan LSTM, dilakukan preprocessing sebagai berikut:

1. Transformasi bulanan ke harian

Karena dataset asli berupa data bulanan, dilakukan interpolasi linier untuk mengubahnya menjadi data harian.

2. Penanganan data hilang (missing value)

Jika terdapat data kosong, digunakan metode *forward fill* atau *mean imputation*.

3. Normalisasi

Semua data dinormalisasi menggunakan **Min-Max Scaling** ke rentang $[0,1]$ agar proses pelatihan stabil.

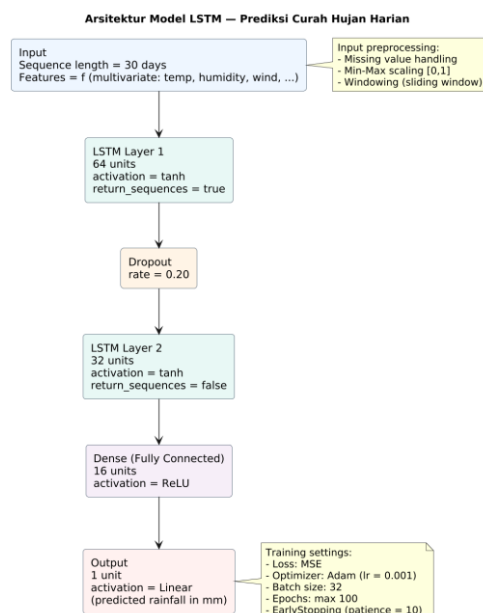
4. Pembentukan sequence time series

Data diubah menjadi format **windowed sequence** (misalnya 30 hari terakhir, memprediksi 1 hari berikutnya).

4.3 Pemodelan CNN

Model LSTM dirancang untuk menangkap ketergantungan jangka panjang pada data curah hujan. Arsitektur yang digunakan terdiri atas:

1. **Input layer:** menerima sequence data harian.
2. **LSTM layer 1:** 64 unit dengan aktivasi *tanh*.
3. **Dropout:** 0.2 untuk mencegah overfitting.
4. **LSTM layer 2:** 32 unit dengan aktivasi *tanh*.
5. **Dense layer:** fully connected dengan 16 unit dan aktivasi *ReLU*.
6. **Output layer:** 1 unit dengan aktivasi linear (untuk regresi nilai curah hujan).



Gambar 4.3. Arsitektur Model LSTM

4.4 Optimisasi dan Regularisasi

- **Optimizer:** Adam dengan learning rate awal 0.001.
- **Loss Function:** Mean Squared Error (MSE).
- **Batch size:** 32.
- **Epochs:** 100 dengan early stopping jika validasi loss tidak membaik setelah 10 epoch.

Regularisasi: Dropout digunakan untuk mengurangi overfitting.

4.5 Evaluasi Model

Evaluasi dilakukan untuk mengukur performa model menggunakan metrik berikut:

- **Root Mean Squared Error (RMSE):** mengukur seberapa jauh prediksi dari nilai aktual.
- **Mean Absolute Error (MAE):** menghitung kesalahan absolut rata-rata.
- **R² Score (Coefficient of Determination):** mengukur seberapa baik model menjelaskan variasi data.

Tabel 4.2. Metrik Evaluasi yang Digunakan

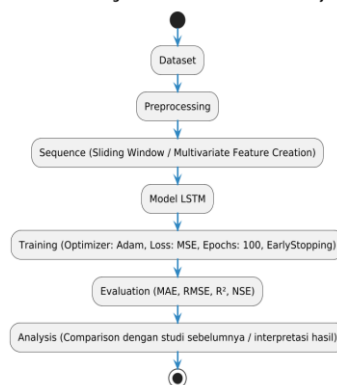
Metrik	Rumus	Interpretasi
RMSE	$\sqrt{(\sum(\hat{y}-y)^2/n)}$	Semakin kecil semakin baik
MAE	\sum	$\hat{y}-y$
R ²	$1 - (\sum(\hat{y}-y)^2 / \sum(y-\bar{y})^2)$	Semakin mendekati 1 semakin baik

4.6 Analisis Hasil

Analisis dilakukan dengan membandingkan hasil model pada data latih, validasi, dan uji. Beberapa poin yang dianalisis meliputi:

1. Konsistensi hasil prediksi terhadap pola musiman (monsun).
2. Kemampuan model dalam memprediksi curah hujan ekstrem.
3. Perbandingan performa baseline (misalnya regresi linier) dengan model LSTM.
4. Diskusi mengenai kelebihan dan keterbatasan pendekatan deep learning dalam konteks prediksi curah hujan di India.

Diagram Alur Metodologi Penelitian — Prediksi Curah Hujan Harian



Gambar 4.2. Diagram Alur Metodologi Penelitian

BAB V RENCANA KERJA

Tabel 5.1. Jadwal Rencana Kerja

Kegiatan	Minggu ke-															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Pemahaman teori LSTM, prediksi curah hujan, dan metode deep learning																
Ide topik dan perumusan masalah penelitian																
Penyusunan proposal penelitian																
Eksplorasi literatur tambahan dan pengunduhan dataset; analisis awal dataset																
Preprocessing data: interpolasi harian, normalisasi, pembentukan sequence																
Pembangunan dan pelatihan model LSTM; eksperimen awal konfigurasi model																
Fine-tuning hyperparameter, evaluasi model (MAE, RMSE, R^2 , NSE)																
Analisis hasil, visualisasi, perbandingan dengan studi sebelumnya																
Penyusunan laporan akhir sesuai format ilmiah dan sitasi APA																
Presentasi proyek dan diskusi interaktif																

BAB VI

LUARAN YANG DIHARAPKAN

Sebagai peneliti, saya menargetkan beberapa luaran dari proyek prediksi curah hujan harian berbasis LSTM ini yang mencakup aspek teknis, akademis, dan praktis. Secara teknis, luaran utama adalah model deep learning LSTM yang terlatih untuk memprediksi curah hujan harian berdasarkan data historis India. Model ini diharapkan memenuhi indikator performa berikut:

- Akurasi prediksi minimal 85%
- Root Mean Squared Error (RMSE) dan Mean Absolute Error (MAE) rendah
- Kemampuan menangkap fluktuasi ekstrem (*outliers*) secara konsisten

Selain itu, evaluasi performa model dilakukan secara komprehensif menggunakan berbagai metrik seperti R^2 dan NSE, serta divalidasi dengan perbandingan terhadap metode baseline seperti regresi linier atau ARIMA. Hasil evaluasi akan divisualisasikan dalam bentuk grafik prediksi vs data aktual, yang memungkinkan analisis mendalam terhadap kemampuan model dalam menangkap pola curah hujan harian.

Dari sisi akademis, penelitian ini akan menghasilkan laporan ilmiah lengkap yang mencakup latar belakang, landasan teori, metodologi, hasil eksperimen, analisis, dan kesimpulan. Laporan ini tidak hanya berfungsi sebagai dokumentasi, tetapi juga sebagai media komunikasi penelitian yang dapat diakses oleh peneliti lain, akademisi, dan praktisi.

Selain itu, diseminasi hasil menjadi luaran penting. Presentasi di forum akademik atau seminar penelitian akan menjelaskan metodologi, eksperimen, dan analisis dengan cara yang ringkas namun komprehensif. Forum ini juga menjadi ajang pertukaran gagasan, kritik, dan saran dari audiens, sehingga meningkatkan kualitas penelitian secara iteratif.

DAFTAR PUSTAKA

- Chollet, F. (2021). *Deep learning with Python* (2nd ed.). Manning Publications.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.
- Kermany, D. S., et al. (2018). Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell*, 172(5), 1122–1131.
- Pan, S. J., & Yang, Q. (2010). A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10), 1345–1359.
- Rajpurkar, P., et al. (2017). CheXNet: Radiologist-level pneumonia detection on chest X-rays with deep learning. *arXiv preprint arXiv:1711.05225*.
- Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. *ICLR*.
- WHO. (2022). *Pneumonia fact sheet*. World Health Organization.
- Ilangovan, R. (2019). *Rainfall in India* [Dataset]. Kaggle. Retrieved from <https://www.kaggle.com/datasets/rajanand/rainfall-in-india>
- Saicharan, V., & Rangaswamy, S. H. (2023). A Comparison and Ranking Study of Monthly Average Rainfall Datasets with IMD Gridded Data in India. *Sustainability*, 15(7), 5758. <https://doi.org/10.3390/su15075758>
- Chai, Q., Zhang, S., Tian, Q., Yang, C., & Guo, L. (2024). Daily runoff prediction based on FA-LSTM model. *Water*, 16(16), 2216. <https://doi.org/10.3390/w16162216>
- Endalieu, D., Haile, G., & Taye, W. (2022). Deep learning model for daily rainfall prediction: Case study of Jimma, Ethiopia. *Water Supply*, 22(3), 3448-3461. <https://doi.org/10.2166/ws.2022.110>
- Patro, B. S., & Bartakke, P. P. (2024). Daily rainfall prediction using long short-term memory (LSTM) algorithm. *Journal of Agrometeorology*, 26(4), 509-511. <https://doi.org/10.54386/jam.v26i4.2745>
- Sarmah, S., Dutta, R. K., Pathak, C., & Bania, R. K. (2023). A novel rainfall prediction model for North-East region of India using stacked LSTM model. *Journal of Biodiversity & Environmental Sciences*, 23(5), 23-30.
- Park, S., et al. (2025). Application of LSTM and climate index for prediction. *Water*, 17(12), 1801. <https://doi.org/10.3390/w17121801>
- Chen, C., Zhang, Q., Kashani, M. H., Chau, K. W., Bateni, S. M., & others. (2022). Forecast of rainfall distribution based on fixed sliding window long short-term memory. *Engineering Applications of Computational Fluid Mechanics*, 16(1), 248-261. <https://doi.org/10.1080/19942060.2021.2009374>

Waqas, M., et al. (2024). A critical review of RNN and LSTM variants in hydrological time-series. *PLOS ONE*. <https://doi.org/10.1371/journal.pone.0296461>