

**ANALISIS *LONG SHORT-TERM MEMORY* UNTUK MEMPREDIKSI
CUACA BERDASARKAN PENYINARAN MATAHARI**

TESIS

**Oleh:
NOR FAIZA
NIM. 240605210010**



**PROGRAM STUDI MAGISTER INFORMATIKA
FAKULTAS SAINS & TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2026**

**ANALISIS *LONG SHORT-TERM MEMORY* UNTUK MEMPREDIKSI
CUACA BERDASARKAN PENYINARAN MATAHARI**

TESIS

**Diajukan kepada:
Universitas Islam Negeri Maulana Malik Ibrahim Malang
Untuk memenuhi Salah Satu Persyaratan dalam
Memperoleh Gelar Magister Komputer (M.Kom)**

**Oleh:
NOR FAIZA
NIM.240605210010**

**PROGRAM STUDI MAGISTER INFORMATIKA
FAKULTAS SAINS & TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2026**

**ANALISIS *LONG SHORT-TERM MEMORY* UNTUK MEMPREDIKSI
CUACA BERDASARKAN PENYINARAN MATAHARI**

TESIS

**Diajukan kepada:
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang
Untuk memenuhi Salah Satu Persyaratan dalam
Memperoleh Gelar Magister Komputer (M.Kom)**

**Oleh:
NOR FAIZA
NIM.240605210010**

**PROGRAM STUDI MAGISTER INFORMATIKA
FAKULTAS SAINS & TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2026**

**ANALISIS LONG SHORT-TERM MEMORY UNTUK MEMPREDIKSI
CUACA BERDASARKAN PENYINARAN MATAHARI**

TESIS

**Oleh:
NOR FAIZA
NIM.240605210010**

Telah diperiksa dan disetujui untuk diuji:
Tanggal: 25 Mei 2026

Pembimbing I,



Dr. Agung Teguh Wibowo A, S.kom., M.T
NIP 19860301 202321 1 016

Pembimbing II,



Dr. Usman Pagalay, M.Si
NIP. 19650414 200312 1 001

Mengetahui,
Ketua Program Studi Magister Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang



Prof. Dr. U. Muhammad Faisal, S.Kom., M.T
NIP. 19740310 200501 1 007

**ANALISIS *LONG SHORT-TERM MEMORY* UNTUK MEMPREDIKSI
CUACA BERDASARKAN PENYINARAN MATAHARI**

TESIS

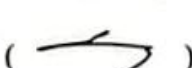
Oleh:
NOR FAIZA
NIM.240605210010

Telah Dipertahankan di Depan Dewan Penguji Tesis
dan Dinyatakan Diterima Sebagai Salah Satu Persyaratan
Untuk Memperoleh Gelar Magister Komputer (M.Kom)
Tanggal: 25 Mei 2026

Susunan Dewan Penguji

Penguji I	:	<u>Dr. Ririen Kusumawati, S.Si., M.Kom</u> NIP 19720309 200501 2 002
Penguji II	:	<u>Dr. Ir. Yunifa Miftachul Arif, S.ST., M.T., SMIEEE</u> NIP. 19830616 201101 1 004
Pembimbing I	:	<u>Dr. Agung Teguh Wibowo Almais, S.Kom., M.T</u> NIP 19860301 202321 1 016
Pembimbing II	:	<u>Dr. Usman Pagalay, M.Si</u> NIP. 19650414 200312 1 001

Tanda Tangan

()
()
()
()

Mengetahui dan Mengesahkan
Ketua Program Studi Magister Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang


Prof. Dr. H. Mubandaz Faisal, S.Kom., M.T
NIP. 19740510 200501 1 007

PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan dibawah ini:

Nama : Nor Faiza
NIM : 240605210010
Program Studi : Magister Informatika
Fakultas : Sains dan Teknologi

Menyatakan dengan sebenarnya bahwa Thesis yang saya tulis ini benar-benar merupakan hasil karya saya sendiri, bukan merupakan pengambilalihan data, tulisan atau pikiran orang lain yang saya akui sebagai hasil tulisan atau pikiran saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar pustaka.

Apabila dikemudian hari terbukti atau dapat dibuktikan thesis ini hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 08 Mei 2026

Yang membuat pernyataan,



Nor Faiza

NIM.240605210010

MOTTO

قُلْ لَّنْ يُصِيبَنَا إِلَّا مَا كَتَبَ اللَّهُ لَنَا

"Tidak akan menimpa kami melainkan apa yang telah ditetapkan Allah bagi kami." (QS. At-Taubah [9]: 51)

“Percayalah bahwa setiap usaha memiliki hasilnya, setiap doa memiliki waktunya, dan setiap takdir akan menemukan pemiliknya. Sebab apa yang telah Allah tetapkan untukmu tidak akan pernah tertukar.”

“Kesuksesan bukan tentang seberapa cepat sampai, tetapi tentang tidak berhenti melangkah.”

PERSEMBAHAN

Segala puji bagi Allah SWT yang telah memberikan limpahan rahmat dan hidayah kepada hamba-Nya, sehingga saya bisa menyelesaikan karya ini. Serta shalawat dan salam semoga selalu tercurahkan kepada nabi Muhammad SAW.

Saya persembahkan karya ini kepada:

Bapak dan Emak saya,

Alm. Musahnan dan Imma Wati.

Yang telah memberikan kasih sayang, doa, pengorbanan, dukungan, serta pelajaran hidup yang tak ternilai. Segala perjuangan, ketulusan, dan cinta yang telah diberikan menjadi sumber semangat dalam setiap langkah perjalanan hidup dan pendidikan saya. Semoga karya sederhana ini dapat menjadi wujud rasa hormat, bakti, dan terima kasih yang mendalam atas segala jasa yang telah diberikan.

Teruntuk adik saya,

Ayyul Hizbaini

Terima kasih atas doa, dukungan, dan semangat yang selalu diberikan. Semoga langkah kecil ini dapat menjadi motivasi bagimu untuk terus berusaha, belajar, dan meraih cita-cita yang diimpikan. Semoga Allah SWT senantiasa memberikan kemudahan dan keberkahan dalam setiap langkahmu.

Tunangan saya,

Muhammad Darwis, S.Ars

Terimakasih telah menemani, memberikan doa, dukungan dan semangat selama proses penyelesaian karya ini. Semoga langkah kecil ini menjadi awal dari perjalanan dan impian yang akan kita raih bersama.

Keluarga besar “Bani Assalbiyah”

Terima kasih atas doa, dukungan, dan motivasi yang telah diberikan. Semoga karya kecil ini dapat menjadi kebanggaan dan membawa manfaat bagi keluarga.

Sahabat dan Teman Seperjuangan,

Maimuna Syarifatussolehah, Teman” Kontrakan Sejahtera dan Seluruh Mahasiswa Magister Informatika khususnya Angkatan 10 yang tidak bisa saya sebutkan satu persatu.

Terima kasih atas kebersamaan, dukungan, motivasi, dan pengalaman berharga yang telah kita lalui bersama.

KATA PENGANTAR

Assalamualaikum Wr. Wb.

Syukur alhamdulillah penulis haturkan kehadiran Allah SWT yang telah melimpahkan Rahmat dan Hidayah-Nya, sehingga penulis dapat menyelesaikan studi di Program Studi Magister Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang sekaligus menyelesaikan Thesis ini dengan baik.

Selanjutnya penulis haturkan ucapan terimakasih seiring do'a dan harapan jazakumullah ahsanal jaza' kepada semua pihak yang telah membantu terselesaikannya Thesis ini. Ucapan terimakasih ini penulis sampaikan kepada:

1. Prof. Dr. Hj. Ilfi Nur Diana, M.Si., selaku Rektor Universitas Islam Negeri Maulana Malik Ibrahim Malang.
2. Dr. H. Agus Mulyono, M.Kes., selaku Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang.
3. Prof. Dr. Ir. Muhammad Faisal, S.Kom., M.T., selaku Ketua Program Studi Magister Informatika Universitas Islam Negeri Maulana Malik Ibrahim Malang.
4. Dr. Agung Teguh Wibowo Almais, S.Kom., M.T., selaku Dosen Pembimbing I yang telah memberikan bimbingan, arahan, dan motivasi selama proses penyusunan tesis.
5. Dr. Usman Pagalay, M.Si., selaku Dosen Pembimbing II, yang telah banyak memberikan pengarahan dan pengalaman yang berharga.
6. Dr. Ririen Kusumawati, S.Si., M.Kom., selaku Dosen Penguji I yang telah memberikan kritik dan saran yang membangun.
7. Dr. Ir. Yunifa Miftachul Arif, S.ST., M.T., SMIEEE., Selaku Dosen Penguji II yang telah memberi kritik dan saran sehingga penulis dapat menuntaskan thesis dengan baik.
8. Segenap sivitas akademika Program Studi Magister Informatika, terutama seluruh Bapak/Ibu dosen, terimakasih atas segenap ilmu dan bimbingannya.

9. Bapak Ahmad Zarkoni, M.Kom dan Mohammad Manba'ul ulum selaku Pembimbing Lapangan yang telah memberikan arahan dan dukungan sehingga penulis dapat menyelesaikan thesis ini dengan baik.
10. Ayahanda dan Ibunda tercinta yang senantiasa memberikan doa dan restunya kepada penulis dalam menuntut ilmu.
11. Adik penulis yang selalu memberikan semangat kepada penulis untuk menyelesaikan Thesis ini.
12. Semua pihak yang ikut membantu dalam menyelesaikan Thesis ini baik berupa materil maupun moril.

Penulis menyadari bahwa dalam penyusunan Thesis ini masih terdapat kekurangan dan penulis berharap semoga Thesis ini bisa memberikan manfaat kepada para pembaca khususnya bagi penulis secara pribadi. *Amin Ya Rabbal Alamin.*

Wassalamualaikum Wr. Wb.

Malang, 08 Mei 2026
Penulis

DAFTAR ISI

COVER	i
HALAMAN PENGAJUAN	ii
HALAMAN PERSETUJUAN.....	iv
HALAMAN PENGESAHAN	v
HALAMAN PERNYATAAN.....	vi
MOTTO	vii
PERSEMBAHAN.....	viii
KATA PENGANTAR.....	ix
DAFTAR ISI.....	xi
DAFTAR TABEL	xiv
DAFTAR GAMBAR	xv
ABSTRAK	xvi
المخلص.....	xviii
BAB I.....	1
PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Pernyataan Masalah	5
1.3 Tujuan Penelitian	5
1.4 Batasan Masalah	5
1.5 Manfaat Penelitian	6
BAB II	7
STUDI PUSTAKA.....	7
2.1 Prediksi Cuaca.....	7
2.2 Kerangka Teori.....	15

2.3 Time Series Data.....	16
2.4 LSTM (Long Short-Term Memory)	18
2.5 Evaluasi Model	21
BAB III.....	24
METODOLOGI PENELITIAN	24
3.1 Kerangka Konsep.....	24
3.2 Desain Penelitian	25
3.2.1 Pengumpulan Data.....	26
3.2.2 Preprocessing Data.....	27
3.2.3 Desain Sistem.....	29
3.2.4 Pemodelan dengan LSTM	31
3.3 Skenario uji coba.....	32
BAB IV.....	34
HASIL DAN PEMBAHASAN	34
4.1 Arsitektur Model LSTM.....	34
4.2 Pre-Processing Data (Persiapan Data)	35
4.2.1 Cleaning Data	36
4.2.2 Normalisasi Data	36
4.2.3 Transformasi Windowing	37
4.2.4 Pembagian Data (Data Splitting).....	38
4.3 Hyperparameter LSTM	39
4.4 Implementasi dan Pelatihan Model LSTM.....	40
4.5 Pengujian Model.....	64
4.6 Model Matematis LSTM	71
4.7 Prediksi Cuaca Dalam Pandangan Islam	74

BAB V	77
PENUTUP.....	77
5.1 KESIMPULAN.....	77
5.2 SARAN	78
DAFTAR PUSTAKA.....	79

DAFTAR TABEL

Tabel 2.1. Prediksi Cuaca.....	13
Tabel 3.1. Dataset Penyinaran Matahari	27
Tabel 3.2. Threshold Perkiraan Berdasarkan Iradiasi Absolut (W/m^2)	29
Tabel 3.3 Rancangan Skenario Uji Coba	32
Tabel 4.1. Persiapan Data.....	36
Tabel 4.2. Hasil Normalisasi Data.....	37
Tabel 4.3. Pembagian Data.....	39
Tabel 4.4. hyperparameter Tuning (Hartono, 2025).....	40
Tabel 4.5. Hasil Evaluasi Model	64
Tabel 4.6. Nilai Aktual vs Prediksi.....	68

DAFTAR GAMBAR

Gambar 2.1 Kerangka Teori	15
Gambar 2.2 Architecture LSTM (Supriyadi, 2020)	19
Gambar 3.1 Kerangka Konsep Penelitian	25
Gambar 3.2 Desain Penelitian.....	26
Gambar 3.3 Desain Sistem.....	30
Gambar 4.1 Arsitektur LSTM	34
Gambar 4.2. Loss Function Model A1	42
Gambar 4.3. Data Aktual vs Prediksi Data Model A1	44
Gambar 4.4. Loss Function Model A2	45
Gambar 4.5. Data Aktual vs Prediksi Data Model A2	46
Gambar 4.6. Loss Function Model A3	48
Gambar 4.7. Data Aktual vs Prediksi Data Model A3	49
Gambar 4.8. Loss Function Model B1	50
Gambar 4.9. Data Aktual vs Prediksi Data Model B1	51
Gambar 4.10. Loss Function Model B2	53
Gambar 4.11. Data Aktual vs Prediksi Data Model B2	54
Gambar 4.12. Loss Function Model B3	55
Gambar 4.13. Data Aktual vs Prediksi Data Model B3	56
Gambar 4.14. Loss Function Model C1	57
Gambar 4.15. Data Aktual vs Prediksi Data Model C1	58
Gambar 4.16. Loss Function Model C2	59
Gambar 4. 17. Data Aktual vs Prediksi Data Model C2	61
Gambar 4.18. Loss Function Model C3	62
Gambar 4.19. Data Aktual vs Prediksi Data Model C3	64
Gambar 4.20. Grafik Perbandingan MAE dan RMSE.....	66
Gambar 4.21. Grafik Akurasi R^2 Antar Model.....	66
Gambar 4.23. Grafik Data Actual Dan Prediksi.....	67

ABSTRAK

Faiza, Nor. 2026. **Analisis Long Short-Term Memory Untuk Memprediksi Cuaca Berdasarkan Penyinaran Matahari**. Thesis. Program Studi Magister Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing: (I) Dr. Agung Teguh Wibowo Almais, S.Kom., M.T. (II) Dr. Usman Pagalay. M.Si.

Kata kunci: Prediksi Cuaca, LSTM, Penyinaran Matahari, MAE, RMSE, Optimasi Hiperparameter, Deep Learning.

Prediksi parameter meteorologi merupakan tantangan yang kompleks dalam bidang informatika karena tingginya variabilitas dan sifat non-linier data atmosfer. Permasalahan utama dalam studi ini adalah fluktuasi data penyinaran matahari yang sulit diidentifikasi oleh model konvensional, serta adanya risiko degradasi performa model akibat konfigurasi arsitektur dan pembagian data yang tidak optimal. Penelitian ini bertujuan untuk mengevaluasi efektivitas arsitektur jaringan *Long Short-Term Memory* (LSTM) dalam memprediksi kondisi cuaca harian melalui variabel *Solar Radiation Average* (SR Avg), *Solar Radiation Maximum* (SR Max), dan durasi penyinaran. Metode eksperimen dilakukan dengan mengevaluasi sembilan skenario arsitektur LSTM melalui proses *hyperparameter tuning* dan variasi komposisi data latih (70%, 80%, dan 90%) dengan keluaran 1 lapisan *Dense*. Proses *tuning* difokuskan pada optimasi kombinasi jumlah *neurons* (3 hingga 30 unit), durasi iterasi (*epoch* 300 dan 500), serta ukuran komputasi (*batch size* 50 hingga 200). Kinerja estimasi diukur menggunakan indikator *Mean Absolute Error* (MAE), *Root Mean Square Error* (RMSE) dan Koefisien Determinasi (R^2_Score). Hasil uji global skala 0–1 menunjukkan nilai MAE total 0.1095, RMSE 0.1392, dan R^2_Score 0.3594 (menjelaskan 35,94% variabilitas). Secara komparatif, Model A1 (data latih 70%) mencatatkan kapabilitas generalisasi terbaik dengan R^2 tertinggi 0.4142 dan RMSE terkecil 0.1193. Sementara itu, Model B2 (data latih 80%) menghasilkan penyimpangan absolut terendah dengan MAE terkecil 0.0911. Sebaliknya, porsi data latih 90% memicu *overfitting* yang menurunkan performa model. Pada batas toleransi error < 0.05 , akurasi kategorial menghasilkan prediksi "Benar" 28,55% (207 hari) dan "Salah" 71,45% (518 hari). Rendahnya persentase kategorial aman ini dipicu oleh *smoothing effect* pada LSTM, model akurat pada fase stabil (Berawan), namun mengalami regresi menuju nilai rata-rata (*underestimate*) dan gagal mengejar nilai puncak saat transisi ekstrem ke cuaca Cerah.

ABSTRACT

Faiza, Nor. 2026. **Long Short-Term Memory Analysis for Weather Prediction Based on Solar Radiation**. Thesis. Master Program of Informatics, Faculty of Science and Technology, Universitas Islam Negeri Maulana Malik Ibrahim Malang. Advisors: (I) Dr. Agung Teguh Wibowo Almais, S.Kom., M.T. (II) Dr. Usman Pagalay. M.Si.

Forecasting meteorological parameters is a complex challenge in the field of informatics due to the high variability and non-linear nature of atmospheric data. The main problems in this study are the fluctuations of solar radiation data that are difficult to identify by conventional models, as well as the risk of degradation of model performance due to architectural configuration and suboptimal data sharing. This study aims to evaluate the effectiveness of the Long Short-Term Memory (LSTM) network architecture in predicting daily weather conditions through the variables of Solar Radiation Average (SR Avg), Solar Radiation Maximum (SR Max), and duration of irradiation. The experimental method was carried out by evaluating nine LSTM architecture scenarios through a hyperparameter tuning process and variations in the composition of the training data (70%, 80%, and 90%) with the output of 1 Dense layer. The tuning process was focused on optimizing the combination of the number of neurons (3 to 30 units), the duration of the iteration (epoch 300 and 500), and the computational size (batch size 50 to 200). The estimated performance was measured using the Mean Absolute Error (MAE), Root Mean Square Error (RMSE) and Coefficient of Determination (R^2 _Score) indicators. The results of the global test on a scale of 0–1 showed a total MAE value of 0.1095, RMSE 0.1392, and R^2 _Score 0.3594 (explaining 35.94% variability). Comparatively, Model A1 (70% training data) recorded the best generalization capabilities with the highest R^2 of 0.4142 and the smallest RMSE of 0.1193. Meanwhile, Model B2 (80% training data) produced the lowest absolute deviation with the smallest MAE of 0.0911. In contrast, 90% of the training data portion triggers overfitting that degrades model performance. At an error tolerance limit of < 0.05 , categorical accuracy results in a "True" prediction of 28.55% (207 days) and "False" of 71.45% (518 days). This low percentage of safe categorizations is triggered by the smoothing effect on LSTMs, which are accurate in the stable phase (Cloudy), but regress to mean values (underestimate) and fail to catch peak values during extreme transitions to sunny weather.

Key words: Weather Prediction, LSTM, Solar Radiation, MAE, RMSE, Hyperparameter Optimization, Deep Learning.

الملخص

فايزة، نور. 2026. تحليل الذاكرة طويلة المدى قصيرة الأجل للتنبؤ بالطقس بناء على الإشعاع الشمسي. أطروحة. برنامج ماجستير المعلوماتية، كلية العلوم والتكنولوجيا، جامعة مولانا مالك إبراهيم الإسلامية، مالانغ. المشرف: (الأول) الدكتور أغونغ تيغو ويوو ألميس، م.م.ت. (2) الدكتور عثمان باغلاي. M.Si.

الكلمات المفتاحية: التنبؤ بالطقس، LSTM، الإشعاع الشمسي، RMSE، MAE، تحسين المعلمات الفائقة، التعلم العميق.

التنبؤ بمعايير الأرصاد الجوية يمثل تحدياً معقداً في مجال المعلوماتية بسبب التغيرات العالية والطبيعة غير الخطية لبيانات الغلاف الجوي. المشاكل الرئيسية في هذه الدراسة هي تقلبات بيانات الإشعاع الشمسي التي يصعب تحديدها بواسطة النماذج التقليدية، بالإضافة إلى خطر تدهور أداء النماذج بسبب التكوين المعماري ومشاركة البيانات غير المثالية. تهدف هذه الدراسة إلى تقييم فعالية بنية شبكة الذاكرة طويلة المدى قصيرة المدى (LSTM) في التنبؤ بظروف الطقس اليومية من خلال متغيرات متوسط الإشعاع الشمسي (SR AVG)، وأقصى إشعاع شمسي (SR Max)، ومدة الإشعاع. تم تنفيذ الطريقة التجريبية من خلال تقييم تسعة سيناريوهات لبنية LSTM من خلال عملية ضبط المعلمات الفائقة وتغيرات في تركيب بيانات التدريب (70٪، 80٪، 90٪) مع إنتاج طبقة كثيفة واحدة. ركزت عملية الضبط على تحسين الجمع بين عدد الخلايا العصبية (من 3 إلى 30 وحدة)، ومدة التكرار (العصر 300 و 500)، والحجم الحسابي (حجم الدفعة من 50 إلى 200). تم قياس أداء التقدير باستخدام مؤشرات متوسط الخطأ المطلق (MAE)، والجذر متوسط التربيع (RMSE)، ومعامل التحديد (R^2_Score). أظهرت نتائج الاختبار العالمي على مقياس 0-1 قيمة MAE إجمالية 0.1095، RMSE 0.1392، و R^2_Score 0.3594 (مما يفسر تباين 35.94٪). بالمقارنة، سجل النموذج A1 (70٪ من بيانات التدريب) أفضل قدرات تعميم بأعلى R^2 0.4142 وأصغر RMSE 0.1193. وفي الوقت نفسه، أنتج النموذج B2 (80٪ من بيانات التدريب) أدنى انحراف مطلق مع أصغر مؤشر متوسط متوسط بلغ 0.0911. على النقيض من ذلك، 90٪ من جزء بيانات التدريب يسبب إفراط في التركيب يؤثر على أداء النموذج. عند حد تحمل الخطأ $0.05 >$ ، تؤدي الدقة التصنيفية إلى توقع "صحيح" بنسبة 28.55٪ (207 أيام) و"خاطئ" بنسبة 71.45٪ (518 يوماً). هذه النسبة المنخفضة من التصنيفات الآمنة تفعل تأثير التنعيم على نماذج LSTM، التي تكون دقيقة في الطور المستقر (الغائم)، لكنها تتراجع إلى القيم المتوسطة (تقلل من التقدير) وتفشل في النقاط قيم الذروة خلال الانتقالات القصوى إلى الطقس المشمس.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi dan meningkatnya kebutuhan akan informasi cuaca yang akurat telah mendorong kemajuan signifikan dalam sistem prediksi cuaca modern. Informasi cuaca yang tepat waktu memegang peran krusial bagi berbagai sektor, mulai dari pertanian, transportasi, energi, hingga manajemen bencana (Dwiyanti & Prianto, 2023). Ketidakakuratan prediksi cuaca dapat menimbulkan dampak sosial dan ekonomi yang serius, terutama di negara tropis seperti Indonesia yang memiliki variabilitas cuaca yang sangat tinggi (Damayanti & Prabowo, 2024). Tingginya frekuensi kejadian cuaca ekstrem seperti banjir bandang, kekeringan, dan angin puting beliung, serta ketergantungan sektor vital seperti pertanian pangan dan aviasi pada kondisi cuaca, menjadikan akurasi prediksi sebagai kebutuhan strategis nasional. Oleh karena itu, diperlukan pendekatan baru yang lebih adaptif dan akurat untuk memodelkan perubahan cuaca berdasarkan variabel klimatologis yang relevan (Sarwana dkk, 2024).

Salah satu parameter penting dalam prediksi cuaca adalah penyinaran matahari. Variabel ini tidak hanya memengaruhi suhu udara, tetapi juga berkontribusi pada pembentukan awan, kelembapan, serta tekanan atmosfer (Wida dkk, 2019). Penyinaran matahari mencerminkan dinamika atmosfer yang kompleks dan sangat dipengaruhi oleh faktor-faktor seperti keberadaan awan, kabut, serta pergerakan rotasi dan revolusi bumi (Hamdi, 2015).

Di wilayah tropis seperti Indonesia, pola penyinaran matahari memiliki karakteristik unik akibat pengaruh monsun, topografi vulkanis, tutupan lautan yang luas, dan aktivitas konvektif yang intens sepanjang tahun. Namun, sifatnya yang fluktuatif dan nonlinier membuat model prediksi tradisional kesulitan dalam mengidentifikasi pola perubahannya secara konsisten (Atina & Prasetyo, 2022). Keterbatasan model fisika-atmosfer konvensional dalam menangkap kompleksitas interaksi lokal dan non-linearitas ini menjadi tantangan utama dalam menghasilkan prediksi penyinaran matahari yang andal untuk skala harian.

Tantangan prediksi di Indonesia semakin kompleks mengingat terbatasnya jaringan pengamatan meteorologi yang merata secara spasial, terutama di wilayah kepulauan terpencil dan area perbatasan (Dwiyanti & Prianto, 2023). Meskipun teknologi satelit memberikan cakupan luas, resolusi temporal dan spasialnya masih memiliki keterbatasan untuk pemodelan hiper-lokal yang dibutuhkan dalam perencanaan *sectoral* (Sains dkk, 2022). Di sisi lain, perkembangan *Internet of Things* (IoT) dan sensor berkemampuan tinggi mulai memungkinkan pengumpulan data penyinaran matahari secara *real-time* dengan densitas spasial yang lebih baik, menciptakan peluang baru untuk pemodelan prediktif berbasis data beresolusi tinggi (Setiawan dkk, 2023).

Dalam bidang ilmu komputer dan teknologi data, pendekatan *machine learning* telah terbukti efektif untuk memodelkan data yang kompleks dan bersifat sekuensial (de Castro dkk, 2022). Di antara berbagai algoritma, *Recurrent Neural Network* (RNN) (Shin dkk, 2021), *Long Short-Term Memory* (LSTM) menunjukkan performa yang sangat baik dalam mengatasi permasalahan data deret waktu.

Berbeda dengan RNN klasik yang rentan terhadap masalah *vanishing gradient* (menghilangnya informasi penting pada urutan panjang), LSTM dirancang untuk mengatasi kelemahan tersebut melalui mekanisme *gate* yang mengatur aliran informasi (Yunita dkk, 2025). Dengan struktur "*cell state*" dan tiga gerbang (*input, forget, output*), LSTM mampu secara selektif menyimpan, melupakan, dan mempertahankan informasi penting dari urutan data masa lalu untuk digunakan dalam prediksi masa depan (Gao dkk, 2022). Kemampuan memodelkan ketergantungan temporal jangka panjang (*long-term dependencies*) inilah yang menjadikan LSTM sangat potensial untuk memecahkan kompleksitas pola cuaca tropis yang dipengaruhi oleh osilasi iklim seperti MJO (*Madden-Julian Oscillation*) dan ENSO (*El Niño-Southern Oscillation*).

Penelitian sebelumnya menunjukkan bahwa model *deep learning*, khususnya arsitektur berbasis LSTM dan *transformer*, memiliki potensi besar dalam meningkatkan akurasi prakiraan cuaca dan variabel lingkungan lainnya. Penelitian oleh Maharatha dkk (2024) menunjukkan bahwa penggunaan arsitektur *sequence-to-sequence stacked LSTM autoencoder* secara signifikan meningkatkan akurasi prediksi temperatur di wilayah Khordha, India. Mereka membandingkan beberapa varian LSTM, dan menemukan bahwa model 9-layer *stacked LSTM* dengan *bidirectional decoder* menghasilkan *error* terendah (MAE terkecil), mengindikasikan keunggulan model ini dalam menangkap pola suhu minimum dan maksimum secara *multivariat*.

Sementara itu, Cao dkk (2025) mengembangkan kerangka kerja prakiraan cuaca lima hari berbasis model segmentasi semantik dan teknik *cascade prediction*,

yang terbukti unggul dalam skenario data terbatas. Penelitian ini mengusulkan tiga solusi kunci: penerapan model segmentasi semantik untuk menangkap fitur spasial dan temporal, pengenalan mekanisme kebisingan Gaussian yang dapat dipelajari (*learnable Gaussian noise*) untuk mengatasi *overfitting*, dan penggunaan strategi prediksi *cascade* untuk meminimalkan akumulasi kesalahan. Hasilnya menunjukkan bahwa setiap komponen memberikan kontribusi yang efektif: segmentasi semantik mengurangi kesalahan prediksi suhu sebesar 9.3%, mekanisme kebisingan meningkatkan F1-score prediksi curah hujan sebesar 6.8%, dan strategi *cascade* mengurangi kesalahan kuadrat rata-rata prediksi kecepatan angin sebesar 12.5%.

Dalam konteks energi terbarukan, R. Zhang dkk (2025) mengusulkan pendekatan *multi-task learning* berbasis *Transformer-LSTM* untuk prediksi deterministik dan probabilistik energi angin. Model ini menggabungkan *Dilated Causal Convolution* (DCC), LSTM, dan *Transformer* untuk menangkap fitur nonlinier dan dependensi jangka panjang dari data deret waktu, serta menerapkan optimisasi *hyperparameter* menggunakan algoritma hibrida (CEHSC). Hasilnya menunjukkan bahwa model tersebut secara konsisten mengungguli 23 model *state of the art* lainnya, baik dalam aspek MAE maupun akurasi interval prediksi *probabilistic*.

Meskipun hasil penelitian tersebut menjanjikan, hingga saat ini masih relatif sedikit studi yang secara spesifik mengkaji penerapan model LSTM untuk prediksi kondisi cuaca harian komprehensif dengan menggunakan penyinaran matahari sebagai variabel kunci utama. Terlebih lagi, penerapannya dalam konteks wilayah

tropis basah seperti Indonesia, dengan karakteristik *diurnal pattern* yang khas dan pengaruh laut-kontinen yang intens, masih belum dieksplorasi secara sistematis. Kebutuhan akan model prediksi yang akurat, adaptif terhadap perubahan lokal, dan kontekstual untuk wilayah kepulauan tropis belum terpenuhi secara memadai.

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk menganalisis performa model LSTM dalam memprediksi kondisi cuaca harian berbasis data penyinaran matahari. Penelitian ini tidak hanya menilai kinerja model berdasarkan metrik numerik seperti *Root Mean Square Error* (RMSE) dan *Mean Absolute Error* (MAE), tetapi juga mengevaluasi kemampuan model dalam menghasilkan prediksi kondisi cuaca yang berbeda. Diharapkan hasil penelitian ini dapat memberikan kontribusi dalam pengembangan sistem prediksi cuaca berbasis data yang lebih akurat, adaptif, dan kontekstual untuk wilayah tropis seperti Indonesia.

1.2 Pernyataan Masalah

Bagaimana performa model *Long Short-Term Memory* (LSTM) dalam memprediksi pola perubahan cuaca berdasarkan data historis penyinaran matahari?

1.3 Tujuan Penelitian

Menganalisis performa model *Long Short-Term Memory* (LSTM) dalam memprediksi pola perubahan cuaca berdasarkan data historis penyinaran matahari.

1.4 Batasan Masalah

Agar penelitian tetap terfokus dan hasilnya dapat diukur secara optimal, batasan masalah ditetapkan sebagai berikut:

1. Data yang digunakan hanya mencakup deret waktu historis penyinaran matahari dari stasiun pengamatan BMKG periode Januari - Desember 2023.
2. Prediksi cuaca terbatas pada kondisi umum (cerah, berawan, hujan).
3. Data yang digunakan dalam penelitian ini mencakup wilayah Kabupaten Malang.

1.5 Manfaat Penelitian

Penelitian ini diharapkan memberikan manfaat sebagai berikut:

1. Mengembangkan penerapan model LSTM dalam pemodelan meteorologi berbasis penyinaran matahari di wilayah tropis.
2. Menyediakan model prediksi cuaca yang dapat dimanfaatkan oleh lembaga meteorologi, pemerintah, dan sektor pertanian untuk mendukung perencanaan berbasis cuaca.
3. Memberikan informasi perubahan cuaca harian dan mengurangi risiko kerugian akibat kondisi cuaca.

BAB II

STUDI PUSTAKA

2.1 Prediksi Cuaca

Dalam beberapa tahun terakhir, penelitian di bidang prediksi cuaca telah mengalami kemajuan pesat berkat penerapan teknologi *machine learning*. Berbagai pendekatan telah dikembangkan untuk meningkatkan akurasi dan efisiensi prediksi kondisi atmosfer, khususnya dalam skala jangka pendek.

Penelitian yang dilakukan oleh Nugroho (2023) mengimplementasikan Jaringan Syaraf Tiruan *Multi-Layer Perceptron* (MLP) untuk memprediksi tingkat penyinaran matahari di Kota Bandung. Penelitian menggunakan 72 data periode 2016–2021, dengan 48 data untuk pelatihan dan 24 data untuk pengujian, serta parameter *learning rate* 0,2 dan maksimum 1000 iterasi. Hasil penelitian menunjukkan nilai *Mean Square Error* (MSE) sebesar 0.086 pada tahap pelatihan dan 0,109 pada tahap pengujian. Temuan ini membuktikan bahwa MLP efektif digunakan dalam memprediksi penyinaran matahari dan berpotensi diterapkan pada berbagai sektor seperti pertanian, kesehatan, serta pemanfaatan energi surya secara lebih optimal.

Penelitian oleh Supriyadi (2020) meneliti penerapan metode *deep learning Long-Short Term Memory* (LSTM) untuk prediksi parameter cuaca, meliputi suhu, kelembaban, kecepatan angin, dan tekanan udara. Penelitian menggunakan data pengamatan Stasiun Meteorologi Maritim Tanjung Priok periode Januari–Februari 2019 dengan pembagian data latih dan uji 9:1 serta 200 *layer* LSTM. Hasil penelitian menunjukkan bahwa model LSTM dengan pembaruan data aktual menghasilkan nilai kesalahan (RMSE) yang lebih rendah dibandingkan tanpa

pembaruan, sehingga prediksi hingga tiga hari ke depan cukup akurat, sedangkan untuk jangka panjang hingga satu bulan akurasi cenderung menurun. Temuan ini menegaskan efektivitas LSTM dalam mendukung prakiraan cuaca jangka pendek di Indonesia.

Penelitian oleh Hennayake dkk (2021) bertujuan untuk mengembangkan model prediksi cuaca jangka pendek menggunakan teknologi machine learning (pembelajaran mesin) untuk diterapkan di Sri Lanka. Peneliti berfokus pada prediksi suhu dengan menggunakan model *Long Short-Term Memory* (LSTM). Hal ini dilakukan untuk mengatasi keterbatasan metode tradisional *Numerical Weather Prediction* (NWP) yang kurang akurat untuk kondisi cuaca tropis. Penelitian ini menggunakan model LSTM multivariat yang dilatih dengan data observasi cuaca harian dari stasiun cuaca Kolombo, Sri Lanka, dari tahun 2010 hingga 2019. Data yang mencakup sepuluh parameter cuaca ini melewati proses *pre-processing* seperti pembersihan, pemilihan fitur, dan normalisasi. Dataset kemudian dibagi menjadi 80% untuk pelatihan dan 20% untuk pengujian. Model LSTM yang dirancang memiliki satu lapisan LSTM dengan 128 *node* dan tiga lapisan *dense*. Model ini dilatih menggunakan *mean squared error* (MSE) sebagai fungsi biaya dan *Adam* sebagai algoritma optimasi. Hasil evaluasi menunjukkan bahwa model LSTM memiliki akurasi yang tinggi dalam memprediksi suhu. Metrik evaluasi pada set pengujian menghasilkan nilai kesalahan yang sangat rendah, yaitu RMSE sebesar 1.37482°F, MAE sebesar 0.9898567°F, dan MAPE sebesar 1.211372%. Nilai kesalahan yang rendah ini menunjukkan kemampuan model LSTM untuk

mengidentifikasi pola nonlinier yang kompleks dari data observasi cuaca masa lalu, sehingga menghasilkan prakiraan yang akurat.

Penelitian oleh Yunita dkk (2025) bertujuan untuk membandingkan kinerja dan keandalan sembilan arsitektur jaringan saraf, termasuk RNN, LSTM, GRU, dan model hibridanya, untuk peramalan deret waktu. Menggunakan evaluasi berbasis Monte Carlo pada tiga dataset, penelitian ini menemukan bahwa secara statistik tidak ada perbedaan yang signifikan di antara model-model tersebut. Namun, analisis deskriptif menunjukkan bahwa model hibrid berbasis LSTM, seperti LSTM-GRU dan LSTM-RNN, secara konsisten menunjukkan kinerja yang lebih unggul dibandingkan model lainnya. Temuan ini menyimpulkan bahwa meskipun banyak arsitektur yang tersedia, model hibrid berbasis LSTM adalah pilihan praktis dan kuat untuk peramalan deret waktu.

Penelitian oleh Boubaker dkk (2021) bertujuan untuk mengembangkan dan mengevaluasi berbagai model jaringan saraf tiruan (*Deep Neural Networks/DNN*) untuk memprediksi radiasi matahari horizontal global (GHI) satu hari ke depan di wilayah Hail, Arab Saudi. Penelitian ini merupakan aplikasi pertama dari jenisnya di wilayah tersebut dan menganalisis secara mendalam parameter yang memengaruhi akurasi model. Metode yang digunakan adalah dengan mengembangkan dan membandingkan beberapa model DNN, termasuk *long short-term memory* (LSTM), *bidirectional*-LSTM (BiLSTM), *gated recurrent unit* (GRU), *bidirectional*-GRU (Bi-GRU), dan *hybrid* seperti CNN-LSTM serta CNN-BiLSTM. Model-model ini dilatih menggunakan data GHI harian historis yang dikumpulkan oleh NASA dari 1 Januari 2000 hingga 30 Juni 2020, tanpa

menggunakan parameter cuaca tambahan. Kinerja model dievaluasi menggunakan metrik seperti koefisien korelasi (r), *Mean Absolute Percent Error* (MAPE), dan *Mean Absolute Error* (MAE). Hasil penelitian menunjukkan bahwa model-model DNN memiliki kinerja yang baik dalam memprediksi GHI, dengan nilai koefisien korelasi (r) maksimum mencapai 96%. Meskipun beberapa model menunjukkan kinerja yang sedikit lebih rendah pada nilai GHI di bawah 4 kWh/m²/hari, secara keseluruhan mereka menunjukkan akurasi yang baik dan tidak ada masalah.

Penelitian yang dilakukan oleh Suleman & Shridevi (2022) bertujuan untuk mengembangkan model *deep learning* baru yang disebut *Spatial Feature Attention-Long Short-Term Memory* (SFA-LSTM). Model ini dibuat untuk memprediksi suhu dengan lebih akurat, terutama saat terjadi perubahan pola cuaca yang tiba-tiba, dengan cara menangkap hubungan spasial dan temporal dari berbagai fitur meteorologi. Metode yang digunakan adalah arsitektur *encoder-decoder* yang memanfaatkan data cuaca aktual dari Bandara Internasional Saskatoon. Model ini diuji dengan data dari tahun 2012-2021 dan dibandingkan dengan model-model *deep learning* lainnya. Hasilnya menunjukkan bahwa model SFA-LSTM yang diusulkan memiliki akurasi prediksi yang lebih baik. Keunggulannya terletak pada kemampuannya untuk menginterpretasi bobot perhatian fitur spasial, yang terbukti secara komputasi tidak mahal dan hanya bergantung pada data historis.

Selanjutnya penelitian oleh Andrew Gahwera dkk (2025) tujuan utama dari penelitian tersebut adalah untuk mengatasi masalah kelangkaan data meteorologi dalam memprediksi jumlah kelas curah hujan jangka pendek dengan memperkenalkan pendekatan *ensemble transfer learning* berbasis multi-stasiun.

Pendekatan ini mentransfer pengetahuan yang diperoleh dari kumpulan data beberapa stasiun meteorologi ke satu stasiun target. Metode yang digunakan meliputi model *deep learning* seperti *Multi-Layer Perceptron* (MLP), *Convolutional Neural Networks* (CNN), dan *Long Short-Term Memory* (LSTM). Model-model ini dilatih menggunakan kumpulan data dari Cekungan Danau Victoria (LVB). Pendekatan *ensemble* dengan algoritma bobot eksponensial digunakan untuk menggabungkan prediksi dari model-model yang telah dilatih sebelumnya. Hasilnya, model LSTM menunjukkan performa yang unggul dengan perolehan nilai F1 yang tinggi. Penyesuaian model (*fine-tuning*) yang telah dilatih *ensemble* yang diusulkan lebih lanjut meningkatkan performa, memberikan hasil klasifikasi yang sangat akurat, dengan model *ensemble* mencapai skor F1 mendekati 100% dan akurasi 0.989. untuk stasiun target berhasil meningkatkan akurasi hingga 5%. Pendekatan.

Penelitian oleh Alghushairy dkk (2025) bertujuan untuk mengembangkan sistem peramalan cuaca dua tahap di Arab Saudi guna mengoptimalkan operasi rantai pasok. Sistem ini dirancang untuk memprediksi kecepatan angin dan suhu udara secara akurat. Metodologi yang diusulkan melibatkan empat langkah: akuisisi data, pra-pemrosesan data, pemilihan fitur menggunakan algoritma *Fox metaheuristics*, dan penggunaan model prediksi berbasis *Bidirectional Gated Recurrent Units* (BiGRU). Dalam model dua tahap ini, BiGRU pertama-tama memprediksi kecepatan angin, dan hasil prediksi tersebut kemudian digunakan bersama fitur-fitur lain untuk memprediksi suhu udara pada tahap kedua. Hasil evaluasi menunjukkan bahwa sistem ini berhasil dengan baik, mencapai *mean*

absolute error (MAE) sebesar 1.4714 untuk prediksi kecepatan angin dan 5.6799 untuk prediksi suhu udara.

Penelitian yang dilakukan oleh Karevan & Suykens (2020) untuk mengembangkan model peramalan berbasis data untuk prediksi cuaca menggunakan *Long Short-Term Memory* (LSTM). Penulis mengusulkan model *Transductive LSTM* (T-LSTM) yang memanfaatkan informasi lokal dalam prediksi deret waktu dengan melokalisasi fungsi tujuan melalui fungsi biaya kuadrat tertimbang. Penelitian ini menggunakan dua skema penimbangan berdasarkan kesamaan kosinus antara sampel pelatihan dan titik uji. Hasil dari eksperimen yang dilakukan pada dua periode waktu yang berbeda menunjukkan bahwa T-LSTM menghasilkan kinerja yang lebih baik dibandingkan metode konvensional dalam tugas prediksi deret waktu.

Penelitian oleh Singhal dkk (2022) mengintegrasikan prakiraan cuaca dengan *machine learning* untuk prediksi produksi energi surya. Dari beberapa algoritma yang digunakan, XGBoost menunjukkan performa terbaik dengan R^2 sebesar 0.92, membuktikan bahwa integrasi ini dapat meningkatkan efisiensi sistem energi terbarukan.

Berdasarkan berbagai studi tersebut, dapat disimpulkan bahwa metode *deep learning*, terutama varian LSTM dan kombinasi *hybrid*-nya, sangat efektif dalam prediksi cuaca jangka pendek maupun ekstrem. Model-model ini menawarkan peningkatan akurasi, efisiensi komputasi, dan fleksibilitas dalam berbagai skenario cuaca dan lokasi geografis.

Tabel 2.1. Prediksi Cuaca

No	Referensi	Topik	Metode	Subjek
1	Nugroho (2023)	Prediksi lama penyinaran matahari di Kota Bandung menggunakan metode jaringan syaraf tiruan.	<i>Multi-Layer Perceptron</i> (MLP)	Data penyinaran matahari Kota Bandung dari tahun 2016 hingga 2021.
2	Supriyadi (2020)	Penerapan metode <i>deep learning</i> berbasis Long-Short Term Memory (LSTM) untuk memprediksi parameter cuaca,	<i>Long-Short Term Memory</i> (LSTM)	Data observasi sinoptik dari Stasiun Meteorologi Maritim Tanjung Priok pada bulan Januari–Februari 2019
3	Hennayake dkk (2021)	Prediksi cuaca jangka pendek di wilayah tropis (Sri Lanka) menggunakan <i>machine learning</i>	Model <i>Long Short-Term Memory</i> (LSTM) multivariat	Data observasi cuaca harian (10 parameter cuaca) dari stasiun cuaca Kolombo, Sri Lanka, periode 2010–2019
4	Yunita dkk (2025)	Perbandingan kinerja berbagai arsitektur jaringan saraf untuk peramalan deret waktu	- <i>Recurrent Neural Network</i> (RNN) - <i>Long Short-Term Memory</i> (LSTM) - <i>Gated Recurrent Unit</i> (GRU)	Tiga dataset deret waktu yang digunakan untuk pengujian kinerja model.
5	Boubaker dkk (2021)	Prediksi radiasi matahari horizontal global (GHI) harian di Hail, Arab Saudi menggunakan jaringan saraf tiruan.	- <i>Long Short-Term Memory</i> (LSTM) - <i>Gated Recurrent Unit</i> (GRU) - <i>Convolutional Neural Network</i> (CNN)	Data historis GHI harian dari NASA (1 Januari 2000 - 30 Juni 2020).
6	Suleman & Shridevi (2022)	Prediksi suhu menggunakan model deep learning baru SFA-LSTM dengan fokus pada pola cuaca yang berubah tiba-tiba.	<i>Spatial Feature Attention-Long Short-Term Memory</i> (SFA-LSTM)	Data cuaca historis Bandara Internasional Saskatoon (2012–2021).
7	Andrew Gahwera dkk (2025)	Prediksi kelas curah hujan jangka pendek dengan mengatasi kelangkaan data meteorologi melalui pendekatan ensemble transfer learning berbasis multi-stasiun.	- <i>Multi-Layer Perceptron</i> (MLP) - <i>Convolutional Neural Network</i> (CNN)	Data meteorologi dari Cekungan Danau Victoria (LVB).

No	Referensi	Topik	Metode	Subjek
			- <i>Long Short-Term Memory (LSTM)</i>	
8	Alghushairy dkk (2025)	Peramalan cuaca dua tahap (kecepatan angin dan suhu udara) untuk optimasi rantai pasok di Arab Saudi.	<i>Bidirectional Gated Recurrent Unit (BGRU)</i>	Data cuaca dari wilayah Arab Saudi.
9	Karevan & Suykens (2020)	Prediksi cuaca berbasis data menggunakan model Transductive LSTM (T-LSTM).	Model <i>Transductive Long Short-Term Memory (T-LSTM)</i>	Data deret waktu cuaca pada dua periode waktu berbeda.
10	Singhal dkk (2022)	Prediksi produksi energi surya berbasis integrasi prakiraan cuaca dan machine learning.	<i>Extreme Gradient Boosting (XGBoost)</i>	Data prakiraan cuaca dan produksi energi surya.
11	Penelitian ini	Prediksi Cuaca	<i>Long Short-Term Memory (LSTM)</i>	Data penyinaran Matahari

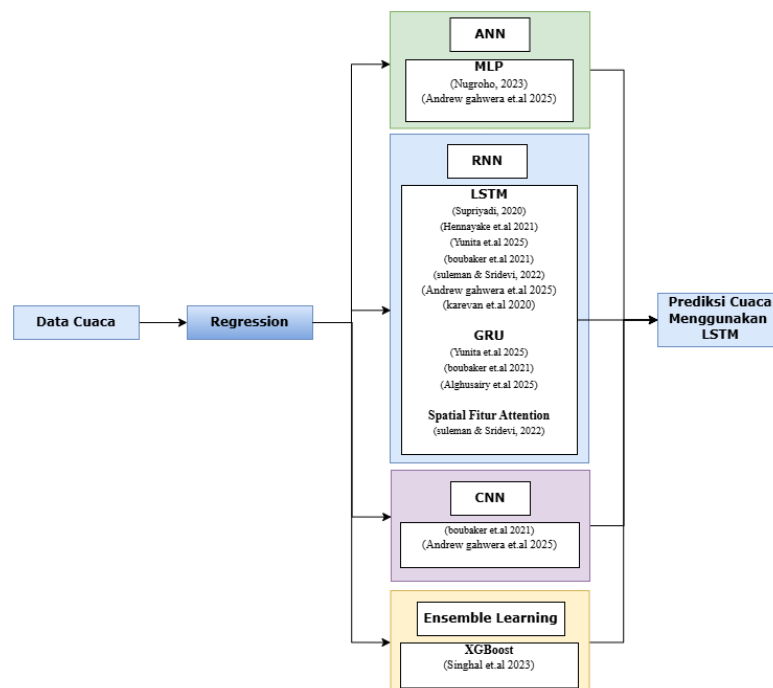
Berbagai penelitian sebelumnya menunjukkan bahwa metode jaringan syaraf tiruan dan *deep learning* telah banyak digunakan dalam bidang prediksi cuaca maupun radiasi matahari. Nugroho (2023) menggunakan *Multi-Layer Perceptron (MLP)* untuk memprediksi lama penyinaran matahari di Kota Bandung dengan data tahun 2016–2021, sedangkan Supriadi (2020) menerapkan *Long Short-Term Memory (LSTM)* pada data sinoptik Stasiun Meteorologi Maritim Tanjung Priok untuk memprediksi parameter cuaca. Hennayake dkk (2021) juga memanfaatkan LSTM multivariat untuk prediksi cuaca jangka pendek di Sri Lanka, sementara Yunita dkk (2025) membandingkan kinerja RNN, LSTM, dan GRU pada tiga dataset deret waktu.

Selanjutnya, Boubaker dkk (2021) mengkombinasikan LSTM, GRU, dan CNN untuk memprediksi radiasi matahari di Arab Saudi, sedangkan Suleman dan Shridevi (2022) mengembangkan model *Spatial Feature Attention-LSTM (SFA-LSTM)* untuk prediksi suhu dengan pola cuaca yang dinamis. Andrew Gahwera dkk

(2025) menerapkan pendekatan *ensemble transfer learning* berbasis multi-stasiun untuk prediksi curah hujan di Cekungan Danau Victoria, dan Alghushairy dkk (2025) menggunakan *Bidirectional GRU* untuk peramalan cuaca di Arab Saudi. Penelitian lain, seperti Karevan dan Suykens (2020), mengusulkan *Transductive LSTM (T-LSTM)*, sementara Singhal dkk (2022) menggunakan XGBoost untuk prediksi energi surya. Dari berbagai penelitian tersebut dapat dilihat bahwa model *deep learning* khususnya berbasis LSTM dan turunannya mampu memberikan hasil prediksi yang lebih baik, sehingga relevan digunakan dalam penelitian ini untuk memprediksi penyinaran matahari.

2.2 Kerangka Teori

Kerangka teori (*Theoretical framework*) yang digunakan dalam penelitian ini akan digambarkan sebagai berikut:



Gambar 2.1 Kerangka Teori

Kerangka teoritis diatas menggambarkan bagaimana data cuaca yang bersifat deret waktu diproses menggunakan berbagai metode pembelajaran mesin dan pembelajaran mendalam untuk menghasilkan prediksi cuaca. Data cuaca tersebut digunakan sebagai input, kemudian dianalisis dengan beberapa algoritma, antara lain *Long Short-Term Memory (LSTM)*, *Recurrent Neural Network (RNN)*, *Gated Recurrent Unit (GRU)*, *Convolutional Neural Network (CNN)*, *Multilayer Perceptron (MLP)*, serta *Extreme Gradient Boosting (XGBoost)*. LSTM dan GRU dipilih karena kemampuannya dalam menangkap pola temporal jangka panjang, sedangkan RNN merepresentasikan model dasar untuk data sekuensial. CNN digunakan untuk mengekstraksi fitur spasial-temporal, sementara MLP dan XGBoost berfungsi sebagai pembanding berbasis model non-sekuensial. Hasil dari seluruh metode tersebut adalah prediksi cuaca, baik dalam bentuk nilai kuantitatif maupun kategori kondisi cuaca, yang kemudian dievaluasi untuk menentukan model paling optimal dalam konteks data cuaca tropis di Indonesia.

2.3 Time Series Data

Data deret waktu (*time series data*) adalah Kumpulan data atau urutan pengamatan yang tersusun secara kronologis berdasarkan waktu, dan dicatat pada interval waktu yang umumnya seragam (Ali dkk, 2019). Berdasarkan strukturnya, data deret waktu dapat diklasifikasikan menjadi empat jenis, yaitu:

- 1) *Univariate*: Data yang terdiri atas satu variabel yang diamati secara berkala sepanjang waktu.
- 2) *Multivariate*: Data yang mencakup beberapa variabel yang diamati secara bersamaan pada setiap titik waktu.

- 3) *Tensor fields*: data dengan dimensi yang lebih kompleks, seperti informasi posisi spasial dari objek yang bergerak seiring waktu.
- 4) *Multifield*: gabungan dari berbagai jenis data yang memiliki struktur atau format berbeda namun diamati dalam kerangka waktu yang sama.

Karakteristik data deret waktu merujuk pada sifat-sifat khusus yang membedakan data ini dari jenis data lainnya. Menurut Cendrero dkk (2022), terdapat lima karakteristik utama data deret waktu, yaitu:

- 1) Tersusun secara kronologis (*time-ordered*)

Data deret waktu disusun berdasarkan urutan waktu, seperti harian, mingguan, bulanan, atau tahunan, sehingga setiap observasi memiliki indeks waktu yang jelas.

- 2) Ketergantungan terhadap waktu (*temporal dependence*)

Nilai pada satu titik waktu dalam data deret waktu biasanya dipengaruhi oleh nilai pada waktu sebelumnya, sehingga observasi dalam deret waktu tidak bersifat independent.

- 3) Bersifat stasioner atau tidak stasioner

- Stasioner: data dikatakan stasioner jika karakteristik statistiknya seperti rata-rata dan varians, tetap konstan sepanjang waktu.
- Tidak stasioner: data menunjukkan perubahan karakteristik statistic dari waktu ke waktu, yang dapat disebabkan oleh adanya tren atau siklus.

4) Dapat dipengaruhi oleh *factor* eksternal

Nilai dalam data deret waktu sering kali dipengaruhi oleh variabel luar (eksternal), seperti peristiwa berita, kebijakan pemerintah, atau perubahan kondisi ekonomi.

5) Dapat memiliki pola musiman atau tren

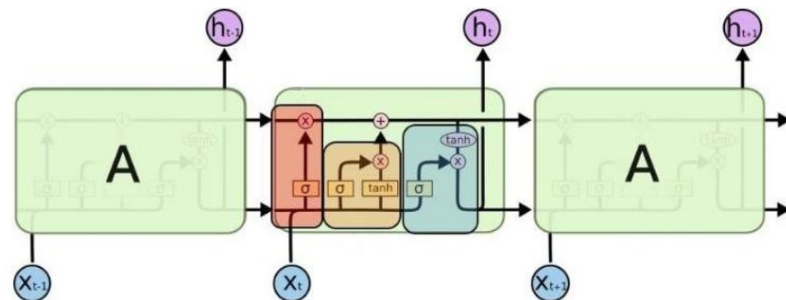
Data deret waktu dapat menunjukkan pola musiman (*seasonal*) yang berulang secara berkala, serta tren jangka panjang yang menunjukkan arah kenaikan atau penurunan nilai dari waktu ke waktu.

2.4 LSTM (Long Short-Term Memory)

Long Short-Term Memory (LSTM) merupakan salah satu jenis *Recurrent Neural Network* (RNN) yang dikembangkan oleh Hochreiter dan Schmidhuber pada tahun 1997. RNN mampu memproses data sekuensial yang panjang dan kompleks. RNN biasanya menggunakan model *backpropagation* yang mirip dengan jaringan saraf, dan rantai yang terlalu panjang dalam topologi. RNN cenderung mengalami penurunan gradien (*vanishing gradient*) (Kristian dkk, 2018: 311; Khumaidi & Nirmala, 2022:8).

Long Short-Term Memory (LSTM) dikembangkan untuk mengatasi *vanishing gradient* pada RNN (Kristian dkk, 2018: 311; Khumaidi & Nirmala, 2022:8). LSTM telah dimodifikasi dengan menambahkan *memory cell* sehingga dapat menyimpan informasi dalam jangka waktu yang lebih lama. Keuntungan dari LSTM adalah adanya *memory block* yang menentukan nilai mana yang akan dipilih sebagai *output* yang relevan untuk *input* yang diberikan. (Wiranda dkk, 2019: 186;

Khumaidi & Nirmala, 2022:8) Contoh arsitektur LSTM dapat dilihat pada gambar 2.2.



Gambar 2.2 Architecture LSTM (Supriyadi, 2020)

Arsitektur LSTM terdiri dari lapisan *input*, lapisan *output*, dan lapisan tersembunyi. Lapisan tersembunyi terdiri dari *memory cell*, masing-masing memori terdiri dari gerbang masukan (*input gate*), gerbang lupa (*forget gate*) dan gerbang keluaran (*output gate*) (Wiranda dkk, 2019:189; Khumaidi & Nirmala, 2022:8).

Masing-masing gerbang yang termasuk pada LSTM dijelaskan sebagai berikut:

a. Forget Gate (f_t)

Forget gate berfungsi untuk mengontrol seberapa baik nilai dipertahankan dalam *memory cell* (Wiranda dkk, 2019: 189; Khumaidi & Nirmala, 2022:9). *Forget gate* adalah lapisan *sigmoid* yang mengambil nilai *output* pada waktu $t - 1$ dan nilai *input* pada waktu t , yang kemudian digabungkan dan diterapkan pada fungsi aktivasi *sigmoid*. Di mana hasil *output sigmoid* bernilai 0 dan 1 (Wildan dkk, 2018: 3550; Khumaidi & Nirmala, 2022:9). Jika nilai $f_t = 1$ maka state sebelumnya tidak akan berubah dan data akan disimpan, sedangkan jika $f_t = 0$ maka state sebelumnya akan dilupakan (Wiranda dkk, 2019: 190; Khumaidi & Nirmala, 2022:9) Persamaan dari f_t yaitu:

$$f_t = \sigma (W_f S_{t-1} + W_f X_t) \quad (1)$$

Keterangan:

W_f = Bobot dari *Forget Gate*.

S_{t-1} = *State* sebelumnya atau *state* pada waktu $t - 1$.

X_t = *Input* pada waktu t .

σ = Fungsi aktivasi *sigmoid*.

b. Input Gate I_t

Peran *input gate* adalah untuk mengontrol sejauh mana nilai baru mengalir di dalam *cell*. Ini dirancang untuk menghindari penyimpanan data yang tidak perlu.

Input gate mengambil nilai *output* sebelumnya dan nilai *input* baru serta melewati lapisan *sigmoid*. Nilai *gate* ini dikembalikan sebagai menjadi 0 atau 1

(Wiranda dkk, 2019: 189; Khumaidi & Nirmala, 2022:9) Persamaan dari I_t yaitu:

$$I_t = \sigma (W_i S_{t-1} + W_i X_t) \quad (2)$$

Keterangan:

W_i = Bobot dari *Input Gate*.

S_{t-1} = *State* sebelumnya atau *state* pada waktu $t - 1$.

X_t = *Input* pada waktu t .

σ = Fungsi aktivasi *sigmoid*.

Kemudian kalikan nilai *input gate* dengan nilai *output gate* dari lapisan kandidat $\check{C}$

(Wiranda dkk, 2019: 190; Khumaidi & Nirmala, 2022:9) Persamaan dari $\check{C}$ yaitu:

$$\check{C} = \tanh(W_c S_{t-1} + W_c X_t) \quad (3)$$

$$C_t = (i_t * \check{C}_t + f_t * C_{t-1}) \quad (4)$$

Keterangan:

$\check{C}$ = *Intermediate cell state*.

W_c = Bobot dari *cell state*.

S_{t-1} = *State* sebelumnya atau *state* pada waktu $t - 1$

X_t = *Input* pada waktu t

c. Output Gate (O_t)

Output Gate digunakan untuk mengendalikan seberapa banyak nilai yang berada di dalam *memory cell* yang berfungsi untuk menghitung nilai *output* (Wiranda dkk, 2019: 190;Khumaidi & Nirmala, 2022:10). Rumus dari (O_t) yaitu:

$$O_t = \sigma (W_o S_{t-1} + W_o X_t) \quad (5)$$

Keterangan:

W_t = Bobot dari *output Gate*

S_{t-1} = *State* sebelumnya atau *state* pada waktu $t - 1$.

X_t = *Input* pada waktu t .

σ = Fungsi aktivasi *sigmoid*.

2.5 Evaluasi Model

Evaluasi model dalam penelitian ini bertujuan untuk mengukur tingkat akurasi dan efektivitas model *Long Short-Term Memory* (LSTM) dalam memprediksi cuaca harian berdasarkan data penyinaran matahari. Tahap ini menjadi krusial untuk memastikan bahwa model yang dikembangkan dapat memberikan hasil prediksi yang andal dan konsisten. Penelitian ini menggunakan dua metrik utama untuk mengevaluasi kinerja model:

- 1) **Mean Absolute Error (MAE):** MAE mengukur rata-rata selisih absolut antara nilai aktual dan nilai prediksi (Shi dkk, 2024). Metrik ini memberikan gambaran yang jelas tentang seberapa besar kesalahan prediksi secara umum, tanpa terpengaruh oleh arah selisihnya. MAE cenderung lebih stabil dan tidak terlalu sensitif terhadap data pencilan (*outlier*). Nilai MAE yang rendah mengindikasikan tingkat kesalahan prediksi yang rendah (Hodson, 2022). Rumus yang digunakan adalah:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (6)$$

Keterangan:

n = jumlah data
 y_i = nilai aktual
 $\hat{y}_i$ = nilai prediksi

- 2) **Root Mean Squared Error (RMSE):** RMSE menghitung akar dari rata-rata kuadrat selisih antara nilai aktual dan prediksi (Shi dkk, 2024). Metrik ini memberikan penekanan lebih besar pada kesalahan yang signifikan karena selisihnya dikuadratkan sebelum dirata-ratakan. Oleh karena itu, RMSE lebih sensitif terhadap data pencilan. Nilai RMSE yang rendah menunjukkan bahwa prediksi model konsisten mendekati nilai aktual (Hodson, 2022). Rumus yang digunakan adalah:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (7)$$

Keterangan:

n = jumlah data
 y_i = nilai aktual
 $\hat{y}_i$ = nilai prediksi

- 3) **Koefisien Determinasi (R^2_Score):** metrik statistik yang digunakan dalam analisis regresi untuk mengukur seberapa baik model prediksi mampu menjelaskan variasi dari data aktual (variabel dependen). Nilai R^2 mencerminkan proporsi total variabilitas data aktual yang berhasil ditangkap atau diprediksi oleh fitur-fitur input (variabel independen) di dalam model (Chicco dkk, 2021). Rumus yang digunakan adalah:

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \quad (8)$$

keterangan:

y_i = Nilai aktual (data observasi di lapangan) pada data ke-i.

$\hat{y}_i$ = Nilai prediksi yang dihasilkan oleh model pada data ke-i.

y = Nilai rata-rata dari seluruh data aktual y .

N = Jumlah sampel data.

- Nilai $R^2 = 1$: Model sempurna. Semua variasi data aktual dapat dijelaskan 100% tanpa kesalahan oleh prediksi model.
- Nilai $R^2 = 0$: Model tidak lebih baik daripada tebakan rata-rata. Model gagal menangkap hubungan atau pola runtun waktu/fitur sama sekali.
- Nilai $R^2 < 0$ (Negatif): Model sangat buruk (bahkan lebih buruk daripada sekadar menarik garis rata-rata). Ini biasanya terjadi jika model *neural network* belum konvergen, salah arsitektur, atau salah melakukan de-normalisasi data.

BAB III

METODOLOGI PENELITIAN

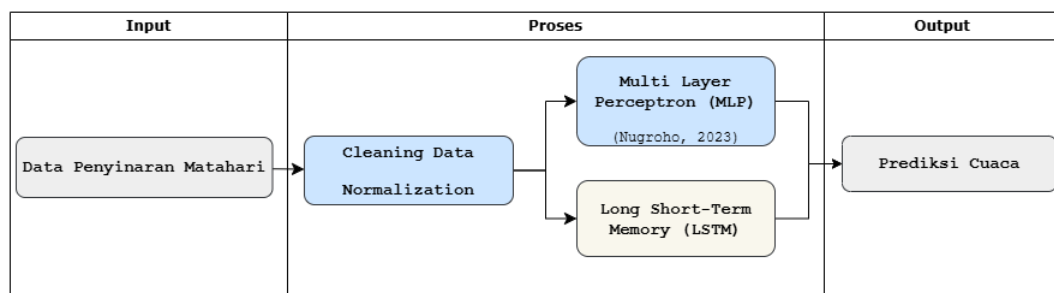
3.1 Kerangka Konsep

Kerangka konsep dalam penelitian ini disusun untuk menggambarkan hubungan antara variabel-variabel yang terlibat dalam proses prediksi cuaca berbasis data penyinaran matahari menggunakan pendekatan *Long Short-Term Memory* (LSTM). Penelitian ini berangkat dari pemahaman bahwa intensitas penyinaran matahari merupakan salah satu indikator penting dalam menentukan kondisi cuaca, khususnya dalam memprediksi variabel-variabel seperti suhu, kelembaban, dan kemungkinan terjadinya hujan. Dalam konteks ini, data penyinaran matahari digunakan sebagai variabel input utama yang merepresentasikan pola radiasi matahari yang diterima permukaan bumi dalam kurun waktu tertentu. Data ini bersifat deret waktu (*time-series*) dan memiliki dinamika yang kompleks serta bersifat non-linear, sehingga diperlukan model prediksi yang mampu menangkap pola jangka panjang dan jangka pendek secara adaptif.

Model LSTM dipilih sebagai metode utama dalam penelitian ini karena memiliki keunggulan dalam mengolah data deret waktu dengan mempertahankan informasi historis yang relevan melalui mekanisme memori jangka panjang dan pendek (Salman dkk, 2018). Dengan struktur jaringan saraf berlapis dan unit-unit memori internal, LSTM mampu mengidentifikasi keterkaitan temporal dalam data penyinaran matahari untuk menghasilkan prediksi cuaca yang lebih akurat. Dalam kerangka ini, data *input* berupa nilai-nilai intensitas penyinaran matahari akan

diproses melalui model LSTM untuk menghasilkan *output* berupa prediksi cuaca. Proses ini juga mencakup tahap *preprocessing* data, pelatihan model, validasi, dan evaluasi performa menggunakan metrik seperti *Mean Absolute Error (MAE)* dan *Root Mean Squared Error (RMSE)*.

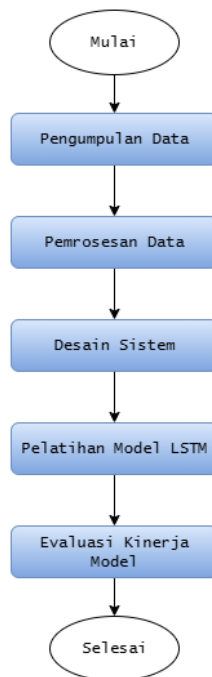
Secara keseluruhan, kerangka konsep ini menggambarkan hubungan kausal dan fungsional antara *input* (data penyinaran matahari), metode pemrosesan (LSTM), dan *output* (prediksi cuaca), serta memperjelas alur kerja penelitian dari pengumpulan data hingga interpretasi hasil. Dengan pendekatan ini, diharapkan penelitian dapat memberikan kontribusi dalam pengembangan sistem prediksi cuaca yang lebih efisien dan responsif terhadap dinamika lingkungan, khususnya dalam konteks wilayah tropis seperti Indonesia yang memiliki karakteristik cuaca yang sangat dipengaruhi oleh radiasi matahari. Berikut merupakan gambar kerangka konsep pada penelitian ini.



Gambar 3.1 Kerangka Konsep Penelitian

3.2 Desain Penelitian

Desain penelitian ini menggunakan pendekatan kuantitatif dengan metodologi pemodelan prediktif *deep learning*, mengikuti alur kerja yang terstruktur dan sistematis, mulai dari akuisisi data hingga evaluasi model akhir. Tahapan penelitian dapat dilihat pada gambar 3.2.



Gambar 3.2 Desain Penelitian

3.2.1 Pengumpulan Data

Penelitian ini memulai dengan tahap pengumpulan data. Data yang akan digunakan dalam penelitian ini merupakan data sekunder. Data sekunder menurut (Sulung & Muspawi, 2024) adalah sumber data penelitian yang tidak diperoleh secara langsung dari objek penelitian melainkan berasal dari dokumen atau media perantara yang telah ada sebelumnya, misalnya buku, laporan, arsip, publikasi, atau sumber lain yang dipublikasikan oleh pihak ketiga. Data yang akan digunakan dalam penelitian ini diperoleh dari situs resmi Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) <https://dataonline.bmkg.go.id/> periode Januari - Desember 2023. Variabel yang digunakan adalah variabel (tanggal, sr avg, sr max dan durasi) yang relevan terhadap permasalahan penelitian. Data yang terkumpul selanjutnya akan direkam dalam format digital dan dijamin memiliki metadata yang

mendukung (waktu, sumber, definisi variabel) agar dapat diproses secara akurat.

Table dataset terdapat pada table 3.1.

Tabel 3.1 Dataset Penyinaran Matahari

No	Date	Sr Avg	Sr Max	Duration
1	1/1/2023	258.2136364	787.1	2.338
2	1/2/2023	215.3416667	3.507	3.507
3	1/3/2023	361.2409091	1107.7	9.686
4	1/4/2023	538.1804878	1203.8	6.847
				
727	1/1/2023	353.759234375	1167	10.688
728	1/2/2023	300.217738461538	1205	10.02
729	1/3/2023	400.3521538	1262	10.855
730	1/4/2023	475.5720909	1274	11.022

Keterangan:

Date: Tanggal Pengamatan.

Sr avg: *Solar Radiation Average* (rata-rata radiasi matahari yang diterima).

Sr max: *Solar Radiation Maximum* (nilai maksimum radiasi matahari).

Duration: Durasi atau lama waktu pengamatan atau pencatatan.

3.2.2 Preprocessing Data

Sebelum data dimasukkan ke dalam model pembelajaran, dilakukan serangkaian langkah *preprocessing* untuk meningkatkan kualitas data dan mempersiapkannya agar sesuai kebutuhan model. Berdasarkan literatur terkini, preprocessing memegang peranan penting dalam meningkatkan kinerja model LSTM dan *model deep learning* lainnya (Lazcano & Jaramillo-Morán, 2025).

Berikut tahap-tahap yang dilakukan:

1) Cleaning Data

- Penanganan nilai hilang (*missing values*), misalnya dengan imputasi (mean, median, atau metode interpolasi) atau jika diperlukan, penghapusan baris/kolom yang sangat banyak kegagalannya.
- Penghapusan data duplikat dan pemeriksaan data tidak konsisten, misalnya data yang bertentangan antar sumber.

- Penanganan *outliers*, jika ditemukan nilai yang sangat ekstrim yang dapat merusak model, melalui teknik seperti *winsorizing*.
- Literatur menunjukkan bahwa data *cleaning* signifikan menurunkan *error* dan meningkatkan keandalan model dalam domain seperti deteksi anomali dan prediksi waktu nyata (Dash dkk, 2025).

2) Normalisasi Data (Normalization)

Setelah data bersih, variabel-variabel dengan skala berbeda dinormalisasi agar berada dalam rentang yang seragam agar adil kontribusinya dalam pelatihan model. Penelitian-penelitian terbaru menyebutkan bahwa teknik normalisasi (misalnya *Min-Max scaling*, *Z-score normalization*, atau teknik lainnya) dapat berdampak signifikan terhadap akurasi model, terutama untuk jaringan saraf dalam (*deep neural networks*) dan LSTM (Kim dkk, 2025). Berikut merupakan rumus yang digunakan:

Rumus:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (8)$$

Keterangan:

X = nilai asli dari data

Xmin = nilai minimum dalam data

Xmax = nilai maksimum dalam data

Xnorm = nilai yang telah dinormalisasi, yang akan berada dalam rentang [0,1].

Nilai prediksi iradiasi matahari rata-rata yang dihasilkan oleh model LSTM dalam skala numerik 0-1 kemudian dikonversi menjadi informasi kategorikal cuaca berdasarkan ambang batas absolut yang telah ditentukan pada table 3.2.

Tabel 3.2. Threshold Perkiraan Berdasarkan Iradiasi Absolut (W/m^2)

Kategori Cuaca	Threshold (W/m^2)	Dasar Ilmiah/Standar	Referensi	Source & notes
Cerah (Clear Sky)	$> 700 \text{ W/m}^2$	Solar Engineering of Thermal Processes" oleh Duffie, J.A., and Beckman, W.A.(BECMAN, 1980)		STC = 1000 W/m^2 Practical = 700 W/m^2
Berawan (Cloudy)	$\sim 120 - 700 \text{ W/m}^2$	A review on the assessment of solar radiation models" (oleh M. Jamil Ahmad, G.N. Tiwari) (Â & Tiwari, 2011)		Awan Tipis (Cirrus): Mungkin hanya mengurangi iradiasi sedikit, sehingga nilainya masih bisa di atas $500\text{-}600 \text{ W/m}^2$. Awan Tebal (Cumulus/Stratus): Dapat mengurangi iradiasi secara drastis hingga di bawah 300 W/m^2 . Implikasi: Rentang antara ambang batas bawah WMO (120 W/m^2) dan ambang batas atas kondisi cerah (sekitar 700 W/m^2) secara logis menjadi representasi untuk kondisi berawan.
Mendung / Hujan	$< 120 \text{ W/m}^2$	WMO 8 Ch (Organization, 2024)		Definisi "Sunshine Duration" dari World Meteorological Organization (WMO)

3) Pemisahan Data (Data Splitting)

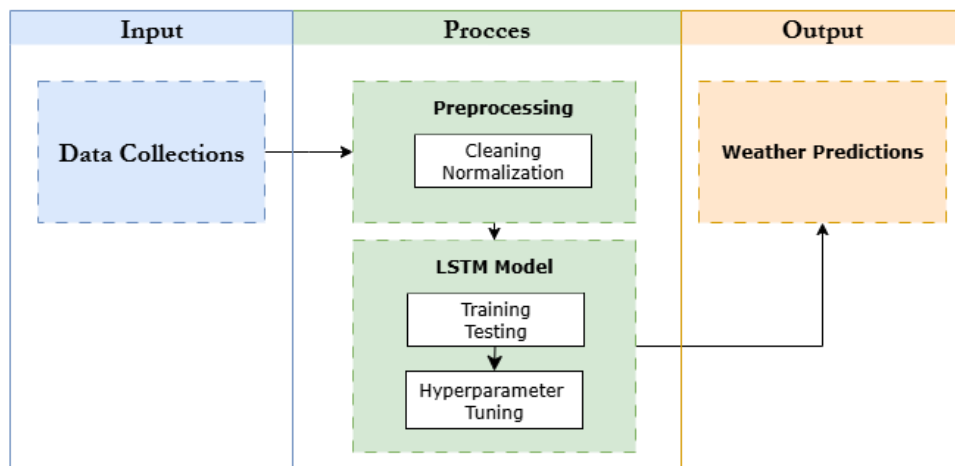
Data dibagi menjadi set pelatihan (*training set*) dan set pengujian (*testing set*).

Rasio umum yang digunakan di literatur antara 70:30, 80:20 atau 90:10 (Shi dkk, 2024).

3.2.3 Desain Sistem

Desain sistem dalam penelitian ini dirancang untuk membentuk alur kerja yang terintegrasi mulai dari pengumpulan data hingga prediksi cuaca menggunakan

model *Long Short-Term Memory* (LSTM). Sistem ini terdiri dari beberapa komponen utama, yaitu pengumpulan data, pra-pemrosesan data dan pelatihan model LSTM.



Gambar 3.3 Desain Sistem

Diagram tersebut dibagi menjadi tiga bagian utama: input, proses dan output. Tahapan tersebut disusun secara sistematis untuk memastikan data yang digunakan valid, model yang diterapkan sesuai, serta hasil penelitian dapat dievaluasi secara objektif. Rincian dari masing-masing tahapan adalah sebagai berikut:

1. Input (*Data Collections*)

Data diperoleh dari situs resmi BMKG (<https://dataonline.bmkg.go.id/>) sebagai sumber utama informasi cuaca yang relevan untuk kebutuhan penelitian.

2. Proses

- a. Preprocessing: meliputi *cleaning data* untuk mengatasi nilai hilang, duplikasi, dan inkonsistensi, serta *normalization data* agar setiap variabel berada pada skala seragam.

- b. Pemodelan dengan LSTM: data hasil *preprocessing* dilatih dan diuji menggunakan model LSTM yang unggul dalam pengolahan data berurutan (*time series*), serta menentukan *hyperparameter tuning*.

3. Output (*Weather Prediction*)

Prediksi cuaca merupakan luaran (*output*) berupa nilai estimasi atau perkiraan dari parameter cuaca yang diuji. Kualitas dari hasil prediksi ini dievaluasi secara kuantitatif menggunakan metrik seperti MAE (*Mean Absolute Error*) dan RMSE (*Root Mean Squared Error*) yang berfungsi untuk mengukur performa model.

3.2.4 Pemodelan dengan LSTM

Model yang akan digunakan adalah *Long Short-Term Memory (LSTM)* tipe RNN yang memiliki sel memori dan mekanisme *gates* untuk menjaga informasi penting dalam urutan (*sequence*) serta untuk mengatasi permasalahan *vanishing gradient*. Pemilihan LSTM didasarkan pada literatur terbaru yang menunjukkan bahwa LSTM sangat efektif pada tugas prediksi deret waktu (*time-series forecasting*), pengenalan aktivitas manusia berbasis sensor, dan masalah yang memerlukan pemahaman dependensi jangka panjang (Khan & Hossni, 2025).

Langkah-langkah pemodelan:

- 1) Menentukan arsitektur LSTM, termasuk jumlah lapisan (*layers*), jumlah unit tiap layer, fungsi aktivasi, dan parameter lainnya (seperti regularisasi).

- 2) Melakukan *training* model dengan data yang sudah dibersihkan dan dinormalisasi. Proses training menggunakan algoritma optimisasi Adam dengan monitoring fungsi loss dan metrik lain untuk menghindari *overfitting*.
- 3) Melakukan *testing* atau prediksi terhadap data uji yang tidak pernah dilihat selama training, untuk mengukur kemampuan generalisasi model.

3.3 Skenario uji coba

Skenario uji coba dalam penelitian ini disusun untuk mengevaluasi kinerja model LSTM dalam memprediksi kondisi cuaca harian berbasis data penyinaran matahari. Evaluasi kinerja model dilakukan menggunakan tiga metrik utama, yaitu *Mean Absolute Error* (MAE), *Root Mean Square Error* (RMSE) serta R^2 (*Koefisien determinasi*). Tahapan uji coba dilakukan melalui beberapa tahap seperti pada tabel 3.2.

Tabel 3.3 Rancangan Skenario Uji Coba

No	Scenario Uji Coba	Variasi/Parameter	Tujuan
1	<i>Baseline</i>	1 LSTM (64 unit), Fully Connected.	Menjadi acuan kinerja dasar model.
2	Kedalaman jaringan	1, 2, 3 lapis LSTM	Menilai pengaruh jumlah lapisan terhadap kinerja prediksi.
3	Kapasitas jaringan	Jumlah unit = 16, 32, 64	Menguji pengaruh kompleksitas arsitektur terhadap akurasi model.
4	Panjang urutan (<i>sequence length</i>)	7, 14, 30 hari	Menilai pengaruh panjang data historis pada kualitas prediksi.
5	Regularisasi	<i>Dropout</i> = 0/0,2; <i>Recurrent dropout</i> = 0 / 0,2	Mengevaluasi dampak regularisasi terhadap <i>overfitting</i> .
6	Optimisasi	<i>Adam optimizer</i> , <i>Learning Rate</i> = 0,001 / 0,0005	Menemukan tingkat pembelajaran yang paling stabil.
7	Batch size & Early stopping	Batch size = 8, 16, 32; Epoch = maks. 500; <i>patience</i> = 10	Menguji pengaruh ukuran <i>batch</i> dan mencegah pelatihan berlebihan.
8	Praproses data	Interpolasi data hilang; normalisasi.	Menjamin data masukan bersih dan konsisten.
9	Post-processing kategori cuaca	Threshold berbasis iradiasi absolut (w/m2)	Mengklasifikasikan data iradiasi absolut (w/m2) ke dalam tiga kategori cuaca: cerah, berawan dan hujan.

No	Scenario Uji Coba	Variasi/Parameter	Tujuan
10	Uji kestabilan	3 kali ulangan dengan <i>random seed</i> berbeda	Mengukur variabilitas hasil dan kestabilan model.

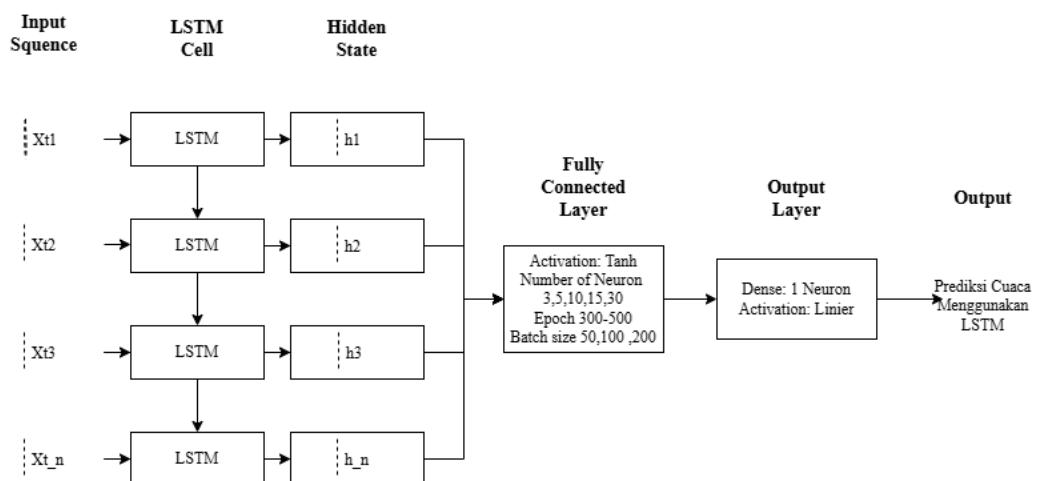
Rancangan skenario uji coba ini disusun untuk mengevaluasi kinerja model LSTM dalam memprediksi kondisi cuaca harian berbasis data penyinaran matahari. Uji coba dilakukan melalui beberapa variasi, mulai dari arsitektur dasar (*baseline*), kedalaman dan kapasitas jaringan, panjang urutan data historis, hingga penerapan regularisasi. Proses optimisasi menggunakan algoritma Adam dengan variasi *learning rate*, sedangkan pengaturan pelatihan melibatkan perbedaan *batch size*, jumlah *epoch*, serta *early stopping*. Seluruh data diproses melalui tahap interpolasi dan normalisasi sebelum digunakan. Setiap konfigurasi dijalankan sebanyak tiga kali dengan *random seed* berbeda untuk memastikan konsistensi, sedangkan hasil evaluasi dianalisis menggunakan metrik MAE, RMSE dan R^2_Score .

BAB 1V

HASIL DAN PEMBAHASAN

4.1 Arsitektur Model LSTM

Arsitektur model LSTM seperti pada gambar 4.1 menunjukkan penerapan *Long Short-Term Memory* untuk melakukan prediksi cuaca berdasarkan data runtun waktu penyinaran matahari, yaitu SR_avg (rata-rata penyinaran), SR_max (penyinaran maksimum), dan durasi penyinaran. Data tersebut disusun dalam bentuk urutan waktu (*time series*), Dimana setiap *time step* x_t merepresentasikan kondisi penyinaran matahari pada hari ke-t dengan tiga variabel tersebut. Setiap input diproses secara berurutan oleh LSTM cell yang memiliki mekanisme *memory internal* melalui *forget gate*, *input gate*, dan *output gate*. Mekanisme ini memungkinkan untuk menyimpan informasi penting dari periode sebelumnya serta menangkap pola temporal, tren, dan perubahan kondisi atmosfer yang tercermin dari variasi penyinaran matahari. Arsitektur LSTM dapat dilihat pada gambar 4.1



Gambar 4.1 Arsitektur LSTM

Keluaran dari setiap LSTM *cell* direpresentasikan dalam bentuk *hidden state* ($h_1, h_2, \dots, h_n$), yang berfungsi sebagai ringkasan informasi historis hingga waktu tertentu. *Hidden state* terakhir mengandung representasi pola keseluruhan dari data runtun waktu dan diteruskan ke *fully connected layer* (*dense layer*). Pada bagian ini dilakukan transformasi *non-linear* menggunakan sejumlah *neuron* tertentu dengan fungsi aktivasi *tanh* serta penerapan *dropout* untuk mengurangi resiko *overfitting*. Lapisan ini berperan dalam mengoptimalkan pemetaan antara pola temporal yang telah dipelajari oleh LSTM dan target prediksi cuaca.

Tahap terakhir dari arsitektur adalah *output layer* yang terdiri atas satu *neuron* dengan fungsi aktivasi *linear*. Lapisan ini menghasilkan satu nilai prediksi yang merepresentasikan kondisi cuaca yang diperkirakan berdasarkan pola penyinaran matahari sebelumnya. Secara keseluruhan, kombinasi LSTM dan *dense layer* dalam model ini memungkinkan sistem untuk mempelajari hubungan *non-linear* dan dinamis antara variabel penyinaran matahari dan kondisi cuaca, sehingga model mampu memberikan prediksi cuaca yang berbasis pada pola historis data meteorologi.

4.2 Pre-Processing Data (Persiapan Data)

Persiapan data meliputi data *cleaning* untuk menghilangkan *noise*, normalisasi untuk menyamakan skala, serta transformasi *windowing* ke format *supervised learning*. Tahapan ini bertujuan agar model LSTM dapat mempelajari pola temporal dari data historis secara optimal dan akurat. Tabel 4.1 menyajikan data mentah sebelum melalui tahap pra-pemrosesan (*preprocessing*).

Tabel 4.1. Persiapan Data

Tanggal	Sr_avg	Sr max	Duration
1/1/2023	258.213636	787.1	2.338
1/2/2023	215.341667	489.6	3.507
1/3/2023	361.240909	1107.7	9.686
1/4/2023	538.180488	1203.8	6.847
1/5/2023	660.191176	1316.1	5.678
1/6/2023	482.268085	1205.8	7.682
1/7/2023	528.494444	1459.2	6.012
1/8/2023	433.253333	1012.7	2.505
1/9/2023	331.266667	1028.9	4.342
1/10/2023	420.221212	1140.5	4.509

4.2.1 Cleaning Data

Tahap awal pemrosesan data dilakukan melalui prosedur data *cleaning* untuk menjamin kualitas dan reliabilitas dataset sebelum diolah lebih lanjut. Proses ini melibatkan identifikasi serta penanganan terhadap nilai yang hilang (*missing values*), penghapusan data ganda (*duplicates*), serta deteksi data anomali atau *outlier* yang tidak merepresentasikan pola sistematis. Tujuan utama dari tahapan ini adalah untuk mereduksi *noise* yang dapat mendistorsi kemampuan belajar model, sehingga informasi yang masuk ke dalam arsitektur LSTM benar-benar merupakan data yang valid. Dengan memastikan dataset bersih dari gangguan administratif maupun teknis, model dapat melakukan ekstraksi fitur secara lebih presisi, yang pada akhirnya akan meningkatkan generalisasi dan akurasi prediksi pada hasil akhir penelitian.

4.2.2 Normalisasi Data

Memasuki tahap pemrosesan inti, data yang telah dibersihkan ditransformasikan menggunakan teknik *Min-Max Scaling* untuk membawa seluruh variabel ke dalam rentang nilai seragam antara 0 dan 1. Standardisasi ini menjadi krusial bagi arsitektur LSTM karena fungsi aktivasi internalnya, seperti *sigmoid*

dan *tanh*, bekerja secara optimal pada rentang nilai tersebut guna mencegah saturasi gradien dan mempercepat konvergensi selama pelatihan. Setelah data berada dalam skala yang konsisten, struktur data diubah ke dalam format *supervised learning* dengan pendekatan *time-series windowing*. Format ini dirancang agar mekanisme *forget gate* dan *input gate* pada sel LSTM dapat secara efektif memilah informasi relevan dari masa lalu ($t-1$) untuk memperbarui *cell state*. Dengan merepresentasikan hubungan temporal antara urutan data historis dan target pada waktu saat ini (t), model dilatih untuk mengenali ketergantungan jangka panjang (*long-term dependencies*) yang menjadi landasan utama dalam memprediksi kondisi di masa depan secara akurat. Hasil normalisasi data dapat dilihat pada *table 4.2*.

Tabel 4.2. Hasil Normalisasi Data

Sr_avg	Sr_max	Duration
0.300687154551216	0.489290063758148	0.212121212121212
0.438731875587631	0.761444229529335	0.969696969696969
0.361374875159544	0.788666809943405	0.909090909090909
0.238745534828263	0.276165914463786	0.318181818181818
0.449541433730191	0.718962676409484	0.878787878787878

4.2.3 Transformasi Windowing

Setelah data melalui tahap normalisasi, dilakukan *transformasi windowing* untuk mengubah struktur data deret waktu yang bersifat sekuensial menjadi format *supervised learning*. Proses ini dilakukan dengan menggeser jendela waktu (*sliding window*) sepanjang *dataset* untuk membentuk pasangan *input* (x) dan target (y). Secara teknis, sejumlah n langkah waktu sebelumnya (*timesteps*) dikelompokkan sebagai fitur masukan untuk memprediksi nilai pada langkah waktu berikutnya. Transformasi ini sangat krusial bagi arsitektur LSTM, karena memungkinkan model untuk mengekstraksi ketergantungan temporal dan memahami konteks

urutan data. Dengan mendefinisikan *window size* yang tepat, model dapat mempelajari seberapa jauh informasi historis memengaruhi perubahan data di masa depan, sehingga LSTM mampu menangkap tren, musiman, maupun fluktuasi jangka pendek yang terdapat dalam data penelitian.

4.2.4 Pembagian Data (Data Splitting)

Pembagian data merupakan tahapan krusial dalam pembangunan model berbasis *Deep Learning* seperti *Long Short-Term Memory* (LSTM). Pada penelitian ini, dataset yang telah melalui tahap prapemrosesan dibagi menjadi dua bagian utama, yaitu data latih (*training data*) dan data uji (*testing data*), dengan skenario rasio yang berbeda yaitu 90:10, 80:20, dan 70:30. Hasil dari pembagian data latih dan data uji dapat dilihat pada *table 4.3*.

a. Data latih

Data latih digunakan sebagai basis bagi arsitektur LSTM untuk mempelajari pola temporal, keterkaitan antar variabel, dan tren dari data historis radiasi matahari. Melalui proses iterasi (*epochs*), model melakukan optimasi bobot internal (*weights*) berdasarkan nilai *Loss Function* yang dihasilkan. Persentase data yang lebih besar pada data latih bertujuan untuk memberikan representasi fitur yang komprehensif agar model mampu menangkap kompleksitas data secara mendalam.

b. Data uji

Data uji digunakan untuk mengevaluasi kemampuan generalisasi model terhadap data baru yang belum pernah dilihat sebelumnya (*unseen data*). Hasil evaluasi pada data uji menjadi indikator performa model yang

sebenarnya. Pengujian pada berbagai rasio *split* dilakukan untuk memastikan stabilitas performa model dan menguji sensitivitas model terhadap ketersediaan volume data latih.

Tabel 4.3. Pembagian Data

Rasio	Train	Test
90:10	650	73
80:20	578	145
70:30	506	217

4.3 Hypermarameter LSTM

Tahap akhir dalam konfigurasi model adalah penentuan *hyperparameter*, yaitu parameter-parameter eksternal yang ditetapkan sebelum proses pelatihan dimulai untuk mengontrol jalannya pembelajaran serta arsitektur model LSTM. Berbeda dengan parameter internal yang dipelajari model secara otomatis, *hyperparameter* seperti jumlah *hidden layers*, jumlah *units* (neuron) pada setiap sel LSTM, *learning rate*, *batch size*, dan jumlah *epochs* ditentukan secara manual guna mendapatkan performa optimal. Penentuan *hyperparameter* yang tepat sangat krusial karena secara langsung memengaruhi kemampuan model dalam menghindari kondisi *overfitting* atau *underfitting*. Sebagai contoh, *learning rate* mengatur kecepatan model dalam memperbarui bobot berdasarkan gradien kesalahan, sementara jumlah *units* menentukan kapasitas model dalam merepresentasikan kompleksitas pola temporal. Melalui optimasi *hyperparameter*, model LSTM dapat diarahkan untuk mencapai titik konvergensi yang lebih stabil dengan tingkat kesalahan (*error*) yang paling minimum. Penentuan *hyperparameter tuning* dapat dilihat pada tabel 4.4.

Tabel 4.4 Hyperparameter Tuning (Hartono, 2025)

Model	Training Data	Hidden Layer (Neuron)	Dense	Epoch	Batch Size
A	70%	3-5	1	500 dan 300	100 dan 200
B	80%	5,10,15	1	300	200
C	90%	10-30	1	300 dan 500	100 dan 50

4.4 Implementasi dan Pelatihan Model LSTM

Pelatihan model dilakukan dengan menguji tiga scenario utama (model A, B dan C) yang dibedakan berdasarkan komposisi data latih (70%, 80% dan 90%) setiap scenario diuji dengan tiga variasi arsitektur untuk mengevaluasi pengaruh hyperparameter seperti jumlah neuron pada hidden layer dan dense layer. Pelatihan dijalankan selama 500 epoch untuk memastikan model mencapai konvergensi yang stabil. Metrik evaluasi yang digunakan untuk mengukur kinerja model adalah *Root Mean Square Error* (RMSE) dan *Mean Absolute Error* (MAE).

Tahap pelatihan model (*model training*) merupakan proses krusial di mana arsitektur LSTM secara iteratif mempelajari pola ketergantungan temporal yang kompleks di dalam data. Selama proses ini, data yang telah diformat ke dalam jendela waktu (*windowing*) diumpankan ke dalam sel-sel LSTM untuk diproses melalui mekanisme *forward propagation*. Di dalam setiap sel, informasi akan disaring oleh tiga gerbang utama: *forget gate* yang membuang informasi tidak relevan, *input gate* yang memperbarui informasi baru ke dalam *cell state*, dan *output gate* yang menentukan hasil prediksi. Model kemudian menghitung selisih antara nilai prediksi dengan nilai aktual menggunakan fungsi kerugian (*loss function*), seperti *Mean Squared Error* (MSE).

Nilai kesalahan tersebut kemudian dialirkan balik melalui mekanisme *Backpropagation Through Time* (BPTT). Tidak seperti jaringan saraf standar, BPTT pada LSTM memungkinkan model untuk memperbarui bobot berdasarkan urutan waktu, sehingga model dapat memahami hubungan antara kejadian di masa lampau dengan target saat ini. Proses ini diulangi dalam sejumlah *epoch* dengan bantuan algoritma optimasi Adam yang bertugas menyesuaikan bobot secara dinamis hingga model mencapai titik konvergensi. Tujuan akhir dari pelatihan ini adalah untuk meminimalkan fungsi kerugian sehingga model memiliki kemampuan generalisasi yang kuat dalam memproyeksikan tren data pada masa mendatang.

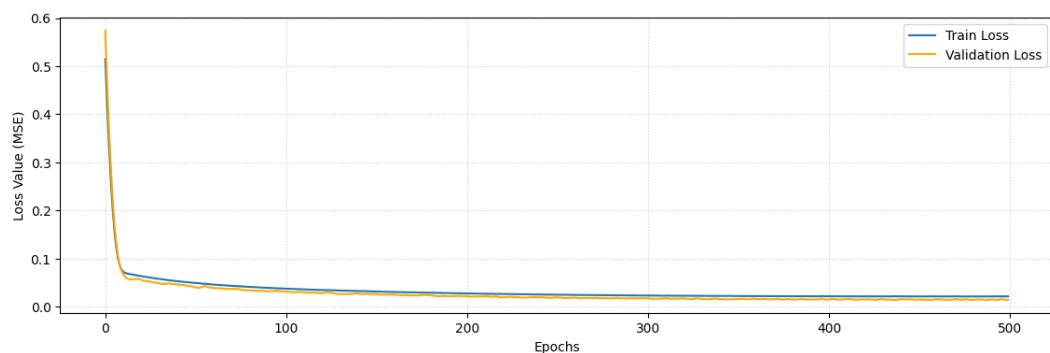
1. Model A1

Berdasarkan hasil visualisasi fungsi kerugian pada gambar 4.2. Berdasarkan grafik nilai Mean Squared Error (MSE) selama proses pelatihan model, terlihat bahwa baik *training loss* maupun *validation loss* mengalami penurunan yang sangat signifikan pada *epoch* awal. Nilai loss awal yang berada di kisaran 0.55–0.58 berhasil turun hingga mendekati 0.02 pada akhir pelatihan. Pola ini menunjukkan bahwa model mampu mempelajari hubungan dan pola yang terdapat dalam data dengan baik sehingga kesalahan prediksi semakin kecil seiring bertambahnya jumlah *epoch*.

Selain itu, kurva *training loss* dan *validation loss* menunjukkan pergerakan yang relatif berdekatan tanpa adanya jarak yang signifikan sepanjang proses pelatihan. Kondisi ini mengindikasikan bahwa model memiliki kemampuan generalisasi yang baik terhadap data yang belum pernah dilihat sebelumnya. Tidak terdapat gejala *overfitting* yang ditandai dengan meningkatnya *validation loss*

ketika *training loss* terus menurun. Sebaliknya, kedua kurva mengalami penurunan secara konsisten dan kemudian mencapai kondisi stabil (*konvergen*) pada *epoch* sekitar 250–300.

Setelah mencapai fase konvergensi, penurunan nilai *loss* cenderung melambat dan berada pada rentang yang sangat kecil. Hal ini menunjukkan bahwa model telah mencapai kondisi optimal dalam mempelajari pola data, sehingga penambahan *epoch* hingga 500 hanya memberikan peningkatan performa yang relatif kecil. Secara keseluruhan, hasil pelatihan berdasarkan metrik MSE menunjukkan bahwa model memiliki tingkat kesalahan prediksi yang rendah, proses pembelajaran yang stabil, serta kemampuan generalisasi yang baik, sehingga model layak digunakan untuk melakukan prediksi pada data penyinaran matahari.



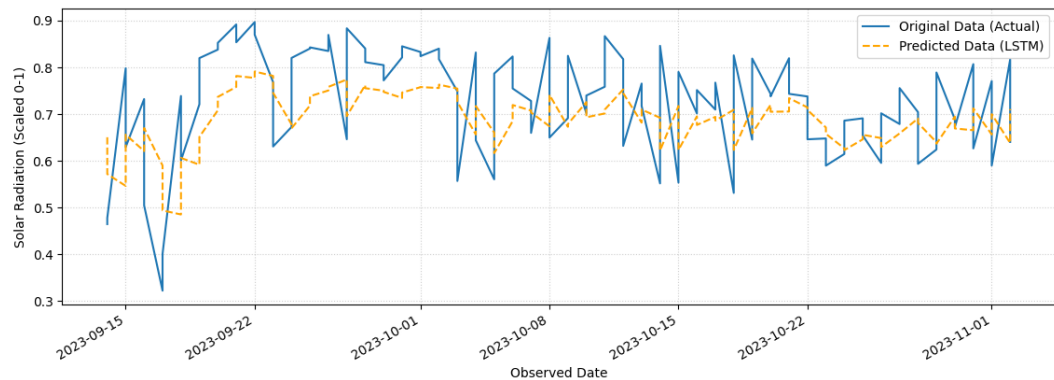
Gambar 4.2. Loss Function Model A1

Berdasarkan grafik perbandingan antara data aktual (Original Data) dan hasil prediksi model LSTM pada gambar 4.3, terlihat bahwa pola prediksi yang dihasilkan model mampu mengikuti tren umum perubahan data penyinaran matahari selama periode pengamatan. Kurva prediksi menunjukkan arah pergerakan yang relatif sejalan dengan data aktual, terutama dalam menangkap kecenderungan kenaikan dan penurunan nilai penyinaran matahari. Hal ini

mengindikasikan bahwa model LSTM berhasil mempelajari pola temporal yang terdapat dalam data historis sehingga dapat merepresentasikan karakteristik utama dari data yang diamati.

Meskipun demikian, hasil prediksi terlihat lebih halus (*smoother*) dibandingkan data aktual. Beberapa fluktuasi tajam yang muncul pada data observasi tidak sepenuhnya dapat diikuti oleh model. Kondisi ini tampak pada beberapa titik ketika data aktual mengalami kenaikan atau penurunan secara signifikan, sedangkan nilai prediksi hanya mengalami perubahan yang lebih moderat. Fenomena tersebut menunjukkan bahwa model cenderung menangkap pola umum (*trend*) dibandingkan variasi ekstrem jangka pendek, yang merupakan karakteristik umum pada model LSTM ketika digunakan untuk prediksi deret waktu dengan tingkat variabilitas yang tinggi.

Secara keseluruhan, kedekatan antara kurva prediksi dan data aktual menunjukkan bahwa model memiliki kemampuan yang baik dalam melakukan prediksi penyinaran matahari. Tidak terdapat penyimpangan yang terlalu besar antara kedua kurva, sehingga dapat disimpulkan bahwa model mampu menghasilkan prediksi yang cukup akurat dan stabil. Hasil ini juga didukung oleh grafik MSE sebelumnya yang menunjukkan nilai kesalahan yang rendah dan proses pelatihan yang konvergen, sehingga model LSTM yang dibangun dapat dianggap cukup efektif untuk memprediksi data penyinaran matahari pada periode pengujian.



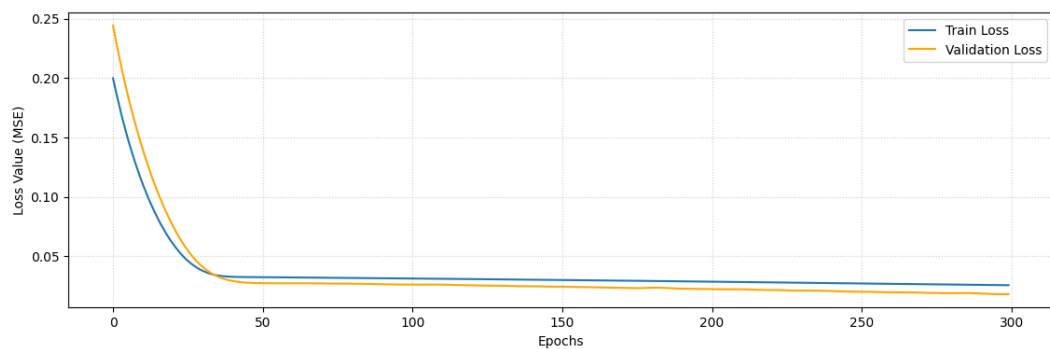
Gambar 4.3. Data Aktual vs Prediksi Data Model A1

2. Model A2

Berdasarkan hasil visualisasi fungsi kerugian pada gambar 4.4, terlihat bahwa nilai *training loss* dan *validation loss* mengalami penurunan yang sangat cepat pada *epoch* awal, yaitu dari sekitar 0.20–0.24 hingga mendekati 0.03 pada *epoch* ke-40. Penurunan yang signifikan ini menunjukkan bahwa model mampu mempelajari pola dasar yang terdapat dalam data secara efektif pada tahap awal pelatihan. Setelah *epoch* ke-40, kedua kurva menunjukkan penurunan yang lebih lambat dan cenderung stabil hingga akhir proses pelatihan pada *epoch* ke-300, yang mengindikasikan bahwa model telah mencapai kondisi konvergen.

Selain itu, jarak antara kurva *training loss* dan *validation loss* relatif kecil sepanjang proses pelatihan. Bahkan pada sebagian besar *epoch*, nilai *validation loss* berada sangat dekat dengan *training loss*. Kondisi ini menunjukkan bahwa model memiliki kemampuan generalisasi yang baik terhadap data validasi dan tidak mengalami gejala *overfitting* yang berarti. Tidak adanya peningkatan *validation loss* selama *training loss* terus menurun menunjukkan bahwa model mampu mempertahankan performa yang konsisten pada data yang tidak digunakan selama proses pelatihan.

Nilai MSE akhir yang berada pada kisaran sekitar 0.02–0.03 menunjukkan bahwa tingkat kesalahan prediksi model tergolong rendah. Hal ini mengindikasikan bahwa prediksi yang dihasilkan memiliki kedekatan yang baik terhadap nilai aktual. Dengan demikian, model yang dibangun dapat dikatakan telah mencapai performa yang stabil, akurat, dan mampu merepresentasikan pola data secara memadai untuk digunakan dalam proses prediksi.



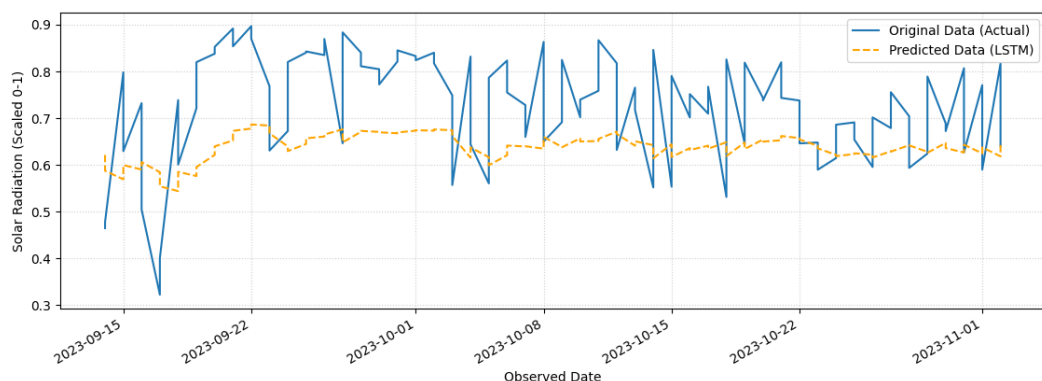
Gambar 4.4. Loss Function Model A2

Berdasarkan grafik perbandingan antara data aktual dan hasil prediksi pada gambar 4.5, terlihat bahwa model mampu mengikuti pola umum pergerakan data penyinaran matahari selama periode pengamatan. Kurva prediksi menunjukkan kecenderungan yang sejalan dengan arah perubahan data aktual, terutama dalam merepresentasikan tren jangka menengah dan jangka panjang. Hal ini mengindikasikan bahwa model LSTM berhasil mempelajari hubungan temporal yang terdapat dalam data historis sehingga mampu menghasilkan prediksi yang konsisten terhadap pola dasar penyinaran matahari.

Namun, jika diamati lebih lanjut, kurva prediksi memiliki tingkat fluktuasi yang jauh lebih rendah dibandingkan data aktual. Nilai prediksi cenderung bergerak pada rentang yang relatif sempit, yaitu sekitar 0.60–0.69, sedangkan data aktual

menunjukkan variasi yang lebih besar dengan nilai yang berfluktuasi antara sekitar 0.32 hingga 0.90. Kondisi ini menunjukkan bahwa model cenderung menghasilkan prediksi yang lebih halus (*smoothing effect*), sehingga perubahan ekstrem atau lonjakan yang terjadi pada data aktual belum dapat direpresentasikan secara optimal. Akibatnya, pada beberapa periode ketika terjadi kenaikan atau penurunan yang tajam, model menghasilkan prediksi yang lebih konservatif dan mendekati nilai rata-rata data.

Meskipun demikian, tidak terdapat penyimpangan yang sangat besar antara pola prediksi dan tren data aktual secara keseluruhan. Model masih mampu mempertahankan arah pergerakan data dan menghasilkan prediksi yang stabil sepanjang periode pengujian. Dari sudut pandang prediksi deret waktu, hasil ini menunjukkan bahwa model memiliki kemampuan yang cukup baik dalam menangkap karakteristik umum data, tetapi masih memiliki keterbatasan dalam memodelkan variabilitas tinggi yang sering muncul pada data penyinaran matahari harian.



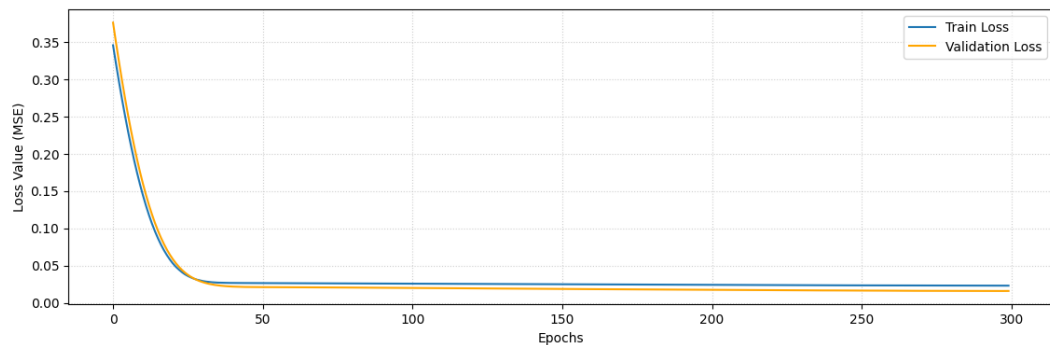
Gambar 4.5. Data Aktual vs Prediksi Data Model A2

3. Model A3

Berdasarkan grafik *Mean Squared Error* (MSE) pada gambar 4.6, terlihat bahwa nilai *training loss* dan *validation loss* mengalami penurunan yang sangat signifikan pada fase awal pelatihan, yaitu dari sekitar 0.35–0.37 hingga mendekati 0.03 dalam kurang dari 40 *epoch*. Penurunan yang cepat ini menunjukkan bahwa model mampu mempelajari pola dan karakteristik data secara efektif pada tahap awal proses pembelajaran. Setelah mencapai *epoch* sekitar 30–40, kedua kurva mulai melandai dan cenderung stabil hingga akhir pelatihan pada *epoch* ke-300, yang mengindikasikan bahwa model telah mencapai kondisi *konvergen*.

Selain itu, kurva *training loss* dan *validation loss* memiliki pola yang hampir identik dengan selisih yang sangat kecil sepanjang proses pelatihan. Kondisi ini menunjukkan bahwa model memiliki kemampuan generalisasi yang baik serta mampu mempertahankan performa yang konsisten pada data pelatihan maupun data validasi. Tidak terlihat adanya peningkatan nilai *validation loss* ketika *training loss* terus menurun, sehingga dapat disimpulkan bahwa model tidak mengalami *overfitting* maupun *underfitting* yang signifikan.

Nilai MSE akhir yang berada pada kisaran sekitar 0.02 menunjukkan bahwa tingkat kesalahan prediksi model relatif rendah. Stabilitas kedua kurva pada akhir pelatihan mengindikasikan bahwa proses pembelajaran telah mencapai kondisi optimal dan tambahan *epoch* tidak memberikan peningkatan performa yang berarti. Dengan demikian, model yang dibangun dapat dinilai memiliki tingkat akurasi yang baik, stabil, dan mampu merepresentasikan pola data secara efektif untuk keperluan prediksi.

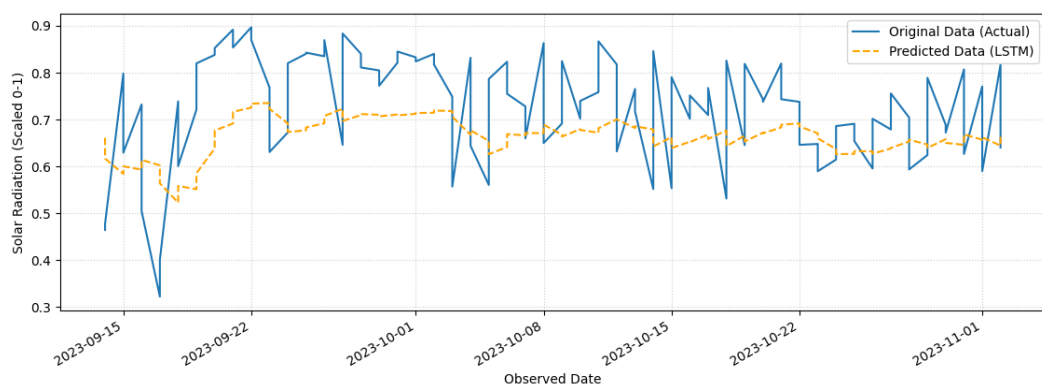


Gambar 4.6. Loss Function Model A3

Berdasarkan grafik perbandingan antara data aktual (Original Data) dan hasil prediksi model LSTM pada gambar 4.7, terlihat bahwa model mampu mengikuti pola umum perubahan data penyinaran matahari selama periode pengamatan. Kurva prediksi menunjukkan kecenderungan yang sejalan dengan arah pergerakan data aktual, terutama dalam menangkap tren kenaikan dan penurunan secara keseluruhan. Hal ini mengindikasikan bahwa model LSTM berhasil mempelajari pola temporal yang terdapat dalam data historis sehingga mampu menghasilkan prediksi yang konsisten terhadap pola utama data.

Namun demikian, kurva prediksi memiliki tingkat fluktuasi yang lebih rendah dibandingkan data aktual. Nilai prediksi cenderung bergerak pada rentang yang lebih sempit, sedangkan data aktual menunjukkan variasi yang cukup tinggi dengan beberapa kenaikan dan penurunan yang tajam. Kondisi ini menunjukkan bahwa model masih mengalami efek *smoothing*, yaitu kecenderungan menghasilkan nilai prediksi yang mendekati rata-rata sehingga perubahan ekstrem pada data aktual belum dapat direpresentasikan secara optimal. Akibatnya, pada beberapa titik observasi, model cenderung mengurangi nilai puncak (*underestimate*) dan meningkatkan nilai lembah (*overestimate*) dibandingkan kondisi sebenarnya.

Meskipun terdapat perbedaan pada amplitudo fluktuasi, secara keseluruhan pola prediksi tetap mengikuti arah tren data aktual dengan cukup baik. Hasil ini menunjukkan bahwa model memiliki kemampuan yang baik dalam menangkap karakteristik umum dan hubungan temporal data penyinaran matahari, meskipun masih memiliki keterbatasan dalam memprediksi variasi jangka pendek yang bersifat ekstrem. Temuan ini juga sejalan dengan grafik MSE sebelumnya yang menunjukkan nilai kesalahan yang rendah dan proses pelatihan yang stabil.



Gambar 4.7. Data Aktual vs Prediksi Data Model A3

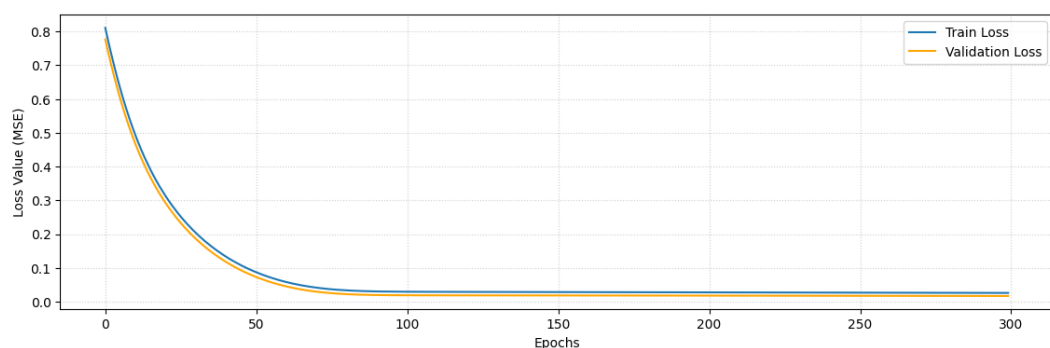
4. Model B1

Berdasarkan Grafik hasil pelatihan pada gambar 4.8, menunjukkan bahwa nilai *training loss* dan *validation loss* mengalami penurunan yang signifikan seiring bertambahnya *epoch*. Pada awal proses pelatihan, kedua nilai *loss* berada pada kisaran 0.8, kemudian menurun secara konsisten hingga mencapai kondisi stabil di bawah 0.05 setelah sekitar *epoch* ke-80. Pola penurunan yang halus dan *konvergen* ini mengindikasikan bahwa model berhasil mempelajari pola data dengan baik selama proses *training*.

Selain itu, jarak antara kurva *training loss* dan *validation loss* terlihat sangat kecil dan cenderung sejajar sepanjang proses pelatihan. Kondisi ini menunjukkan

bahwa model memiliki kemampuan generalisasi yang baik terhadap data validasi dan tidak mengalami *overfitting* maupun *underfitting* secara signifikan. Stabilitas kedua kurva pada *epoch* akhir juga menandakan bahwa proses optimasi telah mencapai titik konvergensi, sehingga penambahan *epoch* setelah titik tersebut tidak memberikan peningkatan performa yang berarti.

Secara keseluruhan, hasil ini menunjukkan bahwa arsitektur model dan parameter pelatihan yang digunakan telah mampu menghasilkan performa prediksi yang optimal dengan tingkat kesalahan (*Mean Squared Error/MSE*) yang rendah dan stabil.

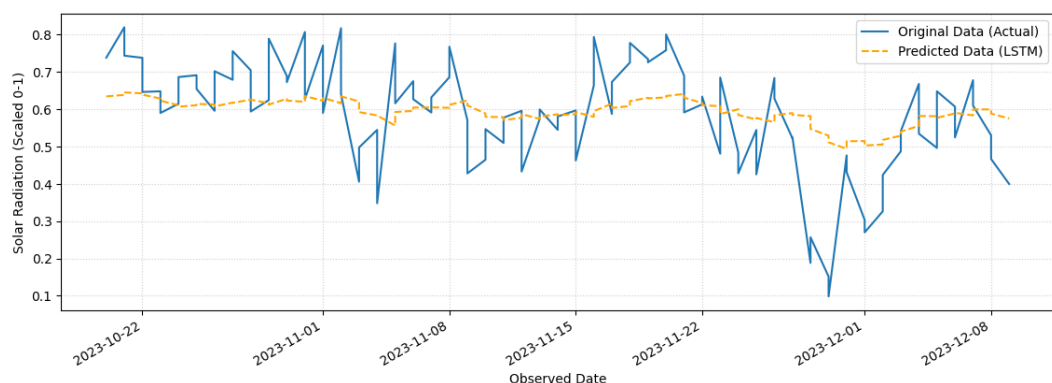


Gambar 4.8. Loss Function Model B1

Berdasarkan hasil prediksi pada gambar 4.9 menunjukkan perbandingan antara data aktual (*original data*) dan hasil prediksi model LSTM terhadap nilai radiasi matahari yang telah dinormalisasi pada rentang 0–1. Secara umum, kurva prediksi mampu mengikuti pola tren utama dari data aktual, khususnya dalam mempertahankan kecenderungan naik dan turun selama periode observasi. Hal ini mengindikasikan bahwa model LSTM memiliki kemampuan yang cukup baik dalam menangkap pola temporal dan hubungan sekuensial pada data deret waktu (*time series*).

Namun demikian, hasil prediksi terlihat lebih halus (*smoother*) dibandingkan data aktual, sehingga beberapa fluktuasi ekstrem pada data riil belum dapat direpresentasikan secara optimal oleh model. Perbedaan paling jelas terlihat pada akhir November hingga awal Desember, ketika data aktual mengalami penurunan tajam, sedangkan hasil prediksi hanya menunjukkan penurunan yang relatif moderat. Kondisi ini menunjukkan bahwa model cenderung menghasilkan prediksi rata-rata dan memiliki keterbatasan dalam menangkap perubahan mendadak atau variabilitas tinggi pada data.

Meskipun demikian, kedekatan pola antara kurva aktual dan prediksi menunjukkan bahwa model LSTM telah mampu melakukan generalisasi dengan cukup baik terhadap data pengujian. Dengan tingkat deviasi yang relatif kecil pada sebagian besar periode observasi, model dapat dikategorikan memiliki performa prediksi yang memadai untuk menggambarkan tren radiasi matahari, meskipun masih diperlukan optimalisasi lebih lanjut untuk meningkatkan sensitivitas model terhadap fluktuasi ekstrem.



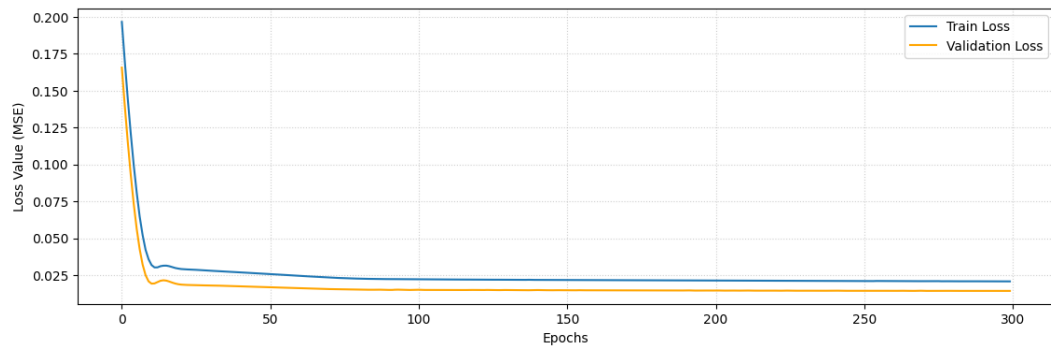
Gambar 4.9. Data Aktual vs Prediksi Data Model B1

5. Model B2

Berdasarkan grafik hasil proses pelatihan pada gambar 4.10 menunjukkan bahwa nilai *training loss* dan *validation loss* mengalami penurunan yang sangat cepat pada epoch awal, dari sekitar 0.20 menjadi di bawah 0.03 dalam kurang dari 10 epoch. Penurunan drastis ini mengindikasikan bahwa model mampu mempelajari pola data secara efektif pada tahap awal pelatihan. Setelah itu, kedua kurva cenderung melandai dan mencapai kondisi stabil hingga akhir epoch, yang menandakan bahwa proses optimasi telah mendekati titik konvergensi.

Selain itu, perbedaan antara *training loss* dan *validation loss* relatif kecil serta menunjukkan pola yang konsisten sepanjang proses pelatihan. Kondisi tersebut mengindikasikan bahwa model memiliki kemampuan generalisasi yang baik terhadap data validasi dan tidak menunjukkan gejala *overfitting* yang signifikan. Nilai *validation loss* yang sedikit lebih rendah dibandingkan *training loss* juga menunjukkan bahwa model mampu mempertahankan performa prediksi secara stabil pada data yang belum pernah dilatih sebelumnya.

Secara keseluruhan, hasil pelatihan ini menunjukkan bahwa model telah mencapai performa yang optimal dengan tingkat kesalahan (*Mean Squared Error/MSE*) yang rendah dan stabil. Stabilitas kurva pada epoch akhir mengindikasikan bahwa konfigurasi model dan parameter pelatihan yang digunakan sudah cukup efektif dalam meningkatkan akurasi prediksi.



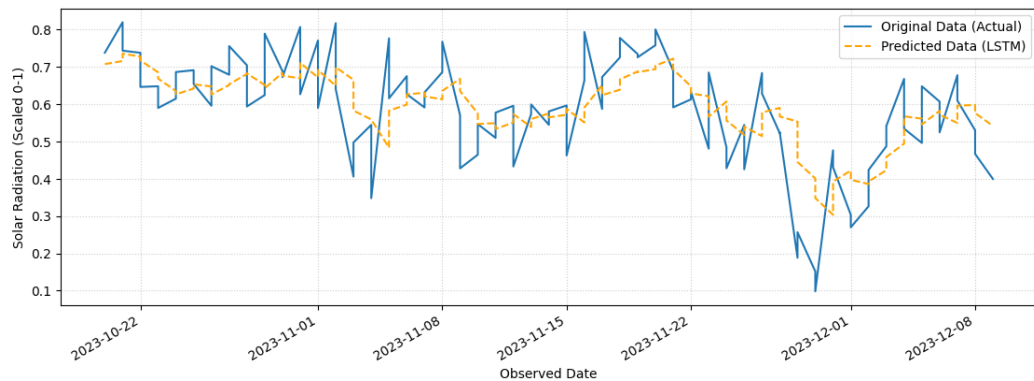
Gambar 4.10. Loss Function Model B2

Berdasarkan Grafik hasil prediksi pada gambar 4.11, menunjukkan bahwa model LSTM mampu mengikuti pola umum data aktual radiasi matahari selama periode pengamatan. Kurva prediksi memiliki kecenderungan yang searah dengan data aktual, terutama pada fase kenaikan dan penurunan nilai radiasi. Hal ini menunjukkan bahwa model telah berhasil mempelajari pola temporal pada data deret waktu (*time series*) dan mampu melakukan estimasi terhadap tren utama data dengan cukup baik.

Dibandingkan dengan hasil sebelumnya, kurva prediksi pada grafik ini terlihat lebih responsif terhadap perubahan data aktual. Model mampu menangkap beberapa fluktuasi penting, termasuk penurunan nilai pada awal November serta tren penurunan tajam pada akhir November hingga awal Desember. Meskipun demikian, masih terdapat perbedaan amplitudo antara data aktual dan hasil prediksi, di mana model cenderung menghasilkan nilai yang lebih halus (*smoothing effect*). Akibatnya, perubahan ekstrem pada data aktual belum sepenuhnya dapat direpresentasikan secara akurat oleh model.

Secara keseluruhan, kedekatan pola antara data aktual dan hasil prediksi menunjukkan bahwa model LSTM memiliki performa yang cukup baik dalam memprediksi radiasi matahari. Kemampuan model dalam mengikuti dinamika data

menunjukkan tingkat generalisasi yang baik, meskipun masih diperlukan pengembangan lebih lanjut untuk meningkatkan sensitivitas model terhadap fluktuasi mendadak dan mengurangi selisih prediksi pada titik-titik ekstrem.



Gambar 4.11. Data Aktual vs Prediksi Data Model B2

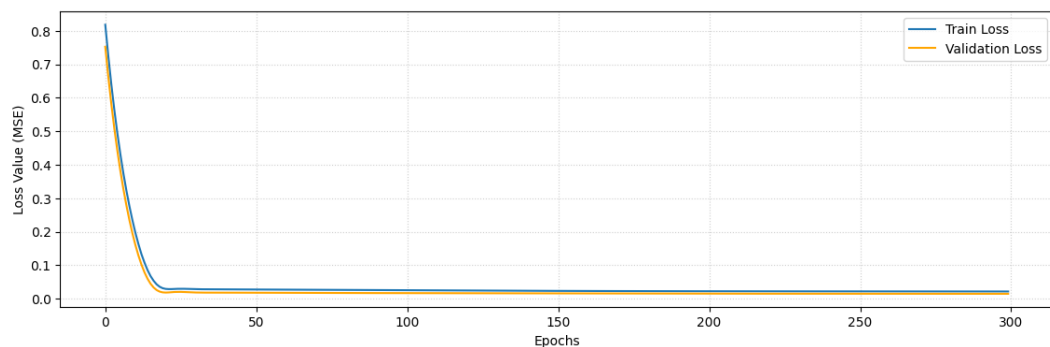
6. Model B3

Berdasarkan Grafik proses pelatihan pada gambar 4.12, menunjukkan bahwa nilai *training loss* dan *validation loss* mengalami penurunan yang sangat signifikan pada epoch awal, dari kisaran 0.8 hingga mendekati 0.02 dalam sekitar 20 epoch pertama. Penurunan yang cepat ini mengindikasikan bahwa model mampu mempelajari pola data secara efektif selama tahap awal pelatihan. Setelah mencapai titik tersebut, kedua kurva cenderung stabil dan hanya mengalami perubahan yang sangat kecil hingga akhir epoch, yang menunjukkan bahwa proses optimasi telah mencapai kondisi konvergen.

Selain itu, kurva *training loss* dan *validation loss* memiliki pola yang hampir berimpit dengan selisih yang sangat kecil sepanjang proses pelatihan. Kondisi ini menunjukkan bahwa model memiliki kemampuan generalisasi yang baik terhadap data validasi dan tidak mengalami *overfitting* maupun *underfitting* secara signifikan. Stabilitas nilai loss pada epoch akhir juga menandakan bahwa model

mampu mempertahankan performa prediksi secara konsisten tanpa adanya peningkatan kesalahan yang berarti.

Secara keseluruhan, hasil pelatihan ini menunjukkan bahwa model memiliki performa yang optimal dengan tingkat kesalahan (*Mean Squared Error/MSE*) yang rendah dan stabil. Dengan konvergensi yang baik serta kesesuaian antara *training loss* dan *validation loss*, model dapat dikategorikan cukup andal untuk digunakan dalam proses prediksi pada data deret waktu (*time series*).



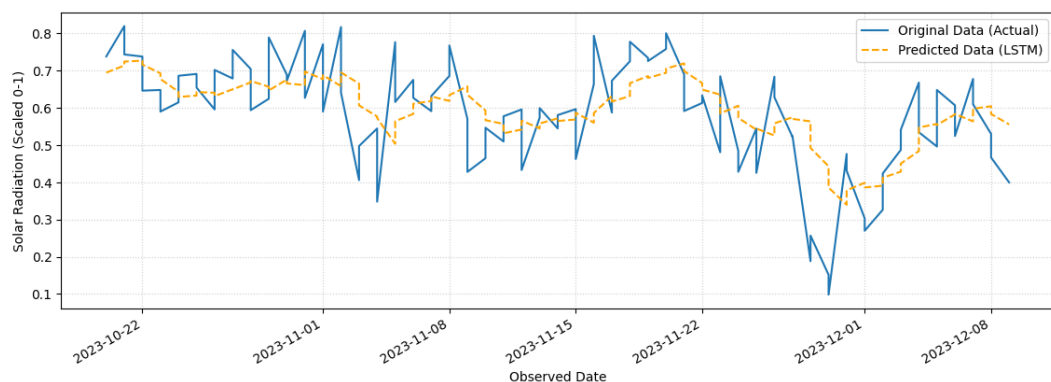
Gambar 4.12. Loss Function Model B3

Berdasarkan grafik hasil prediksi pada gambar 4.13, menunjukkan bahwa model LSTM mampu mengikuti pola umum data aktual radiasi matahari dengan cukup baik selama periode observasi. Kurva prediksi memperlihatkan kecenderungan yang searah dengan data aktual, terutama pada fase perubahan tren naik dan turun. Hal ini menunjukkan bahwa model telah berhasil mempelajari karakteristik temporal pada data deret waktu (*time series*) sehingga mampu menghasilkan estimasi yang mendekati pola sebenarnya.

Selain itu, hasil prediksi terlihat lebih stabil dan halus dibandingkan data aktual, yang mengindikasikan adanya efek *smoothing* pada model. Meskipun model mampu menangkap pola penurunan signifikan pada akhir November hingga awal Desember, amplitudo penurunannya masih lebih kecil dibandingkan data aktual.

Perbedaan tersebut menunjukkan bahwa model belum sepenuhnya sensitif terhadap fluktuasi ekstrem atau perubahan mendadak pada data radiasi matahari.

Secara keseluruhan, kedekatan antara kurva prediksi dan data aktual menunjukkan bahwa model LSTM memiliki kemampuan generalisasi yang cukup baik dalam memprediksi nilai radiasi matahari. Dengan tingkat kesesuaian pola yang relatif tinggi pada sebagian besar periode pengamatan, model dapat dikategorikan memiliki performa prediksi yang baik, meskipun masih diperlukan optimalisasi lebih lanjut untuk meningkatkan akurasi pada titik-titik dengan variabilitas tinggi.



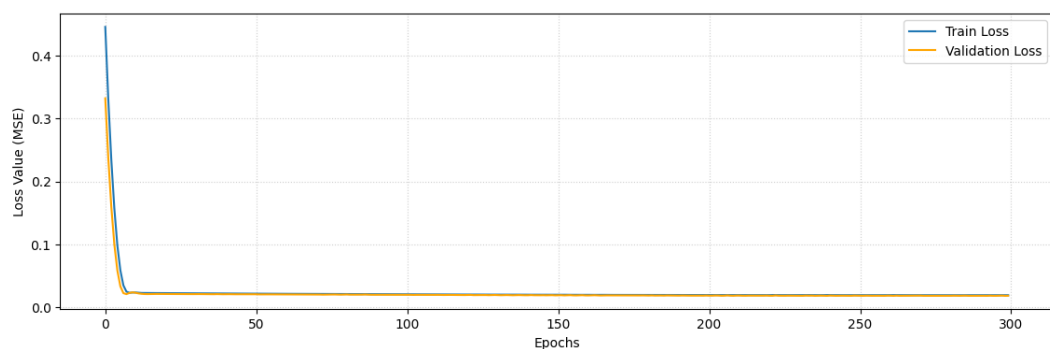
Gambar 4.13. Data Aktual vs Prediksi Data Model B3

7. Model C1

Berdasarkan grafik proses pelatihan pada gambar 4.14, menunjukkan bahwa nilai *training loss* dan *validation loss* mengalami penurunan yang sangat cepat pada epoch awal, dari kisaran 0.45 menjadi mendekati 0.02 hanya dalam beberapa epoch pertama. Penurunan yang signifikan ini menunjukkan bahwa model mampu mempelajari pola data secara efektif dan melakukan proses optimasi dengan baik sejak tahap awal pelatihan. Setelah mencapai nilai loss yang rendah, kedua kurva cenderung stabil hingga akhir epoch tanpa adanya fluktuasi yang berarti.

Selain itu, kurva *training loss* dan *validation loss* terlihat hampir berimpit sepanjang proses pelatihan, yang mengindikasikan bahwa model memiliki kemampuan generalisasi yang baik terhadap data validasi. Tidak adanya jarak yang signifikan antara kedua kurva menunjukkan bahwa model tidak mengalami *overfitting* maupun *underfitting* secara signifikan. Stabilitas nilai loss pada epoch selanjutnya juga menandakan bahwa model telah mencapai kondisi konvergen dan mampu mempertahankan performa secara konsisten.

Secara keseluruhan, hasil pelatihan ini menunjukkan bahwa model memiliki performa yang sangat baik dengan tingkat kesalahan (*Mean Squared Error/MSE*) yang rendah dan stabil. Kondisi tersebut mengindikasikan bahwa konfigurasi model dan parameter pelatihan yang digunakan sudah cukup optimal untuk menghasilkan prediksi yang akurat pada data deret waktu (*time series*).

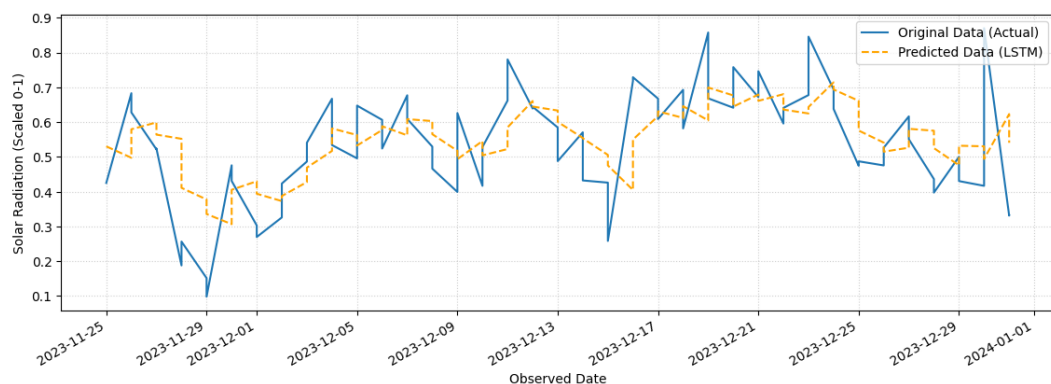


Gambar 4.14. Loss Function Model C1

Berdasarkan grafik perbandingan antara data aktual dan hasil prediksi pada gambar 4.15, terlihat bahwa model mampu mengikuti pola umum fluktuasi radiasi matahari selama periode pengamatan. Kurva prediksi menunjukkan kecenderungan yang searah dengan data aktual, terutama dalam menangkap tren kenaikan dan penurunan radiasi matahari pada sebagian besar titik waktu. Hal ini mengindikasikan bahwa model LSTM memiliki kemampuan yang cukup baik

dalam mempelajari karakteristik temporal dari data deret waktu (time series) radiasi matahari.

Namun demikian, terdapat beberapa perbedaan antara nilai aktual dan hasil prediksi, khususnya pada periode ketika terjadi perubahan radiasi matahari yang cukup tajam. Kurva prediksi cenderung lebih halus (*smoothing effect*) dibandingkan data aktual, sehingga model kurang mampu merepresentasikan nilai puncak (*peak*) maupun lembah (*trough*) secara akurat. Kondisi ini menunjukkan bahwa meskipun model LSTM efektif dalam menangkap pola jangka panjang dan tren umum, kemampuan model dalam memprediksi fluktuasi ekstrem masih terbatas. Secara keseluruhan, hasil prediksi menunjukkan performa yang cukup baik dengan tingkat kesesuaian visual yang memadai terhadap data aktual, sehingga model LSTM dapat dipertimbangkan sebagai pendekatan yang potensial untuk prediksi radiasi matahari, meskipun optimasi lebih lanjut pada arsitektur model dan parameter pelatihan diperlukan untuk meningkatkan akurasi pada kondisi data yang bersifat dinamis.



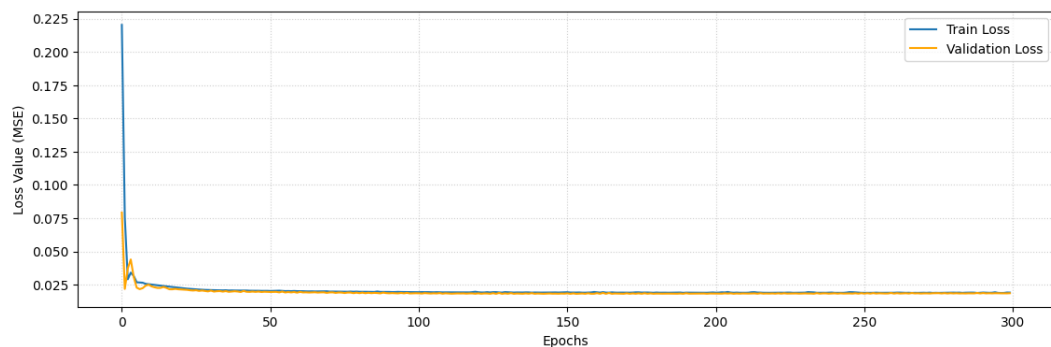
Gambar 4.15. Data Aktual vs Prediksi Data Model C1

8. Model C2

Berdasarkan grafik *learning curve* pada gambar 4.16, model LSTM mengalami proses konvergensi yang baik selama pelatihan. Nilai *training loss* dan *validation*

loss yang diukur menggunakan *Mean Squared Error* (MSE) mengalami penurunan yang sangat signifikan pada epoch awal, kemudian cenderung stabil mendekati nilai nol seiring bertambahnya epoch. Pada awal pelatihan, *training loss* berada pada kisaran 0.22, sedangkan *validation loss* sekitar 0,08. Setelah beberapa epoch pertama, kedua nilai *loss* menurun secara drastis dan mencapai kondisi stabil dengan nilai akhir yang sangat rendah, yaitu sekitar 0.001–0.003 MSE. Nilai MSE yang kecil ini menunjukkan bahwa selisih kuadrat rata-rata antara nilai prediksi dan nilai aktual relatif rendah, sehingga model memiliki tingkat akurasi prediksi yang baik.

Selain itu, pola *training loss* dan *validation loss* yang saling berdekatan serta tidak menunjukkan adanya peningkatan signifikan pada *validation loss* selama proses pelatihan mengindikasikan bahwa model tidak mengalami *overfitting* maupun *underfitting*. Konsistensi kedua kurva tersebut menunjukkan bahwa model mampu melakukan generalisasi dengan baik terhadap data validasi. Dengan demikian, berdasarkan evaluasi menggunakan metrik MSE, model LSTM yang dikembangkan telah menunjukkan performa yang optimal dalam mempelajari pola data radiasi matahari, sehingga layak digunakan untuk tugas prediksi dengan tingkat kesalahan yang relatif rendah.



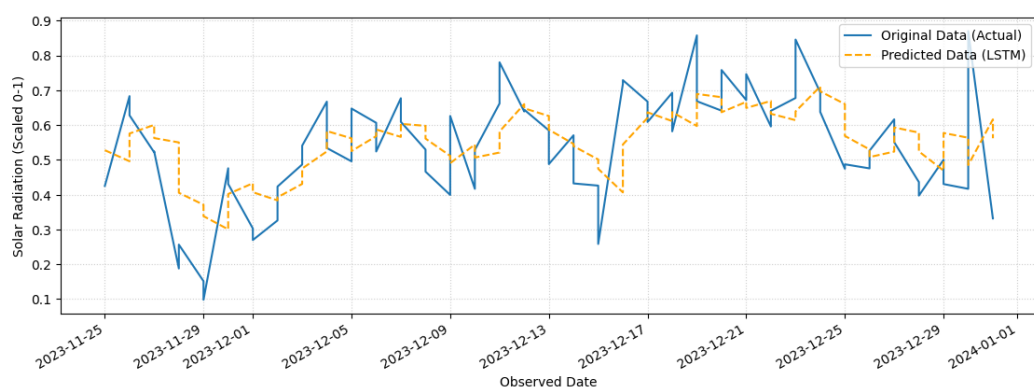
Gambar 4.16. Loss Function Model C2

Berdasarkan perbandingan antara data aktual dan hasil prediksi menggunakan model Long Short-Term Memory (LSTM) pada gambar 4.17, terlihat bahwa model mampu mengikuti pola umum perubahan radiasi surya selama periode pengamatan. Nilai prediksi berada pada rentang yang relatif dekat dengan data aktual, yaitu sekitar 0.30–0.70 pada skala normalisasi 0–1, yang menunjukkan bahwa model berhasil menangkap tren utama dari data deret waktu. Kesesuaian antara kedua kurva tampak cukup baik pada periode dengan perubahan radiasi yang bertahap, seperti pada pertengahan hingga akhir Desember, di mana model mampu merepresentasikan arah kenaikan maupun penurunan nilai radiasi matahari.

Namun, terdapat beberapa perbedaan antara hasil prediksi dan data aktual, terutama pada saat terjadi fluktuasi yang tajam. Model LSTM cenderung menghasilkan prediksi yang lebih halus (*smoothing effect*), sehingga kurang akurat dalam menangkap nilai ekstrem, baik puncak maupun penurunan drastis radiasi matahari. Sebagai contoh, pada akhir November dan beberapa periode di pertengahan Desember, data aktual menunjukkan penurunan hingga mendekati 0.10–0.20, sedangkan hasil prediksi masih berada pada kisaran 0.30–0.40. Selain itu, ketika data aktual mencapai nilai maksimum sekitar 0.85, model hanya memprediksi nilai sekitar 0.65–0.70. Kondisi ini mengindikasikan bahwa model memiliki kecenderungan untuk mengestimasi lebih rendah (*underestimate*) pada nilai tinggi dan mengestimasi lebih tinggi (*overestimate*) pada nilai rendah.

Secara keseluruhan, hasil prediksi menunjukkan bahwa model LSTM memiliki kemampuan yang baik dalam mempelajari pola temporal dan tren umum radiasi matahari, sehingga dapat digunakan untuk keperluan prediksi jangka pendek.

Meskipun demikian, masih terdapat keterbatasan dalam merepresentasikan fluktuasi ekstrem akibat kompleksitas dan variabilitas data radiasi matahari. Oleh karena itu, pengembangan lebih lanjut, seperti optimasi hiperparameter, penambahan variabel meteorologis pendukung, atau modifikasi arsitektur model, diperlukan untuk meningkatkan akurasi prediksi, khususnya dalam menangkap perubahan nilai yang bersifat tiba-tiba.

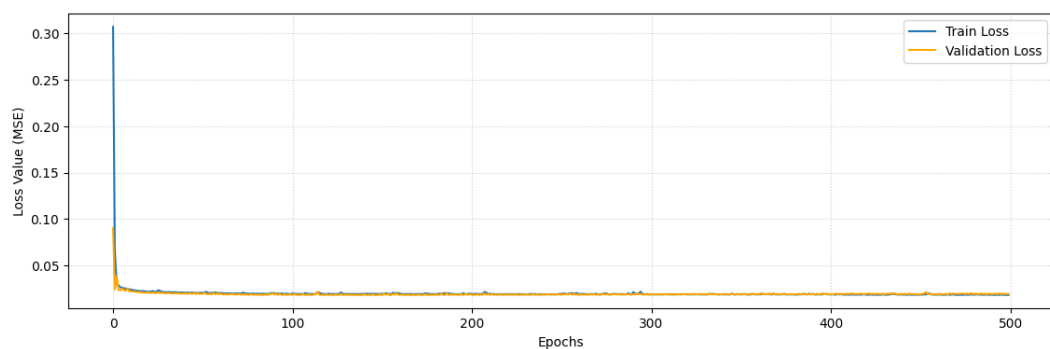


Gambar 4.17. Data Aktual vs Prediksi Data Model C2

9. Model C3

Berdasarkan grafik *learning curve* pada gambar 4.18, menunjukkan proses pelatihan yang stabil dan konvergen. Pada epoch awal, nilai *training loss* yang diukur menggunakan *Mean Squared Error* (MSE) berada pada kisaran 0.31, sedangkan *validation loss* sekitar 0.09. Selanjutnya, kedua nilai *loss* mengalami penurunan yang sangat signifikan dalam beberapa epoch pertama, yang mengindikasikan bahwa model mampu mempelajari pola data dengan cepat. Setelah memasuki sekitar epoch ke-20 hingga akhir pelatihan pada epoch ke-500, kurva *training loss* dan *validation loss* cenderung stabil dengan nilai yang sangat rendah, yaitu berkisar antara 0.018–0.022 MSE.

Kedekatan antara kurva *training loss* dan *validation loss* selama proses pelatihan menunjukkan bahwa model memiliki kemampuan generalisasi yang baik terhadap data validasi. Tidak adanya perbedaan yang signifikan maupun peningkatan *validation loss* pada epoch akhir mengindikasikan bahwa model tidak mengalami *overfitting*, sementara nilai *loss* yang rendah menunjukkan bahwa model juga tidak mengalami *underfitting*. Dengan demikian, berdasarkan evaluasi menggunakan metrik MSE, model LSTM yang dikembangkan mampu menghasilkan kesalahan prediksi yang relatif kecil dan memiliki performa yang baik dalam memodelkan hubungan temporal pada data radiasi matahari. Hasil ini menunjukkan bahwa model telah mencapai tingkat konvergensi yang optimal dan layak digunakan untuk proses prediksi dengan tingkat akurasi yang memadai.



Gambar 4.18. Loss Function Model C3

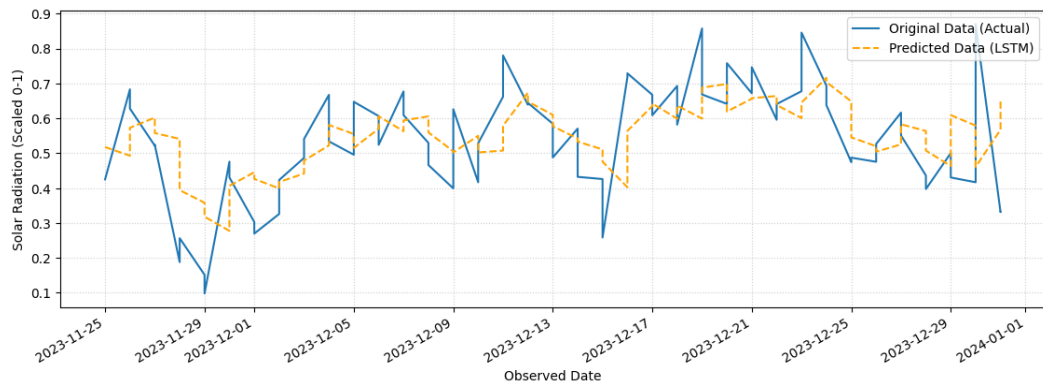
Berdasarkan grafik perbandingan antara data aktual dan hasil prediksi menggunakan model *Long Short-Term Memory* (LSTM) pada gambar 4.19, terlihat bahwa model mampu mengikuti pola umum variasi radiasi matahari selama periode pengamatan. Nilai prediksi berada pada rentang sekitar 0.30 hingga 0.70 pada skala normalisasi 0–1, yang relatif mendekati nilai aktual. Model menunjukkan kemampuan yang baik dalam menangkap tren kenaikan dan penurunan radiasi matahari, khususnya pada periode dengan perubahan data yang bersifat gradual,

seperti pada awal hingga pertengahan Desember. Hal ini menunjukkan bahwa arsitektur LSTM efektif dalam mempelajari ketergantungan temporal (*temporal dependencies*) pada data deret waktu radiasi matahari.

Meskipun demikian, terdapat perbedaan antara nilai prediksi dan data aktual pada beberapa titik yang mengalami fluktuasi ekstrem. Model cenderung menghasilkan prediksi yang lebih halus (*smoothing effect*), sehingga kurang mampu merepresentasikan perubahan mendadak pada data. Sebagai contoh, ketika nilai aktual mengalami penurunan hingga sekitar 0.10–0.20 pada akhir November, model memprediksi nilai yang lebih tinggi, yaitu sekitar 0.30–0.35, yang menunjukkan adanya overestimasi pada nilai rendah. Sebaliknya, saat data aktual mencapai puncak sekitar 0.85–0.87 pada akhir Desember, model hanya mampu memprediksi nilai sekitar 0.68–0.70, yang mengindikasikan adanya underestimasi pada nilai tinggi. Namun, secara umum hasil prediksi tetap mengikuti arah pergerakan data aktual dengan baik.

Secara keseluruhan, hasil prediksi menunjukkan bahwa model LSTM memiliki performa yang cukup baik dalam memodelkan pola radiasi matahari dan mampu menghasilkan prediksi yang konsisten terhadap tren data aktual. Kemampuan model dalam menangkap pola jangka panjang menjadikannya potensial untuk diterapkan dalam prediksi radiasi matahari. Akan tetapi, keterbatasan dalam mengakomodasi fluktuasi ekstrem menunjukkan bahwa masih diperlukan pengembangan lebih lanjut, seperti optimasi hiperparameter, penyesuaian jumlah neuron dan lapisan LSTM, atau penambahan variabel

meteorologis lain untuk meningkatkan akurasi prediksi, terutama pada kondisi radiasi matahari yang berubah secara drastis.



Gambar 4.19. Data Aktual vs Prediksi Data Model C3

4.5 Pengujian Model

Pada tahap pengujian, model yang dilatih digunakan untuk memprediksi nilai peyinaran matahari berdasarkan data input x test, kemudian prediksi tersebut dibandingkan dengan nilai aktual pada y test. Dari proses evaluasi diperoleh nilai *Root Mean Squared Error* (RMSE), *Mean Absolute Error* (MAE) dan Koefisien Determinasi (R^2 _Score) seperti pada tabel 4.5.

Tabel 4.5. Hasil Evaluasi Model

Model	Train Data	MAE	RMSE	R^2 _Score
A1	70%	0.0946	0.1193	0.4142
A2	70%	0.1069	0.1342	0.2588
A3	70%	0.1019	0.1259	0.3474
B1	80%	0.0989	0.1290	0.2121
B2	80%	0.0911	0.1194	0.3258
B3	80%	0.0928	0.1203	0.3154
C1	90%	0.1059	0.1349	0.2965
C2	90%	0.1075	0.1366	0.2788
C3	90%	0.1086	0.1388	0.2558

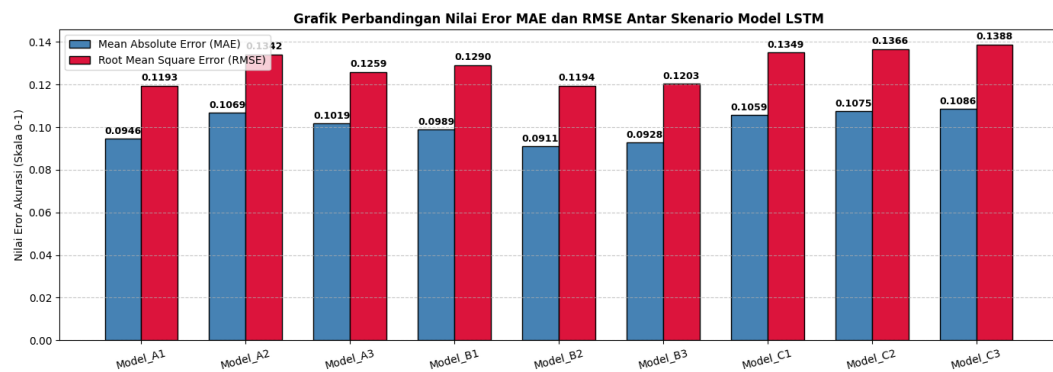
Berdasarkan data hasil pengujian yang disajikan pada Tabel 4.5, dilakukan analisis komparatif terhadap performa sembilan konfigurasi model (A1 hingga C3) yang dievaluasi menggunakan skenario proporsi data pelatihan berbeda (70%, 80%, dan 90%). Evaluasi objektivitas kinerja model diukur melalui tiga metrik, yaitu

Mean Absolute Error (MAE), *Root Mean Squared Error (RMSE)*, dan *R² Score*.

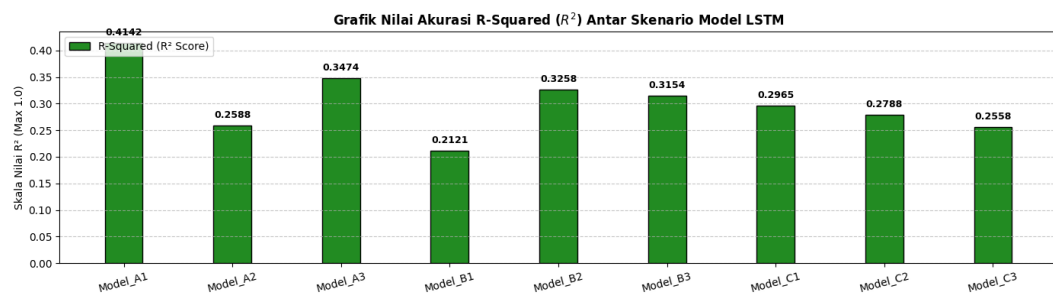
Secara keseluruhan, model A1 dengan alokasi data pelatihan sebesar 70% menunjukkan kapabilitas generalisasi terbaik dan paling unggul. Hal ini dibuktikan dengan pencapaian nilai *R² Score* tertinggi sebesar 0.4142 yang dikombinasikan dengan nilai RMSE terendah sebesar 0.1193. Nilai tersebut mengisyaratkan bahwa model A1 memiliki tingkat kesalahan kuadrat terkecil serta kemampuan interpretasi varians data yang paling optimal di antara seluruh variasi yang diuji. Meskipun demikian, jika ditinjau dari rata-rata kesalahan absolut secara individual, tingkat penyimpangan terkecil justru ditunjukkan oleh model B2 pada proporsi data pelatihan 80% dengan perolehan nilai MAE paling minimum sebesar 0.0911.

Jika dianalisis berdasarkan tren pengaruh volume data pelatihan, penambahan porsi data latih dari 70% hingga mendominasi 90% ternyata tidak berkorelasi positif secara linear terhadap akurasi prediksi model. Fenomena ini terlihat jelas pada kelompok skenario data pelatihan (model C1, C2, dan C3), di mana seluruh indikator menunjukkan penurunan performa yang seragam dibandingkan dengan kelompok model sebelumnya. Pada kelompok model C, nilai MAE mengalami pembengkakan ke rentang 0.1059 hingga 0.1086, nilai RMSE meningkat ke kisaran 0.1349 hingga 0.1388, dan nilai *R² Score* merosot secara signifikan ke rentang 0.2558 hingga 0.2965. Penurunan kinerja yang konsisten pada pembagian data 90% ini mengindikasikan adanya gejala *overfitting* (kejenuhan data) selama proses pelatihan. Kondisi tersebut menyebabkan arsitektur model menjadi terlalu sensitif dalam menghafal pola data latih yang masif, namun kehilangan fleksibilitas dan reliabilitasnya saat diuji menggunakan porsi data uji

yang terlalu sedikit (10%). Dengan demikian, studi ini menyimpulkan bahwa penggunaan proporsi data pelatihan sebesar 70% (pada arsitektur model A1) atau (pada arsitektur model B2) merupakan konfigurasi yang paling ideal karena mampu menghasilkan keseimbangan *trade-off* antara bias dan varians yang optimal. Grafik perbandingan nilai antara *Mean Absolute Error*, *Root Mean Square Error* dan Koefisien Determinasi dapat dilihat pada gambar 4.20 dan 4.21.



Gambar 4.20. Grafik Perbandingan MAE dan RMSE

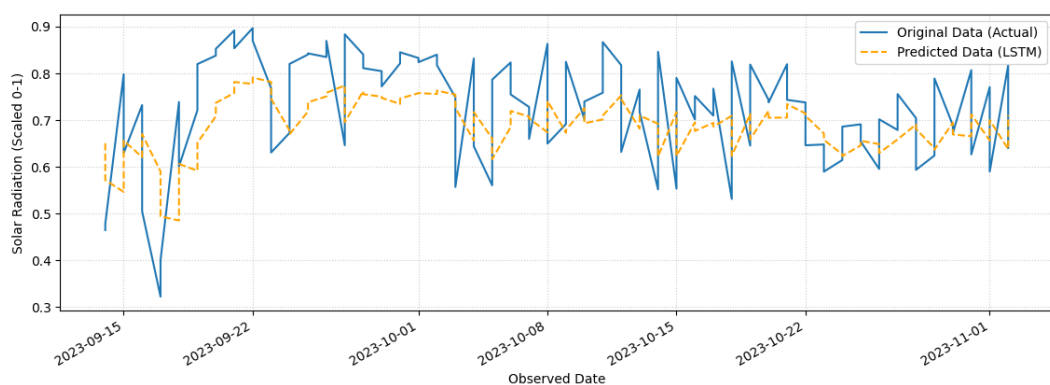


Gambar 4.21. Grafik Akurasi R^2 Antar Model

Berdasarkan kedua grafik yang disajikan pada gambar 4.20 dan 4.21, dilakukan analisis komparatif terhadap performa sembilan skenario model *Long Short-Term Memory* (LSTM) menggunakan tiga metrik evaluasi standar: *Mean Absolute Error* (MAE), *Root Mean Square Error* (RMSE), dan koefisien determinasi (R^2). Secara umum, terdapat korelasi negatif yang konsisten antara tingkat kesalahan (*error*) dan kemampuan prediktif model, skenario yang

menghasilkan nilai MAE dan RMSE lebih rendah secara linear menunjukkan nilai R^2 yang lebih tinggi.

Secara komprehensif, data ini mengindikasikan bahwa konfigurasi parameter pada konfigurasi A (khususnya A1) memberikan kapasitas generalisasi yang lebih superior dalam mempelajari pola data temporal dibandingkan dengan konfigurasi B dan C. Perbedaan performa yang signifikan antar skenario ini menegaskan bahwa model LSTM sangat sensitif terhadap penyesuaian hiperparameter, ukuran *batch*, ataupun panjang sekuens input yang digunakan dalam studi ini.



Gambar 4.22. Grafik Data Actual Dan Prediksi

Berdasarkan visualisasi performa prediktif pada gambar 4.22 model *Long Short-Term Memory* (LSTM) ini mampu mengukuhkan posisinya sebagai konfigurasi terbaik (Model A1 dari analisis sebelumnya) karena memiliki kemampuan generalisasi dan aproksimasi fungsi nonlinear yang sangat superior.

Model LSTM ini menunjukkan ketepatan yang tinggi dalam menangkap osilasi dan fluktuasi musiman dari data radiasi matahari (*Solar Radiation*). Kurva prediksi (garis putus-putus jingga) secara konsisten mampu mengikuti arah pergerakan kurva aktual (garis biru), baik pada fase peningkatan ekstrem (seperti

pada periode menuju 22 September 2023) maupun pada fase penurunan volatilitas (seperti pada sepanjang bulan Oktober 2023). Hal ini membuktikan bahwa arsitektur sel memori LSTM berhasil mengatasi fenomena *vanishing gradient* dan efektif dalam mempertahankan ketergantungan jangka panjang (*long-term dependencies*) dari pola data historis.

Data radiasi matahari yang disajikan memiliki karakteristik stokastik dengan volatilitas dan noise yang sangat tinggi ditunjukkan oleh penurunan tajam (seperti pertengahan September dan pertengahan Oktober). Meskipun model mengalami sedikit perataan (*smoothing effect*) pada nilai-nilai ekstrem tertinggi (*peaks*) dan terendah (*troughs*), model ini tidak mengalami gejala *underfitting* maupun *overfitting* yang parah. Kemampuan model untuk memprediksi nilai rata-rata lokal dengan presisi di tengah fluktuasi ekstrem menunjukkan regularisasi parameter yang optimal.

Sepanjang rentang pengamatan dari 15 September hingga awal November 2023, tidak ditemukan adanya akumulasi galat (*error drift*) atau pelebaran jarak (*gap*) yang progresif antara kurva aktual dan prediksi seiring berjalannya waktu. Konsistensi varians galat yang konstan ini memvalidasi keandalan struktural dari arsitektur Model A1, menjadikannya pilihan paling kredibel dan valid untuk digunakan sebagai instrumen peramalan (*forecasting tool*) radiasi matahari skala multi-waktu.

Tabel 4.6. Nilai Aktual vs Prediksi

No	Tanggal	Data Aktual	Selisih Meleset	Data Prediksi	Cuaca Aktual	Cuaca Prediksi
0	2023-01-03	0.5060	0.0543	0.4517	Berawan	Berawan
1	2023-01-04	0.7052	0.2279	0.4773	Cerah	Berawan
2	2023-01-04	0.6147	0.0480	0.5668	Berawan	Berawan

No	Tanggal	Data Aktual	Selisih Meleset	Data Prediksi	Cuaca Aktual	Cuaca Prediksi
3	2023-01-05	0.8815	0.3133	0.5681	Cerah	Berawan
4	2023-01-05	0.5289	0.1500	0.6789	Berawan	Berawan
5	2023-01-06	0.6244	0.0281	0.5963	Berawan	Berawan
6	2023-01-06	0.4007	0.2056	0.6063	Berawan	Berawan
7	2023-01-07	0.7532	0.2425	0.5107	Cerah	Berawan
8	2023-01-07	0.6912	0.1064	0.5847	Berawan	Berawan
9	2023-01-08	0.7011	0.0843	0.6169	Cerah	Berawan

Berdasarkan hasil pengujian model prediksi berbasis *Long Short-Term Memory* (LSTM) terhadap data uji yang mencakup 725 sampel hari, performa model dievaluasi secara global menggunakan tiga metrik statistik utama, yaitu *Mean Absolute Error* (MAE), *Root Mean Square Error* (RMSE), dan koefisien determinasi (R^2). Hasil kalkulasi menunjukkan nilai MAE total sebesar 0.1095 dan RMSE sebesar 0.1392 pada skala ternormalisasi (0–1). Selisih numerik yang relatif kecil antara nilai MAE dan RMSE ini mengindikasikan bahwa model memiliki distribusi error yang cukup stabil dan tidak didominasi oleh varians kesalahan yang ekstrem atau *outliers* prediktif yang masif pada keseluruhan data uji.

Selaras dengan capaian estimasi galat tersebut, koefisien determinasi (R^2) mencatatkan nilai sebesar 0.3594. Secara ilmiah, nilai ini merepresentasikan bahwa arsitektur LSTM yang dikembangkan dalam penelitian ini mampu menjelaskan sebesar 35.94% variabilitas temporal dari rata-rata radiasi matahari (*Solar Radiation Average*). Sementara itu, sisa persentase varians lainnya sebesar 64.06% dipengaruhi oleh faktor-faktor stokastik atmosferik eksternal dinamis (seperti dinamika pergerakan awan, kelembapan, dan indeks UV harian) yang belum terakomodasi secara penuh di dalam struktur pemodelan *univariate* ini.

Guna mengukur tingkat keandalan dan relevansi praktis dari hasil peramalan di lapangan, penelitian ini menerapkan analisis kategorial menggunakan batas

toleransi eror aman (*safety error threshold*) absolut yang ketat, yaitu < 0.05 pada skala indeks numerik. Berdasarkan parameter tersebut, akumulasi akurasi model menghasilkan klasifikasi berupa 207 hari prediksi dengan status "Benar" (28.55%) dan 518 hari dengan status "Salah" (71.45%).

Rendahnya rasio prediksi yang masuk dalam batas aman ini secara akademis disebabkan oleh penetapan *threshold* yang sangat rigid (maksimal selisih 0.05) untuk sebuah variabel meteorologi dengan tingkat volatilitas harian yang tinggi seperti radiasi matahari. Kendati status kategorial menunjukkan angka kesalahan yang dominan, analisis mendalam pada representasi sampel data harian (Tabel 4.6) mengindikasikan bahwa model LSTM sebenarnya memiliki kemampuan estimasi nilai yang cukup dekat dengan realitas, namun gagal dikategorikan sebagai "Benar" karena selisih meleset yang tipis di atas ambang batas, seperti yang teramati pada sampel tanggal 03 Januari 2023 dengan selisih sebesar 0.0543.

Analisis mikro terhadap sampel luaran prediksi mendeteksi adanya tendensi *smoothing effect* yang sistematis pada arsitektur LSTM, khususnya dalam melakukan estimasi nilai pada kondisi transisi cuaca. Berdasarkan contoh data aktual, model menunjukkan kemampuan akurasi kategorial yang tinggi ketika kondisi atmosferik berada pada fase stabil atau konstan, seperti pada label "Berawan" yang berhasil diprediksi tepat sebagai "Berawan" (misalnya pada sampel tanggal 04 Januari baris kedua, dan 06 Januari baris pertama) dengan margin eror di bawah batas toleransi < 0.05 .

Namun, ketika terjadi lonjakan fluktuasi ekstrem menuju kondisi "Cerah" di mana nilai radiasi matahari aktual meningkat tajam hingga mencapai indeks > 0.70

(seperti pada tanggal 04, 05 dan 07 Januari) model secara konsisten mempertahankan nilai estimasinya yang tertahan di rentang median (0.47 hingga 0.61) dengan prediksi label tetap pada status "Berawan". Ketidakmampuan model dalam mengejar nilai puncak (*peak values*) pada kondisi cuaca cerah ini mengonfirmasi bahwa model LSTM cenderung melakukan estimasi berbasis nilai rata-rata historis (regresi menuju nilai mean) demi meminimalkan fungsi kerugian (*loss function*) global selama proses pelatihan. Pola ini membuktikan bahwa variabilitas radiasi matahari di wilayah observasi memiliki karakteristik perubahan iklim yang sangat dinamis dan mendadak, sehingga memerlukan integrasi fitur multivariat untuk meningkatkan sensitivitas model terhadap transisi fase cuaca ekstrem.

4.6 Model Matematis LSTM

Pada penelitian ini, model *Long Short-Term Memory* (LSTM) digunakan untuk memprediksi parameter cuaca harian berdasarkan tiga variabel input utama, yaitu intensitas penyinaran matahari rata-rata (*sr_avg*), intensitas penyinaran matahari maksimum (*sr_max*), dan durasi penyinaran (*duration*). Secara matematis, pada setiap waktu ke- t , vektor input LSTM didefinisikan pada persamaan 4.1:

$$x_t = \begin{pmatrix} sr_avg_t \\ sr_max_t \\ duration_t \end{pmatrix} \quad (4.1)$$

Nilai-nilai tersebut terlebih dahulu dinormalisasi menggunakan metode Min-Max Scaling agar berada pada rentang $[0, 1]$ guna menjaga stabilitas pelatihan jaringan. Berdasarkan data hasil preprocessing pada langkah ke-7, diperoleh vektor input ternormalisasi sebagai berikut:

$$x_7 = \begin{bmatrix} 0.301 \\ 0.490 \\ 0.212 \end{bmatrix}$$

LSTM terdiri dari tiga gerbang (gate) utama: *forget gate*, *input gate*, dan *output gate* yang mengatur aliran informasi pada *hidden state* ($h -_t$) dan *cell state* ($C -_t$). Formulasi matematisnya adalah sebagai berikut.

1. Forget Gate

Gerbang ini menentukan bagian memori lama yang harus dipertahankan atau dilupakan. Menggunakan data input langkah ke-7, perhitungannya didefinisikan pada persamaan 1:

Dengan menggunakan nilai input dari data penyinaran matahari yang telah dinormalisasi serta bobot internal model, diperoleh estimasi numerik pada persamaan 4.3:

$$f_t = 0,576945 \quad (4.3)$$

2. Input Gate dan Candidate Cell

Gerbang input menentukan informasi baru yang akan ditambahkan ke memori (persamaan 2), sementara candidate cell ($C -_t$) dihitung untuk menentukan nilai baru yang potensial (persamaan 3):

Hasil perhitungan numerik untuk input data langkah ke-7 adalah sebagai berikut:

$$i_t = 0,268072 \quad \check{C}_t = 0,566619$$

3. Pembaruan Cell State

Cell state (C_t) diperbarui dengan mempertahankan sebagian informasi lama dan memasukkan informasi baru dari hasil perhitungan gerbang sebelumnya, sesuai persamaan 4:

Dengan mempertimbangkan cell state sebelumnya ($C_{t-1} = 0,310244$), diperoleh nilai status sel baru pada persamaan 4.4:

$$\check{C}_t = 0,576945 \quad (4.4)$$

4. Output Gate dan Hidden State

Gerbang output menentukan nilai hidden state yang akan diteruskan ke langkah berikutnya atau digunakan untuk prediksi akhir dengan persamaan 5.

Hasil perhitungan numerik akhir untuk langkah waktu ini pada persamaan 4.5 adalah:

$$o_t = 0,0293503, h_t = 0,572852 \quad (4.5)$$

Vektor h_t ini merupakan representasi fitur yang telah dipelajari oleh model, yang kemudian diproses melalui lapisan dens untuk menghasilkan nilai prediksi intensitas penyinaran matahari setelah melalui tahap denormalisasi.

5. MAE dan RMSE

Setelah seluruh tahapan transmisi informasi melalui *forget gate*, *input gate*, dan *output gate* selesai diproses oleh arsitektur LSTM, tahapan selanjutnya adalah melakukan evaluasi performa model menggunakan metrik *Mean Absolute Error* (MAE) dan *Root Mean Square Error* (RMSE). Evaluasi ini bertujuan untuk mengukur tingkat akurasi serta mengidentifikasi nilai deviasi estimasi terkecil dari keseluruhan skenario pemodelan yang diuji.

a. N=1

b. Data aktual- data prediksi: $388,152513 - 545,627441 = -157,474928$

c. Memutlakkan tanda minus: $-157,474928 = 157,474928$

$$MAE = 157,474928 \text{ W/m}^2$$

a. $N=1$

b. Data aktual- data prediksi: $388,152513 - 545,627441 = -157,474928$

c. Hasil pengurangan dikuadratkan: $(-157,474928)^2 = 24.798,352932$

d. Tarik akar kuadrat dari hasil kuadrat tersebut: $\sqrt{24.798,352932} = 157,474928$

$$RMSE = 157,474928 \text{ W/m}^2$$

4.7 Prediksi Cuaca Dalam Pandangan Islam

Dalam perspektif Islam, prediksi cuaca dipahami sebagai bagian dari upaya manusia membaca tanda-tanda alam (ayat kauniyah) yang diciptakan oleh Allah, tanpa mengklaim pengetahuan mutlak atas apa yang akan terjadi. Al-Qur'an menegaskan bahwa hanya Allah yang memiliki pengetahuan sempurna tentang hal-hal gaib, termasuk kondisi cuaca di masa depan. Hal ini dijelaskan dalam firman Allah dalam surah Luqman ayat 34:

إِنَّ اللَّهَ عِنْدَهُ عِلْمُ السَّاعَةِ وَيُنَزِّلُ الْغَيْثَ وَيَعْلَمُ مَا فِي الْأَرْحَامِ وَمَا تَدْرِي نَفْسٌ مَّاذَا تَكْسِبُ غَدًا وَمَا

تَدْرِي نَفْسٌ بِأَيِّ أَرْضٍ تَمُوتُ إِنَّ اللَّهَ عَلِيمٌ خَبِيرٌ

“Sesungguhnya Allah memiliki pengetahuan tentang hari Kiamat, menurunkan hujan, dan mengetahui apa yang ada dalam rahim. Tidak ada seorang pun yang dapat mengetahui (dengan pasti) apa yang akan dia kerjakan besok. Begitu pula, tidak ada seorang pun yang dapat mengetahui di bumi mana dia akan mati. Sesungguhnya Allah Maha Mengetahui lagi Maha Teliti” (Qs. Luqman/31: 34).

Dalam penelitian ini, prediksi cuaca berbasis algoritma *Long Short-Term Memory* (LSTM) merupakan manifestasi dari upaya manusia dalam memahami Sunnatullah, yakni keteraturan hukum alam yang ditetapkan oleh Allah SWT. Secara teologis, fenomena cuaca adalah tanda kekuasaan-Nya yang diperintahkan untuk dipelajari, sebagaimana firman Allah dalam QS. Ar-Rum ayat 48: "*Dialah Allah yang mengirimkan angin, lalu angin itu menggerakkan awan dan Allah membentangkannya di langit menurut yang Dia kehendaki...*".

Pemanfaatan teknologi LSTM untuk menganalisis data masa lalu selaras dengan perintah untuk melakukan *tadabbur* terhadap ayat-ayat *kauniyah* guna memitigasi ketidakpastian di masa depan. Hal ini didukung oleh hadis Rasulullah SAW:

أَنْتُمْ أَغْلَمُ بِأَمْرِ دُنْيَاكُمْ (HR. Muslim)

"Kalian lebih mengetahui tentang urusan dunia kalian" (HR. Muslim), yang memberikan legitimasi bagi pengembangan inovasi teknologi demi kemaslahatan umat. Selain itu, upaya optimasi model hingga mencapai nilai galat (*Root Mean Square Error*) terkecil merupakan bentuk pengamalan sifat Itqan, sesuai dengan sabda Nabi SAW: "Sesungguhnya Allah mencintai jika salah seorang di antara kalian melakukan suatu pekerjaan, maka ia melakukannya dengan Itqan (tekun, teliti, dan sempurna)" (HR. Thabrani).

Secara aksiologis, akurasi prediksi ini bertujuan untuk mewujudkan Masalah Mursalah melalui perlindungan jiwa dan harta benda (Hifdzun Nafs dan Hifdzul Maal) dari risiko cuaca ekstrem, sehingga penelitian ini memposisikan

sains sebagai wasilah bagi manusia untuk menunaikan amanah sebagai Khalifah fil Ardh di bumi.

Secara ilmiah, prediksi cuaca modern berkembang melalui disiplin Meteorologi, yang menggunakan data atmosfer, model matematika, dan teknologi satelit untuk memperkirakan kondisi cuaca. Dalam Islam, penggunaan ilmu pengetahuan seperti ini tidak bertentangan dengan ajaran agama, selama disadari bahwa hukum-hukum alam tersebut merupakan sunnatullah (ketetapan Allah di alam semesta). Hal ini sejalan dengan surah Ar rum ayat 48:

اللَّهُ الَّذِي يُرْسِلُ الرِّيحَ فَتُثِيرُ سَحَابًا فَيَبْسُطُهُ فِي السَّمَاءِ كَيْفَ يَشَاءُ وَيَجْعَلُهُ كِسْفًا فَتَرَى الْوَدْقَ يَخْرُجُ

مِنْ خِلَالِهِ فَإِذَا أَصَابَ بِهِ مَنْ يَشَاءُ مِنْ عِبَادِهِ إِذَا هُمْ يَسْتَبْشِرُونَ

“Allahlah yang mengirim angin, lalu ia (angin) menggerakkan awan, kemudian Dia (Allah) membentangkannya di langit menurut yang dikehendaki-Nya dan Dia menjadikannya bergumpal-gumpal, lalu engkau melihat hujan keluar dari celah-celahnya. Maka, apabila Dia menurunkannya kepada hamba-hamba-Nya yang dikehendaki-Nya, seketika itu pula mereka bergembira” (QS. Ar Rum/30: 48).

BAB V

PENUTUP

5.1 KESIMPULAN

Berdasarkan hasil eksperimen untuk menganalisis performa model *Long Short-Term Memory* (LSTM) dalam memprediksi pola perubahan cuaca melalui data historis penyinaran matahari, dapat disimpulkan beberapa poin utama berikut:

1. Skenario Model B2 dan Model A1 merupakan skenario terbaik dibandingkan arsitektur lainnya dalam memodelkan data penelitian. Rumpun arsitektur A dan B terbukti memberikan performa yang jauh lebih stabil dibandingkan rumpun skenario C.
2. Model B2 menghasilkan akurasi prediksi tertinggi dengan nilai kesalahan terkecil, yaitu *Mean Absolute Error* (MAE) sebesar 0.0911 dan *Root Mean Square Error* (RMSE) sebesar 0.1194.
3. Model A1 menunjukkan tingkat keselarasan terbaik dalam merepresentasikan data dengan nilai koefisien determinasi (R^2) tertinggi sebesar 0.4142.
4. Secara akademis, seluruh skenario model yang diuji dikategorikan belum berhasil (*unsuccessful predictive capability*) untuk melakukan prediksi dengan tingkat kepercayaan yang kuat. Hal ini diindikasikan oleh nilai R^2 yang berada di bawah ambang batas 0.50 (50%), yang menandakan bahwa model LSTM dalam konfigurasi saat ini belum mampu menjelaskan sebagian besar variabilitas data aktual karena adanya pengaruh dari variabel lain yang belum terakomodasi.

5. Skenario Model C3 secara konsisten menunjukkan performa paling tidak optimal dengan akumulasi tingkat eror tertinggi (RMSE 0.1388) dan kemampuan representasi varians data paling lemah (R^2 0.2558).

5.2 SARAN

Berdasarkan temuan penelitian, terdapat beberapa saran yang dapat dipertimbangkan untuk pengembangan studi di masa mendatang.

1. Menambahkan parameter meteorologi pendukung seperti tutupan awan (*cloud cover*), kelembapan udara, tekanan atmosfer, dan kecepatan angin sebagai input tambahan untuk mereduksi efek perataan (*smoothing*).
2. Menerapkan kombinasi model spasiotemporal seperti CNN-LSTM atau mengintegrasikan mekanisme *Attention Block* (LSTM-Transformer) guna menangkap karakteristik temporal data secara lebih kuat.
3. Perlu dilakukan eksperimen lebih mendalam terkait optimasi *hyperparameter* pada model LSTM, seperti pengujian variasi *learning rate*, penyesuaian jumlah *hidden layer*, penggunaan fungsi aktivasi yang berbeda, atau penerapan metode *regularization* yang lebih ketat untuk mengatasi keterbatasan kapabilitas prediktif model.
4. Menambah rentang waktu dataset di atas satu tahun untuk menangkap anomali iklim musiman, serta melakukan validasi lintas stasiun BMKG di luar Malang untuk menguji ketahanan (*robustness*) model.

DAFTAR PUSTAKA

- Ã, M. J. A., & Tiwari, G. N. (2011). *Solar radiation models — A review*. *March 2010*, 271–290. <https://doi.org/10.1002/er>
- Alghushairy, O., Alsini, R., Yafoz, A., Yonbawi, S., & Ma, X. (2025). Two-Stages Weather Forecasting System for Optimizing Supply Chain Operations in Saudi Arabia Using BiGRU and Fox Metaheuristics Algorithms. *Arabian Journal for Science and Engineering*. <https://doi.org/10.1007/s13369-025-10464-8>
- Andrew Gahwera, T., Steven Eyobu, O., Isaac, M., Kakuba, S., & Han, D. S. (2025). Transfer Learning-Based Ensemble Approach for Rainfall Class Amount Prediction. *IEEE Access*, 13(February), 48318–48334. <https://doi.org/10.1109/ACCESS.2025.3551737>
- Atina, A., & Prasetyo, H. (2022). Pengamatan Lamanya Penyinaran Matahari di BMKG Kelas II Kota Palembang Menggunakan Alat Campbell Stokes. *Jurnal Penelitian Fisika Dan Terapannya (JUPITER)*, 3(2), 10. <https://doi.org/10.31851/jupiter.v3i2.7148>
- BECMAN, J. A. D. & W. A. (1980). *SOLAR ENGINEERING OF THERMAL PROCESSES*.
- Boubaker, S., Benghanem, M., Mellit, A., Lefza, A., Kahouli, O., & Kolsi, L. (2021). Deep Neural Networks for Predicting Solar Radiation at Hail Region, Saudi Arabia. *IEEE Access*, 9, 36719–36729. <https://doi.org/10.1109/ACCESS.2021.3062205>
- Chicco, D., Warrens, M. J., & Jurman, G. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Computer Science*, 1–24. <https://doi.org/10.7717/peerj-cs.623>
- Damayanti, A., & Prabowo, A. (2024). Analisis Perubahan Iklim: Frekuensi Bencana Alam dan Kerugian Ekonomi Menggunakan Google. *Prosiding Seminar Nasional Sains Data*, 4(1), 662–668. <https://doi.org/10.33005/senada.v4i1.299>
- Dash, N., Chakravarty, S., Rath, A. K., Giri, N. C., AboRas, K. M., & Gowtham, N. (2025). An optimized LSTM-based deep learning model for anomaly network intrusion detection. *Scientific Reports*, 15(1), 1–21. <https://doi.org/10.1038/s41598-025-85248-z>
- Dwiyanti, Z. A., & Prianto, C. (2023). Prediksi Cuaca Kota Jakarta menggunakan Metode Random Forest : Studi Optimalitas. *Jurnal Tekno Insentif*, 17(2), 127–137.
- Hamdi, S. (2015). Mengenal lama penyinaran matahari sebagai salah satu parameter klimatologi. *Jurnal Berita Dirgantara* 15(1): 7-16. *Berita*

Dirgantara, 15(1), 7–16.

- Hartono, E. D. (2025). *Prediction Analysis of Heat Penetration in Ohmic Heating using Multivariate Long Short- Term Memory Networks*. 15(2), 22527–22537.
- Hennayake, K. M. S. A., DInalankara, R., & Mudunkotuwa, D. Y. (2021). Machine Learning Based Weather Prediction Model for Short Term Weather Prediction in Sri Lanka. *2021 10th International Conference on Information and Automation for Sustainability, ICIAfS 2021*, 299–304. <https://doi.org/10.1109/ICIAfS52090.2021.9606077>
- Hodson, T. O. (2022). Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not. *Geoscientific Model Development*, 15(14), 5481–5487. <https://doi.org/10.5194/gmd-15-5481-2022>
- Karevan, Z., & Suykens, J. A. K. (2020). Transductive LSTM for time-series prediction: An application to weather forecasting. *Neural Networks*, 125, 1–9. <https://doi.org/10.1016/j.neunet.2019.12.030>
- Khan, M., & Hossni, Y. (2025). A comparative analysis of LSTM models aided with attention and squeeze and excitation blocks for activity recognition. *Scientific Reports*, 15(1), 1–20. <https://doi.org/10.1038/s41598-025-88378-6>
- Khumaidi, A., & Nirmala, I. A. (2022). Algoritma Long Short Term Memory dengan Hyperparameter Tuning: Prediksi Penjualan Produk. In *Deepublish*.
- Kim, Y. S., Kim, M. K., Fu, N., Liu, J., Wang, J., & Srebric, J. (2025). Investigating the impact of data normalization methods on predicting electricity consumption in a building using different artificial neural network models. *Sustainable Cities and Society*, 118(October 2023), 105570. <https://doi.org/10.1016/j.scs.2024.105570>
- Lazcano, A., & Jaramillo-Morán, M. A. (2025). Data preprocessing techniques and neural networks for trended time series forecasting. *Applied Soft Computing*, 174(September 2024). <https://doi.org/10.1016/j.asoc.2025.113063>
- Nugroho, B. S., Meteorologi, S., Pangkalpinang, D. A., Bandara, J., Amir, D., Baru, P., Tengah, B., & Belitung, K. B. (2022). PENGARUH ASIMILASI DATA SATELIT HIMAWARI-8 PADA PEMODELAN CUACA WRF-ARW UNTUK PREDIKSI SIKLON TROPIS The Effect of Himawari-8 Satellite Data Assimilation on WRF-ARW Model for Tropical Cyclone Prediction. *Jurnal Sains & Teknologi Modifikasi Cuaca*, 23(1), 11–25.
- Nugroho, P. A. (2023). Implementasi Jaringan Syaraf Tiruan Multi-Layer Perceptron Untuk Prediksi Penyinaran Matahari Kota Bandung. *Komputa : Jurnal Ilmiah Komputer Dan Informatika*, 12(1), 83–90. <https://doi.org/10.34010/komputa.v12i1.9419>
- Organization, W. M. (2024). Guide to Instruments and Methods of Observation (WMO-No. 8): *World Meteorological Organization*, 1.

- Salman, A. G., Heryadi, Y., Abdurahman, E., & Suparta, W. (2018). Single Layer & Multi-layer Long Short-Term Memory (LSTM) Model with Intermediate Variables for Weather Forecasting. *Procedia Computer Science*, 135, 89–98. <https://doi.org/10.1016/j.procs.2018.08.153>
- Sarwana, L., Irma, & Ismail. (2024). Implementasi Metode K – Nearest Neighbors Untuk Prediksi Cuaca Dengan Konsep Data Mining. *Jurnal RISTER : Riset Sistem Cerdas*, 1(2), 56–61. <https://doi.org/10.25126/Rister>
- Setiawan, Z., Hiswara, A., & Muthmainah, H. N. (2023). Mengoptimalkan Jaringan Sensor Nirkabel dalam Aplikasi Monitor Lingkungan dengan Teknologi IoT di Indonesia. *Jurnal Multidisiplin West Science*, 2(10), 858–867. <https://doi.org/10.58812/jmws.v2i10.704>
- Shi, J., Wang, S., Qu, P., & Shao, J. (2024). Time series prediction model using LSTM-Transformer neural network for mine water inflow. *Scientific Reports*, 14(1), 1–16. <https://doi.org/10.1038/s41598-024-69418-z>
- Suleman, M. A. R., & Shridevi, S. (2022). Short-Term Weather Forecasting Using Spatial Feature Attention Based LSTM Model. *IEEE Access*, 10(August), 82456–82468. <https://doi.org/10.1109/ACCESS.2022.3196381>
- Sulung, U., & Muspawi, M. (2024). Memahami sumber data penelitian: primer, sekunder, dan tersier. *Tribal Architecture in Northeast India*, 5(September), 121–125. https://doi.org/10.1163/9789004263925_015
- Supriyadi, E. (2020). Prediksi Parameter Cuaca Menggunakan Deep Learning Long-Short Term Memory (LSTM) Weather Parameters Prediction Using Deep Learning Long-Short Term Memory (LSTM). *Jurnal Meteorologi Dan Geofisika*, 21(2), 55–67. <http://bmkgsoft.database.bmkg.go.id>
- Wida, dewa A. K., Sumaja, K., & Wiguna, H. (2019). The Relationship Analysis of Radiation Intensity and Sunshine Duration With Weather Parameters At Meteorological Station Ngurah Rai and Their Effect on the Potential of Solar Power Plants in South Bali. *Buletin Meteo Ngurah Rai*, 5(1), 1–7.
- Yunita, A., Pratama, M. I., Almuzakki, M. Z., Ramadhan, H., Akhir, E. A. P., Firdausiah Mansur, A. B., & Basori, A. H. (2025). Performance analysis of neural network architectures for time series forecasting: A comparative study of RNN, LSTM, GRU, and hybrid models. *MethodsX*, 15, 103462. <https://doi.org/10.1016/j.mex.2025.103462>