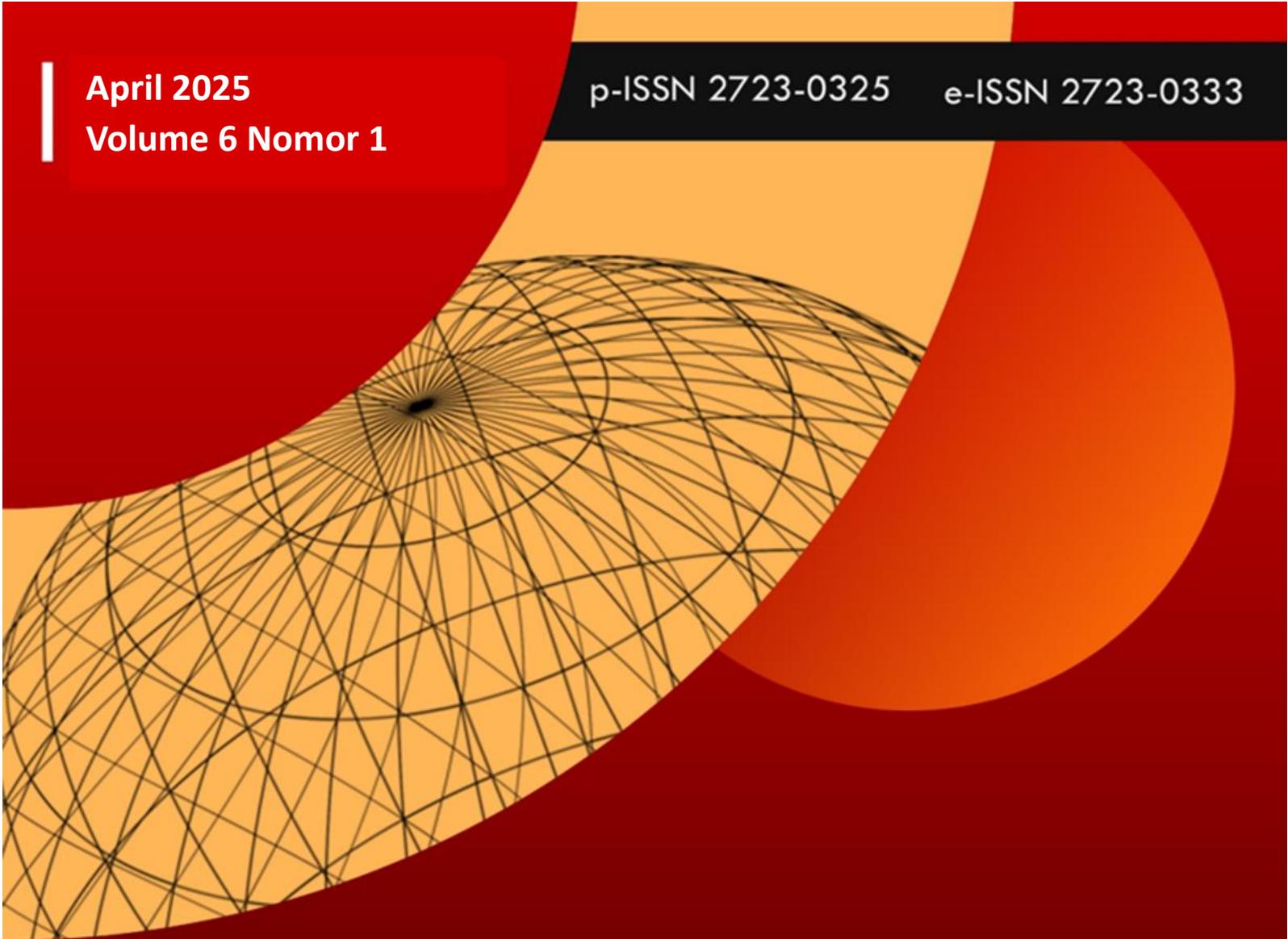


April 2025

Volume 6 Nomor 1

p-ISSN 2723-0325

e-ISSN 2723-0333



TENSOR

Pure and Applied Mathematics Journal

PROGRAM STUDI MATEMATIKA

JURUSAN MATEMATIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS PATTIMURA

TENSOR

Pure and Applied Mathematics Journal

is an international academic open-access journal that gains a foothold in mathematics, and its applications are issued twice a year. The focus is to publish original research and review articles on all aspects of pure and applied Mathematics. Editorial board members of the Journal and reviewers will review submitted papers. All submitted articles should report original, previously unpublished research results, experimental or theoretical, and will be peer-reviewed. Articles submitted to the journal should meet these criteria and must not be under consideration for publication elsewhere. Manuscripts should follow the journal template and are subject to both review and editing.

Published by:

**Department of Mathematics,
Faculty of Science and Technology,
Pattimura University.
Ambon
2025**

Copyright© Program Studi Matematika FST UNPATTI 2025

Person In Charge

Head of Undergraduate Program in Mathematics,
Faculty of Mathematics and Natural Sciences, Pattimura University

Editor in Chief

Dr. H. Batkunde, S.Si, M.Si

Editors

M. I. Tilukay, S.Si, M.Si (Managing and Section Editor)
L. Bakarbesy, S.Si, M.Si (Managing and Section Editor)
Z. A. Leleury, S.Si., M.Si (Copy and Production Editor)
B. P. Tomasouw, S.Si, M.Si (Copy and Production Editor)
Dr. L. K. Beay, S.Pd., M.Si (Proofreader)
N. Dahoklory (Proofreader)

Secretariat and Financial Officer

M. E. Rijoly, S.Si, M.Sc

Graphic Design

V. Y. I. Ilwaru, S.Si, M.Si

Expert Editorial Boards

Prof. Dr. Basuki Widodo, M.Sc (Institut Teknologi Sepuluh November Surabaya, Indonesia)
Prof. Dr. M. Salman A. N, M.Si (Institut Teknologi Bandung, Indonesia)
Dr. H. J. Wattimanela, S.Si., M.Si (Universitas Pattimura, Indonesia)
Dr. Al Azhary Masta, S.Si., M.Si (Universitas Pendidikan Indonesia, Indonesia)
Dr. Muh. Nur, S.Si., M.Si (Universitas Hasanudin, Indonesia)
Dr. Meta Kallista, S.Si., M.Si (Universitas Telkom, Indonesia)
Dr. Teguh Herlambang, S.Si., M.Si (Universitas Nahdlatul Ulama Surabaya, Indonesia)
Asst. Prof. Dr. Anurak Thanyacharoen (Muban Chombueng Rajabhat University, Ratchaburi, Thailand)

Publisher

Department of Mathematics,
Faculty of Mathematics and Natural Sciences,
Pattimura University, Ambon, Indonesia

Editorial Address

Program Studi Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Pattimura
Jln. Ir. M. Putuhena, Kampus Unpatti, Poka - Ambon 97233, Provinsi Maluku, Indonesia
Contact : +62 82397854220
Email : tensormathematics@gmail.com

Modeling the Spread of Hepatitis B Disease from the SEIR Model in East Java Using RKF45	Sakinun Na'malia Faisol Tony Yulianto	1-12
Analisis Perbandingan Optimasi <i>Stochastic Gradient Descent</i> dan <i>Adaptive Moment Estimation</i> dalam Klasifikasi Emosi dari Audio Menggunakan <i>Convolutional Neural Network</i>	Aldelia Jocelyn Tutuhaturunewa Dorteus Lodewyik Rahakbauw Zeth Arthur Leleury	13-22
Kajian Basis dan Dimensi pada Ruang Hipervektor Atas Lapangan	Loisa Genesis Kambu Henry W. M. Patty Lusye Bakarbesy Novita Dahoklory	23-38
Fungsi Trace dan Fungsi Norm Lapangan Perluasan Atas \mathbb{Q}	Novita Dahoklory Henry W. M. Patty	39-48
Optimasi Model LSTM untuk Prediksi Curah Hujan di Kota Ambon: Perbandingan Mean Imputation dan Interpolasi dalam Prediksi Data Time Series	Emanuella M. C. Wattimena Pranaya D. M. Taihuttu Devi V. Waas Citra F. Palembang Victor E. Pattiradjawane ¹	49-56
The Rainbow Vertex Connection Number of Some Amalgamation of Two Cycles	Pranaya D. M. Taihuttu Meilin I. Tilukay Francis Y. Rumlawang E. M. C. Wattimena	57-66

Optimasi Model LSTM untuk Prediksi Curah Hujan di Kota Ambon: Perbandingan Mean Imputation dan Interpolasi dalam Prediksi Data Time Series

Emanuella M. C. Wattimena^{1*}, Pranaya D. M. Taihuttu² Devi V. Waas¹, Citra F. Palembang¹, Victor E. Pattiradjawane¹

¹ Program Studi Ilmu Komputer, Jurusan Matematika, Fakultas Sains dan Teknologi, Universitas Pattimura, Ambon, Indonesia

² Program Studi Matematika, Jurusan Matematika, Fakultas Sains dan Teknologi Universitas Pattimura, Ambon, Indonesia

*Email: emanuellowattimena@gmail.com

Manuscript submitted : March 2025;

Accepted for publication : June 2025.

doi: <https://doi.org/10.30598/tensorvol6iss1pp49-56>

Abstract: Rainfall prediction is an essential aspect of meteorology, agriculture, and disaster management, particularly in regions like Ambon, where rainfall patterns significantly impact daily life. However, one of the major challenges in developing an accurate predictive model is handling missing values in the dataset. This study aims to optimize the Long Short-Term Memory (LSTM) model for rainfall prediction in Ambon by comparing two missing value handling techniques: mean imputation and interpolation. The dataset used in this study consists of daily rainfall data from 2021 to 2024, with approximately 26.89% missing values. Two experimental scenarios were conducted: the first using mean imputation to fill in missing values with the average rainfall, and the second using linear interpolation. Both scenarios utilized the same LSTM architecture to evaluate their impact on model performance. The evaluation metrics used in this study include Root Mean Square Error (RMSE) and R-squared (R^2). The results show that the interpolation-based model achieved a lower RMSE and a slightly higher R^2 value than the mean imputation-based model, indicating better predictive performance. However, both models struggled to capture extreme values, necessitating further improvements. To address this limitation, a more complex LSTM architecture was implemented in the subsequent experiments, incorporating additional layers and optimized hyperparameters. The findings suggest that choosing an appropriate missing value handling method significantly influences the predictive accuracy of LSTM models for rainfall forecasting. This research contributes to the development of more reliable weather prediction models, which can aid in agricultural planning, flood risk assessment, and climate change adaptation in Ambon.

2010 Mathematics Subject Classification: 68T05

Keywords: interpolation, LSTM, value, mean imputation, rainfall prediction

1. Pendahuluan

Curah hujan merupakan salah satu parameter iklim yang krusial dalam menentukan kondisi cuaca dan iklim suatu wilayah. Di Provinsi Maluku, khususnya di Kota Ambon, curah hujan memiliki pola musiman yang khas. Selama periode 2020 hingga 2022, curah hujan bulanan di Kota Ambon menunjukkan variasi yang signifikan. Misalnya, pada tahun 2020, curah hujan tertinggi tercatat pada

bulan Maret dengan 75 hari hujan dan total curah hujan sebesar 97,7 mm, sedangkan pada tahun 2021, puncak curah hujan terjadi pada bulan Januari dengan 54 hari hujan dan total curah hujan mencapai 124,7 mm [1].

Secara umum, Kota Ambon mengalami musim hujan yang berlangsung sekitar 7,9 bulan, dimulai dari akhir Desember hingga pertengahan Agustus, dengan puncak curah hujan terjadi antara bulan Mei hingga Agustus setiap tahunnya. Selama periode ini, kemungkinan terjadinya hari hujan melebihi 37%, dengan puncak tertinggi pada bulan Juli, di mana rata-rata terdapat 16,2 hari hujan dengan intensitas curah hujan minimal 1 mm per hari [2].

Variasi curah hujan yang signifikan ini dipengaruhi oleh letak geografis Kota Ambon yang berada di wilayah kepulauan dengan topografi yang beragam. Kondisi ini menyebabkan distribusi curah hujan yang tidak merata sepanjang tahun, dengan intensitas yang tinggi pada bulan-bulan tertentu. Sebagai contoh, data dari Stasiun Meteorologi Maritim Ambon menunjukkan bahwa pada bulan November 2021 terjadi 22 hari hujan dengan total curah hujan sebesar 180,5 mm, di mana curah hujan harian maksimum tercatat pada tanggal 14 November 2021 sebesar 46,5 mm [3].

Pemahaman terhadap pola curah hujan di Kota Ambon sangat penting, terutama dalam konteks perencanaan pembangunan dan mitigasi bencana. Informasi mengenai distribusi curah hujan dapat membantu dalam pengelolaan sumber daya air, pertanian, serta penanggulangan bencana seperti banjir dan tanah longsor yang sering terjadi akibat curah hujan yang tinggi. Oleh karena itu, analisis yang komprehensif dan akurat mengenai pola curah hujan di wilayah ini menjadi kebutuhan yang mendesak.

Namun, salah satu tantangan dalam pemodelan prediksi curah hujan adalah keberadaan *missing value* dalam dataset yang dapat terjadi karena gangguan sensor, kondisi cuaca ekstrem, atau kesalahan dalam pengumpulan data. Data yang hilang ini dapat mengurangi kualitas model prediksi jika tidak ditangani dengan metode yang tepat. Dalam penelitian terkait, metode imputasi data telah banyak digunakan untuk mengisi nilai yang hilang guna meningkatkan keakuratan analisis statistik [4].

Terdapat berbagai metode yang dapat digunakan untuk menangani *missing value*, di antaranya *mean imputation* dan interpolasi. *Mean imputation* merupakan metode sederhana yang menggantikan nilai yang hilang dengan rata-rata dari data yang tersedia. Metode ini banyak digunakan karena mudah diterapkan, tetapi memiliki kelemahan dalam menangkap pola variabilitas data. Sementara itu, interpolasi menggunakan estimasi nilai berdasarkan tren data sekitarnya, sehingga dapat mempertahankan pola fluktuasi dalam data deret waktu [5].

Seiring dengan perkembangan teknologi kecerdasan buatan, metode berbasis *deep learning*, khususnya Long Short-Term Memory (LSTM), telah terbukti efektif dalam menangani prediksi data berbasis time series. LSTM, sebagai pengembangan dari *Recurrent Neural Network* (RNN), mampu mempertahankan informasi dari data jangka panjang dan mengatasi masalah vanishing gradient yang sering terjadi pada RNN konvensional. Model ini telah banyak digunakan dalam berbagai penelitian untuk prediksi curah hujan dan telah menunjukkan hasil yang lebih baik dibandingkan metode konvensional seperti ARIMA atau regresi linear [6], [7].

Namun, efektivitas model LSTM sangat bergantung pada kualitas data yang digunakan, termasuk penanganan terhadap nilai hilang (*missing value*). Kehadiran *missing value* dalam data pelatihan dapat memengaruhi kinerja model secara signifikan jika tidak ditangani secara tepat [8]. Oleh karena itu, penelitian ini bertujuan untuk menganalisis pengaruh metode *mean imputation* dan interpolasi terhadap performa model LSTM dalam memprediksi curah hujan di Kota Ambon. Perbandingan kedua metode ini akan dilakukan untuk melihat sejauh mana masing-masing metode dapat meningkatkan akurasi prediksi curah hujan.

Dengan memahami pengaruh metode imputasi terhadap hasil prediksi, penelitian ini diharapkan dapat memberikan kontribusi dalam pengembangan strategi optimasi model LSTM serta meningkatkan kualitas prediksi curah hujan yang dapat digunakan untuk mitigasi risiko bencana hidrometeorologi dan perencanaan sumber daya air yang lebih baik.

2. Metode Penelitian

2.1. Pengumpulan Data

Data intensitas curah hujan harian yang digunakan dalam penelitian ini diperoleh dari Stasiun Meteorologi Pattimura, Kota Ambon. Rentang waktu data mencakup periode 1 Januari 2021 hingga 31 Desember 2024. Data tersebut diakses melalui portal resmi Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) dalam format CSV, yang memuat informasi tanggal pengukuran serta nilai intensitas curah hujan harian dalam satuan milimeter [9].

2.2. Praproses Data

Sebelum dilakukan analisis lebih lanjut, data yang diunduh lebih dulu melalui tahapan praproses, yang bertujuan untuk mengubah struktur data sesuai dengan kebutuhan pemodelan. Dalam penelitian ini, tahapan praproses yang harus dilakukan adalah pengecekan *missing value*. Setelah itu, *missing value* yang ditemukan akan ditangani dengan menggunakan dua skenario penanganan *missing value*, yaitu dengan menggunakan *mean imputation* dan interpolasi.

2.3. Pembagian Dataset

Data yang telah diproses sebelumnya kemudian akan dibagi menjadi dua kelompok, yaitu data latih dan data uji. Pada penelitian ini, data curah hujan dari tahun 2021 hingga 2023 digunakan sebagai data latih untuk pembuatan model dengan menggunakan LSTM. Sementara itu, data tahun 2024 digunakan sebagai data uji untuk mengevaluasi kinerja model yang telah dikembangkan.

Pembagian dataset dengan cara ini merupakan praktik standar dalam pemodelan *machine learning* untuk data deret waktu. Tujuannya adalah untuk menghindari *data leakage* dan memastikan bahwa evaluasi performa model dilakukan secara objektif dan tidak bias [10].

2.4. Pemodelan LSTM

Dalam penelitian ini, pemodelan dilakukan dengan menggunakan data univariat. Model yang akan dibuat menggunakan LSTM dengan menggunakan dua skenario pemodelan yang merupakan hasil dari penanganan *missing value* pada tahapan praproses data. Pemodelan dengan menggunakan dua skenario ini bertujuan untuk melihat pengaruh cara mengatasi *missing value* terhadap akurasi dan performa LSTM dalam melakukan prediksi intensitas curah hujan di Kota Ambon.

2.5. Evaluasi Model

Setelah menghasilkan model prediksi, evaluasi dilakukan untuk mengukur akurasi prediksi model tersebut. Penelitian ini memanfaatkan Root Mean Square Error (RMSE) dan R-square sebagai metrik evaluasi. Selanjutnya, nilai RMSE untuk kedua skenario data akan dibandingkan.

3. Hasil dan Pembahasan

3.1. Praproses Data

Praproses data merupakan Langkah awal yang krusial dalam analisis ini untuk memastikan kualitas data yang digunakan dalam pemodelan. Pada penelitian ini, proses ini meliputi beberapa tahapan utama, yaitu pengecekan *missing value*, pengecekan statistik deskriptif dan visualisasi data.

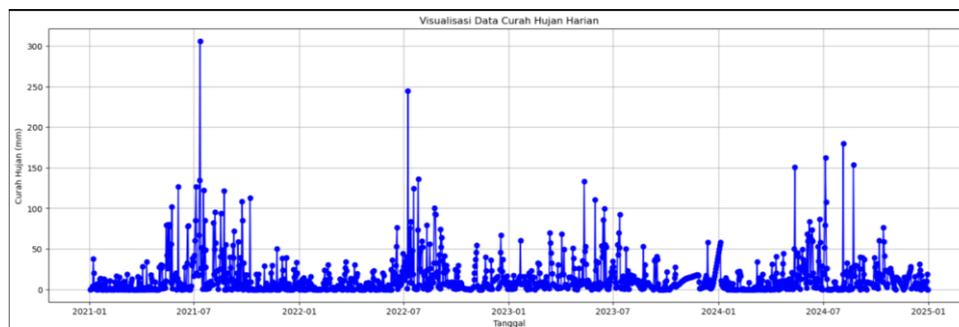
Pengecekan *missing value* dilakukan untuk mengidentifikasi apakah terdapat kekosongan data yang dapat mempengaruhi hasil analisis. Dari pemeriksaan yang dilakukan, data intensitas curah hujan harian di Kota Ambon yang diunduh melalui portal online BMKG menunjukkan terdapat sebanyak 392 data yang kosong. Kekosongan data ini kemudian ditangani dengan menggunakan metode pengisian nilai rata-rata (*imputasi mean*), mengingat distribusi data yang relatif simetris.

Setelah menangani *missing value*, analisis statistik deskriptif dilakukan untuk mendapatkan pemahaman dasar tentang distribusi curah hujan. Statistik deskriptif yang dihasilkan meliputi nilai rata-rata, median, modus, standar deviasi, serta nilai minimum dan maksimum. Statistik ini memberikan gambaran tentang *central tendency* dan disperse data, yang penting untuk menilai kecenderungan umum serta variabilitas dalam data curah hujan. Statistik deskriptif dari data intensitas curah hujan harian di Kota Ambon pada tahun 2021 hingga tahun 2024 dapat dilihat pada Tabel 1.

Tabel 1. Statistik Deskriptif Data Curah Hujan

<i>Deskriptif</i>	<i>Nilai</i>
Mean	15.293152
Standar deviasi	26.58
Min	0.00
25%	0.425
50%	5.4
75%	17.7
Max	306

Selanjutnya, visualisasi data dilakukan untuk memberikan representasi grafis yang memudahkan dalam mengamati pola *trend* yang mungkin tidak terlihat melalui analisis numerik. Visualisasi data intensitas curah hujan pada tanggal 1 Januari 2021 hingga 31 Desember 2024 dapat dilihat pada Gambar 1.



Gambar 1. Visualisasi Intensitas Curah Hujan Harian Kota Ambon tahun 2021-2024

Visualisasi curah hujan di Kota Ambon pada Gambar 1 menunjukkan bahwa terdapat periode dengan intensitas curah hujan yang sangat tinggi serta periode dengan aktivitas hujan yang relatif rendah. Grafik tersebut memperlihatkan beberapa puncak yang sangat tinggi, yang menunjukkan adanya peristiwa hujan lebat pada beberapa hari tertentu. Puncak-puncak ini menunjukkan adanya peristiwa cuaca ekstrem. Tahapan terakhir dari praproses data adalah normalisasi data dengan menggunakan normalisasi min-max.

3.2. Pembagian Data Latih dan Data Uji

Setelah menyelesaikan tahapan praproses, data kemudian dibagi menjadi data latih dan data uji. Pada penelitian ini, data curah hujan harian pada tahun 2021 hingga 2023 akan dijadikan sebagai data latih yang akan digunakan untuk pemodelan. Sementara itu, data curah hujan harian tahun 2024 akan sepenuhnya digunakan sebagai data uji untuk mengukur performa model.

3.3. Pemodelan dengan LSTM

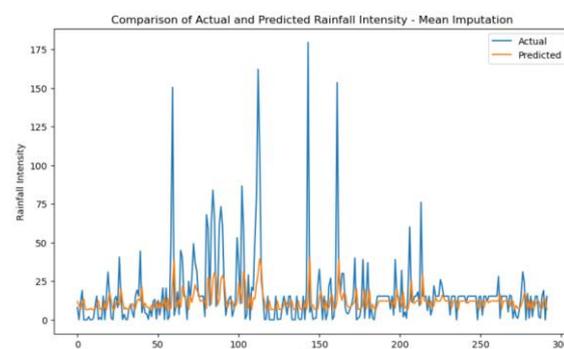
Hasil praproses data menunjukkan bahwa terdapat 392 *missing value* pada dataset curah hujan di Kota Ambon pada tahun 2021-2024. Persentase missing value pada dataset adalah 26,89%. Oleh karena itu, pemodelan pada penelitian ini akan melakukan dua percobaan terhadap dataset untuk mengoptimalkan performa model LSTM berdasarkan dua skenario penanganan missing value, yaitu *mean imputation* dan interpolasi.

Percobaan pertama dilakukan dengan menggunakan *mean imputation*, di mana setiap *missing value* akan diisi dengan rata-rata intensitas curah hujan sepanjang tahun 2021 hingga 2024, yaitu 15.293152. Setelah itu, dilakukan pemodelan dengan menggunakan LSTM. Arsitektur LSTM yang digunakan pada percobaan 1 pada kedua skenario pengisian *missing value* dapat dilihat pada Tabel 2. Jumlah epoch yang digunakan pada percobaan ini adalah sebanyak 100 epoch dan batch size 16.

Tabel 2. Arsitektur LSTM Pada Percobaan 1 dan 2

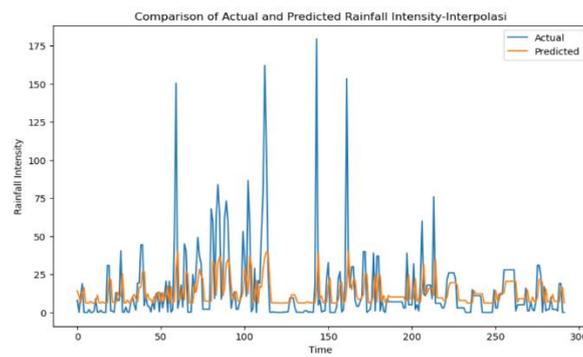
<i>Layer</i>	<i>Units</i>	<i>Activation</i>	<i>Return Sequence</i>	<i>Dropout Rate</i>	<i>Learning Rate</i>	<i>Optimizer</i>	<i>Loss Function</i>
LSTM 1	50	Tanh	True	20%	0.01	Adam	MSE
Dense (Output)	1	tanh	N/A	N/A			

Percobaan pemodelan dengan LSTM dengan menggunakan *mean imputation* sebagai salah satu cara untuk menangani *missing value* menghasilkan RMSE pada data uji sebesar 23.02, sementara *R-square* yang dihasilkan adalah sebesar 0.072. Hasil ini menunjukkan bahwa performa model dalam memprediksi curah hujan untuk satu hari ke depan belum baik. Perbandingan nilai hasil prediksi dan nilai aktual pada dataset dapat dilihat pada Gambar 2. Grafik visualisasi perbandingan tersebut menunjukkan bahwa model yang dihasilkan belum dapat melakukan prediksi dengan baik, dimana model tersebut tidak dapat mengikuti pola data aktual, dan tidak dapat memprediksi nilai-nilai ekstrem.



Gambar 2. Perbandingan Hasil Prediksi dan Aktual Percobaan 1

Selanjutnya, Percobaan 2 dilakukan dengan menggunakan skenario penanganan missing value yang kedua, yaitu interpolasi linear. Pemodelan dilakukan dengan menggunakan arsitektur LSTM yang sama dengan Percobaan 1. RMSE yang dihasilkan dari pemodelan dengan skenario ini menghasilkan RMSE 22.56 dan *R-square* sebesar 0.18. Perbandingan nilai hasil prediksi yang dihasilkan dan data aktual dapat dilihat pada Gambar 3. Nilai prediksi yang dihasilkan juga belum dapat mengikuti fluktuasi data aktual dengan baik dan belum dapat memprediksi nilai-nilai ekstrem. Tetapi *R-square* yang dihasilkan mengalami peningkatan walaupun tidak signifikan.



Gambar 3. Perbandingan Hasil Prediksi dan Aktual Percobaan 2

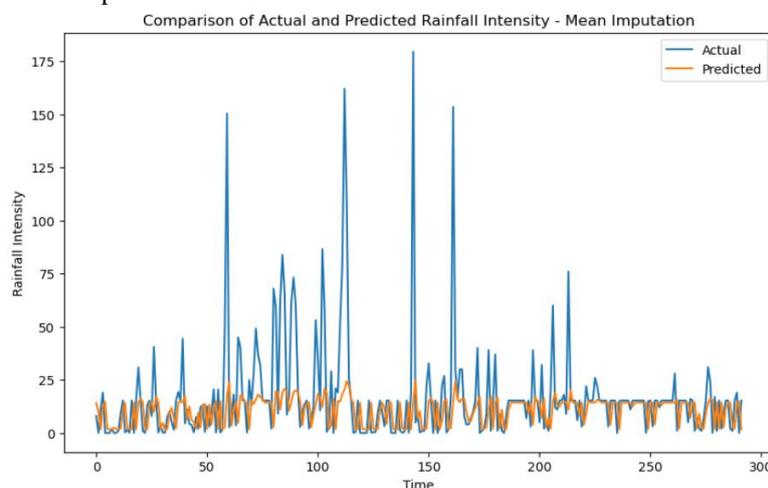
Perbandingan nilai prediksi dan nilai aktual pada Percobaan 1 dan 2 dengan menggunakan kedua skenario pemodelan seperti yang ditunjukkan pada Gambar 3 dan 4 menunjukkan bahwa model prediksi yang dihasilkan belum dapat mengikuti fluktuasi data aktual dengan baik, terutama dalam melakukan prediksi terhadap nilai-nilai ekstrem. Karena itu, pada percobaan berikutnya arsitektur LSTM akan dibuat lebih kompleks dengan melakukan penyesuaian terhadap arsitektur dan hyperparameter yang digunakan untuk meningkatkan performa model. Untuk selanjutnya, Percobaan 3 dan 4 dilakukan dengan menggunakan arsitektur LSTM pada Tabel 3.

Tabel 3. Arsitektur LSTM Pada Percobaan 3 dan 4.

<i>Layer</i>	<i>Units</i>	<i>Activation</i>	<i>Return Sequence</i>	<i>Dropout Rate</i>	<i>Learning Rate</i>	<i>Optimizer</i>	<i>Loss Function</i>
LSTM 1	100	Relu	True	30%	0.001	Adam	MAE
LSTM 2	100	Relu	False	30%	0.001	Adam	MAE
Dense (Intermediate)	50	Relu	N/A	N/A	0.001	Adam	MAE
Dense (Output)	1	relu	N/A	N/A	0.001	Adam	MAE

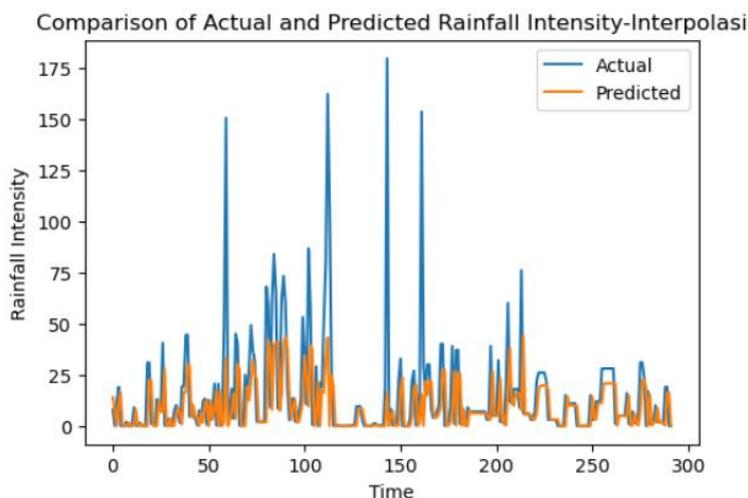
Membuat arsitektur LSTM yang lebih kompleks pada percobaan ini bertujuan agar model dapat menangkap pola-pola yang lebih rumit dalam data sehingga model dapat memahami hubungan yang lebih kompleks antara input dan output.

Arsitektur ini kemudian diterapkan pada Percobaan 3 dengan menggunakan *mean imputation* sebagai skenario penanganan *missing value*. RMSE yang dihasilkan pada data uji adalah sebesar 23.82, dan R-square yang dihasilkan sebesar 0.0063. Perbandingan hasil prediksi dan nilai aktual pada Percobaan 3 dapat dilihat pada Gambar 4.



Gambar 4. Perbandingan Hasil Prediksi dan Aktual pada Percobaan 3

Sementara itu, pada Percobaan 4 yang menggunakan interpolasi linear sebagai skenario penanganan missing value menghasilkan RMSE sebesar 22.61 dan R-square sebesar 0.12 pada data uji.



Gambar 5. Perbandingan Hasil Prediksi dan Aktual pada Percobaan 4

Berdasarkan hasil yang diperoleh dari Percobaan 3 dan 4, membuat arsitektur LSTM yang lebih kompleks pada LSTM tidak memiliki pengaruh yang signifikan terhadap performa model, dimana RMSE yang tinggi pada kedua skenario penanganan *missing value* menunjukkan bahwa model memiliki tingkat kesalahan prediksi yang cukup besar. RMSE yang tinggi menunjukkan bahwa prediksi yang dihasilkan tidak cukup akurat dalam menyesuaikan pola data aktual. Sementara itu, r-square yang dihasilkan pada data latih pada Percobaan 3 adalah sebesar -0.00096 dan mendekati nol pada data uji. Hasil ini mengindikasikan bahwa model yang dihasilkan tidak mampu menjelaskan variabilitas data dengan baik.

Dibandingkan dengan arsitektur yang pertama, penambahan kompleksitas arsitektur LSTM pada Percobaan 3 dan 4 membuat model mampu menangkap pola dasar dari curah hujan yang lebih kecil, tetapi sulit memprediksi lonjakan ekstrem seperti yang ditunjukkan pada Gambar 4 dan 5. Puncak curah hujan yang tinggi tidak dapat direpresentasikan dengan baik oleh model.

Walaupun demikian, hasil dari 4 percobaan yang telah dilakukan menunjukkan bahwa secara keseluruhan, penanganan *missing value* dengan menggunakan interpolasi linear menghasilkan RMSE yang lebih kecil dan R-square yang lebih tinggi dalam melakukan prediksi.

4. Kesimpulan

Berdasarkan hasil penelitian ini, dapat disimpulkan bahwa pemodelan curah hujan menggunakan LSTM masih belum optimal dalam memprediksi curah hujan harian di Kota Ambon. Percobaan pertama menggunakan mean imputation sebagai metode penanganan missing value menghasilkan RMSE sebesar 23.02 dan R^2 sebesar 0.072. Sementara itu, percobaan kedua dengan interpolasi linear menunjukkan sedikit peningkatan dengan RMSE sebesar 22.56 dan R^2 sebesar 0.18. Meskipun interpolasi linear memberikan hasil yang lebih baik dibandingkan *mean imputation*, model tetap mengalami kesulitan dalam menangkap pola fluktuasi data curah hujan yang ekstrem.

Pada percobaan ketiga, dilakukan peningkatan kompleksitas arsitektur LSTM dengan menambahkan lapisan dan menyesuaikan hyperparameter. Namun, hasil yang diperoleh justru menunjukkan performa yang lebih buruk, dengan RMSE meningkat menjadi 23.82 dan R^2 turun drastis menjadi 0.0063. Hal ini menunjukkan bahwa peningkatan kompleksitas model tidak selalu menghasilkan performa yang lebih baik jika tidak diiringi dengan optimasi arsitektur dan hyperparameter yang tepat.

Secara keseluruhan, penelitian ini menunjukkan bahwa metode interpolasi linear lebih baik dibandingkan mean imputation dalam menangani missing value, namun model yang digunakan masih

belum cukup akurat dalam memprediksi curah hujan. Untuk meningkatkan performa model, diperlukan optimasi lebih lanjut seperti tuning hyperparameter menggunakan *grid search* atau *Bayesian Optimization*, mencoba model alternatif seperti GRU atau *transformer*, serta menambahkan fitur lain seperti suhu dan kelembaban untuk meningkatkan akurasi prediksi. Dengan langkah-langkah ini, diharapkan model dapat menangkap pola curah hujan dengan lebih baik dan memberikan prediksi yang lebih akurat.

Daftar Pustaka

- [1] Badan Pusat Statistik (BPS) Kota Ambon. (2023). *Banyaknya Curah Hujan dan Hari Hujan di Kota Ambon Dirinci Per Bulan, 2020-2022*. Diakses dari <https://ambonkota.bps.go.id>
- [2] WeatherSpark. (2023). *Cuaca Rata-rata pada Bulan di Kota Ambon, Indonesia Sepanjang Tahun*. Diakses dari <https://id.weatherspark.com>
- [3] BMKG Stasiun Meteorologi Maritim Ambon. (2021). *Buletin Meteorologi Maritim Desember 2021*. Diakses dari <https://stamar-ambon.bmkg.go.id>
- [4] Widianti, A., & Pratama, I. (2024). *Penanganan Missing Values dan Prediksi Data Timbunan Sampah Berbasis Machine Learning*. Jurnal Teknologi dan Sistem Informasi Univarb. Volume 9 No. 2. Hal 242-251.
- [5] Noor, N. M., Abdullah, M. M. A. B., Yahaya, A. S., & Ramli, N. A. (2007). *Comparison of Linear Interpolation Method and Mean Method to Replace the Missing Values in Environmental Data Set*. ICoSM2007.
- [6] Zhang, Z., Zhang, J., & Zhou, W. (2020). Precipitation forecasting with LSTM networks: A case study in China. *Atmosphere*, 11(3), 253. <https://doi.org/10.3390/atmos11030253>
- [7] Siami-Namini, S., Tavakoli, N., & Siami Namin, A. (2019). A comparison of ARIMA and LSTM in forecasting time series. In *2018 IEEE International Conference on Big Data* (pp. 1394–1401). IEEE. <https://doi.org/10.1109/BigData.2018.8622031>
- [8] Luo, G. J., Zhou, J., & Lu, X. (2020). A comparative study of missing value imputation methods for time series data. *IEEE Access*, 8, 132146–132155. <https://doi.org/10.1109/ACCESS.2020.3009902>
- [9] BMKG. (2024). *Data online: Statistik curah hujan harian Kota Ambon*. Badan Meteorologi, Klimatologi, dan Geofisika. <https://dataonline.bmkg.go.id/>
- [10] Brownlee, J. (2018). *Deep learning for time series forecasting: Predict the future with MLPs, CNNs and LSTMs in Python*. Machine Learning Mastery.