

PENERAPAN METODE LONG SHORT-TERM MEMORY DALAM MEMPREDIKSI DATA METEOROLOGI DI KALIMANTAN TIMUR

David Jansen¹, Teny Handhayani*², Janson Hendryli³

^{1,2,3}Teknik Informatika, Universitas Tarumanagara,

Jl. Letjen S. Parman No. 1, Jakarta Barat 11440, Indonesia

e-mail: ¹david.535190038@stu.untar.ac.id, *²tenyh@fti.untar.ac.id, ³Jansonh@fti.untar.ac.id

Meteorologi adalah cabang dari ilmu atmosfer yang memiliki banyak variabel didalamnya. Variabel seperti curah hujan, kecepatan angin, temperatur, dan kelembapan mempengaruhi aktivitas dan keputusan yang akan diambil pada kehidupan manusia dalam banyak bidang. Variabel pada meteorologi ini dapat membahayakan sebagian bahkan seluruh dunia. Memprediksi variabel tersebut dapat meningkatkan keputusan yang akan diambil sehingga meningkatkan kualitas hidup manusia dalam berbagai bidang. Deep Learning adalah cabang dari Machine Learning yang dapat digunakan untuk mempelajari data inputan yang belum terklasifikasi dengan tujuan untuk melakukan klasifikasi pada data tersebut dan mendeteksi objek atau pola yang ada. Long Short-Term Memory adalah bagian dari Deep Learning yang dapat digunakan untuk memprediksi variabel tersebut. Menggunakan Recurrent Neural Network (RNN) sebagai dasarnya dan menambahkan sel memori didalamnya yang telah terbukti menyelesaikan masalah vanishing gradient yang ada pada Recurrent Neural Network sehingga dapat melakukan prediksi untuk data dengan deret waktu yang besar. Long Short-Term Memory cocok digunakan untuk melakukan klasifikasi dan melakukan prediksi terhadap data deret waktu. Data yang digunakan dalam melakukan prediksi ini adalah data stasiun meteorologi Kalimarau, stasiun meteorologi Sultan Aji Sulaiman Sepingga, dan stasiun meteorologi Aji Pangeran Tumenggung Pranoto di Kalimantan Timur pada periode Januari 2010 – Juni 2022 yang diambil melalui Badan Meteorologi Klimatologi dan Geofisika Indonesia (BMKG)..

Kata Kunci: Deep learning, LSTM, meteorologi

I. PENDAHULUAN

Meteorologi adalah studi yang berfokus pada mempelajari hal-hal yang terjadi di atmosfer seperti temperatur, kelembapan, curah hujan, dan kecepatan angin. Berdasarkan data meteorologi ini manusia menentukan keputusan dan aktivitasnya selama jangka waktu tertentu dikarenakan data meteorologi ini mempengaruhi keberhasilan dan gagalnya suatu hal yang

telah direncanakan. Terlebih data meteorologi ini mempengaruhi berbagai sektor seperti pembangunan, industri, pertanian, dan lain-lain.

Kalimantan Timur sedang menjalani pembangunan untuk menjadi ibu kota Indonesia yang baru yang diperkirakan akan selesai pada tahun 2024. Perubahan ibu kota pada suatu negara tentunya akan berdampak besar kepada masyarakat terlebih untuk mereka yang bekerja untuk pemerintah. Tidak hanya lokasi pekerjaan mereka namun lingkungan hidup mereka pun akan berubah sehingga akan mempengaruhi kehidupan dan aktivitas mereka.

Algoritma Long Short-Term Memory (LSTM) adalah bagian dari Deep Learning yang sebelumnya telah digunakan oleh peneliti untuk melakukan prediksi mengenai berbagai hal seperti pengenalan pola, mendeteksi objek, dan sebagainya. Algoritma LSTM merupakan Recurrent Neural Network (RNN) yang telah dikembangkan dengan menambahkan sel memori kedalamnya yang dikenalkan oleh Hochreiter dan Schmidhuber pada tahun 1997 [1]. Pada tahun 2020, Dina Zatusviya Haq et al menggunakan algoritma LSTM untuk memprediksi curah hujan berdasarkan El Nino dan data IOD. Menggunakan 300 hidden layers, 32, 64, 128, dan 256 batch size, dan 50, 100, 150 learning drop rate. Hasil yang didapat dari penelitian tersebut adalah 0.6218 yang merupakan nilai rata-rata MAAPE [2]. Pada tahun 2022, Tita Lafitta, Putu Wira Buana, dan Ni Kadek Dwi Rusjayanthi menggunakan algoritma LSTM untuk melakukan prediksi temperatur dan curah hujan menggunakan 50, 100, 200, dan 400 epochs. Hasil RMSE dan MAPE terbaik untuk curah hujan adalah 1.7444 dan 1.0164% sedangkan hasil RMSE dan MAPE terbaik untuk temperatur adalah 0.6929 dan 1.9499% [3]. Dengan memprediksi data meteorologi di Kalimantan Timur, pemerintah dan masyarakat diharapkan dapat menggunakan data tersebut sebagai referensi sebelum pembangunan ibu kota Indonesia yang baru selesai.

II. LANDASAN TEORI

Objek dari penelitian ini adalah data meteorologi di Kalimantan Timur yaitu stasiun meteorologi Kalimarau, stasiun meteorologi Aji Pangeran Tumenggung Pranoto,

dan stasiun meteorologi Sultan Aji Sulaiman Sepinggan. Data meteorologi yang diamati adalah temperature, kelembapan, curah hujan, kecepatan angin pada periode Januari 2010 hingga Juni 2022.

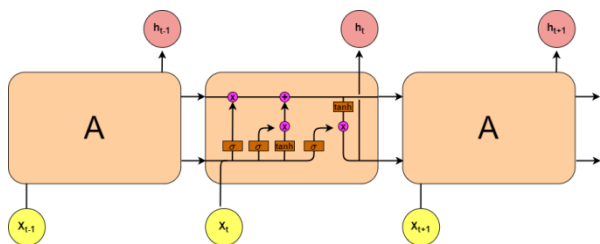
Data dari setiap stasiun meteorologi tersebut didapat melalui Badan Meteorologi Klimatologi dan Geofisika (BMKG). Proses pada penelitian ini adalah pengumpulan data, pre-processing, pelatihan model, dan evaluasi model. Setelah data dikumpulkan akan dilakukan pre-processing untuk mengisi data yang kosong, pre-processing terdiri dari beberapa tahap yaitu membersihkan data, menggabungkan semua data yang telah dibersihkan, dan mengisi data kosong. Setelah itu data yang ada akan dilatih kedalam sebuah model yang kemudian akan dievaluasi.

A. Deep Learning

Deep learning merupakan cabang dari Machine Learning dengan basis Artificial Intelligent (AI) yang berguna untuk meniru cara pikir manusia [4]. Machine Learning sendiri disebut sebagai proses penyelesaian masalah tanpa campur tangan manusia dengan cara mengumpulkan data dan membangun sebuah model dari kumpulan data yang ada [5]. Model tersebut nantinya akan digunakan untuk menyelesaikan masalah yang ada [6]. Arsitektur deep learning memiliki beberapa lapisan tersembunyi yang digunakan untuk mempelajari fitur yang ada pada beberapa tingkatan abstraksi [7]. Algoritma Deep Learning menggunakan lapisan tersebut untuk menemukan representasi yang baik antara fitur yang telah dipelajari pada tingkat yang lebih tinggi maupun lebih rendah [8].

B. Long Short-Term Memory

Algoritma Long Short-Term Memory adalah Recurrent Neural Network yang telah dimodifikasi dengan menambahkan cell state yang berguna untuk mengatasi vanishing gradient problem yang terjadi pada RNN [9]. RNN adalah jaringan yang dirancang untuk menerima data masukan atau keluaran secara sekuensial [10]. RNN berbeda pada jaringan saraf tradisional karena dalam prosesnya RNN menyimpan informasi dari input sebelumnya untuk mempengaruhi output yang akan dikeluarkan [11]. LSTM terdiri dari forget gate, input gate, output gate, dan cell state [12]. Arsitektur LSTM dapat dilihat pada Gambar 1.



Gambar 1. Arsitektur LSTM

Gerbang pertama pada arsitektur LSTM adalah forget gate yang berfungsi untuk menentukan informasi yang akan disimpan di cell state. Perhitungan forget gate dapat dilihat pada rumus (1) – (6).

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{1}$$

- f_t = forget gate
- σ = fungsi sigmoid
- W_f = bobot forget gate
- h_{t-1} = keluaran timestep-1
- x_t = masukan timestep terkini
- b_t = bobot bias forget gate

Gerbang kedua pada arsitektur LSTM adalah input gate yang terdiri dari sebuah fungsi sigmoid yang menentukan informasi mana yang akan diperbaharui serta sebuah vektor akan dibuat yang berguna untuk menyimpan informasi yang berkemungkinan menjadi informasi yang akan diperbaharui [13].

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{2}$$

$$C'_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \tag{3}$$

- i_t = input gate
- σ = fungsi sigmoid
- W_i = bobot input gate
- h_{t-1} = keluaran timestep-1
- x_t = masukan timestep terkini
- b_i = bobot bias input gate
- C'_t = kandidat cell state baru
- W_c = bobot cell state
- b_c = bobot bias cell state
- \tanh = fungsi tanh

Hasil dari forget gate dan input gate akan menjadi komponen untuk menentukan cell state yang baru

$$C_t = f_t * C_{t-1} + i_t * C'_t \tag{4}$$

- C_t = cell state
- f_t = forget gate
- i_t = input gate
- C_{t-1} = cell state timestep-1
- C'_t = kandidat cell state baru

Gerbang terakhir pada arsitektur LSTM adalah output gate yang berfungsi untuk menghasilkan keluaran LSTM pada timestep terakhir.

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{5}$$

$$h_t = o_t * \tanh(C_t) \tag{6}$$

- o_t = output gate
- σ = fungsi sigmoid
- W_o = bobot output gate
- h_{t-1} = keluaran timestep-1
- x_t = masukan timestep terkini
- b_o = bobot bias output gate
- h_t = keluaran timestep terkini
- C_t = cell state
- \tanh = fungsi tanh

C. Metrik Evaluasi

Metrik evaluasi yang digunakan pada penelitian ini adalah MAE, MSE, dan RMSE. MAE menghitung rata-

rata absolut dari kesalahan prediksi yang tidak terlalu terpengaruhi oleh data outlier. Sedangkan MSE dan RMSE digunakan untuk mencerminkan kesalahan model [14],[15]. Perhitungan MAE, MSE dan RMSE dapat dilihat pada rumus (7), (8), dan (9).

$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_i - Y'_i| \tag{7}$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - Y'_i)^2 \tag{8}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_i - Y'_i)^2} \tag{9}$$

- n = total data
- y_i = data aktual
- y'_i = data prediksi

III. METODE PENELITIAN

Penelitian ini menggunakan metode LSTM. Dataset yang digunakan merupakan data meteorologi deret waktu dari Badan Meteorologi Klimatologi dan Geofisika melalui <http://dataonline.bmkg.go.id/>. Data dikumpulkan dari periode Januari 2010 – Juni 2022. Parameter yang diamati yaitu suhu rata-rata, curah hujan rata-rata, kelembaban rata – rata dan kecepatan angin rata – rata. Data latih yaitu data dari Januari 2010 – Desember 2020. Data uji yaitu sampel dari Januari 2021 – Juni 2022.

Data yang telah dikumpulkan akan melalui tahap preprocessing yang terdiri dari dua tahap dimana pada tahap pertama data yang bernilai 8888 dan 9999 serta data yang kosong akan diisi menggunakan teknik interpolasi linear. Perubahan nilai tersebut akan diberlakukan untuk setiap data dengan rentang waktu 1 bulan pada setiap stasiun untuk mengatasi data outlier. Tahap kedua adalah melakukan normalisasi data menggunakan minmaxscaller yang berguna untuk menskalakan nilai yang ada kedalam rentang nilai 0 hingga 1. Normalisasi ini dilakukan agar tidak terdapat perbedaan signifikan terhadap nilai yang satu dengan lainnya. Hasil prediksi dievaluasi menggunakan nilai MAE, MSE, dan RMSE.

IV. HASIL DAN PEMBAHASAN

Penelitian ini menggunakan data pada periode 2010 – 2020 sebagai data latih dan melakukan validasi menggunakan data pada tahun 2021. Terdapat 2 model yang digunakan sebagai perbandingan dimana model pertama terdiri dari 3 layer LSTM dengan unit sebesar 64, 64, dan 32 sedangkan model kedua terdiri dari 3 layer LSTM dengan unit sebesar 128, 64, dan 32. Kedua model ini menggunakan learning rate sebesar 0.001, adam optimizer, batch size dengan nilai 32 sebanyak 50 epochs. Konfigurasi dari kedua model ini dapat dilihat pada Tabel 1 dan Tabel 2.

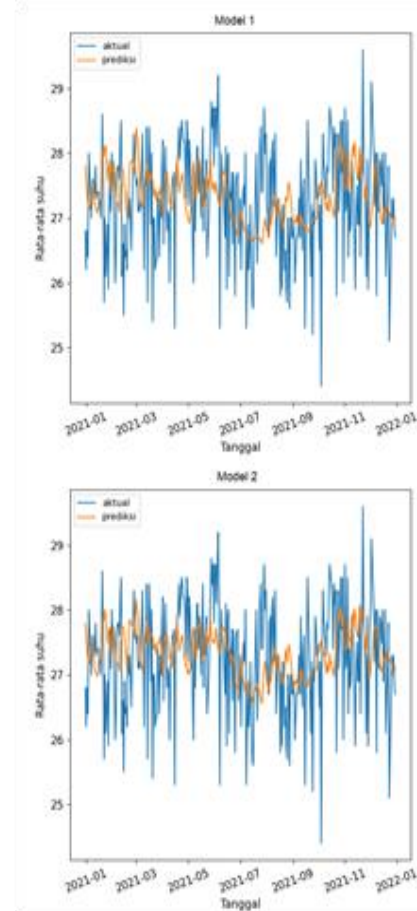
Tabel 1. Konfigurasi Model LSTM 1

Layer	Parameter	Output Shape
LSTM	64, relu	(None, 30, 64)
LSTM_1	64, relu	(None, 30, 64)
Dropout	0.3	(None, 30, 64)
LSTM_2	32, relu	(None, 32)
Dropout_1	0.3	(None, 32)
Dense	1	(None, 1)

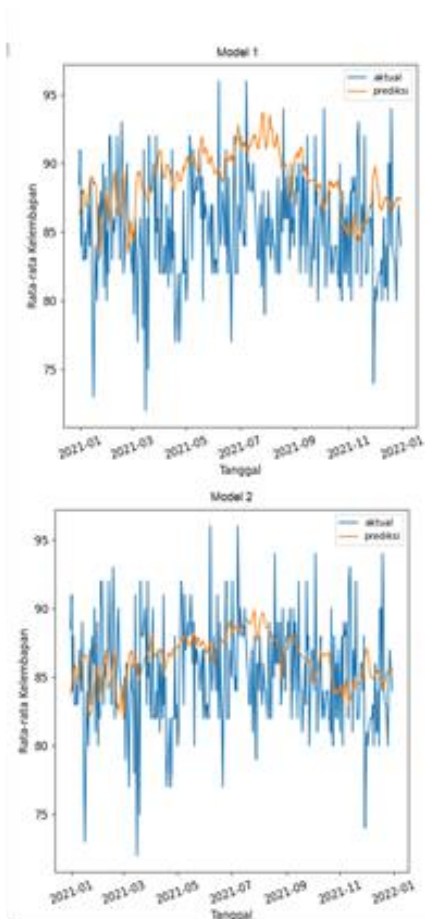
Tabel 2. Konfigurasi Model LSTM 2

Layer	Parameter	Output Shape
LSTM	128, relu	(None, 30, 128)
LSTM_1	64, relu	(None, 30, 64)
Dropout	0.3	(None, 30, 64)
LSTM_2	32, relu	(None, 32)
Dropout_1	0.3	(None, 32)
Dense	1	(None, 1)

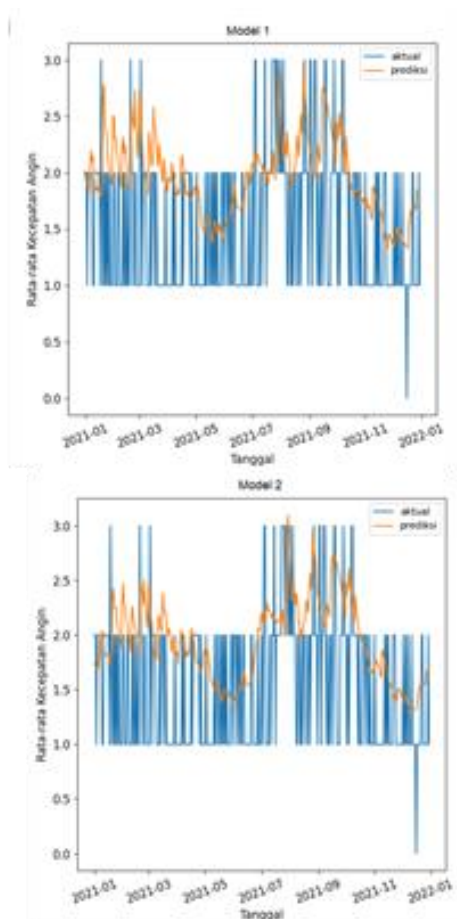
Nilai prediksi yang dihasilkan oleh kedua model tersebut tidak terlalu berbeda, menggunakan data latih dari tahun 2010 hingga 2020 dan melakukan evaluasi menggunakan data tahun 2021 pada stasiun Sultan Aji Sulaiman Sepinggang ditampilkan pada Gambar 2, 3, 4 dan 5. Berdasarkan nilai prediksi dan nilai actual pada tahun 2021 dihasilkan metrik evaluasi untuk setiap model yang dapat dilihat pada Tabel 3 dan Tabel 4.



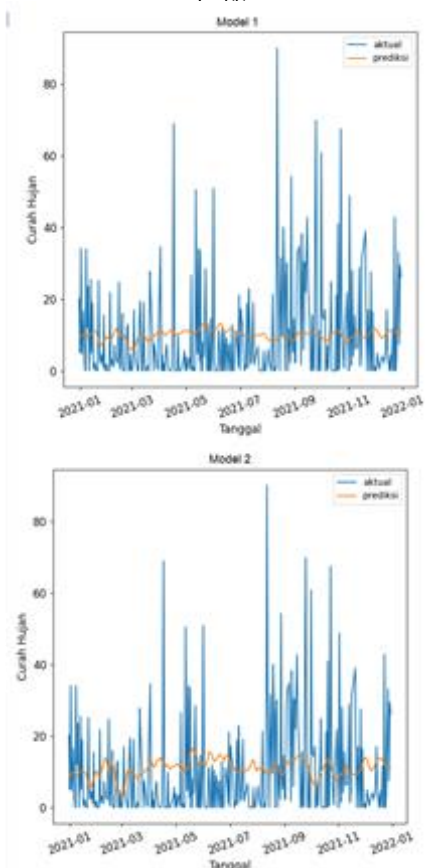
Gambar 2. Perbandingan Suhu Pada Stasiun Sultan Aji Sulaiman Sepinggang



Gambar 3. Perbandingan Kelembapan Pada Stasiun Sultan Aji Sulaiman Sepinggang



Gambar 5 Perbandingan Kecepatan Angin Pada Stasiun Sultan Aji Sulaiman Sepinggang



Gambar 4. Perbandingan Curah Hujan Pada Stasiun Sultan Aji Sulaiman Sepinggang

Tabel 3. Evaluasi Metrik Model LSTM 1

Stasiun	Parameter	MSE	RMSE	MAE
Kalimara	Temperatur	0.84	0.91	0.72
	Kelembapan	23.56	4.85	3.88
	Kecepatan Angin	0.27	0.52	0.38
Sultan Aji Muhammad Sulaiman Sepinggang	Curah Hujan	163.47	12.79	8.97
	Temperatur	0.75	0.86	0.68
	Kelembapan	30.44	5.52	4.59
Aji Pangeran Tumenggung Pranoto	Kecepatan Angin	0.54	0.74	0.62
	Curah Hujan	185.23	13.61	10.10
	Temperatur	1.41	1.19	0.92
Aji Pangeran Tumenggung Pranoto	Kelembapan	23.41	4.84	3.83
	Kecepatan Angin	0.38	0.62	0.47
	Curah Hujan	220.60	14.85	9.44

Tabel 4. Evaluasi Metrik Model LSTM 2

Stasiun	Parameter	MSE	RMSE	MAE
Kalimara	Temperatur	0.83	0.91	0.73
	Kelembapan	24.64	4.96	3.98
	Kecepatan Angin	0.28	0.53	0.39
Sultan Aji Muhammad Sepinggang	Curah Hujan	155.81	12.48	8.25
	Temperatur	0.72	0.85	0.67
	Kelembapan	17.40	4.17	3.37
Aji Pangeran Tumenggung Pranoto	Kecepatan Angin	0.52	0.72	0.61
	Curah Hujan	189.99	13.78	10.55
	Temperatur	1.49	1.22	0.95
Aji Pangeran Tumenggung Pranoto	Kelembapan	20.76	4.56	3.61
	Kecepatan Angin	0.38	0.62	0.50
	Curah Hujan	212.14	14.56	9.66

V. KESIMPULAN

Berdasarkan penelitian ini, dapat disimpulkan bahwa model 2 memiliki nilai RMSE yang lebih rendah dibandingkan model 1 secara keseluruhan namun memiliki nilai MAE yang lebih tinggi dibandingkan dengan model 1. Kedua model memiliki nilai error yang cukup baik jika dilihat melalui metrik evaluasi yang digunakan. Sehingga dapat dikatakan bahwa metode Long Short-Term Memory bekerja dengan baik dalam memprediksi data meteorologi di Kalimantan Timur.

UCAPAN TERIMA KASIH

Penulis mengucapkan terimakasih kepada LPPM Universitas Taarumanagara yang menjadi sponsor dalam penelitian ini

DAFTAR PUSTAKA

[1] F. Masri, D. Saepudin, dan D. Adytia, "Forecasting of Sea Level Time Series using Deep Learning RNN, LSTM, and BiLSTM, Case Study in Jakarta Bay, Indonesia," in *e-Proceeding of Engineering*, vol. 7, no. 2, pp. 8544-8551, 2020.

[2] Dina Zatusiva Haq et al., "Long Short-Term Memory Algorithm for Rainfall Prediction Based on El-Nino and IOD Data" in *5th International Conference on Computer Science and Computational Intelligence*, pp. 829 – 837, 2022.

[3] T. Lattifia, P. W. Buana, dan N. K. D. Rusjyanthi, "Model Prediksi Cuaca Menggunakan Metode LSTM," in *Jurnal Ilmiah Teknologi dan Komputer (JITTER)*, vol. 3, no. 1, 2022.

[4] Asmaa Halbouni et al., "Machine Learning and Deep Learning Approaches for CyberSecurity: A Review," in *IEEE Access*, pp 19572-19588, 2022.

[5] Dur-E-Maknoon Nisar et al., "Healthcare Techniques Through Deep Learning: Issues, Challenges and Opportunities," in *IEEE Access*, pp. 98523-98541, 2021.

[6] Andriy Burkov, *The Hundred-Page Machine Learning Book*, Quebec: Andriy Burkov, 2019.

[7] P.H. Gunawan, D. Munandar, and A.Z. Farabiba, "Long Short-Term Memory Approach for Predicting Air Temperature in Indonesia," in *Jurnal Online Informatika (JOIN)*, vol. 5, no. 2, pp. 161-168, 2020.

[8] M. Arif Wani et al., *Advances in Deep Learning*, Warsaw: Janusz Kacprzyk, 2020.

[9] Y. O. Ouma, R. Cheruyot, dan A. N. Wachera, "Rainfall and runoff time-series trend analysis using LSTM recurrent neural network and wavelet neural network with satellite-based meteorological data: case study of Nzoia hydrologic basin," in *Complex & Intelligent Systems*, pp. 213-236, 2021.

[10] N. M.Rezk, M. Purnaprajna, T. Nordstrom, dan Z. Ul-abdin, "Recurrent Neural Networks: An Embedded Computing Perspective," in *IEEE Access*, pp. 57967-57996, 2020.

[11] M. A. Faishol, Endroyono, dan A. N. Irfansyah, "Prediksi Polusi Udara Perkotaan Di Surabayamenggunakan Recurrent Neural Network – Long Short-Term Memory," in *Jurnal Ilmiah Teknologi Informasi (JUTI)*, vol. 18, no. 2, pp. 102-114, 2020

[12] Chang Jiang Zhang et al., "Correction model for rainfall forecasts using the LSTM with multiple meteorological factors," in *Royal Meteorological Society*, vol. 27, pp. 1-15, 2019.

[13] Qi Tang et al., "Prediction of Financial Time Series Based on LSTM Using Wavelet Transform and Singular Spectrum Analysis," in *Hindawi*, pp. 1-13, 2021.

[14] Qing Tao et al., "Air Pollution Forecasting Using a Deep Learning Model Based on 1D Convnets and Bidirectional GRU," in *IEEE Access*, pp. 76690 -76698, 2019.

[15] H. Freecenta, E. Y. Puspaningrum, dan H. Maulana, "Prediksi Curah Hujan di Kab. Malang Menggunakan LSTM (Long Short-Term Memory)," in *Jurnal Informatika dan Sistem Informasi*, vol. 3, no. 1, pp. 51-55, 2022.