

Pengembangan Model Prediksi Curah Hujan di Kota Denpasar Menggunakan Metode LSTM dan GRU

I Wayan Aditya Suranata
Universitas Pendidikan Nasional
e-mail: aditya@undiknas.ac.id

Diajukan: 7 Oktober 2023; Direvisi: 30 Oktober 2023; Diterima: 3 November 2023

Abstrak

Perubahan iklim menjadi isu penting yang perlu dicermati karena memengaruhi berbagai sektor, termasuk pertanian yang sangat bergantung pada curah hujan karena menentukan jadwal tanam dan panen. Oleh karena itu, penelitian ini mengembangkan model prediksi curah hujan menggunakan metode Long Short-Term Memory (LSTM) dan Gated Recurrent Units (GRU). Data yang digunakan dalam pembuatan model diperoleh dari Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) dan diukur selama satu tahun pada 2019. Dataset terdiri dari 10 atribut yang digunakan sebagai acuan pengukuran cuaca oleh BMKG. Kemudian data tersebut dianalisis untuk memperbaiki nilai yang hilang dan melakukan pelabelan untuk menyamakan tipe data. Dari 10 atribut, hanya 7 atribut yang digunakan dalam proses pemodelan. Hasil pemodelan menunjukkan bahwa metode LSTM menghasilkan nilai RMSE sebesar 8,853, MAE sebesar 4,090, dan MSE sebesar 78,383, sedangkan metode GRU menghasilkan nilai RMSE sebesar 9,698, MAE sebesar 4,291, dan MSE sebesar 94,058. Berdasarkan hasil ini, metode LSTM memiliki tingkat error yang lebih rendah dibandingkan GRU dalam memprediksi curah hujan. Penelitian ini diharapkan dapat memberikan kontribusi pada bidang meteorologi dalam memprediksi curah hujan serta berperan dalam penanganan perubahan iklim di masa depan.

Kata kunci: Curah hujan, LSTM, GRU, Kecerdasan buatan.

Abstract

Climate change is an important issue that needs to be addressed as it affects various sectors, including agriculture, which heavily relies on rainfall to determine planting and harvesting schedules. Therefore, this study develops a rainfall prediction model using Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU) methods. The data used for model development was obtained from the Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) and measured over one year in 2019. The dataset consists of 10 attributes used as weather measurement references by BMKG. The data was then analyzed to correct missing values and perform labeling to standardize data types. Of the 10 attributes, only 7 were used in the modeling process. The modeling results show that the LSTM method produced an RMSE value of 8.853, MAE of 4.090, and MSE of 78.383, while the GRU method produced an RMSE value of 9.698, MAE of 4.291, and MSE of 94.058. Based on these results, the LSTM method has a lower error rate compared to GRU in predicting rainfall. This research is expected to contribute to the field of meteorology in predicting rainfall and play a role in addressing climate change in the future.

Keywords: Rainfall, LSTM, GRU, Artificial Intellegent.

1. Pendahuluan

Perubahan iklim yang menyebabkan pemanasan global menjadi isu lingkungan yang paling banyak dibicarakan saat ini, baik dalam tataran ilmiah maupun dalam kehidupan sehari-hari. Hingga saat ini, bumi telah mengalami peningkatan suhu sebesar 1,1°C, seperti yang dipublikasikan oleh Organisasi Meteorologi Dunia (WMO). Perubahan iklim ini menjadi fenomena yang harus dihadapi oleh berbagai pihak untuk menangani kondisi ini, karena perubahan iklim sangat mengancam kehidupan semua makhluk hidup. Dalam penelitian ini, salah satu fenomena perubahan iklim yang diteliti adalah curah hujan. Dampak dari perubahan curah hujan dapat mengakibatkan kekeringan dan banjir. Selain itu, berbagai sektor juga terdampak, salah satunya adalah pertanian yang sangat bergantung pada kondisi cuaca untuk menentukan jadwal tanam dan panen, serta dalam manajemen sumber daya air. Sektor-sektor lain juga sangat bergantung pada curah hujan, yang berpotensi menyebabkan kerugian ekonomi. Hal ini tentunya tidak

terlepas dari adanya perubahan iklim. Oleh karena itu, penelitian ini mengembangkan sebuah model yang dapat memprediksi kondisi cuaca, khususnya curah hujan, dengan menggunakan konsep kecerdasan buatan melalui metode LSTM dan GRU [1][2].

Kecerdasan buatan (AI) adalah teknologi yang memungkinkan komputer atau perangkat digital untuk belajar dan mengerjakan tugas-tugas seperti halnya manusia. AI memiliki berbagai jenis, termasuk machine learning, natural language processing, robotic, deep learning, dan lainnya [3]. Dalam penelitian ini, yang berfokus pada prediksi curah hujan sehingga metode yang digunakan adalah deep learning. Metode ini efektif dalam memodelkan data dalam bentuk deret waktu seperti data curah hujan, karena mampu mempelajari pola data curah hujan untuk menghasilkan prediksi masa depan. Dalam deep learning, metode yang digunakan dalam penelitian yaitu Long Short-Term Memory (LSTM) dan Gated Recurrent Units (GRU). Kedua metode ini memiliki arsitektur yang mirip dan termasuk dalam kategori Recurrent Neural Network (RNN). Perbedaannya terletak pada kesederhanaan arsitektur GRU dibandingkan dengan LSTM. GRU memiliki komputasi yang lebih efisien dan cocok untuk dataset yang lebih kecil, sedangkan LSTM memiliki proses komputasi yang lebih kompleks dan memerlukan sumber daya komputasi yang lebih besar [4].

Pada penelitian ini, kedua metode LSTM dan GRU dimanfaatkan untuk membuat model prediksi curah hujan. Pemilihan kedua metode ini didasarkan pada kemampuannya dalam menangani data deret waktu yang kompleks dan dalam menangkap hubungan temporal jangka panjang. LSTM dan GRU memiliki arsitektur yang dapat mengatasi masalah vanishing gradient, sehingga lebih akurat dalam memprediksi pola cuaca yang fluktuatif dan tidak teratur seperti curah hujan. Data yang digunakan untuk pemodelan ini diperoleh dari Badan Meteorologi, Klimatologi, dan Geofisika (BMKG), yang memiliki otoritas dalam pengukuran cuaca di Indonesia. Data tersebut mencakup beberapa atribut, salah satunya curah hujan. Setelah data diperoleh, dilakukan analisis untuk persiapan pemodelan. Kemudian hasil dari pemodelan kedua model akan dievaluasi untuk mengetahui performanya. Selanjutnya, performa model dibandingkan untuk menentukan metode yang lebih cocok digunakan dalam konteks prediksi curah hujan. Diharapkan penelitian ini dapat memberikan kontribusi bagi bidang meteorologi dan memberikan wawasan baru dalam pemanfaatan teknologi kecerdasan buatan untuk mengatasi masalah perubahan iklim. Selain itu, penelitian ini dapat menjadi dasar dalam pengembangan prediksi cuaca yang lebih cerdas dan tangguh di masa depan.

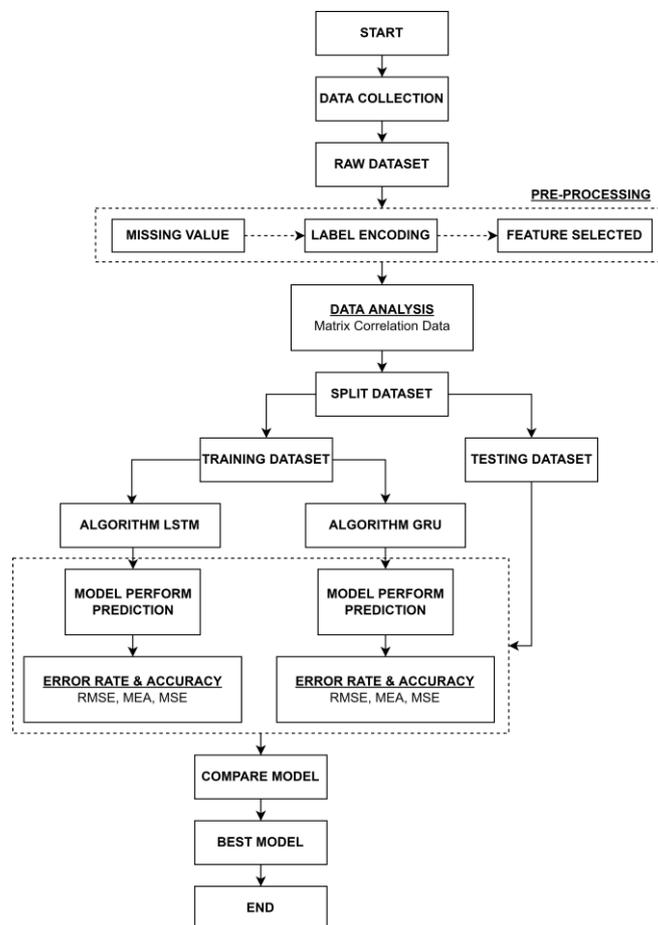
2. Metode Penelitian

2.1. Alur Kerja Penelitian

Alur kerja penelitian ini diperlihatkan pada Gambar 1, yang menunjukkan proses penelitian dari pengumpulan data hingga menemukan model terbaik untuk memprediksi kondisi hujan. Berdasarkan alur kerja tersebut, kami memulai dengan tahap pengumpulan data. Proses pengumpulan data ini dilakukan untuk mendapatkan data terkait perkiraan cuaca yang akan digunakan dalam penelitian. Data dikumpulkan dari pihak terkait atau lembaga resmi yang melakukan pengukuran cuaca sehingga data yang digunakan sudah tervalidasi kebenarannya. Setelah memperoleh data yang dibutuhkan, langkah selanjutnya adalah melakukan proses pre-processing. Karena data yang diperoleh masih dalam bentuk mentah, tahap pre-processing sangat penting dilakukan untuk menghasilkan model yang baik. Pada tahap pre-processing, beberapa langkah dilakukan untuk menyiapkan data agar lebih siap untuk proses pemodelan dan meningkatkan kualitas data. Proses pre-processing ini mencakup beberapa langkah utama, yaitu mengecek data yang hilang (missing values), melakukan proses pelabelan (labeling) untuk menyamakan tipe data agar menjadi seragam, dan melakukan pemilihan atribut yang akan digunakan untuk pemodelan.

Setelah tahap pre-processing, dilanjutkan dengan analisis data menggunakan teknik confusion matrix untuk mengetahui hubungan masing-masing atribut dengan atribut target, yaitu curah hujan. Setelah mengetahui hubungan antar atribut, dilakukan proses pemodelan. Sebelum itu, data dibagi menjadi dua bagian, yaitu untuk proses pelatihan (training) dan pengujian (testing). Proses pelatihan menggunakan dua metode, yaitu LSTM dan GRU. Metode LSTM adalah salah satu metode deep learning yang dapat digunakan untuk memprediksi kondisi cuaca karena memiliki banyak parameter yang dapat menyimpan ingatan dalam waktu yang lama secara bersamaan serta memahami korelasi yang kompleks antara data. Sementara itu, GRU memiliki lebih sedikit parameter dibandingkan LSTM, sehingga cocok untuk data yang lebih sedikit agar tidak terjadi overfitting [5]. Kedua metode ini memiliki kelebihan dan kekurangan masing-masing, sehingga kami ingin menggunakan kedua metode ini untuk melakukan pemodelan dengan dataset yang ada. Dalam proses pemodelan dengan algoritma ini, kami menggunakan bahasa pemrograman Python dengan pustaka yang disediakan oleh Tensorflow, yang memudahkan kami dalam melakukan pemodelan. Setelah proses pemodelan, masing-masing hasil dari algoritma tersebut dievaluasi untuk mengetahui performanya. Proses evaluasi menggunakan data pengujian yang telah dibagi sebelumnya, di

mana hasil evaluasi ini akan menghasilkan nilai error rate dan akurasi dari masing-masing model. Selanjutnya, hasil dari masing-masing model LSTM dan GRU akan dibandingkan untuk mengetahui metode mana yang lebih cocok untuk memprediksi curah hujan. Hasil perbandingan ini bertujuan untuk menentukan model terbaik dari algoritma yang digunakan, sehingga cakupan penelitian ini hanya sampai pada tahap mendapatkan model terbaik tersebut.



Gambar 1. Alur Kerja Penelitian.

2.2. Deskripsi Data

Data yang digunakan untuk membuat model prediksi curah hujan merupakan data kondisi cuaca yang diperoleh dari Badan Meteorologi, Klimatologi, dan Geofisika (BMKG). BMKG bertugas mengamati, mengolah, menganalisis, dan menyebarkan informasi mengenai cuaca, iklim, gempa bumi, hingga tsunami. Sebagai lembaga pemerintah non kementerian di Indonesia, BMKG bertanggung jawab langsung kepada Presiden dan telah berdiri sejak tahun 1841. Lembaga ini memiliki beberapa stasiun pengamatan cuaca yang tersebar di berbagai daerah di Indonesia [6][7]. Namun dalam penelitian ini, kami hanya menggunakan data dari Stasiun Geofisika Denpasar yang berada di Kota Denpasar, Provinsi Bali.

Penelitian ini menggunakan data yang dikumpulkan oleh stasiun tersebut dari tanggal 1 Januari 2019 hingga 31 Desember 2019, mencakup periode satu tahun. Pemilihan data pada tahun ini dilakukan karena data tersebut memiliki kualitas yang konsisten dan lengkap sepanjang tahun, sehingga sangat sesuai untuk pengembangan dan pengujian model. Data cuaca tersebut dicatat oleh BMKG setiap harinya, sehingga dalam satu tahun terdapat 365 baris data. Informasi mengenai data ini diperlihatkan pada Tabel 1, yang menunjukkan bahwa data pengukuran BMKG memiliki 10 atribut utama untuk menggambarkan kondisi cuaca. Sepuluh atribut tersebut adalah suhu minimum (°C), suhu maksimum (°C), suhu rata-rata (°C), kelembaban rata-rata (%), curah hujan (mm), lamanya penyinaran matahari (jam), kecepatan angin maksimum (m/s), arah angin saat kecepatan maksimum (°), kecepatan rata-rata angin (m/s), dan arah angin (°). Dari kesembilan atribut tersebut, semuanya memiliki tipe data numerik, kecuali satu atribut yaitu arah angin (°) yang memiliki tipe data karakter, seperti timur (E), utara (N), timur laut (NE), barat laut (NW), selatan (S), tenggara (SE), barat daya (SW), dan barat (W).

Tabel 1. Informasi Dataset Cuaca.

Atribut	Satuan	Deskripsi
tn	°C	Suhu minimum
tx	°C	Suhu maksimum
tavg	°C	Suhu rata-rata
rh_avg	%	Kelembaban rata-rata
rr	mm	Curah hujan
ss	jam	Lamanya penyinaran matahari
ff_x	m/s	Kecepatan angin maksimum
ddd_x	°	Arah angin saat kecepatan maksimum
ff_avg	m/s	Kecepatan angin rata-rata
ddd_car	°	Arah angin terbanyak

2.3. Pra-Pemrosesan Data

Data yang diperoleh dari BMKG tidak dapat langsung digunakan untuk pemodelan, karena jika langsung digunakan, hasil model yang dihasilkan akan kurang baik dan pemodelan bisa menjadi sulit karena ketidaksesuaian bentuk data. Oleh karena itu, pada tahap ini data tersebut akan melalui proses pra-pemrosesan agar lebih siap digunakan untuk analisis dan pembuatan model prediksi. Ada beberapa langkah penting yang perlu dicermati dalam proses ini. Data yang digunakan memiliki 10 atribut dan dalam pengukuran kondisi cuaca, mungkin terdapat data yang hilang atau tidak terukur oleh sensor, atau sensor mengalami kegagalan dalam pembacaan data sehingga menyebabkan *missing value* [8][9].

Untuk menangani hal ini, dalam penelitian ini data yang hilang diproses dengan dua cara yaitu, menggunakan teknik median dan mengisi dengan nilai 0. Kedua metode ini dipilih berdasarkan konteks data yang bersangkutan. Metode median digunakan untuk atribut yang memiliki nilai numerik (atribut *tn* dan *ss*). Metode Median mencari nilai tengah dari dataset yang telah diurutkan, dan kemudian hasil dari perhitungan ini menggantikan nilai yang hilang pada kolom yang sama. Langkah perhitungan Median dimulai dari mengurutkan data secara *ascending* dan selanjutnya mencari nilai tengahnya dengan persamaan seperti yang diperlihatkan pada Rumus 1. Sedangkan untuk atribut seperti curah hujan (*rr*), di mana sensor hujan akan memberikan nilai 0 jika tidak terdeteksi adanya hujan, maka jika ada nilai yang hilang pada atribut tersebut akan diisi dengan nilai 0 yang artinya tidak ada hujan.

$$median = x_{\left(\frac{n+1}{2}\right)} \tag{1}$$

- x*: nilai dalam dataset.
- n*: jumlah total baris dalam dataset.

Pra-pemrosesan berikutnya adalah melakukan pelabelan data. Proses ini dilakukan untuk mengubah data yang bertipe string atau karakter menjadi numerik sehingga algoritma dapat dengan mudah melakukan pemodelan. Dalam penelitian ini, atribut arah angin (*ddd_car*) memiliki nilai bertipe karakter yang diubah menjadi numerik. Adapun nilai-nilai tersebut diubah sebagai berikut, timur (E) menjadi 0, utara (N) menjadi 1, timur laut (NE) menjadi 2, barat laut (NW) menjadi 3, selatan (S) menjadi 4, tenggara (SE) menjadi 5, barat daya (SW) menjadi 6, dan barat (W) menjadi 7.

Berdasarkan proses yang telah dilakukan tersebut seperti untuk mengatasi missing value dan pelabelan pada dataset, diperoleh hasil yang diperlihatkan pada Tabel 2 dan Tabel 3 dalam perbandingan data. Tabel 2 menunjukkan data sebelum pra-pemrosesan, sementara Tabel 3 menunjukkan data setelah proses penanganan missing value dan pelabelan.

Tabel 2. Data Sebelum Pra-Pemrosesan.

no	tn	rr	ss	ff_x	ddd_x	ff_avg	ddd_car
1	25	NaN	NaN	5	220	2	SW
2	NaN	1.2	0.5	6	230	4	S
3	26	0.2	0	5	270	3	W
4	NaN	NaN	NaN	3	90	2	NW
5	25.2	NaN	5	5	100	3	E
..
364	26	18.2	6.3	4	200	2	N
365	25	0	4	3	140	2	SE

Tabel 3. Data Setelah Pra-Pemrosesan.

no	tn	rr	ss	ff_x	ddd_x	ff_avg	ddd_car
1	25	0.0	7.3	5	220	2	6
2	24	1.2	0.5	6	230	4	4
3	26	0.2	0	5	270	3	7
4	24	0.0	0	3	90	2	3
5	25.2	0.0	5	5	100	3	0
..
364	26	18.2	6.3	4	200	2	1
365	25	0	4	3	140	2	5

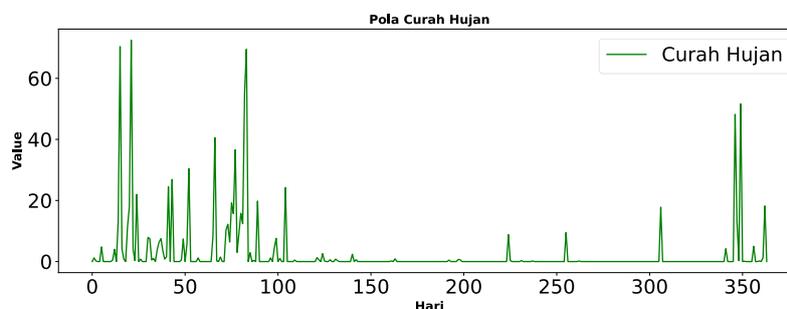
Selanjutnya, dilakukan proses pemilihan atribut yang akan digunakan untuk pemodelan. Pemilihan ini sangat penting karena tidak semua atribut dalam dataset relevan untuk pemodelan. Oleh karena itu, diperlukan pemilihan atribut yang signifikan untuk pemodelan agar dapat membantu proses pemodelan, dan meningkatkan performa model serta juga mengurangi dimensi data. Berdasarkan data yang ada, terdapat 10 atribut, tetapi hanya 7 atribut yang digunakan untuk pemodelan. Adapun 7 atribut tersebut adalah suhu minimum (*tn*), curah hujan (*rr*), lamanya penyinaran matahari (*ss*), kecepatan angin maksimum (*ff_x*), arah angin saat kecepatan maksimum (*ddd_x*), kecepatan rata-rata angin (*ff_avg*), dan arah angin (*ddd_car*). Atribut-atribut ini diproses untuk membuat model prediksi curah hujan (*rr*), sehingga targetnya adalah curah hujan. Pola data dari curah hujan tersebut dapat dilihat pada Gambar 2.

Selain itu, untuk analisis lebih lanjut guna mengetahui relasi antara atribut target dengan atribut lainnya, dilakukan perhitungan untuk mendapatkan nilai matriks korelasi dengan Rumus 2 [10]. Di mana *r* adalah koefisien korelasi, x_i dan y_i adalah nilai individu dari dua variabel *x* dan *y*, dan *n* adalah jumlah observasi. Nilai matriks ini memperlihatkan koefisien korelasi (*r*) antara atribut, di mana koefisien target (*rr*) dengan atribut *tn* memiliki nilai 0.012, dengan atribut *ss* memiliki nilai -0.26, dengan atribut *ff_x* memiliki nilai -0.1, dengan atribut *ddd_x* memiliki nilai 0.33, dengan atribut *ff_avg* memiliki nilai -0.13, dan atribut *ddd_car* memiliki nilai 0.32. Nilai koefisien korelasi ini memberikan informasi mengenai hubungan antara atribut-atribut dalam dataset. Untuk lebih lengkap mengenai nilai koefisien korelasi atribut lainnya, dapat dilihat pada Tabel 4.

$$r = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{\sqrt{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} \sqrt{n \sum_{i=1}^n y_i^2 - (\sum_{i=1}^n y_i)^2}} \tag{2}$$

Tabel 4. Koefisien Korelasi (*r*) Dataset.

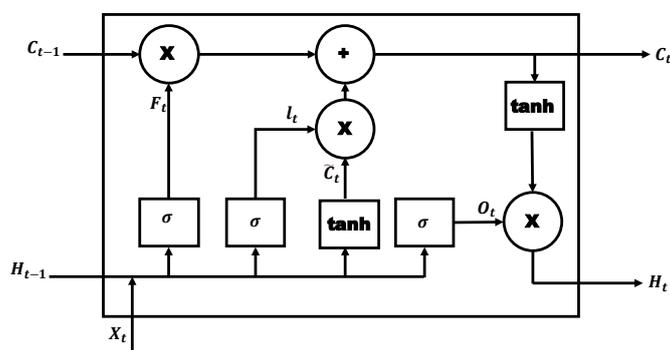
	tn	rr	ss	ff_x	ddd_x	ff_avg	ddd_car
tn	1	0.012	-0.053	-0.061	0.25	-0.19	0.25
rr	0.012	1	-0.26	-0.1	0.33	-0.13	0.32
ss	-0.053	-0.26	1	-0.03	-0.25	0.029	-0.21
ff_x	-0.061	-0.1	-0.03	1	-0.15	0.77	-0.13
ddd_x	0.25	0.33	-0.25	-0.15	1	-0.23	0.72
ff_avg	-0.19	-0.13	0.029	0.77	-0.23	1	-0.27
ddd_car	0.25	0.32	-0.21	-0.13	0.72	-0.27	1



Gambar 2. Pola Data Curah Hujan.

2.4. Long Short-Term Memory (LSTM)

Setelah menyelesaikan proses analisis data pada tahap pra-pemrosesan, langkah berikutnya adalah melakukan pemodelan dengan menggunakan metode Long Short-Term Memory (LSTM). LSTM adalah salah satu metode deep learning yang dikembangkan dari Recurrent Neural Network (RNN) untuk menangani data time series dalam jangka waktu yang panjang serta data yang memiliki korelasi kompleks antara atribut. Dalam penelitian ini, prediksi curah hujan dibuat menggunakan LSTM dengan beberapa lapisan (layer) yang karakteristiknya ditunjukkan pada Tabel 5. Metode LSTM bekerja secara umum dengan tiga gerbang utama, yaitu input gate, forget gate, dan output gate, yang strukturnya dapat dilihat pada Gambar 3. Selain itu, persamaan dari struktur tersebut dijelaskan pada Rumus 3-10 [11][12]. Di mana, F_t merepresentasikan *forget gate*, H_{t-1} adalah *hidden state* sebelumnya, sedangkan X_t adalah *input* saat ini, W adalah bobot, dan bias dilambangkan dengan b . I_t adalah *input gate*, \tilde{C}_t adalah nilai kandidat, dan *cell state* dilambangkan dengan C_t . Terakhir, O_t merepresentasikan *output gate*. Namun, dalam penelitian ini, untuk mempermudah pemrograman pemodelan, kami memanfaatkan pustaka TensorFlow dalam bahasa pemrograman Python, yang memudahkan dalam implementasi dan penyesuaian parameter model.



Gambar 3. Struktur LSTM.

$$F_t = \text{sigmoid}([H_{t-1}, X_t] \cdot W_f + b_f) \tag{3}$$

$$I_t = \text{sigmoid}([H_{t-1}, X_t] \cdot W_i + b_i) \tag{4}$$

$$\tilde{C}_t = \text{tanh}([H_{t-1}, X_t] \cdot W_c + b_c) \tag{5}$$

$$C_t = (F_t \cdot C_{t-1}) + (I_t \cdot \tilde{C}_t) \tag{6}$$

$$O_t = \text{sigmoid}(W_o \cdot [H_{t-1}, X_t] + b_o) \tag{7}$$

$$H_t = O_t \cdot \text{tanh}(C_t) \tag{8}$$

$$\text{sigmoid}(x) = \frac{1}{1 + e^{-x}} \tag{9}$$

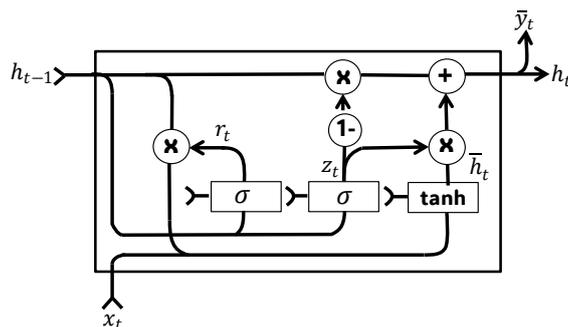
$$\text{tanh}(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{10}$$

Tabel 5. Karakteristik Pemodelan dengan LSTM.

Layer	Propertis
1 st LSTM	Unit = 50, Activation = ReLu
2 nd LSTM	Unit = 10, Activation = ReLu
3 rd LSTM	Unit = 5, Activation = ReLu
1 st Dense	Unit = 5, Activation = ReLu
2 nd Dense	Unit = 1

2.5. Gated recurrent units (GRU)

Metode Gated Recurrent Units (GRU) juga digunakan dalam pemodelan untuk memprediksi curah hujan pada penelitian ini. Metode ini memiliki kemiripan dengan LSTM, namun dengan struktur yang lebih sederhana. GRU adalah varian dari RNN yang dikembangkan untuk mengatasi masalah vanishing gradient dalam data time series. Dibandingkan dengan LSTM, GRU lebih sederhana dan efisien secara komputasi, karena hanya memiliki dua gerbang utama yaitu, reset gate (r_t) dan update gate (z_t). Dari kedua gerbang ini, dilakukan perhitungan untuk menghasilkan hidden gate (\bar{h}_t). Untuk lebih memahami gerbang-gerbang ini, lihat Gambar 5, dan persamaan metode ini dapat dilihat pada Rumus 11-13 serta karakteristik parameter setiap layer yang digunakan dalam pembuatan model dapat dilihat pada Tabel 6. Di mana, W adalah bobot, U adalah bobot untuk hidden state, X_t adalah input saat ini, dan b adalah bias. Berdasarkan hal tersebut, penelitian ini akan membandingkan hasil model LSTM dan GRU menggunakan data kondisi cuaca [13][14].



Gambar 4. Struktur GRU.

$$z_t = \sigma_g(W_z x_t + U_z h_{t-1} + b_z) \tag{11}$$

$$r_t = \sigma_g(W_r x_t + U_r h_{t-1} + b_r) \tag{12}$$

$$h_t = (1 - z_t) h_{t-1} + z_t \phi_h(W_h x_t + U_h(r_t h_{t-1}) + b_h) \tag{13}$$

Tabel 6. Karakteristik Pemodelan dengan GRU.

Layer	Propertis
1 st GRU	Unit = 95, Activation = ReLu
2 nd GRU	Unit = 15, Activation = ReLu
3 rd GRU	Unit = 5, Activation = ReLu
1 st Dense	Unit = 10, Activation = ReLu
2 nd Dense	Unit = 1

2.6. Evaluasi Model

Pada penelitian ini, digunakan metode LSTM dan GRU untuk membuat model prediksi curah hujan. Setiap model yang dihasilkan dievaluasi menggunakan tiga teknik evaluasi yaitu *root mean squared error* (RMSE), *mean absolute error* (MAE), dan *mean squared error* (MSE). Tiga parameter evaluasi ini digunakan untuk menilai performa model. Hasil dari masing-masing evaluasi pada model dengan LSTM dan GRU dibandingkan untuk mendapatkan model terbaik sesuai dengan karakteristik dan parameter dari metode yang digunakan. Setiap teknik evaluasi ini memiliki kelebihan masing-masing. Persamaan untuk RMSE dapat dilihat pada Rumus 14, MAE pada Rumus 15, dan MSE pada Rumus 16 [15][16]. Dimana, n adalah jumlah observasi, y_i adalah nilai sebenarnya dan \tilde{y}_i merupakan nilai prediksi.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \tilde{y}_i)^2}{n}} \tag{14}$$

$$MAE = \frac{\sum_{i=1}^n |y_i - \tilde{y}_i|}{n} \tag{15}$$

$$MSE = \frac{\sum_{i=1}^n |y(i) - \hat{y}_i|^2}{n} \tag{16}$$

3. Hasil dan Pembahasan

Hasil pengujian model prediksi curah hujan dalam penelitian ini terbagi menjadi dua bagian yaitu hasil pengujian dengan metode LSTM dan hasil pengujian dengan metode GRU. Proses pemodelan didasarkan pada perancangan karakteristik dan parameter yang telah ditentukan sebelumnya. Parameter yang diatur untuk mendapatkan model meliputi jumlah neuron (unit), fungsi aktivasi (ReLU), dan nilai *lookback*. Dalam penelitian ini, nilai *lookback* ditetapkan sebesar 3 untuk setiap model. Proses pemodelan mengatur jumlah neuron pada setiap layernya, dengan 3 layer yang digunakan, di mana jumlah neuron di layer pertama diatur, sedangkan di layer kedua dan ketiga tetap konstan. Selain itu, digunakan learning rate sebesar 0,001 dengan optimalisasi Adam, dan nilai epoch sebesar 500.

Berdasarkan hasil pengujian, model dengan metode LSTM memiliki nilai RMSE sebesar 8,853, MAE sebesar 4,090, dan MSE sebesar 78,383. Sedangkan model dengan metode GRU memiliki nilai RMSE sebesar 9,698, MAE sebesar 4,291, dan MSE sebesar 94,058. Dari hasil ini, dapat disimpulkan bahwa metode LSTM memiliki performa yang lebih baik dibandingkan metode GRU. Selain itu, ditemukan bahwa jumlah neuron yang kecil belum tentu menghasilkan *error rate* yang kecil begitu juga sebaliknya yang artinya jumlah neuron yang besar atau kecil mempengaruhi kualitas model. Selanjutnya, untuk hasil percobaan masing-masing metode dapat dilihat pada Tabel 7 untuk metode LSTM dan Tabel 8 untuk metode GRU. Perbandingan prediksi antara nilai aktual dan prediksi untuk model dengan LSTM dan GRU dapat dilihat pada Gambar 5.

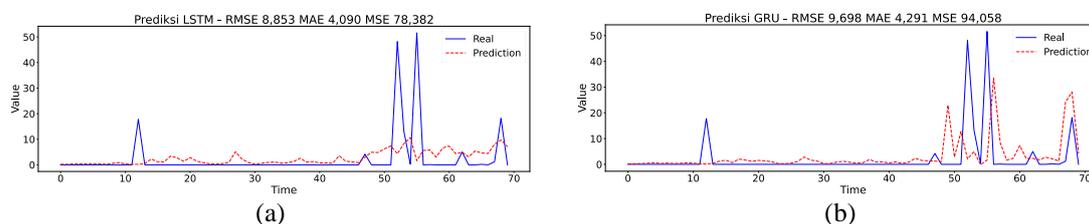
Tabel 7. Hasil Pemilihan Jumlah Neuron pada LSTM.

Lookback	1 st LSTM Neuron	2 nd LSTM Neuron	3 rd LSTM Neuron	1 st Dense	2 nd Dense	RMSE (mm)	MAE (mm)	MSE (mm)
3	5	10	5	5	1	11,669	4,647	136,156
3	10	10	5	5	1	12,025	5,318	144,602
3	15	10	5	5	1	12,211	4,578	149,115
3	20	10	5	5	1	21,835	9,975	476,777
3	25	10	5	5	1	18,245	8,069	332,881
3	30	10	5	5	1	12,325	5,650	151,904
3	35	10	5	5	1	12,728	4,811	162,013
3	40	10	5	5	1	14,715	6,500	216,537
3	45	10	5	5	1	11,265	4,419	126,903
3	50	10	5	5	1	8,853	4,090	78,382
3	55	10	5	5	1	13,180	6,173	173,711
3	60	10	5	5	1	11,911	6,470	141,884
3	65	10	5	5	1	12,510	5,216	156,506
3	70	10	5	5	1	10,327	4,305	106,639
3	75	10	5	5	1	11,847	5,506	140,354
3	80	10	5	5	1	10,268	4,624	105,422
3	85	10	5	5	1	13,236	5,571	175,204
3	90	10	5	5	1	10,177	4,875	103,562
3	95	10	5	5	1	11,620	5,145	135,027
3	100	10	5	5	1	15,200	6,562	233,473

Tabel 8. Hasil Pemilihan Jumlah Neuron pada GRU.

Lookback	1 st GRU Neuron	2 nd GRU Neuron	3 rd GRU Neuron	1 st Dense	2 nd Dense	RMSE (mm)	MAE (mm)	MSE (mm)
3	5	15	5	10	1	14,184	5,831	201,197
3	10	15	5	10	1	9,850	4,660	97,017
3	15	15	5	10	1	12,461	4,851	155,271
3	20	15	5	10	1	10,297	5,217	106,033
3	25	15	5	10	1	13,779	5,491	189,857
3	30	15	5	10	1	14,513	5,456	210,638
3	35	15	5	10	1	12,147	5,969	147,560
3	40	15	5	10	1	12,129	4,931	147,124
3	45	15	5	10	1	15,366	6,403	236,103
3	50	15	5	10	1	14,774	7,653	218,265
3	55	15	5	10	1	11,038	4,599	121,831
3	60	15	5	10	1	11,435	4,532	130,758
3	65	15	5	10	1	12,383	5,395	153,341

3	70	15	5	10	1	11,977	5,452	143,443
3	75	15	5	10	1	14,452	5,465	208,861
3	80	15	5	10	1	17,703	7,989	313,385
3	85	15	5	10	1	9,863	5,089	97,278
3	90	15	5	10	1	18,986	9,040	360,461
3	95	15	5	10	1	9,698	4,291	94,058
3	100	15	5	10	1	11,360	5,165	129,052



Gambar 5. Perbandingan Data Aktual dan Prediksi (a) pemodelan dengan LSTM, (b) pemodelan dengan GRU

4. Kesimpulan

Penelitian ini bertujuan untuk membuat model prediksi curah hujan dengan menggunakan metode LSTM dan GRU. Data yang digunakan untuk pembuatan model diperoleh dari Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) yang diukur selama satu tahun pada tahun 2019. Dataset mentah yang diperoleh memiliki 10 atribut. Data ini kemudian diproses dan dianalisis untuk menangani data yang hilang dan melakukan pelabelan. Dari 10 atribut yang ada, hanya 7 atribut yang digunakan, yaitu suhu minimum, curah hujan, lamanya penyinaran matahari, kecepatan angin maksimum, arah angin saat kecepatan maksimum, kecepatan rata-rata angin, dan arah angin. Berdasarkan percobaan pemodelan, hasil model dengan metode LSTM memiliki nilai RMSE sebesar 8,853, MAE sebesar 4,090, dan MSE sebesar 78,383. Sementara itu, model dengan metode GRU memiliki nilai RMSE sebesar 9,698, MAE sebesar 4,291, dan MSE sebesar 94,058. Dari hasil tersebut, dapat disimpulkan bahwa metode LSTM memiliki performa yang lebih baik dibandingkan dengan GRU dalam hal prediksi curah hujan.

Daftar Pustaka

- [1] I. Salehin, I. M. Talha, M. Mehedi Hasan, S. T. Dip, M. Saifuzzaman, and N. N. Moon, "An Artificial Intelligence Based Rainfall Prediction Using LSTM and Neural Network," in *2020 IEEE International Women in Engineering (WIE) Conference on Electrical and Computer Engineering (WIECON-ECE)*, 2020, pp. 5–8. doi: 10.1109/WIECON-ECE52138.2020.9398022.
- [2] F. Manokij, K. Sarinapakorn, and P. Vateekul, "Forecasting Thailand's Precipitation with Cascading Model of CNN and GRU," in *2019 11th International Conference on Information Technology and Electrical Engineering (ICITEE)*, 2019, pp. 1–6. doi: 10.1109/ICITEED.2019.8929975.
- [3] W. Liu, G. Zhuang, X. Liu, S. Hu, R. He, and Y. Wang, "How do we move towards true artificial intelligence," in *2021 IEEE 23rd Int Conf on High Performance Computing & Communications; 7th Int Conf on Data Science & Systems; 19th Int Conf on Smart City; 7th Int Conf on Dependability in Sensor, Cloud & Big Data Systems & Application (HPCC/DSS/SmartCity/DependSys)*, 2021, pp. 2156–2158. doi: 10.1109/HPCC-DSS-SmartCity-DependSys53884.2021.00321.
- [4] V. B. Kumar, V. M. Nookesh, B. S. Saketh, S. Syama, and J. Ramprabhakar, "Wind Speed Prediction Using Deep Learning-LSTM and GRU," in *2021 2nd International Conference on Smart Electronics and Communication (ICOSEC)*, 2021, pp. 602–607. doi: 10.1109/ICOSEC51865.2021.9591886.
- [5] K. P. Rasheed Abdul Haq and V. P. Harigovindan, "Water Quality Prediction for Smart Aquaculture Using Hybrid Deep Learning Models," *IEEE Access*, vol. 10, pp. 60078–60098, 2022, doi: 10.1109/ACCESS.2022.3180482.
- [6] K. Purwandari, J. W. C. Sigalingging, M. Putri, and M. A. Kusuma, "The Indonesian Surface Weather Observation Monitoring System for Data Availability," in *2022 International Conference on Information Management and Technology (ICIMTech)*, 2022, pp. 207–212. doi: 10.1109/ICIMTech55957.2022.9915178.
- [7] BMKG, "Badan Meteorologi, Klimatologi, dan Geofisika." <https://www.bmkg.go.id/>
- [8] A. K. Tripathi, H. Saini, and G. Rathee, "Missing Values Imputation in Food Consumption: An Analytical Study," in *2021 6th International Conference on Signal Processing, Computing and Control (ISPCC)*, 2021, pp. 303–307. doi: 10.1109/ISPCC53510.2021.9609371.

-
- [9] S. Z. Christopher, T. Siswantining, D. Sarwinda, and A. Bustaman, "Missing Value Analysis of Numerical Data using Fractional Hot Deck Imputation," in *2019 3rd International Conference on Informatics and Computational Sciences (ICICoS)*, 2019, pp. 1–6. doi: 10.1109/ICICoS48119.2019.8982412.
- [10] I. N. K. Wardana, J. W. Gardner, and S. A. Fahmy, "Optimising Deep Learning at the Edge for Accurate Hourly Air Quality Prediction," *Sensors*, vol. 21, no. 4, 2021, doi: 10.3390/s21041064.
- [11] I. W. A. Suranata, I. N. K. Wardana, N. Jawas, and I. K. A. A. Aryanto, "Feature engineering and long short-term memory for energy use of appliances prediction," *Telkomnika (Telecommunication Comput. Electron. Control.)*, vol. 19, no. 3, pp. 920–930, 2021, doi: 10.12928/TELKOMNIKA.v19i3.17882.
- [12] I. Nyoman, K. Wardana, N. Jawas, I. Komang, and A. A. Aryanto, "Prediksi Penggunaan Energi Listrik pada Rumah Hunian Menggunakan Long Short-Term Memory," *TIERS Inf. Technol. J.*, vol. 1, no. 1, pp. 1–11, 2020, [Online]. Available: <http://journal.undiknas.ac.id/index.php/tiers>
- [13] J. R. Ban, Q. Gou, and Y. S. Li, "Study on Rainfall Prediction of Yibin City Based on GRU and XGBoost," in *2022 4th International Conference on Advances in Computer Technology, Information Science and Communications (CTISC)*, 2022, pp. 1–5. doi: 10.1109/CTISC54888.2022.9849730.
- [14] H. F. Nurrohman, D. C. R. Novitasari, F. Setiawan, Rochimah, A. Taufiq, and A. Hamid, "Rainfall Prediction Using Gated Recurrent Unit Based on DMI and Nino3.4 Index," in *2022 IEEE International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology (IAICT)*, 2022, pp. 191–196. doi: 10.1109/IAICT55358.2022.9887474.
- [15] T. O. Hodson, "Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not," *Geosci. Model Dev.*, vol. 15, no. 14, pp. 5481–5487, 2022, doi: 10.5194/gmd-15-5481-2022.
- [16] A. Lawal, S. Rehman, L. M. Alhems, and M. M. Alam, "Wind Speed Prediction Using Hybrid 1D CNN and BLSTM Network," *IEEE Access*, vol. 9, pp. 156672–156679, 2021, doi: 10.1109/ACCESS.2021.3129883.