

# Penerapan Metode Principal Component Analysis (PCA) dan Long Short-Term Memory (LSTM) dalam Memprediksi Prediksi Curah Hujan Harian

Musfiroh<sup>1</sup>, Dian Candra Rini Novitasari<sup>1,\*</sup>, Putroue Keumala Intan<sup>1</sup>, Gede Gangga Wisnawa<sup>2</sup>

<sup>1</sup>Sains dan Teknologi, Jurusan Matematika, UIN Sunan Ampel Surabaya, Surabaya, Indonesia

<sup>2</sup>Peramalan dan Analisis, Badan Meteorologi Klimatologi dan Geofisika, Surabaya, Indonesia

Email: <sup>1</sup>musfiroh.rohmati22@gmail.com, <sup>2</sup>diancrini@uinsby.ac.id, <sup>3</sup>putroue@uinsby.ac.id, <sup>4</sup>gangga.wisnawa@gmail.com

Email Penulis Korespondensi: diancrini@uinsby.ac.id

Submitted: 07/02/2023; Accepted: 18/05/2023; Published: 29/06/2023

**Abstrak**—Sejak tiga tahun terakhir Luwu Utara sering terjadi bencana hidrologi berupa banjir dan tanah longsor. Bencana tersebut berdampak buruk terhadap ketersediaan air bersih, gagal tanam dan bahkan cenderung menurunkan kualitas panen. Kakao merupakan salah satu komoditas unggulan Kabupaten Luwu Utara yang produktivitasnya menurun akibat dampak dari perubahan iklim sehingga akan berpengaruh pada keberlangsungan pendapatan penduduk setempat. Maka dari itu, tujuan penelitian ini adalah mengantisipasi curah hujan yang akan terjadi untuk mencegah atau mengurangi risiko kegagalan dan kerugian. Metode Principal Component Analysis (PCA) digunakan sebagai ekstraksi fitur untuk mengetahui variabel yang paling berpengaruh dan metode Long Short-Term Memory (LSTM) digunakan sebagai metode prediksi. Curah hujan yang akan datang diprediksi menggunakan variabel meteorologi seperti variabel tekanan, evaporasi, suhu maksimal, kelembaban rata-rata, dan lama penyinaran matahari dari tanggal 1 Januari 2017 sampai 30 September 2022. Berdasarkan hasil PCA diperoleh 4 variabel yang paling berpengaruh terhadap curah hujan yaitu: variabel evaporasi, suhu maksimal, kelembaban rata-rata, dan lama penyinaran matahari. Variabel tersebut digunakan sebagai input untuk memprediksi curah hujan menggunakan LSTM. Pada penelitian ini menggunakan parameter uji coba yaitu jumlah hidden, batch size, dan learn rate drop period. Hasil prediksi terbaik diperoleh MAPE sebesar 0,0018 dengan jumlah hidden, batch size dan learn rate drop period masing-masing 100, 32, dan 50. Hasil prediksi menunjukkan curah hujan yang sangat lebat terjadi pada tanggal 28 Agustus 2021 sebesar 101,9734 mm, 21 September 2021 sebesar 108,6528 mm, dan 5 April 2022 sebesar 116,5510 mm. Pada penelitian ini PCA mampu meningkatkan akurasi dalam mempertimbangkan semua parameter dan memilih yang efektif.

**Kata Kunci:** Curah Hujan; Ekstraksi Fitur; Hidrologi; LSTM; Luwu Utara; PCA

**Abstract**—Since the last three years North Luwu has experienced frequent hydrological disasters in the form of floods and landslides. The disaster had a negative impact on the availability of clean water, failed to plant and even tended to reduce the quality of the harvest. Cocoa is one of the leading commodities of North Luwu Regency whose productivity has decreased due to the impact of climate change so that it will affect the sustainability of the local population's income. Therefore, the purpose of this research is to anticipate rainfall that will occur to prevent or reduce the risk of failure and loss. Principal Component Analysis (PCA) Method is used as feature extraction to find out the most influential variables and the Long Short-Term Memory (LSTM) method is used as a prediction method. Future rainfall is predicted using meteorological variables such as pressure, evaporation, maximum temperature, average humidity, and sunshine duration from 1 January 2017 to 30 September 2022. Based on the PCA results, 4 variables are obtained that have the most influence on rainfall, namely: variable evaporation, maximum temperature, average humidity, and length of sunlight. These variables are used as input to predict rainfall using LSTM. In this study using trial parameters, namely the number of hidden, batch size, and learn rate drop period. The best prediction results were obtained for MAPE of 0.0018 with the number of hidden, batch size and learn rate drop periods of 100, 32, and 50 respectively. The prediction results show very heavy rainfall occurring on August 28, 2021 of 101.9734 mm, 21 September 2021 of 108.6528 mm, and 5 April 2022 of 116.5510 mm. In this study PCA was able to increase accuracy in considering all parameters and choosing the most effective.

**Keywords:** Rainfall; Feature Extraction; Hydrology; LSTM; North Luwu; PCA

## 1. PENDAHULUAN

Sulawesi Selatan merupakan salah satu provinsi penyumbang kakao terbesar sehingga menjadikan Indonesia berada di peringkat ketiga dunia sebagai negara penghasil kakao [1]. Luwu Utara merupakan daerah di Sulawesi Selatan yang mayoritas penduduknya bekerja di bidang pertanian, kehutanan, perburuan, dan perikanan dengan komoditas unggulan kakao [2], [3]. Namun, beberapa tahun terakhir akibat dampak perubahan iklim menunjukkan produktivitas tanaman kakao menurun [1], [4]. Tipe hujan ekuatorial menjadi penyebab Luwu Utara rawan terjadi bencana hidrometeorologi dan perubahan iklim [5], [6].

Beberapa tahun terakhir frekuensi kejadian banjir di Luwu Utara meningkat. Pada tahun 2020 Luwu Utara mengalami banjir bandang dan tanah longsor terparah di Indonesia [7], [8]. Banjir tersebut berlangsung selama beberapa hari merendamkan sektor penting dan lahan pertanian hingga mengakibatkan korban jiwa [9]. Pada bulan September 2021 akibat hujan dengan intensitas tinggi menyebabkan 100 ha kebun, 2 ha sawah, dan 538 ha tambak tergenang banjir [10]. Pada Bulan Juli 2022 hujan dengan intensitas yang tinggi kembali terjadi sehingga membanjiri 76 rumah warga dan 87 ha persawahan [11]. Bencana tersebut berdampak buruk pada ketersediaan air bersih, gagal tanam dan bahkan cenderung menurunkan kualitas panen sehingga akan berpengaruh pada keberlangsungan pendapatan penduduk setempat [12].

Berdasarkan bencana yang pernah terjadi di wilayah Luwu Utara, prediksi curah hujan yang akurat sangat diperlukan sebagai informasi penting untuk mengetahui curah hujan yang akan terjadi di masa yang akan datang. Informasi tersebut akan berguna untuk mendukung kegiatan di berbagai sektor terutama sektor pertanian [13]. Tujuan dari penelitian ini diharapkan masyarakat melakukan upaya antisipasi curah hujan yang akan terjadi untuk mencegah atau mengurangi risiko kegagalan dan kerugian dalam proses produksi [14].

Besarnya curah hujan yang akan terjadi dapat diprediksi dengan menggunakan data historis curah hujan [15]. Beberapa algoritma machine learning dalam memprediksi curah hujan telah banyak dilakukan pada penelitian sebelumnya. Penelitian yang dilakukan oleh Yue-Chao Chen, dkk. menggunakan model *Integrated Flood Analysis System* (IFAS) untuk simulasi limpasan dan model *Long Short-Term Memory* (LSTM) untuk prakiraan curah hujan per jam. Penelitian tersebut memperoleh model IFAS yang akurat dengan menggunakan simulasi limpasan di DAS Tokachi menghasilkan NSE 0,75 dan model LSTM menghasilkan NSE 0,86 untuk prakiraan curah hujan per jam. Model IFAS menghasilkan nilai NSE 0,81 dengan menggunakan data curah hujan optimal yang diramalkan oleh model LSTM [16]. Chengcheng Chen, dkk. meramalkan curah hujan bulanan mendapatkan hasil LSTM mengungguli model *Random Forest* (RF) dengan peningkatan RMSE sebesar 12,20–14,90%, RSR sebesar 12,30–14,80%, R sebesar 9,40–13,50% dan NSE sebesar 32,90–33,20% [17]. Zeyi Chao, dkk. memprediksi curah hujan yang lebih detail, tepat waktu, dan akurat dapat dicapai dengan menggunakan LSTM pada sensor cuaca MEMS. Dalam penelitian tersebut metode LSTM unggul dibandingkan dengan metode *Support Vector Machine* (SVM), RF, *Back Propagation Neural Networks* (BPNNs), dan *Autoregressive And Moving Average* (ARMA) [18].

Curah hujan dipengaruhi oleh beberapa variabel meteorologi seperti yang dilakukan oleh Poornima, dkk. menggunakan parameter suhu maksimum, suhu minimum, kelembaban relatif maksimum, kelembaban relatif minimum, kecepatan angin, sinar matahari, dan evapotranspirasi. Penelitian tersebut memprediksi curah hujan dengan membandingkan model *Holt-Winters*, *Autoregressive Integrated Moving Average* (ARIMA), *Extreme Learning Machine* (ELM), *Recurrent Neural Network* (RNN), LSTM, dan *Intensified Long Short-Term Memory* (*Intensified LSTM*) diperoleh model *Intensified LSTM* lebih baik dengan RMSE 0,33 dibandingkan LSTM dengan RMSE 0,35 [19]. Peramalan variabel cuaca di bandara Hang Nadim Indonesia menggunakan visibilitas sebagai data prediksi sedangkan data suhu, tekanan, kelembaban, dan titik embun sebagai perantara. Hasil model terbaik adalah LSTM *multiple layer* dan data perantara terbaik adalah variabel tekanan dengan akurasi validasi 0,81 dan RMSE 0,08 [20]. Jinle Kang menggunakan model LSTM dengan variabel input ketinggian awan rendah, kecenderungan tekanan, suhu, tekanan atmosfer, dan kelembaban relatif. Hasil eksperimen menunjukkan bahwa LSTM bagus dalam memprediksi curah hujan dengan nilai RMSE pada dataset training, validation, dan testing masing-masing sebesar 42,28 mm, 42,03 mm, dan 41,72 mm [21].

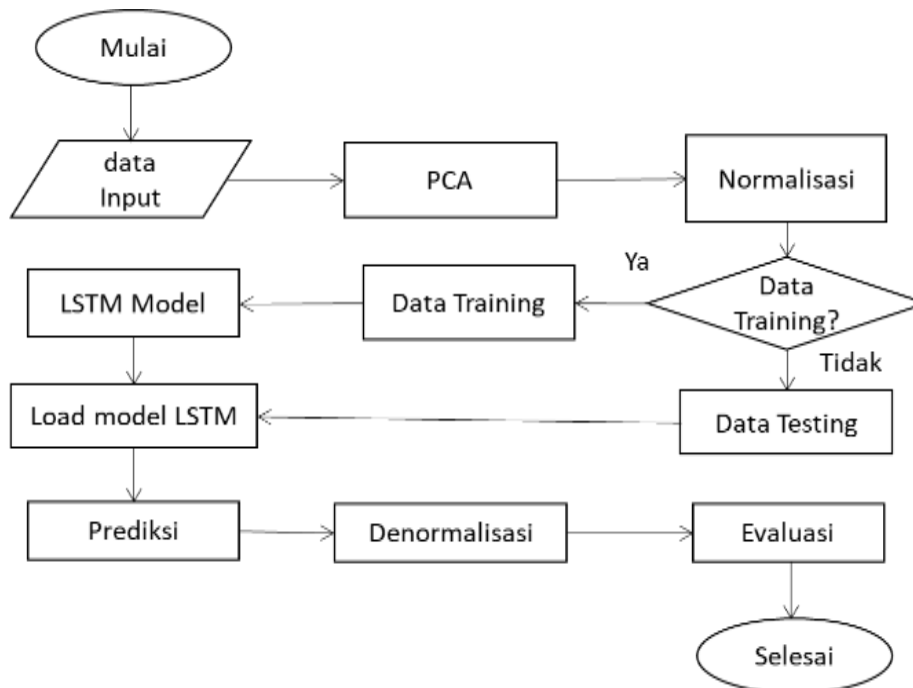
LSTM merupakan pengembangan dari RNN yang memiliki kemampuan bagus dalam menangani interaksi nonlinear yang tinggi dan dapat mengatasi masalah gradien terutama pada kasus prediksi curah hujan [22], [23]. Namun, faktor yang mempengaruhi curah hujan sangat kompleks sehingga reduksi dimensi sangat diperlukan agar model bekerja secara optimal [22]. Reduksi dimensi terdiri dari dua jenis yaitu *feature extraction* dan *feature selection*. *Principal Component Analysis* (PCA) merupakan metode *feature extraction* yang digunakan untuk mereduksi dengan mengkompres atribut yang jumlahnya banyak menjadi lebih sedikit [24]. Bashir, dkk menggunakan PCA untuk menentukan variabel yang paling berpengaruh terhadap kerentanan Akuifer Varamin di Iran. Berdasarkan hasil PCA diperoleh 7 variabel yang paling berpengaruh dari 14 variabel yang ada [25]. Tripathi, dkk. Menggunakan PCA untuk memilih parameter kualitas air Sungai Gangga di India mendapatkan 9 parameter terbaik dari 28 parameter [26]. Penelitian Yulian Wen, dkk. memprediksi data time series harga saham menggunakan model PCA-LSTM, CNN Model, MPL Model, dan *Moving average model* diperoleh model PCA-LSTM memprediksi secara akurat dengan RMSE dan MAPE sebesar 0,22 dan 1,67% [27]. Dawei Geng, dkk. memprediksi *time series* kecepatan angin dengan membandingkan metode PCA-BPNN, PCA-GPR, PCA-LSTM, PCA-RNN, dan PCA-SVR diperoleh model PCA-LSTM memiliki RMSE, MAPE, dan MAE paling kecil, dengan R<sup>2</sup> sebesar 0,9989 [28].

LSTM dengan kemampuannya yang sangat cocok pada data *time series* dan PCA dapat mempercepat pelatihan neural network [29]. Berdasarkan penelitian sebelumnya PCA-LSTM menghasilkan error lebih kecil dibandingkan metode lainnya, maka penelitian ini mengusulkan model deep learning berbasis PCA-LSTM untuk memprediksi curah hujan dengan menggunakan variabel metodologi yaitu: tekanan, evaporasi, suhu maksimal, kelembaban rata-rata, dan lama penyinaran matahari untuk memprediksi curah hujan di daerah Luwu Utara. Variabel tersebut akan dilakukan ekstraksi fitur menggunakan PCA. Hasil PCA akan dijadikan input dalam memprediksi curah hujan harian menggunakan metode LSTM. Hasil prediksi akan dievaluasi menggunakan MAPE.

## 2. METODOLOGI PENELITIAN

### 2.1 Tahapan Penelitian

Langkah-langkah pada penelitian ini adalah preprocessing, ekstraksi fitur, analisis, dan evaluasi yang ditunjukkan pada Gambar 1.



**Gambar 1.** Flowchart Prediksi Curah Hujan Menggunakan PCA-LSTM

Kerangka kerja pada penelitian ini adalah pertama mengumpulkan data. kemudian dilakukan preprocessing dengan mengisi data kosong menggunakan interpolasi. Setelah itu, menerapkan ekstraksi fitur menggunakan metode PCA yang digunakan untuk mencari variabel yang paling berpengaruh terhadap curah hujan. Kemudian dilakukan normalisasi data dan membentuk pola *time series*. Tahap selanjutnya membagi data menjadi data training dan data testing. Data training digunakan untuk mendapatkan model LSTM yang kemudian model tersebut dilakukan prediksi menggunakan data testing. Hasil prediksi akan dikembalikan ke nilai sebenarnya dengan denormalisasi. Model akan dievaluasi menggunakan MAPE.

**2.2 Sumber Data**

Data meteorologi dikumpulkan dari website Stasiun Meteorologi Andi Jemma di Luwu Utara dan sebagian diperoleh dari BMKG perak II dengan menggunakan 2099 data harian berkisar dari 1 Januari 2017 sampai 30 September 2022 (Gambar 2) [30]. Data yang digunakan adalah variabel tekanan (P), evaporasi (E), suhu maksimal (T), kelembaban rata-rata (RH), dan lama penyinaran matahari (LS) sebagai variabel input sedangkan data curah hujan (CH) sebagai variabel target. Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) mengkategorikan intensitas curah hujan berdasarkan 4 kriteria seperti pada Tabel 2 [31]. Dataset penelitian ini ditunjukkan pada Tabel 1.



**Gambar 2.** Peta Lokasi Penelitian

**Tabel 1.** Dataset penelitian

Date	P (hPa)	E (mm)	T (°C)	RH (°C)	LS (jam)	CH (mm)
01-01-2017	1008,76	6,00			5,30	0,00
02-01-2017	1009,02	6,00	32,40	80,00	6,60	0,00
03-01-2017	1008,02	5,00	33,80	78,00	3,00	0,00

04-01-2017	1008,45	4,00	33,20	82,00	7,20	8888
...	...	...	...	...	...	...
27-09-2022	1008,09	7,00	34,60	69,00	10,90	0,00
28-09-2022	1007,61	12,10	33,80	73,00	8,00	0,00
29-09-2022	1008,89	5,60	33,80	79,00	8,60	0,40
30-09-2022	1009,30	5,30	34,60	74,00	7,20	5,20

Tabel 2. Kategori curah hujan

Kategori	Intensitas (mm/hari)
Hujan Ringan	5 – 20
Hujan Sedang	20 – 50
Hujan Lebat	50 – 100
Hujan Sangat Lebat	> 100

### 2.3 Normalisasi Min-Max Scaler

Dataset yang memiliki rentang data yang tinggi akan mempengaruhi hasil prediksi maka perlu mengubah data pada rentang yang sama dengan cara normalisasi [32], [33]. Normalisasi bertujuan mengubah estimasi segmen numerik dalam dataset ke skala reguler, tanpa mendistorsi perbedaan dalam rentang nilai [34]. Salah satu metode normalisasi adalah menggunakan *min-max scaler* dimana metode tersebut mentransformasikan secara linier dari data asli ke dalam skala 0 sampai 1 [35]. Rumus *min-max scaler* dihitung dengan Persamaan (1) [36].

$$\tilde{x} = \frac{x_i - x_{min}}{x_{max} - x_{min}} \quad (1)$$

Dimana  $\tilde{x}$  merupakan data hasil normalisasi,  $x_i$  merupakan data asli,  $x_{min}$  merupakan nilai minimum dari data asli, dan  $x_{max}$  merupakan nilai maksimum dari data asli.

### 2.4 PCA Algorithm

*Principal Component Analysis (PCA)* merupakan teknik statistik *multivariate* yang sangat kuat dalam menganalisis *dataset* dimensi tinggi. PCA berguna untuk mengekstraksi fitur untuk memperoleh variabel yang paling berpengaruh sehingga mempercepat pelatihan *neural network* [25], [29], [37]. PCA juga berguna untuk mengetahui kontribusi individu dari variabel independen [38]. Beberapa cara analisa pada PCA dalam menentukan komponen utama dengan melihat variansi totalnya, nilai *eigen* lebih dari satu, atau dengan melihat *scree plot* [39].

*Dataset* berukuran  $n \times m$  dimana  $n$  mewakili observasi  $x_i$  untuk  $i \in \{1,2,3, \dots, n\}$  dan  $m$  adalah jumlah variabelnya. Proses pertama perhitungan algoritma PCA yaitu menghitung matrik *kovarians* untuk seluruh *dataset* menggunakan Persamaan (2). Nilai *eigen* dan *eigenvector*  $x$  dari matrik kovarians dihitung dengan Persamaan (3) dan (4). Kemudian menghitung tingkat kontribusi varians dan tingkat kontribusi varians kumulatif dari masing-masing komponen utama menurut nilai *eigen* dengan Persamaan (5) dan Persamaan (6) [40].

$$Cov(x_j, x_k) = \frac{\sum_{i=1}^n (x_{ij} - \bar{x}_j)(x_{ik} - \bar{x}_k)}{n - 1} \quad (2)$$

$$\det(C - \lambda I) = 0 \quad (3)$$

$$(C - \lambda I)x = 0 \quad (4)$$

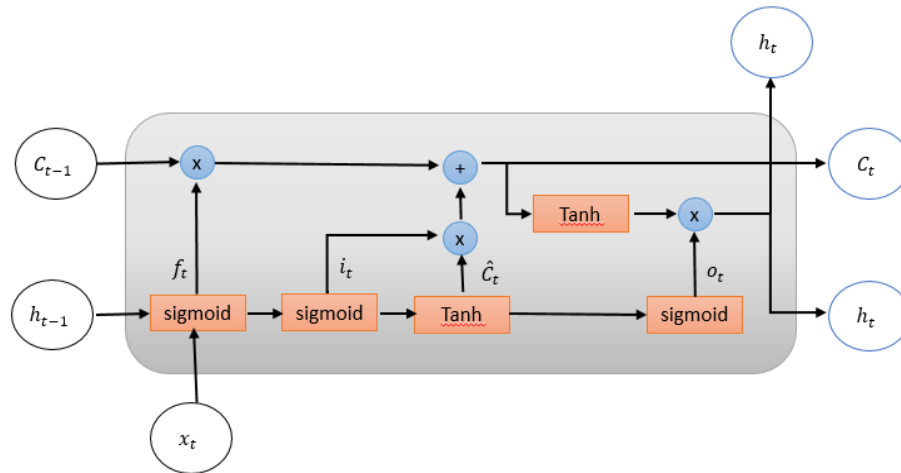
$$p_l = \frac{\lambda_l}{\sum_{l=1}^m \lambda_l}, \quad \text{dimana } l < m, \quad \lambda_1 > \lambda_2 > \dots > \lambda_m \quad (5)$$

$$pk = \sum_{l=1}^m p_l \quad (6)$$

dimana  $x_j$  dan  $x_k$  adalah data,  $\bar{x}_j$  dan  $\bar{x}_k$  adalah rata-rata,  $C$  adalah matrik *kovarians*,  $\lambda$  adalah nilai *eigen*,  $I$  adalah matrik identitas.  $\lambda_l$  adalah nilai *eigen* ke- $l$  yang disusun dari besar ke kecil,  $m$  adalah jumlah variabel.  $p_l$  adalah tingkat kontribusi komponen ke- $l$ ,  $pk$  adalah tingkat kontribusi kumulatif.

### 2.5 LSTM

LSTM merupakan pengembangan dari metode RNN, LSTM mampu menyimpan ingatan jangka pendek dan panjang dengan menggunakan sel memori sehingga dapat menyelesaikan masalah *time series*[32], [41]. Fitur ini memungkinkan LSTM untuk mengelola masalah gradien yang dialami RNN klasik [42]. LSTM memiliki tiga *gate* untuk mengontrol *cell state* ( $C_t$ ) dan *output* ( $h_t$ ), termasuk *forget gate*, *input gate*, dan *output gate*. Arsitektur secara umum LSTM ditunjukkan pada Gambar 3.



Gambar 3. Flowchart LSTM

Pada fase pertama, *forget gate* ( $f_t$ ) akan memutuskan apakah informasi perlu dipertahankan atau dihapus dari sel sebelumnya dengan menggunakan parameter *output* dari sel memori sebelumnya ( $h_{t-1}$ ) dan *input* baru ( $x_t$ ). *Forget gate* dihitung menggunakan Persamaan (8) menghasilkan *output* berupa angka 0 atau 1, di mana 0 berarti membuang informasi dan 1 berarti menyimpan informasi. Langkah kedua, *Input gates* ( $i_t$ ) menentukan informasi baru yang harus disimpan dalam *cell state* menggunakan fungsi aktivasi sigmoid untuk memperoleh informasi terbaru dan fungsi tanh membuat kandidat *cell state* ( $\hat{C}_t$ ) untuk menyimpan nilai baru. Perhitungan *input gate* ditunjukkan pada Persamaan (9) dan (10).

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (8)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (9)$$

$$\hat{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (10)$$

dimana  $W_f$ ,  $W_i$ , dan  $W_c$  masing-masing adalah nilai bobot pada *forget gate*, *input gate*, dan *cell state*.  $b_f$ ,  $b_i$  dan  $b_c$  masing-masing adalah bias pada *forget gate*, *input gate*, dan *cell state*.  $\sigma$  adalah fungsi sigmoid.

Langkah ketiga adalah memperbarui *cell state* lama ( $C_{t-1}$ ) ke *cell state* baru ( $C_t$ ). Rumus perhitungan menggunakan Persamaan (11). Langkah terakhir menentukan *output* dari LSTM. Fungsi sigmoid digunakan untuk menentukan status unit mana yang dihasilkan menggunakan Persamaan (12), kemudian menggunakan Persamaan (13) untuk mendapatkan *state unit output* yang baru [43].

$$C_t = f_t * C_{t-1} + i_t * \hat{C}_t \quad (11)$$

$$O_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (12)$$

$$h_t = O_t \tanh(C_t) \quad (13)$$

dimana  $O_t$  adalah nilai dari output gate,  $W_o$  adalah nilai bobot pada *output gate*,  $b_o$  adalah nilai bias pada *output gate*, dan  $h_t$  adalah *output* final pada waktu ke- $t$ .

## 2.6 Denormalisasi

Denormalisasi berfungsi untuk mengembalikan nilai yang telah dinormalisasikan ke dalam *range* sesungguhnya [44]. Denormalisasi dihitung dengan Persamaan (14) [45].

$$X_t = Y (X_{max} - X_{min}) + X_{min} \quad (14)$$

dimana  $X_t$  nilai data asli,  $Y$  adalah hasil *output*,  $X_{max}$  nilai maksimal data asli,  $X_{min}$  nilai minimal data asli.

## 2.7 MAPE

*Mean Absolute Percentage Error* (MAPE) merupakan cara mengevaluasi akurasi model dengan membandingkan data prediksi dan data *actual* [33]. Nilai MAPE yang semakin kecil menunjukkan semakin baik peramalannya sesuai kriteria yang ditunjukkan pada Tabel 3 [45]. Perhitungan nilai MAPE menggunakan Persamaan (15).

$$MAPE = \frac{\sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right|}{n} \times 100\% \quad (15)$$

dimana  $y_i$  adalah nilai data asli dan  $\hat{y}_i$  adalah nilai hasil prediksi, sedangkan  $n$  adalah jumlah data prediksi.

**Tabel 3.** Kriteria nilai MAPE [46]

MAPE	Kemampuan prediksi
< 10%	Sangat Baik
10 – 20%	Baik
20 – 50%	Cukup Baik
> 50%	Kurang Baik

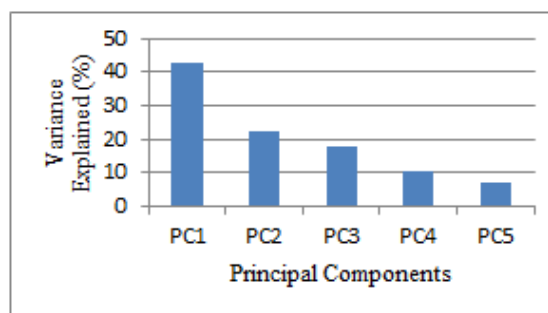
### 3. HASIL DAN PEMBAHASAN

#### 3.1 Hasil Preprocessing Data dan Metode PCA

Data meteorologi pada penelitian ini dikumpulkan memiliki beberapa data yang hilang. Pengisian data yang hilang dilakukan menggunakan metode interpolasi [47]. Pada variabel curah hujan yang memiliki nilai 8888 yang berarti curah hujan di daerah tersebut tidak terukur sehingga nilai tersebut diganti dengan 0,1 mm. Kemudian dilakukan interpolasi pada variabel curah hujan. Pada Variabel tekanan pada penelitian ini diperoleh dari rata-rata tekanan dalam waktu 24 jam.



**Gambar 4.** Plot nilai Eigent



**Gambar 5.** Variance Explained Setiap Principal Components

Variabel yang mempengaruhi curah hujan (variabel independen) dilakukan reduksi fitur menggunakan metode *Principal Component Analysis* (PCA). Reduksi fitur pada penelitian ini bertujuan untuk menemukan variabel yang paling berpengaruh. Proses PCA menggunakan Persamaan (2) sampai (7) menghasilkan *Principal Components* (PC) yang memiliki nilai nilai *eigen* seperti pada Gambar 4. Berdasarkan gambar tersebut menghasilkan 2 PC yang memiliki nilai nilai *eigen* lebih dari satu yaitu pada *Principal Components* pertama (PC1) dan *Principal Components* kedua (PC2). PC3, PC4, dan PC5 memiliki nilai *eigen* kurang dari satu sehingga *Principal Component* tersebut tidak dilakukan analisis lebih lanjut. Pada Gambar 5 PC1 menjelaskan sekitar 42,830% varians dan PC2 sekitar 22,282% varians. PC1-PC2 bersama-sama menjelaskan varians kumulatif sebesar 65,112%.

**Tabel 4.** Component matrix

Variabel	Component	
	1	2
P	-0,444	0,496
E	0,515	<b>0,661</b>
T	<b>0,819</b>	-0,321
RH	<b>-0,740</b>	0,271
LS	<b>0,679</b>	0,505

Nilai positif menunjukkan korelasi bernilai positif pada variabel PC tertentu, begitu pula sebaliknya. Jika nilai negatif menunjukkan korelasi yang bernilai negatif pada variabel PC tertentu [38]. Suatu variabel dikatakan memiliki pengaruh yang kuat jika *component* matrik memiliki nilai lebih dari 0,5. Tabel 4 menunjukkan pada PC1 variabel suhu maksimal dan lama penyinaran matahari berkorelasi positif yang kuat sedangkan variabel kelembaban rata-rata berkorelasi negatif yang kuat. Pada PC2 berkorelasi positif yang kuat pada variabel evaporasi. Oleh karena itu, variabel yang akan digunakan pada tahap prediksi menggunakan LSTM sebanyak 4 variabel yang paling berpengaruh terhadap curah hujan yaitu: variabel evaporasi, suhu maksimal, kelembaban rata-rata, dan lama penyinaran matahari [25], [37].

#### 3.2 Normalisasi dan Pola Time Series

Variabel curah hujan dan variabel hasil PCA dilakukan normalisasi data menggunakan persamaan (1). Normalisasi bertujuan untuk mendapatkan data dalam rentang 0 sampai 1. Hal ini menyebabkan data yang terlalu besar tidak akan mempengaruhi proses pemodelan. Hasil normalisasi pada Tabel 5 dibuat pola *time series* seperti pada Tabel 6 [48]. Pola *time series* pada penelitian ini menggunakan 7 input dengan lima variabel yang diperoleh dari variabel curah



hujan dan variabel hasil PCA kemudian menghasilkan satu output yang merupakan prediksi curah hujan satu hari mendatang. Pola ini mempengaruhi banyaknya data training dan testing berubah menjadi 2092 data dan setiap data memiliki  $7 \times 5$  inputan yang akan digunakan untuk memprediksi curah hujan selanjutnya [49]. Data *time series* dibagi menjadi 80% untuk *training* dan 20% untuk *testing* sehingga terdapat 1673 data training dari 1 Januari 2017 sampai 7 Agustus 2021 dan 419 data testing dari 8 Agustus 2021 sampai 30 September 2022.

Tabel 5. Sampel data normalisasi

Tanggal	E	T	RH	LS	CH
01-01-2017	0,0380	0,2070	0,5127	0,0335	0,0000
02-01-2017	0,0380	0,2051	0,5063	0,0418	0,0000
...	...	...	...	...	...
30-09-2022	0,0335	0,2190	0,4684	0,0456	0,0329

Tabel 6. Pola *time series*

E	T	RH	LS
$E_1, E_2, \dots, E_7$	$T_1, T_2, \dots, T_7$	$RH_1, RH_2, \dots, RH_7$	$LS_1, LS_2, \dots, LS_7$
$E_2, E_3, \dots, E_8$	$T_2, T_3, \dots, T_8$	$RH_2, RH_3, \dots, RH_8$	$LS_2, LS_3, \dots, LS_8$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$E_{2092}, E_{2093}, \dots, E_{2098}$	$E_{2092}, E_{2093}, \dots, E_{2098}$	$RH_{2092}, RH_{2093}, \dots, RH_{2098}$	$LS_{2092}, LS_{2093}, \dots, LS_{2098}$

CH	Target
$CH_1, CH_2, CH_3, CH_4, CH_5, CH_6, CH_7$	$CH_8$
$CH_2, CH_3, CH_4, CH_5, CH_6, CH_7, CH_8$	$CH_9$
$\vdots$	$\vdots$
$CH_{2092}, CH_{2093}, CH_{2094}, CH_{2095}, CH_{2096}, CH_{2097}, CH_{2098}$	$CH_{2099}$

### 3.3 Penerapan LSTM untuk memprediksi curah hujan

Model LSTM pada penelitian ini menggunakan fungsi adam sebagai penentuan optimizernya. Beberapa hyperparameter yang ada di LSTM seperti jumlah *hidden*, *batch size*, dan periode *drop learning rate*. Jumlah *hidden* adalah jumlah perhitungan dalam proses pelatihan, *batch size* digunakan untuk mengatur seberapa sering bobot pada jaringan akan diperbarui, dan periode *drop rate* belajar adalah jumlah pengulangan untuk menentukan kecepatan pembelajaran [50]. Penelitian ini menggunakan beberapa hyperparameter uji coba seperti jumlah *hidden*, *batch size*, dan periode *drop learning rate*. Nilai optimal hyperparameter pelatihan model yang diperoleh pada penelitian ini menggunakan metode uji coba seperti pada Tabel 7.

Berdasarkan Tabel 7 diperoleh rata-rata nilai MAPE pada setiap *hidden layer* tidak jauh berbeda. Rata-rata MAPE terkecil terdapat pada 64 *hidden layer* dengan rata-rata sebesar 0,0272. Pada 64 *hidden layer* yang memiliki MAPE terkecil terletak pada 16 *batch size* dan 100 *learn rate drop period* sebesar 0,0063. Pada 100 *hidden layer* yang memiliki MAPE terkecil terdapat uji coba 32 *batch size* dan 50 *learn rate drop period* sebesar 0,0018. Pada uji coba 200 *hidden layer* yang memiliki MAPE terkecil pada 32 *batch size* dan 150 *learn rate drop period* sebesar 0,0096. Pada 32 *hidden layer* memiliki MAPE terkecil terletak pada 16 *batch size* dan 50 *learn rate drop period* sebesar 0,0050. Secara keseluruhan nilai MAPE terendah adalah 0,0018 yang terletak pada 100 *hidden layer*, 32 *batch size* dan 50 *learn rate drop period*. Hasil prediksi curah hujan dengan nilai MAPE terendah divisualisasikan pada Gambar 6.

Tabel 7. Nilai MAPE menggunakan parameter curah hujan

Hidden Layer	Batch Size	Learn rate	MAPE	Rata-rata MAPE	
100	16	50	0,0086	0,0464	
		100	0,0497		
		150	0,0448		
	32	50	0,0018		
		100	0,0241		
		150	0,0121		
		64	50		0,0261
			100		0,1443
			150		0,1062
200	16	50	0,0639	0,0467	
		100	0,0515		
		150	0,0490		
	32	50	0,0142		
		100	0,0352		
		150	0,0096		



		64	50	0,0116	
			100	0,0190	
			150	0,1660	
32	16	50	0,0050	0,0327	
			100	0,0372	
			150	0,0058	
	32	50	0,0316		
			100	0,0096	
			150	0,0101	
	64	50	0,0320		
			100	0,0452	
			150	0,1174	
64	16	50	0,0523	0,0272	
			100	0,0063	
			150	0,0286	
	32	50	0,0315		
			100	0,0172	
			150	0,0073	
	64	50	0,0420		
			100	0,0465	
			150	0,0135	

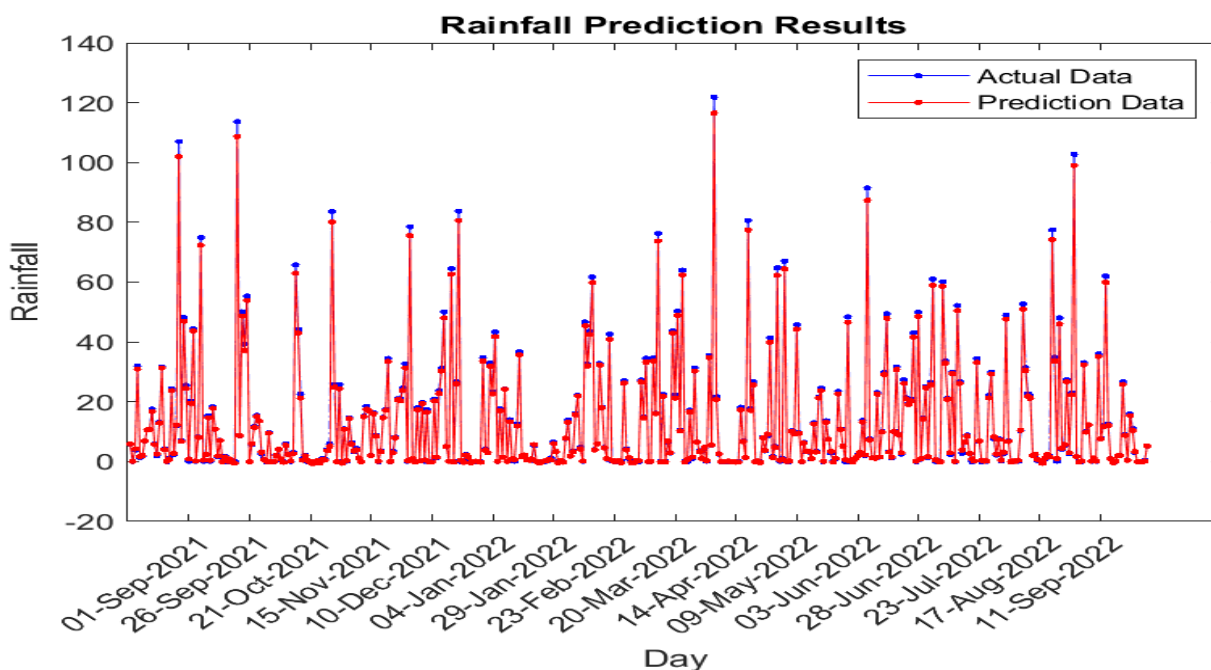
Gambar 6 menunjukkan hasil prediksi curah hujan dari tanggal 8 Agustus 2021 sampai 30 September 2022 menggunakan LSTM dengan parameter uji coba 100 *hidden layer*, 32 *batch size*, dan 50 *learn rate drop period*. Grafik prediksi (merah) memiliki pola sangat mirip dengan grafik *actual* (biru). Prediksi curah hujan menggunakan LSTM mendekati nilai sebenarnya. Berdasarkan hasil prediksi diperoleh informasi bahwa curah hujan di Luwu Utara rata-rata mengalami hujan lebat hingga sangat lebat selama 1 sampai 3 kali per bulan. Curah hujan yang sangat lebat terjadi pada tanggal 28 Agustus 2021 sebesar 101,9734 mm, 21 September 2021 sebesar 108,6528 mm, dan 5 April 2022 sebesar 116,5510 mm. Evaluasi model diperoleh MAPE sebesar 0,0018. Oleh karena itu, model dalam penelitian ini bagus dalam memprediksi curah hujan yang akan datang sehingga dapat memberikan informasi kepada masyarakat khususnya di Kabupaten Luwu Utara untuk mengantisipasi curah hujan. Masyarakat dapat memanfaatkan curah hujan dan mengantisipasi kerugian akibat curah hujan.

**Tabel 8.** Perbandingan metode prediksi curah hujan yang digunakan dalam literatur

Tipe NN	Variabel Prediksi Curah Hujan	Wilayah dan Dataset	Evaluasi	Penulis & Tahun
LSTM	Uji coba parameter El-Nino dan Indian Ocean Dipole (IOD). Uji coba parameter curah hujan.	Sidoarjo, Indonesia.	MAAPE masing-masing uji coba sebesar 0,9644 dan 0,5810.	Dina Zatusiva Haq [50]
LSTM	suhu, titik embun, kelembaban, tekanan angin, kecepatan angin, dan arah angin.	Bangladesh Meteorological Department (BMD).	Akurasi 76%	Imrus Salehin, dkk. [51]
LSTM GRU PCA-LSTM PCA-GRU	Variabel di Sungai Muskegon yaitu maksimum, minimum dan rata-rata di masing-masing variabel temperatur (°C), Titik embun maksimum (°C), kelembaban (%), tekanan permukaan laut (hPa), dan visibilitas (km), serta Kecepatan angin maks (km/h), Kecepatan angin rata-rata (km/h), Pengendapan (mm), Variabel di Sungai Pearl yaitu maksimum, minimum dan rata-rata di masing-masing variable temperature (°C), titik embun (°C), kelembaban (%), dan tekanan permukaan laut (hPa), serta Kecepatan angin maks (km/h), Kecepatan angin rata-rata (km/h).	Sungai Muskegon dan Sungai Pearl di AS.	Nilai WMAPE di Sungai Muskegon pada masing-masing metode 27,180%, 27,820%, 11,410%, 11,530% Nilai WMAPE di Sungai Pearl pada masing-masing metode 57,570%, 58,440%, 53,770%, 55,060%.	Jiangwei Zhang, dkk. [22]

MPL	Suhu Maksimum, Suhu Minimum,	<i>The National Center of Hydrology and Meteorology Department (NCHM) di Wilayah Simtokha, Bhutan, Thimphu.</i>	Nilai R <sup>2</sup> masing-masing model 0,501; 0,583; 0,663; 0,749; 0,777; dan 0,870	Manoj Chhetri, dkk. [48]
CNN	Milimeter Curah Hujan, Persentase			
GRU	Kelembaban Relatif, Lama Sinar			
BLSTM	Matahari, dan Kecepatan Angin.			
LSTM				
LSTM-GRU.				
Proposed method (PCA-LSTM)	Variabel tekanan, evaporasi, suhu maksimal, kelembaban rata-rata, dan lama penyinaran matahari.	Stasiun Meteorologi Andi Jemma di Luwu Utara.	MAPE 0,0018	

Tabel 8 merupakan deskripsi penelitian yang pernah melakukan prediksi curah hujan. Penelitian tersebut menggunakan banyak parameter *input* yang berbeda dan metode yang hampir berbeda. Metode dan parameter *input* pada penelitian kami memiliki *error* paling kecil dibandingkan dengan penelitian sebelumnya. Pada penelitian ini menggunakan variabel input yang paling berpengaruh yang diperoleh dari metode PCA, yaitu: variabel evaporasi, suhu maksimal, kelembaban rata-rata, dan lama penyinaran matahari dalam memprediksi curah hujan yang akan datang menghasilkan nilai MAPE 0,0018. Namun, penelitian ini tidak memaparkan jenis-jenis tanaman yang cocok sebagai komoditas unggulan sesuai dengan kondisi cuaca di wilayah setempat. Penelitian selanjutnya dapat menentukan kalender pola tanam sehingga sektor pertanian bisa memanfaatkan curah hujan secara optimal [52], [53]. Selain itu, penelitian selanjutnya dapat mengembangkan metode hibrida menggunakan model *deep learning* yang berbeda dan metode PCA lebih ditingkatkan lagi agar lebih aplikatif dan efisien [48].



Gambar 6. Plot Data Prediksi dan Data Actual

#### 4. KESIMPULAN

Algoritma LSTM digunakan untuk memprediksi curah hujan di Luwu Utara, Sulawesi Selatan. Prediksi curah hujan menggunakan fitur yaitu: tekanan, evaporasi, suhu maksimal, kelembaban rata-rata, dan lama penyinaran matahari. Variabel tersebut dilakukan ekstraksi fitur menggunakan PCA diperoleh 4 variabel yang paling berpengaruh yaitu: variabel evaporasi, suhu maksimal, kelembaban rata-rata, dan lama penyinaran matahari. Berdasarkan beberapa uji coba menghasilkan MAPE terkecil terdapat pada uji coba 100 *hidden layer*, 32 *batch size*, dan 50 *learn rate drop period* sebesar 0,0018. Berdasarkan hasil prediksi diperoleh informasi bahwa Luwu Utara rata-rata mengalami hujan lebat 1 hingga 3 kali per bulan. Curah hujan yang sangat lebat terjadi pada tanggal 28 Agustus 2021 sebesar 101,9734 mm, 21 September 2021 sebesar 108,6528 mm, dan 5 April 2022 sebesar 116,5510 mm. Hasil prediksi tersebut sangat baik untuk memprediksi curah hujan di Luwu Utara, Sulawesi Selatan.

#### UCAPAN TERIMAKASIH

Kami berterima kasih kepada pihak BMKG perak II dan BMKG andi jemma yang telah membantu dan menyediakan data untuk melakukan penelitian ini.

## REFERENCES

- [1] F. Zul, Kaimuddin, and Rafiuddin, "Analysis of climate and population dynamics of Conomoporpha cramerella pest in North Luwu," *IOP Conf. Ser. Earth Environ. Sci.*, vol. 486, no. 1, 2020, doi: 10.1088/1755-1315/486/1/012084.
- [2] A. Nugraha, M. A. Heryanto, E. Wulandari, and P. Pardian, "Heading towards sustainable cacao agribusiness system (a case study in North Luwu, South Sulawesi, Indonesia)," *IOP Conf. Ser. Earth Environ. Sci.*, vol. 306, no. 1, 2019, doi: 10.1088/1755-1315/306/1/012035.
- [3] A. S. Ahmar, M. A. Tiro, and R. P. Utama, "Rainfall forecasting using backpropagation neural network : a case in of North Luwu , Indonesia," *Jt. Work. KO2PI 1st Int. Conf. Adv. Sci. Innov.*, no. January, pp. 1–7, 2018, doi: 10.4108/eai.23-4-2018.2277580.
- [4] Kaimuddin *et al.*, "Water requirement for cocoa (Theobroma cacao L.) plant and the effect of climate factors on the distribution of the cocoa pod borer attacks (Conopomorpha cramerella Snellen) in North Luwu Regency using Cropwat 8.0," *IOP Conf. Ser. Earth Environ. Sci.*, vol. 575, no. 1, pp. 0–11, 2020, doi: 10.1088/1755-1315/575/1/012116.
- [5] N. H. Muzaki, Eriska Febriati, and Yosafat Donni Haryanto, "Pemanfaatan satelit himawari-8 untuk estimasi curah hujan dengan metode Convective Stratiform Technique (CST) Dan Modified Convective Stratiform Technique (Mcsst) di wilayah ekuatorial dan musonal (studi kasus Sulawesi 2020)," *J Stat.*, vol. 14, no. 2, pp. 67–73, 2022, doi: 10.36456/jstat.vol14.no2.a3860.
- [6] S. Nurdiati, A. Sopaheluwakan, and P. Septiawan, "Joint distribution analysis of forest fires and precipitation in response to ENSO, IOD, and MJO (study case: Sumatra, Indonesia)," *Atmosphere (Basel)*, vol. 13, no. 4, pp. 1–20, 2022, doi: 10.3390/atmos13040537.
- [7] J. A. I. Paski *et al.*, "Analysis of multi-scale hydrometeorological triggering flash flood event of the 13 July 2020 in North Luwu, South Sulawesi," *IOP Conf. Ser. Earth Environ. Sci.*, vol. 893, no. 1, 2021, doi: 10.1088/1755-1315/893/1/012014.
- [8] M. T. R. Abdullah, H. Cangara, and Arianto, "The analysis of communication strategies of North Luwu regency government in socio-economic recovery after landslides and flash flood," *Proc. 2nd Int. Conf. Sci. Technol. Mod. Soc. (ICSTMS 2020)*, vol. 576, no. 24, pp. 453–461, 2021, doi: 10.2991/assehr.k.210909.098.
- [9] M. P. Hatta, F. Fadlin, R. Harun, Y. Elfita, and I. Renreng, "Application of 2D numerical simulation for the analysis of July 2020 North Luwu flood," *IOP Conf. Ser. Earth Environ. Sci.*, vol. 841, no. 1, 2021, doi: 10.1088/1755-1315/841/1/012028.
- [10] Herman, "Banjir terjang pemukiman di Luwu Utara, 982 warga terdampak banjir," *liputan6.com*, 2021. <https://www.liputan6.com/news/read/4655925/banjir-terjang-pemukiman-di-luwu-utara-982-warga-terdampak-banjir> (accessed Aug. 27, 2022).
- [11] A. Setyadi, "Banjir kembali terjang Luwu Utara, 76 KK terdampak," *news.okezone.com/*, 2022. <https://news.okezone.com/read/2022/07/06/340/2624736/banjir-kembali-terjang-luwu-utara-76-kk-terdampak> (accessed Aug. 05, 2022).
- [12] Y. P. Pasaribu, H. Fitrianti, and D. R. Suryani, "Rainfall forecast of Merauke using Autoregressive Integrated Moving Average model," *E3S Web Conf.*, vol. 73, pp. 1–5, 2018, doi: 10.1051/e3sconf/20187312010.
- [13] S. Sunardi, A. Yudhana, and G. Z. Muflih, "Sistem prediksi curah hujan bulanan menggunakan jaringan saraf tiruan backpropagation," *J. Sist. Inf. Bisnis 02(2020)*, vol. 02, pp. 155–162, 2020.
- [14] R. E. Caraka *et al.*, "Generalized Spatio Temporal Autoregressive rainfall-ENSO pattern in East Java Indonesia," *1st 2018 Indones. Assoc. Pattern Recognit. Int. Conf. Ina. 2018 - Proc.*, no. September, pp. 75–79, 2019, doi: 10.1109/INAPR.2018.8627042.
- [15] J. Badriyah, A. Fariza, and T. Harsono, "Prediksi curah hujan menggunakan Long Short Term Memory," *J. Media Inform. Budidarma*, vol. 6, pp. 1297–1303, 2022, doi: 10.30865/mib.v6i3.4008.
- [16] Y. C. Chen, J. J. Gao, Z. H. Bin, J. Z. Qian, R. L. Pei, and H. Zhu, "Application study of IFAS and LSTM models on runoff simulation and flood prediction in the Tokachi River basin," *J. Hydroinformatics*, vol. 23, no. 5, pp. 1098–1111, 2021, doi: 10.2166/hydro.2021.035.
- [17] C. Chen *et al.*, "Forecast of rainfall distribution based on fixed sliding window long short-term memory," *Eng. Appl. Comput. Fluid Mech.*, vol. 16, no. 1, pp. 248–261, 2022, doi: 10.1080/19942060.2021.2009374.
- [18] Z. Chao, F. Pu, Y. Yin, B. Han, and X. Chen, "Research on real-time local rainfall prediction based on MEMS sensors," *J. Sensors*, vol. 2018, 2018, doi: 10.1155/2018/6184713.
- [19] S. Poornima and M. Pushpalatha, "Prediction of rainfall using intensified LSTM based Recurrent Neural Network with Weighted Linear Units," *Atmosphere (Basel)*, vol. 10, no. 11, 2019, doi: 10.3390/atmos10110668.
- [20] A. G. Salman, Y. Heryadi, E. Abdurahman, and W. Suparta, "single layer & multi-layer Long Short-Term Memory (LSTM) model with intermediate variables for weather forecasting," *Procedia Comput. Sci.*, vol. 135, pp. 89–98, 2018, doi: 10.1016/j.procs.2018.08.153.
- [21] J. Kang, H. Wang, F. Yuan, Z. Wang, J. Huang, and T. Qiu, "Prediction of precipitation based on recurrent neural networks in jingdezhen, jiangxi province, China," *Atmosphere (Basel)*, vol. 11, no. 3, pp. 1–17, 2020, doi: 10.3390/atmos11030246.
- [22] J. Zhang *et al.*, "Daily runoff forecasting by deep recursive neural network," *J. Hydrol.*, vol. 596, no. December 2020, p. 126067, 2021, doi: 10.1016/j.jhydrol.2021.126067.
- [23] G. Mao *et al.*, "Comprehensive comparison of artificial neural networks and long short-term memory networks for rainfall-runoff simulation," *Phys. Chem. Earth*, vol. 123, p. 103026, 2021, doi: 10.1016/j.pce.2021.103026.
- [24] R. Pujiyanto, Adiwijaya, and A. A. Rahmawati, "Analisis Ekstraksi Fitur Principle Component Analysis pada Klasifikasi Microarray Data Menggunakan Classification And Regression Trees," *eProceedings ...*, vol. 6, no. 1, pp. 2368–2379, 2019.
- [25] B. Rahmani, S. Javadi, and S. M. H. Shahdany, "Evaluation of aquifer vulnerability using PCA technique and various clustering methods," *Geocarto Int.*, vol. 36, no. 18, pp. 2117–2140, 2021, doi: 10.1080/10106049.2019.1690057.
- [26] M. Tripathi and S. K. Singal, "Use of Principal Component Analysis for parameter selection for development of a novel Water Quality Index: A case study of river Ganga India," *Ecol. Indic.*, vol. 96, no. September 2018, pp. 430–436, 2019, doi:



- 10.1016/j.ecolind.2018.09.025.
- [27] Y. Wen, P. Lin, and X. Nie, "Research of stock price prediction based on PCA-LSTM model," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 790, no. 1, 2020, doi: 10.1088/1757-899X/790/1/012109.
- [28] D. Geng, H. Zhang, and H. Wu, "Short-term wind speed prediction based on principal component analysis and LSTM," *Appl. Sci.*, vol. 10, no. 13, 2020, doi: 10.3390/app10134416.
- [29] Y. Ling, Q. Yue, C. Chai, Q. Shan, D. Hei, and W. Jia, "Nuclear accident source term estimation using Kernel Principal Component Analysis, Particle Swarm Optimization, and Backpropagation Neural Networks," *Ann. Nucl. Energy*, vol. 136, p. 107031, 2020, doi: 10.1016/j.anucene.2019.107031.
- [30] BMKG, "Data online pusat database BMKG." <http://dataonline.bmkg.go.id/home>.
- [31] I. G. W. Sena, J. W. Dillak, P. Leunupun, and A. J. Santoso, "Predicting rainfall intensity using Naïve Bayes and Information Gain methods ( case study : Sleman Regency )," *J. Phys. Conf. Ser. 1577(1)*, 2020, doi: 10.1088/1742-6596/1577/1/012011.
- [32] I. I. Zulfa, D. C. R. Novitasari, F. Setiawan, A. Fanani, and M. Hafiyusholeh, "Prediction of sea surface current velocity and direction using LSTM," *IJEIS (Indonesian J. Electron. Instrum. Syst.)*, vol. 11, no. 1, p. 93, 2021, doi: 10.22146/ijeis.63669.
- [33] E. A. Kusnanti, D. C. R. Novitasari, F. Setiawan, A. Fanani, M. Hafiyusholeh, and G. I. Permata Sari, "Predicting velocity and direction of ocean surface currents using Elman Recurrent Neural Network method," *J. Inf. Syst. Eng. Bus. Intell.*, vol. 8, no. 1, pp. 21–30, 2022, doi: 10.20473/jisebi.8.1.21-30.
- [34] N. Hasan, "A methodological approach for predicting COVID-19 epidemic using EEMD-ANN hybrid model," *Internet of Things (Netherlands)*, vol. 11, 2020, doi: 10.1016/j.iot.2020.100228.
- [35] Z. Chang, Y. Zhang, and W. Chen, "Electricity price prediction based on hybrid model of adam optimized LSTM neural network and wavelet transform," *Energy*, vol. 187, p. 115804, 2019, doi: 10.1016/j.energy.2019.07.134.
- [36] B. Deepa and K. Ramesh, "Epileptic seizure detection using deep learning through min max scaler normalization," *Int. J. Health Sci. (Qassim)*, vol. 6, no. April, pp. 10981–10996, 2022, doi: 10.53730/ijhs.v6ns1.7801.
- [37] M. Abdipour, M. Younessi-Hmazekhanlu, S. H. R. Ramazani, and A. hassan omidi, "Artificial neural networks and multiple linear regression as potential methods for modeling seed yield of safflower (*Carthamus tinctorius L.*)," *Ind. Crops Prod.*, vol. 127, no. November 2018, pp. 185–194, 2019, doi: 10.1016/j.indcrop.2018.10.050.
- [38] S. Manandhar, S. Dev, Y. H. Lee, S. Winkler, and Y. S. Meng, "Systematic study of weather variables for rainfall detection," no. 2, pp. 3035–3038, 2018.
- [39] N. M. Sunariadi, S. N. Fadilah, D. Candra, R. Novitasari, S. Teknologi, and U. I. N. S. Ampel, "Analisis resiko kanker serviks menggunakan PCA-ANFIS berdasarkan historical medical record," *J. Media Inform. Budidarma*, vol. 6, pp. 1349–1355, 2022, doi: 10.30865/mib.v6i3.3901.
- [40] L. N. Desainaini, A. Mualimah, D. C. R. Novitasari, and M. Hafiyusholeh, "Application of Fuzzy K-Nearest Neighbor (FKNN) to Detect the Parkinson's Disease," *Inpr. Indones. J. Pure Appl. Math.*, vol. 1, no. 1, pp. 8–16, 2019, doi: 10.15408/inprime.v1i1.12827.
- [41] B. Tian, G. Wang, Z. Xu, Y. Zhang, and X. Zhao, "Communication delay compensation for string stability of CACC system using LSTM prediction," *Veh. Commun.*, vol. 29, p. 100333, 2021, doi: 10.1016/j.vehcom.2021.100333.
- [42] J. Mejia, L. Avelar-Sosa, B. Mederos, E. S. Ramirez, and J. D. Diaz Roman, "Prediction of time series using an analysis filter bank of LSTM units," *Comput. Ind. Eng.*, vol. 157, p. 107371, 2021, doi: 10.1016/j.cie.2021.107371.
- [43] H. Tian, P. Wang, K. Tansey, J. Zhang, S. Zhang, and H. Li, "An LSTM neural network for improving wheat yield estimates by integrating remote sensing data and meteorological data in the Guanzhong Plain, PR China," *Agric. For. Meteorol.*, vol. 310, no. 17, p. 108629, 2021, doi: 10.1016/j.agrformet.2021.108629.
- [44] T. Lattifia, P. Wira, N. Kadek, and D. Rusjayanthi, "Model prediksi cuaca menggunakan metode LSTM," *J. Ilm.*, vol. 3, no. 1, 2018.
- [45] B. Y. Phiadelvira, D. Z. Haq, D. C. R. Novitasari, and F. Setiawan, "Prediksi besar daya Listrik Tenaga Gelombang Laut Metode Oscillating Water Coloumn (PLTGL-OWC) di Banyuwangi menggunakan Extreme Learning Machine (ELM)," *Unnes J. Math.*, vol. 11, no. 1, pp. 1–7, 2022, doi: 10.15294/ujm.v11i1.50967.
- [46] M. E. Ervina, R. Silvi, and I. R. N. Wissono, "Peramalan jumlah penumpang kereta api di Indonesia dengan Resilient Back-Propagation (Rprop) Neural Network," *J. Mat. "MANTIK"*, vol. 4, no. 2, pp. 90–99, 2018, doi: 10.15642/mantik.2018.4.2.90-99.
- [47] T. Moon, S. Hong, H. Y. Choi, D. H. Jung, S. H. Chang, and J. E. Son, "Interpolation of greenhouse environment data using multilayer perceptron," *Comput. Electron. Agric.*, vol. 166, no. July, p. 105023, 2019, doi: 10.1016/j.compag.2019.105023.
- [48] M. Chhetri, S. Kumar, P. P. Roy, and B. G. Kim, "Deep BLSTM-GRU model for monthly rainfall prediction: A case study of Simtokha, Bhutan," *Remote Sens.*, vol. 12, no. 19, pp. 1–13, 2020, doi: 10.3390/rs12193174.
- [49] 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)*, Jul. 2022, pp. 191–196, doi: 10.1109/IAICT55358.2022.9887474.
- [50] D. Z. Haq et al., "Long Short-Term Memory algorithm for rainfall prediction based on El-Nino and IOD data," *Procedia Comput. Sci.*, vol. 179, no. 2019, pp. 829–837, 2021, doi: 10.1016/j.procs.2021.01.071.
- [51] 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," *Proc. 2020 IEEE Int. Women Eng. Conf. Electr. Comput. Eng. WIECON-ECE 2020*, pp. 5–8, 2020, doi: 10.1109/WIECON-ECE52138.2020.9398022.
- [52] O. Darnius and S. Sitorus, "Plant calendar pattern based on rainfall forecast and the probability of its success in Deli Serdang regency of Indonesia," *J. Phys. Conf. Ser.*, vol. 983, no. 1, 2018, doi: 10.1088/1742-6596/983/1/012113.
- [53] M. R. Arif Yudianto, T. Agustin, R. Morgan James, F. Imannisa Rahma, A. Rahim, and E. Utami, "Rainfall forecasting to recommend crops varieties using moving Average and Naive Bayes methods," *Int. J. Mod. Educ. Comput. Sci.*, vol. 13, no. 3, pp. 23–33, 2021, doi: 10.5815/ijmecs.2021.03.03.