

DROUGHT PREDICTION USING LSTM MODEL WITH STANDARDIZED PRECIPITATION INDEX ON THE NORTH COAST OF CENTRAL JAVA

Aji Supriyanto^{*1}, Eri Zuliarso², Eko Taufiq Suharmanto³, Hana Amalina⁴, Fitri Damaryanti⁵

^{1,2,3,4,5}Faculty of Information Technology and Industry, Master of Information Technology, Universitas Stikubank Semarang, Indonesia
Email: ¹ajisup@edu.unisbank.ac.id

(Article received: November 27, 2024; Revision: December 21, 2024; published: December 29, 2024)

Abstract

Fluctuating weather can trigger hydrometeorological disasters, especially affecting farmers and fishermen on the north coast of Central Java. Weather predictions including drought are very important to anticipate drought disasters. Deep learning-based prediction models such as Long Short Term Memory (LSTM) are used in an effort to reduce the impact of drought. The purpose of this study is to prove the level of accuracy of the LSTM model and determine the drought index with the Standardized Precipitation Index (SPI). The LSTM model is used to predict drought based on the SPI, while the SPI acts as a drought index that considers precipitation (rainfall) for a period of 1, 3, and 6 months. Predictions use rainfall data obtained from online data from the Central Java BMKG UPT Indonesia for the period 2010-2023 in the Tegal City and Semarang City station areas. The results of data treatment with LSTM can effectively analyze and capture complex patterns in meteorological data to predict drought events accurately. The effectiveness of the model is shown by the relatively small MAE and RMSE results, namely MAE 0.163 - 0.352 and RMSE 0.247-0.515. The best prediction result is the 3-month SPI in the Semarang area with MAE 0.163 and RMSE 0.274. While the prediction result with the largest error is the 1-month SPI in the Tegal area. Drought modeling using LSTM has been successfully implemented for the northern coast of Central Java using the Streamlit Framework and can process and visualize the drought prediction system well.

Keywords: *Climate, Drought, LSTM, North Coast of Central Java, Prediction, SPI.*

PREDIKSI KEKERINGAN MENGGUNAKAN MODEL LSTM DENGAN INDEKS CURAH HUJAN TERSTANDAR DI PANTAI UTARA JAWA TENGAH

Abstrak

Cuaca yang fluktuatif dapat memicu bencana hidrometeorologi terutama berdampak bagi para petani dan nelayan di pesisir utara Jawa Tengah. Prediksi cuaca termasuk kekeringan sangat penting untuk mengantisipasi bencana kekeringan. Model prediksi berbasis deep learning seperti *Long Short Term Memory* (LSTM) digunakan dalam upaya mengurangi dampak kekeringan. Tujuan penelitian ini untuk membuktikan tingkat akurasi model LSTM dan menentukan indeks kekeringan dengan *Standardized Precipitation Index* (SPI). Model LSTM digunakan untuk memprediksi kekeringan berdasarkan SPI, sedangkan SPI berperan sebagai indeks kekeringan yang mempertimbangkan presipitasi (curah hujan) untuk rentang waktu 1, 3, dan 6 bulan. Prediksi menggunakan data curah hujan yang diperoleh dari data online UPT BMKG Jawa Tengah Indonesia pada periode tahun 2010-2023 di wilayah stasiun Kota Tegal dan Kota Semarang. Hasil perlakuan data dengan LSTM ini efektif dapat menganalisis dan menangkap pola kompleks dalam data meteorologi untuk memprediksi kejadian kekeringan secara akurat. Efektivitas model ditunjukkan dengan hasil MAE dan RMSE yang relatif kecil yakni MAE 0,163 - 0,352 dan RMSE 0,247-0,515. Hasil prediksi terbaik adalah SPI 3 bulan di wilayah Semarang dengan MAE 0,163 dan RMSE 0,274. Sedangkan hasil prediksi dengan error terbesar adalah SPI 1 bulan di wilayah Tegal. Pemodelan kekeringan menggunakan LSTM berhasil diimplementasikan untuk wilayah pantai utara Jawa Tengah menggunakan Framework Streamlit dan dapat mengolah serta memvisualisasikan sistem prediksi kekeringan dengan baik.

Kata kunci: *Iklim, Kekeringan, LSTM, Pesisir utara Jawa Tengah, Prediksi, SPI.*

1. PENDAHULUAN

Indonesia merupakan negara dengan letak geografis pada daerah tropis yang memiliki dua

musim yaitu kemarau dan hujan. Musim kemarau identik dengan berkurangnya curah hujan yang berkepanjangan atau terjadinya penurunan intensitas curah hujan yang signifikan dibandingkan dengan kondisi normal. Kekeringan mengurangi ketersediaan air dan menurunkan produksi pertanian. Bahkan di wilayah pesisir utara Jawa terutama di propinsi Jawa Tengah Indonesia terjadi tingkat kekeringan sedang hingga tinggi untuk daerah bentang alam karst[1]. Bencana kekeringan sering terjadi di Indonesia[1][2][3] penyebab utamanya karena Indonesia sebagai negara tropis yang terletak di antara 95 derajat (95°) hingga 141 derajat (141°) Bujur Timur, serta terletak antara 6 derajat (6°) lintang utara sampai 11 derajat (11°) lintang selatan. Ini termasuk propinsi Jawa Tengah yang terletak di sepanjang garis khatulistiwa antara 5°40' sampai 8°30' Lintang Selatan dan 108°30' sampai 111°30' Bujur Timur (termasuk Pulau Karimunjawa)[4]. Di Indonesia, curah hujan merupakan input paling besar dalam sistem hidrologi[1][5]. Curah hujan yang lebih rendah dari rata-ratanya merupakan tanda awal terjadinya kekeringan[6][7][8].

Perubahan iklim juga dapat menyebabkan curah hujan lebih fluktuatif, sehingga memicu terjadinya bencana hidrometeorologi[1][5][9][5]. Dibandingkan dengan bencana lainnya, kekeringan memiliki dampak yang paling besar[1][10][11]. Hal ini dikarenakan sulitnya memprediksi awal dan akhir periode kekeringan[9]. Kekeringan dapat terjadi selama berbulan-bulan bahkan bertahun-tahun[1][12][13]. Pulau Jawa memiliki frekuensi kekeringan tertinggi dibandingkan pulau lain di Indonesia selama 30 tahun (1979-2009) dan hingga saat ini masih rawan mengalami kekeringan[11]. Kejadian kekeringan di Indonesia terkait erat dengan fenomena iklim global, khususnya El Niño dan *Indian Ocean Dipole* (IOD)[14]. Kejadian El Niño yang kuat juga dapat menyebabkan kebakaran hebat di wilayah hutan dan lahan gambut[15] yang menyebabkan kabut asap atau asap lintas batas. Diperlukan pemantauan berkelanjutan dan perkiraan kondisi meteorologi yang akurat untuk mengantisipasi dan meminimalkan dampak negatif dari peristiwa kekeringan.

Pulau Jawa memiliki frekuensi kekeringan tertinggi dibandingkan pulau-pulau lain di Indonesia selama 30 tahun (1979-2009) dan hingga saat ini masih rawan mengalami kekeringan[11]. Analisis kekeringan di Pulau Jawa Indonesia telah dilakukan dengan menggunakan *Standardized Precipitation and Evapotranspiration Index* (SPEI) untuk menentukan tingkat risiko kekeringan spasial dan temporal. Selain itu penelitian juga menggunakan *Standardized Precipitation Index* (SPI) sebagai pembanding dalam mendeteksi kekeringan dan juga divalidasi dengan kejadian kekeringan historis dengan data antara tahun 1979- 2020[11]. Selanjutnya Siswanto dkk (2022) menganalisis kekeringan meteorologis berbasis SPI di pulau Jawa khususnya

di kabupaten penghasil beras terbesar di Indonesia dengan data tahun 1981-2021[14].

Namun kedua penelitian tersebut tidak menggunakan RNN-LSTM untuk analisis prediksi dalam pola urutan (*time-series*) yang panjang yang berfungsi agar dapat menangkap pola temporal dalam data SPI guna menghasilkan peramalan yang lebih akurat. Jaringan saraf berulang (*Recurrent Neural Network/ RNN*) umumnya digunakan untuk prediksi deret waktu[16][7]. Sedangkan LSTM (*Long Short-Term Memory*) merupakan arsitektur khusus RNN yang secara khusus dikembangkan untuk merepresentasikan urutan temporal. LSTM memiliki saling ketergantungan yang luas, sehingga membuatnya lebih presisi daripada RNN tradisional[7]. Selain itu kinerja model RNN-LSTM yang unggul dibandingkan metode tradisional. Arsitektur unik LSTM dalam studi prediksi, khususnya untuk kekeringan, karena unggul dalam mengklasifikasikan proses dan memprediksi rangkaian waktu dengan tingkat akurasi tinggi[7].

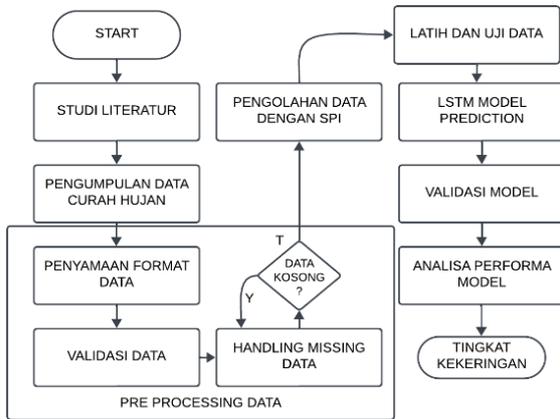
Kekeringan juga dapat meningkatkan suhu udara. Perkiraan peningkatan suhu dapat digunakan model machine learning, salah satunya adalah metode Long Short-Term Memory (LSTM). Untuk kinerja prediksi suhu jangka pendek (1 hingga 7 hari) LSTM memiliki kinerja lebih baik daripada ConvLSTM, dan *Multi-Layer Perceptron* (MLP)[17]. Terkait dengan cuaca dan bencana alam LSTM juga untuk memprediksi akurasi tinggi gelombang laut[18], bencana gempa bumi[16], dan kenaikan permukaan air[18]. Pada penelitian Kadam et., al.(2024) penggunaan LSTM dan SPI untuk mengukur kekeringan, namun studi kasusnya di India dengan data dari tahun 1980 – 2020 atau rentang data selama 31 tahun. Penelitian serupa oleh Taylan et.al., (2024) menggunakan LSTM dan SPI untuk mengukur kekeringan dengan studi kasus di Turkiye dengan data tahun 1991-2023 atau rentang data selama 23 tahun[19].

Berdasarkan penelitian tersebut maka tujuan penelitian ini akan membuktikan tingkat akurasi model LSTM dan indeks kekeringan dengan SPI setelah diuji dengan data antara tahun 2010 – 2023 atau rentang data selama 14 tahun. Hal ini juga untuk membuktikan bahwa metode SPI dengan data presipitasi ideal disusun untuk periode tanpa henti minimal 30 tahun[20], menunjukkan apakah konsep tersebut masih berlaku atau tidak. Penelitian ini juga mengukur kinerja atau akurasi model terhadap nilai aktual dengan menghitung *Coefficient of determination* (R^2), *Mean Absolute Error* (MAE), *Root Mean Square Error* (RMSE), dan *Mean Absolute Percentage Error* (MAPE) [21][22]. Penelitian ini selain untuk menguji akurasi model LSTM, juga dapat berkontribusi dalam perencanaan tata kelola sumber daya air agar lebih efektif dan efisien. Selain itu sebagai kelanjutan dapat digunakan sebagai instrumen dalam strategi mitigasi menghadapi dampak kekeringan dimasa mendatang

khususnya untuk wilayah pantai utara di Jawa Tengah Indonesia.

2. METODE PENELITIAN

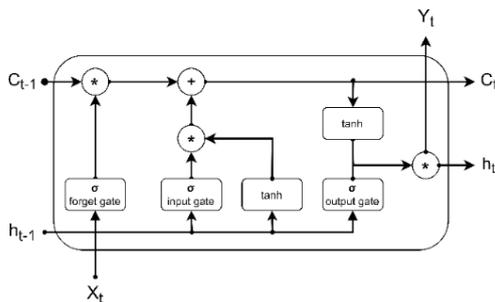
Metode penelitian yang digunakan dilakukan dalam beberapa tahapan. Tahapan yang dilakukan adalah studi literatur, pengumpulan data curah hujan, melakukan preprosesing data, pengolahan data dengan SPI, pelatihan dan pengujian data, penerapan prediksi dengan model LSTM, melakukan validasi data, dan melakukan analisis performa model. Tahapan tersebut digambarkan seperti pada Gambar 1 berikut.



Gambar 1 Tahapan Penelitian

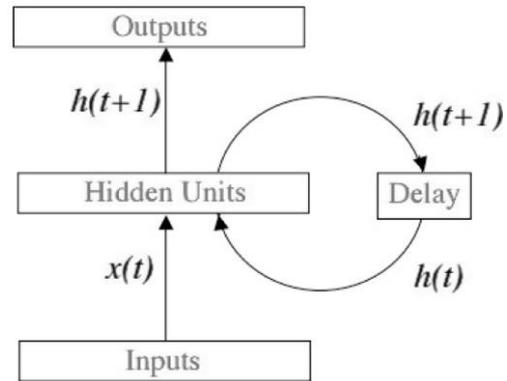
2.1. Studi Literatur

Jaringan saraf berulang (*Recurrent Neural Network/ RNN*) umumnya digunakan untuk prediksi deret waktu. Sedangkan *Long Short-Term Memory (LSTM)* merupakan arsitektur khusus RNN yang secara khusus dikembangkan untuk merepresentasikan urutan temporal. LSTM memiliki saling ketergantungan yang luas, sehingga membuatnya lebih presisi daripada RNN tradisional[7]. Selain itu kinerja model RNN-LSTM yang unggul dibandingkan metode tradisional[7]. LSTM merupakan salah satu jenis RNN yang dapat mengingat nilai dari tahap awal untuk penggunaan di masa mendatang. Algoritma LSTM memiliki lapisan tersembunyi yang terdiri dari beberapa operasi matematika yang melibatkan penghitungan pada tiga gerbang utama: **Forget Gate**, **Input Gate**, dan **Output Gate**, serta pengelolaan **Cell State** seperti ditunjukkan pada Gambar 2[16].



Gambar 2. Arsitektur LSTM

Setiap sel memori di LSTM menggunakan 3 komponen unit *gate* yang berusaha menghilangkan disparitas bobot nilai input[16]. Dengan deret waktu masukan $x(t)$ dan deret waktu keluaran $y(t)$, maka $h(t)$ merepresentasikan status sistem dinamis pada waktu t , merangkum semua informasi penting tentang perilaku masa lalunya, yang penting untuk memprediksi perilaku masa depannya seperti yang ditunjukkan pada Gambar 3[7].



Gambar 3 Proses Deret Waktu LSTM

Proses deret waktu yang terjadi pada arsitektur LSTM dapat dilakukan sebagai berikut:

1. Forget gate (f_t), ini menentukan informasi dari cell state sebelumnya (C_{t-1}) yang akan dilupakan atau dipertahankan.

$$f_t = \sigma(W_f \cdot [h_{t-1}, X_t] + b_f) \tag{1}$$
2. Input gate (i_t), ini menentukan informasi baru apa yang akan disimpan di cell state.

$$i_t = \sigma(W_i \cdot [h_{t-1}, X_t] + b_i) \tag{2}$$
3. Candidate Cell Get (C_t), Ini adalah nilai kandidat yang akan ditambahkan ke cell state setelah dipertimbangkan oleh input gate.

$$N_t = \tanh(W_c \cdot [h_{t-1}, X_t] + b_c) \tag{3}$$
4. Setelah forget gate dan input gate diproses, cell state diperbarui:

$$C_t = C_{t-1}f_t + N_t i_t \tag{4}$$
5. Output gate (o_t), ini menentukan hidden state baru dan output

$$O_t = \sigma(W_o \cdot [h_{t-1}, X_t] + b_o) \tag{5}$$
6. Hidden state baru dihitung menggunakan cell state yang diperbarui dan output gate:

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

Prediksi cuaca termasuk kekeringan di Indonesia merupakan bagian tugas utama dari Badan Meteorologi, Klimatologi, Geofisika (BMKG). Akurasi prediksi sangat penting untuk informasi para stakeholder di Jawa Tengah bagian utara terutama bagi petani dan nelayan. Untuk evaluasi kekeringan regional menggunakan *Reconnaissance Drought Index (RDI)* dan *Standardized Precipitation Index (SPI)*[23][8]. penggunaan SPI sebagai parameter utama dalam pemantauan kekeringan oleh BMKG sebagai indeks dalam menganalisis kekeringan yang berupa curah hujan[11][8].

Indeks kekeringan SPI yang dievaluasi berdasarkan catatan curah hujan jangka panjang untuk durasi tertentu, yaitu masing-masing 3 bulan, 6 bulan, 9 bulan, dan 12 bulan. Ketika nilai SPI secara konsisten negatif atau kurang dari (-1) maka dikatakan bahwa kekeringan telah terjadi dan berakhir ketika nilai SPI positif[24]. Situasi di mana nilai SPI di bawah nol dianggap sebagai periode kekeringan, sedangkan situasi ketika nilainya naik di atas nol dianggap sebagai periode di mana kekeringan telah berakhir, yaitu periode hujan. Dalam metode SPI, data presipitasi disusun untuk periode tanpa henti minimal 30 tahun[20]. Tabel 1 berikut kategori tingkat keparahan kekeringan menurut nilai SPI[24].

World Meteorological Organization (WMO) menyatakan bahwa tanda awal terjadinya kekeringan dapat diidentifikasi melalui penurunan intensitas curah hujan di bawah rata-rata normal, yang dikenal sebagai kekeringan meteorologis. Ini dapat dimanfaatkan sebagai sistem peringatan dini untuk mengantisipasi kekeringan yang lebih parah, dengan salah satu metode yaitu SPI. SPI dapat diterapkan dalam rentang waktu dari 1 bulan hingga 72 bulan[25]. Perhitungan indeks SPI didasarkan pada distribusi gamma, yang didefinisikan sebagai fungsi frekuensi atau probabilitas kejadian. Penggunaan indeks ini memungkinkan pemantauan dan prediksi kekeringan yang lebih akurat dan komprehensif. Berikut adalah perhitungan indeks SPI berdasarkan sebaran Gamma :

$$G(x) = \int_0^x g(x)dx = \frac{1}{\beta^\alpha \Gamma(\alpha)} \int_0^x t^{\alpha-1} e^{-\frac{x}{\beta}} dt \tag{7}$$

dengan:

$\alpha > 0$ = Parameter bentuk,

$\beta > 0$ = Parameter skala,

$x > 0$ = Jumlah curah hujan

Selanjutnya mencari nilai α dan β dengan rumus sebagai berikut:

$$\alpha = \frac{1}{4A} \left(1 + \sqrt{1 + \frac{4A}{3}} \right) \tag{8}$$

atau

$$\alpha = \frac{x^2}{s^2} \tag{9}$$

$$\beta = \frac{x}{\alpha}$$

dengan:

n = total data curah hujan

Jika fungsi gamma tidak diketahui untuk $x = 0$, maka nilai $H(x)$:

$$H(x) = q + (1 - q) \cdot G(x) \tag{10}$$

dengan:

q = total curah hujan = 0 (m) / total data (n)

Kemudian probabilitas kumulatif $H(x)$ diubah menjadi standar normal dari variabel acak Z , yang memiliki nilai rata-rata 0 dan variansi 1. Nilai yang diperoleh Z adalah nilai SPI. Perhitungan Z atau SPI untuk $0 < H(x) \leq 0,5$,

$$Z = SPI = -\left(t - \frac{c_0 + c_1 t + c_2 t^2}{1 + d_1 t + d_2 t^2 + d_3 t^3} \right) \tag{11}$$

dengan:

$$t = \sqrt{\ln \ln \left(\frac{1}{(H(x))^2} \right)} \tag{12}$$

perhitungan Z atau SPI untuk $0,5 < H(x) \leq 1,0$

$$Z = SPI = +\left(t - \frac{c_0 + c_1 t + c_2 t^2}{1 + d_1 t + d_2 t^2 + d_3 t^3} \right) \tag{13}$$

dengan:

$$t = \sqrt{\ln \ln \left(\frac{1}{(1-H(x))^2} \right)} \tag{14}$$

Kriteria tingkat kekeringan digunakan untuk menilai klasifikasi kekeringan yang ditunjukkan dalam Tabel 1.

Tabel 1 Klasifikasi Indeks SPI (McKee,1993)

Klasifikasi Kekeringan	Nilai SPI
Sangat-basah	≥ 2
Basah	1,5 - 1,99
Agak-Basah	1 - 1,49
Normal	(-)0,99 - (-) 0,99
Agak-Kering	(-)1 - (-)1,49
Sangat-Kering	(-)1,5 - (-)1,99
Kering	$\leq (-)2$

Metrik MAE dan RMSE dapat digunakan untuk mengukur seberapa baik model prediktif untuk perkiraan nilai SPI dibandingkan dengan nilai SPI aktual. MAE mengukur selisih antara hasil observasi (sebenarnya) dengan hasil model (prediksi). Sedangkan RMSE merupakan teknik statistik untuk menilai keakuratan model prakiraan dengan membandingkan nilai prediksi dengan nilai teramati. RMSE umumnya digunakan untuk menentukan keakuratan data, dengan mempertimbangkan rentang relatifnya, dan untuk menilai tingkat kesamaan antara nilai yang diharapkan dan nilai teramati[7][16]. Nilai lain adalah MAPE untuk mengukur kesalahan absolut sebagai persentase dari nilai aktual. Metrik ini menunjukkan seberapa besar kesalahan rata-rata dari model dalam bentuk persentase, sehingga lebih mudah untuk diinterpretasikan[16]. Sedangkan R^2 berfungsi untuk mengukur tingkat replikasi hasil oleh model. Nilainya berkisar antara $[-\infty, 1]$, di mana untuk prediksi model yang optimal, skor R^2 mendekati 1[18].

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

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

$$MAPE = \frac{1}{n} \sum_{t=1}^n \frac{|y_t - x_t|}{x_t} \times 100. \tag{17}$$

$$R^2 = \frac{\sum_{i=1}^N (y_i - \bar{y})^2 - \sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \tag{18}$$

2.2. Lokasi Penelitian

Lokasi penelitian ini dikhususkan pada wilayah pantai utara Jawa Tengah Indonesia. Data yang dianalisis berupa catatan curah hujan harian yang dikumpulkan selama periode 2010 hingga 2023 dari Unit Pelaksana Teknis (UPT) BMKG Jawa Tengah yang stasiun pencatatannya berada di wilayah Kota Tegal dan Kota Semarang.



Gambar 4 Peta Lokasi Stasiun Pencatatan Curah Hujan BMKG

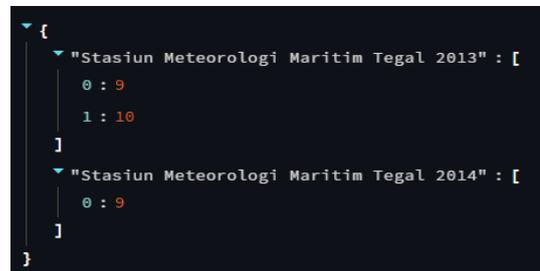
Dua lokasi utama menjadi sumber data penelitian ini. Pertama, Stasiun Klimatologi Jawa Tengah yang terletak di Kota Semarang dengan kode pos hujan 33741301a, yang terletak pada koordinat 6.98470° Lintang Selatan dan 110.318° Bujur Timur, dengan elevasi 6 meter di atas permukaan laut. Kedua, Stasiun Meteorologi Maritim Tegal yang terletak di Kota Tegal dengan kode pos hujan 33760101a, yang berada pada koordinat 6.868° Lintang Selatan dan 109.121° Bujur Timur, dengan ketinggian 1 meter di atas permukaan laut. Pemilihan kedua lokasi ini bertujuan untuk memberikan representasi yang komprehensif tentang kondisi curah hujan di wilayah pantai utara Jawa Tengah.

2.3. Data Penelitian

Pengumpulan data dari dua stasiun pengamatan (Kota Semarang dan Kota Tegal) dengan mengambil pada laman web melalui portal resmi BMKG https://dataonline.bmkg.go.id/data_iklim. Data yang diakses berupa catatan curah hujan harian yang disajikan dalam format bulanan, tersimpan dalam file berekstensi .xlsx. Total file yang diunduh dari kedua stasiun mencapai 336 buah. Seluruh file data tersebut diintegrasikan menjadi satu dataset komprehensif. Dataset gabungan tersebut menjadi sumber utama atau sebagai data primer untuk dilakukan analisis dan pengolahan lebih lanjut dalam penelitian ini. Proses tersebut digunakan peneliti untuk menunjukkan gambaran menyeluruh tentang pola curah hujan di wilayah yang diteliti (pantai utara Jawa Tengah) selama periode waktu 14 tahun (2010 – 2023). Pendekatan ini memastikan bahwa data yang diperoleh berasal dari sumber resmi dan terpercaya, serta memudahkan proses analisis dengan menggabungkan semua informasi ke dalam satu dataset yang terstruktur. Pada penelitian ini proses pengolahan dataset dilakukan menggunakan bahasa pemrograman Python melalui platform Jupyter Notebook.

2.4. Data Preprocessing

Tahap pertama pengolahan melibatkan transformasi data curah hujan dari format harian menjadi bulanan. Setelah konversi, dilakukan identifikasi terhadap data bulanan yang tidak lengkap atau kosong. Guna mengatasi permasalahan data yang hilang atau kosong digunakan metode pengisian dengan nilai rata-rata normal bulanan yang dihitung dari bulan yang sama pada tahun-tahun lainnya. Pendekatan ini memungkinkan peneliti untuk memperoleh dataset yang lebih komprehensif dan konsisten, mengurangi potensi bias dalam analisis akibat data yang hilang, serta mempertahankan integritas temporal dari data curah hujan yang dikumpulkan. Data kosong pada proses identifikasi terlihat pada Gambar 5.



Gambar 5 Hasil pengecekan dataset kosong

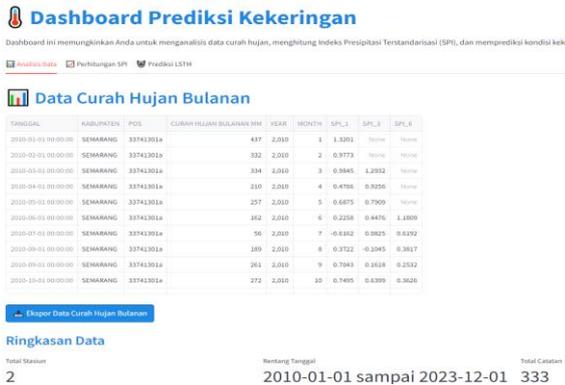
2.5. Penerapan Pada Framework Stramlit

Kemunculan aplikasi web dengan Stramlit semakin umum seiring dengan perkembangan metode komputasi baru, terutama yang melibatkan analisis hasil dari kampanye eksperimen berskala besar. Streamlit merupakan sebuah framework pengkodean Python sumber terbuka, telah menjadi alat yang semakin penting bagi para peneliti untuk membangun aplikasi *web* yang memungkinkan berbagi dataset besar dan sumber daya penelitian lainnya [26]. Kemampuannya memungkinkan cara yang intuitif dan mudah bagi peneliti untuk mendapatkan wawasan ilmiah baru melalui eksplorasi dataset besar dan memiliki fitur visualisasi dasar seperti grafik, diagram pencar, dan juga format tabuler seperti *dataframe* dan tabel.

3. HASIL DAN PEMBAHASAN

Guna mendapatkan akurasi prediksi yang optimal, penelitian ini mengimplementasikan strategi *preprocessing* data yang komprehensif. Pertama, dilakukan identifikasi nilai-nilai yang hilang melalui proses *scanning* dataset. Selanjutnya, metode pengisian data berbasis rata-rata (*mean*) normal bulanan yang berseduaian diterapkan untuk mengisi nilai-nilai yang kosong tersebut. Dataset yang telah dilakukan pengisian kemudian ditransformasi menjadi SPI. Penelitian ini berfokus pada analisis SPI untuk periode 1, 3, dan 6 bulanan. Pemilihan periode ini didasarkan pada pertimbangan relevansi terhadap

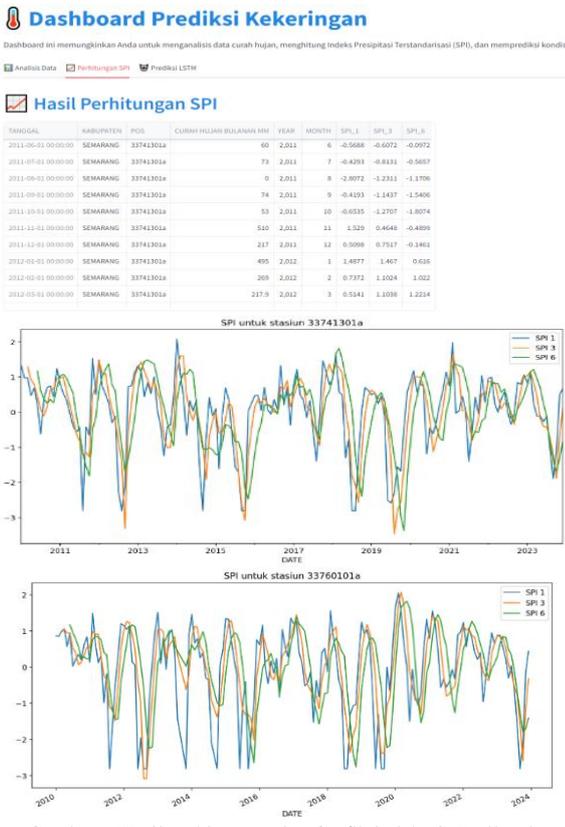
tujuan penelitian dan karakteristik iklim wilayah studi.



Gambar 6. Preprocessing Data Curah Hujan Bulanan.

3.1. Penentuan Indeks SPI

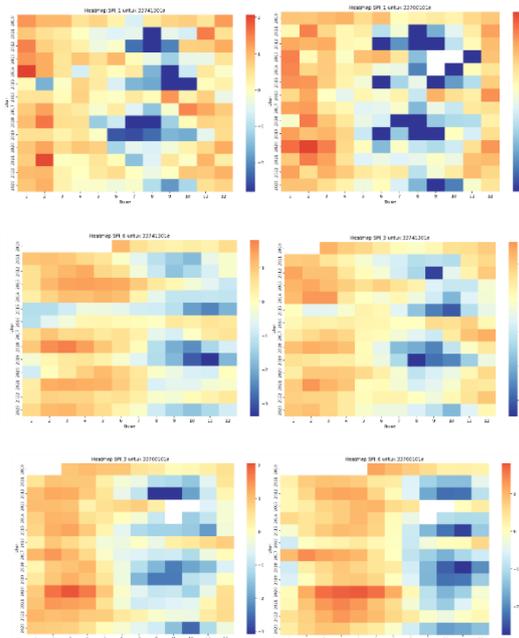
Pendekatan indeks SPI dari dataset ini tidak hanya meningkatkan integritas dataset, tetapi juga memungkinkan analisis yang lebih mendalam terhadap pola dan tren curah hujan, yang krusial untuk pemahaman dan prediksi fenomena kekeringan. Perhitungan SPI digunakan metode distribusi *gamma* dan transformasi ke distribusi normal standar. SPI dihitung untuk setiap stasiun dan untuk tiga interval waktu yang berbeda (1, 3, dan 6 bulan). Nilai-nilai tak terhingga yang mungkin muncul dalam perhitungan SPI diganti dengan nilai minimum atau maksimum yang valid.



Gambar 6 Hasil perhitungan dan Grafik indeks SPI wilayah Semarang dan Tegal

Setelah data hujan diolah menjadi data indeks SPI dalam periode 1,3,6 bulanan, selanjutnya data ini dijadikan sumber untuk melakukan proses perhitungan prediksi menggunakan LSTM. Untuk mengoptimalkan kinerja model, dilakukan proses *hyperparameter tuning* menggunakan *Keras Tuner* dengan metode *Random Search*, sebuah pendekatan yang memungkinkan eksplorasi efisien dari ruang *hyperparameter* yang luas.

Pada *heatmap* SPI untuk wilayah semarang dan Tegal selama periode 2010-2023 mengungkapkan pola temporal yang kompleks dalam variabilitas curah hujan. Visualisasi SPI-1, SPI-3, dan SPI-6 menunjukkan gradasi yang signifikan dalam intensitas dan durasi episode kering dan basah. SPI-1 memperlihatkan fluktuasi yang lebih tinggi, mencerminkan variabilitas jangka pendek curah hujan bulanan. Sedangkan SPI-3 dan SPI-6 menampilkan pola yang lebih koheren dan persisten, mengindikasikan tren curah hujan jangka menengah hingga panjang. Terlihat adanya periode kering yang signifikan (ditandai dengan warna biru tua) yang cenderung terjadi pada pertengahan hingga akhir tahun, terutama pada SPI-6, menunjukkan kecenderungan kekeringan musiman. Sebaliknya, periode basah (ditandai dengan warna merah) lebih sering teramati pada awal tahun. Perbedaan pola antara kedua stasiun juga terlihat, mengindikasikan variabilitas spasial dalam distribusi curah hujan. Analisis ini menyoroti kompleksitas pola curah hujan di wilayah tersebut dan pentingnya mempertimbangkan berbagai skala waktu dalam evaluasi kondisi kekeringan. Gambar 7 menunjuk plot *heatmap* SPI dari dua wilayah data yang diolah.



Gambar 7 Heat Map SPI Wilayah Semarang dan Tegal.

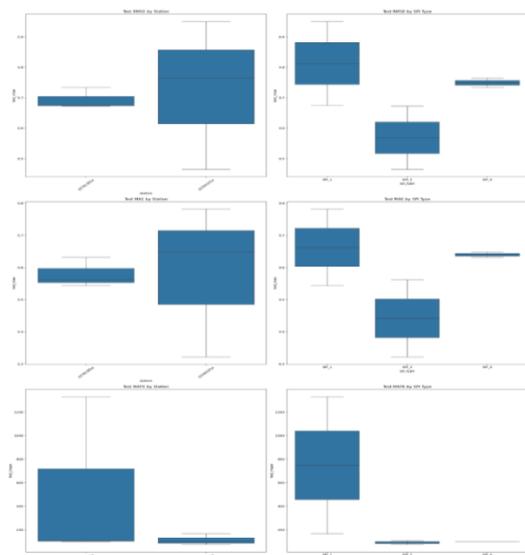
3.2. Menghitung Metrik Kinerja

Model kemudian dilatih dan dievaluasi secara sistematis untuk setiap kombinasi stasiun pengamatan dan tipe SPI, memastikan *robustness* dan generalisasi model *across* berbagai kondisi geografis dan temporal. Evaluasi komprehensif dilakukan dengan mengkalkulasi beragam metrik kinerja, meliputi RMSE, MAE, R², MAPE, baik untuk data *training* maupun *testing*. Hasil evaluasi diorganisir dalam format tabel yang terstruktur, menyediakan panorama kuantitatif dari metrik-metrik kinerja kunci seperti terlihat pada Tabel 2.

Tabel 2 Hasil Uji Performa Model LSTM

SEMARANG					
	Metrik	RMSE	MAE	R ²	MAPE
SPI-1	Train	0,844	0,676	0,444	210,849
	Test	0,687	0,606	0,109	1865,686
SPI-3	Train	1,142	0,956	-0,098	405,190
	Test	0,738	0,615	-0,122	181,085
SPI-6	Train	0,994	0,846	0,135	473,963
	Test	0,704	0,593	0,163	98,418
TEGAL					
	Metrik	RMSE	MAE	R ²	MAPE
SPI-1	Train	1,238	0,949	0,033	372,430
	Test	0,872	0,638	0,104	130,529
SPI-3	Train	0,504	0,397	0,780	67,200
	Test	0,468	0,373	0,697	69,127
SPI-6	Train	0,478	0,404	0,797	150,740
	Test	0,396	0,356	0,782	138,052

Sebagai komplemen, visualisasi boxplot diimplementasikan untuk memfasilitasi analisis komparatif terhadap performa model di antara berbagai stasiun pengamatan dan tipe SPI. Pendekatan ini memungkinkan identifikasi pola dan variabilitas dalam akurasi prediktif model *across* spektrum kondisi geografis dan temporal yang berbeda, tersaji pada Gambar 8.

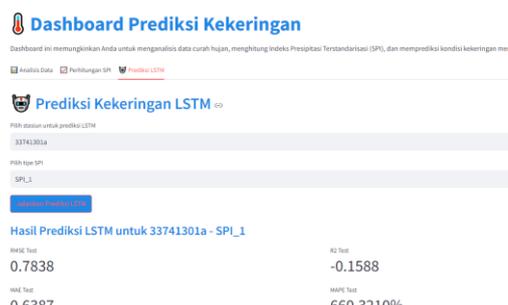


Gambar 8 Hasil uji performa model LSTM variasi silang berdasarkan stasiun pengamat dan tipe SPI

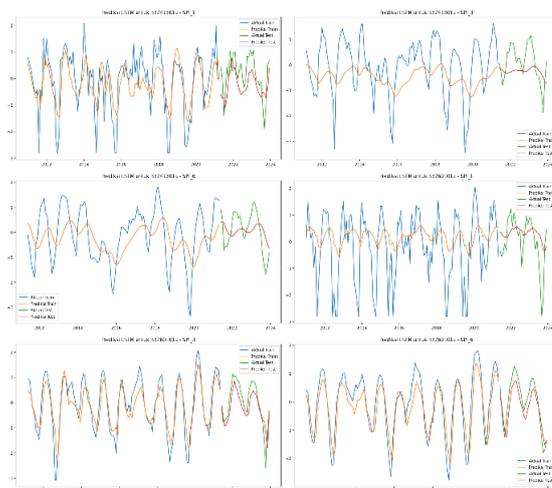
Fitur ini juga memastikan bahwa model yang disimpan adalah versi dengan performa terbaik.

Pendekatan ini mencerminkan kompleksitas dan ketelitian dalam proses pemodelan data SPI, menggabungkan teknik pengolahan data lanjutan dengan strategi pelatihan model yang sophisticated untuk menghasilkan prediksi yang akurat dan reliable.

Setelah mengetahui performa kinerja model LSTM dalam membuat prediksi tingkat kekeringan di wilayah Pantura Jawa Tengah, prediksi kekeringan berdasarkan indeks SPI untuk periode waktu 1, 3 dan 6 bulanan tersaji dalam bentuk visual dashboard seperti terlihat pada Gambar 9 dan disajikan dalam bentuk grafik seperti Gambar 10.



Gambar 9 Prediksi SPI menggunakan model LSTM



Gambar 10 Prediksi SPI menggunakan model LSTM

Hasil uji performa model LSTM seperti pada Tabel 2, dapat di jelaskan sebagai berikut :

- Wilayah Semarang : Kinerja LSTM kurang stabil, terutama pada SPI-1 dan SPI-3. Untuk itu model perlu dioptimalkan, terutama untuk SPI-1 dan SPI-3. Sedangkan SPI-6 memiliki kinerja terbaik di antara tipe SPI lainnya, dengan R²=0.163 dan MAPE terendah (98.418). Namun, performa keseluruhan masih jauh dari ideal.
- Wilayah Tegal: Kinerja LSTM performanya sangat baik pada SPI-3 dan SPI-6 dengan R² tinggi serta MAPE yang rendah. Sedangkan SPI-1 menunjukkan performa lebih lemah dibandingkan SPI-3 dan SPI-6. SPI-3 memberikan performa terbaik untuk Tegal dengan R²=0.780 pada training, dan R²=0.697

pada testing, serta MAPE terendah di antara semua tipe (69.127 pada testing).

Sementara itu prediksi model LSTM seperti Gambar 10 dapat dijelaskan sebagai berikut:

- a. Performa pada data pelatihan menunjukkan model mampu menangkap pola dengan baik, hal ini terlihat dari garis biru dan oranye yang cukup selaras.
- b. Performa pada data pengujian menunjukkan bahwa pada SPI-3 dan SPI-6 umumnya memiliki prediksi yang lebih baik dibandingkan SPI-1, baik untuk stasiun pertama maupun kedua. Model cenderung kesulitan menangkap fluktuasi tajam pada data pengujian (terlihat dari deviasi antara garis merah dan hijau).

4. DISKUSI

Berdasarkan hasil tersebut, maka model LSTM dapat digunakan dan bahkan direkomendasikan dalam prediksi cuaca oleh BMKG untuk menghasilkan prediksi yang lebih akurat. Hal ini dikarenakan hasil dari analisis dan pengujian di atas. Namun demikian diperlukan beberapa catatan guna diskusi lebih lanjut yaitu:

1. Data penelitian masih perlu ditinjau kembali tingkat kelengkapan data harian curah hujan harian. Hal ini dikarenakan imputansi data menggunakan instrumen yang berbeda dan tergantung kontinuitas isi data harian. Perlu adanya model imputansi yang akurat untuk mengatasi data harian yang kosong atau tidak terisi.
2. Rentang waktu dari data yang digunakan penelitian dapat dilakukan lebih panjang seperti dua puluh tahunan dan bahkan lebih untuk membuktikan tingkat kesetabilan model yang diuji dalam rentang waktu 14 tahun (2010-2023) ini. Sehingga hasil matrik kerja seperti pada Tabel 2, dapat membuktikan kebenarannya.
3. Pemodelan LSTM dan SPI dalam penelitian ini menggunakan satu variabel curah hujan, sehingga dengan penerapan model SPEI dapat digunakan untuk menambah variabel lain yang mempengaruhi cuaca seperti suhu (evaporasi), hal ini akan lebih dapat digunakan untuk memprediksi anomali cuaca.

5. KESIMPULAN

Berdasarkan hasil analisis pemodelan LSTM untuk prediksi indeks SPI di wilayah Semarang dan Tegal telah terungkap pola performa yang kompleks dan bervariasi. Di wilayah Semarang, model menunjukkan konsistensi relatif dalam RMSE untuk semua skala waktu SPI pada data uji, namun dengan indikasi overfitting pada SPI-1 dan SPI-3. Sebaliknya, di wilayah Tegal memperlihatkan peningkatan performa yang signifikan untuk skala waktu yang lebih panjang, dengan SPI-6 menunjukkan hasil terbaik (RMSE 0,396, R^2 0,782

pada data uji). Fenomena ini menggaris bawahi pentingnya faktor geografis dan temporal dalam pemodelan curah hujan. Perbandingan antar lokasi mengungkapkan bahwa model LSTM secara konsisten menunjukkan performa superior di Tegal, terutama untuk SPI-3 dan SPI-6, dengan tingkat generalisasi yang lebih baik antara data latih dan uji. Hal ini kontras dengan wilayah Semarang, di mana nilai R^2 yang rendah bahkan negatif untuk SPI-3 mengindikasikan keterbatasan kemampuan prediktif model. Menariknya, SPI-6 menunjukkan akurasi tertinggi di kedua lokasi berdasarkan metrik RMSE dan MAE, sementara SPI-1 konsisten menampilkan performa terlemah, menggambarkan kompleksitas dalam memprediksi variabilitas curah hujan jangka pendek. Meskipun demikian, nilai MAPE yang tinggi, terutama di Semarang, mengindikasikan adanya kesalahan persentase yang substansial dalam prediksi, menunjukkan area potensial untuk perbaikan model. Temuan-temuan ini menyoroti efektivitas model LSTM dalam memprediksi SPI jangka panjang dibandingkan jangka pendek, serta signifikansi konteks geografis dalam performa model. Model LSTM dapat dengan baik diimplementasikan pada framework Streamline, terbukti dengan tampilnya visualisasi yang baik pada proses dan evaluasi model. Perlu uji coba jangka panjang untuk mengetahui performa sistem yang dibangun terkait fleksibilitas dan keandalan *user interface*.

Penelitian berikutnya dapat dilakukan dengan membuat perbandingan menggunakan model yang sama datanya pada sumber stasiun BMKG yang sama namun dengan rentang data antara dua puluhan dan tiga puluhan tahun. Selanjutnya pada data yang sama tersebut dibandingkan dengan model prediksi yang lainnya termasuk model *Standardized Precipitation Evapotranspiration Index* (SPEI) dengan input beberapa variabel.

DAFTAR PUSTAKA

- [1] D. Auliyani and N. Wahyuningrum, "Rainfall variability based on the Climate Hazards Group InfraRed Precipitation with Station Data (CHIRPS) in Lesti watershed, Java Island, Indonesia," *IOP Conf. Ser. Earth Environ. Sci.*, vol. 874, no. 1, 2021, doi: 10.1088/1755-1315/874/1/012003.
- [2] M. I. Supriyatna, A. A. N. S. D. Negara, and N. F. Margini, "Prediksi dan Pemetaan Kekeringan Menggunakan Metode Thomas Fiering dan Standardized Precipitation Indexs (SPI) di Kabupaten Tuban," *J. Apl. Tek. Sipil*, vol. 21, no. 3, p. 229, 2023, doi: 10.12962/j2579-891x.v21i3.17239.
- [3] N. A. Prakoso Indaryono, "Analisa Perbandingan Algoritma Random Forest Dan Naive Bayes Untuk Klasifikasi Curah Hujan Berdasarkan Iklim Di Indonesia," *JIFI*

- (*Jurnal Ilm. Penelit. dan Pembelajaran Inform.*, vol. 9, no. 1, pp. 158–167, 2024, doi: 10.29100/jipi.v9i1.4421.
- [4] S. Koesuma, D. A. Kusuma, and I. W. Harmoko, “Mitigation strategic of drought in central Java Indonesia during Covid-19 pandemic,” *J. Phys. Conf. Ser.*, vol. 1825, no. 1, 2021, doi: 10.1088/1742-6596/1825/1/012024.
- [5] G. Prayitno, G. N. Kaunang, and A. Amiruddin Natsir, “Mapping Rainfall [Isohyet] in Areas With a Shortage of Rain Stations Using Manual Station Technique and Precipitation Merra-2 Data Analysis,” *J. Tek. Inform.*, vol. 4, no. 4, pp. 819–830, 2023, doi: 10.52436/1.jutif.2023.4.4.943.
- [6] D. G. C. Kirono, N. J. Tapper, and J. L. McBride, “Documenting Indonesian rainfall in the 1997/1998 EL Niño event,” *Phys. Geogr.*, vol. 20, no. 5, pp. 422–435, 1999, doi: 10.1080/02723646.1999.10642687.
- [7] W. Shah, J. Chen, I. Ullah, M. H. Shah, and I. Ullah, “Application of RNN-LSTM in Predicting Drought Patterns in Pakistan: A Pathway to Sustainable Water Resource Management,” *Water (Switzerland)*, vol. 16, no. 11, pp. 1–19, 2024, doi: 10.3390/w16111492.
- [8] O. M. Baez-Villanueva *et al.*, “On the timescale of drought indices for monitoring streamflow drought considering catchment hydrological regimes,” *Hydrol. Earth Syst. Sci.*, vol. 28, no. 6, pp. 1415–1439, 2024, doi: 10.5194/hess-28-1415-2024.
- [9] D. Auliyani, N. Wahyuningrum, A. Budi Supangat, and T. Mutiara Basuki, “Prediction of drought vulnerability and its mitigation measures in the Northern Coast of Java, Indonesia,” *IOP Conf. Ser. Earth Environ. Sci.*, vol. 1016, no. 1, 2022, doi: 10.1088/1755-1315/1016/1/012003.
- [10] G. G. Haile, Q. Tang, W. Li, X. Liu, and X. Zhang, “Drought: Progress in broadening its understanding,” *Wiley Interdiscip. Rev. Water*, vol. 7, no. 2, 2020, doi: 10.1002/WAT2.1407.
- [11] Suroso, D. Nadhilah, Ardiansyah, and E. Aldrian, “Drought detection in Java island based on standardized precipitation and evapotranspiration index (SPEI),” *J. Water Clim. Chang.*, vol. 12, no. 6, pp. 2734–2752, 2021, doi: 10.2166/wcc.2021.022.
- [12] Renggo Mike Al’Aziz; Muhammad Rifai, “Rancang Bangun Purwa Rupa Sistem Peringatan Dini Curah Hujan Tinggi Dan Banjir Berbasis Aplikasi Blynk 2.0,” *J. Teknol.*, vol. 2, no. 1, pp. 19–25, 2013, [Online]. Available: www.jurnalteknologi.utm.my
- [13] H. N. Irmanda, Ermatita, M. K. Awang, and M. Adrezo, “Enhancing Weather Prediction Models through the Application of Random Forest Method and Chi-Square Feature Selection,” *Int. J. Informatics Vis.*, vol. 8, no. 3–2, pp. 1506–1514, 2024, doi: 10.62527/joiv.8.3.2356.
- [14] S. Siswanto *et al.*, “Satellite-based meteorological drought indicator to support food security in Java Island.” *PLOS ONE*, pp. 1–20, 2021.
- [15] M. Taufik, B. I. Setiawan, and H. A. J. Van Lanen, “Increased fire hazard in human-modified wetlands in Southeast Asia,” *Ambio*, vol. 48, no. 4, pp. 363–373, 2019, doi: 10.1007/s13280-018-1082-3.
- [16] C. S. Alfredo and D. Adytia, “Time Series Forecasting of Significant Wave Height using GRU, CNN-GRU, and LSTM,” *Resti*, vol. 6, no. 5, pp. 776–781, 2022.
- [17] F. Rasyid and D. A. Adytia, “Time Series Temperature Forecasting by using ConvLSTM Approach, Case Study in Jakarta,” *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 6, no. 4, pp. 563–569, 2022, doi: 10.29207/resti.v6i4.4162.
- [18] Namora and Jan Everhard Riwurohi, “Prediction of Water Levels on Peatland using Deep Learning,” *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 6, no. 2, pp. 234–239, 2022, doi: 10.29207/resti.v6i2.3919.
- [19] E. D. Taylan, “An Approach for Future Droughts in Northwest Türkiye: SPI and LSTM Methods,” *Sustainability*, vol. 16, no. 16, p. 6905, 2024, doi: 10.3390/su16166905.
- [20] T. McKee, N. Doesken, J. K.-P. of the 8th, and U. 1993, “The relationship of drought frequency and duration to time scales Drought in Africa,” *Climate.Colostate.Edu*, no. January, pp. 17–22, 1993, [Online]. Available: <https://climate.colostate.edu/pdfs/relationshipofdroughtfrequency.pdf>
- [21] N. A. A. B. S. Bahari *et al.*, “Predicting Sea Level Rise Using Artificial Intelligence: A Review,” *Arch. Comput. Methods Eng.*, vol. 30, no. 7, pp. 4045–4062, 2023, doi: 10.1007/s11831-023-09934-9.
- [22] S. S. Baek, J. Pyo, and J. A. Chun, “Prediction of water level and water quality using a cnn-lstm combined deep learning approach,” *Water (Switzerland)*, vol. 12, no. 12, 2020, doi: 10.3390/w12123399.
- [23] R. Rezaei and A. Shabri, “Improving Drought Prediction Accuracy: A Hybrid EEMD and Support Vector Machine

Approach with Standardized Precipitation Index,” *Water Resour. Manag.*, 2024, doi: 10.1007/s11269-024-03912-x.

- [24] D. J. Mehta and S. M. Yadav, “Meteorological drought analysis in Pali District of Rajasthan State using standard precipitation index,” *Int. J. Hydrol. Sci. Technol.*, vol. 15, no. 1, pp. 1–10, 2023, doi: 10.1504/IJHST.2021.10037751.
- [25] WMO-No.1090, *Standardized Precipitation Index User Guide*, vol., no. 2012.
- [26] J. M. Nápoles-Duarte, A. Biswas, M. I. Parker, J. P. Palomares-Baez, M. A. Chávez-Rojo, and L. M. Rodríguez-Valdez, “Stmol: A component for building interactive molecular visualizations within streamlit web-applications,” *Front. Mol. Biosci.*, vol. 9, no. September, pp. 1–10, 2022, doi: 10.3389/fmolb.2022.990846.