



Optimalisasi Prediksi Parameter Lingkungan Menggunakan Model LSTM Multivariat dan Univariat

Chandra Nilasari Yunantara^{1,*}, April Firman Daru²

Teknologi Informasi dan Komunikasi, Teknik Informatika, Universitas Semarang, Semarang, Indonesia

Email: ^{1,*}chndrnila@gmail.com, ²firman@usm.ac.id

Email Penulis Korespondensi: chndrnila@gmail.com

Abstrak—Prediksi parameter lingkungan berperan penting dalam mendukung pemantauan kondisi cuaca serta pengambilan keputusan berbasis data, khususnya di wilayah perkotaan seperti Semarang. Namun, akurasi prediksi sering kali dipengaruhi oleh keterbatasan model dalam menangkap hubungan antarparameter lingkungan. Penelitian ini bertujuan untuk menganalisis dan membandingkan kinerja dua pendekatan Long Short-Term Memory (LSTM), yaitu Multivariat dan Univariat, dalam memprediksi suhu udara (temperature) sebagai variabel dependen. Pada model Multivariat, prediksi suhu dipengaruhi oleh variabel independen lain seperti kelembaban (humidity), tekanan udara (pressure), dan ketinggian (altitude), sedangkan pada model Univariat, prediksi suhu dilakukan berdasarkan data suhu masa lalu. Arsitektur model terdiri atas tiga lapisan utama yaitu, input layer, dua hidden layer, dan output layer. Evaluasi kinerja dilakukan menggunakan Root mean Square Error (RMSE), Mean Square Error (MSE), dan Mean Absolute percentage Error (MAPE). Hasil pengujian menunjukkan bahwa model LSTM Multivariat menghasilkan nilai error yang lebih rendah untuk parameter temperature dan pressure, sedangkan model LSTM Univariat lebih unggul pada humidity dan altitude. Dengan demikian, model Multivariat lebih sesuai digunakan ketika hubungan antarparameter lingkungan berpengaruh signifikan terhadap hasil prediksi.

Kata Kunci: LSTM; Multivariat; Univariat; Prediksi

Abstract—Environmental parameter prediction plays an essential role in supporting weather monitoring and data-driven decision-making, particularly in urban areas. However, prediction accuracy is often limited by a model's ability to capture the interrelationships among environmental parameters. This study aims to analyze and compare the performance of two Long Short-Term Memory (LSTM) approaches Multivariate and Univariate in predicting air temperature as the dependent variable. In the Multivariate model, temperature prediction is influenced by other independent variables such as humidity, pressure, and altitude, whereas in the Univariate model, temperature prediction is based solely on its historical data. The model architecture consists of three main layers an input layer, two hidden layers, and an output layer. Model performance was evaluated using Root Mean Square Error (RMSE), Mean Square Error (MSE), and Mean Absolute Percentage Error (MAPE). The experimental results show that the multivariate LSTM model produces lower error values for temperature and pressure parameters, while the univariate LSTM model performs better for humidity and altitude. Therefore, the multivariate model is more suitable when the interrelationships among environmental parameters significantly influence prediction outcomes.

Keywords: LSTM; Multivariate; Univariate; Prediction

1. PENDAHULUAN

Fenomena perubahan iklim dan dinamika lingkungan perkotaan membawa dampak besar terhadap aktivitas manusia, terutama dikota-kota besar seperti Semarang. Kota ini dikenal sebagai salah satu wilayah dengan tingkat urbanisasi tinggi dan kerentanan terhadap perubahan cuaca ekstrem. Faktor-faktor lingkungan seperti suhu udara, kelembaban, tekanan udara, dan ketinggian memiliki pengaruh langsung terhadap berbagai aspek kehidupan, mulai dari kesehatan masyarakat, perencanaan transportasi, hingga tata kelola kota. Cuaca yang panas dan lembab misalnya, dapat meningkatkan resiko penyakit berbasis lingkungan, sedangkan perubahan tekanan udara dapat memengaruhi keselamatan penerbangan. Oleh karena itu, prediksi cuaca mikro menjadi salah satu elemen penting dalam mendukung tata kelola perkotaan yang berkelanjutan. Namun, kondisi cuaca yang fluktuatif dan penuh ketidakpastian seringkali menjadi hambatan dalam melakukan prediksi yang akurat. Metode prediksi konvensional, seperti *Autoregressive Integrated Moving Average (ARIMA)* dan *Seasonal Autoregressive Integrated Moving Average (SARIMA)* masih memiliki keterbatasan dalam menangkap pola data non-linear dan kompleks. Model-model tersebut umumnya lebih sesuai digunakan untuk data dengan tren linear dan variasi musiman yang sederhana. Dalam praktiknya, variabel lingkungan di kawasan urban sering kali dipengaruhi oleh banyak faktor eksternal yang saling berinteraksi, sehingga menghasilkan pola data yang kompleks dan dinamis. Ketika metode prediksi tidak dapat menangkap keterkaitan antarvariabel tersebut, maka hasil yang diperoleh berpotensi memiliki akurasi rendah. Akibatnya, kesalahan dalam prediksi dapat menimbulkan risiko serius pada berbagai sektor, misalnya pada mitigasi bencana hidrometeorologi, perencanaan energi, sistem transportasi, perkembangan teknologi, hingga pengambilan kebijakan publik berbasis data.

Sebagai jawaban atas keterbatasan tersebut, perkembangan teknologi kecerdasan buatan menghadirkan pendekatan baru untuk pemodelan prediksi. Salah satu metode yang relevan adalah *Long Short-Term Memory (LSTM)*, yaitu model jaringan saraf tiruan yang termasuk dalam kategori *Recurrent Neural Network (RNN)*. LSTM dirancang untuk mengatasi kelemahan RNN konvensional, terutama masalah *vanishing gradient*.

Kemampuan LSTM untuk mengingat informasi jangka panjang menjadikannya sangat sesuai digunakan dalam pemodelan deret waktu yang memiliki pola non-linear dan kompleks. Dalam konteks penelitian ini, LSTM digunakan dengan dua pendekatan yaitu univariat yang berfokus pada satu variabel target, serta multivariat yang memanfaatkan beberapa variabel input sekaligus. Perbandingan kedua pendekatan tersebut menjadi penting karena keduanya memiliki keunggulan masing-masing, dan hasil penelitian dapat menjadi dasar pemilihan model terbaik sesuai kebutuhan.

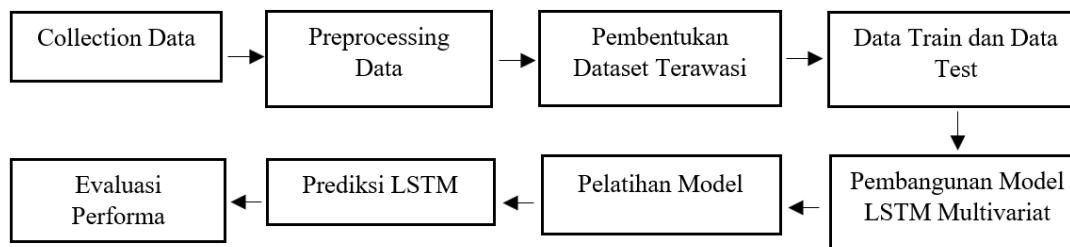


Penelitian-penelitian sebelumnya menunjukkan efektivitas LSTM pada berbagai bidang. Menerapkan LSTM univariat dan multivariat untuk memprediksi kasus aktif COVID-19 di Malaysia dengan integrasi data meteorologi. Hasil penelitian mereka menunjukkan bahwa univariat lebih unggul di sebagian wilayah, sedangkan multivariat lebih sesuai untuk wilayah ini. Hal ini mengindikasikan bahwa pemilihan model sangat bergantung pada karakteristik data yang digunakan [1]. Meneliti LSTM univariat dan multivariat untuk memprediksi *significat wave height* di sektor kelautan. Mereka memanfaatkan variabel meteorologi seperti kecepatan dan arah angin, tekanan udara, serta suhu permukaan laut. Hasilnya menunjukkan bahwa model multivariat mampu meningkatkan performa prediksi dibandingkan metode tradisional, sehingga lebih sesuai digunakan dalam kondisi dengan variabel eksternal yang kompleks [2]. Dalam bidang energi membandingkan efektivitas LSTM univariat dan multivariat pada peramalan daya energi surya fotovoltaik. Mereka menemukan bahwa model univariat cukup efektif untuk peramalan jangka pendek (*one-step forecasting*), sementara model multivariat lebih unggul untuk peramalan jangka panjang (*multistep forecasting*), karena mampu mempertimbangkan variabel seperti radiasi surya dan suhu [3]. Membandingkan pendekatan serupa pada prediksi pasar saham Malaysia. Mereka menemukan bahwa LSTM mengungguli metode lain seperti MLP dan CNN, baik dalam horizon jangka pendek maupun jangka panjang. Temuan ini memperkuat pandangan bahwa LSTM dapat menangani data non-linear yang kompleks, mirip dengan pola data lingkungan [4]. Penelitian terbaru memfokuskan pada prediksi harga minyak kelapa sawit dunia. Mereka menemukan bahwa LSTM univariat lebih efektif ketika menghadapi data dengan volatilitas tinggi, namun, multivariat tetap memiliki keunggulan karena dapat mengintegrasikan variabel eksternal seperti harga minyak kedelai dan minyak bumi, temuan ini memperlihatkan bahwa keunggulan kedua model bersifat kontekstual bergantung pada jenis data yang digunakan [5].

Berdasarkan uraian diatas, terlihat jelas bahwa univariat maupun multivariat memiliki keunggulan sekaligus keterbatasan. Univariat cenderung lebih efektif ketika pola historis cukup kuat dan stabil, sedangkan multivariat lebih sesuai digunakan ketika terdapat variabel eksternal yang berpengaruh signifikan. Meski demikian, sebagian besar penelitian terdahulu berfokus pada kesehatan, kelautan, energi, atau pasar keuangan. Masih sedikit penelitian yang membandingkan kedua pendekatan ini secara langsung dalam konteks prediksi parameter lingkungan perkotaan. Kota Semarang, dengan iklim tropis lembab serta tingkat urbanisasi yang tinggi, menjadi kasus menarik sekaligus menantang untuk menguji efektivitas kedua pendekatan ini. Kesenjangan penelitian (*gap analysis*) muncul dari kenyataan bahwa sebagian besar studi terdahulu tidak menyoroti parameter lingkungan perkotaan secara komprehensif. Sebagian besar penelitian lebih menekankan pada variabel tunggal dan dominan tertentu, sementara integrasi beberapa parameter lingkungan di kawasan urban masih jarang dilakukan. Dengan menempatkan penelitian pada konteks tersebut, studi ini tidak hanya menguji kinerja LSTM, tetapi juga berupaya memberikan kontribusi praktis dalam mendukung sistem peringatan dini, perencanaan tata kota, dan pengambilan keputusan berbasis data. Hasil penelitian ini diharapkan dapat memberikan rekomendasi yang relevan, baik bagi pemerintah daerah, lembaga penelitian, maupun masyarakat umum. Dengan demikian, penelitian ini bertujuan untuk menganalisis kinerja LSTM multivariat dan univariat dalam memprediksi parameter lingkungan di wilayah perkotaan. Secara khusus, penelitian ini berusaha mengidentifikasi parameter lingkungan mana yang lebih akurat diprediksi dengan model tertentu, serta menyusun rekomendasi model prediksi yang relevan dengan kebutuhan aplikasi. Harapannya, penelitian ini tidak hanya memberikan kontribusi teoritis bagi pengembangan *metode time series forecasting*, tetapi juga manfaat praktis dalam mendukung perencanaan dan kebijakan berbasis data dikota-kota besar seperti Semarang.

2. METODOLOGI PENELITIAN

Dalam penelitian ini, proses pembuatan model LSTM multivariat dan univariat dilakukan dalam beberapa tahap, seperti yang digambarkan pada Gambar 1.



Gambar 1. Metode Penelitian LSTM multivariat dan univariat

Berikut ini penjelasan singkat setiap fase dalam pendekatan studi untuk model LSTM multivariat dan univariat, diambil dari Gambar 1.

a. *Collection data*

Pada tahap pengumpulan data dalam penelitian ini dilakukan untuk memperoleh informasi yang relevan dengan fokus penelitian. Data diperoleh melalui meteorologi dan lingkungan yang menyediakan catatan parameter iklim harian di Kota Semarang. Proses pengumpulan dilakukan dengan cara mengunduh dataset historis melalui permintaan resmi ke



lembaga terkait yang digunakan terdiri dari parameter (temperature, humidity, pressure, dan altitude) dijadikan sebagai fitur input.

b. *Preprocessing data*

Pada tahapan ini dilakukan pembersihan data dan dinormalisasi. Karena terdapat lebih dari satu variabel, dilakukan standarisasi agar setiap variabel memiliki skala yang seimbang, serta penanganan *missing values*.

c. Pembentukan dataset terawasi

Data runtun waktu diubah menjadi format *supervised learning*, di mana nilai data pada periode sebelumnya digunakan untuk memprediksi periode selanjutnya.

d. Data latih dan uji

Pada tahapan ini dibagi menjadi dua yaitu, data latih (80%) dan data uji (20%) untuk menghindari *overfitting* serta memastikan model dapat diuji secara objektif.

e. Pembangunan model LSTM multivariat dan univariat

Model LSTM dibangun menggunakan arsitektur berlapis untuk menangkap pola temporal pada data deret waktu lingkungan. Model LSTM multivariat menggunakan empat variabel input (temperature, humidity, pressure, dan altitude) dengan time step 7 hari untuk memprediksi suhu udara keesokan harinya. Arsitektur terdiri atas dua hidden layer LSTM masing-masing berisi 64 dan 32 neuron, serta output layer tunggal dengan aktivasi linear. Sedangkan model LSTM univariat hanya menggunakan data suhu masa lalu dengan konfigurasi dua hidden layer berisi 50 dan 25 neuron, dan time step yang sama yaitu 7 hari.

f. Pelatihan model

Pada tahapan ini proses pelatihan dilakukan dengan optimisasi menggunakan Adam dengan learning rate 0.001 dan fungsi kehilangan berupa MSE. Model dilatih selama 100 *epoch* dengan batch size 32 untuk mempercepat konvergensi dan menjaga stabilitas pelatihan. Untuk mencegah *overfitting*, diterapkan teknik *early stopping* dengan patience selama 10 *epoch*, serta penggunaan dropout layer sebesar 0.2 pada masing-masing hidden layer. Karena kompleksitas hubungan antarvariabel pada model multivariat lebih tinggi dibandingkan univariat, dilakukan penyesuaian jumlah neuron dan regularisasi dropout agar model tidak terlalu menyesuaikan diri terhadap data pelatihan.

g. Hasil prediksi

Kemudian pada tahapan ini dilakukan denormalisasi agar kembali ke skala semula, sehingga dapat dibandingkan secara langsung dengan data aktual.

h. Evaluasi performa

Pada kinerja model dievaluasi menggunakan metrik RMSE, MSE, dan MAPE. Hasil perbandingan akan menunjukkan keunggulan maupun keterbatasan masing-masing pendekatan dalam memprediksi parameter lingkungan.

2.1 Long Short-Term Memory (LSTM)

Long short-Term Memory (LSTM) merupakan pengembangan dari *Recurrent Neural Network (RNN)* yang dirancang untuk menangani keterbatasan model RNN konvensional, terutama pada masalah *vanishing gradient* saat memproses data sekunsial. LSTM memiliki kemampuan untuk mengingat informasi dalam jangka panjang (*long-term dependencies*) melalui struktur arsitektur khusus yang terdiri atas tiga gerbang utama, *yaitu forget gate, input gate, dan output gate*. LSTM mampu menggunakan ukuran data yang besar dan memanfaatkan semua informasi historis sebagai masukan dalam proses pembelajaran. Model ini banyak digunakan untuk peramalan (*forecasting*) data deret waktu, seperti curah hujan, suhu, dan kualitas udara karena mampu menangkap hubungan non-linear antar variabel temporal secara efektif [6]. Model LSTM pendekatan sekunsial yang digunakan dalam penelitian ini memiliki empat lapisan, yaitu:

a. Input Layer

Pada model LSTM, input layer berfungsi menerima data sekunsial dalam format tiga dimensi yaitu samples, timesteps, dan features. Menggambarkan bahwa input layer menerima beberapa variabel historis dan mengatur data agar sesuai dengan format input LSTM [7].

b. Hidden Layer

Hidden layer terdiri dari satu atau lebih lapisan LSTM yang memiliki unit memori dengan tiga gerbang utama yaitu, *forget gate, input gate, dan output gate*. Lapisan ini bertugas mempelajari pola temporal antara langkah waktu masalalu dan masa depan [8].

c. Output Layer

Output layer menghasilkan prediksi akhir berdasarkan fitur yang telah diproses melalui hidden layer. Jumlah neuron pada output layer tergantung pada prediksi yang ingin diciptakan bisa satu neuron untuk *single step* atau lebih jika model memprediksi *multistep*. Fungsi aktivasi biasanya linear dalam kasus regresi tetapi bisa juga penggunaan output layer dalam berbagai model *forecasting*, termasuk LSTM, dan bagaimana model menggunakan fungsi *loss* seperti MSE untuk pelatihan dan validasi [9].

2.1.1 Metode Multivariat LSTM

LSTM multivariat Adalah varian dari LSTM yang menggunakan lebih dari satu variable input untuk memprediksi satu atau lebih target variable. Pendekatan ini memungkinkan model mempelajari hubungan antarvariabel (*inter-variabel dependencies*) yang memengaruhi hasil prediksi. Misalnya, dalam peramalan kualitas udara, variable seperti suhu, kelembaban, kecepatan angin digunakan secara bersamaan untuk memprediksi tingkat polutan PM10 atau CO [10]. Setiap fitur dalam model multivariat memberikan kontribusi terhadap pembentukan pola yang kompleks dan saling bergantung



antarvariabel. Dengan demikian, LSTM multivariat lebih representative dibandingkan model univariat karena mampu menangkap keterkaitan antar faktor yang mempengaruhi fenomena yang diamati [11].

2.1.2 Metode Multivariat LSTM

Pendekatan univariat LSTM digunakan untuk menganalisis deret waktu dengan satu variabel input tunggal (*single future*) dan satu target output. Model ini belajar dari urutan nilai masalalu dari variabel yang sama untuk memprediksi nilai masa depan tanpa melibatkan variabel eksternal. Pendekatan ini umum digunakan pada penelitian peramalan kasus COVID-19, suhu udara, atau harga saham karena struktur datanya sederhana dan pola temporalnya kuat [12], [13].

2.2 Root Mean Square Error (RMSE)

Root Mean Square Error (RMSE) merupakan salah satu metrik evaluasi yang digunakan untuk mengukur tingkat kesalahan antara nilai hasil prediksi model dan nilai aktual. Dalam konteks model LSTM multivariat, RMSE digunakan untuk menilai seberapa baik model memprediksi nilai target, seperti volatilitas pasar atau parameter lingkungan. Semakin kecil nilai RMSE, maka semakin baik akurasi model dalam melakukan prediksi [14]. RMSE secara umum dinyatakan dengan persamaan berikut:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=0}^n (Y_{actual} - Y_{pred})^2} \quad (1)$$

Dimana N mewakili jumlah data, Y_{actual} mewakili nilai aktual, dan Y_{pred} mewakili nilai prediksi.

2.3 Root Mean Square Error (MSE)

Mean Squared Error (MSE) merupakan salah satu metrik evaluasi yang digunakan untuk mengukur rata-rata kuadrat dari selisih antara nilai aktual dan nilai prediksi. MSE digunakan secara luas dalam penelitian berbasis *machine learning* dan *deep learning* karena memberikan ukuran seberapa jauh hasil prediksi model menyimpang dari nilai sebenarnya. Nilai MSE yang semakin kecil menunjukkan bahwa hasil prediksi model semakin mendekati data aktual, sehingga akurasi model semakin tinggi [15]. MSE secara umum dinyatakan dengan persamaan berikut:

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_{pred,i} - Y_{actual,i})^2 \quad (2)$$

Dimana N mewakili jumlah data, $Y_{pred,i}$ mewakili nilai prediksi ke-i, dan $Y_{actual,i}$ mewakili nilai aktual ke-i.

2.4 Root Mean Square Error (MAPE)

Mean Absolute Percentage Error (MAPE) adalah salah satu metrik evaluasi paling populer yang digunakan untuk mengukur tingkat kesalahan pada model prediksi atau peramalan, terutama dalam regresi berbasis *machine learning* dan *deep learning*. Dalam penelitian tersebut, MAPE digunakan untuk menilai akurasi prediksi pada data pelatihan, pengujian, dan peramalan masa depan. Nilai MAPE yang rendah menunjukkan kemampuan model dalam mengikuti pola tren harga saham meskipun terdapat fluktuasi pasar yang dinamis [16].

$$MAPE = \frac{100\%}{n} \sum_{t=1}^n \left| \frac{Y_{actual,t} - Y_{pred,t}}{Y_{actual,t}} \right| \quad (3)$$

Dimana N mewakili jumlah data, $Y_{actual,t}$ mewakili nilai aktual pada waktu ke-t, dan $Y_{pred,t}$ mewakili nilai prediksi pada waktu ke-t.

3. HASIL DAN PEMBAHASAN

3.1 Pengumpulan Data

Pengumpulan data dalam penelitian ini dilakukan untuk memperoleh informasi yang diperlukan dalam menganalisis kinerja model *Long Short-Term Memory (LSTM)* baik multivariat maupun univariat dalam memprediksi parameter lingkungan di wilayah perkotaan. Data yang digunakan mencakup beberapa variabel lingkungan seperti temperature, humidity, pressure, dan altitude yang relevan dengan kondisi iklim di wilayah penelitian.

Data diperoleh melalui meteorologi dan lingkungan yang menyediakan catatan parameter iklim harian di Kota Semarang. Proses pengumpulan dilakukan dengan cara mengunduh dataset historis melalui permintaan resmi ke lembaga terkait. Periode waktu observasi yang digunakan adalah Januari 2024 hingga April 2025, sehingga mencakup dinamika cuaca dalam berbagai musim dan kondisi atmosfer yang berbeda. Sebelum digunakan dalam proses pemodelan, data terlebih dahulu melalui tahap pemeriksaan kelengkapan dan validasi untuk memastikan tidak adanya nilai yang hilang, duplikasi, maupun data yang tidak wajar. Selanjutnya, data yang telah divalidasi disimpan dalam format digital seperti CSV agar mudah diolah dalam tahap pra-pemrosesan dan analisis menggunakan model LSTM. Dengan demikian, tahap pengumpulan data ini menjadi dasar penting untuk mencapai tujuan penelitian, yaitu menganalisis sejauh mana perbedaan kinerja antara model LSTM multivariat dan univariat dalam memprediksi parameter lingkungan, serta menentukan model mana yang paling relevan digunakan untuk memprediksi parameter lingkungan di kota-kota besar seperti Semarang.



3.2 Preprocessing Data

Pada tahapan ini menunjukkan hubungan yang saling berkaitan antarparameter lingkungan. Suhu (20-37°C) memperlihatkan pola musiman dengan nilai tertinggi pada musim kemarau dan terendah pada saat musim hujan, disertai fluktuasi harian akibat pengaruh angin laut dan penyinaran matahari. Kelembaban (30-100%) menunjukkan pola berlawanan dengan suhu, meningkat pada musim hujan dan menurun saat suhu tinggi. Tekanan udara (931-947 hPa) relatif stabil dengan tren mengikuti suhu dan berlawanan dengan kelembaban. Sementara itu, ketinggian (589-720 m) tidak menunjukkan pola musiman dan lebih dipengaruhi oleh variasi sensor, namun tetap relevan karena berkaitan dengan perubahan tekanan dan suhu.

3.3 Pelatihan Model

Tahap pelatihan model LSTM dilakukan untuk membangun sistem prediksi parameter lingkungan berdasarkan data deret waktu. Dua jenis model digunakan, yaitu LSTM multivariat yang melibatkan beberapa variabel seperti temperature, humidity, pressure, dan altitude, dan LSTM univariat yang memprediksi satu variabel seperti suhu udara berdasarkan data historisnya sendiri. Data dibagi menjadi 80% data latih dan 20% data uji. Proses pelatihan dilakukan menggunakan *TensorFlow* dan *Keras*, dengan parameter utama seperti *optimizer Adam*, serta *loss function MSE*. Pelatihan dilakukan selama beberapa *epoch* dengan penerapan *early stopping* untuk mencegah *overfitting*. Setiap model dilatih selama 100 epoch dengan batch size 32, dan verbose 1.

3.3.1 Model LSTM Multivariat

Model LSTM multivariat dirancang untuk mempelajari keterkaitan dinamis antarvariabel input, baik dalam jangka pendek maupun jangka panjang guna menghasilkan prediksi yang lebih presisi dan stabil. Model ini menggunakan empat variabel utama, yakni temperature, humidity, pressure, dan altitude. Arsitektur jaringan yang digunakan untuk tiga variabel pertama memiliki total 794.277 parameter dengan ukuran sekitar 3,03 MB, sedangkan model yang dibangun untuk variabel altitude memuat 249.153 parameter berukuran 973,24 KB. Seluruh parameter bersifat *trainable*, memastikan setiap lapisan berperan aktif dalam proses pembelajaran. Hasil pelatihan ini menunjukkan grafik *loss* yang konvergen dan stabil sepanjang *epoch*, menandakan kemampuan model dalam melakukan generalisasi tanpa mengalami *overfitting* maupun *underfitting*.

3.3.2 Model LSTM Univariat

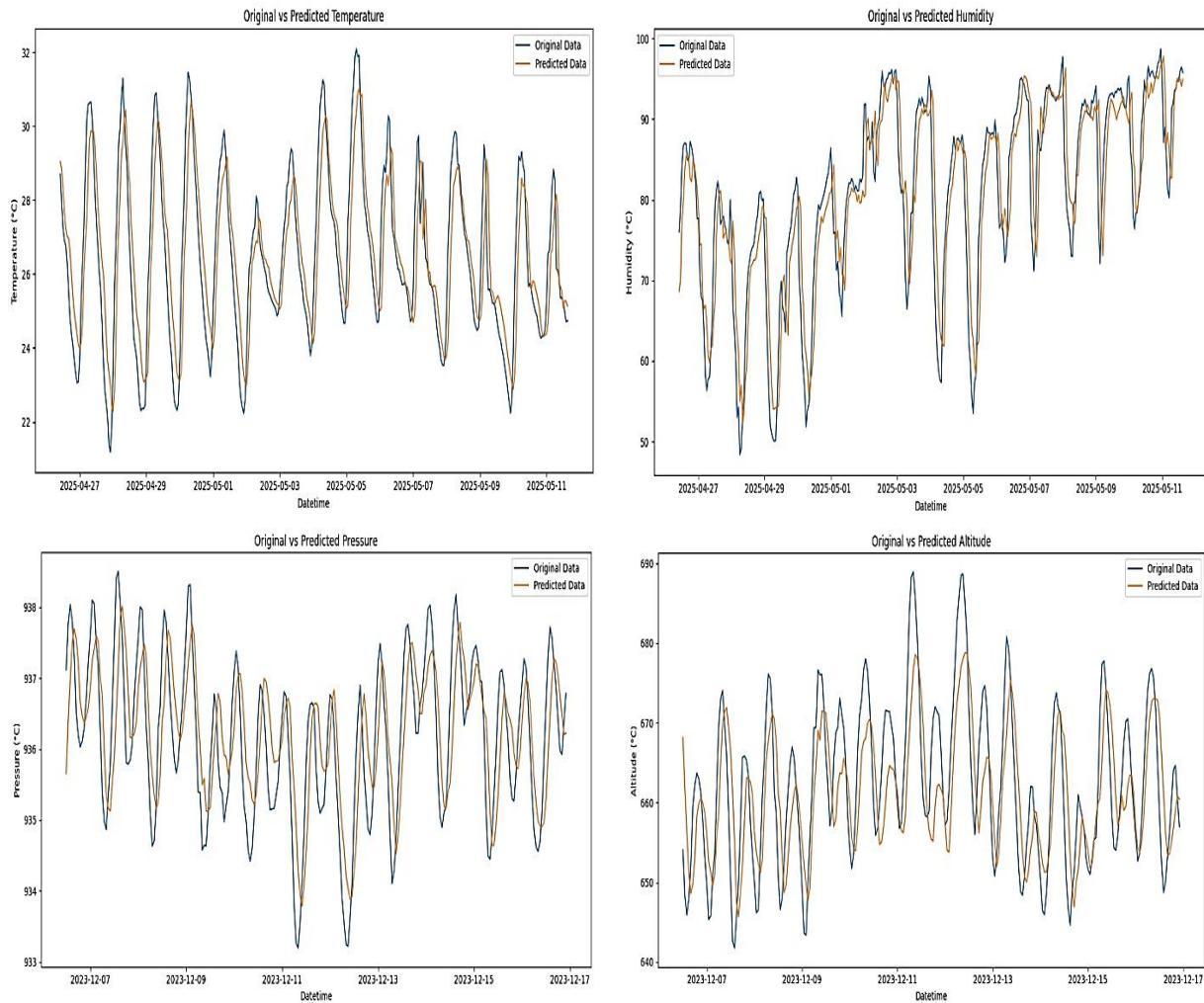
Model LSTM multivariat dikembangkan untuk menganalisis pola temporal dari satu variabel tunggal, yakni temperature, tanpa mempertimbangkan pengaruh variabel eksternal lainnya. Pendekatan ini berfokus pada pemanfaatan data historis suhu untuk menghasilkan prediksi masa depan secara spesifik. Arsitektur jaringan terdiri dari dua hidden layer LSTM dengan total 412.000 parameter berukuran sekitar 1,6 MB, seluruhnya bersifat *trainable*. Proses pelatihan memperlihatkan nilai *loss* yang menurun secara konsisten pada setiap *epoch*, menunjukkan bahwa model dapat menyesuaikan bobot secara efektif dan mencapai keseimbangan antara akurasi pelatihan serta validasi.

3.4 Hasil Prediksi

Hasil prediksi menunjukkan bahwa model LSTM multivariat maupun LSTM univariat mampu mengikuti pola fluktuasi data aktual parameter lingkungan. Model multivariat menunjukkan kinerja lebih baik pada variabel yang saling berkorelasi, seperti suhu, kelembaban, dan tekanan udara, sedangkan model univariat menghasilkan prediksi yang cukup stabil untuk variabel tunggal seperti suhu udara. Nilai error yang diukur menggunakan metrik seperti RMSE, MAPE, dan MSE menunjukkan bahwa model multivariat memiliki tingkat kesalahan yang lebih rendah dibandingkan univariat.

3.4.1 Hasil Prediksi LSTM Multivariat

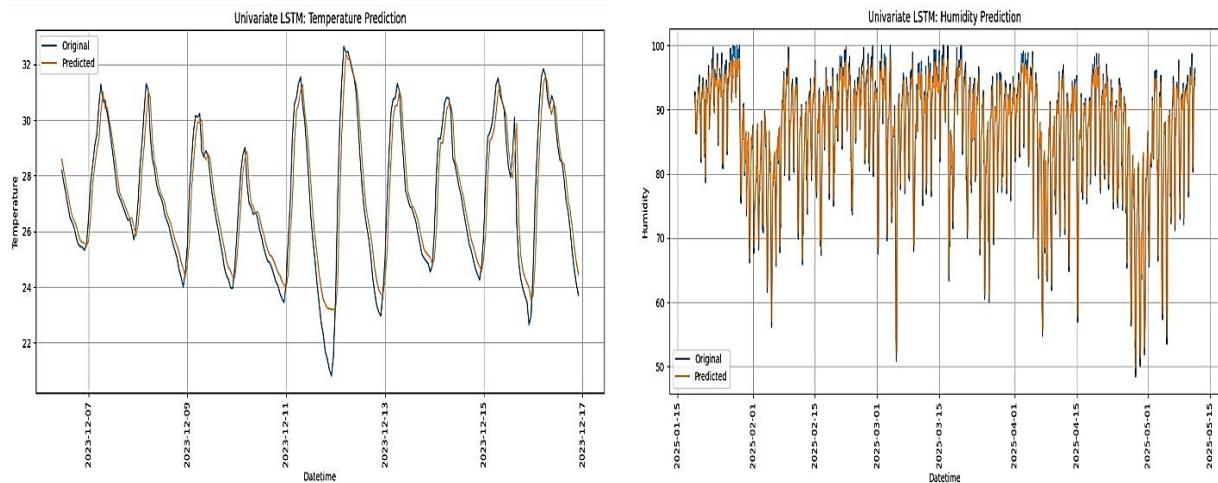
Berdasarkan Gambar 2 model LSTM multivariat menunjukkan hasil prediksi yang sangat mendekati data aktual pada keempat variabel lingkungan (temperature, humidity, pressure, dan altitude). Garis prediksi yang hampir berimpit dengan data asli menandakan model mampu mengenali pola perubahan antarvariabel dengan baik, sehingga menghasilkan prediksi yang akurat dan stabil. Kondisi ini menunjukkan bahwa model tidak hanya belajar dari pola tunggal setiap variabel, tetapi juga memahami hubungan dinamis di antara variabel-variabel tersebut. Ketepatan bentuk kurva antara data aktual dan data hasil prediksi juga menandakan bahwa proses pelatihan model telah berjalan efektif dalam menangkap karakteristik data waktu yang kompleks. Selain itu, kinerja model yang stabil pada setiap variabel lingkungan mengindikasikan bahwa LSTM multivariat memiliki kemampuan generalisasi yang tinggi dalam menghadapi variasi data. Model ini mampu merespons fluktuasi nilai tanpa menimbulkan perbedaan signifikan antara data prediksi dan data sebenarnya. Dengan demikian, dapat disimpulkan bahwa pendekatan multivariat memberikan hasil yang lebih akurat dibandingkan pendekatan univariat karena mempertimbangkan korelasi antarparameter lingkungan secara simultan. Hal ini menjadikan LSTM multivariat sebagai metode yang potensial untuk diterapkan dalam sistem prediksi kondisi lingkungan yang membutuhkan ketepatan dan konsistensi tinggi. Secara keseluruhan, hasil ini membuktikan bahwa integrasi variabel lingkungan dalam satu model prediksi mampu meningkatkan performa dan reliabilitas sistem peramalan berbasis LSTM.

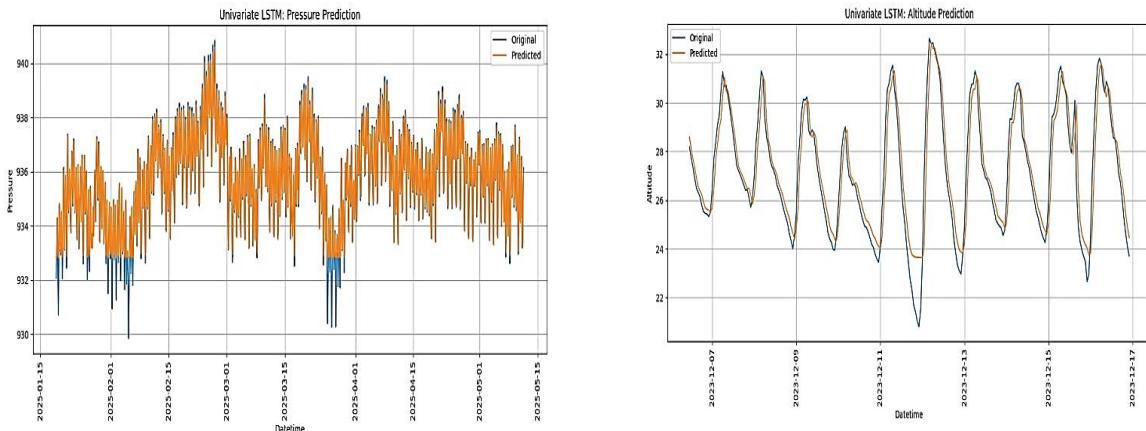


Gambar 2. Grafik data asli dan hasil prediksi LSTM multivariat

3.4.2 Hasil Prediksi LSTM Univariat

Sementara itu, hasil pada Gambar 3 menunjukkan bahwa LSTM univariat juga mampu mempelajari pola temporal tiap variabel secara efektif. Prediksi yang dihasilkan mengikuti data aktual secara konsisten, menandakan model dapat menangkap dinamika perubahan dari waktu ke waktu. Secara keseluruhan, LSTM multivariat unggul dalam mengenali keterkaitan antarvariabel, sedangkan LSTM univariat efektif untuk pola temporal tunggal yang kuat.





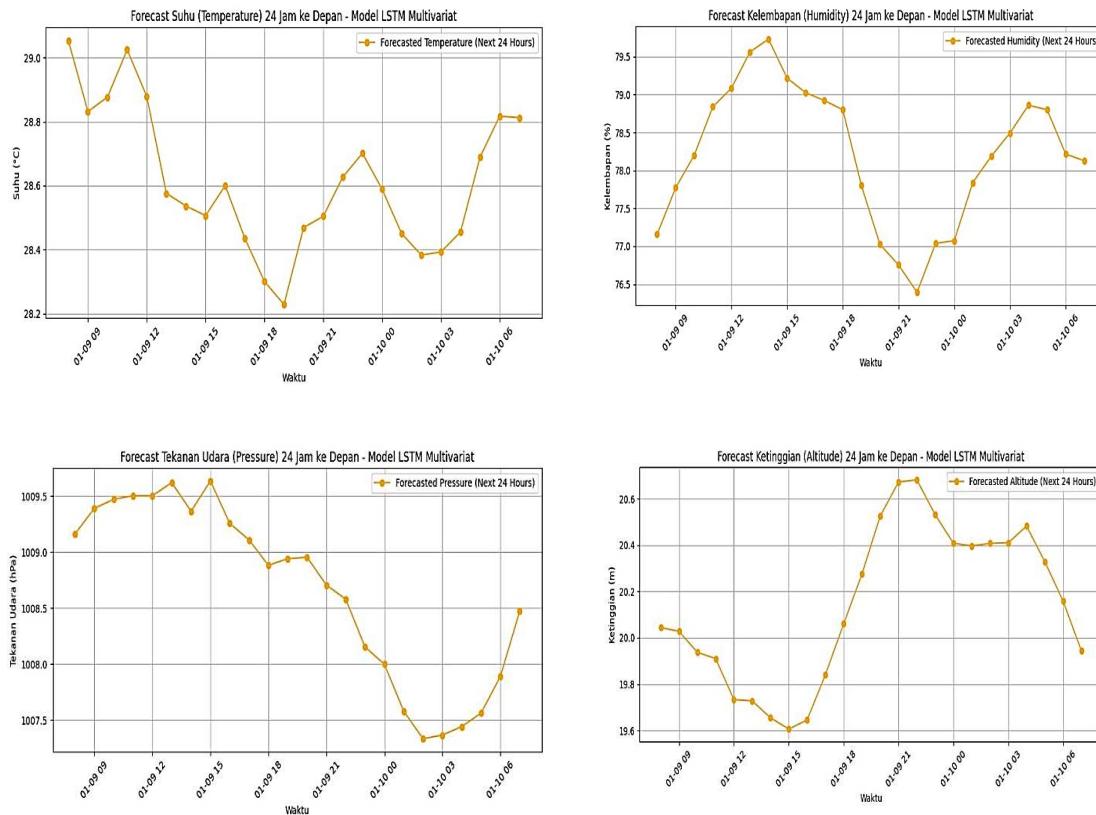
Gambar 3. Grafik data asli dan hasil prediksi LSTM univariat

3.5 Hasil Forecasting

Hasil forecasting menunjukkan bahwa model LSTM multivariat dan LSTM univariat mampu memprediksi tren parameter lingkungan dengan pola yang mendekati data aktual. Secara keseluruhan, hasil forecasting menunjukkan bahwa pendekatan LSTM multivariat lebih unggul dalam memprediksi parameter lingkungan di Kota Semarang, sehingga dapat dijadikan dasar dalam pengembangan sistem pemantauan dan perencanaan lingkungan berbasis data.

3.5.1 Hasil Forecasting LSTM Multivariat

Hasil forecasting pada Gambar 4 menggunakan model LSTM multivariat menunjukkan bahwa model ini mampu memprediksi perubahan nilai parameter lingkungan dengan pola yang sangat mendekati data aktual. Secara visual, grafik hasil prediksi memperlihatkan bahwa nilai prediksi mengikuti tren data sebenarnya dengan deviasi yang relatif kecil, terutama pada periode fluktuasi cuaca yang tinggi.



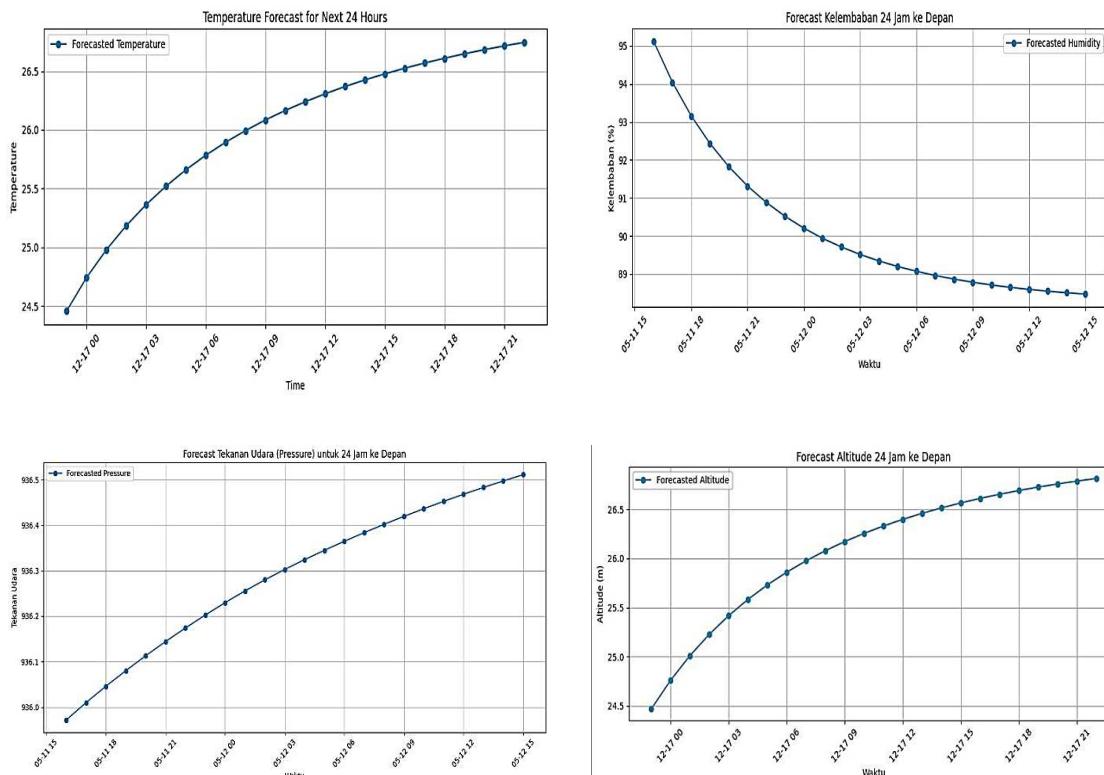
Gambar 4. Grafik hasil forecasting model LSTM multivariat

3.5.2 Hasil Prediksi LSTM Univariat

Hasil forecasting Gambar 5 menggunakan model LSTM univariat menunjukkan bahwa model mampu memprediksi pola perubahan nilai parameter lingkungan berdasarkan data historis satu variabel, seperti suhu udara. Model ini dapat mengikuti tren umum data aktual, terutama pada kondisi stabil atau fluktuasi yang tidak terlalu tajam. Namun, pada



periode dengan perubahan cuaca yang ekstrem, hasil prediksi menunjukkan sedikit deviasi data sebenarnya karena model hanya mempertimbangkan satu variabel tanpa memperhitungkan pengaruh variabel lain seperti kelembaban atau tekanan udara.



Gambar 5. Grafik hasil forecasting model LSTM univariat

3.6 Evaluasi Model

Hasil evaluasi kinerja kedua model ditunjukkan pada Tabel 1. Sejumlah penelitian menunjukkan bahwa model LSTM multivariat cenderung lebih unggul dibandingkan LSTM univariat dalam menangkap korelasi antarvariabel dan menghasilkan prediksi yang lebih akurat untuk langkah awal [17]. Namun, keunggulan tersebut tidak selalu konsisten, karena akurasi model multivariat dapat menurun pada horizont prediksi yang lebih panjang akibat kompleksitas hubungan antarvariabel dan akumulasi kesalahan. Hasil ini sejalan dengan temuan penelitian ini, dimana model multivariat mampu mengenali keterkaitan antarvariabel cuaca dengan baik, sedangkan model univariat lebih stabil pada variabel dengan fluktuasi rendah [18].

Tabel 1. Perbandingan kinerja model LSTM multivariat dan univariat

Variabel	Model	RMSE	MSE	MAPE (%)	Keterangan
Temperature	Multivariat	0.61964	0.38396	1.92	Lebih akurat
	Univariat	0.876	0.768	2.55	
Humidity	Multivariat	3.231	10.675	3.14	-
	Univariat	3.047	10.058	2.97	
Pressure	Multivariat	0.48271	0.23300	0.04	Lebih akurat
	Univariat	0.565	0.319	0.05	
Altitude	Multivariat	4.91641	24.17111	0.61	Relatif akurat (MAPE rendah)
	Univariat	0.899	0.808	2.60	

4. KESIMPULAN

Berdasarkan hasil pengujian yang telah dilakukan, diperoleh bahwa model LSTM multivariat dan LSTM univariat menunjukkan tingkat akurasi yang bervariasi pada setiap parameter lingkungan, yaitu temperature, humidity, pressure, dan altitude. Pada parameter temperature, model LSTM multivariat memberikan hasil yang lebih akurat dengan nilai RMSE sebesar 0.61964, MSE 0.38396, dan MAPE 1.92%, dibandingkan model univariat yang menghasilkan RMSE 0.876, MSE 0.768, dan MAPE 2.55%. Hal ini menunjukkan bahwa mempertimbangkan variabel lain seperti humidity, pressure, dan altitude mampu meningkatkan kemampuan model dalam mengenali pola suhu. Untuk parameter humidity, model LSTM univariat menunjukkan performa yang sedikit lebih baik dengan RMSE 3.047, MSE 9.282, dan MAPE



2.71%, menandakan bahwa kelembaban udara bersifat lebih independen terhadap variabel lain sehingga model univariat lebih sesuai digunakan. Pada parameter pressure, model LSTM multivariat menghasilkan performa terbaik dengan RMSE 0.48271, MSE 0.233000, dan MAPE 0.04%, membuktikan bahwa tekanan udara sangat dipengaruhi oleh suhu dan ketinggian sehingga model multivariat lebih efektif dalam menangkap keterkaitan antarparameter. Sedangkan parameter altitude, model multivariat memperoleh RMSE 4.91641, MSE 24.17111, dan MAPE 0.61%, sementara model univariat menghasilkan RMSE 0.899, MSE 0.808, dan MAPE 2.60%. Meskipun nilai RMSE dan MSE model multicariat lebih tinggi, nilai MAPE yang lebih rendah menunjukkan bahwa kesalahan relatif terhadap data aktual masih kecil sehingga model tetap mampu merepresentasikan pola perubahan ketinggian dengan baik. Secara keseluruhan, penelitian ini berhasil mencapai tujuannya, yaitu menganalisis dan membandingkan kinerja model LSTM multivariat dan univariat dalam memprediksi parameter lingkungan. Berdasarkan hasil evaluasi, model LSTM multivariat direkomendasikan sebagai model terbaik secara keseluruhan karena mampu menangkap hubungan antarvariabel dengan lebih efektif dan memberikan hasil prediksi yang lebih stabil pada sebagian besar parameter yang diuji.

REFERENCES

- [1] N. W. Shen, A. A. Bakar, and H. Mohamad, “Univariate and Multivariate Long Short Term Memory (LSTM) Model to Predict Covid-19 Cases in Malaysia Using Integrated Meteorological Data,” *Malaysian J. Fundam. Appl. Sci.*, vol. 19, no. 4, pp. 653–667, 2023, doi: 10.11113/mjfas.v19n4.2814.
- [2] V. Domala and T. W. Kim, “A Univariate and multivariate machine learning approach for prediction of significant wave height.,” *Ocean. Conf. Rec.*, vol. 2022-Octob, no. December, 2022, doi: 10.1109/OCEANS47191.2022.9977028.
- [3] T. Limouniac, R. Yaagoubib, K. Bouzianec, K. Guissia, and E. H. Baalia, “Univariate and Multivariate LSTM Models for One Step and Multistep PV Power Forecasting | Limouni | International Journal of Renewable Energy Development,” *Int. J. Renew. Energy Dev.*, vol. 11, no. 3, pp. 815-828., 2022, [Online]. Available: <https://ijred.cbiore.id/index.php/ijred/article/view/43953>
- [4] R. A. B. Khalil and A. A. B. U. Bakar, “A Comparative Study of Deep Learning Algorithms in Univariate and Multivariate Forecasting of the Malaysian Stock Market,” *Sains Malaysiana*, vol. 52, no. 3, pp. 993–1009, 2023, doi: 10.17576/jsm-2023-5203-22.
- [5] N. Izzany, M. Masjkur, and A. Rizki, “Application of Univariate and Multivariate Long Short Term Memory for World Crude Palm Oil Price Prediction,” *Indones. J. Stat. Its Appl.*, vol. 9, no. 1, pp. 10–20, 2025, doi: 10.29244/ijsa.v9i1p10-20.
- [6] P. Sugiartawan and S. Gusprio Santoso, “Multivariate Forecasting Curah Hujan Menggunakan Algoritma LSTM Di Kota Denpasar,” *Semin. Nas. Corisindo*, pp. 580–585, 2022.
- [7] M. J. Abbass, R. Lis, and W. Rebizant, “A Predictive Model Using Long Short-Time Memory (LSTM) Technique for Power System Voltage Stability,” *Appl. Sci.*, vol. 14, no. 16, 2024, doi: 10.3390/app14167279.
- [8] A. Mahmoudi, “Investigating LSTM-based time series prediction using dynamic systems measures,” *Evol. Syst.*, vol. 16, no. 2, pp. 1–18, 2025, doi: 10.1007/s12530-025-09703-y.
- [9] X. Song, L. Deng, H. Wang, Y. Zhang, and Y. He, “Deep learning-based time series forecasting,” 2025.
- [10] A. A. L. Rahmah, “Analisis Model Multivariate Long Short-Term Memory Untuk Prakiraan Kualitas Udara Dki Jakarta Berdasarkan Data Tahun 2010-2022,” *Repository.Uinjkt.Ac.Id*, pp. 1–87, 2024.
- [11] I. F. Efaranti, “Peramalan Data Kualitas Udara Menggunakan Multivariat LSTM di Wilayah Kota Surabaya,” 2024.
- [12] A. Kharel, Z. Zarean, and D. Kaur, “Long Short-Term Memory (LSTM) Based Deep Learning Models for Predicting Univariate Time Series Data,” *Int. J. Mach. Learn.*, vol. 14, no. 1, 2024, doi: 10.18178/ijml.2024.14.1.1154.
- [13] T. A. Prasetyo *et al.*, “Evaluating the efficacy of univariate LSTM approach for COVID-19 data prediction in Indonesia,” *Indones. J. Electr. Eng. Comput. Sci.*, vol. 34, no. 2, pp. 1353–1366, 2024, doi: 10.11591/ijeecs.v34.i2.pp1353-1366.
- [14] O. Assaf, G. Di Fatta, and G. Nicosia, “Multivariate LSTM for Stock Market Volatility Prediction,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 13164 LNCS, no. February, pp. 531–544, 2022, doi: 10.1007/978-3-030-95470-3_40.
- [15] R. Tanjung, A. Listiani, and F. Lestari, “Prediksi Multivariate Time Series Parameter Cuaca Menggunakan Long Short - Term Memory (LSTM),” *Pros. Semin. Nas. Sains Data*, vol. 4, no. 1, pp. 445–456, 2024, doi: 10.33005/senada.v4i1.253.
- [16] Furizal, A. Ritonga, A. Ma’arif, and I. Suwarno, “Stock Price Forecasting with Multivariate Time Series Long Short-Term Memory: A Deep Learning Approach,” *J. Robot. Control*, vol. 5, no. 5, pp. 1322–1335, 2024, doi: 10.18196/jrc.v5i5.22460.
- [17] S. Salehi, M. Kavgic, H. Bonakdari, and L. Begnoche, “Comparative study of univariate and multivariate strategy for short-term forecasting of heat demand density: Exploring single and hybrid deep learning models,” *Energy AI*, vol. 16, no. January, p. 100343, 2024, doi: 10.1016/j.egyai.2024.100343.
- [18] G. Davies, “Evaluating the effectiveness of predicting covariates in LSTM Networks for Time Series Forecasting,” ArXiv, 2024, [Online]. Available: <http://arxiv.org/abs/2404.18553>
- [19] A. K. Salem and A. A. Abokifa, “A Multivariate LSTM Model for Short-Term Water Demand Forecasting †,” *Eng. Proc.*, vol. 69, no. 1, pp. 1–4, 2024, doi: 10.3390/engproc2024069167.
- [20] E. Yang, H. Zhang, X. Guo, Z. Zang, Z. Liu, and Y. Liu, “A multivariate multi-step LSTM forecasting model for tuberculosis incidence with model explanation in Liaoning Province, China,” *BMC Infect. Dis.*, vol. 22, no. 1, pp. 1–13, 2022, doi: 10.1186/s12879-022-07462-8.
- [21] A. W. Saputra, A. P. Wibawa, U. Pujiyanto, A. B. Putra Utama, and A. Nafalski, “LSTM-based Multivariate Time-Series Analysis: A Case of Journal Visitors Forecasting,” *Ilk. J. Ilm.*, vol. 14, no. 1, pp. 57–62, 2022, doi: 10.33096/ilkom.v14i1.1106.57-62.
- [22] M. A. A. Bakar, N. M. Mohd Ariff, M. S. Mohd Nadzir, O. L. Wen, and F. N. A. Suris, “Prediction of Multivariate Air Quality Time Series Data using Long Short-Term Memory Network,” *Malaysian J. Fundam. Appl. Sci.*, vol. 18, no. 1, pp. 52–59, 2022, doi: 10.11113/MJFAS.V18N1.2393.