

Analisis Prediksi Penjualan Isi Ulang Air Galon menggunakan Metode LSTM dan SARIMA

Ulfarida Miftakhul Jannah ^{1*}, Nurmatalasari ¹, Ridwan Dwi Irawan ¹

¹ Program Studi Sistem Informasi, Universitas Duta Bangsa Surakarta, Indonesia

* Korespondensi: 210101084@mhs.udb.ac.id

Situs: Jannah, U. M.; Nurmatalasari, N.; and Irawan, R. D. (2025). Analisis Prediksi Penjualan Isi Ulang Air Galon menggunakan Metode LSTM dan SARIMA. *JTIM: Jurnal Teknologi Informasi Dan Multimedia*, 7(3), 626-639.

<https://doi.org/10.35746/jtim.v7i3.802>

Diterima: 23-06-2025

Direvisi: 01-08-2025

Disetujui: 11-08-2025

Abstract: Refillable drinking water depots often face challenges in dealing with unpredictable customer demand on a daily basis. This uncertainty complicates the process of stock management, production planning, and overall operations. Without accurate sales forecasts, depots risk losing potential sales and experiencing a decline in service quality to customers. Therefore, a solution is needed that can accurately predict daily sales. The first step in this research is to collect relevant data. Once the data is available, pre-processing is conducted to prepare the data before entering the modeling process. The Long Short-Term Memory (LSTM) model has the advantage of remembering historical patterns in time series data. Meanwhile, the Seasonal Autoregressive Integrated Moving Average (SARIMA) model is an extension of ARIMA that can handle data with seasonal characteristics. In this study, the LSTM model demonstrated better performance than SARIMA. This is evidenced by the performance evaluation values: MAPE of 9.54%, RMSE of 0.17, and MAE of 0.14 for the LSTM model, which are lower than MAPE of 10.51%, RMSE of 0.19, and MAE of 0.16 for SARIMA. These values indicate that LSTM is capable of providing more accurate prediction results. Based on these results, it can be concluded that the LSTM model is more effective and recommended for use in predicting daily sales of refillable water at the Manshurin Water depot.

Keywords: Sales Forecast; LSTM; SARIMA; Time Series; Refillable Drinking Water.



Copyright: © 2025 oleh para penulis. Karya ini dilisensikan di bawah Creative Commons Attribution-ShareAlike 4.0 International License. (<https://creativecommons.org/licenses/by-sa/4.0/>).

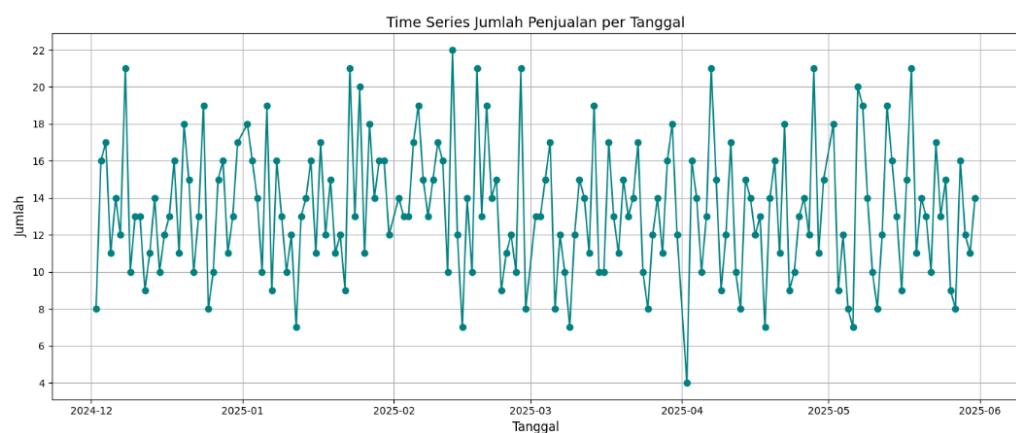
Abstrak: Depot air minum isi ulang sering menghadapi tantangan dalam menghadapi permintaan pelanggan yang tidak menentu setiap harinya. Ketidakpastian ini menyulitkan proses pengelolaan stok, perencanaan produksi, serta operasional secara keseluruhan. Tanpa adanya prediksi penjualan yang akurat, depot berisiko kehilangan potensi penjualan dan mengalami penurunan kualitas layanan terhadap pelanggan. Oleh karena itu, diperlukan solusi yang mampu memperkirakan penjualan harian secara tepat. Langkah awal dalam penelitian ini adalah mengumpulkan data yang relevan. Setelah data tersedia, dilakukan tahapan *pre-processing* untuk mempersiapkan data sebelum masuk ke proses pemodelan. Model Long Short-Term Memory (LSTM) memiliki keunggulan dalam mengingat pola historis pada data deret waktu. Sementara itu, model Seasonal Autoregressive Integrated Moving Average (SARIMA) merupakan pengembangan dari ARIMA yang mampu menangani data dengan karakteristik musiman. Dalam penelitian ini, model LSTM menunjukkan kinerja yang lebih baik dibandingkan SARIMA. Hal ini dibuktikan melalui nilai evaluasi performa: MAPE sebesar 9,54%, RMSE sebesar 0,17, dan MAE sebesar 0,14 pada model LSTM, yang lebih rendah dibandingkan MAPE 10,51%, RMSE 0,19, dan MAE 0,16 pada SARIMA. Nilai-nilai tersebut menunjukkan bahwa LSTM mampu memberikan hasil prediksi yang lebih akurat. Berdasarkan hasil tersebut, dapat disimpulkan bahwa model LSTM lebih efektif dan direkomendasikan untuk digunakan dalam memprediksi penjualan harian air isi ulang di depot Manshurin Water.

Kata kunci: Prediksi Penjualan; LSTM; SARIMA; Time Series; Air Minum Isi Ulang.

1. Pendahuluan

Di Indonesia, depot air minum UMKM merupakan pilihan utama bagi masyarakat untuk memenuhi kebutuhan air minum bersih, terutama di tempat-tempat yang sulit diakses oleh usaha besar. Selain itu, UMKM ini melengkapi inisiatif pemerintah untuk meningkatkan kesejahteraan masyarakat dengan membantu perluasan akses air minum bersih di berbagai daerah. Salah satu industri yang memiliki potensi pasar yang signifikan dan pertumbuhan yang berkelanjutan adalah UMKM Depo Air Minum Isi Ulang [1]. Salah satu depot air minum isi ulang yang berkembang pesat adalah Manshurin Water, yang telah berdiri sejak tahun 2018 di Polokarto, Sukoharjo, Jawa Tengah. Manshurin Water menawarkan air minum berkualitas dengan harga bersaing untuk memenuhi kebutuhan pelanggannya di daerahnya.

Depot air minum isi ulang menghadapi tantangan signifikan terkait ketidakstabilan permintaan pelanggan yang bervariasi setiap hari. Ketidakpastian ini menyulitkan pengelolaan persediaan, perencanaan produksi, dan pengelolaan operasional secara keseluruhan. Tanpa perkiraan penjualan yang akurat, depot berisiko kehilangan peluang penjualan berharga, dan berpotensi mengganggu kualitas layanan kepada pelanggan. Oleh karena itu, sangat penting untuk mengembangkan solusi yang dapat memprediksi penjualan harian dengan akurasi yang memadai [2]. Berdasarkan penelitian sebelumnya [3], peramalan penjualan dapat memberikan dukungan signifikan bagi manajemen perusahaan dalam mengambil keputusan yang lebih terinformasi. Proyeksi teknis permintaan konsumen di masa depan memungkinkan perusahaan merencanakan strategi yang lebih efektif. Dalam konteks ini, pemilik Manshurin Water Depot perlu menerapkan metode peramalan penjualan yang tepat untuk mengantisipasi fluktuasi permintaan yang potensial. Dengan demikian, mereka dapat meningkatkan efisiensi operasional dan memastikan tingkat persediaan selalu mencukupi untuk memenuhi kebutuhan pelanggan. Untuk melakukan prediksi penjualan dibutuhkan data harian dari pemilik Depot Isi Ulang Manshurin water, dari hasil wawancara dengan pemilik Depot Isi Ulang Manshurin water didapatkan data penjualan khususnya bulan Desember 2024 sampai dengan bulan Mei 2025. Hasil perhitungan *Time Series* dari data yang diperoleh sebagai berikut:



Gambar 1. *Time Series* Penjualan Isi Ulang air Galon

Berdasarkan gambar 1 di atas, visualisasi data deret waktu (*time series*) penjualan isi ulang air galon harian di depot Manshurin Water selama rentang waktu dari Desember 2024 hingga Mei 2025. Terlihat bahwa pola penjualan cenderung berfluktuasi dengan variasi naik turun yang cukup tajam dari hari ke hari. Beberapa titik menunjukkan puncak penjualan di atas 20 unit per hari, sementara titik terendah berada di kisaran 4 unit. Karakteristik ini mencerminkan adanya variasi musiman dan komponen acak, yang umum ditemukan pada data penjualan harian. Dalam konteks deret waktu seperti ini,

diperlukan metode peramalan yang mampu menangkap pola musiman, tren, dan fluktuasi acak. Salah satu metode statistik yang banyak digunakan adalah Seasonal ARIMA (SARIMA), yang mampu mengakomodasi pola musiman dan non-stasioner pada data penjualan. SARIMA bekerja efektif dalam kasus yang memiliki musiman yang berulang secara periodik dan cocok untuk data yang telah diubah menjadi stasioner terlebih dahulu [4].

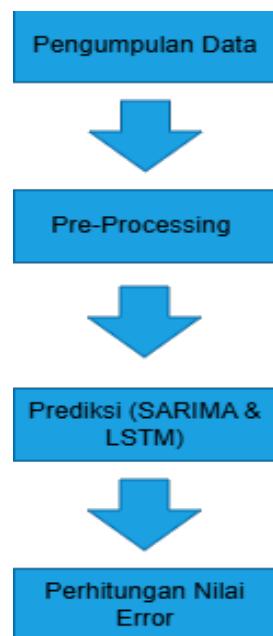
Model Seasonal ARIMA (SARIMA) merupakan suatu bentuk dari ARIMA yang dirancang khusus untuk menganalisis data deret waktu yang menunjukkan pola musiman yang jelas. Dalam penelitian [5] menerapkan SARIMA pada peramalan permintaan e-commerce dan berhasil menangkap pola musiman harian secara akurat dengan evaluasi menggunakan metrik RMSE dan MAPE yang baik. Selain itu, studi oleh [6] pada harga sayuran e-commerce menunjukkan bahwa SARIMA mampu memodelkan siklus musiman dan tren secara efektif dalam platform perdagangan elektronik. Namun, SARIMA memiliki keterbatasan dalam mengenali pola non-linier dan dependensi jangka panjang yang kompleks dalam data yang variatif. Oleh karena itu, penelitian ini juga menguji metode Long Short-Term Memory (LSTM) yang merupakan model jaringan saraf dalam dan telah diterapkan secara efektif dalam konteks deret waktu non-stasioner di Indonesia[7].

Dalam penelitian oleh [8], LSTM berhasil diterapkan untuk memprediksi penjualan sediaan farmasi di Apotek Suganda, menunjukkan kemampuan model ini dalam menangkap pola nonlinier dan fluktuasi data yang signifikan. Selain itu, studi oleh [9] menunjukkan bahwa LSTM mampu memberikan akurasi prediksi yang lebih tinggi dibandingkan metode ARIMA dalam konteks peramalan kebutuhan bahan baku pada industri makanan.

Penelitian ini memiliki kebaruan pada penerapan perbandingan metode LSTM dan SARIMA secara spesifik untuk memprediksi penjualan harian air galon pada UMKM depot air minum isi ulang di Indonesia dengan data musiman dan fluktuasi harian yang belum banyak diteliti sebelumnya. Secara teoritis, penelitian ini memperkaya literatur forecasting dengan menunjukkan bagaimana LSTM dapat mengatasi pola musiman dan fluktuasi harian secara lebih akurat dibandingkan SARIMA pada data penjualan granular harian. Secara praktis, penelitian ini memberikan kontribusi nyata bagi UMKM depot isi ulang air minum dalam mengoptimalkan manajemen stok, perencanaan produksi, dan distribusi harian berdasarkan hasil prediksi yang akurat. Penerapan model prediksi ini dapat membantu depot mengurangi risiko kehabisan stok, menekan biaya operasional akibat kelebihan stok, dan meningkatkan pelayanan kepada pelanggan melalui pengelolaan distribusi yang lebih terencana. Beberapa penelitian sebelumnya telah menguji dua model yaitu LSTM dan SARIMA dalam berbagai konteks. Penelitian [7] menunjukkan bahwa meskipun ARIMA memiliki performa yang cukup baik dengan nilai MAPE di bawah 3% pada data suhu permukaan laut, model LSTM menunjukkan akurasi yang lebih tinggi dengan nilai MAPE yang lebih rendah pada data latih dan uji, serta kemampuan yang lebih baik dalam menangkap pola fluktuatif dan hubungan jangka panjang pada data non-linier. Sementara itu, [10] dalam studi peramalan emisi karbon menemukan bahwa LSTM menghasilkan akurasi terbaik dengan MAPE hanya sebesar 0,540%, jauh lebih baik dibanding SARIMA yang mencapai 1,995%. Temuan ini menjadi dasar bagi penelitian ini untuk menguji kedua model tersebut dalam konteks prediksi penjualan isi ulang air galon di depot Manshurin Water, guna memperoleh model yang paling sesuai dengan karakteristik data yang digunakan.

2. Bahan dan Metode

Untuk mencapai hasil peramalan penjualan yang optimal, terdapat serangkaian tahap penelitian yang harus dilakukan secara berurutan. Tahap-tahap ini dapat dilihat pada Gambar 2.



Gambar 2. Tahapan Penelitian.

2.1. Pengumpulan Data

Tahap awal penelitian ini adalah pengumpulan data aktual melalui pengumpulan informasi yang relevan dengan objek penelitian. Setelah data dikumpulkan, langkah selanjutnya adalah melakukan observasi untuk menganalisis data tersebut [11]. Data penjualan isi ulang air galon diperoleh melalui wawancara dengan pemilik Depot Mansjurin dalam konteks penelitian prediktif ini. Set data yang digunakan mencakup periode dari Desember 2024 hingga Mei 2025, dengan rentang waktu 6 bulan. Tabel 1 di bawah ini menyajikan data tersebut :

Tabel 1. Data Penjualan Isi Ulang Galon

Tanggal	Produk	Jumlah	Bulan
01/12/2024	Isi Ulang Galon	11	Desember2024
02/12/2024	Isi Ulang Galon	8	Desember2024
03/12/2024	Isi Ulang Galon	16	Desember2024
04/12/2024	Isi Ulang Galon	17	Desember2024
...
27/05/2025	Isi Ulang Galon	8	Mei
28/05/2025	Isi Ulang Galon	16	Mei
29/05/2025	Isi Ulang Galon	12	Mei
30/05/2025	Isi Ulang Galon	11	Mei
31/05/2025	Isi Ulang Galon	14	Mei

2.2. Pre-Processing

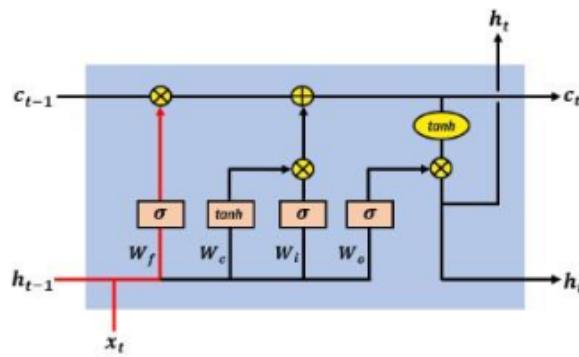
Langkah *pre-processing* adalah tahap awal yang dilakukan sebelum pengolahan data. Pada tahap ini, akan dilakukan pemeriksaan terhadap variabel-variabel yang ada dalam dataset. Analisis data ini bertujuan untuk menyederhanakan data dengan mengurangi kompleksitas yang tidak perlu, menghindari adanya gangguan seperti noise, serta mendeteksi dan menghapus elemen yang tidak relevan[12]. Adapun tahapan-tahapan

preprocessing data dalam penelitian ini adalah cek dan menangani *missing value*, konversi kolom tanggal, analisis statistik deskriptif, deteksi dan penanganan outlier dan normalisasi data (khusus untuk LSTM).

2.3. Long Short-Term Memory Neural Network (LSTM)

Long Short-Term Memory Neural Network (LSTM) merupakan salah satu bentuk dari Recurrent Neural Network (RNN). LSTM dapat menyimpan informasi terkait pola-pola dalam data, dengan kemampuan untuk mempelajari bagian data yang perlu disimpan atau dihapus. Hal ini dimungkinkan karena setiap neuron dalam LSTM memiliki beberapa gerbang (gates) yang mengendalikan memori di dalam neuron tersebut[9].

LSTM sangat cocok untuk masalah dengan ketergantungan jangka panjang. LSTM dirancang untuk mengingat informasi jangka panjang secara default. Struktur LSTM mencakup blok memori dengan 4 gerbang: *input gate*, *forget gate*, *memory-cell state gate*, dan *output gate*. LSTM memiliki unit memori yang mengontrol aliran informasi ke tahap selanjutnya [12].



Gambar 3. Struktur LSTM

Gambar 3 menggambarkan struktur LSTM. Dalam konteks pemrosesan data time series, LSTM telah terbukti efektif dalam menghubungkan informasi dari masa lalu dengan informasi dari masa depan, serta memiliki kemampuan untuk menyimpan data dalam jangka waktu yang lama. Tahap pertama dalam proses ini melibatkan perhitungan nilai fungsi sigmoid dan tanh. Selanjutnya, data dikonversi ke format pembelajaran terawasi dengan tujuan mempelajari fungsi pemetaan dari input ke output. Data yang digunakan berbentuk seri waktu, di mana nilai output untuk hari sebelumnya ditandai dengan (t-1) dan untuk hari ini dengan (t). Langkah ketiga adalah menerapkan normalisasi menggunakan metode min-max scaling. Setelah itu, variabel yang tidak relevan dihilangkan. Proses selanjutnya adalah membagi data menjadi dataset pelatihan dan dataset uji. Selanjutnya, model LSTM dibangun dengan memilih jumlah neuron, lapisan tersembunyi, dan epoch yang sesuai. Langkah berikutnya melibatkan perhitungan nilai gerbang lupa, gerbang input, pembaruan sel memori, perhitungan gerbang output, dan nilai output akhir. Sebagai langkah terakhir, model LSTM yang dirancang dievaluasi [13].

2.4. Seasonal Autoregressive Integrated Moving Average (SARIMA)

Seasonal Autoregressive Integrated Moving Average (SARIMA) merupakan pengembangan dari model ARIMA yang digunakan untuk menganalisis data deret waktu musiman. Model ini terdiri dari dua komponen utama, yaitu bagian non-musiman dan bagian musiman[14].

Dalam proses peramalan, data yang akan diramalkan harus memenuhi asumsi stasioner, artinya data tersebut harus berfluktuasi di sekitar nilai rata-rata yang stabil dan memiliki varians yang konstan yang tidak bergantung pada waktu. Jika data tersebut tidak stasioner, maka perlu dilakukan langkah differencing, yaitu dengan menghitung

selisih nilai data yang diobservasi untuk menghilangkan ketidak-stasioneran pada data [15]. Hal ini dilakukan menggunakan persamaan berikut :

- a. Untuk pembeda pertama

$$\Delta Z_t = Z_t - Z_{t-1} = Z_t - BZ_t = (1 - B)Z_t \quad (1)$$

- b. Untuk pembeda orde ke -d

$$\Delta^d Z_t = (1 - B)^d Z_t \quad (2)$$

Di mana Z_t merupakan nilai observasi pada periode ke-t, Z_{t-1} adalah nilai observasi pada periode ke-(t-1), dan d adalah tingkat differencing yang digunakan. Model SARIMA sendiri mengintegrasikan komponen non-musiman dan musiman dalam satu model dengan sifat multiplikatif. Notasi SARIMA dituliskan sebagai ARIMA(p,d,q) x (P,D,Q)S, yang dapat dituliskan sebagai berikut :

$$\Phi_p(B)\Phi_p(B^S)(1 - B)^d(1 - B^S)^D Z_t = \theta_q(B)\Theta_q(B^S)a_t \quad (3)$$

2.5. Pengukuran Performa Model

Evaluasi terhadap tingkat akurasi peramalan digunakan untuk mengukur sejauh mana hasil prediksi berbeda dengan data aktual. Beragam metode telah diterapkan untuk mengidentifikasi kesalahan yang timbul akibat penggunaan suatu pendekatan peramalan tertentu. Secara umum, metode yang digunakan menghitung rata-rata fungsi yang menggambarkan selisih antara nilai aktual dan nilai prediksi. Selisih antara data aktual dan hasil prediksi disebut residual. Sebuah model dianggap akurat jika selisih antara nilai aktual dan prediksi kecil, sebaliknya jika selisihnya besar, akurasi dianggap rendah[16]. Pengukuran performa model pada pebelitian ini dilakukan menggunakan nilai *Mean Absolute Percentage Error* (MAPE), *Root Mean Square Error* (RMSE), dan *Mean Absolute Error* (MAE).

MAPE adalah kesalahan persentase yang dihasilkan dari selisih rata-rata antara data aktual dan data prediksi. Nilai MAPE yang lebih kecil menunjukkan kinerja model prediksi yang lebih baik [17]. Tabel 2 menunjukkan kriteria nilai MAPE:

Tabel 2. Kriteria nilai MAPE (Sumber :[18])

Nilai MAPE	Kriteria
<10%	Sangat baik
<10%-20%	Baik
<20%-50%	Cukup baik
<50%	Buruk

Adapun rumus MAPE adalah sebagai berikut :

$$MAPE = 100\% \times \sum_{t=1}^n \left| \frac{A_t - F_t}{A_t} \right| \quad (4)$$

Root Mean Square Error (RMSE) mengukur akurasi perhitungan dengan membandingkan tingkat kesalahan [19]. Untuk menentukan nilai RMSE, gunakan persamaan berikut :

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (5)$$

Mean Absolute Error (MAE) adalah selisih rata-rata antara nilai aktual dan nilai yang diperkirakan. Nilai MAE yang lebih rendah menunjukkan model yang lebih akurat dalam memprediksi nilai aktual [20]. Untuk menentukan nilai MAE, gunakan persamaan berikut :

$$MAE = \frac{1}{n} \sum_{t=1}^n |Y_t - \tilde{Y}_t| \quad (6)$$

3. Hasil dan Pembahasan

Pada bagian ini akan dijelaskan hasil dari analisis data penjualan isi ulang air galon menggunakan model SARIMA dan LSTM. Proses analisis SARIMA mencakup tahapan *pre-processing* data, uji stasioneritas, analisis ACF dan PACF, pembangunan model SARIMA, evaluasi model SARIMA. Sedangkan proses analisis LSTM mencakup pembangunan model LSTM dan evaluasi model LSTM. Serta proses peramalan 7 hari ke depan dari kedua model tersebut, lalu uji model terbaik. Setiap tahap diikuti dengan visualisasi data dalam bentuk plot, yang memungkinkan pemahaman yang lebih mendalam terhadap hasil analisis. Hasil tersebut disajikan secara sistematis melalui subbab berikut.

3.1. *Pre-processing*

Tahap *preprocessing* merupakan proses penting sebelum data dianalisis atau digunakan dalam pemodelan. Pada penelitian ini, preprocessing diawali dengan pembersihan data dengan menghapus baris yang memiliki nilai null, kemudian eksplorasi data menggunakan analisis statistik deskriptif untuk memahami karakteristik dasar data.

Tabel 3. Analisis Statistik Deskriptif

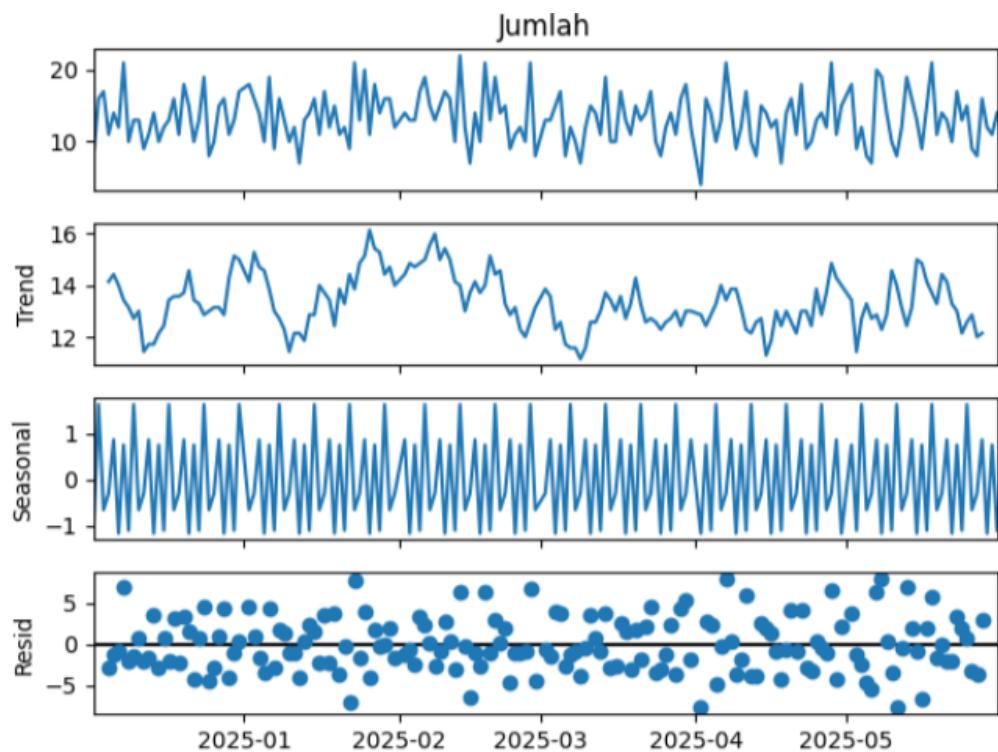
Analisis Statistik Deskriptif	
count	176.00
mean	13.32
std	3.57
min	4.00
25%	11.00
50%	13.00
75%	16.00
max	22.00

Tabel di atas menunjukkan ringkasan statistik dari atribut *Jumlah*, yang merepresentasikan volume penjualan isi ulang air galon per hari. Berdasarkan hasil analisis, jumlah total data yang diamati adalah sebanyak 176 entri, dengan nilai penjualan harian rata-rata sebesar 13,32 dan standar deviasi sebesar 3,58. Nilai minimum tercatat sebesar 4, sedangkan nilai maksimum mencapai 22. Nilai kuartil pertama (Q1) adalah 11, median (Q2) sebesar 13, dan kuartil ketiga (Q3) sebesar 16.

Analisis deskriptif ini dilakukan untuk mengidentifikasi sebaran data, mendeteksi potensi keberadaan outlier, serta mengevaluasi kebutuhan akan transformasi data seperti normalisasi. Data dinyatakan siap digunakan untuk pemodelan LSTM setelah dinormalisasi menggunakan *Min-Max Scaler* untuk mempercepat proses konvergensi dan menjaga stabilitas selama pelatihan [12]. Sementara itu, untuk pemodelan SARIMA, data harus memenuhi asumsi stasioneritas yang diuji menggunakan Augmented Dickey-Fuller (ADF) dengan $p < 0,05$, dan dilakukan *differencing* jika diperlukan [14]. Dengan demikian, data yang telah melalui preprocessing siap digunakan untuk proses modeling.

3.2. *Seasonal Autoregressive Integrated Moving Average* (SARIMA)

Langkah awal dalam prediksi menggunakan SARIMA adalah melakukan uji dekomposisi pada data training di *Google Colab* dengan bantuan *library statsmodels* dan *pandas* untuk mengidentifikasi tren dan pola. Hasil dekomposisi ini digunakan sebagai dasar dalam menentukan kelayakan penggunaan model ARIMA atau SARIMA.



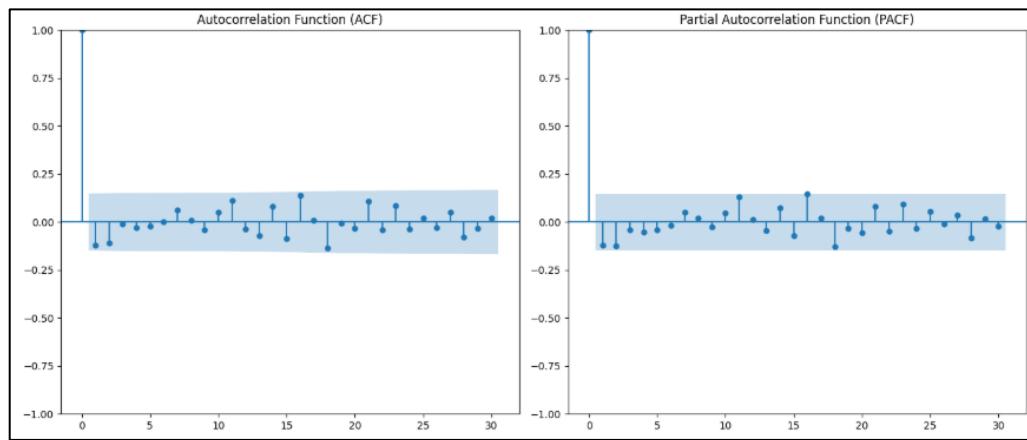
Gambar 4. Dekomposisi Time Series

Berdasarkan hasil dekomposisi, komponen tren menunjukkan adanya pola peningkatan dan penurunan penjualan secara perlahan selama periode observasi. Nilai rata-rata penjualan selama periode Desember 2024 hingga Mei 2025 tercatat sebesar 13,25 unit per hari, dengan standar deviasi 3,58, yang menunjukkan adanya variasi penjualan namun tetap berada dalam rentang stabil. Komponen musiman menunjukkan pola fluktuasi naik turun secara konsisten dalam periode mingguan, dengan puncak musiman biasanya terjadi pada awal dan pertengahan minggu. Sementara itu, komponen residual memperlihatkan variasi acak yang tidak mengikuti pola tertentu, mengindikasikan adanya pengaruh faktor eksternal yang tidak dapat dijelaskan oleh tren dan musiman.

Berdasarkan hasil analisis kuantitatif dan visual dari dekomposisi ini, dapat disimpulkan bahwa model SARIMA layak diterapkan pada dataset ini karena mampu menangkap karakteristik musiman dan tren yang telah teridentifikasi, serta dapat mengakomodasi fluktuasi penjualan secara akurat untuk mendukung pengambilan keputusan manajemen stok dan produksi di depot air minum isi ulang.

3.2.1. Analisis Autocorrelation Function (ACF) dan Partial Autocorrelation Function (PACF)

ACF dan PACF digunakan untuk membantu proses identifikasi model deret waktu yang tepat, khususnya dalam penentuan orde parameter *autoregressive* (AR) dan *moving average* (MA) pada model SARIMA. Dengan menerapkan analisis ACF dan PACF, peneliti dapat mengenali pola keterkaitan antar data berdasarkan lag waktu tertentu, sehingga menghasilkan model prediktif yang sesuai dengan karakteristik data penjualan isi ulang air galon yang digunakan dalam penelitian ini. Berdasarkan hasil dekomposisi sebelumnya yang menunjukkan adanya pola musiman dan tren, maka analisis ACF dan PACF menjadi tahap krusial untuk menentukan struktur model yang optimal. Berikut ditampilkan plot ACF dan PACF untuk mendukung proses identifikasi model SARIMA.



Gambar 5. Analisis ACF dan PACF

Berdasarkan plot ACF dan PACF pada gambar 5, menunjukkan bahwa hanya lag pertama yang menunjukkan korelasi signifikan, sementara lag-lag berikutnya berada dalam batas kepercayaan dan tidak menunjukkan pola musiman yang kuat. Pola ini mengindikasikan adanya tren jangka pendek tanpa pengaruh musiman yang dominan. Plot ACF menunjukkan penurunan cepat menuju nol (dies down), yang menunjukkan bahwa model Autoregressive (AR) dengan lag rendah sesuai untuk digunakan dalam prediksi penjualan isi ulang air galon. Hasil identifikasi ini menunjukkan bahwa lag optimal berada pada lag ke-1, sesuai dengan pola pada plot ACF dan PACF, sehingga parameter AR dan MA dengan lag rendah dapat dipilih dalam pembangunan model prediktif berbasis SARIMA untuk data penjualan isi ulang air galon.

3.2.2. Penentuan Parameter Optimal Model SARIMA

Penentuan model yang paling optimal dilakukan melalui pendekatan *grid search*, yaitu metode yang menguji berbagai set parameter pada model SARIMA. Setiap konfigurasi model dievaluasi menggunakan nilai *Akaike's Information Criterion* (AIC), di mana model dengan nilai AIC terendah dipilih sebagai yang paling tepat. Penggunaan AIC bertujuan untuk menyeimbangkan antara tingkat kesesuaian model terhadap data dan tingkat kerumitan model tersebut.

Tabel 4. Model SARIMA

Model Konfigurasi	AIC
(0,0,1)x(0,1,1,7)	877.93
(1,0,1)x(0,1,1,7)	879.45
(0,0,2)x(0,1,1,7)	878.82
(1,0,0)x(0,1,1,7)	881.12
(0,0,1)x(0,1,0,7)	880.07

Berdasarkan hasil analisis pada tabel 4, model SARIMA dengan konfigurasi $(0, 0, 1)x(0, 1, 1, 7)$ dipilih sebagai yang paling baik dengan nilai AIC sebesar 877.93. Hal ini menunjukkan bahwa model tersebut paling efektif dalam merepresentasikan pola harian penjualan isi ulang galon air dengan memperhitungkan efek musiman mingguan (periode = 7). Model ini diharapkan mampu memberikan hasil prediksi yang lebih presisi, serta menghindari overfitting karena tetap sederhana namun cukup kuat dalam menangkap pola utama pada data.

3.2.3. Implementasi Peramalan SARIMA

Setelah berhasil mengidentifikasi model yang paling sesuai, yaitu SARIMA $(0, 0, 1)x(1, 1, 1, 7)$, tahap selanjutnya adalah menerapkan model tersebut untuk proses prediksi. Proses prediksi dilakukan menggunakan *Google Colab* sebagai platform eksekusi kode

Python, untuk mempermudah komputasi dan visualisasi hasil prediksi. Prediksi dilakukan untuk periode 7 hari ke depan dengan tujuan memperkirakan jumlah penjualan harian isi ulang air galon sebagai dasar perencanaan manajemen stok dan produksi. Tabel 5 menunjukkan hasil prediksi SARIMA :

Tabel 5. Hasil Peramalan SARIMA

Tanggal	Prediksi SARIMA
2025-06-01	14.52
2025-06-02	12.37
2025-06-03	13.25
2025-06-04	14.37
2025-06-05	12.14
2025-06-06	13.77
2025-06-07	12.34

Berdasarkan hasil peramalan dari tabel 5, diperoleh bahwa jumlah penjualan isi ulang air galon yang diprediksi berkisar antara 12 hingga 14 unit per hari. Nilai ini menunjukkan adanya pola yang relatif stabil dan berulang, yang sesuai dengan karakteristik data historis yang memiliki musiman mingguan. Hasil prediksi ini menunjukkan adanya pola yang relatif stabil dengan fluktuasi mingguan yang berulang, sesuai dengan karakteristik musiman mingguan pada data historis. Informasi prediksi ini dapat digunakan oleh manajemen depot air minum untuk merencanakan jumlah produksi dan distribusi harian secara lebih akurat, sehingga dapat meminimalkan risiko kehabisan stok saat permintaan tinggi, serta mengurangi kelebihan produksi pada hari dengan permintaan lebih rendah. Dengan demikian, penerapan model SARIMA pada penelitian ini memberikan kontribusi praktis dalam pengelolaan operasional depot air minum isi ulang secara efisien.

3.3. Long Short-Term Memory Neural Network (LSTM)

Pelatihan model *Long Short-Term Memory* (LSTM) dilakukan untuk memprediksi penjualan isi ulang air galon berbasis data deret waktu harian. Sebelum dilatih, data penjualan dinormalisasi menggunakan MinMaxScaler agar berada pada rentang 0 hingga 1 untuk mempercepat proses konvergensi. Data kemudian dibagi menjadi data latih (80%) dan data uji (20%), serta diolah menjadi bentuk berurutan menggunakan parameter *look_back* sebanyak 7 hari.

Pelatihan dilakukan menggunakan pendekatan *hyperparameter tuning* menggunakan *Random Search* dilakukan dengan mengeksplorasi beberapa kombinasi hyperparameter, antara lain jumlah unit LSTM per layer (range 32 hingga 512 unit), jumlah layer LSTM (1–3 layer), tingkat dropout (0.1 hingga 0.5), serta nilai learning rate (1e-2, 1e-3, 1e-4) untuk optimasi model. Proses tuning ini bertujuan menemukan kombinasi parameter terbaik yang dapat menghasilkan nilai validation loss terendah dan akurasi prediksi yang optimal, dan disertai mekanisme *early stopping* untuk menghentikan pelatihan saat model tidak lagi menunjukkan peningkatan performa validasi selama beberapa epoch. Teknik *Early Stopping* dapat digunakan untuk meningkatkan efisiensi waktu dan mengurangi overfitting dalam klasifikasi deep learning [21]. Model akhirnya berhenti pada epoch ke-21 yang diterapkan selama proses hyperparameter tuning dengan Random Search, dimana pemantauan nilai validation loss secara otomatis menghentikan pelatihan setelah tidak ada penurunan signifikan selama 10 epoch berturut-turut, meskipun pelatihan direncanakan hingga 100 epoch. Hasil epoch pelatihan LSTM adalah sebagai berikut :

```

Epoch 1/100
5/5 6s 343ms/step - loss: 0.4238 - mae: 0.5716 - val_loss: 3.2852 - val_mae: 1.7906
Epoch 2/100
5/5 2s 133ms/step - loss: 1.5399 - mae: 1.0344 - val_loss: 0.0840 - val_mae: 0.2544
Epoch 3/100
5/5 0s 96ms/step - loss: 0.5423 - mae: 0.5509 - val_loss: 0.3308 - val_mae: 0.5318
Epoch 4/100
5/5 1s 89ms/step - loss: 0.1827 - mae: 0.3288 - val_loss: 0.0684 - val_mae: 0.2115
Epoch 5/100
5/5 0s 100ms/step - loss: 0.0570 - mae: 0.1913 - val_loss: 0.0526 - val_mae: 0.1775
Epoch 6/100
5/5 1s 89ms/step - loss: 0.0449 - mae: 0.1680 - val_loss: 0.0581 - val_mae: 0.1890
Epoch 7/100
5/5 0s 95ms/step - loss: 0.0495 - mae: 0.1757 - val_loss: 0.0569 - val_mae: 0.1763
Epoch 8/100
5/5 1s 91ms/step - loss: 0.0499 - mae: 0.1807 - val_loss: 0.0866 - val_mae: 0.2184
Epoch 9/100
5/5 1s 90ms/step - loss: 0.0519 - mae: 0.1796 - val_loss: 0.0522 - val_mae: 0.1916
Epoch 10/100
5/5 1s 92ms/step - loss: 0.0439 - mae: 0.1643 - val_loss: 0.0518 - val_mae: 0.1860
Epoch 11/100
5/5 1s 95ms/step - loss: 0.0421 - mae: 0.1637 - val_loss: 0.0446 - val_mae: 0.1758
Epoch 12/100
5/5 0s 93ms/step - loss: 0.0435 - mae: 0.1656 - val_loss: 0.0474 - val_mae: 0.1796
Epoch 13/100
5/5 0s 90ms/step - loss: 0.0412 - mae: 0.1603 - val_loss: 0.0498 - val_mae: 0.1882
Epoch 14/100
5/5 1s 99ms/step - loss: 0.0405 - mae: 0.1592 - val_loss: 0.0480 - val_mae: 0.1755
Epoch 15/100
5/5 1s 96ms/step - loss: 0.0412 - mae: 0.1600 - val_loss: 0.0501 - val_mae: 0.1867
Epoch 16/100
5/5 1s 93ms/step - loss: 0.0427 - mae: 0.1677 - val_loss: 0.0479 - val_mae: 0.1776
Epoch 17/100
5/5 1s 92ms/step - loss: 0.0443 - mae: 0.1675 - val_loss: 0.0496 - val_mae: 0.1835
Epoch 18/100
5/5 0s 88ms/step - loss: 0.0403 - mae: 0.1604 - val_loss: 0.0491 - val_mae: 0.1799
Epoch 19/100
5/5 0s 90ms/step - loss: 0.0374 - mae: 0.1523 - val_loss: 0.0491 - val_mae: 0.1787
Epoch 20/100
5/5 1s 126ms/step - loss: 0.0420 - mae: 0.1600 - val_loss: 0.0490 - val_mae: 0.1786
Epoch 21/100
5/5 1s 160ms/step - loss: 0.0402 - mae: 0.1606 - val_loss: 0.0492 - val_mae: 0.1791
Epoch 21: early stopping

```

Gambar 6. Hasil Evaluasi Model LSTM

Hasil pelatihan menunjukkan bahwa pada epoch pertama, model masih mengalami kesalahan prediksi yang cukup tinggi, dengan nilai validation loss sebesar 3.2852 dan validation MAE mencapai 1.7906. Namun, pada epoch-epoch selanjutnya, performa model meningkat secara signifikan. *Validation loss* menurun drastis menjadi 0.0840 pada epoch kedua dan terus membaik hingga mencapai nilai yang relatif stabil mulai sekitar epoch kelima, yakni 0.0526. Stabilitas performa model semakin terlihat sejak epoch ke-10 hingga epoch ke-21, dengan nilai *validation loss* dan validation MAE yang konsisten rendah dan tidak menunjukkan fluktuasi signifikan. Pelatihan dihentikan pada epoch ke-21 karena pada titik tersebut tidak terlihat adanya penurunan lebih lanjut pada validation loss, sehingga proses pelatihan dinilai telah mencapai titik konvergensi. Selain itu, tidak ditemukan indikasi *overfitting*, karena perbedaan antara training loss dan validation loss tetap kecil dan stabil hingga akhir pelatihan. Secara keseluruhan, hasil ini menunjukkan bahwa model LSTM telah berhasil mempelajari pola dari data historis secara efektif dan memiliki performa yang cukup baik untuk digunakan dalam prediksi penjualan isi ulang air galon pada periode berikutnya.

3.3.1. Implementasi peramalan LSTM

Model LSTM yang telah dilatih dan dievaluasi sebelumnya digunakan untuk memprediksi penjualan isi ulang air galon selama periode 7 hari ke depan, dimulai dari tanggal 1 Juni 2025 hingga 7 Juni 2025. Hasil peramalan tersebut menunjukkan nilai prediksi yang relatif stabil dari hari ke hari, yang menandakan bahwa model mampu menangkap pola penjualan yang tidak fluktuatif secara ekstrem. Tabel 6 menunjukkan hasil peramalan LSTM :

Tabel 6. Hasil Peramalan LSTM

Tanggal	Prediksi LSTM
2025-06-01	13.46
2025-06-02	13.02
2025-06-03	13.88
2025-06-04	12.57
2025-06-05	13.21
2025-06-06	13.19
2025-06-07	13.28

Berdasarkan hasil peramalan dari Tabel 6, diperoleh bahwa jumlah penjualan isi ulang air galon yang diprediksi berada pada kisaran 12.57 hingga 13.88 unit per hari. Hasil prediksi yang menunjukkan pola stabil dan berulang dari hari ke hari menegaskan kemampuan LSTM dalam menangkap tren musiman secara efektif. Penelitian lain oleh [13] menunjukkan bahwa LSTM lebih unggul dibandingkan ARIMA dan GRU dalam memprediksi penjualan aerosol, dengan nilai akurasi yang lebih tinggi dan kestabilan model yang lebih baik. Hasil ini sejalan dengan studi oleh [10] yang menyatakan bahwa LSTM mampu menghasilkan prediksi yang lebih akurat dalam deret waktu harian emisi karbon, dengan nilai MAPE yang jauh lebih rendah dibandingkan SARIMA. Meskipun beberapa penelitian sebelumnya mengandalkan teknik pelatihan standar, pendekatan dalam penelitian ini memanfaatkan mekanisme *early stopping* dan *random search* untuk optimasi hyperparameter, yang memberikan peningkatan performa model dan mencegah terjadinya *overfitting*.

3.4. Pengukuran Performa Model

Evaluasi performa model merupakan tahap penting dalam proses peramalan untuk mengetahui seberapa akurat hasil prediksi yang dihasilkan oleh masing-masing metode. Pengukuran performa model pada pebelitian ini dilakukan menggunakan nilai *Mean Absolute Percentage Error* (MAPE), *Root Mean Square Error* (RMSE), dan *Mean Absolute Error* (MAE). Tabel 7 menyajikan nilai *Mean Absolute Percentage Error* (MAPE), *Root Mean Square Error* (RMSE), dan *Mean Absolute Error* (MAE) dari dua model yang digunakan dalam prediksi penjualan isi ulang air galon, yaitu LSTM dan SARIMA:

Tabel 7. Perbandingan Pengukuran Performa Model

Model	MAPE	RMSE	MAE
LSTM	9.54%	0.17	0.14
SARIMA	10.51%	0.19	0.16

Berdasarkan Tabel 7, dapat dilihat bahwa model LSTM menunjukkan performa yang lebih baik dibandingkan dengan model SARIMA. Hal ini ditunjukkan oleh nilai MAPE sebesar 9,54%, RMSE sebesar 0,17, dan MAE sebesar 0,14 pada model LSTM, yang semuanya lebih rendah dibandingkan dengan nilai MAPE sebesar 10,51%, RMSE sebesar 0,19, dan MAE sebesar 0,16 pada model SARIMA. Nilai-nilai tersebut menunjukkan bahwa model LSTM memiliki tingkat kesalahan prediksi yang lebih kecil, karena dengan MAPE yang kurang dari 10% dinyatakan sangat baik untuk akurasi peramalan seperti yang dijelaskan pada subbab 2.2, dan nilai pengujian RMSE dan MAE dikatakan baik apabila mendekati 0 (nol). Sehingga, LSTM pada prediksi penjualan isi ulang air galon pada depot manshurin water dikatakan lebih baik daripada metode SARIMA.

4. Kesimpulan

Penelitian ini dilakukan untuk meramalkan penjualan harian isi ulang air galon pada depot manshurin water dengan menggunakan dua metode peramalan, yaitu SARIMA dan LSTM. Tujuan utama dari penelitian ini adalah untuk mengevaluasi performa kedua model dalam menghasilkan prediksi yang akurat dan stabil, guna mendukung

pengambilan keputusan yang lebih efektif dalam manajemen stok dan perencanaan distribusi. Berdasarkan hasil peramalan, diperoleh bahwa model LSTM menghasilkan nilai nilai MAPE sebesar 9,54%, RMSE sebesar 0,17, dan MAE sebesar 0,14, sedangkan model SARIMA memiliki nilai MAPE sebesar 10,51%. Nilai akurasi pengukuran performa model yang lebih rendah pada model LSTM menunjukkan bahwa model ini mampu memberikan prediksi yang lebih akurat dibandingkan SARIMA.

Dari temuan ini, dapat disimpulkan bahwa LSTM merupakan model yang lebih efektif dan direkomendasikan untuk digunakan dalam memprediksi penjualan harian isi ulang air galon di depot Manshurin Water. Implementasi model ini diharapkan dapat membantu pihak depot dalam mengatur jumlah produksi, menghindari kekurangan stok, serta meningkatkan efisiensi operasional secara keseluruhan. Ke depan, penelitian ini dapat dikembangkan lebih lanjut dengan mempertimbangkan faktor eksternal seperti cuaca, promosi, atau hari libur nasional untuk meningkatkan ketepatan prediksi.

Referensi

- [1] P. Rupisiay, V. N. Pelletimu, A. M. Nyawikuhy, H. M. Detek, N. N. I. Salampessy, and O. H. Nendissa, "Analisis Peramalan Permintaan Pada Depot Air Minum Isi Ulang Menggunakan Metode Least Square," *J. Tagalaya Pengabdi. Kpd. Masy.*, vol. 1, no. 3, pp. 257–262, Aug. 2024, <https://doi.org/10.71315/jtpkm.v1i3.59>
- [2] K. Sussolaikah, P. Jayadi, and D. Andrianto Putra, "Prediksi Penjualan Harian dengan Menggunakan Metode Trend Moment pada Depot Air Minum Isi Ulang," *KLIK: Kajian Ilmiah Informatika dan Komputer*, vol. 4, no. 4, pp. 1953–1961, 2024, <https://djournals.com/kliek/article/view/1629>
- [3] H. P. Hutrim and L. Bachtiar, "Aplikasi Pendukung Keputusan Prediksi Penjualan Air Minum Isi Ulang Menggunakan Metode Trend Moment," *Jutisi J. Ilm. Tek. Inform. dan Sist. Inf.*, vol. 11, no. 3, p. 583, 2022, <https://doi.org/10.35889/jutisi.v11i3.948>
- [4] T. Falatouri, F. Darbanian, P. Brandtner, and C. Udokwu, "Predictive Analytics for Demand Forecasting - A Comparison of SARIMA and LSTM in Retail SCM," *Procedia Comput. Sci.*, vol. 200, no. 2019, pp. 993–1003, 2022, <https://doi.org/10.1016/j.procs.2022.01.298>
- [5] Y. Zhao, "E-Commerce Demand Forecasting Using SARIMA Model and K-means Clustering Analysis," *J. Innov. Dev.*, vol. 7, no. 1, pp. 1–6, 2024, <https://doi.org/10.54097/ctfb0379>
- [6] K. Y. Choong *et al.*, "Application of SARIMA model in forecasting vegetable price in Malaysia: E-commerce platform domain," *AIP Conf. Proc.*, vol. 2898, no. 1, 2024, <https://doi.org/10.1063/5.0196278>
- [7] Y. Angraini *et al.*, "Comparative Analysis of ARIMA and LSTM Methods for Sea Surface Temperature Forecasting in the Sunda Strait," *J. Mat. Stat. dan Komputasi*, vol. 21, no. 3, pp. 868–885, May 2025, <https://doi.org/10.20956/j.v21i3.42565>
- [8] M. I. Anshory, Y. Priyandari, and Y. Yuniaristanto, "Peramalan Penjualan Sediaan Farmasi Menggunakan Long Short-term Memory: Studi Kasus pada Apotik Suganda," *Performa Media Ilm. Tek. Ind.*, vol. 19, no. 2, pp. 159–174, 2020, <https://doi.org/10.20961/performa.19.2.45962>
- [9] E. S. Putri and M. Sadikin, "Prediksi Penjualan Produk Untuk Mengestimasi Kebutuhan Bahan Baku Menggunakan Perbandingan Algoritma LSTM dan ARIMA," *Format J. Ilm. Tek. Inform.*, vol. 10, no. 2, p. 162, 2021, <https://doi.org/10.22441/format.2021.v10.i2.007>
- [10] S. I. N. Suwandi, Raras Tyasnurita, and Hanifan Muhyat, "Peramalan Emisi Karbon Menggunakan Metode SARIMA dan LSTM," *J. Comput. Sci. Informatics Eng.*, vol. 6, no. 1, pp. 73–80, 2022, <https://doi.org/10.29303/jcosine.v6i1.436>
- [11] Muhammad Rais Ramadhani, Nurmatalasari, and Nurchim, "Pengembangan Model Prediksi Penjualan Ice Cream UMKM Jogja Menggunakan Metode Autoregressive," *J. Deriv. J. Mat. dan Pendidik. Mat.*, vol. 11, no. 2, pp. 113–122, 2024, <https://doi.org/10.31316/jderivat.v10i2.6289>
- [12] T. A. Faisal Muhammad and M. I. Irawan, "Implementasi Long Short-Term Memory (LSTM) untuk Prediksi Intensitas Curah Hujan (Studi Kasus: Kabupaten Malang)," *J. Sains dan Seni ITS*, vol. 12, no. 1, May 2023, <https://doi.org/10.12962/j23373520.v12i1.106892>
- [13] Nendi Sunendar, Harjono P. Putro, and Rizki Hesananda, "Prediksi Penjualan Aerosol Menggunakan Algoritma ARIMA, LSTM Dan GRU," *INSOLOGI J. Sains dan Teknol.*, vol. 4, no. 1, pp. 113–126, Feb. 2025, <https://doi.org/10.55123/insologi.v4i1.4868>
- [14] Marisa, La Pimpi, and Alfian, "ANALISIS METODE SEASONAL AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (SARIMA) DAN PENERAPANNYA UNTUK MERAMALKAN PENJUALAN MOTOR YAMAHA DI INDONESIA," *J. Mat. Komputasi dan Stat.*, vol. 5, no. 1, pp. 848–856, Apr. 2025, <https://doi.org/10.33772/jmks.v5i1.117>
- [15] I. Fitriani, Prizka Rismawati Arum, and S. Amri, "Pemodelan Seasonal Autoregressive Integrated Moving Average (SARIMA) untuk Meramalkan Volume Angkutan Barang Kereta Api di Pulau Jawa Tahun 2021," *J. Data Insights*, vol. 2, no. 1, pp. 26–35, Jun. 2024, <https://doi.org/10.26714/jodi.v2i1.167>

- [16] L. Sarifah, S. Kamilah, and S. Khotijah, "Penerapan Metode Single Moving Average Dalam Memprediksi Jumlah Penduduk Miskin Pada Perencanaan Pembangunan Daerah Kabupaten Pamekasan," *Zeta - Math J.*, vol. 8, no. 2, pp. 47–54, 2023, <https://doi.org/10.31102/zeta.2023.8.2.47-54>
- [17] W. I. Butarbutar *et al.*, "Peramalan Harga Taksiran Emas Dan Uang Pinjaman Pada Pt. Pegadaian Cabang Passo Dengan Menggunakan Metode Double Moving Average," *Var. J. Stat. Its Appl.*, vol. 5, no. 1, pp. 1–12, 2023, <https://doi.org/10.30598/variancevol5iss1page1-12>
- [18] A. T. Nurani, A. Setiawan, and B. Susanto, "Perbandingan Kinerja Regresi Decision Tree dan Regresi Linear Berganda untuk Prediksi BMI pada Dataset Asthma," *J. Sains dan Edukasi Sains*, vol. 6, no. 1, pp. 34–43, May 2023, <https://doi.org/10.24246/juses.v6i1p34-43>
- [19] R. Al Kiramy, I. Permana, A. Marsal, M. R. Munzir, and M. Megawati, "Perbandingan Performa Algoritma RNN dan LSTM dalam Prediksi Jumlah Jamaah Umrah pada PT. Hajar Aswad," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 4, no. 4, pp. 1224–1234, Jul. 2024, <https://doi.org/10.57152/malcom.v4i4.1373>
- [20] A. R. Wijaya, "Model Prediksi Data Harga Minyak Mentah Dunia Dengan Metode Exponential Smoothing," *Bul. Ilm. Math. Stat. dan Ter.*, vol. 12, no. 1, pp. 21–28, 2023, <https://jurnal.untan.ac.id/index.php/jbmstr/article/viewFile/61977>
- [21] I. M. Andika Surya, T. A. Cahyanto, and L. A. Muharom, "Deep Learning dengan Teknik Early Stopping untuk Mendeteksi Malware pada Perangkat IoT," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 12, no. 1, pp. 21–30, Feb. 2025, <https://doi.org/10.25126/jtiik.2025128267>