

PREDIKSI PERMINTAAN PRODUK MENGGUNAKAN LONG SHORT-TERM MEMORY (LSTM) UNTUK OPTIMASI MANAJEMEN PERSEDIAAN PADA BISNIS RITEL

Dwi Hartanti, Sri Lestari

Teknik Informatika, Universitas Duta Bangsa Surakarta
Jl. Bhayangkara No.55, Tipes, Kec. Serengan, Kota Surakarta, Jawa Tengah
dwhartanti@udb.ac.id

ABSTRAK

Ketidakpastian permintaan pada sektor ritel sering menyebabkan ketidakseimbangan persediaan, baik dalam bentuk kelebihan stok maupun kekurangan stok. Kondisi tersebut berdampak pada meningkatnya biaya operasional serta potensi kehilangan peluang penjualan. Penelitian ini bertujuan untuk memprediksi penjualan produk ritel menggunakan metode Long Short-Term Memory (LSTM) berbasis data deret waktu (time series). Tahapan penelitian meliputi preprocessing data, pembentukan dataset menggunakan teknik sliding window, pembangunan model LSTM, serta evaluasi kinerja model menggunakan Root Mean Square Error (RMSE) dan Mean Absolute Percentage Error (MAPE). Dataset yang digunakan terdiri dari 730 data penjualan harian dengan ukuran window selama 7 hari. Hasil penelitian menunjukkan bahwa model LSTM mampu menghasilkan prediksi dengan tingkat kesalahan yang relatif rendah, ditunjukkan oleh nilai RMSE sebesar 4,16 dan MAPE sebesar 1,64%. Nilai tersebut menunjukkan bahwa model mampu menangkap pola temporal dan dinamika fluktuasi penjualan secara efektif. Hasil penelitian ini menunjukkan bahwa metode LSTM memiliki potensi untuk mendukung proses pengambilan keputusan dalam perencanaan dan pengendalian persediaan pada sektor ritel.

Kata kunci : Deep Learning, LSTM, Manajemen Persediaan, Prediksi Permintaan, Time Series.

1. PENDAHULUAN

Pertumbuhan teknologi digital telah mendorong transformasi besar dalam berbagai industri, khususnya pada sektor distribusi dan ritel. Salah satu tantangan utama yang dihadapi industri tersebut adalah ketidakpastian permintaan produk yang berdampak langsung terhadap efektivitas manajemen persediaan[1].

Ketidaktepatan dalam memprediksi permintaan dapat menyebabkan terjadinya overstock maupun stockout sehingga meningkatkan biaya operasional dan menurunkan tingkat kepuasan pelanggan[2][3].

Penelitian sebelumnya menunjukkan bahwa kesalahan dalam peramalan permintaan dapat meningkatkan biaya rantai pasok secara signifikan serta mengganggu stabilitas operasional perusahaan[1][4].

Selain itu, dalam konteks transformasi digital, perusahaan dituntut untuk mampu memanfaatkan data historis secara optimal guna mendukung pengambilan keputusan berbasis data yang lebih akurat [5].

Menanggapi permasalahan tersebut, berbagai metode prediksi telah dikembangkan, mulai dari pendekatan statistik konvensional hingga metode berbasis kecerdasan buatan. Metode seperti ARIMA masih banyak digunakan dalam prediksi data deret waktu, namun metode tersebut memiliki keterbatasan dalam menangkap pola non-linear dan hubungan kompleks pada data time series[6][7].

Dalam beberapa tahun terakhir, pendekatan berbasis deep learning menjadi perhatian utama karena dinilai mampu memodelkan hubungan data yang lebih kompleks[8][9].

Salah satu model deep learning yang banyak digunakan adalah Long Short-Term Memory (LSTM), yang memiliki kemampuan dalam menangkap dependensi jangka panjang pada data sekuensial dibandingkan metode tradisional[10][11].

Penelitian sebelumnya juga menunjukkan bahwa model LSTM mampu menghasilkan performa prediksi yang lebih stabil dan akurat pada berbagai jenis data deret waktu[10][12].

Sejumlah penelitian telah mengkaji penerapan model LSTM dalam prediksi permintaan produk. Penelitian sebelumnya menyebutkan bahwa model deep learning memiliki kemampuan yang lebih baik dalam menangani pola data yang kompleks dan fluktuatif[10].

Namun demikian, sebagian besar penelitian tersebut masih berfokus pada peningkatan performa model dari sisi akurasi tanpa mengkaji secara mendalam bagaimana hasil prediksi dapat diinterpretasikan dan dimanfaatkan dalam mendukung pengambilan keputusan operasional. Selain itu, masih terbatas penelitian yang mengintegrasikan tahapan preprocessing data, proses pemodelan, serta evaluasi performa model secara komprehensif dalam satu kerangka penelitian yang sistematis.

Berdasarkan uraian tersebut, permasalahan yang dapat diidentifikasi adalah perlunya pendekatan prediksi yang tidak hanya berfokus pada peningkatan akurasi model, tetapi juga mampu memberikan analisis hasil prediksi yang mendukung pengambilan keputusan berbasis data. Selain itu, diperlukan penelitian yang mengimplementasikan model LSTM secara menyeluruh dengan tahapan yang terstruktur

mulai dari preprocessing data hingga evaluasi performa model.

Oleh karena itu, penelitian ini bertujuan untuk menerapkan model Long Short-Term Memory (LSTM) pada prediksi data deret waktu penjualan harian serta mengevaluasi kinerjanya menggunakan metrik Mean Absolute Percentage Error (MAPE) dan Root Mean Squared Error (RMSE). Penelitian ini juga bertujuan untuk menganalisis hasil prediksi yang dihasilkan serta mengidentifikasi potensi pemanfaatannya dalam mendukung pengambilan keputusan pada manajemen persediaan.

Rencana penyelesaian masalah dalam penelitian ini dilakukan melalui beberapa tahapan, yaitu preprocessing data untuk membersihkan dan menyiapkan data penjualan, pembentukan sequence data menggunakan pendekatan sliding window berbasis pola mingguan, pembangunan model LSTM, proses pelatihan dan pengujian model, serta evaluasi performa model menggunakan metrik MAPE dan RMSE. Kebaruan penelitian ini terletak pada penerapan model LSTM pada data penjualan harian dengan pendekatan sliding window berbasis pola mingguan yang tidak hanya berfokus pada tingkat akurasi prediksi, tetapi juga pada analisis interpretatif hasil prediksi sebagai pendukung pengambilan keputusan berbasis data.

2. TINJAUAN PUSTAKA

Dalam beberapa tahun terakhir, penelitian tentang metode LSTM untuk prediksi data deret waktu telah berkembang pesat, terutama dalam hal peningkatan akurasi prediksi dibandingkan dengan konvensional.

Penelitian juga menunjukkan bahwa model LSTM memiliki kemampuan yang lebih baik dibandingkan metode statistik seperti ARIMA dalam memodelkan data deret waktu. Dalam penelitian tersebut, dilakukan perbandingan langsung antara LSTM dan ARIMA pada beberapa dataset, dan hasilnya menunjukkan bahwa LSTM mampu menghasilkan error yang lebih rendah secara konsisten. Hal ini disebabkan oleh kemampuan LSTM dalam menangkap pola non-linear serta dependensi jangka panjang yang tidak dapat dimodelkan secara optimal oleh ARIMA[13].

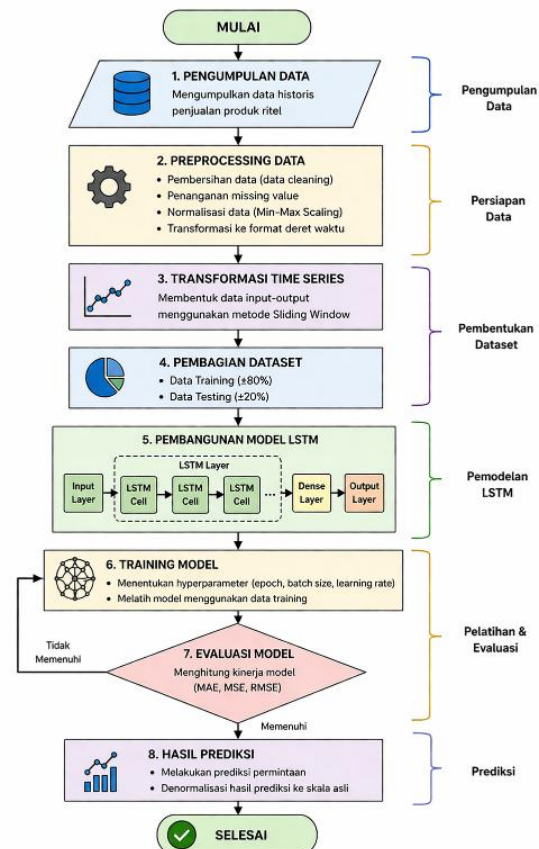
Penelitian secara komprehensif penggunaan deep learning untuk data sekuensial. Dalam studi ini dijelaskan bahwa model berbasis Recurrent Neural Network (RNN), khususnya LSTM, memiliki keunggulan dalam memproses data yang memiliki ketergantungan temporal. LSTM mampu menyimpan informasi penting dalam jangka panjang melalui mekanisme memory cell, sehingga menghasilkan performa yang lebih baik dibandingkan metode konvensional dalam memahami pola data kompleks[10].

Selanjutnya, penelitian menyoroti perkembangan metode deep learning dalam forecasting time series modern. Penelitian ini menjelaskan bahwa model

seperti LSTM mampu menangkap dependensi jangka panjang serta interaksi kompleks antar variabel dalam data. Selain itu, pendekatan deep learning terbukti lebih fleksibel dan adaptif dibandingkan metode tradisional[11].

3. METODE PENELITIAN

Metode penelitian disusun secara sistematis untuk mendukung proses pembangunan model prediksi menggunakan LSTM. Tahapan pengolahan data, pemodelan, dan evaluasi hasil prediksi termasuk dalam tahapan.



Gambar 1. Tahapan penelitian

Penelitian dimulai dengan melakukan pengumpulan data penjualan produk ritel dikumpulkan untuk dasar analisis. Setelah itu, data diproses melalui tahap preprocessing, yang mencakup pembersihan data (pembersihan data), penanganan nilai yang hilang (penanganan nilai yang hilang), normalisasi data menggunakan metode Min-Max Scaling, dan transformasi data ke dalam deret waktu. Selanjutnya, pasangan input dan output data yang akan digunakan dalam pemodelan dibentuk melalui transformasi seri waktu menggunakan metode sliding window.

Selanjutnya, dataset dilakukan pembagian menjadi sekitar 80% data pelatihan (training) dan sekitar 20% data pengujian. Kemudian dibangun model Long Short-Term Memory (LSTM) yang terdiri dari lapisan input, beberapa LSTM lapisan, dan lapisan output. Parameter seperti epoch, ukuran batch, dan

tingkat pembelajaran digunakan untuk melatih model yang telah dirancang.

Setelah dilakukan proses pelatihan sampai selesai, model diuji untuk menilai kinerjanya dengan data tes. Nilai kesalahan rata-rata akurat (MAE), rata-rata akurat kuadrat (MSE), dan rata-rata akurat kuadrat (RMSE) digunakan untuk menilai model. Jika hasil evaluasi tidak memenuhi kriteria, proses pelatihan akan diulang dengan parameter disesuaikan.

Jika model telah memenuhi kriteria evaluasi, maka dilakukan tahap prediksi untuk menghasilkan nilai permintaan di masa mendatang. Hasil prediksi kemudian dikembalikan ke skala asli (*denormalisasi*).

4. HASIL DAN PEMBAHASAN

4.1. Pengumpulan Data

Tahapan yang dilakukan mulai dari mendapatkan data yang akan digunakan dalam proses prediksi menggunakan metode Long Short-Term Memory (LSTM). Data yang digunakan terdiri dari 730 data harian dari penjualan produk ritel selama periode Januari 2025 hingga Desember 2026. Dataset terdiri dari tanggal dan jumlah penjualan yang menunjukkan permintaan produk selama periode waktu tertentu.

Tabel 1. Data penjualan

Tanggal	Jumlah Penjualan
01-01-2025	201.49
02-01-2025	207.78
03-01-2025	212.44
04-01-2025	210.03
05-01-2025	196.45
06-01-2025	191.42
.....	
31-12-2026	227.01

4.2. Persiapan Data

Tahap persiapan data (*data preprocessing*) dilakukan untuk memastikan bahwa data yang digunakan memiliki kualitas yang baik dan sesuai untuk pemodelan Long Short-Term Memory (LSTM)[14][15].

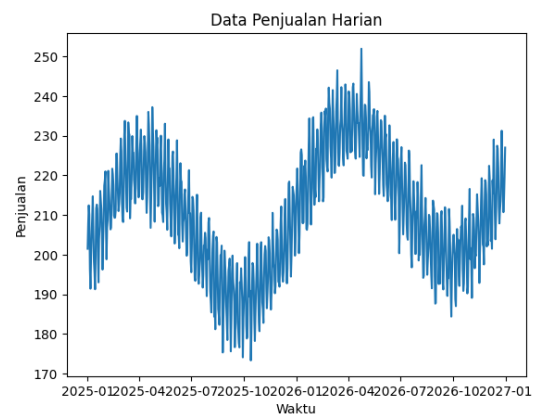
Dataset yang digunakan terdiri dari 730 data penjualan harian dari Januari 2025 hingga Desember 2026. Proses diawali dengan *data cleaning* untuk memeriksa duplikasi, inkonsistensi, dan anomali nilai[16].

Hasil pemeriksaan menunjukkan tidak terdapat data duplikat, namun validasi tetap dilakukan untuk memastikan tidak adanya *outlier* ekstrem. Selanjutnya, dilakukan penanganan *missing value*. Meskipun tidak ditemukan nilai hilang, metode interpolasi linear disiapkan sebagai antisipasi karena mampu mempertahankan pola tren data deret waktu.

Untuk meningkatkan kinerja dan stabilitas model LSTM, langkah selanjutnya adalah normalisasi dengan metode skala Min-Max untuk mengubah rentang data ke skala 0–1. Proses normalisasi dilakukan dengan persamaan:

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (1)$$

Selanjutnya, data ditransformasikan ke dalam bentuk *time series* menggunakan teknik *sliding window*, di mana sejumlah data sebelumnya (misalnya 7 hari) digunakan sebagai input untuk memprediksi nilai berikutnya. Terakhir, dataset dibagi menjadi data training dan testing dengan rasio 80:20 untuk proses pelatihan dan evaluasi model. Tahapan ini memastikan data yang digunakan telah bersih, terstandarisasi, dan terstruktur, sehingga mampu meningkatkan akurasi dan stabilitas hasil prediksi model LSTM.



Shape X: (723, 7)
Shape y: (723,)

Gambar 2. Hasil Windowing

Proses transformasi data menggunakan metode *sliding window* dengan time step 7 menghasilkan 723 data sekuens yang digunakan sebagai input model.

4.3. Pembentukan Dataset

Tahap pembentukan dataset dilakukan untuk mengubah data historis penjualan yang telah melalui proses *preprocessing* menjadi format yang sesuai untuk pemodelan Long Short-Term Memory (LSTM). Dataset yang digunakan terdiri dari 730 data penjualan harian selama periode Januari 2025 hingga Desember 2026. Pada tahap ini, data deret waktu ditransformasikan ke dalam bentuk *supervised learning* menggunakan teknik *sliding window*. Teknik ini membentuk pasangan data input dan output berdasarkan sejumlah periode sebelumnya. Dalam penelitian ini digunakan *window size* sebanyak 7 hari, sehingga data penjualan selama 7 hari sebelumnya digunakan sebagai input untuk memprediksi nilai penjualan pada hari ke-8.

Tabel 2. Pembentukan dataset menggunakan sliding window

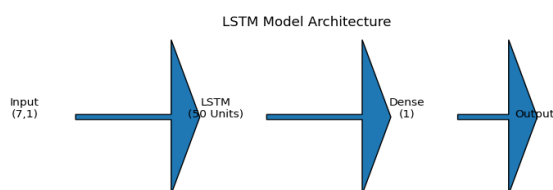
No	Input (7 Hari Sebelumnya)	Target (Hari ke-8)
1	201.49, 207.85, 210.32, 215.67, 220.11, 218.45, 216.90	222.34
2	207.85, 210.32, 215.67, 220.11, 218.45, 216.90, 222.34	225.10

Berdasarkan proses transformasi, dari total 730 data penjualan harian (tanpa data hilang) dihasilkan 723 data sekuens yang terdiri dari pasangan input (X) dan target (y). Setiap input memiliki panjang 7 periode (time step) yang merepresentasikan pola historis penjualan, dengan target berupa nilai pada hari ke-8. Dataset kemudian data dilakukan pembagian menjadi = data training dan data testing menggunakan rasio 80:20 secara berurutan untuk menjaga karakteristik deret waktu. Dari pembagian tersebut diperoleh 578 data sebagai training dan 145 data sebagai testing (hasil pembulatan dari proporsi 80:20). Selanjutnya, data input diubah ke dalam bentuk tiga dimensi sesuai kebutuhan model LSTM, yaitu [jumlah sampel, time step, jumlah fitur]. Dalam penelitian ini digunakan satu fitur (univariate), yaitu data penjualan harian. Bentuk akhir data training adalah (578, 7, 1), yang menunjukkan 578 sampel, 7 time step, dan 1 fitur. Dengan demikian, tahap pembentukan dataset menghasilkan data sekuens time series yang terstruktur dan siap digunakan dalam proses pelatihan model LSTM.

4.4. Pemodelan LSTM

Untuk menghasilkan model prediksi, tahap pemodelan dilakukan dengan menggunakan Long Short-Term Memory (LSTM). LSTM dipilih karena kemampuan untuk menangkap pola temporal dan dependensi jangka panjang pada data deret waktu. Model yang digunakan terdiri dari satu lapisan LSTM yang memiliki 50 unit dan satu lapisan dense sebagai outputnya. Lapisan LSTM digunakan untuk mengekstraksi pola sekuensial dari data historis, dan output dikirim ke lapisan dense untuk menghasilkan nilai prediksi. Data input memiliki dimensi (578, 7, 1), yang menunjukkan 578 sampel dengan 1 fitur (univariate) dan 7 langkah waktu.

Untuk meminimalkan nilai kehilangan, optimizer Adam dan fungsi kehilangan mean squared error (MSE) digunakan dalam proses pelatihan. Model dilatih selama 50 epoch dengan ukuran batch 16 titik. Model digunakan untuk melakukan prediksi pada data tes setelah pelatihan selesai. Kemudian, inverse transform dari normalisasi Min-Max digunakan untuk mengembalikan hasil prediksi ke skala asli. Kinerja model kemudian dapat dievaluasi dengan menggunakan metrik seperti Error Rata Rata Bulat (RMSE) dan Error Rata Rata Bulat (MAE). Oleh karena itu, model LSTM yang dibangun memiliki kemampuan untuk membuat prediksi berbasis deret waktu dan mempelajari pola historis data penjualan.



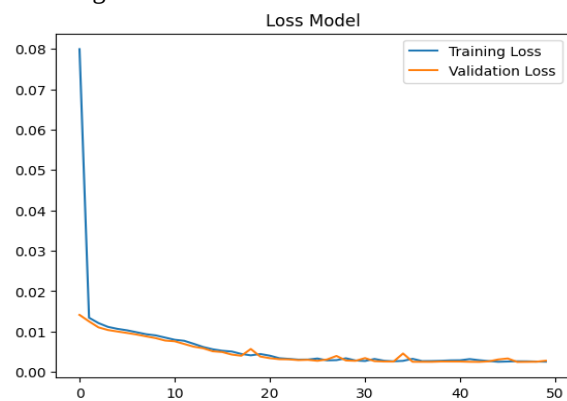
Gambar 3. Arsitektur Model LSTM

Gambar 3 menunjukkan arsitektur model LSTM yang digunakan dalam penelitian ini. Model terdiri mulai dari lapisan input dengan 7 time step, diikuti oleh lapisan LSTM dengan 50 neuron yang berfungsi untuk mengekstraksi pola temporal. Selanjutnya, output diteruskan ke lapisan dense untuk menghasilkan nilai prediksi.

4.5. Pelatihan dan Evaluasi

Tahap pelatihan model dilakukan untuk melatih model LSTM dalam mempelajari pola data penjualan berdasarkan dataset yang telah diproses sebelumnya. Proses pelatihan menggunakan data training sebanyak 578 data dengan parameter berupa 50 epoch dan batch size sebesar 16. Model dilatih menggunakan optimizer Adam dan fungsi loss MSE untuk meminimalkan kesalahan prediksi.

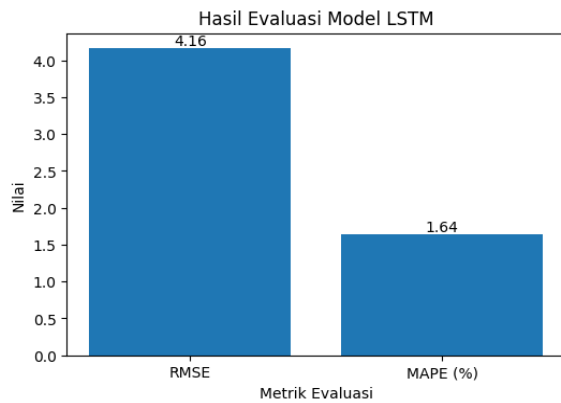
Selama proses pelatihan, model melakukan pembaruan bobot secara iteratif pada setiap epoch. Hasil pelatihan menunjukkan bahwa nilai loss mengalami penurunan secara bertahap, yang menandakan bahwa model mampu mempelajari pola data dengan baik.



Gambar 4. Loss model

Gambar 4 menunjukkan bahwa seiring bertambahnya waktu, nilai kehilangan pelatihan dan validasi menurun. Model memiliki kemampuan untuk mempelajari pola data dengan benar dan tidak mengalami overfitting yang signifikan. Setelah pelatihan selesai, fase evaluasi model dilakukan dengan 145 data pengujian. Tujuan dari evaluasi ini adalah untuk mengetahui seberapa baik model melakukan prediksi pada data yang belum pernah dilihat sebelumnya. Penelitian ini menggunakan metrik RMSE dan MAPE untuk mengukur kinerja model. RMSE mengukur kesalahan rata-rata prediksi dalam satuan asli, sedangkan MAPE digunakan untuk mengukur tingkat kesalahan.

Berdasarkan hasil evaluasi, diperoleh nilai RMSE sebesar 4.16 dan MAPE sebesar 1.64%. Nilai MAPE yang berada jauh di bawah 10% berarti dapat disimpulkan bahwa model yang dihasilkan memiliki tingkat akurasi yang tinggi. Serta model LSTM mampu untuk menangkap pola data penjualan dengan baik.

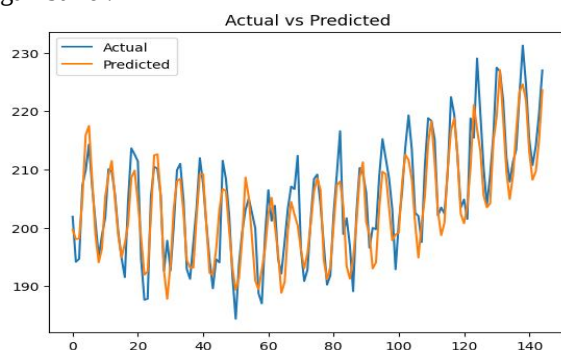


Gambar 5. Hasil Evaluasi Model LSTM

4.6. Prediksi

Tahap prediksi dilakukan setelah model LSTM selesai melalui proses pelatihan dan evaluasi. Pada tahap ini, model digunakan untuk melakukan prediksi terhadap data pengujian sebanyak 145 data. Input yang digunakan merupakan data hasil transformasi *time series* dengan *window size* 7, sehingga setiap nilai prediksi dihasilkan berdasarkan 7 data historis sebelumnya. Hasil prediksi yang diperoleh masih berada dalam skala normalisasi (0–1), sehingga dilakukan proses denormalisasi untuk mengembalikan nilai ke skala asli penjualan. Proses ini bertujuan agar hasil prediksi dapat diinterpretasikan secara langsung dalam satuan jumlah produk.

Berdasarkan hasil yang diperoleh, model LSTM mampu menghasilkan prediksi yang mengikuti pola data aktual dengan baik, baik pada kondisi tren naik maupun turun. Meskipun terdapat perbedaan pada beberapa titik, selisih yang terjadi relatif kecil dan masih dalam batas toleransi, yang ditunjukkan dengan nilai MAPE sebesar 1.64%. Dengan demikian, model LSTM terbukti mampu menangkap pola temporal dalam data historis dan dapat membantu dalam melakukan prediksi permintaan produk pada periode berikutnya secara akurat. Hasil prediksi kemudian dibandingkan dengan data aktual untuk melihat tingkat kesesuaian antara keduanya, yang ditampilkan pada gambar 6.



Gambar 6. Perbandingan data aktual dan hasil prediksi

Grafik pada Gambar 6 menunjukkan bahwa pola data aktual dengan baik, baik pada kondisi tren

meningkat maupun menurun yang dihasilkan dari model LSTM. Hal ini terlihat dari kedekatan antara garis prediksi dan data aktual pada sebagian besar titik pengamatan. Meskipun terdapat perbedaan pada beberapa titik, selisih yang terjadi relatif kecil dan masih dalam batas toleransi, yang diperkuat dengan nilai MAPE sebesar 1.64%. Dengan demikian model yang dihasilkan memiliki kemampuan yang sangat baik dalam menangkap pola temporal serta fluktuasi data penjualan.

5. KESIMPULAN DAN SARAN

Kesimpulan dari penelitian ini bahwa bahwa model LSTM efektif digunakan untuk memprediksi penjualan produk ritel berbasis data deret waktu. Tahapan pengolahan data yang meliputi preprocessing, pembentukan dataset dengan teknik sliding window, serta pemodelan LSTM mampu menghasilkan model yang dapat menangkap pola temporal dari data historis. Hasil evaluasi menunjukkan nilai RMSE sebesar 4,16 dan MAPE sebesar 1,64%. Nilai MAPE yang berada di bawah 10% menunjukkan model memiliki tingkat kesalahan yang sangat rendah dan kinerja prediksi yang sangat baik. Selain itu, hasil prediksi mampu mengikuti pola data aktual, baik pada kondisi tren meningkat maupun menurun. Dengan demikian, model LSTM yang dibangun dapat dimanfaatkan sebagai alat bantu prediksi permintaan produk secara akurat, sehingga mendukung pengambilan keputusan dalam pengelolaan dan perencanaan persediaan.

DAFTAR PUSTAKA

- [1] E. Fatmawati and E. Sudarmilah, "Forecasting Permintaan Persediaan Berbasis Long Short-Term Memory untuk Penentuan Safety Stock dan Reorder Point," vol. 7, 2026, doi: 10.30865/json.v7i3.9648.
- [2] K. Sweeney, J. Riley, and Y. Duan, "Product variety in retail: the moderating influence of demand variability," *Int. J. Phys. Distrib. Logist. Manag.*, vol. 52, no. 4, pp. 351–369, 2021, doi: <https://doi.org/10.1108/IJPDLM-12-2020-0407>.
- [3] S. F. W. T. Lim, E. Rabinovich, S. Lee, and S. Park, "Estimating Stockout Costs and Optimal Stockout Rates: A Case on the Management of Ugly Produce Inventory," *Manage. Sci.*, vol. 71, no. 4, pp. 3106–3126, Jul. 2024, doi: 10.1287/mnsc.2021.03174.
- [4] F. Fikri Al Khafid, F. Bilwildi Emyu, S. Indhah Ariescahyani, and R. Maulana Effendi, "Analisis Kinerja Rantai Pasok Studi Kasus Pada Pt Nidec Instruments Bekasi," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 9, no. 1, pp. 1011–1016, 2024, doi: 10.36040/jati.v9i1.12572.
- [5] A. Maulana, "Transformasi Digital, Inovasi Terbuka, dan Pengambilan Keputusan Berbasis Data: Implikasi bagi Strategi Bisnis," *Koaliansi Coop. J.*, vol. 5, no. 1, pp. 55–70, 2025.
- [6] A. Martono, Padel, M. iip Suhaepi, and A. R. T.

- Wulandari, "Evaluasi Performa Model ARIMA dan LSTM dengan Permintaan Inventarisasi Produk pada Perusahaan Produksi," vol. 11, no. 10, pp. 165–173, 2025.
- [7] U. J. I. Performa, D. A. N. Perbandingan, and R. Mysql, "Jurnal Informatika Terpadu HIVE-HADOOP," *J. Inform. Terpadu*, vol. 8, no. 1, pp. 1–7, 2022.
- [8] N. J. Salsabila, "Eksistensi Metode ARIMA, SARIMA dan LSTM dalam Memprediksi Penjualan," *JUSTIN J. Sist. dan Teknol. Inf.*, vol. 13, no. 04, pp. 399–408, 2025, doi: 10.26418/justin.v13i4.76892.
- [9] X. Kong *et al.*, *Deep learning for time series forecasting: a survey*, vol. 16, no. 7–8. Springer Berlin Heidelberg, 2025. doi: 10.1007/s13042-025-02560-w.
- [10] H. Ismail Fawaz, G. Forestier, J. Weber, L. Idoumghar, and P. A. Muller, "Deep learning for time series classification: a review," *Data Min. Knowl. Discov.*, vol. 33, no. 4, pp. 917–963, 2019, doi: 10.1007/s10618-019-00619-1.
- [11] B. Lim, S. Zohren, and B. Lim, "Time-series forecasting with deep learning : a survey Subject Areas : Author for correspondence :," no. April, 2021, doi: 10.1098/rsta.2020.0209/248905/rsta.2020.0209.pdf.
- [12] K. Bandara, C. Bergmeir, and H. Hewamalage, "LSTM-MSNet: Leveraging Forecasts on Sets of Related Time Series With Multiple Seasonal Patterns," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 32, no. 4, pp. 1586–1599, 2021, doi: 10.1109/TNNLS.2020.2985720.
- [13] S. Wang, C. Li, and A. Lim, "Why Are the ARIMA and SARIMA not Sufficient," pp. 1–10, 2021, [Online]. Available: <http://arxiv.org/abs/1904.07632>
- [14] K. M. Saini and A. Sharma, "Impact of data preprocessing on Prophet–LSTM hybrid models," *Educ. Adm. Theory Pract.*, vol. 30, no. 5, pp. 14647–14650, 2024, doi: 10.53555/kuey.v30i5.7207.
- [15] F. Grimm, G. Bitsch, and C. van Dinther, "The Influence of Preprocessing on Time Series Prediction," *Procedia Comput. Sci.*, vol. 277, pp. 739–748, 2026, doi: <https://doi.org/10.1016/j.procs.2026.02.115>.
- [16] A. S. AlSalehy and M. Bailey, "Improving Time Series Data Quality: Identifying Outliers and Handling Missing Values in a Multilocation Gas and Weather Dataset," *Smart Cities*, vol. 8, no. 3, pp. 1–39, 2025, doi: 10.3390/smartcities8030082.