



ANALISIS LITERATUR PREDIKSI TREN PENJUALAN E-COMMERCE BERBASIS DATA TIME-SERIES: METODE STATISTIK & MACHINE LEARNING

Maya Rafika Utami

mayarapikautami@gmail.com

Universitas Islam Negeri Sumatera Utara

Muhammad Irwan Padli Nasution

irwannst@uinsu.ac.id

Universitas Islam Negeri Sumatera Utara

Alamat: Jl. IAIN No.1, Gaharu, Kec. Medan Tim., Kota Medan, Sumatera Utara 20235

Korespondensi penulis: *mayarapikautami@gmail.com*

Abstract. *The rapid development of the e-commerce industry in Indonesia has generated a huge amount of time-series sales data, creating a significant opportunity to predict sales trends. This study aims to analyze the existing literature on methods used to predict e-commerce sales trends by utilizing time-series data from databases. The analysis shows that various time-series techniques, such as ARIMA, LSTM, and hybrid models, have been applied with varying degrees of success. Factors such as data volume, data pattern complexity, and the need for real-time prediction play an important role in the selection and effectiveness of prediction methods. Overall, this study concludes that hybrid models and machine learning, especially LSTM, show great potential in improving the accuracy of sales trend prediction in the e-commerce domain.*

Kata Kunci: E-commerce, Prediksi Tren Penjualan, Data Time-Series, Analisis Literature, Machine Learning

Abstrak. Perkembangan pesat industri *e-commerce* di Indonesia telah menghasilkan jumlah data penjualan *time-series* yang sangat besar, menciptakan peluang yang signifikan untuk memprediksi tren penjualan. Penelitian ini bertujuan untuk menganalisis literatur yang ada mengenai metode-metode yang digunakan untuk memprediksi tren penjualan *e-commerce* dengan memanfaatkan data *time-series* dari database. Hasil analisis menunjukkan bahwa berbagai teknik *time-series*, seperti ARIMA, LSTM, dan model *hibryd*, telah diterapkan dengan tingkat keberhasilan yang bervariasi. Faktor-faktor seperti volume data, kompleksitas pola data, dan kebutuhan akan prediksi *real-time* memainkan peran penting dalam pemilihan dan efektivitas metode prediksi. Secara keseluruhan, penelitian ini menyimpulkan bahwa model *hybrid* dan *machine learning*, terutama LSTM, menunjukkan potensi yang besar dalam meningkatkan akurasi prediksi tren penjualan di ranah *e-commerce*.

Keywords: *E-commerce, Sales Trend Prediction, Time-Series Data, Literature Analysis, Machine Learning*

PENDAHULUAN

Perkembangan teknologi informasi yang sangat pesat, khususnya internet, telah membawa perubahan besar dalam lanskap bisnis global, terutama di sektor perdagangan digital atau *e-commerce*. Di Indonesia, *e-commerce* telah menjadi salah satu pilar utama dalam ekonomi digital, dengan kontribusi terhadap PDB yang terus menunjukkan peningkatan. Namun, tantangan seperti kesenjangan akses digital dan regulasi yang belum memadai masih perlu segera diatasi agar tren pertumbuhan positif ini dapat terus berlanjut. Pada tahun 2021, *e-commerce* menyumbang sebesar 6,1% terhadap PDB,

meningkat menjadi 7,8% pada tahun 2022, dan diperkirakan mencapai 8,9% pada tahun 2023. Pertumbuhan ini mencerminkan kemajuan *e-commerce* di Indonesia yang didorong oleh pengembangan infrastruktur digital dan meningkatnya kepercayaan masyarakat. (Lubis et al., n.d.)

Dengan berkembangnya *e-commerce*, aktivitas bisnis yang berlangsung di dalamnya menghasilkan volume data yang sangat besar, yang merefleksikan pola transaksi, perilaku pelanggan, serta tren pasar yang sedang berlangsung. Salah satu jenis data yang paling krusial dalam analisis bisnis *e-commerce* adalah data deret waktu (*time-series*), yaitu kumpulan data yang dicatat secara berkala untuk memantau perkembangan suatu aktivitas dari waktu ke waktu. Data ini memiliki peran penting dalam mengungkap tren penjualan, meramalkan permintaan produk, serta menyempurnakan strategi pemasaran dan pengelolaan persediaan. Meski demikian, analisis data *time-series* memiliki tantangan tersendiri karena sifatnya yang kompleks, sebab pola penjualan dapat dipengaruhi oleh beragam faktor seperti musiman (*seasonality*), tren jangka panjang, serta elemen eksternal seperti promosi atau perubahan dalam kebijakan pasar. Untuk mengatasi tantangan ini, diperlukan solusi teknologi berupa *Time Series Database* (TSDB), yang dirancang untuk mengelola dan menganalisis data *time-series* dalam skala besar. (Noor et al., 2017)

Sejumlah pendekatan telah dikembangkan untuk memprediksi tren penjualan berbasis data deret waktu (*time-series*), mulai dari teknik statistik konvensional seperti *Autoregressive Integrated Moving Average* (ARIMA), hingga pendekatan berbasis kecerdasan buatan seperti *Long Short-Term Memory* (LSTM). Meskipun sejumlah studi telah dilakukan dalam bidang ini, tantangan masih muncul dalam menentukan metode yang paling efektif dan akurat, khususnya dalam konteks *e-commerce* yang ditandai oleh karakteristik data yang sangat dinamis dan cepat mengalami perubahan. Di samping itu, kajian yang secara sistematis membahas kelebihan dan kekurangan dari berbagai metode prediksi dalam beragam skenario bisnis *e-commerce* masih relatif terbatas.

Oleh sebab itu, tujuan dari kajian ini adalah untuk melakukan analisis literatur yang berkaitan dengan metode prediksi tren penjualan *e-commerce* berbasis data deret waktu (*time-series*) yang bersumber dari berbagai database. Fokus utama kajian literatur ini meliputi identifikasi metode prediksi yang paling efektif dan dominan digunakan dalam *e-commerce*, pengungkapan faktor-faktor yang memengaruhi tingkat akurasi

prediksi, membandingkan efektivitas masing-masing metode berdasarkan konteks. (Ikhlasil Akmal et al., 2024)

METODE PENELITIAN

Penelitian ini menggunakan metode analisis literatur yang memanfaatkan media informasi publikasi pendidikan seperti google scholar, E-Book, jurnal ilmiah dan berbagai artikel yang berkenaan dengan judul jurnal ini. Seluruh sumber yang diperoleh ditelaah secara cermat berdasarkan relevansi, kredibilitas, dan kualitas informasi yang terkandung di dalamnya. Kajian literatur dilakukan secara komprehensif untuk membangun pemahaman menyeluruh tentang topik penelitian sekaligus memperkuat argumen yang diajukan dalam artikel ini.

HASIL DAN PEMBAHASAN

Hasil analisis literatur menunjukkan bahwa beragam metode deret waktu (*time-series*) telah digunakan dalam upaya memprediksi tren penjualan pada sektor *e-commerce*. Metode-metode tersebut mencakup:

A. Metode statistika

Terdapat berbagai metode statistika yang digunakan dalam prediksi penjualan. Namun, dari sekian banyak metode yang tersedia, kajian ini akan berfokus pada pencarian metode yang paling sesuai melalui perbandingan antara masing-masing pendekatan. Hal ini dikarenakan setiap metode menghasilkan prediksi serta tingkat variasi dalam penyimpangan. Dalam kajian ini, metode yang akan dibandingkan berdasarkan tingkat akurasi dalam memprediksi penjualan adalah *Autoregressive Integrated Moving Average* (ARIMA) dan *Exponential Smoothing*, mengingat keduanya memiliki kemampuan dalam menganalisis data historis yang bersifat stasioner, musiman, maupun bersiklus.

a. Metode Autoregressive Integrated Moving Average (ARIMA)

Merupakan salah satu model regresi linier yang paling banyak digunakan untuk meramalkan deret waktu yang bersifat statis, ARIMA juga dikenal sebagai metode deret waktu Box-Jenkins. Metode ini dikembangkan oleh George Box dan Gwilym Jenkins pada tahun 1970. ARIMA sangat efektif dalam melakukan peramalan deret waktu untuk jangka pendek, namun tingkat akurasinya cenderung menurun apabila diterapkan pada jangka waktu yang lebih panjang. Model ARIMA mencakup tiga komponen utama, yaitu

autoregressive (AR), *moving average* (MA), serta kombinasi dari keduanya yang dikenal sebagai *autoregressive moving average* (ARIMA). (Rahmadayanti et al., 2015) Terdapat beberapa tahapan yang perlu dilakukan sebelum menerapkan metode ARIMA untuk peramalan, antara lain:

1) Identifikasi Data (Uji Kestasioneran Data)

Kestasioneran mengacu pada kondisi di mana data deret waktu tidak menunjukkan keunikan atau penurunan nilai secara tajam dari waktu ke waktu. Pengujian ini dilakukan sebagai langkah awal untuk memastikan apakah data yang digunakan bersifat stasioner atau tidak. Suatu data dikatakan stasioner apabila memenuhi dua syarat, yaitu kestasioneran dalam varian dan kestasioneran dalam rata-rata (*mean*).

2) Estimasi Parameter

Estimasi parameter pada metode ARIMA dilakukan dengan menganalisis pola dari plot *Autocorrelation Function* (ACF) dan *Partial Autocorrelation Function* (PACF). Secara umum, nilai parameter AR (p), MA (q), atau kombinasi ARMA (p,q) jarang melebihi angka 2. Setelah nilai p dan q diidentifikasi, proses peramalan menggunakan metode ARIMA dapat dilanjutkan.

3) Peramalan

Setelah model yang paling sesuai berhasil diidentifikasi, tahap selanjutnya adalah melakukan peramalan. Dalam berbagai studi, metode ini sering kali dianggap lebih andal dibandingkan model-model ekonometrika konvensional. Meski demikian, efektivitas metode ini tetap memerlukan pengkajian lebih lanjut, khususnya bagi peneliti yang berniat mengadopsinya. Berdasarkan karakteristiknya, model deret struktural dinilai lebih tepat untuk peramalan dalam jangka panjang.

b. Metode *Exponential Smoothing*

Merupakan suatu metode peramalan yang berbasis pada proses penghitungan rata-rata data historis secara eksponensial, di mana perhitungan dilakukan secara berulang menggunakan data terbaru yang tersedia. Setiap titik data dalam deret waktu diberikan bobot tertentu, dengan penekanan lebih besar pada data yang paling baru. *Exponential Smoothing* sejatinya merupakan bentuk dari metode rata-rata bergerak, namun dengan alokasi bobot yang lebih besar pada observasi terkini dibandingkan data sebelumnya. Pendekatan ini sangat berguna apabila perubahan terbaru dalam data mencerminkan

perubahan nyata (seperti pola musiman), bukan sekedar fluktuasi acak dalam kasus terakhir, metode rata-rata bergerak sederhana sudah mencukupi.

Dalam *Exponential Smoothing*, terdapat satu atau lebih parameter pemulusan yang ditentukan secara eksplisit, dan pemilihan parameter ini akan menentukan distribusi bobot terhadap nilai-nilai observasi. Metode ini memiliki kemampuan untuk secara adaptif mengikuti tren yang terdapat dalam data aktual, kesalahan peramalan terakhir. Kesalahan dari periode sebelumnya digunakan sebagai dasar untuk mengoreksi ramalan selanjutnya ke arah yang berlawanan. Proses penyesuaian ini akan terus berlangsung hingga kesalahan tersebut berhasil diperbaiki.

Meskipun prinsip dasar dari metode ini terkesan sederhana, peranannya sangat penting dalam praktik peramalan. Bila diterapkan secara tepat, prinsip tersebut memungkinkan pengembangan suatu sistem peramalan yang bersifat *self-adjusting*, yakni mampu melakukan koreksi otomatis terhadap kesalahan peramalan yang terjadi sebelumnya.

B. Metode *Machine Learning*: LSTM & GRU

Merupakan kumpulan algoritma komputer yang dirancang untuk mengoptimalkan kinerja sistem atau komputer berdasarkan data sampel yang tersedia. Keunggulan utama dari pendekatan *machine learning* terletak pada kemampuannya untuk secara dinamis menyesuaikan dan memodifikasi keputusan sebagai respons terhadap perubahan kondisi. Seiring perkembangan teknologi, metode prediksi kini banyak dikembangkan menggunakan pendekatan *machine learning* karena performanya yang secara signifikan melampaui metode statistik tradisional.

a. *Long Short-Term Memory* (LSTM)

LSTM merupakan salah satu arsitektur *Recurrent Neural Network* (RNN) yang diperkenalkan oleh Hochreiter dan Schmidhuber pada tahun 1997, sebagai solusi atas permasalahan *vanishing gradient* dan *exploding gradient* yang sering terjadi pada RNN konvensional. Fenomena *vanishing gradient* terjadi ketika nilai *gradien* mengecil secara drastis di setiap lapisan hingga akhirnya tidak berdampak pada pembaruan bobot, menyebabkan model gagal konvergen dan tidak mampu mempelajari pola jangka panjang. Oleh karena itu, LSTM sangat ideal digunakan dalam tugas prediksi dan klasifikasi berbasis waktu. Arsitektur LSTM terdiri dari lapisan input, lapisan tersembunyi, dan lapisan output, di mana setiap blok menyimpan beberapa *memory cell*

dan tiga gates penting: *input*, *output*, dan *forget gate*. Komponen ini memungkinkan model untuk mempertahankan atau menghapus informasi, sehingga menentukan seberapa banyak informasi yang harus ditransmisikan ke sel selanjutnya. (Pambudi et al., 2020)

b. *Gated Recurrent Unit* (GRU)

GRU adalah salah satu varian populer dari LSTM yang diperkenalkan oleh Cho pada tahun 2014. Dibandingkan dengan LSTM, GRU menawarkan struktur komputasi yang lebih sederhana namun tetap memiliki akurasi yang kompetitif. GRU mengatasi masalah *vanishing gradient* melalui dua gates dan tiga fungsi aktivasi yang lebih efisien. Karena kompleksitas arsitektur yang lebih rendah, GRU dapat memproses data dalam jumlah besar lebih cepat, menjadikannya sangat sesuai untuk dataset yang jumlahnya kecil dengan tetap mempertahankan akurasi tinggi. (Agusmawati, N. K et al., 2020)

C. Metode *Hybrid*: Kombinasi Statistik dan Machine Learning

Metode hybrid yang mengombinasikan metode *Autoregressive Integrated Moving Average* (ARIMA) dan LSTM menjadi solusi yang menjanjikan dalam peramalan data deret waktu. ARIMA dikenal sebagai metode statistik yang efektif untuk peramalan jangka pendek dengan karakteristik data linier. Metode ini fleksibel, memiliki akurasi tinggi, dan dapat diandalkan, tetapi kinerjanya menurun secara signifikan dalam peramalan jangka panjang. Untuk mengatasi keterbatasan ini, ARIMA dikombinasikan dengan LSTM adalah kemampuannya menangani data non-linear, namun memiliki kekurangan dalam waktu pelatihan yang lama dan parameter yang kompleks.

ARIMA dan LSTM masing-masing memiliki kekuatan dan kelemahan. ARIMA cenderung kurang optimal dalam menangani data non-linear, sementara LSTM dan jaringan saraf tiruan lainnya mampu memproses baik data linear maupun non-linear. Oleh karena itu, penggabungan keduanya dirancang untuk memanfaatkan keunggulan masing-masing serta mengatasi keterbatasan yang ada, di mana ARIMA digunakan untuk memodelkan trend dan LSTM menangani residual. Proses pelatihan model gabungan ARIMA-LSTM melibatkan beberapa tahapan seperti menyiapkan data dan melatih model ARIMA, lalu gunakan *residual error* ARIMA yang telah di pra-proses sebagai data untuk membangun dan melatih model LSTM. (Rowan et al., 2022)

Setelah fase pelatihan model selesai, langkah berikutnya adalah melakukan peramalan dan pengujian menggunakan data pengujian. Beberapa faktor krusial yang

memengaruhi seberapa baik prediksi tren penjualan dalam *e-commerce* berbasis data deret waktu bekerja adalah:

1. **Kuantitas Data:** Semakin besar himpunan data dan historis yang tersedia, potensi akurasi prediksi akan meningkat. Model *machine learning* dan model statistika, seperti *Long Short-Term Memory* (LSTM), memerlukan volume sampel yang signifikan untuk dapat mengenali pola musiman, tren jangka panjang, serta lonjakan permintaan yang tidak terduga.
2. **Kerumitan Pola Data:** Pola data dalam konteks *e-commerce* seringkali bersifat non-linear dan dipengaruhi oleh beragam elemen, termasuk tren musiman, kampanye diskon, atau perubahan perilaku konsumen yang terjadi dengan cepat. Metode sederhana seperti regresi linear menjadi kurang memadai untuk menangkap kompleksitas pola ini. Oleh karena itu, diterapkan *machine learning* atau bahkan model *hybrid*.
3. **Urgensi Prediksi Waktu Nyata:** Dalam ekosistem *e-commerce* yang dinamis, bisnis memerlukan prediksi yang dapat dieksekusi secara *real-time* untuk mengoptimalkan penetapan harga, manajemen stok, dan pelaksanaan kampanye pemasaran. Model prediksi yang digunakan harus memiliki latensi yang rendah dan mampu memproses informasi dalam orde detik atau menit. (Ralahallo & Muhrim, 2024)

Analisi literatur menunjukkan bahwa metode *machine learning*, terutama LSTM, memperlihatkan potensi signifikan dalam meningkatkan keakuratan prediksi tren penjualan *e-commerce*. LSTM memiliki kemampuan untuk menangkap pola temporal yang kompleks dalam data deret waktu, yang seringkali sulit dicapai dengan metode statistik tradisional. Metode *hybrid*, yang mengkombinasikan metode statistik dan *machine learning*, juga menunjukkan hasil yang menjanjikan. Pendekatan-pendekatan ini memanfaatkan kekuatan unik dari setiap metode untuk menghasilkan prediksi yang lebih tepat.

KESIMPULAN

Dalam kajian ini ditemukan bahwa ada berbagai cara untuk memperkirakan tren penjualan *e-commerce* menggunakan data deret waktu, baik dengan metode statistika maupun *machine learning*. Metode statistika seperti ARIMA dan *Exponential Smoothing* masih berguna, terutama untuk data yang polanya lurus dan stabil. Namun, karena data

menjadi lebih rumit dan kebutuhan akan prediksi yang lebih cepat dan adaptif meningkat, metode *machine learning*, khususnya LSTM lebih efektif. Selain itu, menggabungkan metode statistika dan *machine learning* menjadi model *hybrid*, dapat mengurangi kekurangan masing-masing metode. Model ini memanfaatkan keunggulan ARIMA dalam memproses pola lurus jangka pendek dan kekuatan LSTM dalam memahami dinamika data non-linear dan jangka panjang. Faktor-faktor seperti kuantitas data, kerumitan perilaku konsumen, dan kebutuhan akan prediksi waktu nyata adalah hal penting yang harus diperhatikan saat memilih metode prediksi. Oleh karena itu, pemahaman yang baik tentang karakteristik data dan konteks bisnis sangat diperlukan agar keputusan dapat didukung oleh sistem prediksi yang sesuai.

DAFTAR PUSTAKA

- Ikhlasul Akmal, R., Hekmatiar Ayyubi, R., & Hidayansyah, I. (2024). Analisis Peramalan Jumlah Pengguna E-Commerce Dengan Metode Time Series. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(4), 7674–7679.
<https://doi.org/10.36040/jati.v8i4.9898>
- Lubis, N., Harahap, A. Y., Tantawi, R., Aslami, N., Sitanggang, T. N., Samudra, U., Islam, U., Sumatera, N., & Indonesia, U. P. (n.d.). *Dampak Perkembangan Ekonomi Digital terhadap Pertumbuhan Sektor E-Commerce di Indonesia : Perspektif*. 348–359.
- Noor, S., Naqvi, Z., Yfantidou, S., & Zi, E. (2017). *Universi e libre de Bruxelles Advanced Databases Time Series Databases and InfluxDB*. 000455274.
http://cs.ulb.ac.be/public/_media/teaching/influxdb_2017.pdf
- Pambudi, H. K., Kusuma, P. G. A., Yulianti, F., & Julian, K. A. (2020). Prediksi Status Pengiriman Barang Menggunakan Metode Machine Learning. *Jurnal Ilmiah Teknologi Infomasi Terapan*, 6(2), 100–109.
<https://doi.org/10.33197/jitter.vol6.iss2.2020.396>
- Pratama, A. S., Sari, S. M., Hj, M. F., Badwi, M., & Anshori, M. I. (2023). Pengaruh Artificial Intelligence, Big Data Dan Otomatisasi Terhadap Kinerja SDM Di Era Digital. *Jurnal Publikasi Ilmu Manajemen (JUPIMAN)*, 2(4), 108–123.
- Rahmadayanti, R., Susilo, B., & Puspitaningrum, D. (2015). Perbandingan Keakuratan Metode Autoregressive Integrated Moving Average Dan Exponential Smoothing

Pada Peramalan Penjualan Semen Di PT Sinar Abadi. *Jurnal Rekursif*, 3(1), 23–36.

Ralahallo, B. A. B., & Muhrim, M. R. (2024). Dampak Strategi Pemasaran Digital Terhadap Pertumbuhan Bisnis E-Commerce Di Indonesia. *HIPOTESA- Jurnal Ilmu-Ilmu Sosial*, 18(1), 71–83.

Rowan, Muflikhah, L., & Cholissodin, I. (2022). Peramalan Kasus Positif COVID-19 di Jawa Timur menggunakan Metode Hybrid ARIMA-LSTM. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 6(9), 4146–4153.
<http://j-ptiik.ub.ac.id>

Agusmawati, N. K et al., (2020). Prediksi Harga Emas Menggunakan Metode LSTM & GRU. *JITET*, 8(1), 32–36. <https://doi.org/10.23960/jitet.v11i3.3250>