

Peramalan Volume Transaksi Uang Elektronik Menggunakan Hybrid SARIMA-LSTM dengan Integrasi Klasifikasi Berita IndoBERT

Nabila Sya'bani Wardana*, Wiwik Anggraeni

Fakultas Teknologi Elektro dan Informatika Cerdas, Program Studi Sistem Informasi, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

Email: ^{1,*}6026241014@student.its.ac.id, ²wiwik@its.ac.id

Email Penulis Korespondensi: 6026241014@student.its.ac.id

Submitted: 18/06/2026; Accepted: 30/06/2026; Published: 30/06/2026

Abstrak—Di tengah percepatan transformasi pembayaran digital, volume transaksi uang elektronik di Indonesia tumbuh pesat dengan pola yang fluktuatif dan kerap mengalami perubahan struktural akibat faktor eksternal. Kondisi ini menjadikan peramalan sebagai tantangan tersendiri, padahal proyeksi yang akurat dibutuhkan untuk mendukung pemantauan dan perencanaan ke depan. Pendekatan yang ada umumnya bergantung pada model univariat atau variabel makroekonomi yang mengalami keterlambatan publikasi, sehingga kurang responsif terhadap guncangan ekonomi yang cepat. Untuk mengatasi keterbatasan tersebut, penelitian ini mengusulkan kerangka hybrid SARIMA-LSTM yang diperkaya dengan hasil klasifikasi berita daring berbasis IndoBERT sebagai variabel eksogen, karena teks berita mampu menyediakan sinyal yang lebih timely dibandingkan indikator makroekonomi konvensional. Sekitar 13.000 judul berita dikumpulkan melalui *web scraping* dan diklasifikasikan ke dalam tiga kategori dampak pendorong, penghambat, dan informatif menggunakan IndoBERT yang di-fine-tune untuk bahasa Indonesia. Hasil klasifikasi diagregasi bulanan menjadi fitur eksogen, dengan SARIMA menangkap tren dan pola musiman sedangkan LSTM memodelkan residual nonliniernya. Model terbaik, yaitu hybrid SARIMA-LSTM dengan variabel berita informatif, mencapai MAPE 14,42%, lebih baik dibandingkan hybrid tanpa eksogen (15,24%), SARIMA tunggal (17,77%), dan LSTM tunggal (23,37%). Kontribusi utama penelitian ini adalah menguji performa arsitektur *hybrid* SARIMA-LSTM sekaligus mengevaluasi integrasi klasifikasi berita daring sebagai indikator pelengkap dalam peramalan, guna mengurangi ketergantungan pada data yang memiliki kendala *publication lag*. Meskipun perbaikannya tergolong marginal dan lebih menonjol pada periode gejolak, keunggulannya yang konsisten mengindikasikan bahwa berita informatif mengandung sinyal nonlinier yang bermanfaat bagi model peramalan.

Kata Kunci: IndoBERT; Klasifikasi Berita; Peramalan; SARIMA-LSTM; Uang Elektronik

Abstract—Amid the accelerating transformation of digital payments, the volume of electronic money transactions in Indonesia has grown rapidly, exhibiting volatile patterns and frequent structural breaks. These characteristics make forecasting particularly challenging, even though accurate projections are needed to support monitoring and planning. Existing approaches generally rely on univariate models or macroeconomic variables that suffer from publication lag, rendering them less responsive to rapidly evolving economic shocks. To address these limitations, this study proposes a hybrid SARIMA-LSTM framework that incorporates online news classification, generated using IndoBERT, as an exogenous variable providing more timely signals than macroeconomic indicators. Approximately 13,000 news headlines were collected through web scraping and classified into three impact categories *pendorong* (driving), *penghambat* (inhibiting), and *informatif* (informative) using a fine-tuned IndoBERT model. The classification results were aggregated monthly into exogenous features, with SARIMA capturing the linear trend and seasonal patterns while LSTM modeled the nonlinear residuals. The best-performing model, a hybrid SARIMA-LSTM with the informative news variable, achieved a MAPE of 14.42%, outperforming the hybrid model without exogenous variables (15.24%), the standalone SARIMA (17.77%), and the standalone LSTM (23.37%). The main contribution is to evaluate the hybrid SARIMA-LSTM architecture and assess whether integrating online news classification can serve as a complementary indicator, reducing reliance on lag-affected data. Although the improvement from the news variable was marginal and more pronounced during volatile periods, its consistent superiority indicates that informative news carries nonlinear signals beneficial for the forecasting model.

Keywords: IndoBERT; News Classification; Forecasting; SARIMA-LSTM; Electronic Money

1. PENDAHULUAN

Pesatnya perkembangan teknologi dalam beberapa tahun terakhir telah mendorong transformasi digital, salah satunya pada sektor teknologi finansial (*fintech*) [1]. Kemudahan yang ditawarkan oleh instrumen digital tersebut mendorong adopsi uang elektronik secara masif, sekaligus mengubah preferensi masyarakat yang semula berbasis tunai beralih ke transaksi nontunai [2]. Laporan *Blueprint Sistem Pembayaran Indonesia* menunjukkan volume transaksi ritel melalui kanal digital mencapai 25,3 miliar transaksi pada tahun 2023, lebih dari dua kali lipat dibandingkan akhir 2019 [3]. Tren pertumbuhan ini diproyeksikan terus berlanjut. Riset tahunan CELIOS memperkirakan nilai pembayaran digital akan mencapai Rp2.491 triliun pada 2024 dan meningkat hingga Rp2.908 triliun pada 2025 [4]. Peningkatan penggunaan uang elektronik ini memiliki peran penting karena terbukti mampu mempercepat perputaran uang, mempermudah aktivitas ekonomi masyarakat, serta memberikan manfaat pada produktivitas keuangan dan pertumbuhan ekonomi [5]. Mengingat dampaknya yang cukup signifikan terhadap perekonomian, analisis dan peramalan tren volume transaksi uang elektronik di masa depan perlu dilakukan. Langkah ini diharapkan dapat digunakan sebagai landasan dan alat bantu bagi otoritas keuangan dalam mengevaluasi kebijakan serta merencanakan mitigasi risiko pada sistem pembayaran [6], [7], [8].

Meskipun memiliki tren pertumbuhan yang masif, data historis volume transaksi uang elektronik tidak selalu linier dan sangat rentan terhadap fluktuasi. Beberapa penelitian mencatat bahwa pola transaksi uang elektronik di Indonesia cenderung tidak stabil akibat berbagai guncangan eksternal, seperti perubahan drastis pada mobilitas

masyarakat dan kebijakan pembatasan selama masa pandemi COVID-19 [8], [9]. Ketidakstabilan pola data tersebut menyebabkan proses peramalan menjadi semakin kompleks. Di sisi lain, pendekatan peramalan ekonomi konvensional masih bergantung pada indikator statistik resmi yang memiliki keterlambatan publikasi (*publication lag*). Data volume transaksi sistem pembayaran yang dipublikasikan Bank Indonesia melalui portal Statistik Sistem Pembayaran (SPIP), misalnya, baru tersedia sekitar 3 bulan setelah periode transaksinya berlangsung. Bank Indonesia sendiri mengakui adanya kendala *time lag* pada produksi statistik resmi, yang mendorong pemanfaatan indikator alternatif yang lebih *timely*. Keterlambatan ini menjadi sangat problematik pada masa krisis yang bergerak cepat, ketika data yang tertinggal kurang andal untuk memantau kondisi ekonomi secara tepat waktu, sehingga menyulitkan otoritas moneter dan agen ekonomi dalam merumuskan langkah mitigasi secara responsif [10].

Sejauh ini, sejumlah penelitian telah memodelkan dan meramalkan volume transaksi uang elektronik melalui berbagai pendekatan. Pradnyandita dan Rohmawati [6], memodelkan transaksi uang elektronik di Indonesia menggunakan *Support Vector Regression (SVR)* dan *Vector Autoregressive Moving Average (VARMA)*, dengan memanfaatkan variabel moneter sebagai prediktor; namun, variabel makroekonomi semacam ini memiliki keterlambatan publikasi sehingga model sulit merespons perubahan secara cepat. Andreas dkk. [8], menerapkan model ARIMAX-GARCH untuk menangkap pola volatilitas transaksi uang elektronik dengan memasukkan variabel struktural dan *dummy* peristiwa, tetapi pendekatan *dummy* bersifat statis dan *post-hoc* sehingga hanya mampu merepresentasikan satu peristiwa spesifik yang telah diketahui sebelumnya. Sementara itu, Mahmudah dkk [7] dan Maulida dkk [11] menggunakan pendekatan berbasis *machine learning* dan *neural network*, seperti *Extreme Learning Machine (ELM)*, *Multilayer Perceptron (MLP)*, dan *Neural Network Autoregression (NNETAR)*, yang terbukti akurat dalam menangkap pola *non-linear*; meskipun demikian, kedua studi tersebut sepenuhnya mengandalkan data historis univariat tanpa memanfaatkan sinyal informasi eksternal. Berdasarkan uraian tersebut, penelitian peramalan transaksi uang elektronik di Indonesia hingga saat ini masih menyisakan dua keterbatasan utama. Pertama, mayoritas model bergantung pada data historis univariat berupa variabel makroekonomi yang memiliki *publication lag*, sehingga kurang responsif terhadap guncangan yang bergerak cepat. Kedua, penelitian yang telah melibatkan variabel eksogen masih menggunakan sumber yang bersifat statis (*dummy* peristiwa) atau *lagging* (indikator makroekonomi), dan belum memanfaatkan sumber informasi yang bersifat *timely* serta mampu menangkap beragam peristiwa secara dinamis. Untuk mengatasi celah tersebut, pemanfaatan teks berita daring (*online*) dapat menjadi alternatif solusi. Berbagai literatur terkini membuktikan bahwa ekstraksi sentimen dari artikel berita dapat berfungsi sebagai indikator awal (*leading indicator*) yang mampu menangkap dinamika kondisi ekonomi serta sentimen publik secara lebih cepat, bahkan melampaui kemampuan indikator statistik tradisional [12], [13].

Dalam mengekstraksi informasi dan polaritas arah dampak dari teks berita berbahasa Indonesia, pemanfaatan model berbasis *transformer* seperti IndoBERT dinilai sebagai pendekatan yang optimal. IndoBERT dikembangkan secara spesifik menggunakan korpus bahasa Indonesia berskala masif [14] dan telah terbukti efektif pada berbagai tugas klasifikasi teks berbahasa Indonesia [15], [16]. Di sisi lain, untuk menangani peramalan data volume transaksi uang elektronik yang kompleks dan volatil, penggunaan model tunggal sering kali memiliki keterbatasan [17]. Oleh karena itu, pendekatan *hybrid* yang menggabungkan metode statistik konvensional dan *deep learning* menjadi solusi yang direkomendasikan. Model statistik seperti SARIMA efektif dalam menangkap komponen linier dan pola musiman pada data deret waktu [18], sementara jaringan *Long Short-Term Memory (LSTM)* unggul dalam memodelkan sisaan (residu) *non-linier* serta dependensi jangka panjang [19]. Keunggulan pendekatan *hybrid* semacam ini telah dikonfirmasi pada berbagai domain. Du dkk. [20] menunjukkan bahwa model ARIMA-LSTM lebih unggul dibandingkan model tunggal ARIMA maupun LSTM dalam peramalan penyakit bawaan makanan, sementara Dzulfikar [21] memperoleh hasil serupa menggunakan SARIMA-LSTM pada konteks rekayasa industri. Meskipun demikian, kedua studi tersebut belum melibatkan variabel eksogen dan tidak diterapkan pada domain ekonomi maupun keuangan. Pada ranah pemanfaatan teks, Barbaglia dkk. [22] membuktikan bahwa sentimen yang diekstraksi dari korpus berita berbahasa Inggris memiliki kandungan informasi yang relevan sebagai prediktor variabel makroekonomi, sedangkan Salsabiila dan Anggraeni [23] menunjukkan bahwa pendekatan dua tahap yang menggabungkan klasifikasi teks dengan model peramalan memberikan hasil menjanjikan pada prediksi penyebaran demam berdarah berbasis media sosial. Berbagai temuan tersebut menunjukkan keunggulan masing-masing pendekatan, sekaligus menegaskan adanya celah penelitian yang belum terjawab. Studi yang memanfaatkan pendekatan *hybrid* umumnya belum mengintegrasikan variabel eksogen, sementara studi yang memanfaatkan teks berita umumnya berhenti pada konteks makroekonomi atau domain non-keuangan. Hingga saat ini, belum ditemukan studi yang mengintegrasikan klasifikasi teks berita berbasis model *transformer* sebagai variabel eksogen ke dalam kerangka *hybrid* SARIMA-LSTM, khususnya dalam konteks peramalan transaksi keuangan digital di Indonesia.

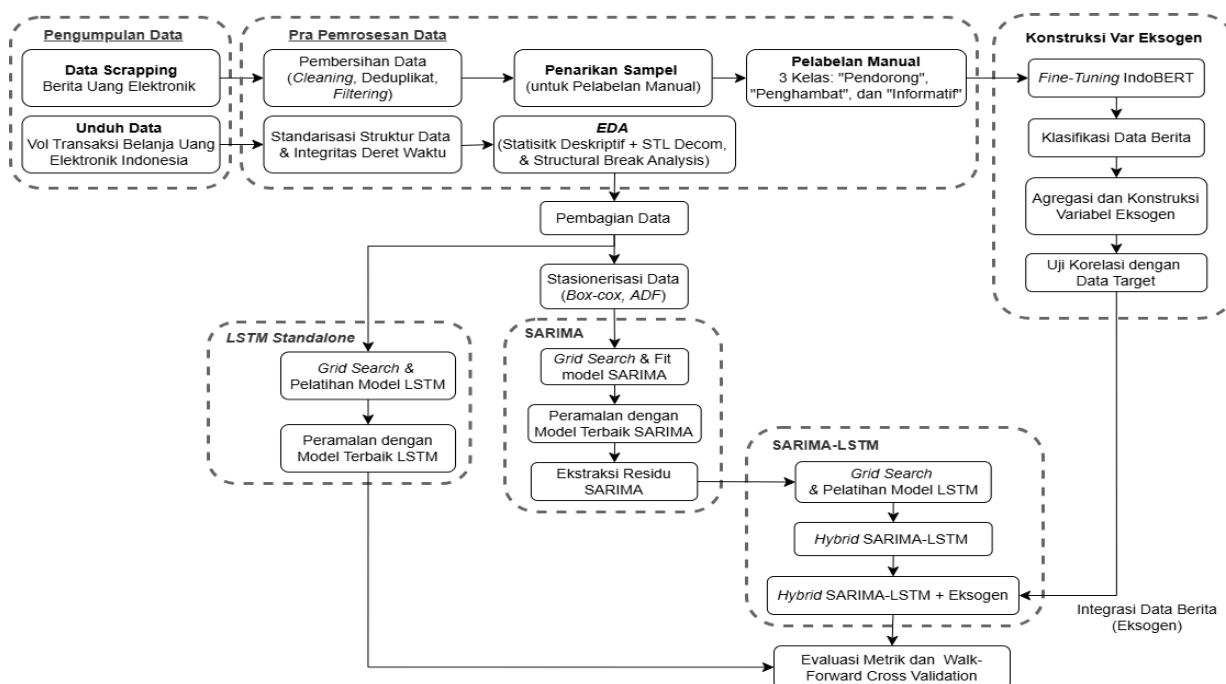
Berdasarkan celah penelitian dan potensi yang telah diuraikan, penelitian ini bertujuan untuk mengusulkan dan mengevaluasi kinerja model *hybrid* SARIMA-LSTM yang diintegrasikan dengan variabel eksogen berupa hasil klasifikasi teks berita daring berbasis IndoBERT untuk meramalkan volume transaksi uang elektronik di Indonesia. Kebaruan penelitian ini terletak pada integrasi klasifikasi teks berita berbasis model *transformer* sebagai variabel eksogen ke dalam kerangka *hybrid* SARIMA-LSTM, yang belum pernah dilakukan dalam konteks peramalan transaksi keuangan digital di Indonesia. Berbeda dengan pada studi sebelumnya yang mengandalkan pendekatan berbasis kamus sentimen (*lexicon-based*) atau variabel makroekonomi yang *lagging*, penelitian ini memanfaatkan model *transformer* yang telah di-*fine-tune* secara khusus untuk konteks bahasa Indonesia, dengan skema klasifikasi tiga kategori, yaitu “pendorong”, “penghambat”, dan “informatif” yang dirancang sesuai karakteristik dampak berita

terhadap perilaku transaksi uang elektronik. Adapun kontribusi utama penelitian ini antara lain sebagai berikut. *Pertama*, penelitian ini menguji performa arsitektur *hybrid* SARIMA-LSTM dalam meramalkan volume transaksi uang elektronik yang bersifat volatil. *Kedua*, penelitian ini mengembangkan dan mengevaluasi skema klasifikasi teks berita tiga kategori (pendorong, penghambat, dan informatif) berbasis IndoBERT sebagai variabel eksogen yang lebih *timely*. *Ketiga*, penelitian ini mengevaluasi apakah integrasi klasifikasi teks berita daring dapat berfungsi sebagai indikator pelengkap dalam peramalan, guna mengurangi ketergantungan pada data yang memiliki kendala keterlambatan publikasi. Melalui integrasi tersebut, model diharapkan mampu menghasilkan proyeksi volume transaksi uang elektronik yang lebih akurat sekaligus mendukung evaluasi kebijakan secara lebih responsif di tengah volatilitas ekonomi digital. Batasan penelitian ini mencakup penggunaan data agregat bulanan pada periode Januari 2015 hingga Maret 2026, serta ekstraksi polaritas arah dampak berita yang dibatasi pada ketiga kategori tersebut.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Metodologi penelitian ini terdiri dari lima tahapan utama, sebagaimana diilustrasikan pada Gambar 1. Tahap pertama adalah pengumpulan data, yang mencakup data volume transaksi uang elektronik sebagai variabel target serta data headline berita sebagai sumber variabel eksogen. Tahap kedua adalah prapemrosesan, yang meliputi pembersihan data, Analisis Eksplorasi Data (EDA), deteksi structural break, dan stasionerisasi data deret waktu. Tahap ketiga adalah konstruksi variabel eksogen melalui fine-tuning model IndoBERT untuk mengklasifikasikan headline berita ke dalam tiga kategori, yang kemudian diagregasi menjadi fitur bulanan. Tahap keempat adalah pemodelan komparatif yang membandingkan tiga arsitektur, yaitu LSTM standalone, SARIMA murni, dan Hybrid SARIMA-LSTM yang diintegrasikan dengan variabel eksogen. Tahap kelima adalah evaluasi dan validasi, yang dilakukan menggunakan metrik kesalahan peramalan serta walk-forward cross-validation untuk menguji ketahanan model. Setiap tahapan saling terhubung, di mana luaran dari satu tahap menjadi masukan bagi tahap berikutnya.



Gambar 1. Alur Penelitian

2.2 Pengumpulan Data

Penelitian ini menggunakan dua jenis data dengan frekuensi bulanan untuk periode Januari 2015–Maret 2026. Data target berupa volume transaksi uang elektronik diperoleh dari publikasi resmi Bank Indonesia, yaitu Statistik Sistem Keuangan Indonesia (SSKI) dan Statistik Sistem Pembayaran dan Infrastruktur Pasar Keuangan (SPIP). Sementara itu, data *headline* berita dikumpulkan melalui *web scraping* menggunakan Google News API. Ekstraksi data (waktu publikasi, URL, dan *headline*) didasarkan pada lima kata kunci: "uang elektronik", "e-money", "dompet digital", "QRIS", dan "pembayaran digital", yang kemudian dideduplikasi berdasarkan URL.

2.3 Pra-pemrosesan Data Berita

Tahap pra-pemrosesan diawali dengan eliminasi *headline* di bawah lima kata dan penyaringan berbasis kata kunci untuk memastikan relevansi topik. Selanjutnya, ditarik sampel data menggunakan *keyword-based sampling* untuk dilabeli secara manual ke dalam tiga kelas (Tabel 1).

Tabel 1. Kriteria Pelabelan dan Contoh

Kelas	Kriteria	Contoh <i>Headline</i>
Pendorong	Utilitas positif, inovasi, kemudahan, fitur baru, atau insentif finansial/promo	“Empat Bank Raksasa Gratiskan Biaya Transaksi Uang Elektronik”
Penghambat	Hambatan, risiko keamanan gangguan sistem, atau sentimen negatif.	“Gangguan Sistem Berimbas ke Top Up Uang Elektronik Bank Mandiri”
Informatif	Regulasi, panduan teknis, laporan statistik, atau informasi netral.	“10 Cara Cek Saldo e-Toll lewat HP yang Praktis dan Mudah”

Tabel 1 menyajikan kriteria pelabelan ketiga kelas beserta contoh headline representatif. Pembagian kategori didasarkan pada arah dampak pemberitaan terhadap perilaku transaksi: kelas pendorong diasosiasikan dengan faktor yang meningkatkan adopsi, kelas penghambat dengan faktor yang menurunkan kepercayaan, dan kelas informatif dengan pemberitaan netral. Kriteria anotasi diadaptasi dari studi empiris adopsi uang elektronik di Indonesia [2], [24].

2.4 Fine-Tuning IndoBERT

Model pre-trained IndoBERT (12 hidden layers, dimensi 768, 12 attention heads) memproses input teks melalui penggabungan token, segment, dan positional embeddings ke dalam lapisan transformer bidireksional [14]. Proses *fine-tuning* mengintegrasikan *classification head* di atas representasi token [CLS] dari IndoBERT. Fungsi softmax digunakan untuk menghasilkan probabilitas prediksi tiga kelas, yang diformulasikan pada persamaan (1) :

$$\hat{y} = \text{softmax}(W \cdot h_{[CLS]} + b) \quad (1)$$

Dengan $h_{[CLS]}$ merepresentasikan token [CLS], W adalah bobot *classification head*, dan b adalah bias. Fungsi loss yang digunakan adalah categorical cross-entropy, didefinisikan pada persamaan (2):

$$\mathcal{L} = -\sum_{i=1}^N \sum_{c=1}^3 y_{i,c} \log(\hat{y}_{i,c}) \quad (2)$$

Nilai $y_{i,c}$ adalah label aktual dan $\hat{y}_{i,c}$ adalah probabilitas prediksi untuk sampel i pada kelas c . Guna memperoleh performa klasifikasi yang optimal, dilakukan komparasi empat arsitektur *pre-trained* (IndoBERT varian $p1$ dan $p2$, BERT-base-Indonesian (cahya/bert-base-indonesian-522M), serta XLM-RoBERTa). Model paling optimal akan digunakan untuk mengklasifikasikan sisa headline tak berlabel.

2.5 Konstruksi Variabel Eksogen

Hasil klasifikasi IndoBERT diintegrasikan ke dalam frekuensi bulanan (t) untuk membentuk empat variabel eksogen, yang diformulasikan pada persamaan (3), (4), (5), dan (6).

$$\text{rasio_pendorong}_t = \frac{n_{\text{pendorong},t}}{n_{\text{total},t}} \quad (3)$$

$$\text{rasio_penghambat}_t = \frac{n_{\text{penghambat},t}}{n_{\text{total},t}} \quad (4)$$

$$\text{rasio_informatif}_t = \frac{n_{\text{informatif},t}}{n_{\text{total},t}} \quad (5)$$

$$\text{skor_neto}_t = \frac{n_{\text{pendorong},t} - n_{\text{penghambat},t}}{n_{\text{total},t}} \quad (6)$$

Nilai $n_{\text{pendorong},t}$, $n_{\text{penghambat},t}$, dan $n_{\text{informatif},t}$ masing-masing merepresentasikan jumlah headline yang diklasifikasikan ke dalam kelas pendorong, penghambat, dan informatif pada bulan t , sedangkan $n_{\text{total},t}$ adalah total headline pada bulan tersebut. Variabel skor neto dihitung untuk merepresentasikan dominasi arah pemberitaan terhadap volume berita secara keseluruhan. Sebelum diintegrasikan ke dalam arsitektur peramalan, data dievaluasi secara statistik menggunakan uji korelasi Spearman.

2.6 Prapemrosesan Data Variabel Target (Volume Transaksi Uang Elektronik)

Praproses data volume transaksi diawali dengan Analisis Eksplorasi Data (EDA) dilanjutkan dengan uji Bai-Perron untuk mendeteksi *structural breaks* [25]. Untuk menstabilkan varians yang meningkat seiring waktu, diterapkan transformasi Box-Cox [26]. Selanjutnya, stasioneritas data divalidasi menggunakan uji Augmented Dickey-Fuller (ADF) [27], [28], diikuti proses differencing jika data masih bersifat non-stasioner. Tahap ini diakhiri dengan analisis plot ACF dan PACF [28].

2.7 Pemodelan SARIMA

Model Seasonal Autoregressive Integrated Moving Average (SARIMA) dinotasikan sebagai SARIMA $SARIMA(p, d, q)(P, D, Q)_s$ untuk menangkap pola kemusiman deret waktu [18]. Secara formal, model ini dirumuskan pada Persamaan (7):

$$\Phi_p(B^s)\phi_p(B)(1-B)^d(1-B^s)^D y_t = \theta_Q(B^s)\theta_Q(B)\varepsilon_t \quad (7)$$

Nilai $\phi_p(B)$ dan $\Phi_p(B^s)$ adalah polinomial autoregressive (non-musiman dan musiman), $\theta_Q(B)$ dan $\theta_Q(B^s)$ adalah polinomial moving average, B adalah backshift operator, d dan D adalah orde differencing, s adalah periode musiman, dan ε_t adalah white noise [18].

Spesifikasi orde terbaik ditentukan melalui grid search berdasarkan nilai AIC/BIC terendah. Residual dari model terbaik kemudian divalidasi melalui serangkaian uji diagnostik: Ljung-Box (autokorelasi), Jarque-Bera dan Shapiro-Wilk (normalitas), ARCH-LM (heteroskedastisitas), serta Brock-Dechert-Scheinkman/BDS (non-linearitas). Residu ini kemudian diekstraksi sebagai target pelatihan LSTM pada kerangka peramalan *hybrid*.

2.8 Pemodelan LSTM dan Hybrid SARIMA-LSTM

Model *Long Short-Term Memory (LSTM)* merupakan varian *Recurrent Neural Network (RNN)* yang memetakan dependensi jangka panjang melalui *memory cell* yang dikendalikan oleh *input*, *forget*, dan *output gate* [29]. Pembaruan informasi pada *memory cell* dikontrol oleh fungsi sigmoid (σ) dan perkalian elemen ($\odot$) dengan formulasi pada persamaan (8) hingga persamaan (12):

$$i_t = \sigma(w_i[h_{t-1}, X_t] + b_i) \quad (8)$$

$$\tilde{c}_t = \tanh(w_c[h_{t-1}, X_t] + b_c) \quad (9)$$

$$f_t = \sigma(w_f[h_{t-1}, X_t] + b_f) \quad (10)$$

$$o_t = \sigma(w_o[h_{t-1}, X_t] + b_o) \quad (11)$$

$$h_t = o_t \odot \tanh(c_t) \quad (12)$$

di mana h_{t-1} adalah *hidden state* sebelumnya, x_t adalah input pada timestep t , sedangkan W dan b merepresentasikan bobot dan bias pada masing-masing gerbang [29]. Sebelum pelatihan, seluruh data input dinormalisasi menggunakan *MinMaxScaler*. Evaluasi mencakup tiga arsitektur utama: *LSTM Standalone* (dilatih pada data target), *Hybrid SARIMA-LSTM* (dilatih pada residual SARIMA), dan *Hybrid SARIMA-LSTM + Eksogen* (9 skenario kombinasi variabel, dirinci pada Tabel 9 di bagian Hasil). Konfigurasi optimal dicari melalui skema *grid search* yang mengevaluasi kombinasi panjang *window* (6, 9, 12 bulan), jumlah hidden units (16, 32, 64), dan rasio dropout (0.2 dan 0.3). Seluruh model dibangun dengan arsitektur 1 *layer LSTM* dan *dense layer* (1 unit) menggunakan optimizer Adam. Mengingat ukuran data yang terbatas serta potensi variabel eksogen berbasis berita bersifat *noisy*, beberapa langkah mitigasi *overfitting* diterapkan. Pertama, arsitektur *LSTM* yang sederhana untuk membatasi kompleksitas model. Kedua, mekanisme *early stopping (patience 10)* diterapkan untuk menghentikan pelatihan ketika performa pada data validasi tidak lagi membaik. Ketiga, kontribusi setiap variabel eksogen dievaluasi secara bertahap melalui skema *ablation* yang membandingkan sembilan skenario kombinasi variabel (dirinci pada Tabel 9). Skema ini memungkinkan identifikasi apakah sinyal eksogen memberikan perbaikan yang konsisten atau justru berperan sebagai *noise* terhadap performa peramalan. Prediksi final model *hybrid* diperoleh dengan menjumlahkan luaran (*output*) dari kedua model, yaitu hasil peramalan linier SARIMA dan estimasi residual oleh jaringan LSTM [20], [21], dinyatakan pada Persamaan (13):

$$\hat{Y}_t = \hat{Y}_{SARIMA,t} + \hat{e}_{LSTM,t} \quad (13)$$

Di mana Y_t merupakan hasil peramalan akhir model *hybrid* pada waktu t , $\hat{Y}_{SARIMA,t}$ Adalah hasil peramalan dari model SARIMA, dan $\hat{e}_{LSTM,t}$ Adalah nilai residual yang diprediksi oleh model LSTM.

2.9 Metrik Evaluasi dan Validasi Model

Evaluasi performa model peramalan dihitung berdasarkan selisih antara nilai aktual observasi (y_t) dan nilai prediksi ($\hat{y}_t$). Tiga metrik kuantitatif utama yang digunakan meliputi *Mean Absolute Error (MAE)*, *Root Mean Squared Error (RMSE)*, dan *Mean Absolute Percentage Error (MAPE)* [30], [31], yang secara matematis didefinisikan pada persamaan (14), (15), dan (16):

$$MAE = \frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t| \quad (14)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2} \quad (15)$$

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{y_t - \hat{y}_t}{y_t} \right| \times 100\% \quad (16)$$

di mana n merepresentasikan jumlah observasi pada kumpulan data uji. Model dengan perolehan nilai MAPE terkecil ditetapkan sebagai model peramalan terbaik. Lalu, sebagai tahap validasi akhir, diterapkan *Walk-Forward Validation* dengan pendekatan *Time Series Split* menggunakan 5 *fold* dan ukuran data uji 12 bulan pada setiap *fold*. Pendekatan ini menggunakan skema *expanding window*, di mana data latih diperluas secara bertahap: setiap *fold* menggunakan

seluruh data hingga titik waktu tertentu sebagai data latih, dan 12 bulan berikutnya sebagai data uji. Dengan demikian, model selalu dilatih pada data masa lalu dan diuji pada data masa depan, sehingga meniru kondisi peramalan sesungguhnya sekaligus mencegah penggunaan informasi masa depan (*look-ahead bias*) [32]. Untuk mencegah *data leakage*, pemisahan data dilakukan berbasis waktu. Parameter normalisasi (*MinMaxScaler*) dihitung hanya pada data latih di setiap *fold*, kemudian diterapkan pada data uji, sehingga tidak ada informasi dari periode uji yang memengaruhi proses pelatihan. Khusus untuk variabel eksogen berbasis berita, agregasi bulanan dihitung secara independen per periode sehingga nilai fitur pada bulan t hanya berasal dari berita yang dipublikasikan hingga bulan t . Pada peramalan *multi-step*, fitur eksogen masa depan tidak diasumsikan tersedia, melainkan menggunakan nilai periode sebelumnya (*lag-1*), konsisten dengan kondisi penerapan riil di mana berita masa depan belum dapat diobservasi.

3. HASIL DAN PEMBAHASAN

3.1 Fine-Tuning Model IndoBERT

Proses pengumpulan data menghasilkan 13.758 *headline* relevan terkait ekosistem uang elektronik. Dari populasi tersebut, 2.705 sampel diekstraksi melalui proses filter kata kunci tertentu untuk pelabelan manual. Lalu, distribusi data diseimbangkan melalui teknik *undersampling*, dengan komposisi setara 814 sampel per kelas. Tahapan *fine-tuning* model IndoBERT, menggunakan proporsi pembagian data 80:10:10 untuk data berita yang akan diklasifikasikan, menghasilkan 1.953 data latih, 244 data validasi, dan 245 data uji. Tahap pemodelan diawali dengan membandingkan kinerja beberapa arsitektur model *pre-trained* dengan konfigurasi *hyperparameter baseline* (5 *epoch*, *batch size* 16, *learning rate* 2×10^{-5} , dan *weight decay* 0.01) untuk menentukan model terbaik. Perbandingan performa dari ketiga model tersebut dirangkum pada Tabel 2.

Tabel 2. Tabel Pelatihan Model

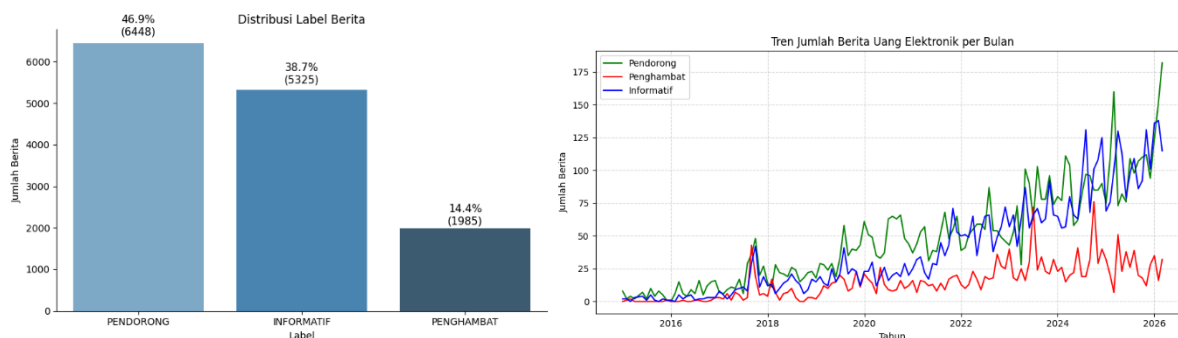
Model	Epoch Optimal	Akurasi Validasi	F1-Macro Validasi
indobenchmark/indobert-base-p1	5	91.8%	91.8%
indobenchmark/indobert-base-p2	2	90.2%	90.2%
xlm-roberta-base	3	88.5%	88.4%
cahya/bert-base-indonesian-522M	5	88.5%	88.4%

Berdasarkan Tabel 2, arsitektur IndoBERT-base-p1 menunjukkan akurasi dan F1-macro validasi tertinggi sebesar 91,8% pada epoch optimal ke-5, sehingga arsitektur tersebut dipilih sebagai model klasifikasi final. Model final selanjutnya dievaluasi pada data uji, dengan hasil performa per kategori dirangkum pada Tabel 3.

Tabel 3. Performa Model Final pada Data Uji

Kategori	Precision (%)	Recall (%)	F1-Score (%)	Akurasi (%)
Pendorong	93%	88%	90%	93%
Penghambat	95%	99%	97%	
Informatif	89%	91%	90%	

Tabel 3 menyajikan performa model final IndoBERT-base-p1 pada data uji. Model mencapai akurasi keseluruhan sebesar 93% dengan performa yang konsisten di ketiga kategori. Kelas Penghambat memperoleh F1-score tertinggi (97%), diikuti Pendorong dan Informatif yang masing-masing sebesar 90%. Hasil ini menunjukkan model mampu mengklasifikasikan ketiga kategori dengan baik pada data yang belum pernah dilihat sebelumnya. Model ini selanjutnya digunakan untuk mengklasifikasikan seluruh 13.758 *headline* berita guna menghasilkan variabel eksogen.



Gambar 2. Distribusi Headline Berita Hasil Klasifikasi

Dari hasil klasifikasi pada Gambar 2, pemberitaan ekosistem uang elektronik di Indonesia didominasi kategori "Pendorong" sebanyak 6.448 berita (46,9%), diikuti "Informatif" 5.325 berita (38,7%), dan "Penghambat" 1.985 berita (14,4%). Setelah diagregasi per bulan, tren frekuensi menunjukkan peningkatan signifikan sejak 2020 dengan kategori

pendorong konsisten memimpin. Proporsi dan fluktuasi ketiga kelas inilah yang merepresentasikan dinamika pemberitaan secara kuantitatif, sekaligus menjadi bahan baku utama dalam konstruksi variabel eksogen selanjutnya.

3.2 Konstruksi Variabel Eksogen (*Headline Berita*)

Hasil agregasi bulanan data berita berlabel kemudian dikonstruksi menjadi empat variabel eksogen: *rasio_pendorong*, *rasio_penghambat*, *rasio_informatif*, dan *skor_net* sebagaimana didefinisikan pada Persamaan 3 hingga 6. Langkah ini dilakukan untuk menormalisasi data dan menghilangkan bias volume pemberitaan yang terus meningkat dari tahun ke tahun. Untuk mengevaluasi hubungan awal antara variabel eksogen berita dengan deret waktu volume transaksi uang elektronik, dilakukan uji korelasi Spearman pada waktu bersamaan (*Nilai r*) dan juga memperhitungkan efek tunda hingga tiga periode sebelumnya (*Lag 1* hingga *Lag 3*). Rangkuman hasil uji korelasi disajikan pada Tabel 4.

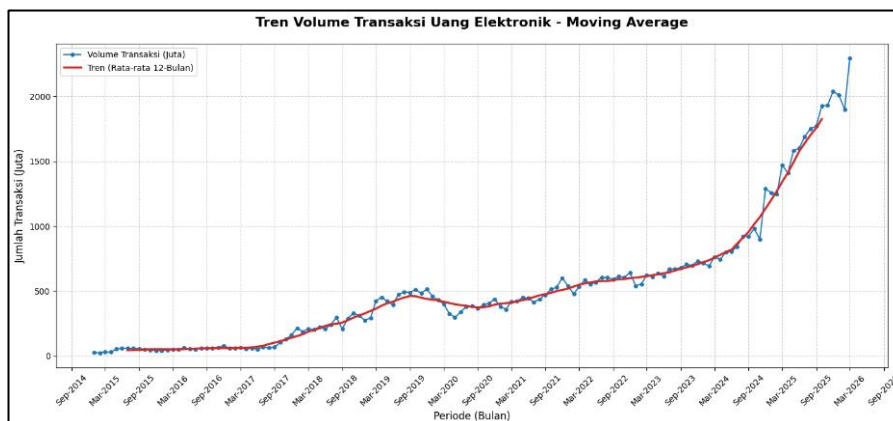
Tabel 4. Uji Korelasi Spearman terhadap data Target (Vol Transaksi Uang Elektronik)

Variabel	Nilai r	Nilai r (Lag 1)	Nilai r (Lag 2)	Nilai r (Lag 3)
<i>rasio_pendorong</i>	0.0308	-0.1370	-0.0966	-0.0926
<i>rasio_penghambat</i>	-0.0489	-0.0023	0.1035	0.0180
<i>rasio_informatif</i>	-0.1493	0.0643	-0.0338	0.0141
<i>skor_net</i>	0.0253	-0.0995	-0.1155	-0.0756

Berdasarkan Tabel 4, seluruh variabel eksogen menunjukkan koefisien korelasi *Spearman* yang tergolong sangat lemah terhadap fluktuasi volume transaksi uang elektronik. Nilai korelasi absolut tertinggi ditunjukkan oleh *rasio_informatif* pada waktu yang bersamaan sebesar -0,1493, disusul oleh *rasio_pendorong* pada *Lag 1* sebesar -0,1370, dan *skor_net* pada *Lag 2* sebesar -0,1155. Meskipun uji korelasi menunjukkan hubungan yang lemah, seluruh variabel ini tetap dipertahankan untuk tahapan pemodelan.

3.3 Analisis Eksplorasi Data Variabel Target (Volume Transaksi Uang Elektronik)

Selama periode Januari 2015–Maret 2026, volume transaksi uang elektronik di Indonesia tumbuh eksponensial hingga lebih dari 94 kali lipat. Transaksi rata-rata mencapai 528,3 juta per bulan, melonjak dari titik terendahnya 24,1 juta (Januari 2015) menjadi 2,29 miliar (Maret 2026). Dinamika tren dan *Moving Average* dari data ini divisualisasikan pada Gambar 3.



Gambar 3. Plot Data Volume Transaksi Uang Elektronik

Gambar 3 menunjukkan tren non-linear dengan percepatan adopsi pada 2017–2018 dan akselerasi tajam menjelang 2024. Pergeseran pola ini mengindikasikan adanya perubahan struktural akibat intervensi eksternal. Melalui uji Bai-Perron, teridentifikasi 4 titik break optimal berdasarkan nilai BIC terkecil (35,3266), yang membagi deret waktu menjadi lima segmen rezim. Rincian pemisahan ini disajikan pada Tabel 5.

Tabel 5. Ringkasan Perubahan Rezim Struktural (Uji Bai-Perron)

Segmen	Periode Waktu	Rata-rata Transaksi (Juta)	Indikasi Eksternal
1	Jan 2015 – Sep 2017	53.5	Fase Awal Pengenalan Ekosistem
2	Okt 2017 – Jan 2020	321.5	Akselerasi Uang Elektronik Berbasis Server
3	Feb 2020 – Sep 2021	398.5	Dampak dan Adaptasi Pandemi COVID-19
4	Okt 2021 – Mei 2024	625.8	Pemulihan dan Digitalisasi Pascapandemi
5	Jun 2024 – Mar 2026	1479.9	Lonjakan Masif Adopsi QRIS

Tabel 5 menunjukkan bahwa deret waktu terbagi menjadi lima segmen dengan karakteristik rata-rata pertumbuhan yang melompat drastis pada setiap pergeseran rezim. Identifikasi pergeseran struktural ini menjadi justifikasi bahwa data tidak stasioner secara rata-rata maupun varians. Oleh karena itu, diperlukan tahapan

prapemrosesan lanjutan, seperti transformasi atau *differencing*, sebelum data siap dimodelkan menggunakan arsitektur peramalan.

3.4 Pra-Pemrosesan Variabel Target (Volume Transaksi Uang Elektronik)

Data dibagi dengan metode *time series split* menjadi 90% data latih (Jan 2015–Jan 2025) dan 10% data uji (Feb 2025–Mar 2026). Pemilihan data uji ini difokuskan untuk mengevaluasi kinerja model pada rezim pertumbuhan terbaru pasca-*structural break* terakhir (Juni 2024). Kemudian, untuk menstabilkan varians akibat tren eksponensial, diterapkan transformasi Box-Cox dengan parameter lambda yang diestimasi dari data latih melalui maximum likelihood (interval kepercayaan 95% pada [0.25, 0.63]). Setelah varians stabil, evaluasi menggunakan Uji *Augmented Dickey-Fuller* (ADF) menunjukkan bahwa data belum stasioner ($p = 0.4783$). Oleh karena itu, dilakukan *first differencing* ($d=1$) yang berhasil mencapai stasioneritas dengan signifikansi 1% (ADF = -3.5819, $p = 0.0061$). Selanjutnya, komponen *Autoregressive* (AR) dan *Moving Average* (MA) diidentifikasi melalui plot ACF dan PACF pada deret waktu yang telah stasioner.

3.5 Hasil Grid Search dan Pemilihan Model Dasar

Pemilihan model dilakukan melalui *grid search* berdasarkan *Akaike Information Criterion* (AIC), yang mengevaluasi parameter non-musiman (ARIMA) dan musiman (SARIMA). Rangkuman kandidat disajikan pada Tabel 6.

Tabel 6. Hasil Grid Search 5 Terbaik

Model	AIC	BIC	Ljung-Box p
SARIMA(0,1,0)(0,1,1) ₁₂	1754.05	1759.41	0.7806
SARIMA(2,1,0)(0,1,1) ₁₂	1755.43	1766.16	0.7307
SARIMA(0,1,1)(0,1,1) ₁₂	1755.94	1763.98	0.7857
ARIMA(0,1,0)	1947.82	1950.61	0.2047
ARIMA(2,1,2)	1948.09	1962.02	0.6304
ARIMA(0,1,1)	1948.38	1953.96	0.4391

Berdasarkan Tabel 6, integrasi komponen musiman terbukti meningkatkan performa. Model SARIMA(0,1,0)(0,1,1,12) mencatatkan AIC terendah (1754,05), turun signifikan dibandingkan model ARIMA terbaik. Uji Ljung-Box ($p > 0.05$) juga memastikan residual model terdistribusi acak tanpa autokorelasi. Selanjutnya, model dievaluasi menggunakan data uji untuk mengukur kemampuan generalisasinya

Tabel 7. Evaluasi Metrik Peramalan Kandidat Model pada Data Uji

Model	MAE	RMSE	MAPE	Akurasi
SARIMA(0,1,0)(0,1,1) ₁₂	330,792,231	368,801,026	17.77%	82.23%
SARIMA(0,1,1)(0,1,1) ₁₂	333,487,643	371,583,106	17.91%	82.09%
SARIMA(2,1,0)(0,1,1) ₁₂	397,834,798	432,514,467	21.61%	78.39%
ARIMA(2,1,2)	441,036,336	514,382,298	23.27%	76.73%
ARIMA(0,1,0)	503,501,293	571,015,723	26.82%	73.18%
ARIMA(0,1,1)	504,685,917	572,231,121	26.89%	73.11%

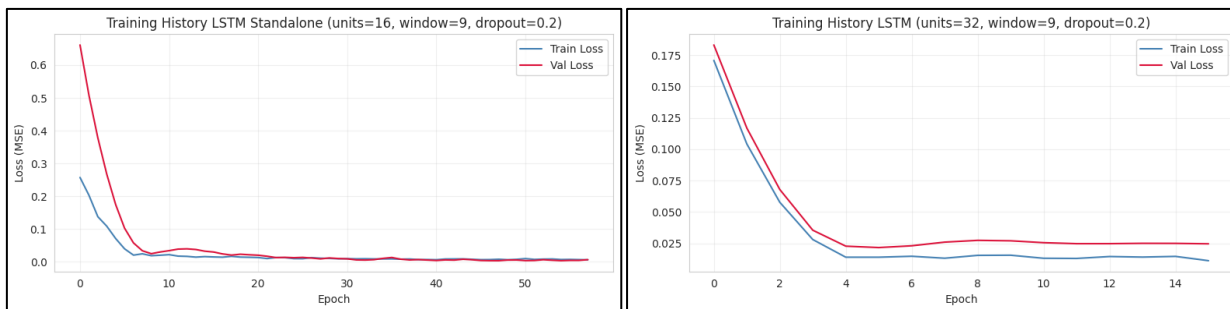
Tabel 7 menyajikan evaluasi keenam kandidat model pada data uji. Model SARIMA(0,1,0)(0,1,1,12) mengonfirmasi posisinya sebagai baseline terbaik dengan MAPE terendah (17,77%), unggul atas seluruh varian ARIMA non-musiman yang mencatatkan MAPE di atas 23%. Konsistensi peringkat antara kriteria AIC (Tabel 6) dan performa pada data uji (Tabel 7) menunjukkan pemilihan model tidak mengalami overfitting terhadap data latih.

3.6 Diagnostik Residual Model SARIMA

Diagnostik ini bertujuan memverifikasi ekstraksi pola linear SARIMA dan mendeteksi sisa komponen non-linear. Hasil diagnosa menunjukkan mayoritas data berada dalam batas $\pm 2\sigma$ dengan rata-rata 54,22, meskipun terdapat beberapa *outliers*. Uji statistik mengonfirmasi bahwa residual bebas dari autokorelasi (Uji Ljung-Box: $p = 0.7806$) dan memiliki varians yang stabil (Uji ARCH LM: $p = 0.9993$). Meskipun asumsi linear terpenuhi, distribusi data tidak normal (Uji Jarque-Bera & Shapiro-Wilk: $p < 0.05$). Uji *Brock-Dechert-Scheinkman* (BDS) pada dimensi $m=5$ ($p = 0.0390$) secara formal membuktikan bahwa residual masih mengandung komponen non-linear yang gagal ditangkap oleh SARIMA. Temuan ini menjadi justifikasi diterapkannya arsitektur LSTM pada tahap selanjutnya.

3.7 Pelatihan Model LSTM dan Hybrid SARIMA-LSTM

Sebagai pembanding, *LSTM Standalone* dilatih menggunakan data historis Volume Transaksi. Berdasarkan optimasi *grid search*, konfigurasi optimal diperoleh pada *window size* 9, 16 *units*, dan *dropout* 0.2. Sebagaimana ditunjukkan pada Gambar 4, proses pelatihan dengan *optimizer* Adam dan *Mean Squared Error* (MSE) berhenti otomatis (*early stopping*) pada *epoch* 58. Konvergensi tercapai stabil tanpa *overfitting* yang signifikan (*training loss* 0.0060; *validation loss* 0.0065).



Gambar 4. Kurva Pelatihan LSTM Standalone (Target) dan LSTM untuk *Hybrid* (Residu)

Pada skema *hybrid*, LSTM difungsikan sebagai pengoreksi *error* untuk mempelajari sisa komponen non-linear dari residual SARIMA(0,1,0)(0,1,1,12). Dengan ruang pencarian parameter yang sama, konfigurasi optimal *hybrid* membutuhkan *window size* 9, 32 *units*, dan *dropout* 0.2. Kebutuhan *units* yang lebih besar ini mengindikasikan bahwa deret residual memiliki pola volatilitas non-linear yang lebih kompleks dibandingkan deret aslinya. Berdasarkan Gambar 4, model *hybrid* mencapai titik konvergensi optimal lebih cepat pada *epoch* 16 (*training loss* 0.0112; *validation loss* 0.0246). Meskipun durasi pelatihannya lebih singkat, penurunan fungsi *loss* tetap stabil, membuktikan bahwa model telah mengekstraksi bobot sisa informasi non-linear SARIMA dengan baik

3.8 Evaluasi Metrik Model LSTM dan Hybrid SARIMA-LSTM

Evaluasi generalisasi model pada data uji dengan skala asli hasil *inverse transform* dirangkum pada Tabel 8.

Tabel 8. Evaluasi Metrik Model LSTM dan Hybrid (Data Uji)

Model	MAE	RMSE	MAPE	Akurasi
LSTM <i>Standalone</i>	421.826.222	443.622.769	23.37%	76.63%
SARIMA (<i>Baseline</i>)	330.792.230	368.801.025	17.77%	82.23%
<i>Hybrid</i> SARIMA-LSTM	286.104.513	325.958.667	15.24%	84.76%

Berdasarkan Tabel 8, model *Hybrid* SARIMA-LSTM terbukti paling unggul dengan akurasi 84.76% (MAPE 15.24%), berhasil memperbaiki tingkat kesalahan *baseline* SARIMA sebesar 2.53 poin persentase. Peningkatan ini membuktikan efektivitas pembagian tugas: SARIMA menangkap tren linear, sementara LSTM mengoreksi sisa fluktuasi non-linear. Di sisi lain, ketiadaan komponen linear eksplisit membuat *LSTM Standalone* berkinerja paling buruk (MAPE 23.37%) karena kesulitan mengekstrapolasi tren pertumbuhan jangka panjang secara mandiri.

3.9 Integrasi Variabel Eksogen pada Model Hybrid

Untuk menguji apakah berita dapat mengoptimalkan akurasi, variabel eksogen diintegrasikan ke dalam arsitektur *baseline* Hybrid SARIMA-LSTM. Integrasi ini dilakukan melalui pendekatan *feature stacking*, yakni menggabungkan deret residual SARIMA dan variabel rasio berita yang telah dinormalisasi. Dengan demikian, setiap *sequence input* memuat fluktuasi residual masa lalu sekaligus vektor variabel eksogen. Guna mengevaluasi kontribusinya, dikonstruksikan sembilan skenario eksperimen: pengujian individu (Skenario 1–4), kombinasi berpasangan (Skenario 5a–5d), dan penggabungan seluruh variabel (Skenario 6). Agar komparasi arsitektural seimbang, seluruh skenario menggunakan *hyperparameter baseline* yang identik (*window* = 9, *units* = 32, *dropout* = 0.2). Hasil proyeksi model pada data uji (Februari 2025–Maret 2026) disajikan pada Tabel 9.

Tabel 9. Evaluasi Metrik Model Hybrid SARIMA-LSTM + Eksogen

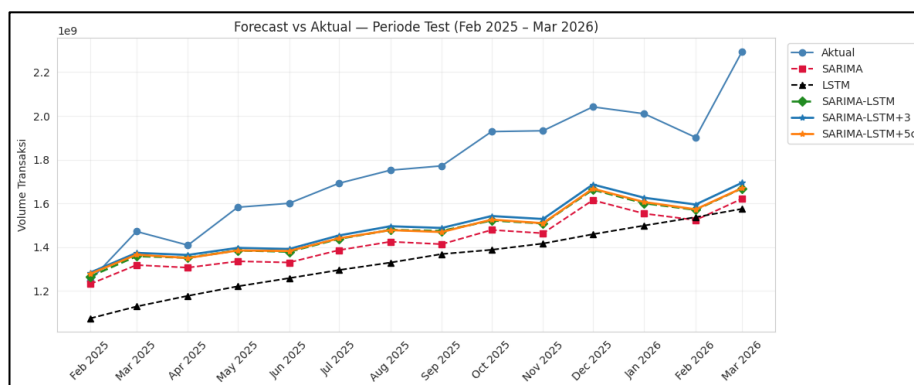
Skenario	Keterangan Fitur Eksogen	MAPE	Akurasi
3	Rasio Informatif (Tunggal)	14.42%	85.58%
5c	Rasio Penghambat + Rasio Informatif	15.22%	84.78%
	SARIMA-LSTM (<i>Baseline</i>)	15.24%	84.76%
5b	Rasio Pendorong + Rasio Informatif	15.41%	84.59%
6	Seluruh Eksogen (Gabungan)	15.91%	84.09%
2	Rasio Penghambat (Tunggal)	16.08%	83.92%
5d	Rasio Informatif + Skor Neto	16.98%	83.02%
4	Skor Neto (Tunggal)	17.11%	82.89%
5a	Rasio Pendorong + Rasio Penghambat	17.14%	82.86%
1	Rasio Pendorong (Tunggal)	17.27%	82.73%

Berdasarkan Tabel 9, dari kesembilan konfigurasi, hanya Skenario 3 dan 5c yang berhasil melampaui performa *baseline* (MAPE 15.24%). Skenario 3 yang mengikutsertakan "Rasio Informatif" secara tunggal tampil sebagai model terbaik dengan tingkat kesalahan terkecil (MAPE 14.42%) dan akurasi tertinggi (85.58%). Sebaliknya, seluruh skenario yang melibatkan sentimen positif "Rasio Pendorong" atau "Skor Neto" secara konsisten menghasilkan

akurasi di bawah *baseline*. Hal ini mengindikasikan bahwa *headline* berita opini positif tidak serta-merta mendorong volume transaksi, melainkan justru memberikan beban *noise* matematis pada algoritma LSTM.

3.10 Visualisasi Hasil Peramalan Pada Data Uji

Visualisasi komparasi peramalan pada data uji (Gambar 5) menunjukkan bahwa seluruh model mampu menangkap arah tren aktual, meskipun persisten mengalami *underestimation* yang melebar menjelang akhir periode (Maret 2026). *LSTM Standalone* memperlihatkan penyimpangan lintasan terjauh. Sebaliknya, model SARIMA *baseline*, Hybrid SARIMA-LSTM, dan Hybrid + Eksogen Skenario 3 tampil lebih presisi. Sesuai dengan metrik evaluasinya (MAPE 14.42%), lintasan Skenario 3 terbukti paling mendekati data riil.



Gambar 5. Visualisasi Peramalan pada *Test Set* (Semua Model)

Secara visual pada Gambar 5, lintasan model *hybrid* hanya tampak sebagai pergeseran kurva SARIMA ke atas secara tipis. Karakteristik ini mengonfirmasi mekanisme kerja arsitektur *hybrid*: SARIMA tetap mengendalikan pembentukan pola dasar (tren linear dan musiman), sementara LSTM murni bertugas memprediksi sisa residu sebagai faktor koreksi akhir, tanpa merusak bentuk lekukan tren utamanya.

3.11 Walk-Forward Validation

Untuk menguji stabilitas model di luar skenario utama, diterapkan *walk-forward validation* sebanyak 5 *folds* (12 bulan per *fold*). Hasil evaluasi lintas periode ini disajikan pada Tabel 10.

Tabel 10. Perbandingan Kinerja Walk-Forward Validation (MAPE %)

Fold	Periode Test	SARIMA	LSTM	SARIMA-LSTM	SARIMA-LSTM+Eksogen
1	Apr 2021 – Mar 2022	6.72%	20.91%	6.39%	6.62%
2	Apr 2022 – Mar 2023	6.69%	21.46%	6.82%	6.60%
3	Apr 2023 – Mar 2024	1.94%	25.08%	2.33%	2.36%
4	Apr 2024 – Mar 2025	15.18%	16.71%	15.46%	14.69%
5	Apr 2025 – Mar 2026	11.09%	16.39%	7.90%	7.54%
Mean		8.32%	20.11%	7.78%	7.56%
Std		4.49%	3.24%	4.28%	3.99%

Berdasarkan Tabel 10, pada periode pertumbuhan stabil sebelum *structural break* (Fold 1–3), komponen linear SARIMA murni sudah memadai, sehingga koreksi dari LSTM berdampak minor. Namun, ketika terjadi *break* pada pertengahan 2024 (Fold 4), tingkat kesalahan seluruh model melonjak, di mana SARIMA-LSTM+Eksogen mampu bertahan dengan MAPE terendah (14.69%). Keunggulan *hybrid* ini makin signifikan pada masa pasca-*break* (Fold 5), berhasil menekan MAPE menjadi 7.54% dibanding SARIMA murni (11.09%). Secara keseluruhan, model SARIMA-LSTM+Eksogen mencatatkan rata-rata kesalahan terendah (MAPE 7.56%) dengan standar deviasi terkecil (3.99%). Angka ini membuktikan bahwa pendekatan *hybrid* tidak hanya memberikan akurasi rata-rata yang lebih baik, tetapi juga lebih konsisten dan tangguh menghadapi perubahan rezim waktu dibandingkan model tunggal.

3.12 Pembahasan

Secara empiris, arsitektur Hybrid SARIMA-LSTM + Eksogen (Skenario 3) terbukti sebagai model paling efektif dengan MAPE terendah (14,42%). Capaian ini mengonfirmasi landasan teoretis bahwa integrasi pemodelan statistik untuk komponen tren linier dan *deep learning* untuk dinamika residual non-linier memberikan hasil yang lebih komprehensif dibandingkan model tunggal [20], [21]. Peningkatan akurasi dari penambahan variabel eksogen tergolong marginal (0,82 poin persentase dari baseline hybrid tanpa eksogen). Hal ini diduga karena model LSTM telah menyerap sebagian besar pola volatilitas dari lag residual historis, sehingga ruang perbaikan dari sinyal eksternal menjadi terbatas. Meski demikian, variabel "Rasio Informatif" tetap memberikan nilai tambah prediktif.

Salah satu temuan yang didapat penelitian ini adalah variabel dengan korelasi Spearman sangat lemah (Tabel 4) justru memberi kontribusi terbaik pada pemodelan (Skenario 3). Hal ini karena Spearman hanya mengukur

hubungan searah antara dua variabel, sehingga tidak menangkap pola non-linier atau yang bergantung konteks waktu yang justru mampu diolah LSTM melalui sekuens sembilan bulan bersama residual SARIMA. Terpilihnya kategori 'Informatif' diduga karena sifatnya yang lebih faktual dan kurang bermuatan sentimen dibanding berita beropini, sehingga sinyalnya lebih stabil. Kontribusi ini bersifat marginal dan kondisional (lebih menonjol pada periode gejolak, Tabel 10), sehingga tidak diklaim sebagai hubungan kausal. Pemanfaatan klasifikasi teks berbasis IndoBERT [14] menjadi kunci dalam mengekstraksi sinyal informasi ini dari headline berita.

Jika dibandingkan dengan penelitian terdahulu yang mencatatkan MAPE lebih rendah, yakni berkisar 3,57% hingga 11,33% [6], [8], performa absolut model pada penelitian ini tampak lebih tinggi. Namun, perbedaan tersebut sebagian besar dipengaruhi oleh rentang data observasi (2015–2026) yang memuat empat titik *structural break* ekstrem, dengan periode data uji berada tepat pada fase pasca-lonjakan transaksi tahun 2024. Ketika diuji pada periode yang lebih stabil melalui *walk-forward validation* (Fold 1–3), model ini mencatatkan MAPE di bawah 7%, sebanding dengan performa studi terdahulu. Secara metodologis, penelitian ini memberikan kontribusi kebaruan melalui integrasi deteksi *structural break* Bai-Perron, klasifikasi teks berbasis transformer IndoBERT, dan arsitektur hybrid SARIMA-LSTM, yang secara kolektif mengatasi keterbatasan studi terdahulu terkait ketergantungan pada variabel makroekonomi yang bersifat *lagging* [10].

4. KESIMPULAN

Penelitian ini menyimpulkan bahwa integrasi pendekatan statistik linier dan *deep learning* melalui model *Hybrid SARIMA-LSTM* yang diperkaya variabel eksogen berbasis berita mampu meningkatkan akurasi peramalan volume transaksi uang elektronik di Indonesia. Model terbaik, yaitu *Hybrid SARIMA-LSTM* dengan variabel "Rasio Informatif" (Skenario 3), mencapai MAPE 14,42% dan mengungguli SARIMA murni maupun LSTM standalone, dengan proses klasifikasi berita IndoBERT yang mencapai akurasi 93%. Berdasarkan kontribusi penelitian yang telah ditetapkan, hasil ini menegaskan bahwa kerangka *hybrid SARIMA-LSTM* yang mengintegrasikan klasifikasi teks berita berbasis *transformer* sebagai variabel eksogen berhasil dikembangkan dan menunjukkan performa pemanfaatan sinyal informasi yang lebih *timely*, berbeda dengan penelitian terdahulu yang mengandalkan variabel makroekonomi *lagging* atau data univariat dalam konteks peramalan transaksi keuangan digital di Indonesia [6], [8]. Selain kontribusi metodologis tersebut, penelitian ini juga mengungkap temuan bahwa berita bertipe informatif (edukasi dan regulasi) memberikan sinyal prediktif yang lebih relevan dibandingkan berita beropini bersentimen positif, meskipun kontribusinya bersifat marginal dan lebih menonjol pada periode gejolak. Sebagai catatan, penelitian ini dibatasi oleh frekuensi agregasi bulanan yang membatasi jumlah sampel pelatihan. Penelitian mendatang dapat mempertimbangkan data berfrekuensi lebih tinggi untuk memperbanyak observasi, dengan catatan bahwa data teks harian menghadirkan tantangan *noise*, kualitas data, dan beban komputasi yang lebih besar sehingga memerlukan strategi penyaringan yang lebih ketat. Selain itu, pengembangan klasifikasi berbasis *Aspect-Based Sentiment Analysis (ABSA)* maupun eksplorasi arsitektur berbasis *Attention* dapat menjadi arah lanjutan untuk meningkatkan granularitas sinyal dan adaptasi terhadap *structural break*.

REFERENCES

- [1] M. Barroso and J. Laborda, "Digital transformation and the emergence of the Fintech sector: Systematic literature review," *Digital Business*, vol. 2, no. 2, Jan. 2022, doi: 10.1016/j.digbus.2022.100028.
- [2] W. Widayat, I. Masudin, and N. R. Satiti, "E-Money payment: Customers' adopting factors and the implication for open innovation," *Journal of Open Innovation: Technology, Market, and Complexity*, vol. 6, no. 3, Sep. 2020, doi: 10.3390/JOITMC6030057.
- [3] Bank Indonesia, *Blueprint Sistem Pembayaran Indonesia 2030: Mengakselerasi Ekonomi Digital Nasional untuk Generasi Mendatang*. Jakarta: Bank Indonesia, 2024. [Online]. Available: <https://www.bi.go.id/id/publikasi/kajian/Documents/Blueprint-Sistem-Pembayaran-Indonesia-2030.pdf>
- [4] P. Nailul, H. • Dyah, A. • Rani, S. E. Bhima, and Y. Adhinegara, "Outlook Ekonomi Digital 2025," 2024. [Online]. Available: www.celios.co.id
- [5] A. Yuli Wijaya, I. Mukhlis, and L. Seprillina, "Analisis pengaruh E-money, volume transaksi elektronik dan suku bunga terhadap jumlah uang beredar di Indonesia pada masa sebelum dan sesudah pandemi COVID-19," *Jurnal Ekonomi, Bisnis dan Pendidikan*, vol. 1, no. 2, pp. 135–145, Feb. 2021, doi: 10.17977/um066v1i22021p135-145.
- [6] I. Nengah Dharma Pradnyandita and A. A. Rohmawati, "Electronic Money Transactions Forecasting with Support Vector Regression (SVR) and Vector Autoregressive Moving Average (VARMA)," *Intl. Journal on ICT*, vol. 8, no. 1, pp. 69–85, 2022, doi: 10.21108/ijoint.v8i1.632.
- [7] U. Mahmudah, S. R. Anggita, and N. A. Suryaningtyas, "Forecasting electronic money trends in Indonesia using neural network models: A comparative analysis," *Majalah Ilmiah Matematika dan Statistika (MIMS)*, vol. 25, no. 1, pp. 1–12, Mar. 2025, doi: 10.19184/mims.v25i1.53688.
- [8] C. Andreas, S. Sediono, E. Ana, S. Suliyanto, and M. F. F. Mardianto, "Penerapan Model ARIMAX-GARCH dalam Pemodelan dan Peramalan Volume Transaksi Uang Elektronik di Indonesia," *MUST: Journal of Mathematics Education, Science and Technology*, vol. 6, no. 2, p. 241, Dec. 2021, doi: 10.30651/must.v6i2.11214.

- [9] C. Fakhri, F. Purwoko, P. N. Farida, S. N. Ramadhani, and V. M. Erson, "Modeling Indonesia's inflation based on e-money transaction volume with local polynomial estimator approach," *Indonesian Journal of Statistics and Its Applications*, vol. 1, no. 5, Feb. 2022, doi: 10.55324/iss.v1i5.120.
- [10] J. T. E. Chapman and A. Desai, "Macroeconomic Predictions Using Payments Data and Machine Learning," *Forecasting*, vol. 5, no. 4, pp. 652–683, Dec. 2023, doi: 10.3390/forecast5040036.
- [11] M. D. Rizqi, U. Mahmudah, and R. D. Aulia, "Prediksi Tren Penggunaan Electronic Money (E-Money) di Indonesia menggunakan pendekatan Extreme Learning Machines," *Emerging Statistics and Data Science Journal*, vol. 3, no. 3, pp. 640–649, Oct. 2025, doi: 10.20885/esds.vol3.iss.3.art17.
- [12] E. Chong, C. C. Ho, Z. F. Ong, and H. H. Ong, "Using news sentiment for economic forecasting: a Malaysian case study," in *Machine learning in central banking*, vol. 57, Bank for International Settlements, 2022. Tersedia: https://www.bis.org/ifc/publ/ifcb57_17.pdf
- [13] J. J. C. Ying, C. C. Liu, V. S. Tseng, W. Zhang, and J. Zhang, "Forecasting leading economic indicators in the US from financial news using multi-task learning," *Soft comput.*, vol. 29, no. 7, pp. 3641–3657, Apr. 2025, doi: 10.1007/s00500-025-10629-5.
- [14] B. Wilie *et al.*, "IndoNLU: Benchmark and Resources for Evaluating Indonesian Natural Language Understanding," in *Proc. 1st Conf. Asia-Pacific Chapter Assoc. Comput. Linguistics and 10th Int. Joint Conf. Natural Language Processing*, Suzhou, China, Dec. 2020, pp. 843–857, doi: 10.18653/v1/2020.aacl-main.85.
- [15] B. Juarto and Yulianto, "Indonesian News Classification Using IndoBERT," *Int. J. Intell. Syst. Appl. Eng.*, vol. 11, no. 2, pp. 454–460, Feb. 2023. Tersedia: <https://ijisae.org/index.php/IJISAE/article/view/2654>
- [16] D. S. Katemba, F. Bisilisin, Heni, and Sumarlin, "Klasifikasi Judul Berita Online Menggunakan BERT dan IndoBERT," *TelKa*, vol. 16, no. 1, pp. 80–87, 2026, doi: 10.36342/3sqzca39.
- [17] W. A. Pratiwi, I. M. Sumertajaya, and K. A. Notodiputro, "Stacking Ensemble RNN-LSTM Models for Forecasting the IDR/USD Exchange Rate with Nonlinear Volatility," *Jurnal Teknik Informatika (Jutif)*, vol. 6, no. 4, pp. 2331–2347, Aug. 2025, doi: 10.52436/1.jutif.2025.6.4.5057.
- [18] M. Rizal and S. U. Zuliana, "Forecasting using SARIMA and Bayesian structural time series method for range seasonal time," in *Proc. Int. Conf. Data Science and Official Statistics*, vol. 2023, no. 1, pp. 382–391, 2023, doi: 10.34123/icdsos.v2023i1.402.
- [19] M. Baek and Y. Seo, "Hybrid forecasting of university electricity demand using time series and deep learning," *Energy Build.*, vol. 347, Nov. 2025, doi: 10.1016/j.enbuild.2025.116400.
- [20] X. Du *et al.*, "Application and comparison of ARIMA, LSTM, and ARIMA-LSTM models for predicting foodborne diseases in Liaoning Province," *Front. Big Data*, vol. 8, 2025, doi: 10.3389/fdata.2025.1666962.
- [21] A. Aiman Dzulfikar, "Hybrid Sarima-LSTM Model for Predicting Erosion In Butterfly Valve," *Jurnal Ilmu dan Inovasi Fisika*, vol. 9, no. 1, pp. 75–88, Feb. 2025, doi: 10.24198/jiif.v9i1.61007.
- [22] L. Barbaglia, S. Consoli, and S. Manzan, "Forecasting with Economic News," *Journal of Business and Economic Statistics*, vol. 41, no. 3, pp. 708–719, 2023, doi: 10.1080/07350015.2022.2060988.
- [23] N. J. Salsabiila and W. Anggraeni, "CNN+BiGRU-Attention Classification and TiDE-PSO Forecasting Approach for Social Media-based Predictive Analysis of Dengue," *International Journal of Intelligent Engineering and Systems*, vol. 18, no. 4, pp. 716–735, 2025, doi: 10.22266/ijies.2025.0531.47.
- [24] H. Aldiah and S. V. V. Lena, "Analisis Faktor-Faktor Yang Mempengaruhi Minat Penggunaan Uang Elektronik Pada Generasi Z," *Jurnal Minfo Polgan*, vol. 13, no. 1, pp. 794–803, Jul. 2024, doi: 10.33395/jmp.v13i1.13808.
- [25] G. Altansukh and D. R. Osborn, "Using structural break inference for forecasting time series," *Empir. Econ.*, vol. 63, no. 1, pp. 1–41, Jul. 2022, doi: 10.1007/s00181-021-02137-w.
- [26] R. Ningombam, C. Singh, S. Sreekumar, R. Bhakar, and S. Padmanaban, "Box-Cox integrated sARIMA model for day-ahead inertia forecasting," *Electrical Engineering*, vol. 107, no. 7, pp. 9135–9153, Jul. 2025, doi: 10.1007/s00202-025-02965-4.
- [27] Z. Guo, "Research on the Augmented Dickey-Fuller Test for Predicting Stock Prices and Returns," *Advances in Economics, Management and Political Sciences*, vol. 44, no. 1, pp. 101–106, Nov. 2023, doi: 10.54254/2754-1169/44/20232198.
- [28] N. A. Achسانی, T. Irawan, A. Holis, and H. Hasanah, *EKONOMETRIKA Dengan Aplikasi dalam Bidang Ekonomi dan Keuangan*. Bogor: PT Penerbit IPB Press, 2023. Accessed: Jun. 15, 2026. [Online]. Available: https://www.researchgate.net/publication/396099756_EKONOMETRIKA_Dengan_Aplikasi_dalam_Bidang_Ekonomi_dan_Kuangan
- [29] H. Okut, "Deep Learning for Subtyping and Prediction of Diseases: Long-Short Term Memory," in *Deep Learning Applications*, IntechOpen, 2021. doi: 10.5772/intechopen.96180.
- [30] V. Kotu and B. Deshpande, "Time Series Forecasting," in *Data Science*, Elsevier, 2019, pp. 395–445. doi: 10.1016/B978-0-12-814761-0.00012-5.
- [31] H. Khoshvaght, R. R. Permala, A. Razmjou, and M. Khiadani, "A critical review on selecting performance evaluation metrics for supervised machine learning models in wastewater quality prediction," Dec. 01, 2025, Elsevier Ltd. doi: 10.1016/j.jece.2025.119675.
- [32] E. P. Wahyuddin, R. E. Caraka, R. Kurniawan, W. Caesarendra, P. U. Gio, and B. Pardamean, "Improved LSTM hyperparameters alongside sentiment walk-forward validation for time series prediction," *Journal of Open Innovation: Technology, Market, and Complexity*, vol. 11, no. 1, Mar. 2025, doi: 10.1016/j.joitmc.2024.100458.