Jurnal Riset Aplikasi Ilmu Data dan Sistem 2 (1), 1–16 DOI: 0000
Januari 2027

Analisis Sentimen Publik terhadap Saham Sektor Pertambangan di Media Sosial X Menggunakan IndoBERT dan Implementasinya untuk Prediksi Harga Berbasis LSTM

Muhammad Abduh[1], Atik Wintarti[1], Ike Fitriyaningsih[2,*]

[1]*Program Studi Sarjana Sains Data, Universitas Negeri Surabaya, Indonesia*
[2]*Program Studi Sarjana Kecerdasan Artifisial, Universitas Negeri Surabaya, Indonesia*

Abstrak

Saham sektor pertambangan di Indonesia memiliki volatilitas tinggi karena sensitif terhadap harga komoditas global dan opini publik di media sosial, khususnya pada emiten emas PT Aneka Tambang Tbk (ANTM). Penelitian ini mengembangkan model prediksi harga saham hibrida yang mengintegrasikan sentimen publik dari Platform X (Twitter) berbasis IndoBERT dengan model deret waktu *Long Short-Term Memory* (LSTM). Model `IndoBERT-base-p2` dengan *class weighting* dilatih melalui tiga skenario (*fine-tuning* pada tweet, *fine-tuning* pada berita, dan *domain adaptation* dua tahap) untuk mengklasifikasikan teks ke dalam kelas positif, netral, dan negatif. Pada validasi manual terhadap 300 sampel data riil Platform X, model berbasis tweet memberikan generalisasi terbaik (akurasi 83,33%, F1-macro 0,8292, F1-weighted 0,8375) sehingga dipilih untuk melabeli 27.181 tweet periode Januari–Desember 2025. Skor keyakinan klasifikasi diagregasikan menjadi skor sentimen harian (S_{daily}) dan digabungkan dengan lima fitur teknikal (*Open, High, Low, Close, Volume*) untuk memprediksi harga penutupan ANTM satu hari ke depan dengan pembagian data kronologis 90:10. Pada data uji (19 hari bursa, Desember 2025), model hibrida IndoBERT-LSTM menghasilkan galat lebih kecil dibandingkan LSTM *baseline*: R^2 meningkat dari 0,4950 menjadi 0,7443 dan MAPE pada skala harga asli turun dari 2,55% menjadi 2,05%. Hasil ini menunjukkan bahwa skor sentimen publik berbasis Transformer berpotensi menjadi fitur tambahan yang informatif bagi prediksi harga saham, meskipun peran sentimen sebagai indikator awal (*leading indicator*) masih memerlukan pengujian pada periode uji yang lebih panjang.

Kata Kunci *Analisis sentimen; ANTM; IndoBERT; LSTM; Prediksi harga saham; Aplikasi X*

Abstract

Mining sector stocks in Indonesia are highly volatile due to their sensitivity to global commodity prices and public opinion on social media, particularly for the gold-mining issuer PT Aneka Tambang Tbk (ANTM). This study develops a hybrid stock price prediction model that integrates IndoBERT-based public sentiment from Platform X (Twitter) with a Long Short-Term Memory (LSTM) time-series model. `IndoBERT-base-p2` with class weighting was trained under three scenarios (fine-tuning on tweets, fine-tuning on news, and two-stage domain adaptation) to classify texts as positive, neutral, or negative. In a manual validation on 300 real Platform X samples, the tweet-trained model generalized best (accuracy 83.33%, macro-F1 0.8292, weighted-

*Corresponding author. Email: muhammadabduh.22032@mhs.unesa.ac.id or atikwintarti@unesa.ac.id or ike-fitriyaningsih@unesa.ac.id.

Received 28 Juli 2026; Revised 16 November 2026; Accepted 17 November 2026

F1 0.8375) and was used to label 27,181 tweets from January to December 2025. Classification confidence scores were aggregated into a daily sentiment score (S_{daily}) and combined with five technical features (*Open, High, Low, Close, Volume*) to predict ANTM's next-day closing price using a chronological 90:10 split. On the test set (19 trading days, December 2025), the hybrid IndoBERT-LSTM model produced lower errors than the baseline LSTM: R^2 increased from 0.4950 to 0.7443 and MAPE on the original price scale decreased from 2.55% to 2.05%. These results suggest that Transformer-based public sentiment scores can be an informative additional feature for stock price prediction, although the role of sentiment as a leading indicator requires testing over longer evaluation periods.

Keywords *Sentiment analysis; ANTM; IndoBERT; LSTM; Stock price prediction; Platform X*

1 Pendahuluan

Pasar modal berperan penting dalam perekonomian sebagai sarana pendanaan perusahaan sekaligus peluang investasi bagi masyarakat. Pergerakan harga saham bersifat dinamis dan dipengaruhi oleh faktor fundamental perusahaan maupun faktor eksternal seperti kondisi ekonomi, kebijakan pemerintah, dan sentimen investor (Adeleye dkk., 2024). Perkembangan media sosial telah mengubah cara investor memperoleh informasi. Platform X (Twitter) banyak digunakan untuk menyampaikan opini terhadap kondisi perusahaan dan pergerakan saham, dan informasi yang menyebar cepat melalui media sosial dapat memengaruhi persepsi investor serta harga saham dalam jangka pendek (Jiang dan Zeng, 2025).

Analisis sentimen digunakan untuk mengidentifikasi opini publik ke dalam kategori positif, netral, atau negatif. Model berbasis Transformer menunjukkan performa unggul pada berbagai tugas *Natural Language Processing* (NLP), dan untuk Bahasa Indonesia tersedia IndoBERT yang dilatih pada korpus berbahasa Indonesia (Koto dkk., 2020; Wilie dkk., 2020). Pada domain keuangan Indonesia, IndoBERT telah digunakan untuk mengklasifikasikan sentimen judul berita saham CNBC Indonesia (Jiwandono dkk., 2025), dan tersedia pula korpus tweet saham berlabel ID-SMSA (Hartanto dkk., 2025).

Di sisi lain, prediksi harga saham umumnya dilakukan menggunakan data historis deret waktu. *Long Short-Term Memory* (LSTM) banyak digunakan karena mampu mempelajari ketergantungan jangka panjang dan mengatasi *vanishing gradient* pada *Recurrent Neural Network* (RNN) (Fahira dan Prianto, 2023). Pada saham ANTM, LSTM berbasis harga penutupan mencapai MAPE 2,64% (Budiprasetyo dkk., 2022). Namun, data numerik saja sering belum cukup menggambarkan kondisi pasar, dan sejumlah studi menunjukkan bahwa penambahan sentimen berita atau media sosial dapat memperbaiki akurasi prediksi harga saham (Srinivas dkk., 2022; Jiang dan Zeng, 2025; Ruan dan Jiang, 2025; Anggraini dkk., 2025).

Terdapat dua celah yang belum banyak dibahas pada konteks pasar modal Indonesia. Pertama, strategi *fine-tuning* IndoBERT umumnya dipilih hanya berdasarkan data uji internal yang berasal dari distribusi yang sama dengan data latih, padahal model diterapkan pada data riil media sosial yang karakteristik bahasanya berbeda. Kedua, sentimen umumnya diintegrasikan sebagai label diskrit atau proporsi kelas, bukan sebagai skor kontinu yang mencerminkan keyakinan model. Oleh karena itu, kebaruan (*novelty*) penelitian ini terletak pada: (1) pemilihan strategi *fine-tuning* IndoBERT (tweet, berita, dan *domain adaptation*) yang divalidasi langsung pada data riil Platform X, bukan hanya pada data uji internal; dan (2) pemanfaatan skor sentimen berbasis keyakinan model sebagai indikator sentimen kontinu harian (S_{daily}) yang diintegrasikan ke dalam arsitektur *sliding window* LSTM untuk memprediksi harga saham komoditas pertambangan (ANTM) di

Bursa Efek Indonesia. Penelitian ini bertujuan (i) membangun dan memilih model klasifikasi sentimen IndoBERT terbaik berdasarkan validasi pada data riil, serta (ii) mengukur kontribusi fitur sentimen harian terhadap kinerja LSTM dibandingkan model yang hanya menggunakan fitur teknikal.

2 Metode

2.1 Dataset

Penelitian ini menggunakan dua kelompok data, yaitu data tekstual untuk pelatihan dan penerapan model sentimen (Tabel 1) dan data historis harga saham untuk pemodelan deret waktu (Tabel 2).

2.1.1 Dataset Tekstual (Sentimen)

Dataset tekstual terdiri atas dua korpus berlabel terbuka dan satu korpus hasil *scraping*. Korpus berlabel pertama adalah judul berita saham CNBC Indonesia yang telah dianotasi sentimennya (Jiwandono dkk., 2025), dan korpus kedua adalah *Indonesian Stock Market Sentiment Analysis* (ID-SMSA) yang berisi tweet saham berbahasa Indonesia berlabel (Hartanto dkk., 2025). Keduanya digunakan untuk pelatihan dan evaluasi IndoBERT. Korpus ketiga adalah tweet hasil *scraping* Platform X terkait saham ANTM dan harga emas yang tidak berlabel dan digunakan sebagai sumber sentimen publik.

Tabel 1: Rincian Dataset Tekstual (Sentimen)

Dataset (sumber)	**Jenis teks**	**Pemanfaatan**	**Mentah**	**Bersih**
CNBC Indonesia (Jiwandono dkk., 2025)	Judul berita, berlabel	Pelatihan & evaluasi IndoBERT (Skenario 2 & 3)	9.819	9.738
ID-SMSA (Hartanto dkk., 2025)	Tweet saham, berlabel	Pelatihan & evaluasi IndoBERT (Skenario 1 & 3)	3.288	3.204
Scraping X ANTM (penelitian ini)	Tweet, tidak berlabel	Pelabelan otomatis & pembentukan S_{daily}	29.074	27.181
Total dataset teks			**42.181**	**40.123**

Data X dikumpulkan menggunakan *tweet-harvest* v2.6.1 (mode *Latest*, filter `lang:id`) dengan kueri "*Saham sektoral pertambangan emas*" OR "*saham PT Aneka Tambang Tbk*" OR "*saham antm*" OR "*emiten antm*" OR "*PT Aneka Tambang Tbk*" OR "*harga emas*" OR "*saham antam*", disertai pengecualian kata kunci *spam* (`-slot -togel -judi -judol -gacor -casino -kasino -bigwin -chundai -bandar`). Periode pengamatan mencakup 1 Januari–31 Desember 2025. Pengambilan dilakukan bertahap per rentang tanggal (`since`/`until`, maksimal 900 tweet per kueri) dan menghasilkan 12 berkas yang kemudian digabungkan. Tweet tersedia pada 364 dari 365 hari, atau rata-rata ±74,7 tweet unik per hari. Sebaran lokasi tidak dianalisis karena atribut lokasi pada Platform X merupakan isian profil yang opsional dan tidak terverifikasi; cakupan populasi ditentukan melalui filter bahasa Indonesia. Distribusi label dan contoh sampel tiap label disajikan pada Subbab 3.1.

2.1.2 Data Historis Harga Saham

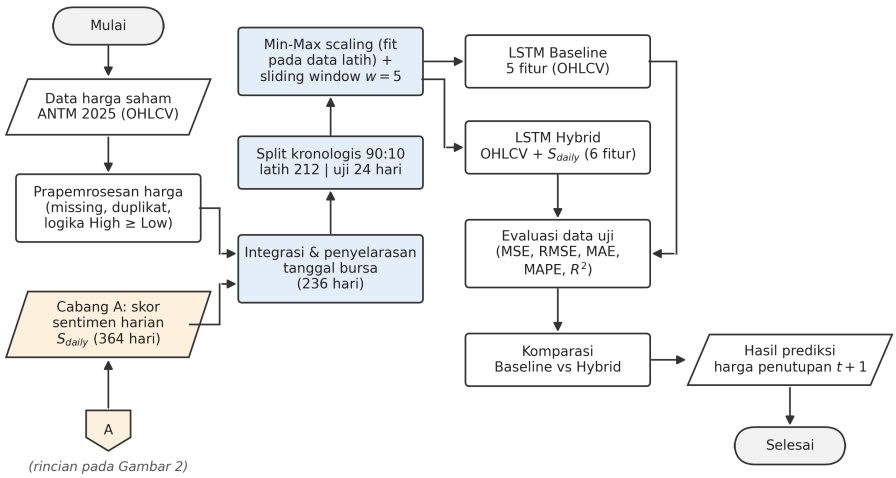
Data harga harian ANTM diperoleh dari Mirae Asset Sekuritas dengan atribut *Open*, *High*, *Low*, *Close*, dan *Volume* (OHLCV). Setelah difilter pada periode 2 Januari–30 Desember 2025 dan diselaraskan dengan hari bursa, diperoleh 236 baris data (Tabel 2).

Tabel 2: Rincian Data Historis Harga Saham

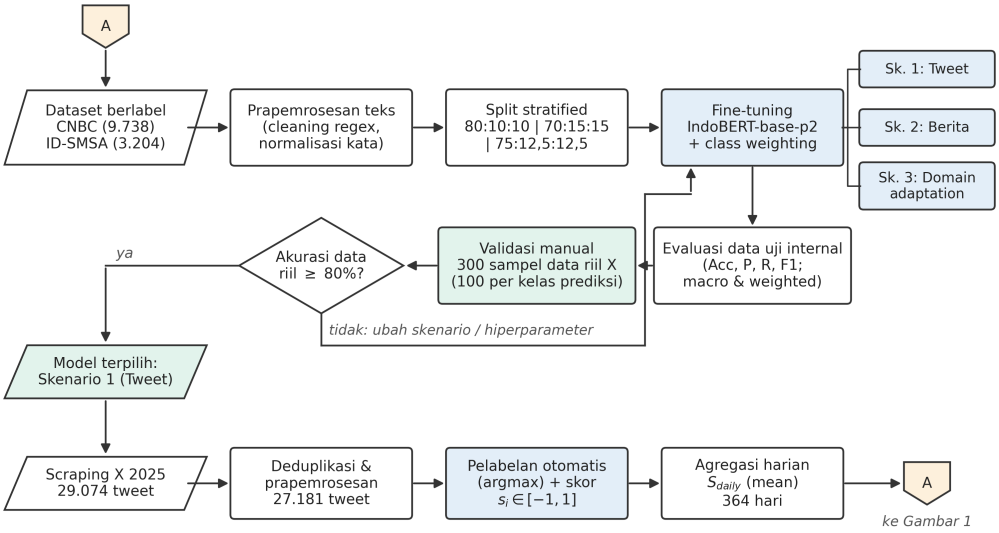
Dataset (sumber)	**Atribut**	**Pemanfaatan**	**Jumlah**
Harga harian ANTM (Mirae Asset Sekuritas)	Date, Open, High, Low, Close, Volume	Fitur teknikal & target (*Close* $t+1$) LSTM	260 baris mentah
Data integrasi (penelitian ini)	OHLCV + S_{daily}	Input model *baseline* dan *hybrid*	236 hari bursa 2025

2.2 Alur Penelitian

Model hibrida yang diusulkan terdiri atas dua cabang, yaitu cabang analisis sentimen teks dan cabang peramalan deret waktu (Gambar 1). Konektor 'A' merujuk pada cabang analisis sentimen (Gambar 2) yang menghasilkan S_{daily}. Hasil tersebut diselaraskan dengan hari bursa, digabungkan dengan fitur OHLCV, dibagi secara kronologis, dinormalisasi, dan dibentuk menjadi sekuens untuk melatih LSTM *baseline* dan *hybrid*, lalu kinerja keduanya dibandingkan. Pada cabang sentimen, IndoBERT di-*fine-tune* melalui tiga skenario, dievaluasi pada data uji internal, lalu divalidasi manual pada 300 sampel data riil. Model dengan generalisasi terbaik (target akurasi minimal 80% pada data riil) digunakan untuk melabeli seluruh data *scraping*, dan hasilnya diagregasikan menjadi skor harian.



Gambar 1: Diagram Alir Penelitian Utama



Gambar 2: Diagram Alir Cabang Analisis Sentimen (Konektor A)

2.3 Arsitektur dan Konfigurasi IndoBERT

IndoBERT adalah model Transformer *encoder* dua arah berbasis BERT (Devlin dkk., 2019) yang dilatih pada korpus Bahasa Indonesia (Koto dkk., 2020; Wilie dkk., 2020). Teks ditokenisasi menjadi *input IDs* dan *attention masks* dengan *truncation* dan *dynamic padding*, lalu diproses melalui *scaled dot-product self-attention* (Vaswani dkk., 2017):

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{1}$$

dengan Q, K, dan V masing-masing *query*, *key*, dan *value*, serta d_k dimensi vektor *key*. Representasi token *[CLS]* diteruskan ke lapisan klasifikasi, dan probabilitas kelas dihitung dengan softmax, $P(y_i) = e^{z_i} / \sum_{j=1}^{C} e^{z_j}$, dengan z_i logit kelas ke-i dan $C = 3$. Pelatihan menggunakan *weighted cross-entropy loss*:

$$L = -\sum_{i=1}^{C} w_i \, y_i \log(\hat{y}_i) \tag{2}$$

dengan y_i label sebenarnya, $\hat{y}_i$ probabilitas prediksi, dan w_i bobot kelas *balanced* dari data latih (bobot kelas netral dikalikan 1,1). Konfigurasi utama: *learning rate* $1{,}5 \times 10^{-5}$ (diturunkan menjadi 1×10^{-5} pada tahap kedua Skenario 3), *effective batch size* 32, maksimal 8 *epoch* (5 pada Skenario 2) dengan *early stopping* (*patience* 2–3), *dropout hidden/attention/classifier* 0,20/0,10/0,20 (0,25/0,15/0,25 pada Skenario 2 dan tahap kedua Skenario 3), *weight decay* 0,03–0,04, *warmup* 10–15% dengan *cosine scheduler*, *label smoothing* 0,1, panjang maksimum 128 token untuk tweet serta 32 token (Skenario 2) dan 64 token (tahap pertama Skenario 3) untuk berita, dan pembekuan tiga lapisan *encoder* terbawah. *Checkpoint* terbaik dipilih berdasarkan F1-macro data validasi.

Tiga skenario dibandingkan: Skenario 1 (*fine-tuning* pada tweet ID-SMSA), Skenario 2 (*fine-tuning* pada berita CNBC), dan Skenario 3 (*domain adaptation*: *fine-tuning* pada berita,

lalu pada tweet). Setiap skenario diuji dengan pembagian *stratified train:validation:test* 80:10:10 dan 70:15:15 (Skenario 3B: 75:12,5:12,5).

2.4 Perhitungan dan Agregasi Skor Sentimen Harian

Label akhir ditentukan berdasarkan probabilitas tertinggi (*argmax*). Agar intensitas sentimen tertangkap, digunakan penskoran berbasis keyakinan (*confidence-based scoring*) (Ruan dan Jiang, 2025):

$$s_i = \begin{cases} +P_{\max}, & \text{jika label} = \text{positif} \\ 0, & \text{jika label} = \text{netral} \\ -P_{\max}, & \text{jika label} = \text{negatif} \end{cases} \tag{3}$$

dengan $P_{\max}$ probabilitas softmax tertinggi, sehingga $s_i \in [-1, 1]$. Skor harian pada hari ke-d dihitung sebagai rata-rata skor seluruh n_d unggahan pada hari tersebut:

$$S_{\text{daily}}(d) = \frac{1}{n_d} \sum_{i=1}^{n_d} s_i \tag{4}$$

Agregasi rata-rata dipilih untuk menangkap kecenderungan sentimen agregat dan mengurangi bias akibat perbedaan jumlah unggahan antarhari. Proses ini menghasilkan 364 titik data harian; hari tanpa tweet diberi nilai 0. S_{daily} kemudian diselaraskan dengan kalender hari bursa melalui *left join* terhadap data harga (236 baris), digabungkan dengan fitur OHLCV (Anggraini dkk., 2025), dan dinormalisasi ke rentang $[0, 1]$ dengan *Min-Max Scaling* (Faqih dan Sugihartono, 2025).

2.5 Arsitektur LSTM

LSTM merupakan pengembangan RNN yang mampu mempelajari ketergantungan jangka panjang pada data deret waktu (Adeleye dkk., 2024). Pada setiap langkah waktu t, *forget gate* (f_t), *input gate* (i_t), kandidat memori ($\tilde{C}_t$), *cell state* (C_t), *output gate* (o_t), dan *hidden state* (h_t) dihitung dengan Persamaan 5:

$$\begin{aligned} f_t &= \sigma(W_f[h_{t-1}, x_t] + b_f), & i_t &= \sigma(W_i[h_{t-1}, x_t] + b_i), \\ \tilde{C}_t &= \tanh(W_c[h_{t-1}, x_t] + b_c), & C_t &= f_t \odot C_{t-1} + i_t \odot \tilde{C}_t, \\ o_t &= \sigma(W_o[h_{t-1}, x_t] + b_o), & h_t &= o_t \odot \tanh(C_t), \end{aligned} \tag{5}$$

dengan x_t input, W dan b bobot dan bias, σ fungsi sigmoid, dan $\odot$ perkalian elemen demi elemen. *Hidden state* lapisan terakhir digunakan untuk mengestimasi harga penutupan hari berikutnya $(t + 1)$ (Ruan dan Jiang, 2025).

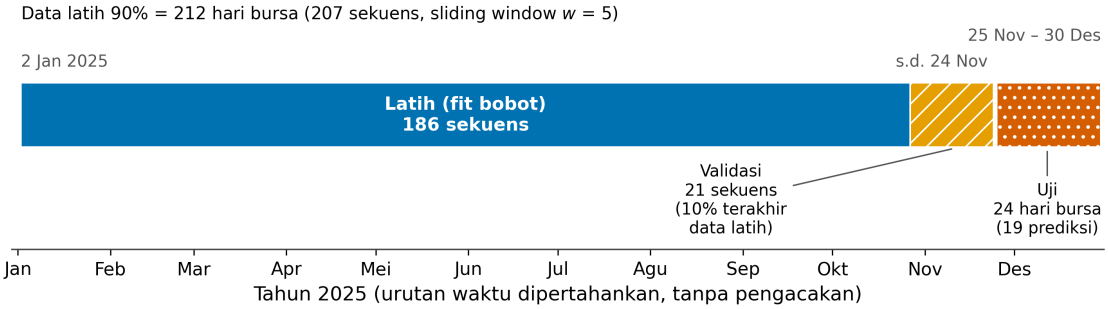
Detail arsitektur jaringan. Model *baseline* dan *hybrid* menggunakan arsitektur dan hiperparameter identik; perbedaannya hanya pada jumlah fitur input F ($F = 5$ untuk OHLCV; $F = 6$ untuk OHLCV + S_{daily}). Rincian lapisan disajikan pada Tabel 3. Karena LSTM tidak menggunakan operasi konvolusi, tidak terdapat parameter ukuran *kernel*; kapasitas jaringan ditentukan oleh jumlah unit tersembunyi, dan jumlah parameter lapisan LSTM dihitung dengan $4\,[u(u + F) + u]$ untuk u unit.

Prosedur pembagian data deret waktu. Untuk mencegah kebocoran informasi masa depan (*look-ahead bias*), pembagian data dilakukan secara kronologis tanpa pengacakan (Hyndman dan Athanasopoulos, 2021) (Gambar 3). Sebanyak 90% data pertama (212 hari; 2 Januari–24 November 2025) menjadi data latih dan 10% terakhir (24 hari; 25 November–30 Desember 2025)

Tabel 3: Detail Arsitektur dan Hiperparameter Model LSTM

No.	Lapisan (jenis)	Konfigurasi	Bentuk output	Parameter (F=**5** / F=**6**)
0	Input	*window* 5 hari × F fitur	(5, F)	–
1	LSTM-1	50 unit, aktivasi tanh, *return_sequences*=True	(5, 50)	11.200 / 11.400
2	Dropout	$p = 0{,}2$	(5, 50)	0
3	LSTM-2	50 unit, aktivasi tanh, *return_sequences*=False	(50)	20.200
4	Dropout	$p = 0{,}2$	(50)	0
5	Dense (output)	1 unit, aktivasi linear	(1)	51
Total parameter terlatih				**31.451 / 31.651**
Optimizer Adam ($\eta = 0{,}001$), *loss* MSE, 50 *epoch*, *batch size* 32, tanpa *early stopping*.				

menjadi data uji. Pembagian dilakukan **sebelum** normalisasi, dan *Min-Max Scaler* hanya di-*fit* pada data latih. *Sliding window* 5 hari diterapkan terpisah pada data latih dan uji untuk memprediksi *Close* hari berikutnya, menghasilkan 207 sekuens latih dan 19 sekuens uji. Data validasi diambil dari 10% sekuens latih paling akhir secara kronologis (21 sekuens; `validation_split`=0,1 pada Keras), sehingga model dilatih pada 186 sekuens. Data uji tidak digunakan selama pelatihan maupun pemilihan konfigurasi.



Gambar 3: Skema Pembagian Data Deret Waktu Secara Kronologis (Latih, Validasi, dan Uji)

2.6 Evaluasi Model

Karena penelitian ini mencakup tugas klasifikasi (sentimen) dan regresi (harga saham), evaluasi dilakukan dengan dua kelompok metrik.

2.6.1 Metrik Klasifikasi Sentimen

Kinerja IndoBERT diukur dengan *accuracy*, *precision*, *recall*, dan F1-*score* dari *confusion matrix*. Untuk kelas c dengan TP_c, FP_c, dan FN_c (*true positive*, *false positive*, *false negative*) dan N total data:

$$Acc = \frac{\sum_c TP_c}{N}, \quad P_c = \frac{TP_c}{TP_c + FP_c}, \quad R_c = \frac{TP_c}{TP_c + FN_c}, \quad F1_c = \frac{2P_cR_c}{P_c + R_c} \tag{6}$$

Nilai agregat dilaporkan sebagai *macro average* ($\frac{1}{C}\sum_c M_c$), yang memberi bobot sama pada setiap kelas sehingga sensitif terhadap kelas minoritas, dan *weighted average* ($\sum_c \frac{n_c}{N} M_c$, dengan n_c jumlah data aktual kelas c), yang mencerminkan proporsi kelas. Keduanya dilaporkan karena distribusi kelas tidak seimbang.

2.6.2 Metrik Regresi (Prediksi Harga)

Prediksi harga dievaluasi dengan MSE, RMSE, MAE, MAPE, dan koefisien determinasi R^2 (Adeleye dkk., 2024):

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|, \qquad RMSE = \sqrt{MSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{7}$$

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right|, \qquad R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2} \tag{8}$$

dengan y_i nilai aktual, $\hat{y}_i$ prediksi, $\bar{y}$ rata-rata aktual, dan n jumlah data uji. Seluruh metrik dihitung pada skala ternormalisasi, sedangkan MAPE juga dihitung pada skala harga asli (Rupiah) setelah *inverse-transform* agar mudah diinterpretasikan dan dapat dibandingkan dengan penelitian lain. Interpretasi MAPE mengacu pada Lewis (1982) ($< 10\%$: sangat akurat).

2.7 Prosedur Validasi Manual pada Data Riil

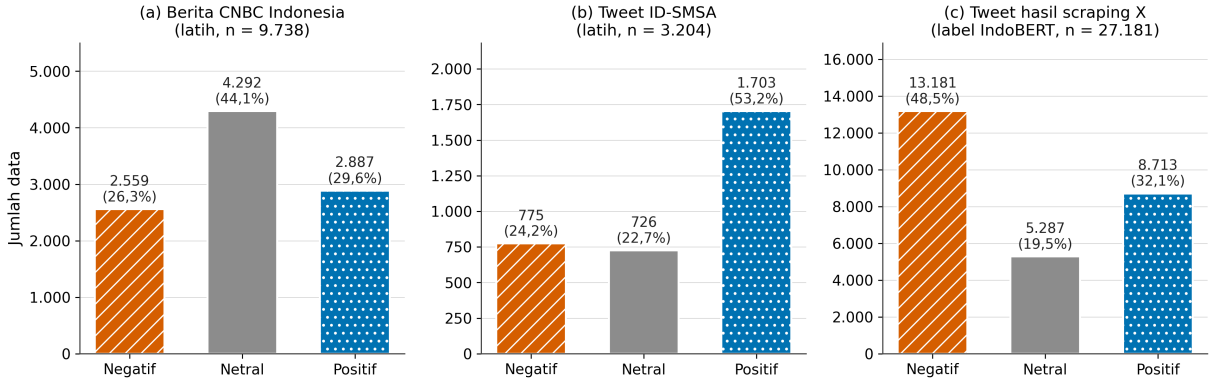
Data uji internal setiap skenario berasal dari distribusi yang sama dengan data latihnya, sehingga belum tentu mencerminkan kinerja pada data target. Oleh karena itu, model terbaik setiap skenario (Skenario 1 dan 3A rasio 80:10:10; Skenario 2 rasio 70:15:15) digunakan untuk melabeli seluruh 27.181 tweet. Dari hasil pelabelan setiap model diambil *stratified random sampling* 100 tweet per kelas prediksi (total 300; *random seed* 42) dengan urutan diacak, agar kelas netral yang paling sedikit tetap terwakili. Setiap sampel dianotasi manual oleh peneliti dengan pedoman: **positif** bila menyatakan optimisme, kenaikan harga, atau rekomendasi beli; **negatif** bila menyatakan kekhawatiran, penurunan harga, kerugian, atau keluhan; dan **netral** bila bersifat informatif tanpa opini (misalnya daftar harga harian). Label aktual dibandingkan dengan label prediksi menggunakan metrik per kelas, *macro*, dan *weighted*, disertai *confusion matrix*. Perlu ditegaskan bahwa 300 sampel ini berperan sebagai **data pemilihan model** (*model selection set*), **bukan** data uji independen, sehingga akurasi model terpilih berpotensi sedikit optimistis. Karena sampel seimbang per kelas prediksi, akurasi populasi diestimasi dengan $\sum_c \pi_c P_c$, dengan π_c proporsi kelas prediksi c pada seluruh 27.181 tweet.

3 Hasil dan Pembahasan

3.1 Pengumpulan dan Eksplorasi Data

Distribusi label ketiga korpus teks ditunjukkan pada Gambar 4. Dataset berita CNBC didominasi kelas netral (44,1%), diikuti positif (29,6%) dan negatif (26,3%), sesuai karakter berita finansial yang informatif dan objektif. Sebaliknya, ID-SMSA didominasi sentimen positif (53,2%), diikuti negatif (24,2%) dan netral (22,7%), yang mengindikasikan optimisme (*bullish sentiment*) investor ritel. Data X terkumpul sebanyak 29.074 tweet dari 12 berkas *scraping*; setelah 1.893 duplikat dihapus diperoleh 27.181 tweet unik. Contoh tweet hasil pelabelan untuk setiap label disajikan pada Tabel 4. Tweet negatif umumnya memuat pemberitaan penurunan harga yang ditandai kata ekspresif seperti “turun tajam” dan “merosot”; tweet netral berupa pelaporan harga harian atau pembahasan pasar tanpa kecenderungan opini; sedangkan tweet positif berkaitan dengan kenaikan harga seperti “menguat” dan “melonjak”. Contoh tersebut juga menunjukkan bahwa

sebagian tweet membahas harga emas secara umum (misalnya emas di Pegadaian), bukan hanya saham ANTM, sebagai konsekuensi kata kunci "harga emas".



Gambar 4: Distribusi Label Sentimen: (a) Berita CNBC Indonesia, (b) Tweet ID-SMSA, dan (c) Tweet Hasil *Scraping* X yang Dilabeli Model IndoBERT Terpilih

Tabel 4: Contoh Hasil Pelabelan Tweet *Scraping* X untuk Setiap Label (setelah prapemrosesan)

No.	**Teks (*clean_text*)**	**Label**
1	harga emas di abdya turun tajam segini harga emas per mayam dan emas antam	Negatif
2	harga emas merosot hari ini jadi segini	Negatif
3	lihat harga emas galeri 24 dan ubs serta buyback antam di pegadaian selasa ...	Netral
4	market itu semua manusia bang kalau misalnya ada kabar geopolitik yang membingungkan	Netral
5	logam mulia buatan antam naik seiring harga emas dunia yang kembali menguat	Positif
6	harga emas xau hari ini 26 november 2025 melonjak ke 4161	Positif

3.2 Karakteristik Dataset dan Prapemrosesan

Prapemrosesan teks mencakup: (1) penghapusan *tag [USERNAME]/[URL]/[HASHTAG]*, URL, *mention*, penanda *retweet*, *hashtag*, simbol *ticker* ($), dan emoji dengan *regular expression*; (2) penggabungan pemisah ribuan (100,000 → 100000) agar angka finansial tetap utuh; (3) penghapusan karakter non-alfanumerik dan *case folding*; (4) reduksi huruf berulang ("naikkk" → "naik"); dan (5) normalisasi kata tidak baku dengan kamus alay dan kamus singkatan. *Stemming* dan *stopword removal* tidak dilakukan karena IndoBERT memanfaatkan *contextual embedding*, sehingga penghapusan imbuhan dan kata fungsi (misalnya negasi "tidak") berpotensi menghilangkan makna. Kondisi data sebelum dan sesudah prapemrosesan disajikan pada Tabel 5 dan Tabel 6.

Setelah prapemrosesan, panjang tweet ID-SMSA rata-rata 26,7 token IndoBERT (persentil ke-95: 54; maksimum: 71), sehingga batas 128 token tidak memotong teks. Pada data harga tidak ditemukan *missing value*, duplikat, maupun anomali High < Low. Harga penutupan ANTM selama 2025 berkisar Rp1.380–Rp3.670 (rata-rata Rp2.623,79), sedangkan S_{daily} pada 236 hari bursa memiliki rata-rata 0,0551 (simpangan baku 0,1960; rentang −0,4186 hingga 0,4746), yang menunjukkan variasi opini pasar dari negatif hingga positif sepanjang periode.

Tabel 5: Jumlah Data Sebelum dan Sesudah Prapemrosesan

Dataset	**Sebelum**	**Sesudah**	**Dihapus**	**Keterangan**
Berita CNBC Indonesia	9.819	9.738	81	Duplikat teks & teks $\leqslant 5$ karakter
Tweet ID-SMSA	3.288	3.204	84	Duplikat teks & teks $\leqslant 5$ karakter
Tweet *scraping* X	29.074	27.181	1.893	Duplikat teks
Harga saham ANTM	260	236	24	Difilter ke periode 2025 & hari bursa; tidak ada *missing value*, duplikat, atau High < Low

Tabel 6: Contoh Teks Sebelum dan Sesudah Prapemrosesan

Sebelum	**Sesudah**
Gk muluk muluk, 100,000 lot saham BBCA aja	tidak muluk muluk 100000 lot saham bbca saja
[USERNAME] Drpd bunga mending hal^2 yg berguna, ...	daripada bunga mending hal yang berguna ...
[USERNAME] Begitu byk saham kamu memilih saham itu kalau uangnya gede bisa nyicil ...	begitu banyak saham kamu memilih saham itu kalau uangnya gede bisa mencicil ...
Jogj harga emas hari ini masih tinggi?	jogj harga emas hari ini masih tinggi

3.3 Evaluasi Performa Klasifikasi Model IndoBERT

Hasil seluruh skenario pada data uji internal disajikan pada Tabel 7.

Tabel 7: Hasil Evaluasi IndoBERT pada Data Uji Internal Setiap Skenario

Skenario	**Split**	n **uji**	**Accuracy**	**Precision**		**Recall**		**F1-macro / F1-weighted**
				Macro	**Weighted**	**Macro**	**Weighted**	
1 *Fine-tuning* Tweet	80:10:10	321	80,06%	0,7770	0,8057	0,7949	0,8006	0,7847 / 0,8019
1 *Fine-tuning* Tweet	70:15:15	481	79,63%	0,7689	0,7983	0,7777	0,7963	0,7723 / 0,7964
2 *Fine-tuning* Berita	80:10:10	974	87,06%	0,8656	0,8790	0,8848	0,8706	0,8711 / 0,8704
2 *Fine-tuning* Berita	70:15:15	1.461	87,61%	0,8700	–*	0,8792	0,8761	0,8739 / –*
3A *Domain Adaptation*	80:10:10	321	80,37%	0,7763	0,8086	0,7930	0,8037	0,7837 / 0,8052
3B *Domain Adaptation*	75:12,5:12,5	401	79,55%	0,7683	0,8090	0,7960	0,7955	0,7788 / 0,7989

Data uji Skenario 1 dan 3 berupa tweet ID-SMSA; data uji Skenario 2 berupa berita CNBC. *Recall weighted* selalu sama dengan *accuracy*.
**Precision* dan F1 *weighted* pada *run* ini tidak tercatat.

Beberapa pola dapat dijelaskan dari Tabel 7. **Pertama**, Skenario 2 memperoleh nilai tertinggi (87,06–87,61%) karena data latihnya tiga kali lebih besar (6.816–7.790 vs. 2.242–2.562 sampel), bahasanya formal dan konsisten, serta distribusi kelasnya relatif seimbang; namun angka ini tidak dapat dibandingkan langsung dengan Skenario 1 dan 3 karena diuji pada domain teks berbeda. **Kedua**, pada data tweet, nilai *weighted* (F1 $\approx$ 0,80) konsisten lebih tinggi daripada nilai *macro* (F1 $\approx$ 0,77–0,78) akibat ketidakseimbangan kelas: kelas positif yang dominan memiliki F1 tertinggi (0,84–0,85), sedangkan kelas netral terendah (0,69–0,72) karena tweet informatif sering memuat kata bermuatan sentimen (misalnya "naik" atau "turun") tanpa opini. **Ketiga**, *recall-macro* selalu lebih tinggi daripada *precision-macro* sebagai efek *class weighting* yang mendorong model lebih sering memprediksi kelas minoritas. **Keempat**, *domain adaptation* (3A) hanya menaikkan

akurasi 0,31 poin persentase dibandingkan Skenario 1 dengan F1-macro hampir sama (0,7837 vs. 0,7847), karena irisan kosakata tweet dan berita kecil (*Jaccard similarity* 26,9%) sehingga sedikit pengetahuan yang dapat ditransfer.

Hasil validasi manual pada 300 sampel data riil disajikan pada Tabel 8, metrik per kelas pada Tabel 9, dan *confusion matrix* pada Gambar 5.

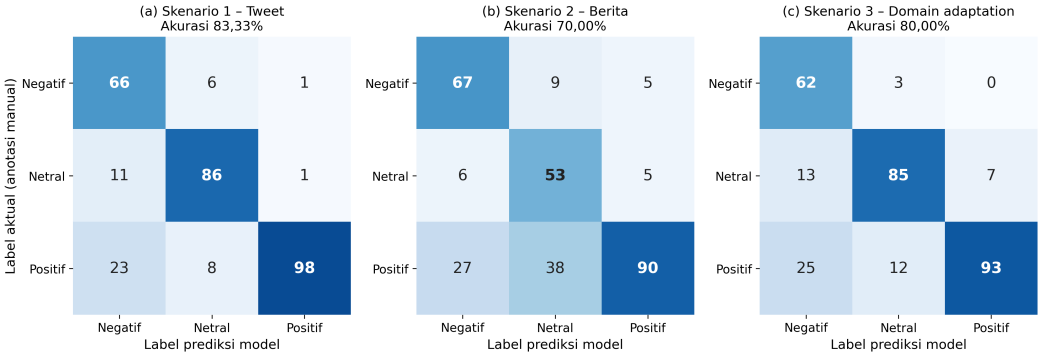
Tabel 8: Hasil Validasi Manual Setiap Skenario pada 300 Sampel Data Riil Platform X

Skenario	**Accuracy**	**Precision**		**Recall**		**F1-score**		**Estimasi akurasi**
		Macro	**Weighted**	**Macro**	**Weighted**	**Macro**	**Weighted**	**populasi**
1 Tweet (80:10:10)	**83,33%**	**0,8333**	**0,8629**	**0,8471**	**0,8333**	**0,8292**	**0,8375**	**80,15%**
2 Berita (70:15:15)	70,00%	0,7000	0,7590	0,7453	0,7000	0,6975	0,7025	64,27%
3A Domain Adaptation (80:10:10)	80,00%	0,8000	0,8348	0,8263	0,8000	0,7965	0,8035	77,78%

Sampel diambil 100 per kelas prediksi sehingga *precision-macro* sama dengan *accuracy*.

Tabel 9: Metrik per Kelas pada Validasi Manual 300 Sampel

Skenario	**Kelas**	**Precision**	**Recall**	**F1**	**Support (aktual)**
1 Tweet	Negatif	0,66	0,90	0,76	73
	Netral	0,86	0,88	0,87	98
	Positif	0,98	0,76	0,86	129
2 Berita	Negatif	0,67	0,83	0,74	81
	Netral	0,53	0,83	0,65	64
	Positif	0,90	0,58	0,71	155
3A Domain Adaptation	Negatif	0,62	0,95	0,75	65
	Netral	0,85	0,81	0,83	105
	Positif	0,93	0,72	0,81	130



Gambar 5: *Confusion Matrix* Validasi Manual pada 300 Sampel Data Riil: (a) Skenario 1, (b) Skenario 2, dan (c) Skenario 3A

Hasil validasi menunjukkan pola yang berbeda dari evaluasi internal. Skenario 1 (*Tweet*) mencatat kinerja tertinggi pada semua metrik (akurasi 83,33%, F1-macro 0,8292, F1-weighted 0,8375), diikuti Skenario 3A (80,00%) dan Skenario 2 (70,00%), dengan urutan yang sama pada estimasi akurasi populasi (80,15%, 77,78%, 64,27%). Model Skenario 2 yang unggul pada data internal justru turun paling besar akibat *domain shift*: model berita cenderung memprediksi netral (56,5% dari seluruh tweet) sehingga *precision* netral hanya 0,53 dan 38 dari 155 tweet positif aktual diprediksi netral (Gambar 5b). Pada ketiga model, kesalahan terbesar adalah tweet positif yang

diprediksi negatif (23–27 kasus), sehingga *precision* negatif rendah (0,62–0,67) meskipun *recall*-nya tinggi (0,83–0,95). Kemungkinan penyebabnya adalah ambiguitas sudut pandang: kenaikan harga emas bermakna positif bagi investor ANTM tetapi dikeluhkan calon pembeli emas (misalnya "harga emas hari ini masih tinggi"), ditambah efek *class weighting*. Berdasarkan hasil ini, model Skenario 1 (80:10:10) dipilih sebagai model final.

3.4 Analisis Sentimen Harian dan Integrasi Finansial

Model final melabeli 27.181 tweet dengan hasil didominasi sentimen negatif 48,5% (13.181), disusul positif 32,1% (8.713) dan netral 19,5% (5.287) (Gambar 4c). Tingginya proporsi negatif mengindikasikan banyaknya keluhan dan kekhawatiran, misalnya terhadap harga emas yang tinggi dari sudut pandang konsumen. Cuplikan data integrasi (Tabel 10) menunjukkan harga ANTM bergerak dari kisaran Rp1.400-an pada awal periode hingga Rp3.100–3.300 pada akhir periode, dengan S_{daily} yang berfluktuasi dari positif ke negatif menjelang akhir tahun.

Tabel 10: Cuplikan Dataset Hasil Integrasi Harga Saham ANTM dan Fitur S_{daily}

Date	**Open**	**High**	**Low**	**Close**	**Volume**	$S_{\textbf{daily}}$
2025-01-03	1.560	1.560	1.520	1.520	181.988	0,2584
2025-01-06	1.525	1.540	1.460	1.465	281.695	0,0035
2025-01-08	1.445	1.465	1.425	1.430	344.148	0,2385
			⋮			
2025-12-22	3.110	3.220	3.110	3.190	2.636.258	0,2733
2025-12-29	3.240	3.420	3.240	3.300	2.839.004	−0,1241
2025-12-30	3.270	3.270	3.130	3.150	2.094.175	−0,3197

3.5 Eksperimen Pemodelan LSTM dan Pembahasan

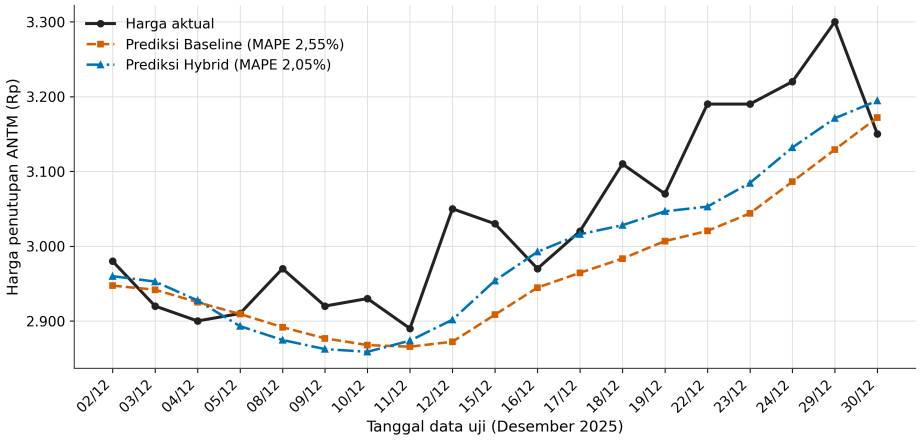
Kedua model LSTM dilatih pada 186 sekuens, divalidasi pada 21 sekuens, dan diuji pada 19 sekuens (2–30 Desember 2025). Hasil evaluasi disajikan pada Tabel 11, dengan kolom selisih (Hybrid − Baseline) dan perubahan relatif terhadap *baseline*.

Tabel 11: Hasil Evaluasi Model LSTM pada Data Uji ($n = 19$)

Metrik	**Baseline**	**Hybrid**	**Selisih (H − B)**	**Perubahan relatif**
Skala ternormalisasi				
MSE	0,001402	0,000710	−0,000692	−49,36%
RMSE	0,037450	0,026649	−0,010801	−28,84%
MAE	0,030552	0,023025	−0,007527	−24,64%
R^2	0,4950	0,7443	+0,2493	+50,36%
MAPE	4,09%	3,19%	−0,90 poin persentase	−22,00%
Skala harga asli (Rupiah)				
MAPE	2,55%	2,05%	−0,50 poin persentase	−19,61%

Integrasi S_{daily} menurunkan seluruh ukuran galat: MSE turun 49,36%, RMSE 28,84%, dan MAE 24,64%, sedangkan R^2 meningkat dari 0,4950 menjadi 0,7443. Artinya, model *hybrid* mampu menjelaskan sekitar 74,4% variasi harga pada periode uji, meningkat sekitar 25 poin persentase dibandingkan *baseline* yang hanya menjelaskan 49,5%. MAPE pada skala ternormalisasi turun

dari 4,09% menjadi 3,19%, dan pada skala harga asli turun dari 2,55% menjadi 2,05%; kedua nilai MAPE skala Rupiah tersebut tergolong sangat akurat menurut Lewis (1982).

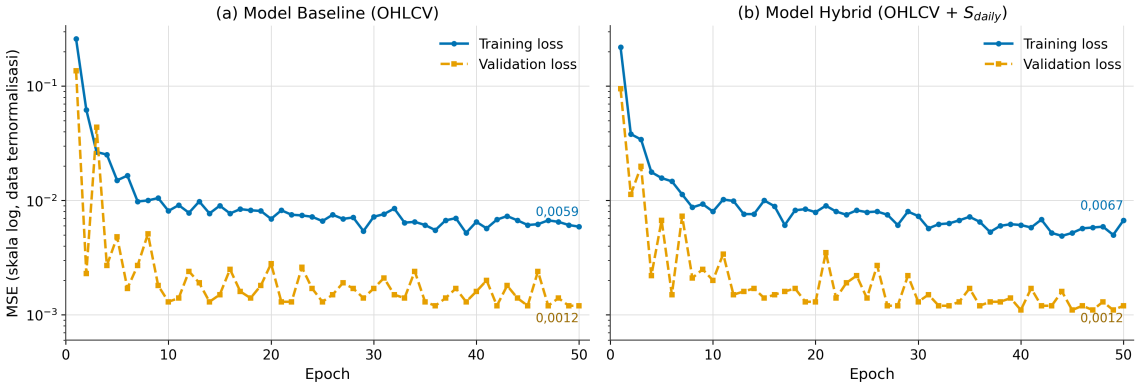


Gambar 6: Perbandingan Harga Aktual dan Prediksi Model *Baseline* serta *Hybrid* pada Data Uji (2–30 Desember 2025)

Gambar 6 menunjukkan bahwa kedua model cenderung tertinggal (*time-lag*) dari harga aktual, karakteristik umum model berbasis *window* historis. Namun, kurva *hybrid* (biru) lebih rapat dengan harga aktual dibandingkan *baseline* (merah), terutama pada 12–24 Desember 2025 saat harga naik: setelah titik terendah 11 Desember, model *hybrid* lebih cepat menyesuaikan prediksinya ke arah kenaikan, sedangkan *baseline* tetap landai. Meskipun demikian, kedua model belum mampu mengantisipasi lonjakan tajam (12 dan 29 Desember) maupun pembalikan arah pada 30 Desember.

3.6 Analisis Kurva *Loss* Pelatihan LSTM

Kurva *training* dan *validation loss* selama 50 *epoch* ditunjukkan pada Gambar 7a (*baseline*) dan Gambar 7b (*hybrid*) dengan sumbu vertikal logaritmik.



Gambar 7: Kurva *Training Loss* dan *Validation Loss* (MSE): (a) Model *Baseline* dan (b) Model *Hybrid*

Pada model *baseline* (Gambar 7a), *training loss* turun dari 0,2588 menjadi sekitar 0,010 pada

epoch ke-7 dan berakhir pada 0,0059, sedangkan *validation loss* turun dari 0,1363 menjadi 0,0023 pada *epoch* ke-2, sempat melonjak ke 0,0436 pada *epoch* ke-3, lalu stabil di kisaran 0,0012–0,0028 setelah *epoch* ke-10. Pada model *hybrid* (Gambar 7b), *training loss* turun dari 0,2192 menjadi sekitar 0,009 pada *epoch* ke-8 dan berakhir pada 0,0067, sedangkan *validation loss* turun dari 0,0948 ke 0,0022 pada *epoch* ke-4 dengan lonjakan awal terbesar yang lebih kecil (0,0199 pada *epoch* ke-3) dibandingkan *baseline*; setelah *epoch* ke-10 rentang fluktuasi kedua model sebanding dan keduanya berakhir pada 0,0012. Pada kedua gambar, *validation loss* berada di bawah *training loss* karena *dropout* hanya aktif saat pelatihan dan *training loss* merupakan rata-rata sepanjang *epoch*. Karena tidak ada tren kenaikan *validation loss* saat *training loss* terus turun, kedua model tidak menunjukkan indikasi *overfitting* yang signifikan; fluktuasi antar-*epoch* juga dipengaruhi kecilnya data validasi (21 sekuens).

3.7 Perbandingan dengan Penelitian Terdahulu

Tabel 12 membandingkan kinerja klasifikasi sentimen penelitian ini dengan Dhendra dan Utomo (2025), sedangkan Tabel 13 membandingkan kinerja prediksi harga ANTM menggunakan MAPE yang bersifat *scale-invariant*.

Tabel 12: Perbandingan Kinerja Model Analisis Sentimen Berbasis Transformer

Peneliti	Model	Karakteristik data	Jumlah data	Split	Akurasi
Dhendra dan Utomo (2025)	IndoBERT	Ulasan *e-government* (*Google Play*)	±11.000 ulasan	80:10:10	88,10%
Dhendra dan Utomo (2025)	XLM-RoBERTa	Ulasan *e-government* (*Google Play*)	±11.000 ulasan	80:10:10	88,00%
Dhendra dan Utomo (2025)	mBERT	Ulasan *e-government* (*Google Play*)	±11.000 ulasan	80:10:10	86,80%
Penelitian ini (Skenario 2)	**IndoBERT-base-p2**	**Judul berita finansial CNBC**	**9.738 berita**	**70:15:15**	**87,61%**
Penelitian ini (Skenario 1)	**IndoBERT-base-p2**	**Tweet saham ID-SMSA**	**3.204 tweet**	**80:10:10**	**80,06%**

Tabel 13: Perbandingan Kinerja Model Prediksi Harga Saham ANTM

Peneliti	Deskripsi data	Jumlah data	MAPE
Budiprasetyo dkk. (2022)	Harga penutupan harian ANTM (Yahoo Finance), Des 2016–Des 2021; LSTM 8 lapisan, *time step* 25, tanpa fitur sentimen	±5 tahun data harian; latih 80% / uji 20%	2,64%
Penelitian ini (LSTM *baseline*)	**OHLCV harian ANTM (Mirae Asset), 2 Jan–30 Des 2025; LSTM 2 lapisan, *window* 5**	**236 hari bursa; 207 sekuens latih, 19 sekuens uji**	**2,55%**
Penelitian ini (LSTM + sentimen)	**OHLCV + S_{daily} dari 27.181 tweet X, periode sama**	**236 hari bursa; 207 sekuens latih, 19 sekuens uji**	**2,05%**

Model Skenario 2 (87,61%) sedikit di bawah IndoBERT pada Dhendra dan Utomo (2025) (88,10%) dengan jumlah data sebanding. Perbedaan ini dapat dijelaskan oleh karakteristik domain: ulasan aplikasi cenderung eksplisit, sedangkan judul berita keuangan bergaya netral sehingga sentimennya lebih implisit. Akurasi Skenario 1 (80,06%) lebih rendah karena data latih lebih

kecil dan bahasa tweet informal, namun model inilah yang dipilih karena generalisasinya terbaik pada data riil. Pada prediksi harga, Budiprasetyo dkk. (2022) memperoleh MAPE 2,64% dengan LSTM delapan lapisan pada data ±5 tahun. Model *baseline* penelitian ini (2,55%) sebanding, dan model *hybrid* mencapai 2,05%. Perbandingan ini bersifat indikatif karena periode data, panjang periode uji, dan variabel input berbeda.

3.8 Keterbatasan Penelitian

Pertama, peningkatan kinerja model *hybrid* diukur pada satu kali pelatihan dan periode uji singkat (19 hari bursa), tanpa pengulangan *random seed*, validasi *walk-forward*, maupun uji signifikansi statistik (misalnya Diebold–Mariano). Kedua, penelitian ini belum menguji hubungan kausal maupun keterdahuluan waktu antara sentimen dan harga (misalnya uji kausalitas Granger, korelasi silang dengan beberapa *lag*, atau pembanding naif/*persistence*). Dengan demikian, hasil ini menunjukkan bahwa S_{daily} merupakan fitur tambahan yang bermanfaat bagi LSTM pada periode uji, tetapi belum cukup untuk menyimpulkan bahwa sentimen publik merupakan *leading indicator* pergerakan harga saham. Ketiga, sebagian tweet membahas harga emas secara umum, bukan saham ANTM, dan sentimen negatif konsumen terhadap harga emas tidak selalu bermakna negatif bagi emiten tambang emas. Keempat, validasi manual dilakukan oleh peneliti dan sampelnya juga digunakan untuk pemilihan model, sehingga akurasinya berpotensi optimistis.

4 Kesimpulan

Klasifikasi sentimen publik terhadap saham ANTM dari Platform X berhasil diimplementasikan menggunakan `IndoBERT-base-p2` melalui tiga skenario pelatihan. Meskipun Skenario 2 (berita CNBC) mencatat kinerja tertinggi pada data uji internalnya (akurasi 87,61%, F1-macro 0,8739), validasi manual pada 300 sampel data riil menunjukkan bahwa Skenario 1 (tweet) memiliki generalisasi terbaik (akurasi 83,33%, F1-macro 0,8292, F1-weighted 0,8375), sedangkan model berita turun paling besar akibat *domain shift*. Temuan ini menegaskan pentingnya memilih strategi *fine-tuning* berdasarkan data domain target. Model terpilih melabeli 27.181 tweet periode 2025 dengan dominasi sentimen negatif (48,5%), diikuti positif (32,1%) dan netral (19,5%).

Pada prediksi harga, LSTM *baseline* dengan lima fitur teknikal menghasilkan RMSE 0,0375, MAE 0,0306, dan R^2 0,4950 pada skala ternormalisasi serta MAPE 2,55% pada skala harga asli. Integrasi S_{daily} dengan arsitektur dan hiperparameter identik menurunkan RMSE menjadi 0,0266 (sekitar 29%) dan MAE menjadi 0,0230, meningkatkan R^2 menjadi 0,7443 (sekitar 25 poin persentase), serta menurunkan MAPE skala harga asli menjadi 2,05%. Hasil ini menunjukkan bahwa opini publik di media sosial yang diekstraksi dengan model Transformer membawa informasi tambahan yang dapat meningkatkan akurasi prediksi harga saham sektor pertambangan pada periode uji. Namun, karena periode uji singkat serta belum dilakukan uji kausalitas dan signifikansi, peran sentimen sebagai *leading indicator* masih perlu dibuktikan. Penelitian selanjutnya disarankan menggunakan validasi *walk-forward* dengan periode uji lebih panjang, pengulangan pelatihan dengan beberapa *random seed*, uji kausalitas Granger dan analisis *lag* sentimen, serta penyaringan tweet yang lebih spesifik terhadap emiten.

Ucapan Terima Kasih

Penulis mengucapkan terima kasih kepada Program Studi Sarjana Sains Data Universitas Negeri Surabaya serta seluruh pihak yang telah memberikan dukungan dan masukan selama pelaksanaan

penelitian ini.

References

Adeleye, R. A., Tubokirifuruar, T. S., Bello, B. G., Ndubuisi, N. L., Asuzu, O. F., dan Owolabi, O. R. (2024). Machine learning for stock market forecasting: A review of models and accuracy. *Finance & Accounting Research Journal*, 6(2):112–124.

Anggraini, N., Prasetya, D. A., dan Trimono (2025). Prediksi harga saham sektor energi menggunakan metode spatial temporal attention-based convolutional network berdasarkan data teks dan numerik. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 9(3).

Budiprasetyo, G., Hani'ah, M., dan Aflah, D. Z. (2022). Prediksi harga saham syariah menggunakan algoritma long short-term memory (LSTM). *Jurnal Nasional Teknologi dan Sistem Informasi (TEKNOSI)*, 8(3):164–172.

Devlin, J., Chang, M.-W., Lee, K., dan Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of NAACL-HLT 2019*, pages 4171–4186.

Dhendra dan Utomo, V. G. (2025). Benchmarking IndoBERT and transformer models for sentiment classification on Indonesian e-government service reviews. *Jurnal Transformatika*, 23(1):86–95.

Fahira, F. dan Prianto, C. (2023). Prediksi pola kedatangan turis mancanegara dan menganalisis ulasan tripadvisor dengan LSTM dan LDA. *Jurnal Tekno Insentif*, 17(2):69–83.

Faqih, A. dan Sugihartono, T. (2025). Perbandingan algoritma XGBoost dan LSTM dalam prediksi harga saham tesla menggunakan data tahun 2025. *Jurnal Pendidikan dan Teknologi Indonesia*, 5(6):1563–1573.

Hartanto, J., Liundi, T., Sutoyo, R., dan Andangsari, E. W. (2025). ID-SMSA: Indonesian stock market dataset for sentiment analysis. *Data in Brief*, 60:111571.

Hyndman, R. J. dan Athanasopoulos, G. (2021). *Forecasting: Principles and Practice*. OTexts, Melbourne, 3 edition.

Jiang, T. dan Zeng, Q. (2025). Financial sentiment analysis using FinBERT with application in predicting stock movement. *arXiv preprint arXiv:2306.02136*.

Jiwandono, T. A., Achmad, M. S. H., dan Suhirman (2025). Fine-tuned IndoBERT for stock market sentiment analysis: Evidence from CNBC Indonesia news. *Indonesian Journal of Electrical Engineering and Computer Science*, 42(3):774–785.

Koto, F., Rahimi, A., Lau, J. H., dan Baldwin, T. (2020). IndoLEM and IndoBERT: A benchmark dataset and pre-trained language model for Indonesian NLP. In *Proceedings of the 28th International Conference on Computational Linguistics (COLING 2020)*, pages 757–770.

Lewis, C. D. (1982). *Industrial and Business Forecasting Methods: A Practical Guide to Exponential Smoothing and Curve Fitting*. Butterworth Scientific, London.

Ruan, L. dan Jiang, H. (2025). Stock price prediction using FinBERT-enhanced sentiment with SHAP explainability and differential privacy. *Mathematics*, 13(17):2747.

Srinivas, R. S., Das, A., Gedela, R. B., Nath, G., Sabu, R., dan Datla, V. (2022). Effects of daily news sentiment on stock price forecasting. Laporan teknis.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., dan Polosukhin, I. (2017). Attention is all you need. In *Advances in Neural Information Processing Systems 30 (NeurIPS 2017)*, pages 5998–6008.

Wilie, B., Vincentio, K., Winata, G. I., Cahyawijaya, S., Li, X., Lim, Z. Y., Soleman, S., Mahendra, R., Fung, P., Bahar, S., dan Purwarianti, A. (2020). IndoNLU: Benchmark and resources for evaluating Indonesian natural language understanding. In *Proceedings of the 1st Conference of the Asia-Pacific Chapter of the ACL and the 10th International Joint Conference on NLP (AACL-IJCNLP 2020)*, pages 843–857.