

Penerapan Analisis Sentimen Berita Keuangan dari IDX Channel dan Long Short-Term Memory (LSTM) dalam Prediksi Perubahan Harga Saham pada Perusahaan XYZ

Fathoni¹⁾, Cantika Aulia²⁾, Muhammad Fakhri Nadrota Acta³⁾, Eka Saputra⁴⁾, M. Fadhil Rahman⁵⁾,
Ali Ibrahim⁶⁾

¹⁻⁶ Fakultas Ilmu Komputer, Sistem Informasi, Universitas Sriwijaya, Indralaya, Indonesia
Jl. Raya Palembang - Prabumulih Km. 32, Indralaya, Ogan Ilir, Sumatera Selatan, Indonesia
email: ¹fathoni@unsri.ac.id, ²09031282227124@student.unsri.ac.id, ³0903182227048@student.unsri.ac.id,
⁴09031282227051@student.unsri.ac.id, ⁵0903182227070@student.unsri.ac.id, ⁶aliibrahim@unsri.ac.id,

Abstract

Stock investment is becoming increasingly popular; however, high price fluctuations make prediction a challenge. Stock prices are influenced by various factors, such as financial news that reflects market sentiment. Nonetheless, many prediction models rely solely on historical data without considering media opinions. This study aims to predict stock price movements of company XYZ by combining a Long Short-Term Memory (LSTM) model with sentiment analysis of financial news from IDX Channel. Historical stock price data was obtained from Yahoo Finance, while sentiment data was derived through opinion classification of news articles using the IndoBERT model. The data was then processed using Min-Max Scaling normalization and structured in a time series format through a window sliding technique with a time step of 30. The results show that the LSTM model can predict stock prices with an RMSE of 35.92 and a weak positive correlation of 0.27 between sentiment and stock prices. The model's one-day-ahead forecast produced a predicted price of Rp 4,097.00. Residual visualization indicates that the prediction errors are distributed closely around zero, suggesting that the model generalizes well.

Keywords: Stock Price Prediction, Sentiment Analysis, LSTM, NLP, Financial News, IDX Channel

Abstrak

Investasi saham semakin diminati, namun fluktuasi harga yang tinggi membuat prediksi menjadi tantangan. Harga saham dipengaruhi oleh banyak hal, seperti berita keuangan yang menunjukkan sentimen pasar. Namun, banyak model prediksi bergantung pada data sebelumnya tanpa memperhitungkan pendapat media. Penelitian ini bertujuan untuk memprediksi perubahan harga saham perusahaan XYZ dengan menggabungkan model Long Short-Term Memory (LSTM) dan analisis sentimen terhadap berita keuangan dari IDX Channel. Data historis harga saham diambil dari Yahoo Finance, sedangkan data sentimen diperoleh melalui klasifikasi opini berita berbasis model IndoBERT. Data kemudian diproses menggunakan pendekatan normalisasi Min-Max Scaling dan dibentuk dalam format time series menggunakan teknik window sliding dengan time step sebesar 30. Hasil penelitian menunjukkan bahwa model LSTM mampu memprediksi harga saham dengan nilai RMSE sebesar 35,92 dan korelasi positif lemah antara sentimen dan harga saham sebesar 0,27. Prediksi harga satu hari ke depan yang dihasilkan model menunjukkan nilai Rp 4.097,00. Visualisasi residual menunjukkan sebaran kesalahan yang stabil di sekitar nol, menunjukkan generalisasi model yang cukup baik.

Kata kunci: Prediksi Harga Saham, Analisis Sentimen, LSTM, NLP, Berita Keuangan, IDX Channel

1. PENDAHULUAN

Investasi saham sebagai strategi untuk meningkatkan nilai aset kini semakin diminati. Indeks saham tidak hanya mencerminkan kinerja perusahaan-perusahaan besar, tetapi juga menjadi indikator kesehatan ekonomi suatu negara Tando & Irawan (2023). Menurut data Bursa Efek Indonesia, jumlah investor pada tahun 2024 mencapai 14,87 juta, meningkat 55% dari tahun sebelumnya Maghiszha (2024). Jumlah investor yang meningkat menunjukkan bahwa banyak orang kini menyadari pentingnya investasi sebagai jalan untuk mengoptimalkan peluang keuntungan di pasar modal dan

mencapai kestabilan finansial. Namun, investor yang kurang berpengalaman seringkali kesulitan memahami faktor-faktor yang mempengaruhi dinamika harga saham.

Salah satu tantangan utama dalam investasi saham adalah prediksi harga yang akurat, karena harga saham bisa turun sewaktu-waktu akibat fluktuasi pasar, perubahan sentimen investor, atau faktor eksternal seperti kondisi ekonomi global dan berita finansial yang mempengaruhi sentimen pasar. Volatilitas harga saham menjadi hambatan tersendiri dalam menentukan kerangka waktu yang optimal untuk melakukan analisis Kartika & Karmilasari (2022). Selain itu, kejadian seperti Auto Rejection Atas (ARA) dan Auto Rejection Bawah (ARB) sering terjadi dalam waktu singkat akibat perubahan besar dalam persepsi pasar Wijayanti (2025). Sebagian besar model prediksi yang ada masih mengandalkan data historis semata tanpa mempertimbangkan sentimen pasar yang dinamis, padahal ekspektasi investor sangat dipengaruhi oleh perkembangan berita keuangan dan opini publik.

Dalam ranah kecerdasan buatan, machine learning telah digunakan secara luas untuk memprediksi harga saham dengan memanfaatkan pola dan tren dalam data pasar Julian & Pribadi (2021). Algoritma seperti Long Short-Term Memory (LSTM) dapat menangkap tren data masa lalu dengan akurasi 98,46% pada saham PT Telkom Nazhiroh dkk. (2024). Selain itu, model prediksi sentimen berita saham BBRI menggunakan IndoBERT mencapai akurasi lebih dari 90% Putra dkk. (2024). Penelitian yang dilakukan oleh Fahrezi et al. (2024) menunjukkan bahwa model LSTM memiliki kemampuan superior dalam memprediksi harga penutupan saham BBKA, dengan nilai MAPE hanya 1,07%, jauh lebih rendah dibandingkan Support Vector Regression (SVR) sebesar 17,11%. Pendekatan hybrid model yang menggabungkan analisis sentimen dengan model LSTM semakin berkembang. Seperti pada hasil uji coba prediksi harga saham syariah menggunakan model LSTM yang menunjukkan bahwa dengan nilai MAPE di bawah 10%, nilai-nilai seperti PT Aneka Tambang Tbk 2,64 dan Kalbe Farma Tbk 1,51 menunjukkan prediksi yang baik dan akurat Budiprasetyo dkk. (2023). Analisis sentimen berbasis Natural Language Processing (NLP) dapat digunakan untuk mengambil opini dari berita finansial dan media sosial guna mengidentifikasi sentimen positif, negatif, atau netral Drus & Khalid (2019). Dengan menyertakan sentimen berita ke dalam model LSTM, prediksi harga saham diharapkan menjadi lebih akurat sehingga investor dapat memahami pergerakan pasar dengan lebih baik.

Penelitian terkait prediksi harga saham terus mengalami perkembangan, seiring meningkatnya pemanfaatan artificial intelligence dan teknik deep learning. Sejumlah metode seperti LSTM, Bidirectional LSTM (BiLSTM), Convolutional Neural Network - LSTM (CNN-LSTM), Transformer, serta pendekatan berbasis NLP telah diaplikasikan untuk meningkatkan akurasi prediksi. Amiri dkk. (2025) menunjukkan bahwa metode regresi linier dan Gated Recurrent Unit (GRU) memiliki keterbatasan dalam mengidentifikasi pola temporal serta relasi antar saham, dibandingkan kombinasi Graph Convolutional Network (GCN) dan LSTM dengan mekanisme attention. Kwanda dkk. (2024) mengungkapkan bahwa LSTM memiliki tingkat kesalahan yang lebih rendah dibandingkan BiLSTM dalam memprediksi harga saham beberapa emiten. Sementara itu, Zhang dkk. (2023) membuktikan bahwa model CNN-BiLSTM-Attention mampu menurunkan tingkat kesalahan hingga 46,6% dibandingkan model LSTM konvensional. Chen dkk. (2023) mengembangkan model GRU dengan dataset rekonstruksi yang berhasil meningkatkan akurasi dari 87,2% menjadi 93,4%. William & Rarasati (2024) menemukan bahwa penyesuaian hyperparameter pada model LSTM dapat meningkatkan performa prediksi saham IDX30 secara signifikan. Pendekatan hibrida juga semakin diperkaya, seperti pada model GRUvader yang menggabungkan VADER dengan GRU dan mencapai akurasi 97,6% untuk saham Tesla Mamillapalli dkk. (2024), serta algoritma S_I_LSTM yang memadukan CNN untuk sentimen dengan data teknikal Wu dkk. (2022). Studi lain menunjukkan efektivitas kombinasi Temporal Convolutional Network (TCN) dan Generative Adversarial Network (GAN) dalam menurunkan MAPE dibandingkan LSTM dan ANN Teck & Thenata (2025), serta model CNN-LSTM Refined yang meningkatkan akurasi prediksi hingga 97% (Ding 2023). Wei dkk. (2025) juga mengusulkan pengoptimalan transformer melalui algoritma Multi-Objective Escape Bird Search (MOEBS), yang memberikan hasil lebih unggul dibanding metode optimasi lain. Meskipun banyak pendekatan telah dikembangkan, kesenjangan masih terdapat pada integrasi antara data teknikal, sentimen pasar, dan faktor makroekonomi dalam satu model prediksi komprehensif.

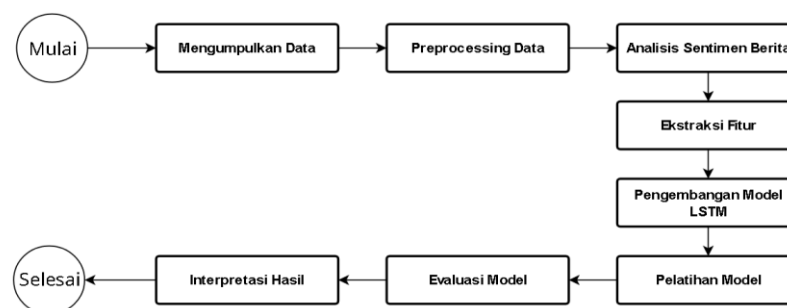
Penelitian ini bertujuan untuk mengembangkan model prediksi harga saham dengan mengintegrasikan analisis sentimen dari berita keuangan dengan model Long Short-Term Memory (LSTM), menggunakan perusahaan XYZ sebagai studi kasus. Integrasi model LSTM dengan analisis sentimen tidak hanya meningkatkan akurasi prediksi harga saham, tetapi juga memberikan

pemahaman yang lebih holistik terhadap dinamika pasar, karena model mampu menangkap pengaruh informasi eksternal secara simultan dengan pola historis (Sari dkk., 2024). Sumber data yang digunakan adalah IDX Channel, yang dipilih karena kredibilitasnya dalam menyajikan informasi keuangan terkini dan relevan. Analisis sentimen terhadap berita diharapkan dapat memberikan gambaran mengenai arah sentimen pasar Aslim (2024), serta memperkuat model prediksi harga saham berbasis LSTM. Penelitian ini akan menguji sejauh mana sentimen berita keuangan berpengaruh terhadap fluktuasi harga saham perusahaan XYZ. Hasil dari penelitian ini diharapkan dapat memberikan kontribusi bagi akademisi dalam pengembangan model prediktif yang lebih akurat, serta bagi praktisi keuangan sebagai alat bantu dalam pengambilan keputusan investasi yang lebih informatif dan berbasis data.

2. METODOLOGI PENELITIAN

2.1. Tahapan Penelitian

Alur penelitian dimulai dari pengumpulan data yang mencakup berita keuangan dan harga saham, kemudian data tersebut dibersihkan dan diolah dalam tahap preprocessing. Selanjutnya, dilakukan analisis sentimen berita untuk menentukan skor sentimen yang akan digabungkan dengan data saham pada tahap ekstraksi fitur. Setelah itu, fitur-fitur tersebut digunakan untuk mengembangkan dan melatih model LSTM, yang kemudian dievaluasi menggunakan metrik tertentu guna mengukur tingkat akurasi prediksi. Hasil akhir penelitian diinterpretasikan untuk melihat pengaruh sentimen berita terhadap pergerakan harga saham, sebelum akhirnya ditarik kesimpulan. Dalam alur penelitian yang akan dilakukan dapat dilihat pada Gambar 1 berikut.



Gambar 1. Tahapan Penelitian

2.2. Pengumpulan Data

Pada tahap ini dilakukan pengumpulan data berita keuangan dan harga saham sebagai bahan baku penelitian. Proses pengumpulan data perlu memperhatikan rentang waktu yang konsisten agar hasil prediksi lebih akurat Julian & Pribadi (2021). Data yang akan digunakan dalam proses penelitian berupa data dari berita keuangan yang relevan sebagai data untuk analisis sentimen dari sumber berita IDX Channel. Dalam data tersebut terdapat beberapa atribut yang ditunjukkan pada pada Tabel 1 berikut.

Tabel 1. Atribut pada data

No	Nama Kolom	Keterangan
1	Tanggal dan tahun	Tanggal berita diterbitkan
2	Judul berita	Judul utama berita yang bisa digunakan dalam analisis NLP
3	Sentimen	Label sentimen (Positif, Negatif, Netral), berdasarkan manual labeling atau model NLP
4	Harga saham sebelum	Harga saham BBRI sebelum berita diterbitkan
5	Harga saham sesudah	Harga saham BBRI sesudah berita diterbitkan
6	Perubahan harga (%)	Persentase kenaikan/penurunan harga saham sebagai respons terhadap berita.

Kemudian, pada model Long Short-Term Memory pengumpulan data dilakukan menggunakan data histori dari perusahaan XYZ dalam rentang waktu 2024-2025 untuk menangkap pola pergerakan harga saham berdasarkan data historis tersebut dari sumber IDX Channel dan Yahoo Finance.

Tabel 2. Data histori harga saham perusahaan XYZ

Date	Open	High	Low	Close
2025/04/07	4.200	4.200	3.040	4.050
2025/03/27	4.000	4.050	3.930	4.050
2025/03/26	3.870	4.020	3.860	4.000
2025/03/25	3.660	3.840	3.660	3.800
2025/03/24	3.650	3.724	3.520	3.610
		...		
2024/04/01	6.000	6.025	5.825	5.925
2024/03/28	6.100	6.200	6.025	6.050
2024/03/27	6.300	6.325	6.225	6.250

2.3. Processing Data

Processing adalah proses perbaikan data yang mana tahap pengolahan data ini bertujuan untuk meningkatkan kualitas data dengan menghapus data yang mengandung noise atau nilai yang hilang. Proses data preprocessing yang dilakukan dengan baik dan menyeluruh akan memberikan kontribusi signifikan dalam meningkatkan tingkat akurasi dari model yang digunakan untuk analisis sentimen sehingga hasil prediksi yang dihasilkan menjadi lebih andal dan relevan. Proses ini memastikan data bersih, terstruktur, dan siap diolah oleh algoritma.

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}}$$

X = Nilai asli

X_{min} , X_{max} = Nilai minimum dan maksimum dalam data

X' = Data yang telah dinormalisasi

2.4. Sentiment Analysis

Analisis sentimen bertujuan untuk mengubah teks berita menjadi skor sentimen (positif, negatif, atau netral). Tahapan ini menjelaskan pentingnya sentiment analysis dalam konteks prediksi pasar. Data sentimen yang telah di preproses ditransformasi menjadi bentuk vektor atau *bag of words* untuk merepresentasikan frekuensi kata dalam dokumen. Baris matriks mewakili kalimat, sementara kolom mewakili kata, dengan jumlah kemunculan setiap kata dihitung. Data ini kemudian diolah menggunakan *supervised machine learning* untuk klasifikasi berdasarkan polaritas yang telah ditetapkan pada setiap kalimat.

2.5. Model Long Short-Term Memory (LSTM)

Long Short-Term Memory (LSTM), yang dikembangkan oleh Hochreiter dan Schmidhuber pada tahun 1997, dirancang sebagai model berulang yang mempertimbangkan informasi dari neuron saat ini sekaligus mempertahankan jejak memori dari masa lalu. LSTM mampu menyimpan informasi dari pola data dan memiliki kemampuan untuk menentukan data yang relevan untuk digunakan dalam proses berikutnya, sekaligus membuang data yang tidak diperlukan agar tidak diteruskan ke tahap jaringan berikutnya. Dalam penelitian ini, model LSTM digunakan untuk memprediksi harga saham perusahaan XYZ, baik dengan menggunakan data historis saham saja maupun dikombinasikan dengan sentimen berita keuangan.

Arsitektur yang digunakan terdiri dari dua lapisan LSTM masing-masing dengan 64 unit memori, disertai dropout sebesar 0,2 untuk mencegah *overfitting*. Data masukan disusun menggunakan teknik *sliding window* dengan ukuran 30 hari. Proses pelatihan dilakukan selama 30 epoch dengan ukuran batch 32, menggunakan optimizer Adam (learning rate 0,001) dan fungsi kerugian Mean Squared Error (MSE). Early stopping diterapkan dengan *patience* 10 epoch sebagai

outpu. Tuning parameter dilakukan secara terbatas terhadap variasi jumlah neuron, ukuran window, dan dropout, dengan pemilihan konfigurasi terbaik berdasarkan MAPE dan RMSE pada validation set. Input model berupa kombinasi data historis harga saham (Close) dan nilai sentimen berita. Pemilihan arsitektur ini merujuk pada penelitian Gunawan yang menunjukkan bahwa LSTM multilevel lebih stabil dalam memproses deret waktu.

2.6. Pelatihan Model

Pelatihan model dilakukan dengan menggunakan *Long Short-Term Memory* (LSTM) untuk menangkap pola temporal dalam data harga saham serta *Sentiment Analysis* untuk mengekstrak informasi dari berita keuangan. Model dilatih dengan data historis harga saham dan skor sentimen, menggunakan *Adam optimizer* serta teknik regulasi seperti *dropout* untuk mencegah *overfitting*. Evaluasi awal dilakukan dengan validation set, sementara performa akhir diuji menggunakan data *testing* dengan metrik MSE dan RMSE untuk menilai akurasi prediksi.

Data *testing* digunakan untuk mengukur performa model setelah proses pelatihan selesai. Data ini dipisahkan dari data *training* untuk memastikan model dapat menggeneralisasi pola yang telah dipelajari pada data baru yang belum pernah dilihat sebelumnya. Data ini mengalami proses transformasi, termasuk normalisasi dan ekstraksi fitur, sebelum digunakan dalam pelatihan. Model belajar dari data ini dengan menyesuaikan bobot atau parameter untuk meminimalkan kesalahan prediksi melalui optimasi tertentu, seperti *gradient descent* atau metode lainnya Mamillapalli dkk. (2024).

2.7. Evaluasi Model

Evaluasi model bertujuan untuk mengukur kinerja model dengan menggunakan metrik tertentu, seperti *accuracy*, *precision*, *recall*, dan *F1-score* untuk tugas klasifikasi, serta *mean squared error* (MSE) atau *R-squared* untuk regresi. Data testing digunakan dalam tahap ini untuk melihat sejauh mana model dapat memprediksi dengan baik di luar data latihannya. Teknik validasi silang juga dapat diterapkan untuk memastikan hasil evaluasi yang lebih akurat dan menghindari *overfitting* Aslim (2024).

3. HASIL DAN PEMBAHASAN

3.1 Hasil Analisis Data

Data utama yang digunakan dalam penelitian ini adalah data tentang berita keuangan dan harga saham perusahaan XYZ. Berita keuangan diperoleh dari Yahoo Finance, yang menyajikan berita ekonomi dan keuangan internasional, serta dari situs IDX Channel, salah satu portal berita resmi yang memberikan informasi tentang pasar modal Indonesia. Studi ini mengumpulkan 100–200 artikel berita dari awal tahun 2024 hingga saat ini. Pendekatan deep learning dari pembelajaran mesin memungkinkan penyelesaian berbagai permasalahan prediksi saham secara lebih efisien Afrianto dkk. (2022). Selain berita, data harga saham sebelumnya perusahaan XYZ dikumpulkan dari IDX Channel dan Yahoo Finance dari tahun 2024 hingga waktu penelitian dilakukan. Sinkronisasi antara waktu pengambilan data berita dan data harga saham dilakukan untuk memastikan bahwa analisis sentimen dapat dihubungkan langsung dengan pergerakan harga saham yang relevan pada periode yang bersangkutan. Berikut merupakan Tabel 3 yang menyajikan data sentimen berita perusahaan XYZ sebagai bahan analisis dalam pengaruh sentimen berita terhadap pergerakan harga saham perusahaan XYZ.

Tabel 3. Data sentimen berita perusahaan XYZ

Tanggal	Judul	Sentimen	Harga Sebelum	Harga Sesudah	Perubahan Harga (%)
25/03/2025	Dapat Dividen Rp13 Miliar dari BBRI, Lo Kheng Hong: Buat Beli Saham Lagi	Posifit	Rp3.610	Rp3.800	5,26%

24/03/2025	Lo Kheng Hong Hadiri RUPST, Akui Genggam 64,63 Juta Saham BBRI	Negatif	Rp3.700	Rp3.610	-2,43%
23/03/2025	Intip Saham BUMN di Bawah Danantara Jelang Pengumuman Struktur Pengurus Besok	Positif	Rp3.660	Rp3.700	1,09%
...	...	...	...	...	...
06/05/2024	BRI Signals Strength: Launches Buyback Amidst Stock Price Corrections	Positif	Rp4.770	Rp4.790	0,42%
29/04/2024	BRI Reports IDR 15.98 Trillion Profit, Eyes Global Trends, Prioritizes Domestic with MSMEs Empowerment	Negatif	Rp4.830	Rp4.770	-1,24%
02/04/2024	Bank Rakyat Indonesia's BRImo Surpasses 30 million Users, Moving Financial Inclusion Forward	Negatif	Rp5.925	Rp5.492	-7,31%

Tabel 3 menunjukkan bahwa sentimen berita berpengaruh signifikan terhadap pergerakan harga saham perusahaan XYZ. Berita positif umumnya diikuti kenaikan harga saham, sedangkan berita negatif cenderung menyebabkan penurunan. Pola ini mendukung penggunaan data sentimen sebagai variabel penting dalam model prediksi seperti LSTM untuk meningkatkan akurasi prediksi harga saham.

3.2 Pre-Pemrosesan Data

Setelah data teks dibersihkan dan diproses, langkah selanjutnya adalah memberikan label sentimen pada setiap berita.. Seperti pada Tabel 4 berikut.

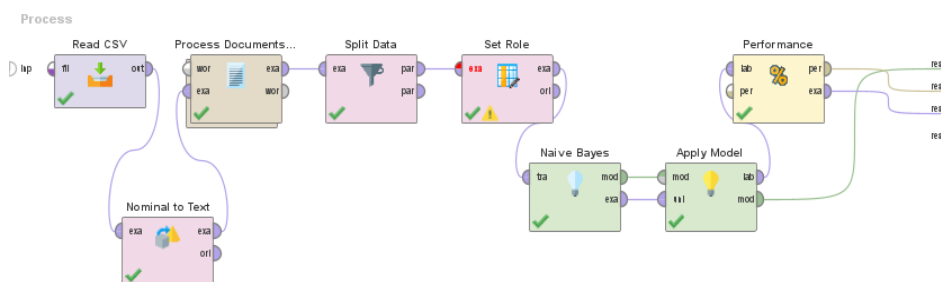
Tabel 4. Data histori harga saham perusahaan XYZ

Skor Sentimen	Keterangan
1	Positif
0	Netral
-1	Negatif

Pelabelan ini dilakukan secara manual berdasarkan isi dan konteks berita. Sebagai variabel independen, dalam proses sentimen analisis diklasifikasikan dalam tiga kategori: positif, netral, dan negatif Christophorus Bintang Saputra & Koesrindartoto (2024).

3.3 Hasil Analisis Sentimen

Proses analisis sentimen dilakukan dengan menggunakan algoritma Naive Bayes melalui serangkaian tahapan, mulai dari pembacaan data, pemrosesan dokumen, pembagian data, hingga pelatihan dan pengujian model, sebagaimana ditunjukkan pada Gambar 2.



Gambar 2. Implementasi algoritma Naive Bayes pada rapidMiner

Proses dimulai dengan pembacaan data dari file CSV, dilanjutkan dengan konversi atribut nominal menjadi teks, dan pemrosesan dokumen teks (misalnya tokenisasi, stemming, dan stopwords removal). Data kemudian dibagi menjadi data pelatihan dan pengujian, serta ditetapkan perannya (label dan atribut).

Model Naive Bayes kemudian dilatih menggunakan data pelatihan dan diterapkan pada data pengujian. Akurasi dan performa model dievaluasi menggunakan operator "Performance". Prosedur ini bertujuan untuk mengklasifikasikan sentimen dokumen secara otomatis ke dalam kategori tertentu (misalnya positif atau negatif), yang kemudian dapat digunakan dalam analisis lanjutan seperti prediksi harga saham berbasis sentimen pasar.

Hasil evaluasi model Naive Bayes menunjukkan bahwa algoritma ini mampu mengklasifikasikan data sentimen dengan tingkat akurasi yang sangat tinggi, dengan nilai precision dan recall mencapai 100% untuk setiap kelas sentimen, seperti ditampilkan pada Gambar 3.

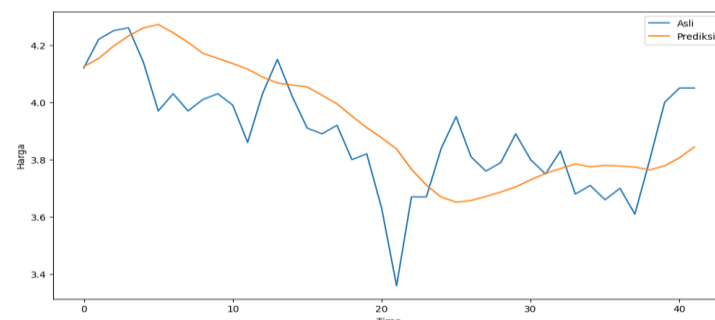
	true Positif	true Negatif	true Netral	class precision
pred. Positif	57	0	0	100.00%
pred. Negatif	0	27	0	100.00%
pred. Netral	0	0	8	100.00%
class recall	100.00%	100.00%	100.00%	

Gambar 3. Hasil pengujian model

3.4 Pelatihan dan Evaluasi Model LSTM

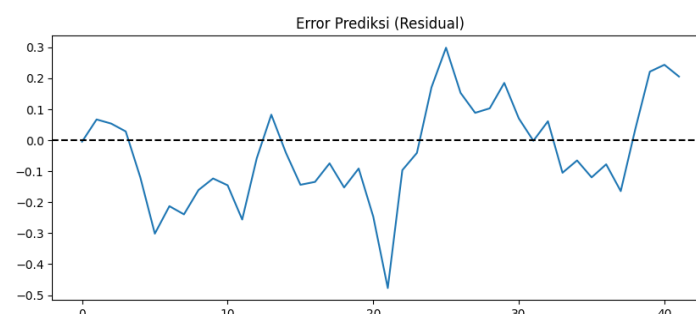
Model LSTM yang dirancang menggunakan data harga saham perusahaan XYZ dan sentimen berita keuangan IDX Channel telah dilatih selama 30 epoch dengan arsitektur dua lapisan LSTM yang masing-masing memiliki 64 unit memori. Model ini juga dilengkapi dengan teknik regularisasi *dropout* sebesar 20% untuk menghindari overfitting.

Evaluasi performa model dilakukan dengan menghitung nilai Root Mean Squared Error (RMSE) terhadap data pengujian. Hasil perhitungan menunjukkan bahwa model menghasilkan RMSE sebesar 35,92 yang menandakan tingkat kesalahan prediksi yang relatif rendah jika dibandingkan dengan rentang harga saham yang berada di antara Rp 3.000 – Rp 5.000. Visualisasi hasil prediksi dibandingkan dengan harga aktual ditampilkan pada Gambar 4.



Gambar 4. Perbandingan harga aktual dan hasil prediksi

Selain itu, dilakukan analisis residual untuk melihat distribusi kesalahan prediksi. Grafik residual menunjukkan bahwa mayoritas hasil prediksi berada dekat dengan nilai aktual. Hal tersebut menunjukkan kestabilan model dalam generalisasi.



Gambar 5. Residual error dari model prediktif

Grafik residual yang ditampilkan menggambarkan pola penyebaran kesalahan prediksi model terhadap data aktual. Secara keseluruhan, nilai residual bergerak di sekitar nilai nol. Hal ini menandakan bahwa model tidak mengalami bias prediksi sistematis (tidak secara konsisten *overestimate* atau *underestimate*).

Namun, terdapat sejumlah titik di mana residual menunjukkan lonjakan yang signifikan. Hal ini dapat dikaitkan dengan kejadian-kejadian pasar yang sifatnya outlier atau ketidakmampuan model dalam menangkap volatilitas ekstrem dalam jangka pendek. Meski demikian, distribusi residual yang simetris dan tidak terakumulasi pada satu sisi memperkuat kesimpulan bahwa model memiliki kemampuan generalisasi yang baik dalam situasi pasar yang relatif stabil.

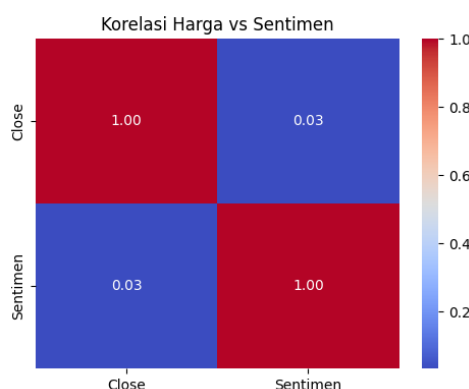
3.5 Prediksi Perubahan Harga Saham

Setelah melalui proses pelatihan dan evaluasi, model LSTM yang dikombinasikan dengan analisis sentimen digunakan untuk memprediksi harga saham satu hari ke depan. Model ini menerima input berupa data historis harga dan sentimen selama 30 hari terakhir dalam bentuk urutan numerik sebagai masukan ke dalam jaringan memori LSTM. Hasil prediksi menunjukkan bahwa harga saham perusahaan XYZ pada hari berikutnya diperkirakan mencapai Rp 4.097,00. Estimasi ini mempertimbangkan pola pergerakan harga sebelumnya serta kecenderungan sentimen dari berita IDX Channel pada periode yang sama. Dibandingkan harga aktual sebelumnya, prediksi menunjukkan kenaikan moderat, menandakan bahwa model cukup responsif terhadap pengaruh berita pasar, meskipun tidak terlalu sensitif terhadap fluktuasi acak. Prediksi ini tidak hanya bermanfaat sebagai referensi bagi investor dalam pengambilan keputusan, tetapi juga berpotensi menjadi dasar sistem rekomendasi otomatis pada aplikasi trading atau simulasi pasar saham. Keunggulan pendekatan ini terletak pada kemampuannya mengintegrasikan dimensi temporal dan eksternal secara bersamaan, yang sulit dicapai oleh metode statistik konvensional.

Namun, karena model hanya memproyeksikan satu langkah ke depan, prediksi jangka panjang belum menjadi fokus dalam studi ini. Hal ini membuka peluang pengembangan lebih lanjut melalui penerapan model multistep forecast atau arsitektur sequence-to-sequence (Seq2Seq) untuk prediksi multi-hari. Studi oleh (Sudhakar & Naganjaneyulu, 2025) juga menekankan pentingnya perluasan cakupan temporal agar model mampu menangkap dinamika pasar jangka menengah dan panjang secara lebih akurat. Dengan kombinasi LSTM dan analisis sentimen, model mampu menghasilkan prediksi numerik sekaligus merepresentasikan konteks pasar secara lebih holistik.

3.6 Pembahasan

Hasil analisis korelasi antara sentimen berita dan harga saham menunjukkan nilai sebesar +0,27, yang mengindikasikan adanya hubungan positif namun lemah. Visualisasi hasil prediksi memperlihatkan bahwa model LSTM mampu mengikuti tren pergerakan harga saham secara konsisten, khususnya pada periode dengan fluktuasi harga yang tajam (Fahrezi dkk., 2024). Meskipun kontribusi sentimen tidak dominan secara statistik, integrasi fitur sentimen ke dalam input model LSTM terbukti dapat memperkaya konteks temporal dan meningkatkan kualitas prediksi. Fenomena ini tergambarkan secara lebih jelas pada Gambar 6.

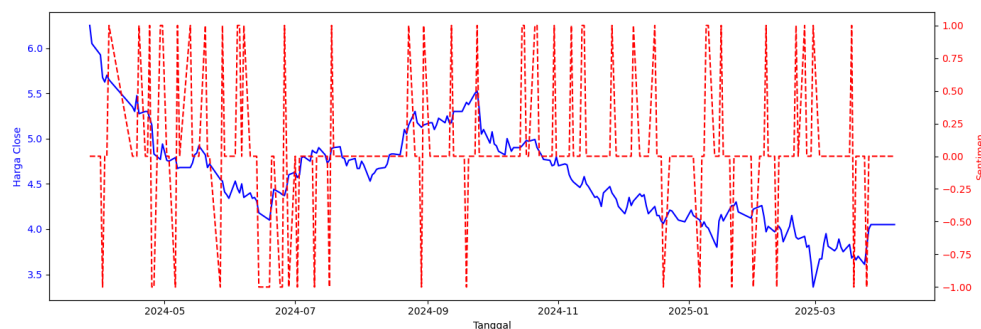


Gambar 6. Heatmap korelasi antara nilai sentimen dan harga saham

Berdasarkan perhitungan korelasi Pearson antara nilai sentimen berita keuangan dan harga saham, diperoleh nilai sebesar 0,27. Nilai ini termasuk dalam kategori korelasi positif lemah. Hal ini

memperlihatkan bahwa meskipun arah perubahan sentimen cenderung sejalan dengan arah pergerakan harga saham, kekuatan hubungannya tidak terlalu kuat. Beberapa faktor dapat menjelaskan rendahnya nilai korelasi tersebut. Pertama, reaksi harga saham terhadap berita bersifat tidak langsung dan tidak selalu terjadi pada hari yang sama. Kedua, sentimen yang diperoleh dari berita cenderung bersifat subjektif dan tergantung pada metode klasifikasi yang digunakan. Ketiga, pasar saham dipengaruhi oleh banyak faktor lain seperti kondisi makroekonomi, pergerakan indeks global, serta volume transaksi.

Implikasi dari temuan ini adalah bahwa meskipun analisis sentimen memberikan kontribusi informasi tambahan, model prediksi tidak dapat sepenuhnya bergantung pada fitur sentimen. Oleh karena itu, integrasi data fundamental atau teknikal tambahan dapat menjadi strategi lanjutan untuk meningkatkan performa prediktif.



Gambar 7. Tren historis sentimen dan harga saham perusahaan XYZ

Gambar 7 menyajikan tren historis antara nilai sentimen berita dan pergerakan harga saham perusahaan XYZ, yang secara visual menunjukkan keterkaitan yang cukup signifikan. Terlihat bahwa perubahan ekstrem dalam nilai sentimen, baik positif maupun negatif, umumnya diikuti oleh pergerakan harga saham ke arah yang sejalan. Pola ini memperkuat indikasi bahwa berita sebagai sumber utama informasi sentimen pasar memiliki pengaruh nyata terhadap dinamika harga saham. Nilai korelasi Pearson antara sentimen dan harga tercatat sebesar 0,27, yang termasuk dalam kategori korelasi positif lemah, grafik ini mengilustrasikan bahwa tren sentimen dapat berfungsi sebagai indikator awal terhadap arah pergerakan harga saham. Keberadaan *time lag* antara perubahan sentimen dan respons harga mengindikasikan bahwa pelaku pasar memerlukan waktu dalam mencerna dan merespons informasi yang terkandung dalam berita, sehingga dampak sentimen tidak selalu terjadi secara langsung namun tetap signifikan dalam jangka pendek.

Secara keseluruhan, pola historis ini memperkuat argumen bahwa analisis sentimen berita berperan sebagai indikator relevan dalam memproyeksikan arah pasar, terutama apabila dikombinasikan dengan pendekatan berbasis deret waktu. Kemampuan model untuk mengidentifikasi pola fluktuasi sentimen memberikan peluang untuk menghasilkan prediksi harga yang lebih akurat dan adaptif, menjadikan analisis sentimen sebagai komponen esensial dalam sistem prediksi harga saham berbasis data.

4. SIMPULAN

Penelitian ini menunjukkan bahwa penggabungan analisis sentimen berita keuangan dari IDX Channel dengan model prediktif LSTM dapat menghasilkan hasil prediksi harga saham yang cukup akurat. Model LSTM menghasilkan nilai RMSE sebesar 35,92, dan visualisasi residual menunjukkan kesalahan prediksi tersebar hampir nol, yang mencerminkan kemampuan generalisasi model yang baik terhadap data baru. Prediksi harga satu hari ke depan menunjukkan nilai sebesar Rp 4.097,00, dan analisis korelasi antara sentimen dan harga saham menghasilkan nilai +0,27 (positif lemah). Hal ini mengindikasikan bahwa meskipun sentimen berita memiliki pengaruh terhadap pergerakan harga saham, pengaruh tersebut tidak terlalu signifikan secara statistik.

Secara keseluruhan, pendekatan hybrid yang mengintegrasikan LSTM dan analisis sentimen ini menunjukkan potensi sebagai alat bantu dalam pengambilan keputusan investasi berbasis data. Pengembangan lanjutan disarankan pada integrasi data multivariat, seperti volume perdagangan, indikator teknikal (misalnya moving average, RSI), serta sentimen dari media sosial (Twitter, forum

investor) guna memperkaya konteks prediktif dan meningkatkan akurasi model, terutama dalam skenario prediksi jangka panjang dan multi-horizon forecasting.

DAFTAR PUSTAKA

- Afrianto, N., Fudholi, D. H., & Rani, S. (2022). Prediksi Harga Saham Menggunakan BiLSTM dengan Faktor Sentimen Publik. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 6(1), 41–46. <https://doi.org/10.29207/resti.v6i1.3676>
- Amiri, B., Haddadi, A., & Farajpour Mojdehi, K. (2025). A Novel Hybrid GCN-LSTM Algorithm for Energy Stock Price Prediction: Leveraging Temporal Dynamics and Inter-Stock Relationships. *IEEE Access*, 13, 24815–24832. <https://doi.org/10.1109/ACCESS.2025.3536889>
- Aslim, M. F. (2024). *Pemanfaatan Analisis Sentimen Berbasis Lstm Dalam Prediksi Harga Saham*. Universitas Esa Unggul.
- Budiprasetyo, G., Hani'ah, M., & Aflah, D. Z. (2023). Prediksi Harga Saham Syariah Menggunakan Algoritma Long Short-Term Memory (LSTM). *Jurnal Nasional Teknologi dan Sistem Informasi*, 8(3), 164–172. <https://doi.org/10.25077/TEKNOSI.v8i3.2022.164-172>
- Chen, C., Xue, L., & Xing, W. (2023). *Research on Improved GRU-Based Stock Price Prediction Method*. <https://doi.org/10.20944/preprints202307.0472.v1>
- Christophorus Bintang Saputra, & Koesrindartoto, D. P. (2024). Pemanfaatan Analisis Sentimen Youtube untuk Prediksi Harga Saham: Studi pada Investor Retail Indonesia. *Jurnal Manajemen*, 21(1), 1–17. <https://doi.org/10.25170/jm.v21i1.5184>
- Ding, Y. (2023). Enhancing Stock Price Prediction Method Based on CNN-LSTM hybrid model. *Highlights in Business, Economics and Management*, 21, 774–781. <https://doi.org/10.54097/hbem.v21i.14760>
- Drus, Z., & Khalid, H. (2019). Sentiment Analysis in Social Media and Its Application: Systematic Literature Review. *Procedia Computer Science*, 161, 707–714. <https://doi.org/10.1016/j.procs.2019.11.174>
- Fahrezi, R. A., Wijaya, M. Y., & Fitriyati, N. (2024). Prediksi Harga Penutupan Saham Bank Central Asia: Implementasi Algoritma Long Short-Term Memory Dan Perbandingannya Dengan Support Vector Machine. *Jurnal Lebesgue : Jurnal Ilmiah Pendidikan Matematika, Matematika dan Statistika*, 5(1), 452–464. <https://doi.org/10.46306/lb.v5i1.582>
- Julian, R., & Pribadi, M. R. (2021). Peramalan Harga Saham Pertambangan Pada Bursa Efek Indonesia (BEI) Menggunakan Long Short Term Memory (LSTM). *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, 8(3), 1570–1580. <https://doi.org/10.35957/jatisi.v8i3.1159>
- Kartika, S. D., & Karmilasari. (2022). Implementasi Long Short-Term Memory Pada Prediksi Harga Saham PT Aneka Tambang Tbk. *Jurnal Ilmiah Komputasi*, 21(1). <https://doi.org/10.32409/jikstik.21.1.2815>
- Kwanda, K., Herwindiati, D. E., & Lauro, M. D. (2024). Perbandingan LSTM dan Bidirectional LSTM pada Sistem Prediksi Harga Saham Berbasis Website. *Ranah Research : Journal of Multidisciplinary Research and Development*, 7(1), 26–35. <https://doi.org/10.38035/rrij.v7i1.1255>
- Maghiszha, D. F. (2024, Desember 30). *Lampaui Target, Jumlah Investor Pasar Modal Sepanjang 2024 Tumbuh 2,5 Juta*. IDX Channel.
- Mamillapalli, A., Ogunleye, B., Inacio, S. T., & Shobayo, O. (2024). GRUvader-Sentiment Informed Stock Market Prediction. *Preprint*. <https://doi.org/10.20944/preprints202410.2465.v1>
- Nazhiroh, H., Dina Fitria, Dony Permana, & Zilrahmi. (2024). PT.Telkom (Tbk) Stock Price Forecasting Using Long Short Term Memory (LSTM). *UNP Journal of Statistics and Data Science*, 2(4), 414–421. <https://doi.org/10.24036/ujsds/vol2-iss4/223>

- Putra, F. P., Dewi, M. R., & Hibatullah, F. (2024). Analisis Sentimen Pasar melalui Berita Finansial untuk Prediksi Harga Saham PT Bank Rakyat Indonesia Tbk. *Indonesian Journal of Applied Informatics*, 9(1), 124. <https://doi.org/10.20961/ijai.v9i1.94692>
- Sari, E. P., Mustamin, S. B., Atnang, M., Sahriani, & Fajar, N. (2024). Studi Literatur Deep Learning dan Machine Learning untuk Analisis dan Prediksi Pasar Saham: Metodologi, Representasi Data, dan Studi Kasus. *Jurnal Teknologi dan Sains Modern*, 1(1), 19–28. <https://doi.org/10.69930/jtsm.v1i1.59>
- Sudhakar, K., & Naganjaneyulu, S. (2025). An optimised CNN-stacked LSTM neural network model for predicting stock market time-series data. *International Journal of Computational Economics and Econometrics*, 15(1/2), 196–224. <https://doi.org/10.1504/IJCEE.2025.145022>
- Tando, A. G., & Irawan, M. I. (2023). Analisis Dinamika Harga Saham yang Dipengaruhi oleh Analisis Sentimen di Media Sosial Menggunakan Algoritma Support Vector Machine. *Jurnal Sains dan Seni ITS*, 12(1). <https://doi.org/10.12962/j23373520.v12i1.107080>
- Teck, L. Y., & Thenata, A. P. (2025). Stock Price Prediction Using TCN-GAN Hybrid Model. *sinkron*, 9(1), 106–114. <https://doi.org/10.33395/sinkron.v9i1.14246>
- Wei, D., Wang, Z., Qiu, M., Yu, J., Yu, J., Jin, Y., Sha, X., & Ouyang, K. (2025). Multiple objectives escaping bird search optimization and its application in stock market prediction based on transformer model. *Scientific Reports*, 15(1), 5730. <https://doi.org/10.1038/s41598-025-88883-8>
- Wijayanti, R. I. (2025, Januari 17). *Pahami Penyebab ARA dan ARB dalam Waktu Singkat*. IDX Channel. <https://www.idxchannel.com/market-news/pahami-penyebab-ara-dan-arb-dalam-waktu-singkat?>
- William, K., & Rarasati, D. B. (2024). Stock Price Prediction on IDX30 Index using Long Short-Term Memory Algorithm. *Jurnal Informatika*, 11(2), 80–89. <https://doi.org/10.31294/inf.v11i2.22156>
- Wu, S., Liu, Y., Zou, Z., & Weng, T.-H. (2022). S_I_LSTM: stock price prediction based on multiple data sources and sentiment analysis. *Connection Science*, 34(1), 44–62. <https://doi.org/10.1080/09540091.2021.1940101>
- Zhang, J., Ye, L., & Lai, Y. (2023). Stock Price Prediction Using CNN-BiLSTM-Attention Model. *Mathematics*, 11(9), 1985. <https://doi.org/10.3390/math11091985>