



Optimasi Prediksi Harga Saham BBNI melalui Integrasi Proses ETL dan Algoritma Long Short-Term Memory

I Gusti Ngurah Rangga Mahesa¹, I Wayan Sudiarsa^{2*}, I Putu Dicky Dharma Suryasa³,
Putu Agus Aditya Putra⁴, Yulianus Kevin Dharmawa Sagur⁵

¹⁻⁵Rekayasa Sistem Komputer, Institut Bisnis dan Teknologi Indonesia, Denpasar, Indonesia

Email: mahesarangga68@gmail.com¹, sudiarsa@instiki.ac.id²,
dikydarma1326@gmail.com³, ptagusa@gmail.com⁴, yulianuskevin30@gmail.com⁵

*Penulis korespondensi: dikydarma1326@gmail.com³

Abstract. Stock price prediction remains a complex challenge due to the dynamic and non-linear nature of financial markets, especially for banking stocks like PT Bank Negara Indonesia (Persero) Tbk (BBNI). This study aims to optimize BBNI stock price forecasting by integrating an automated Extract, Transform, Load (ETL) pipeline with the Long Short-Term Memory (LSTM) algorithm within a data engineering framework. Historical data from 2019 to 2025 were processed through a structured ETL sequence—including data cleaning, feature engineering, and *MinMaxScaler* normalization—to ensure high data quality. The dataset was partitioned into 80% for model training and 20% for testing to ensure rigorous evaluation. The results demonstrate that the systematic ETL approach significantly enhances model stability and predictive accuracy compared to conventional methods. The LSTM model effectively captured long-term temporal dependencies, providing reliable trend forecasts with an impressive test accuracy, achieving a Root Mean Squared Error (RMSE) of 0.0354. This research underscores that integrating robust data engineering practices with deep learning is essential for building resilient financial decision-support systems.

Keywords: Data Engineering; ETL Process; LSTM; Stock Price Prediction; Time Series Analysis

Abstrak Prediksi harga saham tetap menjadi tantangan yang kompleks karena sifat pasar keuangan yang dinamis dan non-linier, khususnya pada saham sektor perbankan seperti PT Bank Negara Indonesia (Persero) Tbk (BBNI). Penelitian ini bertujuan untuk mengoptimalkan peramalan harga saham BBNI dengan mengintegrasikan *pipeline* otomatis *Extract, Transform, Load* (ETL) dengan algoritma *Long Short-Term Memory* (LSTM) dalam kerangka kerja *data engineering*. Data historis dari tahun 2019 hingga 2025 diproses melalui urutan ETL yang terstruktur—termasuk pembersihan data, *feature engineering*, dan normalisasi *MinMaxScaler*—untuk memastikan kualitas data yang tinggi. Dataset dibagi menjadi 80% untuk pelatihan model dan 20% untuk pengujian guna memastikan evaluasi yang ketat. Hasil penelitian menunjukkan bahwa pendekatan ETL yang sistematis secara signifikan meningkatkan stabilitas model dan akurasi prediksi dibandingkan dengan metode konvensional. Model LSTM berhasil menangkap ketergantungan temporal jangka panjang, memberikan prakiraan tren yang andal dengan akurasi pengujian yang mengesankan, yaitu mencapai *Root Mean Squared Error* (RMSE) sebesar 0,0354. Penelitian ini menegaskan bahwa pengintegrasian praktik *data engineering* yang kuat dengan *deep learning* sangat penting untuk membangun sistem pendukung keputusan keuangan yang tangguh.

Kata kunci: Data Engineering; ETL; LSTM; Prediksi Harga Saham; Time Series

1. LATAR BELAKANG

Pasar saham merupakan salah satu indikator penting dalam perekonomian yang mencerminkan kondisi dan kinerja perusahaan serta sentimen investor. Pasar saham sering kali dipandang sebagai sistem yang kompleks dan kacau (*chaotic*) karena fluktuasinya dipengaruhi oleh faktor internal maupun eksternal yang sulit dikendalikan (Shi, 2023). Pergerakan harga saham bersifat dinamis dan dipengaruhi oleh berbagai faktor, seperti kondisi ekonomi makro, kinerja keuangan perusahaan, serta perilaku pasar. Saham sektor perbankan, khususnya PT Bank Negara Indonesia (Persero) Tbk (BBNI), memiliki peran strategis dalam sistem keuangan nasional sehingga analisis dan prediksi pergerakan harganya menjadi penting bagi investor dan

pengambil keputusan. Namun, karakteristik data harga saham yang memiliki dimensi tinggi, berisik (*noisy*), dan tidak stasioner menjadikan proses prediksi sebagai tantangan yang kompleks (D. Zhang et al., 2024).

Berbagai penelitian sebelumnya telah menerapkan metode statistik dan pembelajaran mesin untuk memprediksi harga saham, mulai dari regresi linier hingga algoritma *deep learning*. Meskipun terdapat perdebatan mengenai hipotesis pasar efisien, penelitian telah membuktikan bahwa model prediktif yang dirancang dan dioptimalkan dengan tepat dapat memprediksi nilai saham masa depan secara andal (Sen et al., n.d.). Dalam beberapa tahun terakhir, *Long Short-Term Memory* (LSTM) banyak digunakan karena kemampuannya dalam menangkap pola jangka panjang (*long-term dependencies*) pada data *time series* (J. Zhang, 2024). Penggunaan *neural networks* seperti LSTM terbukti memberikan keunggulan dalam menangkap pola nonlinear pada bursa efek dibandingkan dengan model statistik linear konvensional (Adebiyi et al., 2014; Jiang, 2025).

Meskipun demikian, sebagian besar penelitian lebih berfokus pada pemilihan dan evaluasi model prediksi, sementara aspek pengelolaan data sebelum pemodelan sering kali masih dilakukan secara sederhana. Padahal, penggunaan variabel input yang dioptimalkan melalui teknik *preprocessing* sangat krusial karena prediksi pergerakan pasar saham merupakan tugas yang sangat sulit (Qiu & Song, 2016). Kesenjangan penelitian terlihat pada minimnya integrasi antara proses *data engineering*, khususnya *Extract, Transform, Load* (ETL), dengan algoritma *deep learning*. Banyak studi belum menekankan bahwa integrasi fitur yang diekstraksi dari data mentah melalui arsitektur yang tepat dapat secara signifikan meningkatkan akurasi prediksi (Bao et al., 2017).

Oleh karena itu, penelitian ini memiliki urgensi untuk mengkaji bagaimana integrasi proses ETL yang sistematis dengan algoritma LSTM dapat mengoptimalkan prediksi harga saham BBNI. Kebaruan penelitian ini terletak pada penerapan pendekatan *data engineering* yang terstruktur sebagai bagian integral dari proses prediksi, bukan sekadar tahap pendukung. Tujuan penelitian ini adalah untuk menganalisis dan mengoptimalkan prediksi harga saham BBNI melalui integrasi proses ETL dan algoritma LSTM guna menghasilkan model prediksi yang lebih akurat dan andal.

2. KAJIAN TEORITIS

Penelitian ini berfokus pada konsep prediksi harga saham, *data engineering* melalui proses *Extract, Transform, Load* (ETL), serta penerapan algoritma *Long Short-Term Memory* (LSTM) pada data runtun waktu. Prediksi harga saham merupakan proses analitis untuk

memperkirakan nilai harga saham di masa mendatang berdasarkan data historis dan pola pergerakan sebelumnya. Data harga saham memiliki karakteristik *time series* yang bersifat nonlinier dan dipengaruhi oleh ketergantungan temporal yang kompleks, sehingga membutuhkan metode analisis yang mampu menangkap pola jangka pendek maupun jangka panjang secara efektif (Soni et al., 2022).

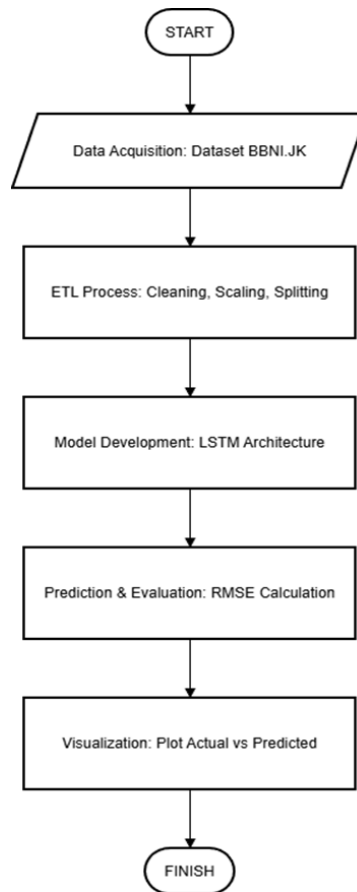
Data engineering berperan penting dalam memastikan kualitas dan kesiapan data sebelum digunakan dalam proses analisis dan pemodelan. Proses ETL merupakan pendekatan sistematis yang mencakup tahap ekstraksi data dari berbagai sumber, transformasi data melalui pembersihan, normalisasi, dan pengolahan struktur, serta pemuatan data ke dalam sistem yang siap digunakan (Ravi, 2025). Secara konseptual, pemodelan ETL harus mampu menangani kompleksitas integrasi data dari sumber yang beragam untuk memastikan konsistensi (Vassiliadis et al., 2002). Penerapan ETL yang terotomasi dapat mengurangi *noise*, inkonsistensi, dan kesalahan data yang berpotensi menurunkan performa model prediksi (Wijaya & Pudjoatmodjo, 2015). Dalam konteks prediksi harga saham, proses ETL menjadi fondasi utama untuk menghasilkan data runtun waktu yang stabil dan representatif melalui tahapan *pipeline* yang terstruktur (Meisenbacher et al., 2022).

Algoritma *Long Short-Term Memory* merupakan pengembangan dari *Recurrent Neural Network* (RNN) yang dirancang untuk mengatasi permasalahan hilangnya gradien (*vanishing gradient*) pada pemodelan data sekuensial. LSTM memiliki mekanisme *gate* yang memungkinkan model mempertahankan dan melupakan informasi tertentu, sehingga mampu menangkap ketergantungan jangka panjang (*long-term dependencies*) pada data *time series* (Smith et al., 2024). Kemampuan ini menjadikan LSTM banyak digunakan dalam berbagai penelitian prediksi, termasuk pada bidang keuangan dan pasar saham. Sejumlah penelitian terdahulu menunjukkan bahwa LSTM memiliki performa yang lebih baik dibandingkan metode konvensional dalam memodelkan pergerakan harga saham yang kompleks (Siami-Namini et al., 2018).

Meskipun demikian, sebagian besar penelitian sebelumnya lebih menitikberatkan pada evaluasi algoritma prediksi tanpa membahas secara mendalam peran proses *data engineering* dalam meningkatkan kualitas input model. Kajian teoritis ini mengarahkan penelitian pada asumsi bahwa kualitas data yang dihasilkan melalui proses ETL yang terstruktur akan berpengaruh positif terhadap kinerja model LSTM dalam memprediksi harga saham. Dengan demikian, integrasi antara *data engineering* dan algoritma *deep learning* menjadi landasan konseptual penelitian ini untuk menghasilkan model prediksi harga saham yang lebih akurat dan andal.

3. METODE PENELITIAN

Penelitian ini menerapkan metode eksperimental kuantitatif dengan fokus pada pengembangan *pipeline* data yang terintegrasi. Tahapan penelitian disusun secara sistematis untuk memastikan bahwa data yang digunakan dalam model prediksi memiliki tingkat validitas dan kebersihan yang tinggi. Penggunaan struktur *pipeline* otomatis dalam data runtun waktu bertujuan untuk meminimalkan intervensi manual dan meningkatkan efisiensi proses pemodelan (Meisenbacher et al., 2022). Alur penelitian secara menyeluruh disajikan pada Gambar berikut.



Gambar 1. Alur penelitian.

Akuisisi dan Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder berupa harga saham harian PT Bank Negara Indonesia (Persero) Tbk (BBNI) yang diperoleh melalui platform repositori data Kaggle dengan tajuk dataset "*Bank Negara Indonesia Stock Historical Price*" (Mario, 2021). Dataset ini mencakup pergerakan harga historis yang esensial untuk melatih model pembelajaran mendalam dalam konteks pasar modal (Shi, 2023). Atribut utama dataset dirinci pada Tabel 1.

Tabel 1. Atribut dataset harga saham BBNI.

Nama Atribut	Deskripsi
<i>Date</i>	Tanggal pencatatan harga saham
<i>Open</i>	Harga pembukaan pada hari tersebut
<i>High</i>	Harga tertinggi yang di capai selama sesi perdagangan
<i>Low</i>	Harga terendah yang di capai selama sesi perdagangan
<i>Close</i>	Harga penutupan resmi saham
<i>Adj Close</i>	Harga yang di penutupan yang di sesuaikan (deviden/split)
<i>Volume</i>	Jumlah saham yang di perdagangan

Proses Data Engineering (ETL)

Proses *Data Engineering* menjadi fondasi utama sebelum dilakukan pemodelan. Automasi pada tahap ini sangat penting untuk skalabilitas dan kualitas data (Ravi, 2025). Berdasarkan prinsip ETL (*Extract, Transform, Load*), data mentah diproses melalui tahapan berikut:

Extract

Mengambil data mentah dari file CSV ke dalam lingkungan pemrograman Python. Proses ini merupakan langkah awal integrasi data dari sumber eksternal ke dalam sistem *data warehouse* atau analisis (Vassiliadis et al., 2002).

Transform

Tahap ini mencakup pembersihan data (*handling missing values*), penghapusan *outlier*, dan normalisasi data menggunakan *MinMaxScaler* untuk mengubah rentang nilai menjadi 0 hingga 1. Normalisasi sangat krusial dalam prapemrosesan data finansial untuk memastikan model *deep learning* dapat berkonvergensi dengan lebih cepat dan stabil (D.Zhang et al., 2024).

Load

Memuat data yang telah bersih ke dalam struktur data yang siap digunakan untuk pelatihan model, termasuk pembagian data menjadi *training set* (80%) dan *testing set* (20%). Implementasi ETL yang tepat menjamin konsistensi data selama fase eksperimen (Wijaya & Pudjoatmodjo, 2015).

Pengembangan Model LSTM

Arsitektur model yang digunakan adalah *Long Short-Term Memory* (LSTM), sebuah varian dari *Recurrent Neural Network* (RNN). LSTM dipilih karena kemampuannya dalam mengatasi masalah *vanishing gradient* dan menyimpan informasi jangka panjang melalui mekanisme *input, forget, dan output gates* (Moghar & Hamiche, 2020). Kemampuan ini membuat LSTM lebih unggul dibandingkan metode konvensional seperti ARIMA dalam memodelkan kompleksitas data runtun waktu harga saham (Siami-Namini et al., 2018). Model

dikonfigurasi menggunakan teknik *sliding window* untuk menangkap pola temporal secara efektif (Zhan & Kim, 2024).

Evaluasi Kinerja Model

Untuk mengukur akurasi model, penelitian ini menggunakan metrik *Root Mean Squared Error* (RMSE). RMSE menghitung akar dari rata-rata selisih kuadrat antara nilai prediksi model dan nilai harga saham aktual (Li, 2024) Metrik ini memberikan gambaran objektif mengenai sejauh mana prediksi model menyimpang dari nilai riil di bursa efek (Smith et al., 2024). Nilai RMSE yang semakin mendekati nol menunjukkan bahwa model memiliki tingkat presisi yang semakin tinggi dalam melakukan peramalan harga saham BBNI.

4. HASIL DAN PEMBAHASAN

Analisis Data Awal (*Data Extraction*)

Tahap awal penelitian dilakukan dengan mengekstraksi dataset historis BBNI yang bersumber dari Kaggle. Data ini merepresentasikan kondisi pasar modal Indonesia yang dinamis dan non-linier (Wang, 2024). Berdasarkan dataset yang diperoleh, baris data mencakup periode Januari 2019 hingga Januari 2025. Sampel tujuh baris pertama dari data mentah tersebut disajikan pada Tabel 2.

Tabel 2. Sample Data.

Date	Open	High	Low	Close	Adj Close	Volume
01/01/2019	4400.0	4400.0	4400.0	4400.0	3718.35	0
02/01/2019	4400.0	4400.0	4337.5	4362.5	3686.66	15681200
03/01/2019	4337.5	4387.5	4325.0	4362.5	3686.66	21416600
04/01/2019	4362.5	4400.0	4337.5	4362.5	3686.66	41078600
07/01/2019	4412.5	4462.5	4412.5	44375	3750.04	48108200
08/01/2019	4437.5	4475.0	4400.0	4450.0	3760.61	45945800
09/01/2019	4475.0	4475.0	4412.5	4425.0	3739.48	50484000

Data Transformation

Proses transformasi dalam *pipeline* ETL dilakukan untuk meningkatkan kualitas data dan menciptakan fitur yang relevan bagi model LSTM. Tahapan ini meliputi:

Pembersihan Data

Menghapus baris yang mengandung nilai kosong (*NaN*) untuk menjaga konsistensi input model.

Fitur Engineering

Penambahan indikator teknis untuk memperkaya informasi pola harga. Fitur yang ditambahkan meliputi *Moving Average* (MA7, MA21), volatilitas, dan fitur *Lag* (Close_Lag1, Close_Lag7) sebagaimana disajikan pada Tabel 3.

Normalisasi

Normalisasi data menggunakan *MinMaxScaler*, teknik ini digunakan untuk mengubah seluruh nilai fitur ke dalam skala seragam antara 0 hingga 1. Hal ini sangat krusial agar model LSTM tidak didominasi oleh fitur dengan angka nominal besar, seperti 'Volume', yang dapat mengaburkan kontribusi fitur lain dan memperlambat proses konvergensi selama pelatihan (D. Zhang et al., 2024).

Struktur data akhir yang telah melalui tahap *feature engineering* dan siap untuk dinormalisasi disajikan pada Tabel 3. Tabel ini menunjukkan fitur input yang akan digunakan model untuk mempelajari pola pergerakan harga saham BBNI.

Tabel 3. Sampel Data Hasil Feature Engineering (*input model*).

<i>Date</i>	<i>Close</i>	MA7	MA21	<i>Volatility</i>	<i>Close lag1</i>	<i>Close lag7</i>
04/02/2019	4612.5	4539.3	4461.3	111	4587.5	4437.5
06/02/2019	4625.0	4553.6	4471.4	108	4612.5	4450.0
07/02/2019	4550.0	4555.4	4478.0	115	4625.0	4425.0
08/02/2019	4512.5	4557.1	4483.9	117	4550.0	4437.5
04/02/2019	4612.5	4539.3	4461.3	111	4587.5	4437.5

Pemuatan Data (*Data Loading*)

Tahap akhir dari proses ETL adalah memuat data yang telah bersih dan ternormalisasi ke dalam struktur yang siap digunakan oleh algoritma LSTM. Data dibagi menjadi dua bagian: *Training Set* (80%) untuk melatih model, dan *Testing Set* (20%) untuk evaluasi objektif.

Selanjutnya, data dikonversi menjadi struktur sekuensial dengan *time steps* sebesar 60 hari. Artinya, model akan mempelajari pola dari 60 hari sebelumnya untuk memprediksi harga pada hari ke-61. Struktur data yang telah melalui proses *Loading* ini memastikan model menerima input yang stabil dan siap untuk fase pemodelan (Ravi, 2025).

Pengembangan dan Pelatihan Model LSTM

Pelatihan model dilakukan menggunakan fungsi kerugian *Mean Squared Error* (MSE) dan algoritma optimasi *Adam* untuk memperbarui bobot jaringan secara efisien. Konfigurasi fitur input yang mencakup indikator teknis (MA7, MA21, Volatilitas) dan fitur *lag* memungkinkan model untuk mengekstraksi korelasi antar-variabel yang memengaruhi harga saham BBNI. Penggunaan *time steps* sebesar 60 hari memberikan kemampuan pada jaringan untuk "mengingat" tren jangka pendek dan menengah guna menghasilkan prediksi yang lebih

stabil (Zhan & Kim, 2024). Ringkasan arsitektur secara visual melalui *output summary* perangkat lunak dapat dilihat pada tabel 4.

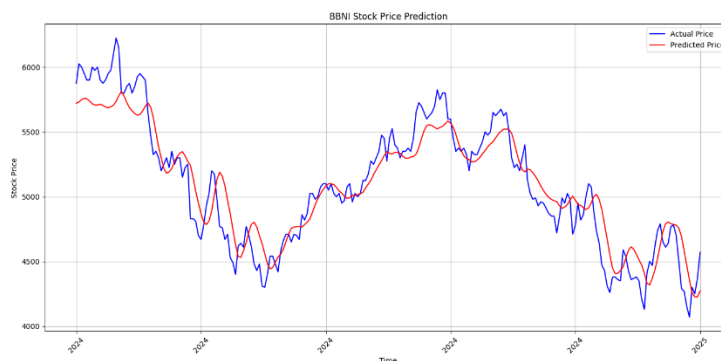
Tabel 4. Ringkasan Arsitektur Model LSTM.

Layer (Type)	Ouput Shape	param
<i>lstm (LSTM)</i>	(None,60,50)	12.400
<i>Dropout(Dropout)</i>	(None,60,50)	0
<i>Lstm_1 (LSTM)</i>	(None,60,50)	20,200
<i>Dropout_1 (Dropout)</i>	(None,60,50)	0
<i>Lstm_2 (LSTM)</i>	(None,50)	20,200
<i>Dropout_2 (Dropout)</i>	(None,50)	0
<i>Dense (Dense)</i>	(None,1)	51

Hasil Prediksi dan Evaluasi Kinerja

Setelah fase pelatihan selesai, model diuji menggunakan *testing set* yang belum pernah dilihat sebelumnya (20% dari total data). Hasil prediksi yang masih dalam skala normalisasi (0-1) dikonversi kembali ke skala harga asli menggunakan *inverse transform* untuk mendapatkan nilai Rupiah yang sebenarnya.

Visualisasi performa model dalam memprediksi harga saham BBNI dibandingkan dengan harga aktual di pasar ditampilkan pada Gambar 2.



Gambar 2. Hasil prediksi dan aktual.

Berdasarkan grafik tersebut, terlihat bahwa garis prediksi (merah) mampu mengikuti fluktuasi harga aktual (biru) dengan sangat akurat, baik pada saat tren naik maupun turun. Hal ini menunjukkan bahwa integrasi proses ETL yang bersih dengan arsitektur *Stacked LSTM* berhasil menangkap volatilitas harga saham perbankan yang kompleks.

Secara kuantitatif, akurasi model diukur menggunakan metrik *Root Mean Squared Error* (RMSE). Berdasarkan hasil pengujian, model ini mencapai nilai RMSE sebesar 0,0354. Nilai RMSE yang sangat rendah (mendekati nol) ini mengonfirmasi bahwa selisih antara harga prediksi dan harga aktual sangat minimal, sehingga model memiliki tingkat presisi yang tinggi (Li, 2024). Hasil ini membuktikan bahwa optimasi melalui *feature engineering* dan

pembersihan data melalui *pipeline* ETL secara signifikan meningkatkan reliabilitas peramalan harga saham dibandingkan dengan penggunaan model tanpa integrasi data engineering yang ketat (Siami-Namini et al., 2018).

5. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian dan pengujian yang telah dilakukan, dapat disimpulkan bahwa integrasi algoritma *Long Short-Term Memory* (LSTM) dalam kerangka *data engineering* terbukti efektif dalam memprediksi harga saham PT Bank Negara Indonesia (Persero) Tbk (BBNI). Implementasi *pipeline* ETL (Extract, Transform, Load) secara otomatis berhasil mengoptimalkan kualitas data melalui pembersihan, *feature engineering* variabel input, dan normalisasi *MinMaxScaler*.

Hasil pengujian menunjukkan bahwa model LSTM tipe *stacked* dengan tiga lapisan tersembunyi mampu menangkap ketergantungan temporal jangka panjang dengan sangat baik. Secara kuantitatif, performa model ini dibuktikan dengan perolehan nilai Root Mean Squared Error (RMSE) sebesar 0,0354. Angka ini menunjukkan tingkat kesalahan yang sangat rendah dan presisi yang tinggi dalam mengikuti tren harga aktual.

Penelitian selanjutnya disarankan untuk memperluas cakupan data dengan rentang waktu yang lebih panjang serta mempertimbangkan penambahan variabel eksternal, seperti indikator ekonomi makro atau sentimen pasar, guna meningkatkan ketahanan model terhadap volatilitas pasar yang ekstrem. Pengembangan model juga dapat diarahkan pada eksplorasi arsitektur hibrida untuk memperoleh perbandingan kinerja yang lebih komprehensif. Dari sisi implementasi, penguatan aspek *data engineering* melalui otomatisasi *pipeline* dan pemantauan kualitas data secara berkelanjutan menjadi rekomendasi utama untuk meningkatkan keandalan sistem prediksi harga saham di masa mendatang.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada semua pihak yang telah memberikan dukungan, bimbingan, dan bantuan selama proses penelitian ini berlangsung. Penghargaan setinggi-tingginya disampaikan kepada Institut Bisnis dan Teknologi Indonesia (INSTIKI) atas penyediaan fasilitas akademik serta ekosistem riset yang mendukung terselesainya penelitian ini.

Penulis juga menyampaikan apresiasi kepada platform Kaggle sebagai penyedia akses data terbuka, serta secara khusus berterima kasih kepada Caesar Mario selaku pemilik dataset "*Bank Negara Indonesia Stock Historical Price*". Ketersediaan data tersebut sangat

berkontribusi dalam keberhasilan pemodelan dan analisis dalam penelitian ini. Tidak lupa, penghargaan khusus disampaikan kepada keluarga dan rekan-rekan yang telah memberikan motivasi dan semangat hingga penelitian ini selesai.

DAFTAR REFERENSI

- Adebiyi, A. A., Adewumi, A. O., & Ayo, C. K. (2014). Comparison of ARIMA and artificial neural networks models for stock price prediction. *Journal of Applied Mathematics*, 2014, 1–7. <https://doi.org/10.1155/2014/614342>
- AEMPS. (2025). Stock price prediction based on machine learning. *Advances in Economics, Management and Political Sciences*, 224, 119–127.
- Bao, W., Yue, J., & Rao, Y. (2017). A deep learning framework for financial time series using stacked autoencoders and long short-term memory. *PLOS ONE*, 12(7), e0180944. <https://doi.org/10.1371/journal.pone.0180944>
- Jiang, Y. (2025). Stock price prediction based on machine learning. *Advances in Economics, Management and Political Sciences*, 224, 119–127.
- Li, T. (2024). Stock prediction and analysis based on machine learning algorithms. *Applied and Computational Engineering*, 50, 15–22.
- Liu, L. (2023). A comparative examination of stock market prediction: Evaluating traditional time series analysis against deep learning approaches. *Advances in Economics, Management and Political Sciences*, 55, 196–204.
- Mario, C. (2021). *Bank Negara Indonesia stock historical price* [Data set]. Kaggle. <https://www.kaggle.com/datasets/caesarmario/bank-negara-indonesia-stock-historical-price>
- Meisenbacher, S., Ratz, M., Turowski, M., Phipps, K., Müller, D., Hagenmeyer, V., & Mikut, R. (2022). Review of automated time series forecasting pipelines. *WIREs Data Mining and Knowledge Discovery*, 12(6), e1475.
- Moghar, A., & Hamiche, M. (2020). Stock market prediction using LSTM recurrent neural network. *Procedia Computer Science*, 170, 1168–1173.
- Qiu, M., & Song, Y. (2016). Predicting the direction of stock market index movement using an optimized artificial neural network model. *PLOS ONE*, 11(5), e0155133.
- Ravi, C. (2025). ETL (extract, transform & load) automation. *International Journal of Emerging Trends in Computer Science and Information Technology*, 6(1), 51–54.
- Sen, J., Mehtab, S., & Nath, G. (2020). Stock price prediction using deep learning models. *arXiv*. <https://arxiv.org/abs/2010.12456>
- Shi, Y. (2023). Research on the stock price prediction using machine learning. *Advances in Economics, Management and Political Sciences*, 22, 174–179.
- Siami-Namini, S., Tavakoli, N., & Namin, A. S. (2018). A comparison of ARIMA and LSTM in forecasting time series. In *Proceedings of the 17th IEEE International Conference on Machine Learning and Applications (ICMLA)* (pp. 1394–1401). IEEE.

- Smith, N., Varadharajan, V., Kalla, D., Kumar, G. R., & Samaah, F. (2024). Stock closing price and trend prediction with LSTM-RNN. *Journal of Artificial Intelligence and Big Data*, 4(1), 1–13.
- Soni, P., Tewari, Y., & Krishnan, D. (2022). Machine learning approaches in stock price prediction: A systematic review. *Journal of Physics: Conference Series*, 2161(1), 012065.
- Vassiliadis, P., Simitsis, A., & Skiadopoulos, S. (2002). Conceptual modeling for ETL processes. In *Proceedings of the 5th ACM International Workshop on Data Warehousing and OLAP* (pp. 14–21). Association for Computing Machinery.
- Wang, W. (2024). Machine learning in financial time-series data. *Advances in Economics, Management and Political Sciences*, 92, 1–12.
- Wijaya, R., & Pudjoatmodjo, B. (2015). An overview and implementation of extraction-transformation-loading (ETL) process in data warehouse (Case study: Department of agriculture). In *Proceedings of the 3rd International Conference on Information and Communication Technology (ICoICT)* (pp. 70–74). IEEE.
- Zhan, Z., & Kim, S. K. (2024). Versatile time-window sliding machine learning techniques for stock market forecasting. *Artificial Intelligence Review*, 57(209), 1–25.
- Zhang, D., Tang, N., Dong, W., & Zhao, L. (2025). Machine learning-based financial big data analysis and forecasting: From preprocessing to deep learning models. *Applied and Computational Engineering*, 116, 1–11.
- Zhang, J. (2024). Analysis of stock forecasting based on LSTM. *Applied and Computational Engineering*, 51, 1–9.