



(ANTM) Stock Price Prediction and Modelling Using Long Short Term Memory (LSTM)

Pemodelan dan Prediksi Harga Saham ANTM Menggunakan Long Short-Term Memory (LSTM)

Muhammad Naufal Luthfi^{1*}, Assad Hidayat², Ivan Chatissa³, Frisca Martha Veronica⁴

^{1,2,3,4} Department of Information Technology, Caltex Riau Polytechnic, Indonesia
E-Mail : naufal@pcr.ac.id¹, assad@pcr.ac.id², ivan@pcr.ac.id³, frisca@pcr.ac.id⁴

Makalah: Diterima 19 juli 2026; Diperbaiki 11 Agustus 2026; Disetujui 20 Agsutus 2026
Corresponding Author: Muhammad Naufal Luthfi

Abstrak

Saham merupakan salah satu perangkat atau alat investasi yang digunakan oleh banyak orang, Saham mempunyai pilihan menarik yaitu saham bisa disimpan atau diinvestasikan dan mampu memberikan return yang menarik berdasarkan pertumbuhan korporasi atau perusahaan. Banyak faktor yang dapat mempengaruhi pergerakan harga saham, diantaranya faktor internal perusahaan, maupun faktor dari external perusahaan. Penelitian ini menggunakan salah satu perusahaan terkenal di Indonesia. Perusahaan pertambangan dan logam mulia terkemuka di Indonesia, bergerak dalam eksplorasi, penambangan, pengolahan, serta pemasaran komoditas mineral seperti emas, nikel dan bauksit. Machine Learning adalah salah satu topik populer untuk memberikan wawasan dan rekomendasi mengenai pergerakan harga saham. Dalam penelitian ini, peneliti akan menggunakan salah satu machine learning Long Short Term Memory dalam memprediksi pergerakan harga saham. Dengan menggunakan data harga saham ANTM dari tahun 2022 hingga 2026. Menunjukkan hasil bahwa LSTM mampu memberikan hasil evaluasi R-squared sebesar 0.92, RMSE sebesar 128.91 dan MAPE sebesar 0.027, dan memiliki kesimpulan bahwa LSTM menunjukkan performa yang sangat bagus dengan akurasi tinggi yang dilihat dari R2, Nilai toleransi yang wajar dengan nilai RMSE.

Kata Kunci : Artificial Intelligence, Long Short Term Memory, Machine Learning, Prediksi Pergerakan Harga Saham

Abstract

Stocks are an investment tool used by many people. Stocks offer an attractive option: they can be held or invested, and they can provide attractive returns based on corporate or company growth. Many factors can influence stock price movements, including internal and external factors. This study uses a well-known company in Indonesia, precious metals company in Indonesia, engaged in the exploration, mining, processing, and marketing of mineral commodities such as gold, nickel, and bauxite. Machine learning is a popular topic for providing insights and recommendations regarding stock price movements. In this study, researchers will use machine learning Long Short Term Memory (LST) to predict. By using ANTM stock price data from 2022 to 2026. Showing the results that LSTM is able to provide R-squared evaluation results of 0.92, RMSE of 128.91 and MAPE of 0.027, and has the conclusion that LSTM shows very good performance with high accuracy as seen from R2, reasonable tolerance value with RMSE value.

Keyword: Artificial Intelligence, Long Short Term Memory, Machine Learning

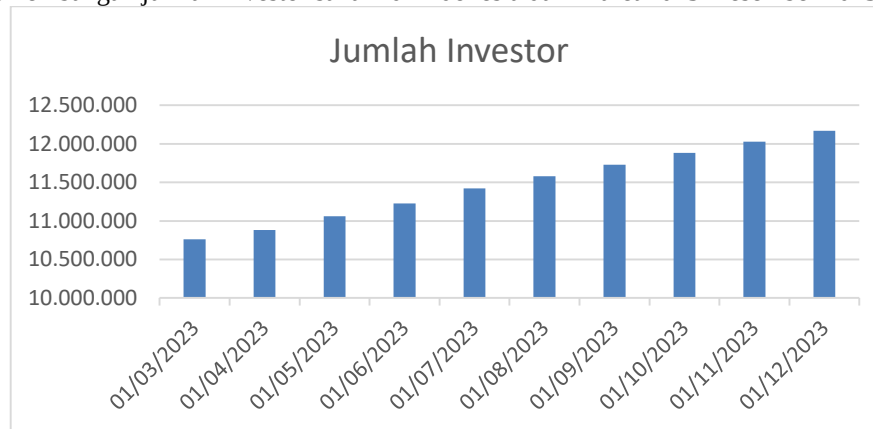
1. PENDAHULUAN

Saham merupakan alat atau instrumen investasi yang sering digunakan banyak orang karena dapat memberikan imbal hasil yang menarik. Saham dapat diartikan sebagai partisipasi seseorang atau badan hukum dalam suatu perusahaan atau perseroan terbatas (PT). Rata-rata surplus yang diperoleh setiap orang yang menginvestasikan modal mencapai antara 11-13% per tahun, didukung oleh kinerja IHSG. Namun, banyak investor di Indonesia ingin ikut serta dalam pembelian saham tetapi mendapatkan hasil yang tidak sesuai harapan atau bahkan mengalami kerugian. Keuntungan membeli saham antara lain adanya dividen, yang memberikan persentase keuntungan kepada pemegang saham; jumlah yang diberikan bergantung pada hasil rapat umum pemegang saham. Keuntungan membeli saham adalah adanya dividen, yang memberikan persentase keuntungan kepada pemegang saham; jumlah yang diberikan bergantung pada hasil. Anda

mendapatkan keuntungan modal ketika pemegang saham menjual saham mereka dengan harga yang lebih menguntungkan daripada harga biasanya. Selain keuntungan, ada juga kerugian dalam membeli saham, seperti tidak mendapatkan dividen karena perusahaan tempat kita berinvestasi tidak memiliki kinerja yang baik atau kinerja menurun. Yang kedua adalah kebalikan dari keuntungan modal, yaitu kerugian modal, di mana pemegang saham menjual saham dengan kondisi yang lebih rendah dari harga pembelian.

Adapula faktor-faktor yang dapat mempengaruhi harga saham dibagi menjadi 2 diantaranya ada faktor external dan faktor internal dimana kondisi external dan internal perusahaan menjadi pengaruh dalam pergerakan harga saham sebuah perusahaan, adapun faktor pergerakan nilai saham pada faktor external diantaranya kondisi fundamental ekonomi makro, fluktuasi kurs rupiah terhadap mata uang asing, kebijakan pemerintah, faktor panik dan manipulasi pasar, faktor faktor ini terjadi diluar kendali internal perusahaan. Dan berikut faktor internal yang dapat mempengaruhi harga saham dampak dari aktivitas perusahaan diantaranya faktor fundamental perusahaan, aksi korporasi perusahaan, proyeksi kinerja perusahaan pada masa mendatang.[1]

Yang terakhir adalah risiko likuiditas, yaitu ketika emiten bangkrut atau dilikuidasi sehingga probabilitas terburuknya adalah pemegang saham tidak akan memperoleh keuntungan. Menurut berita CNBC Indonesia pada 25 Januari 2024, yang ditulis oleh (Puspadini,2024), jumlah investor di Indonesia mencapai 12,2 juta Pemegang Kartu Identitas Investor Tunggal (SID) sebagaimana dinyatakan oleh direktur pengembangan bursa saham Indonesia (Gambar. 1), Bapak Jeffrey Hendrick, yang mengatakan bahwa jumlah investor di Indonesia hanya mencakup 5% dari total populasi di Indonesia. Angka ini sangat kecil dibandingkan dengan Singapura, yang telah mencapai 30-40% dari total populasinya. Berikut adalah grafik yang menunjukkan perkembangan jumlah investor saham di Indonesia dari Maret 2023-Desember 2023 [2]



Gambar 1. Jumlah Investor di Indonesia

Di era teknologi saat ini, kecerdasan buatan dapat membantu memprediksi harga saham dengan mempelajari pola pada setiap variabel yang lebih baik daripada metode statistik sederhana. Pembelajaran Mesin (Machine Learning/ML) dan Pembelajaran Mendalam (Deep Learning/DL) dapat membuat prediksi berdasarkan data deret waktu. Menurut Suresh Chavhan dkk *deep learning* telah secara signifikan meningkatkan minat dalam mengembangkan model yang akurat dan andal untuk peramalan di pasar keuangan, khususnya untuk prediksi harga saham [3]. ML diimplementasikan untuk membaca pola data guna menentukan pergerakan data saham dan mengurangi risiko investasi saat mengambil keputusan. Sementara itu, DL adalah teknologi yang digunakan untuk mensimulasikan jaringan saraf seperti manusia dan memecahkan masalah non-linier yang kompleks[2]

Memprediksi data deret waktu umumnya sangat sulit karena perubahan yang belum pernah terjadi sebelumnya yang disebabkan oleh perubahan tren ekonomi. Oleh karena itu, penilaian akurasi prediksi sangat diperlukan ketika menggunakan berbagai bentuk model pembelajaran mesin, karena kita tahu bahwa setiap model memiliki beberapa keterbatasan. Beberapa contoh model yang digunakan untuk menganalisis data deret waktu salah satunya adalah Long Short-Term Memory (LSTM).

Data Mining adalah metode yang digunakan untuk mengekstrak sejumlah besar data dan mencari pola atau wawasan dari kumpulan data tersebut. Teknik yang digunakan dalam *data mining* adalah teknik statistik untuk meringkas dan menganalisis pola dan tren dalam data, kemudian pembelajaran mesin untuk membangun algoritma yang belajar dari data dan membuat prediksi, dan akhirnya, sistem basis data untuk menyimpan, mengelola, dan mengambil data secara otomatis dan efisien. Terdapat beberapa metodologi yang digunakan dalam melakukan proses data mining, termasuk CRISP-DM (*Cross Industry Standard Process for Data Mining*), KDD (*Knowledge Discovery in Database*), dan SEMMA (*Sample, Explore, Modify, Model, Assess*). Namun, dalam penelitian ini, kami akan menggunakan kerangka kerja CRISP-DM sebagai proses *data mining*.

Machine Learning adalah teknik untuk mempelajari pola dan bentuk menggunakan data dan statistik. Model pembelajaran mesin bekerja dengan memberikan masukan berupa data sehingga model yang dibuat

dapat memberikan keluaran berupa keputusan, rekomendasi, dan prediksi. Ada tiga kategori bagaimana pembelajaran mesin belajar:

1. *Supervised Learning* adalah di mana model yang dibuat mempelajari pola dari data yang telah diberi label. Model ini dapat memetakan input dan output.
2. *Unsupervised Learning* adalah model yang digunakan pada data yang tidak memiliki label. *Unsupervised Learning* bertugas meninjau data sehingga pola tersembunyi atau pengelompokan data dapat dieksplorasi. Biasanya digunakan dalam analisis pengelompokan atau clustering.
3. *Reinforcement Learning* adalah metode yang digunakan untuk pengguna yang harus membuat keputusan dan tindakan dalam keadaan tertentu dengan tujuan mendapatkan imbalan maksimal.

Ketiga kategori di atas digunakan dalam kondisi dan bentuk data yang berbeda untuk mempelajari pola.

Recurrent Neural Network (RNN) memiliki fitur dasar yaitu model ini bekerja dalam setidaknya satu loop umpan balik sehingga fungsi label yang dipilih dapat bekerja dalam setidaknya satu loop. RNN juga dirancang untuk memproses data sekuensial seperti data teks, ucapan, atau data deret waktu. Modul RNN akan melakukan setidaknya satu pengulangan [4]. *Long Short-Term Memory* (LSTM) adalah jenis RNN yang dirancang untuk mengatasi kelemahan RNN standar. LSTM memperkenalkan portal khusus yang dapat mengontrol aliran informasi masuk dan keluar dari lapisan tersembunyi, yang merupakan unit penyimpanan jaringan. Portal ini dapat menentukan informasi mana yang dapat disimpan dan dibuang, serta berapa banyak yang dapat dimasukkan.

Cross Industry Process for Data Mining (CRISP-DM) adalah metodologi dengan pendekatan dan deskripsi proses untuk proyek penambangan data. Terdapat enam proses implementasi dalam metodologi CRISP-DM, yaitu pemahaman bisnis, pemahaman data, persiapan data, pemodelan, evaluasi, dan penerapan [5]. Brzozowska et al. (2023), menyatakan bahwa CRISP-DM mampu mencapai hasil model berkualitas baik; model yang ditunjukkan dapat digunakan untuk dukungan analitis, perencanaan, dan pengambilan keputusan. CRISP-DM juga memfasilitasi kelancaran operasional suatu organisasi dengan meningkatkan akurasi pengambilan keputusan dan mengidentifikasi proses pengambilan keputusan [6].

Penelitian Bansal, dkk (2022), ini memprediksi harga saham menggunakan lima model pembelajaran mesin, termasuk K-Nearest Neighbors, regresi linier, regresi vektor pendukung, regresi pohon keputusan, dan memori jangka pendek panjang. Data diambil dari 12 perusahaan India selama lebih dari 7 tahun untuk analisis. Hasil penelitian ini menunjukkan bahwa LSTM mengungguli semua algoritma dalam hal tingkat akurasi, kemudian diikuti oleh SVR, yang berada di urutan kedua dalam hal kinerja, Regresi Linier, dan Regresi Pohon Keputusan menunjukkan hasil yang sama, dan algoritma terakhir, K-NN, menunjukkan hasil yang kurang memuaskan dalam memprediksi harga saham.[7]

Dalam jurnal Ketut dkk. (2023) ini membandingkan model optimasi (Adam, SGD, RMSprop) pada LSTM yang bertujuan untuk memprediksi harga saham Telkom Indonesia, Tbk dari 1 Januari 2019 hingga 11 Januari 2023. LSTM menunjukkan hasil dengan akurasi prediksi yang sangat baik dengan nilai rendah menggunakan Mean Absolute Percentage Error (MAPE). Dan optimasi Adam menunjukkan akurasi sebesar 98%.[8]

Jurnal Yan dkk. (2021), ini membahas penggunaan jaringan saraf dalam LSTM untuk memprediksi harga saham berdasarkan data dari N hari sebelumnya. Membandingkan LSTM dengan jaringan saraf lain dan model statistik tradisional, tujuannya adalah untuk meningkatkan akurasi prediksi pada data deret waktu pasar keuangan. Penelitian ini juga berfokus pada pengoptimalan proses pelatihan, membandingkan perbedaan dalam metode pengoptimalan, dan mengeksplorasi konsekuensi dari input yang diberikan. Hasil penelitian yang diperoleh dari penelitian ini adalah bahwa jaringan saraf dalam LSTM sangat efektif dalam memprediksi harga saham dan dalam mengatasi masalah gradien yang menghilang pada RNN tradisional. [9].

Studi Yohannes E, dkk (2025), ini mengevaluasi kinerja model *Long Short-Term Memory* (LSTM) untuk memprediksi harga saham, menggunakan data dari Bursa Efek Pakistan. Hasil eksperimen menyoroti kemampuan prediksi model yang kuat, sebagaimana dibuktikan oleh beberapa metrik evaluasi utama. Temuan menunjukkan bahwa model LSTM adalah alat yang berharga untuk prediksi harga saham, menawarkan wawasan berbasis data yang dapat membantu investor dalam membuat keputusan yang lebih tepat [10].

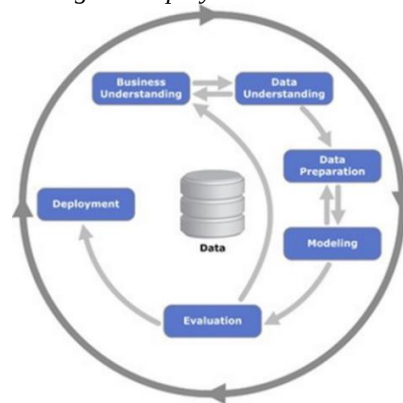
Makalah Farhadi, dkk (2025) ini memperkenalkan model hibrida baru yang menggabungkan *Long Short-Term Memory* (LSTM) dan *Gated Recurrent Unit* (GRU) untuk prediksi harga saham, yang dievaluasi secara khusus pada Bursa Efek Tehran (TSE). Model ini bertujuan untuk meningkatkan akurasi prediksi dengan memanfaatkan keunggulan jaringan LSTM maupun GRU serta mengintegrasikan serangkaian indikator teknikal yang komprehensif. Menunjukkan hasil dan memvalidasi model hibrida LSTM-GRU yang meningkatkan akurasi prediksi harga saham dengan memadukan secara efektif kemampuan khas jaringan GRU dan LSTM serta mengintegrasikan beragam indikator teknikal [11].

Jurnal Gao (2021) ini untuk membahas penggunaan model Recurrent Neural Network (RNN) dan Sequence to Sequence (Seq2Seq) dalam pembelajaran mesin, khususnya untuk tugas-tugas seperti penerjemahan mesin dan prediksi indeks saham. Jurnal ini menjelaskan fungsi dan struktur RNN, Long Short-Term Memory (LSTM), dan lapisan perhatian dalam model ini, serta menyediakan contoh kode untuk mengimplementasikan Seq2Seq dalam memprediksi indeks harga saham. Penelitian ini juga membandingkan kinerja model data deret waktu yang berbeda untuk memprediksi pergerakan harga saham dari Dow Jones Industrial Average (DIJA) dan menemukan bahwa model LSTM dan Seq2Seq mengungguli model lain dalam menghitung mean squared error. Fitur-fiturnya meliputi latensi rendah, relativitas, dan prediksi sekuensial. [12]

Penelitian Dwivedi dkk. (2021), yang menganalisis dan memprediksi data deret waktu, yaitu data indeks Nifty-500 di pasar saham menggunakan S ARIMA, CNN, dan LSTM. Penelitian ini juga melakukan perbandingan dan evaluasi kinerja masing-masing model, yang menunjukkan hasil yang menjanjikan. Alat yang digunakan dalam penelitian ini adalah TensorFlow dan Keras untuk implementasi. MSE digunakan untuk mengukur akurasi prediksi yang dibuat oleh model. Hasil penelitian ini menunjukkan bahwa model deep learning mengungguli model machine learning tradisional.[13]

2. METODE DAN BAHAN

Peneliti menggunakan kerangka kerja CRISP-DM untuk memprediksi pergerakan harga saham. Bab ini akan menjelaskan kerangka kerja tersebut dalam enam tahapan yang terdiri dari *Business understanding, data understanding, data preparation, modeling and deployment*.

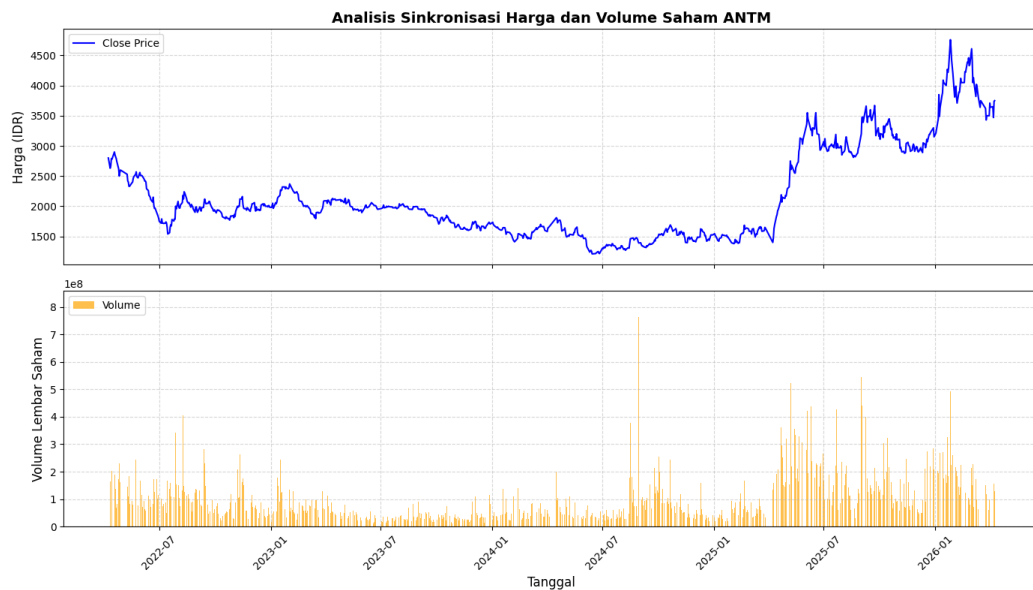


Gambar 2. CRISP-DM Methodology

2.1. Business Understanding

PT Aneka Tambang Tbk atau dikenal Antam, merupakan perusahaan tambang yang ada di Indonesia. Perusahaan ini memproduksi nikel dan emas, Antam juga merupakan produsen nikel terbesar di Indonesia. Antam merupakan perusahaan milik negara sebelum kepemilikannya dialihkan ke PT Mineral Industri Indonesia (Persero) atau MIND ID, salah satu perusahaan holding milik pemerintah.

Menganalisa saham ANTM sangat penting karena emiten sedang mencatat kinerja keuangan tertinggi sepanjang sejarah perusahaan. Memahami dinamika fundamental dan teknikalnya akan membantu melihat potensi keuntungan maupun resiko dari salah satu saham komoditas paling likuid di Bursa Efek Indonesia. Alasan utama menganalisa ANTM saat ini yaitu terdapat nya anomali lonjakan volume (Pertengahan 2024) pada **gambar 3**.



Gambar 3. Analisis Sinkronisasi Harga dan Volume Saham Antam

2.2. Data Understanding

Data Understanding disini menjelaskan bagaimana kita mendapatkan data dan menjelaskan setiap atribut yang terdapat dalam data dan mendefinisikan data yang akan digunakan dalam penelitian ini. Data dikumpulkan melalui web <https://id.investing.com/> dan mencari saham ANTAM dengan memasukkan ticker name “ANTM” dan melakukan filter date dari bulan april 2022 hingga april 2026 agar kita memiliki banyak data untuk kita training kan. Data yang kita peroleh untuk penelitan ini berjumlah 953 data.

Tanggal	Terakhir	Pembukaan	Tertinggi	Terendah	Vol.	Perubahan%
09/04/2026	3.750	3.680	3.800	3.630	129,29M	+0.27%
08/04/2026	3.740	3.630	3.750	3.580	157,47M	+7.78%
07/04/2026	3.470	3.620	3.630	3.450	129,42M	-4.93%
06/04/2026	3.650	3.640	3.650	3.570	41,21M	+0.27%
02/04/2026	3.640	3.680	3.690	3.590	71,72M	-1.89%
01/04/2026	3.710	3.620	3.760	3.570	140,58M	+6.00%
31/03/2026	3.500	3.540	3.620	3.430	99,83M	0.00%
30/03/2026	3.500	3.450	3.550	3.390	60,55M	0.00%
27/03/2026	3.500	3.350	3.590	3.350	91,49M	+2.04%

Gambar 4. Website Investing.com

Jenis data tersebut akan dijelaskan dalam Tabel (1) untuk tabel data pergerakan stok.

Tabel 1. Deskripsi Masing Masing Attribute Data Saham

Attribute Name	Description
Date	Tanggal harga open dan close saham
Open	Harga buka saham pada satu hari
High	Harga tertinggi saham pada satu hari
Close	Harga tutup saham pada satu hari
Low	Harga saham terendah pada satu hari
Volume	Total jumlah lembar saham jual beli yang terjadi
Perubahan	Persen perubahan harga close saham

Dan pada Gambar 3, terlihat grafik harga tutup saham dan jumlah volume saham yang terjual, kita bisa melihat tren pergerakan saham ANTM pada saat naik secara agresif di tahun 2025 volume nya juga ikut naik secara konsisten, dan pada saat 2026 setelah menyentuh puncak tertinggi harga saham mengalami penurunan diiringi dengan volume perdagangan mulai mereda (menurun). Dan kita bisa melihat tahun 2024 bahwa saat harga turun, volume mengalami peningkatan yang sangat signifikan.

2.3. Data Preprocessing

Di sini, proses yang dilakukan terdiri dari, pembersihan data, *splitting the data* dan pengujian data. Dalam penelitian ini, pra-pemrosesan data dilakukan di Google Colab menggunakan Python. Tujuan pra-pemrosesan data adalah untuk menciptakan kualitas data terbaik sehingga dapat dilanjutkan ke tahap berikutnya. Data dalam penelitian ini di namakan "ANTM_cleaned," yang merupakan hasil dari preprocessing data yang akan kita laksanakan.

2.3.1. Pembersihan Data

Pada proses ini, kita melakukan pembersihan data yaitu dengan mengganti huruf M berarti juta (*Million* pada bahasa inggris) pada kolom volume menjadi angka, dan disini kita juga akan drop kolom perubahan karena korelasi dengan attribute lain sangat kecil tujuan pembersihan data agar data bisa digunakan dalam pelatihan model. Berikut dibawah syntax yang digunakan

Data cleaning

```
# Bersihkan kolom Vol. (Ganti koma ke titik, hilangkan 'M', kalikan 1.000.000)
df['Vol.']= df['Vol.'].str.replace(',','').str.replace('M','').astype(float)
* 1_000_000
# Menghapus kolom 'Perubahan%'
df = df.drop(columns=['Perubahan%'])
```

2.3.2 Splitting the Data

Sebelum menerapkan model pada dataset yang berjumlah 953 sampel, dataset tersebut perlu dibagi menjadi dua bagian dengan proporsi 70% atau 667 sampel untuk data pelatihan dan 30% atau 286 sampel untuk data pengujian. Menurut (Bichri dkk., 2024), peningkatan ukuran data pelatihan hingga 70% dari keseluruhan dataset diperlukan pada tahap pelatihan untuk mencapai kinerja yang lebih baik [14]. Dengan syntax berikut

Split Data

```
# Split data into training and testing sets
training_data_len = int(len(data) * 0.7)
train_data = scaled_data[0:training_data_len, :]
```

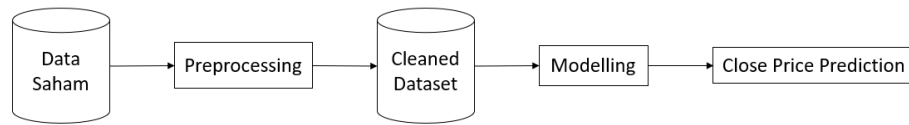
Baris ini menghitung berapa banyak baris yang membentuk 70% dari total seluruh dataset Anda. Hasilnya dibulatkan ke bilangan bulat (int), dan mengambil 70% data pertama dari data yang sudah diskalakan (*scaled_data*) untuk dijadikan sebagai data latih (*train_data*). Sisa 30%-nya nanti akan digunakan sebagai data uji (*testing data*).

2.3.3 Pengujian Data

Pengujian data merupakan komponen penting dari proses CRISP DM. Pengujian data diperlukan sebelum kita menggunakan data tersebut dalam model kita. Kita menggunakan metrik Korelasi untuk menentukan apakah hubungan antar setiap variabel merupakan hubungan linier atau saling berkorelasi. Proses ini mencakup pra-pemrosesan data—yaitu memeriksa adanya nilai yang hilang *missing value* dan melakukan standarisasi serta pengujian data sebelum digunakan untuk melatih model. Kita akan melihat distribusi data dan matrix korelasi terhadap data yang kita punya pada bab hasil dan pembahasan.

2.4 Modelling

Pada tahap modelling kita akan menggunakan *Long Short Term Memory* untuk mempelajari pergerakan harga saham dan pola yang terjadi pada data saham kita. Pada tahap ini, model LSTM digunakan untuk mempelajari dataset training data sehingga mampu memprediksi harga saham. Implementasi dilakukan untuk mencapai tujuan akhir yaitu menghasilkan prediksi harga saham. Demikianlah alur kerja yang akan diterapkan.



Gambar 5. Alur Kerja Penelitian

Sebagaimana ditunjukkan pada Gambar 5, alur kerja penelitian yang melibatkan data dimulai dengan tahap pra-pemrosesan data—yang pemeriksaan nilai yang hilang *missing value*, serta standarisasi data—hingga diperoleh dataset utama atau *cleaned dataset*. Selanjutnya, kemudian LSTM digunakan untuk melatih model dan mempelajari pola data, hingga akhirnya menghasilkan prediksi harga saham.

Rangkaian *syntax* memiliki tahapan lengkap dan terstruktur dalam membangun serta melatih model *deep learning* berbasis *Stacked LSTM* 2 lapis untuk prediksi harga saham ANTM harian, yang diawali dengan *windowed sized* ditentukan dengan nilai 30 hari, kemudian pada proses normalisasi data memiliki 2 tahap, tahap awal menggunakan fungsi `MinMaxScaler(feature_range=(0,1))` untuk mengubah seluruh rentang harga saham asli menjadi skala relatif antara 0 hingga 1, kemudian untuk tahap akhir *inverse_transform* setelah model menghasilkan luaran berupa data yang masih dalam rentang 0–1, kode menjalankan `scaler.inverse_transform(predictions)` untuk membalikkan nilainya kembali ke rentang nominal mata uang asli (Rupiah) agar metrik evaluasi (seperti RMSE dan MAE) bernilai realistis. penyesuaian dimensi data input menjadi matriks 3 dimensi [Jumlah Sampel, Timesteps, 1 Fitur] agar dapat diterima oleh lapisan LSTM. Melalui model arsitektur *Sequential*, data diproses berturut-turut melewati lapisan LSTM pertama (100 neuron) yang mengekstrak informasi temporal secara terus-menerus (`return_sequences=True`), dilanjutkan ke lapisan LSTM kedua (100 neuron) untuk mengambil kesimpulan *Sequential* akhir, hingga dipersempit oleh *Dense layer* (25 neuron) sebelum menghasilkan satu output nilai prediksi harga saham di ujungnya. Seluruh jaringan saraf berlapis ini kemudian dirakit (*compile*) menggunakan fungsi *Mean Squared Error* (MSE) dan dilatih secara masif sebanyak 150 *epochs* menggunakan penggerak *Stochastic Gradient Descent* (SGD) dengan *momentum* 0.9.

Menurut Yohannes, dkk (2025) SGD tetap menjadi pilihan yang andal untuk melatih *deep neural network*, terutama ketika keterbatasan memori menjadi perhatian, namun memerlukan penyetelan hyperparameter yang cermat dan seringkali mengalami konvergensi yang lambat [15]. Tuned batch size yang digunakan pada penelitian ini sebesar 16, Angka ini menunjukkan bahwa selama proses pelatihan (*training*), algoritma optimasi SGD akan memasukkan data sebanyak 16 sampel sekaligus ke dalam memori komputer, menghitung rata-rata *loss*, dan memperbarui bobot model setiap kali 16 sampel tersebut selesai diproses.

2.5 Evaluation

Pada bagian ini, kita akan melihat nilai prediksi dan nilai aktual pada setiap model dalam grafik, serta kinerja model berdasarkan evaluasi metrik *R-squared*, *Mean Absolute Percentage Error* dan *Root Mean Squared Error* yang dimana:

2.5.1 R-Squared(R^2)

Evaluasi R-Squared memiliki fungsi untuk mengukur seberapa baik variasi data aktual dapat dijelaskan oleh model. Nilai mendekati 1 menunjukkan model bekerja dengan sangat baik. Berikut fungsi R-Squared.

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (1)$$

2.5.2 Mean Absolute Percentage Error

Mengukur kesalahan dalam bentuk persentase. Sangat berguna untuk membandingkan performa model di beberapa saham yang rentang harganya berbeda. Nilai MAPE < 10% menunjukkan kemampuan prediksi yang sangat baik. Fungsi MAPE dapat dilihat pada fungsi berikut.

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (2)$$

2.5.3 Root Mean Squared Error

Mengukur rata-rata besarnya kesalahan prediksi. RMSE memberikan bobot lebih tinggi untuk kesalahan yang besar. Berikut bentuk fungsi dari perhitungan evaluasi RMSE.

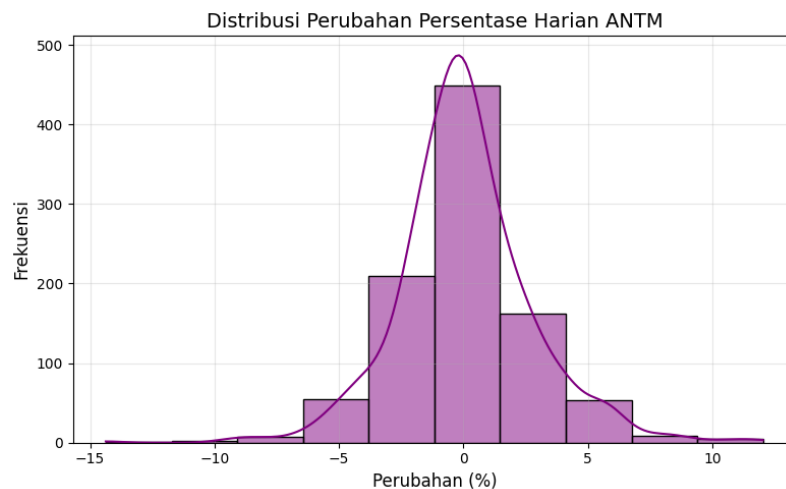
$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (3)$$

3. HASIL DAN PEMBAHASAN

Bagian ini merupakan bagian kelima dari Evaluasi CRISP-DM, bagian ini akan menunjukkan hasil analisis data dan hasil data uji menggunakan matriks korelasi. Sebagai uraian lebih lanjut mengenai hasil untuk setiap model pada bagian sebelumnya, metrik yang akan mengevaluasi prediksi harga saham adalah R-Squared, MSE (Mean Square Error) dan RMSE (Root Mean Square Error) untuk setiap data.

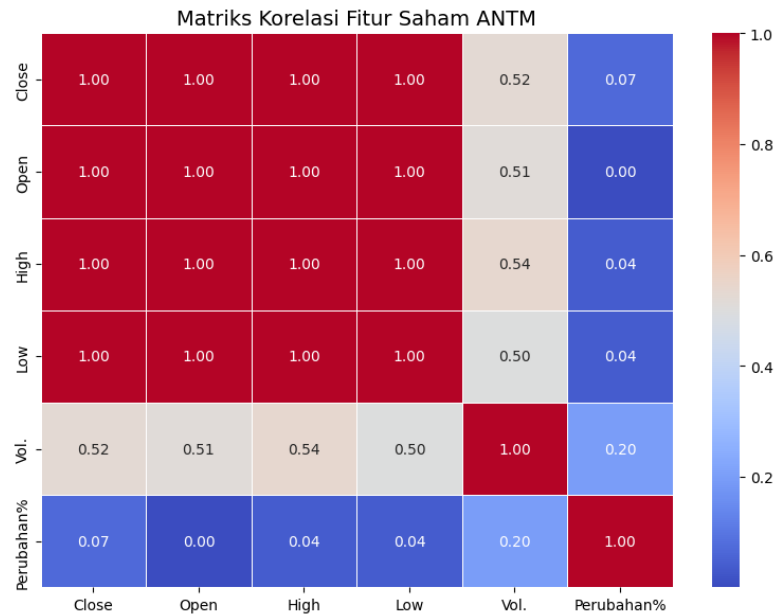
3.1. Hasil Pengujian Data

Pada tahap ini akan dijelaskan hasil dari pengujian data yang telah kita lakukan sebelumnya.



Gambar 6. Distribusi Perubahan Persentase Harian ANTM

Pada **gambar 6** Grafik ini membuktikan secara visual bahwa pergerakan harian saham ANTM mayoritas bersifat stabil dan berfluktuasi tipis di sekitar nilai rata-ratanya (dekat 0%). Kurva ini menyerupai bentuk lonceng (*bell-shaped curve*). Ini menandakan bahwa data perubahan harian saham ANTM berdistribusi normal secara wajar. Semakin ekstrem persentase perubahan harganya (baik naik sangat tinggi atau turun sangat dalam), maka frekuensi kejadiannya akan semakin jarang (batang histogram semakin pendek ke arah kanan dan kiri).

**Gambar 7.** Matrix Korelasi Antar Attribute

Pada gambar 6, Grafik ini memberikan gambaran kuantitatif yang sangat jelas mengenai bagaimana fitur-fitur (kolom data) pada saham ANTM saling memengaruhi satu sama lain. Kita bisa melihat attribute harga tutup atau close berkorelasi dengan attribute lainnya dengan rentang dari 0.07 hingga 1.00.

3.2. Hasil Performa Model

Pada bagian ini, kita akan melihat nilai prediksi dan nilai aktual pada setiap model dalam grafik, serta kinerja model berdasarkan evaluasi metrik R-squared, Mean Absolute Percentage Error, Mean Squared Error, dan Root Mean Squared Error yang ditunjukkan pada tabel 2.

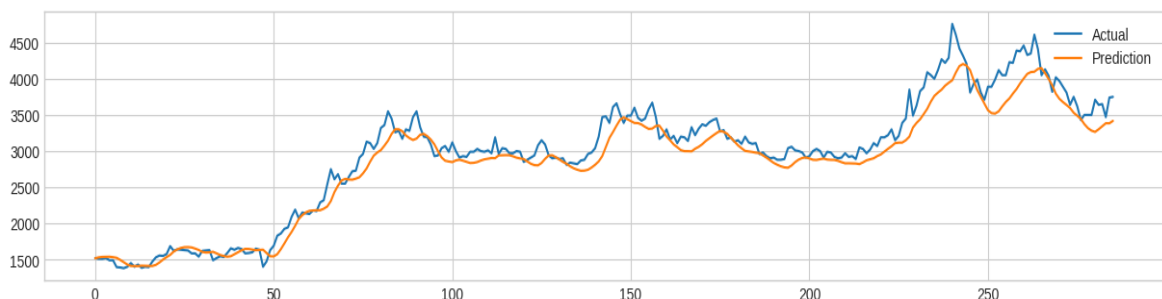
Tabel 2. Hasil Evaluasi Model

Evaluasi	Hasil
R^2	0.9368
MAPE	4.92%
RMSE	206.26

Dapat dilihat pada tabel 2 bahwa hasil dari R^2 menunjukkan bahwa model LSTM kita mampu menjelaskan 93,6% variasi pada data target, dan untuk hasil dari *Mean Absolute Percentage Error* menunjukkan di angka 2,72% dimana artinya rata rata kesalahan sangat kecil dari hasil MAPE yang kita peroleh, hasil RMSE juga menunjukkan angka yang bagus di angka 206.26 untuk data pergerakan saham.

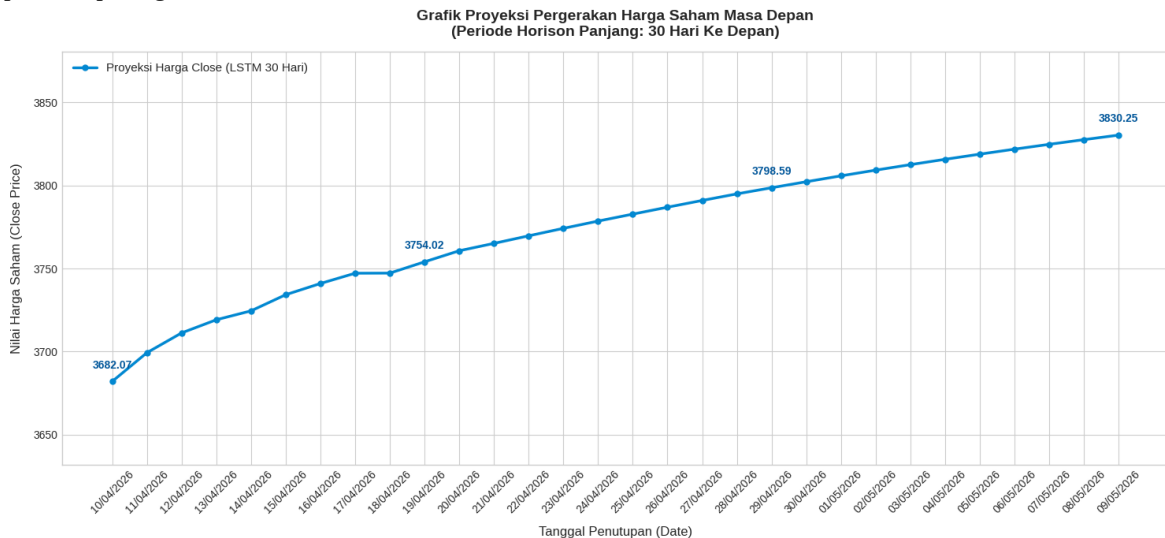
3.3. Hasil Visualisasi Prediksi

Kita bisa melihat garis antara nilai aktual dan nilai prediksi yang menggunakan model *machine learning Long Short Term Memory* hampir berjalan beriringan dan bukti dari keakuratan nilai R^2 yang kita peroleh sebelumnya.

**Gambar 8.** Hasil perbandingan pergerakan actual dan prediksi

4. KESIMPULAN

Berdasarkan pemaparan yang telah dijelaskan sebelumnya, dapat kita simpulkan bahwa *Long Short Term Memory* mampu mempelajari pola dan pergerakan harga saham dengan hasil dari evaluasi R^2 sebesar 0.92, *Mean Absolute Percentage Error* sebesar 0.027 dimana nilai ini sangat kecil dan hasil dari *Root Mean Squared Error* sebesar 128.91, angka ini dinilai sangat kecil dimana kita menggunakan parameter mata uang dalam menggunakan LSTM ini. Dan kita bisa melihat grafik perbandingan antara hasil nilai aktual dan nilai prediksi pada grafik dibawah.



Gambar 9. Hasil prediksi pergerakan saham 30 hari kedepan

Dapat dilihat pada gambar 9 bahwa hasil prediksi menunjukkan pergerakan harga tutup saham akan naik pada 30 hari kedepan walaupun tidak signifikan, dan terlihat juga garis lurus dimana harga saham tidak berubah pada hari tersebut. Ini merupakan hasil prediksi dari model LSTM yang sudah di gunakan pada penelitian ini.

REFERENSI

- [1] Bansal, M., Goyal, A., & Choudhary, A. (2022). Stock Market Prediction with High Accuracy using Machine Learning Techniques. *Procedia Computer Science*, 215, 247–265. <https://doi.org/10.1016/j.procs.2022.12.028>
- [2] Bichri, H., Chergui, A., & Hain, M. (2024). Investigating the Impact of Train / Test Split Ratio on the Performance of Pre-Trained Models with Custom Datasets. In *IJACSA) International Journal of Advanced Computer Science and Applications* (Vol. 15, Number 2). www.ijacsa.thesai.org
- [3] Brzozowska, J., Pizoń, J., Baytikenova, G., Gola, A., Zakimova, A., & Piotrowska, K. (2023). DATA ENGINEERING IN CRISP-DM PROCESS PRODUCTION DATA – CASE STUDY. *Applied Computer Science*, 19(3), 83–95. <https://doi.org/10.35784/acs-2023-26>
- [4] Chavhan, S., Raj, P., Raj, P., Dutta, A. K., & Rodrigues, J. J. P. C. (2024). Deep Learning Approaches for Stock Price Prediction: A Comparative Study of LSTM, RNN, and GRU Models. *2024 9th International Conference on Smart and Sustainable Technologies (SpliTech)*, 01–06. <https://doi.org/10.23919/SpliTech61897.2024.10612666>
- [5] Dattatreya Barve, T., Yograj Samant, A., & Suresh Kulaye, M. (2025). International Journal of Computer Technology and Electronics Communication (IJCTEC) A Comparative Study of Optimization Algorithms in Deep Learning: SGD, Adam, And Beyond. | *An ISO, 9001, 11018*. <https://doi.org/10.15680/IJCTECE.2025.0804003>
- [6] Dwivedi, S. A., Attry, A., Parekh, D., & Singla, K. (2021). Analysis and forecasting of Time-Series data using S-ARIMA, CNN and LSTM. *Proceedings - IEEE 2021 International Conference on Computing, Communication, and Intelligent Systems, ICC CIS 2021*, 131–136. <https://doi.org/10.1109/ICCCIS51004.2021.9397134>
- [7] Farhadi, A., Zamanifar, A., Alipour, A., Taheri, A., & Asadolahi, M. (2025). A Hybrid LSTM-GRU Model for Stock Price Prediction. *IEEE Access*, 13, 117594–117618. <https://doi.org/10.1109/ACCESS.2025.3586558>
- [8] Gao, Z. (2021). Stock Price Prediction with ARIMA and Deep Learning Models. *2021 IEEE 6th International Conference on Big Data Analytics, ICBDA 2021*, 61–68. <https://doi.org/10.1109/ICBDA51983.2021.9403037>

- [9] Ketut, I., Enriko, A., Nizar Gustiyana, F., Putra, R. H., & Kunci, K. (2023). *JURNAL MEDIA INFORMATIKA BUDIDARMA* Komparasi Hasil Optimasi Pada Prediksi Harga Saham PT. Telkom Indonesia Menggunakan Algoritma Long Short Term Memory. <https://doi.org/10.30865/mib.v7i2.5822>
- [10] Mittal, S., & Chauhan, A. (2021). A RNN-LSTM-Based Predictive Modelling Framework for Stock Market Prediction Using Technical Indicators. *International Journal of Rough Sets and Data Analysis*, 7(1), 1–13. <https://doi.org/10.4018/ijrdsda.288521>
- [11] Otoritas Jasa Keuangan. (2019, November 25). *PENYEBAB NAIK TURUN HARGA SAHAM SUATU PERUSAHAAN*. <https://Sikapiuangmu.Ojk.Go.Id/FrontEnd/CMS/Article/10507>. <https://sikapiuangmu.ojk.go.id/FrontEnd/CMS/Article/10507>
- [12] Pete, C., Julian, C., Randy, K., Thomas, K., Thomas, R., Colin, S., & Wirth, R. (2000). Crisp-Dm 1.0. *CRISP-DM Consortium*, 76.
- [13] Puspadini, M. (2024, January 25). *BEI Sebut Investor RI Dua Kali Lipat dari Singapura, Tapi...* CNBC Indonesia.
- [14] Yan, X., Weihang, W., & Chang, M. (2021). Research on financial assets transaction prediction model based on LSTM neural network. *Neural Computing and Applications*, 33(1), 257–270. <https://doi.org/10.1007/s00521-020-04992-7>
- [15] Yohannes, E., Febriansyah, A., Septiyanti, N. D., Suparji, Wiyono, A., Indriyanti, A. D., Utaminigrum, F., Lin, C.-Y., Muchtar, K., & Enkhbat, A. (2025). Forecasting Stock Prices with Sequential Deep Learning: A Long Short-Term Memory Approach. *2025 Eight International Conference on Vocational Education and Electrical Engineering (ICVEE)*, 177–183. <https://doi.org/10.1109/ICVEE66651.2025.11281457>