

BAB III

METODOLOGI

3.1 Objek Penelitian

Penelitian ini berfokus pada prediksi harga saham PT Alamtri Resources Indonesia Tbk (ADRO.JK), sebuah perusahaan sektor batu bara, menggunakan algoritma *Long Short-Term Memory (LSTM)*. Data historis harga saham diambil dari *Yahoo Finance*, mencakup periode 1 Januari 2020 hingga 30 Desember 2024. Penelitian dilakukan selama rentang waktu 20 September 2024 hingga 15 Januari 2025, yang mencakup tahap studi literatur, pengumpulan data, *preprocessing*, pengembangan model, dan evaluasi.

3.1.1 Gambaran Objek Penelitian

1. Fokus Penelitian

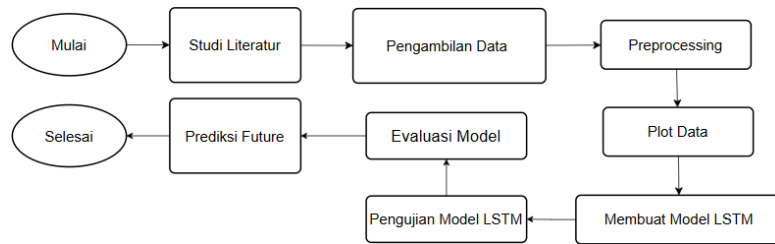
Penelitian ini berfokus pada prediksi pergerakan harga saham PT Alamtri Resources Indonesia Tbk (ADRO.JK) menggunakan algoritma *LSTM*. Algoritma ini dipilih karena kemampuannya dalam menganalisis pola data deret waktu yang *kompleks* dan *dinamis*.

2. Lingkup Batasan

1. Penelitian ini hanya mencakup saham PT Alamtri Resources Indonesia Tbk (ADRO.JK) yang bergerak di sektor tambang batu bara.
2. Data yang digunakan berasal dari *Yahoo Finance*, mencakup periode dari tanggal 1 Januari 2020 hingga 30 Desember 2024. *Variabel* utama yang dianalisis adalah harga *close*.
3. Pendekatan yang digunakan dalam penelitian ini adalah *machine learning* dengan algoritma *LSTM*, yang dirancang untuk menangani data *deret waktu*.

3.2 Prosedur Penelitian

Berikut ini adalah langkah-langkah penelitian yang dijelaskan berdasarkan *flowchart* yang ditampilkan:



Gambar 3. 1 *Flowchart* prosedur penelitian

3.2.1 Studi Literatur

Pada tahap ini, dilakukan pengumpulan *referensi* yang relevan dengan topik penelitian. Literatur yang digunakan mencakup buku, jurnal ilmiah, artikel, dan laporan terkait *machine learning*, algoritma *LSTM*, serta analisis data saham. Studi ini bertujuan untuk memperoleh landasan teori yang kuat dalam mendukung pengembangan model prediksi.

3.2.2 Pengambilan Data dan Analisis Data

Pengambilan Data dan Analisis Data Data harga saham ADRO diperoleh dari *Yahoo Finance*, yang menyediakan data historis harga saham dalam bentuk harian. Data mencakup periode dari 1 Januari 2020 hingga 30 Desember 2024. Analisis awal dilakukan untuk memahami karakteristik data, seperti tren, musiman, dan *volatilitas*. Berikut adalah gambar 3.2 yang menunjukan data.

Jakarta - Delayed Quote • IDR

PT Alamtri Resources Indonesia Tbk (ADRO.JK) ★ Following Compare

2,400.00 +40.00 +(1.69%)
As of January 15 at 4:14:57 PM GMT+7. Market Open.

Jan 01, 2020 - Dec 31, 2024 Historical Prices Daily

Currency in IDR

Date	Open	High	Low	Close	Adj Close	Volume
Dec 30, 2024	2,440.00	2,470.00	2,410.00	2,430.00	2,430.00	75,087,400
Dec 27, 2024	2,530.00	2,540.00	2,490.00	2,540.00	2,540.00	64,253,200
Dec 24, 2024	2,540.00	2,610.00	2,510.00	2,510.00	2,510.00	70,990,000
Dec 23, 2024	2,580.00	2,590.00	2,520.00	2,540.00	2,540.00	49,626,700
Dec 20, 2024	2,540.00	2,580.00	2,510.00	2,550.00	2,550.00	65,183,500
Dec 19, 2024	2,570.00	2,580.00	2,520.00	2,540.00	2,540.00	57,029,200
Dec 18, 2024	2,590.00	2,630.00	2,540.00	2,590.00	2,590.00	91,709,300
Dec 17, 2024	2,650.00	2,660.00	2,510.00	2,510.00	2,510.00	94,626,300
Dec 16, 2024	2,580.00	2,660.00	2,480.00	2,600.00	2,600.00	137,736,700
Dec 13, 2024	2,550.00	2,630.00	2,540.00	2,580.00	2,580.00	110,671,800
Dec 12, 2024	2,610.00	2,660.00	2,510.00	2,530.00	2,530.00	120,595,500

Gambar 3. 2 *Data historical*

3.2.3 Pra-Processing Data

Pada tahap ini, data yang diperoleh dibersihkan untuk memastikan kualitasnya sebelum digunakan dalam model prediksi. Proses *preprocessing* dilakukan melalui beberapa tahap berikut:

1. *Normalisasi Data:*

1. Data harga saham dinormalisasi menggunakan *Min-Max Scaling* untuk mengubah nilai ke dalam rentang $[0,1]$. Hal ini dilakukan untuk memastikan bahwa skala data seragam sehingga model *LSTM* dapat lebih mudah mengenali pola.
2. Hasil *normalisasi* menunjukkan bahwa nilai harga saham dalam dataset telah terkonversi ke dalam skala yang lebih kecil dan seragam, yang siap digunakan dalam pemodelan.

2. *Transformasi Data ke Format Time Series:*

1. Data dikonversi menjadi *format time series* menggunakan pendekatan *sliding window*, di mana setiap 30 hari terakhir digunakan sebagai *input* untuk memprediksi harga saham pada hari ke-31.
2. Proses ini memungkinkan model untuk memahami pola jangka pendek dan panjang dalam data saham ADRO, sehingga prediksi yang dihasilkan lebih akurat.

3. *Pembagian Data (Train-Test Split):*

1. Dataset dibagi menjadi 80% untuk pelatihan dan 20% untuk pengujian guna mengukur performa model secara objektif.
2. Data pelatihan digunakan untuk melatih model dalam mengenali pola historis, sedangkan data pengujian digunakan untuk mengevaluasi akurasi prediksi.

4. *Hasil Preprocessing (Shape Data):*

Setelah dilakukan *preprocessing*, dataset memiliki bentuk sebagai berikut:

Tabel 3. 1 Tabel pembagian dataset

Dataset	Jumlah Data	Jadwal
Data latih (<i>Training Set</i>)	935 sampel	2020-01-02 00:00:00 sampai 2023-12-14 00:00:00
Data uji (<i>Testing Set</i>)	212 sampel	2023-12-15 00:00:00 sampai 2024-12-27 00:00:00

3.2.4 Plot Data

Setelah data selesai diproses, langkah berikutnya adalah memvisualisasikan data dalam bentuk *grafik*. *Visualisasi* ini bertujuan untuk:

1. Mengidentifikasi pola *tren*, musiman, atau *volatilitas* dalam data harga saham.
2. Membantu dalam memahami distribusi dan *outlier* pada data.

Plot data dilakukan menggunakan *library Python* seperti *Matplotlib* atau *Seaborn*. *Grafik* yang dihasilkan mencakup *grafik* harga penutupan harian dan *moving average* untuk menunjukkan *tren* jangka pendek dan jangka panjang.

Rumus *Moving Average*:

$$MAE = \frac{1}{N} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (3.1)$$

Dimana:

1. MA_t : *Moving Average* pada waktu t .
2. P_{t-i} : Harga penutupan pada hari ke- i .
3. n : Periode *moving average*.

3.2.5 Membuat Model *Forecasting* Menggunakan *LSTM*

Pada tahap ini, model *Long Short-Term Memory (LSTM)* dirancang untuk memprediksi harga saham. *LSTM* adalah algoritma berbasis *Recurrent Neural Network (RNN)* yang unggul dalam memproses data deret waktu (*time series*).

Langkah Utama:

3. Pemilihan *Parameter Model*:

1. Jumlah *neuron* pada *hidden layer*.
2. *Epoch* (jumlah iterasi pelatihan).
3. *Batch size* (jumlah data yang diproses dalam satu iterasi).
4. Fungsi aktivasi seperti *ReLU (Rectified Linear Unit)* atau *sigmoid*.

2. *Arsitektur Model*:

Model terdiri dari beberapa lapisan *LSTM* yang diikuti oleh *dense layer* untuk memprediksi nilai akhir. Fungsi *loss* yang digunakan adalah *Mean Squared Error (MSE)*.

Rumus *MSE*:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (3.2)$$

Dimana:

- y_i : Nilai aktual.
 $\hat{y}_i$: Nilai prediksi.
 n : Jumlah sampel data.

3.2.6 Pengujian Model *LSTM*

Setelah proses pelatihan model *LSTM* selesai, langkah selanjut nya adalah melakukan pengujian terhadap model tersebut. Pengujian ini bertujuan untuk memecahkan seberapa baik model mampu memprediksi harga penutupan saham PT Alamtri Resources Indonesia Tbk (ADRO.JK) berdasarkan data yang belum pernah digunakan selama pelatihan. berikut *Pseudocode Algoritma LSTM* :

1. *Inisialisasi Model LSTM*

1. Membangun arsitektur model *LSTM* dengan beberapa lapisan.

2. *Menyempurnakan Lapisan LSTM*

1. Menentukan jumlah unit *LSTM* pada setiap lapisan untuk menangkap pola data historis.
2. Menggunakan beberapa lapisan *LSTM* untuk meningkatkan kemampuan model dalam memahami *tren* jangka panjang dan pendek.

3. Membakar Lapisan *Dropout*

1. Mengurangi risiko overfitting dengan menerapkan *dropout* pada lapisan *LSTM*.
2. *Dropout* membantu dalam menonaktifkan sejumlah neuron secara acak selama pelatihan sehingga model tidak terlalu bergantung pada pola tertentu dalam data latih.

4. Menambahkan Lapisan *Dense*

1. Menambahkan lapisan *fullyconnected* (*Dense*) sebagai lapisan output yang menghasilkan prediksi harga saham berdasarkan pola yang telah dipelajari.
2. Lapisan ini akan menerima *output* dari lapisan *LSTM* terakhir dan menghasilkan satu nilai *numerik* sebagai harga prediksi.

5. Kompilasi Model

1. Menggunakan pengoptimal Adam untuk meningkatkan kecepatan dan stabilitas model *konvergensi*.
2. Menggunakan fungsi *loss Mean Squared Error (MSE)* sebagai metrik utama untuk menghitung perbedaan antara nilai aktual dan nilai prediksi.

3.2.7 Evaluasi Model

Model evaluasi dilakukan untuk mengukur sejauh mana model *LSTM* mampu menghasilkan prediksi yang akurat. Proses evaluasi dilakukan dengan membandingkan hasil prediksi dengan data aktual, baik pada data uji maupun data latih.

1. Evaluasi pada Uji Data

Setelah model diuji, dilakukan prediksi menggunakan data uji. Hasil prediksi dibandingkan dengan harga aktual menggunakan *metrik* evaluasi berikut:

1. *Mean Absolute Error (MAE)*

Mengukur rata-rata kesalahan absolut antara prediksi dan data aktual.

2. *Mean Squared Error (MSE)*

Menghitung selisih kuadrat rata-rata antara prediksi dan data aktual.

3. *Root Mean Squared Error (RMSE)*

Mengukur akar dari *MSE* untuk menilai tingkat error dalam satuan asli.

4. *Mean Absolute Percentage Error (MAPE)*

Mengukur kesalahan dalam bentuk persentase dari nilai aktual.

Selain evaluasi menggunakan metrik di atas, dilakukan visualisasi hasil prediksi dengan membandingkan harga aktual dan hasil prediksi model dalam bentuk *grafik*. Hal ini bertujuan untuk melihat sejauh mana model dapat mengikuti tren pergerakan harga saham.

