

## Perbandingan Algoritma XGBoost dan LSTM dalam Prediksi Harga Saham Tesla Menggunakan Data Tahun 2025

Ahmad Faqih Alkayes<sup>\*1</sup>, Tri Sugihartono<sup>2</sup>

<sup>1,2</sup>Program Studi Teknik Informasi, Fakultas Teknologi Informasi, ISB Atma Luhur, Indonesia  
Email: <sup>1</sup>[2111500114@mahasiswa.atmaluhur.ac.id](mailto:2111500114@mahasiswa.atmaluhur.ac.id) <sup>2</sup>[trisugihartono@atmaluhur.ac.id](mailto:trisugihartono@atmaluhur.ac.id)

### Abstrak

Prediksi harga saham merupakan tantangan dalam analisis keuangan yang memerlukan metode yang akurat dan andal. Penelitian ini membandingkan kinerja algoritma XGBoost dan *Long Short-Term Memory* (LSTM) dalam memprediksi harga saham Tesla menggunakan data tahun 2025. XGBoost, sebagai model berbasis pohon keputusan yang dioptimalkan, dikenal dengan kecepatan dan interpretabilitasnya, sedangkan LSTM, sebagai bagian dari jaringan saraf tiruan, memiliki kemampuan menangkap pola temporal yang kompleks dalam data time series. Ketidakstabilan pasar saham menjadikan prediksi harga saham sebagai aspek krusial dalam pengambilan keputusan investasi, khususnya bagi perusahaan-perusahaan teknologi seperti Tesla yang pergerakan harganya sangat fluktuatif. Oleh karena itu, penelitian ini bertujuan untuk membandingkan efektivitas dua pendekatan prediksi yang populer—XGBoost dan *Long Short-Term Memory* (LSTM)—dalam memprediksi harga saham Tesla menggunakan data historis selama satu dekade terakhir. XGBoost, algoritma berbasis pohon keputusan, unggul dalam kecepatan komputasi dan interpretabilitas, sementara LSTM, jaringan saraf berulang, dirancang untuk mengenali pola temporal jangka panjang dalam data deret waktu. Penelitian ini menerapkan *preprocessing* termasuk normalisasi, *splitting time series*, dan pembentukan *sliding window* untuk LSTM, serta mengevaluasi performa kedua model menggunakan metrik RMSE, MAE, dan koefisien determinasi ( $R^2$ ). Hasil menunjukkan bahwa LSTM secara konsisten menghasilkan prediksi yang lebih akurat dalam menangkap tren dan fluktuasi harga saham, sedangkan XGBoost lebih efisien dalam hal waktu pelatihan dan kemudahan interpretasi hasil. Temuan ini memberikan kontribusi terhadap literatur prediksi finansial dengan menunjukkan bahwa pemilihan model harus disesuaikan dengan tujuan analisis—baik untuk presisi jangka panjang maupun efisiensi operasional—serta membuka peluang pengembangan hibridisasi model untuk pengambilan keputusan investasi yang lebih cerdas.

**Kata kunci:** prediksi harga saham, XGBoost, LSTM, tesla, machine learning.

## Comparison of XGBoost and LSTM Algorithms in Tesla Stock Price Prediction Using 2025 Data

### Abstract

Stock price prediction is a significant challenge in financial analysis that requires accurate and reliable methods. This study compares the performance of two widely used algorithms—XGBoost and Long Short-Term Memory (LSTM)—in predicting Tesla's stock prices using historical data from the past decade. XGBoost, an optimized decision tree-based model, is well known for its computational speed and interpretability, whereas LSTM, a type of recurrent neural network, excels at capturing complex temporal patterns in time series data. The inherent volatility of the stock market makes accurate forecasting crucial for informed investment decisions, especially for technology companies like Tesla, whose stock prices are highly fluctuating. This research involves a series of preprocessing steps including data normalization, time series splitting, and the construction of sliding windows for LSTM, to ensure the models receive structured and well-prepared input. The performance of both models is evaluated using Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and the coefficient of determination ( $R^2$ ). The results show that LSTM consistently delivers more accurate predictions by effectively capturing long-term trends and short-term fluctuations in stock prices, while XGBoost remains superior in terms of training efficiency and interpretability. These findings contribute to the growing body of financial forecasting literature by demonstrating that model selection should align with the intended analytical goals—favoring LSTM for long-term pattern recognition and XGBoost for rapid, interpretable analysis—and highlight the potential for developing hybrid models to support more intelligent investment decision-making.

**Keywords:** stock price prediction, XGBoost, LSTM, tesla, machine learning.

## 1. PENDAHULUAN

Perkembangan pesat dalam dunia keuangan dan teknologi telah mendorong penggunaan kecerdasan buatan dan pembelajaran mesin untuk meningkatkan akurasi prediksi harga saham. Salah satu tantangan utama dalam analisis pasar saham adalah volatilitas tinggi yang menyebabkan fluktuasi harga yang sulit diprediksi dengan metode konvensional. Oleh karena itu, penerapan model pembelajaran mesin seperti Extreme Gradient Boosting (XGBoost) dan Long Short-Term Memory (LSTM) menjadi semakin populer dalam analisis data keuangan. XGBoost, sebagai algoritma berbasis pohon keputusan yang dioptimalkan, dikenal karena kemampuannya dalam menangani dataset besar dan mengurangi overfitting, sementara LSTM, sebagai jenis jaringan saraf berulang (RNN), unggul dalam memahami pola temporal dari data historis [1]. Penelitian ini bertujuan untuk membandingkan kinerja kedua algoritma tersebut dalam memprediksi harga saham Tesla berdasarkan dataset tahun 2025. Hasil dari penelitian ini diharapkan dapat memberikan wawasan tentang efektivitas masing-masing metode dalam konteks prediksi saham, serta menjadi referensi bagi investor dan analis keuangan dalam memilih pendekatan yang lebih optimal untuk strategi perdagangan saham [2].

Perbandingan Algoritma XGBoost dan LSTM dalam Prediksi Harga Saham Menggunakan Data Tahun 2025

Tujuan dari penelitian ini adalah untuk membandingkan kinerja algoritma XGBoost dan Long Short-Term Memory (LSTM) dalam memprediksi harga saham Tesla berdasarkan dataset tahun 2025. Penelitian ini bertujuan untuk mengevaluasi efektivitas kedua algoritma dalam menangani data deret waktu serta menentukan metode yang lebih akurat dalam memprediksi pergerakan harga saham [3]. Dengan memanfaatkan teknik machine learning berbasis pohon keputusan seperti XGBoost dan pendekatan deep learning berbasis jaringan saraf seperti LSTM, penelitian ini akan menganalisis keunggulan dan kelemahan masing-masing model berdasarkan metrik evaluasi seperti Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), dan R-squared ( $R^2$ ). Hasil penelitian ini diharapkan dapat memberikan wawasan bagi investor, analis keuangan, serta akademisi dalam memilih metode terbaik untuk prediksi harga saham, sehingga dapat mendukung pengambilan keputusan investasi yang lebih optimal [4].

Ruang lingkup penelitian ini mencakup analisis dan perbandingan performa algoritma XGBoost dan Long Short-Term Memory (LSTM) dalam memprediksi harga saham Tesla menggunakan dataset tahun 2025.[5] Penelitian ini berfokus pada pengolahan data deret waktu saham, termasuk proses pengumpulan, pembersihan, dan normalisasi data sebelum diterapkan pada model pembelajaran mesin. Selain itu, penelitian ini akan mencakup eksperimen dengan berbagai parameter dan teknik optimasi pada kedua algoritma guna memperoleh hasil prediksi yang optimal [6]. Evaluasi kinerja dilakukan menggunakan metrik seperti Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), dan R-squared ( $R^2$ ) untuk mengukur akurasi serta efektivitas masing-masing model dalam memprediksi pergerakan harga saham. Ruang lingkup penelitian ini terbatas pada data saham Tesla dalam kurun waktu satu tahun (2025) dan tidak mencakup faktor eksternal seperti kondisi ekonomi makro, kebijakan pasar modal, atau sentimen pasar yang dapat mempengaruhi harga saham.[7] Hasil dari penelitian ini diharapkan dapat memberikan wawasan mengenai keunggulan dan kelemahan kedua algoritma dalam konteks prediksi saham serta membantu investor dan analis dalam pengambilan keputusan berbasis data [8].

## 2. METODE PENELITIAN

Metode penelitian ini menggunakan pendekatan eksperimen kuantitatif untuk membandingkan performa algoritma XGBoost dan Long Short-Term Memory (LSTM) dalam memprediksi harga saham Tesla berdasarkan dataset tahun 2025 [9]. Proses penelitian diawali dengan pengumpulan data historis harga saham Tesla, yang mencakup harga pembukaan, harga penutupan, volume perdagangan, serta indikator teknikal lainnya.[10] Data yang diperoleh kemudian dibersihkan dan dinormalisasi untuk memastikan kualitas serta mengurangi noise yang dapat mempengaruhi hasil prediksi [11]. Selanjutnya, data akan dibagi menjadi training set dan testing set dengan rasio yang sesuai untuk melatih dan menguji model. Model XGBoost dan LSTM akan diterapkan dengan berbagai konfigurasi parameter untuk menemukan kombinasi terbaik yang menghasilkan prediksi paling akurat. Evaluasi dilakukan dengan menggunakan metrik Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), dan R-squared ( $R^2$ ) untuk menilai performa masing-masing algoritma. Hasil yang diperoleh akan dianalisis secara komparatif untuk menentukan model yang lebih unggul dalam menangani data saham Tesla. Penelitian ini menggunakan Python dan pustaka machine learning seperti Scikit-learn, TensorFlow, dan XGBoost untuk implementasi model serta visualisasi hasil [12].



. Gambar 1. Metode Penelitian

### 2.1. Pengumpulan Data

Dalam penelitian ini dilakukan dengan mengambil data historis harga saham Tesla untuk tahun 2025 dari sumber terpercaya seperti Yahoo Finance, Google Finance, atau Nasdaq. Data yang dikumpulkan mencakup berbagai variabel penting dalam analisis pasar saham, termasuk harga pembukaan (open), harga penutupan (close), harga tertinggi (high), harga terendah (low), dan volume perdagangan [13]. Selain itu, penelitian ini juga akan mempertimbangkan berbagai indikator teknikal yang relevan, seperti moving average (MA), relative strength index (RSI), dan Bollinger Bands, yang sering digunakan dalam analisis pergerakan harga saham. Data yang diperoleh akan diolah dan disiapkan untuk tahap preprocessing, guna memastikan kualitas dan akurasi sebelum diterapkan pada model XGBoost dan LSTM dalam proses prediksi harga saham Tesla [14].

### 2.2. Preprocessing Data

Penelitian ini dilakukan serangkaian tahapan preprocessing untuk memastikan bahwa data yang digunakan bersih, terstruktur, dan sesuai dengan kebutuhan algoritma XGBoost dan LSTM. Langkah pertama adalah menghilangkan data yang tidak lengkap atau mengandung nilai kosong (missing values), karena data yang hilang dapat menyebabkan bias atau kesalahan dalam prediksi [15]. Selanjutnya, dilakukan normalisasi atau standarisasi data agar seluruh variabel memiliki skala yang seragam, sehingga dapat meningkatkan performa model machine learning. Normalisasi ini sangat penting terutama untuk algoritma LSTM, yang sensitif terhadap perbedaan skala data dalam pemrosesan deret waktu. Selain itu, data diubah ke dalam format yang sesuai dengan karakteristik masing-masing algoritma, khususnya dalam hal struktur deret waktu, di mana data akan dibentuk dalam bentuk windowed sequences agar LSTM dapat memahami pola tren harga saham. Untuk proses pelatihan dan pengujian model, data dibagi menggunakan rasio 80:20 antara data latih (training set) dan data uji (testing set), guna memastikan evaluasi performa model dilakukan pada data yang belum pernah dilihat sebelumnya. Khusus untuk model XGBoost, validasi silang (cross-validation) juga dilakukan menggunakan skema 5-fold untuk mengevaluasi stabilitas dan generalisasi model secara lebih menyeluruh. Dengan preprocessing yang optimal dan strategi evaluasi yang tepat, data menjadi lebih siap untuk diterapkan dalam model XGBoost dan LSTM guna memperoleh hasil prediksi yang lebih akurat dan reliable [16].

### 2.3. Penerapan Algoritma

Penelitian ini akan menggunakan XGBoost dan Long Short-Term Memory (LSTM) untuk memprediksi harga saham Tesla berdasarkan dataset tahun 2025. XGBoost akan dilatih sebagai model berbasis pohon keputusan yang dioptimalkan menggunakan teknik boosting, yang memungkinkan model untuk memperbaiki kesalahan prediksi dari iterasi sebelumnya, sehingga meningkatkan akurasi dan ketahanan terhadap overfitting.

Sementara itu, LSTM akan diterapkan sebagai model jaringan saraf dalam (deep learning) yang dirancang khusus untuk menangani data deret waktu, dengan kemampuan mempertimbangkan hubungan temporal antar data agar dapat mengenali pola dalam pergerakan harga saham secara lebih efektif. Untuk meningkatkan performa kedua model, penelitian ini juga akan menggunakan teknik hyperparameter tuning, seperti Grid Search dan Bayesian Optimization, guna menemukan kombinasi parameter terbaik yang dapat memaksimalkan akurasi prediksi [17]. Dengan pendekatan ini, diharapkan penelitian dapat menghasilkan perbandingan yang objektif mengenai efektivitas kedua algoritma dalam memprediksi harga saham Tesla.

a. Rumus XGBoost

$$\gamma i \sum_{k=1}^k \mathcal{F}k(\mathbf{x}i) \quad (1)$$

Dengan  $\gamma i$  sebagai prediksi untuk data ke- $i$ , dan  $\mathcal{F}k(\mathbf{x}i)$  adalah pohon keputusan ke- $k$  dalam model.

b. Rumus LSTM

$$F_t = \sigma(W_f \cdot [h_t - 1, x_t] + b_f) \quad (2)$$

Dimana  $F_t$  adalah vektor forget gate,  $W_f$  adalah bobot,  $h_t - 1$  adalah hidden state sebelumnya,  $x_t$  adalah input saat ini,  $b_f$  adalah bias, dan  $\sigma$  adalah fungsi aktivasi sigmoid.

#### 2.4. Evaluasi Model

Dalam penelitian ini bertujuan untuk menilai performa algoritma XGBoost dan LSTM dalam memprediksi harga saham Tesla berdasarkan dataset tahun 2025. Untuk mengukur tingkat akurasi dan kesalahan prediksi, penelitian ini akan menggunakan beberapa metrik evaluasi utama. Mean Absolute Error (MAE) digunakan untuk menghitung rata-rata kesalahan absolut antara nilai prediksi dan nilai aktual, sehingga memberikan gambaran mengenai tingkat kesalahan dalam satuan harga saham. Root Mean Squared Error (RMSE) akan digunakan untuk mengevaluasi tingkat kesalahan dengan memberi bobot lebih besar pada kesalahan besar, yang penting dalam menganalisis model prediksi keuangan. Selain itu, R-squared ( $R^2$ ) akan digunakan untuk mengukur sejauh mana model dapat menjelaskan variabilitas harga saham, di mana nilai  $R^2$  yang lebih tinggi menunjukkan model yang lebih baik dalam menangkap pola tren harga saham. Setelah mendapatkan hasil dari masing-masing metrik, penelitian ini akan membandingkan performa XGBoost dan LSTM berdasarkan akurasi prediksi yang dihasilkan, guna menentukan model yang lebih unggul dalam memprediksi harga saham Tesla. Analisis ini diharapkan dapat memberikan wawasan mendalam mengenai kelebihan dan kekurangan kedua algoritma dalam konteks prediksi pasar saham.

a. Mean Absolute Error

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

Dimana  $y_i$  adalah nilai aktual,  $\hat{y}_i$  adalah nilai prediksi, dan  $n$  adalah jumlah data. MAE memberikan gambaran langsung tentang rata-rata kesalahan dalam satuan harga saham.

b. Root Mean Squared Error

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (4)$$

MSE lebih sensitif terhadap kesalahan besar dibandingkan MAE, sehingga cocok untuk mengidentifikasi model yang cenderung melakukan kesalahan besar dalam prediksi.

c.  $R^2$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \bar{y}_i)^2}{\sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (5)$$

Dimana  $\bar{y}_i$  adalah rata-rata nilai aktual. Nilai  $R^2$  berkisar antara 0 hingga 1, dengan nilai yang lebih tinggi menunjukkan bahwa model lebih baik dalam menangkap pola data.

## 2.5. Analisis dan Interpretasi Hasil

Penelitian ini akan menyajikan hasil evaluasi performa algoritma XGBoost dan LSTM dalam bentuk tabel dan visualisasi grafis, seperti plot perbandingan nilai prediksi dan aktual, error distribution, serta kurva loss function untuk memahami pola kesalahan prediksi. Dengan bantuan visualisasi ini, akan lebih mudah untuk mengidentifikasi bagaimana masing-masing model menangani pergerakan harga saham Tesla serta sejauh mana akurasi prediksinya. Selain itu, penelitian ini akan menganalisis berbagai faktor yang mempengaruhi akurasi model, seperti kualitas data, pemilihan fitur, parameter model, serta ketepatan dalam menangani pola fluktuasi harga saham. Berdasarkan hasil evaluasi, penelitian ini akan menarik kesimpulan mengenai algoritma mana yang lebih unggul dalam memprediksi harga saham Tesla tahun 2025, dengan mempertimbangkan kelebihan dan keterbatasan masing-masing model. Dengan demikian, penelitian ini diharapkan dapat memberikan wawasan yang bermanfaat bagi investor, trader, dan akademisi dalam memilih model prediksi saham yang paling efektif [18].

## 2.6. Kesimpulan dan Rekomendasi

Penelitian ini akan merangkum hasil perbandingan performa XGBoost dan LSTM dalam memprediksi harga saham Tesla berdasarkan dataset tahun 2025. Kesimpulan akan didasarkan pada analisis metrik evaluasi, seperti MAE, RMSE, dan  $R^2$ , untuk menentukan algoritma yang lebih akurat dan efisien dalam menangani pola pergerakan harga saham. Berdasarkan temuan ini, penelitian akan memberikan rekomendasi mengenai pemanfaatan model prediksi saham bagi investor, trader, dan peneliti di bidang keuangan, guna membantu mereka dalam mengambil keputusan investasi yang lebih berbasis data. Selain itu, penelitian ini akan menyusun saran untuk penelitian lanjutan, seperti eksplorasi fitur tambahan (misalnya indikator makroekonomi dan sentimen pasar) atau penerapan pendekatan hybrid yang mengombinasikan keunggulan XGBoost dan LSTM untuk meningkatkan akurasi prediksi. Dengan rekomendasi ini, diharapkan penelitian selanjutnya dapat lebih mengembangkan metode prediksi saham yang lebih presisi dan adaptif terhadap dinamika pasar keuangan.

Penelitian terkait prediksi harga saham dan penerapan algoritma kecerdasan buatan telah banyak dilakukan dalam berbagai bidang. Hendrawan (2022) membandingkan algoritma Naïve Bayes, SVM, dan XGBoost dalam klasifikasi teks sentimen masyarakat terhadap produk lokal di Indonesia, yang menunjukkan bahwa XGBoost unggul dalam akurasi dibandingkan dengan metode lainnya dalam pengolahan data teks [19]. Permata dan Nababan (2023) menerapkan teori permainan (game theory) untuk menentukan strategi pemasaran yang optimal dalam marketplace, yang menunjukkan bahwa pendekatan matematis dapat membantu perusahaan dalam pengambilan keputusan strategis berbasis data [20]. Dalam penelitian yang lebih spesifik terkait prediksi saham, Abdupatah dan Rozikin (2025) menggunakan metode Long Short-Term Memory (LSTM) untuk memprediksi harga saham Tesla, menyoroti keunggulan LSTM dalam menangkap pola jangka panjang dalam data time series [21]. Sementara itu, Ningrum dan Ismawardi (2025) mengevaluasi efektivitas algoritma kecerdasan buatan dalam implementasi kesehatan mental melalui tinjauan sistematis literatur, yang menyoroti bagaimana AI dapat dimanfaatkan dalam berbagai aspek kehidupan, termasuk sektor kesehatan [22]. Dengan mengacu pada penelitian-penelitian tersebut, studi ini berkontribusi dengan membandingkan XGBoost dan LSTM dalam prediksi harga saham Tesla berdasarkan dataset tahun 2025, memberikan wawasan mengenai efektivitas kedua algoritma dalam analisis pasar saham serta optimalisasi model melalui teknik hyperparameter tuning.

## 2.7. Arsitektur LSTM

Arsitektur model LSTM yang digunakan terdiri dari dua lapisan LSTM berturut-turut yang masing-masing memiliki 50 unit neuron. Lapisan pertama diatur dengan `return_sequences=True` agar dapat mengembalikan urutan lengkap ke lapisan berikutnya, sedangkan lapisan kedua diatur untuk hanya mengembalikan output akhir. Setelah kedua lapisan LSTM, ditambahkan satu lapisan Dense dengan 25 neuron sebagai lapisan intermediate, dan satu lapisan Dense akhir dengan 1 neuron untuk menghasilkan output prediksi harga saham. LSTM menggunakan fungsi aktivasi tanh untuk memproses input dan sigmoid untuk fungsi gating internal, sementara lapisan Dense menggunakan aktivasi linear, yang sesuai untuk tugas regresi. Model dioptimasi menggunakan algoritma Adam dan menggunakan fungsi loss Mean Squared Error (MSE). Di sisi lain, model XGBoost dikonfigurasi dengan parameter utama yaitu `learning_rate` sebesar 0.1 untuk mengontrol laju pembelajaran setiap pohon, `max_depth` sebesar 3 untuk membatasi kompleksitas struktur pohon agar menghindari overfitting, serta `n_estimators` sebanyak 100 yang menentukan jumlah total pohon yang dibangun dalam proses boosting. Fungsi

objektif yang digunakan adalah 'reg:squarederror' karena tugas yang dihadapi adalah regresi kontinu terhadap harga saham. Kombinasi arsitektur dan parameter ini dirancang untuk membandingkan secara optimal kemampuan model LSTM dalam memahami pola sekuensial jangka panjang dan kekuatan XGBoost dalam efisiensi dan interpretabilitas prediksi.

### 3. HASIL DAN PEMBAHASAN

#### 3.1. Table Isi Dataset

Tabel 1. Table Dataset

Date	Open	High	Low	Close	Adj Close	Volume
2010-06-29	1.266667	1.666667	1.169333	1.592667	1.592667	281494500
2010-06-30	1.719333	2.028000	1.553333	1.588667	1.588667	257806500
2010-07-01	1.666667	1.728000	1.351333	1.464000	1.464000	123282000
2010-07-02	1.533333	1.540000	1.247333	1.280000	1.280000	77097000
2010-07-06	1.333333	1.333333	1.055333	1.074000	1.074000	103003500

Tabel Dataset menampilkan data historis harga saham yang mencakup beberapa kolom utama: Date (tanggal transaksi), Open (harga pembukaan), High (harga tertinggi), Low (harga terendah), Close (harga penutupan), Adj Close (harga penutupan yang disesuaikan), dan Volume (jumlah saham yang diperdagangkan). Data ini digunakan dalam analisis pasar saham untuk mengamati tren harga, volatilitas, dan likuiditas suatu aset dari waktu ke waktu. Harga pembukaan dan penutupan mencerminkan pergerakan harian, sementara harga tertinggi dan terendah menunjukkan kisaran fluktuasi dalam satu hari perdagangan. Adj Close mempertimbangkan penyesuaian seperti dividen atau pemecahan saham agar lebih akurat dalam analisis historis. Volume perdagangan mengindikasikan seberapa aktif saham diperjualbelikan, yang dapat menjadi indikator sentimen pasar dan potensi perubahan tren harga.

#### 3.2. Exploratory Data Analysis



Gambar 2. Grafik Exploratory Data Analysis

Grafik Histograms of Numerical Features menyajikan distribusi variabel numerik dalam dataset melalui histogram, yang berguna untuk eksplorasi data awal (Exploratory Data Analysis/EDA). Setiap subplot menampilkan histogram dari berbagai fitur numerik, memungkinkan analisis pola distribusi, apakah data bersifat normal, miring (skewed), atau memiliki pencilan (outliers). Dengan memahami distribusi data ini, kita dapat mengidentifikasi karakteristik penting seperti penyebaran, konsentrasi nilai, dan kemungkinan adanya data ekstrem yang dapat memengaruhi pemodelan atau analisis lebih lanjut. Visualisasi ini membantu dalam menentukan transformasi data yang sesuai serta memastikan kualitas data sebelum masuk ke tahap analisis statistik atau machine learning.

### 3.3. Time Series Analysis



Gambar 3. Grafik Time Series Analysis

Grafik Close over Time menyajikan analisis deret waktu terhadap harga penutupan (Close) dan harga penutupan yang telah disesuaikan (Adj Close), serta transformasi logaritmiknya (log\_Close dan log\_Adj Close). Dengan menggunakan subplot, setiap grafik memperlihatkan pola pergerakan harga dari waktu ke waktu, baik dalam skala linear maupun logaritmik. Transformasi logaritmik diterapkan untuk menstabilkan variabilitas data dan mengatasi perbedaan skala, sehingga membantu dalam analisis tren jangka panjang serta mengidentifikasi perubahan harga yang lebih halus. Visualisasi ini sangat berguna dalam memahami karakteristik data, mendeteksi pola musiman atau tren, serta mengevaluasi volatilitas harga untuk keperluan analisis investasi atau pemodelan lebih lanjut.

### 3.4. Trend

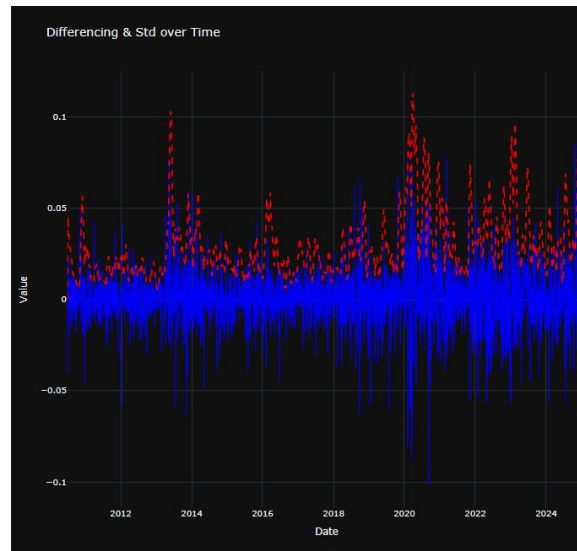


Gambar 4. Grafik Trend

Grafik Close & Trend over Time menampilkan pergerakan harga penutupan (Close) beserta tren jangka panjangnya menggunakan rata-rata bergerak 180 hari (Rolling Mean Trend). Setiap subplot dalam grafik merepresentasikan harga penutupan dari berbagai aset atau indeks yang dianalisis. Garis biru menunjukkan pergerakan harian harga penutupan, sementara garis merah putus-putus menunjukkan tren jangka panjang, yang dihasilkan dari perataan menggunakan jendela waktu 180 hari untuk mengurangi fluktuasi harian dan mengidentifikasi arah pasar yang lebih stabil. Dengan visualisasi ini, dapat diidentifikasi apakah harga suatu aset

mengalami tren naik, turun, atau bergerak sideways dalam periode yang lebih luas, membantu dalam analisis investasi dan strategi perdagangan.

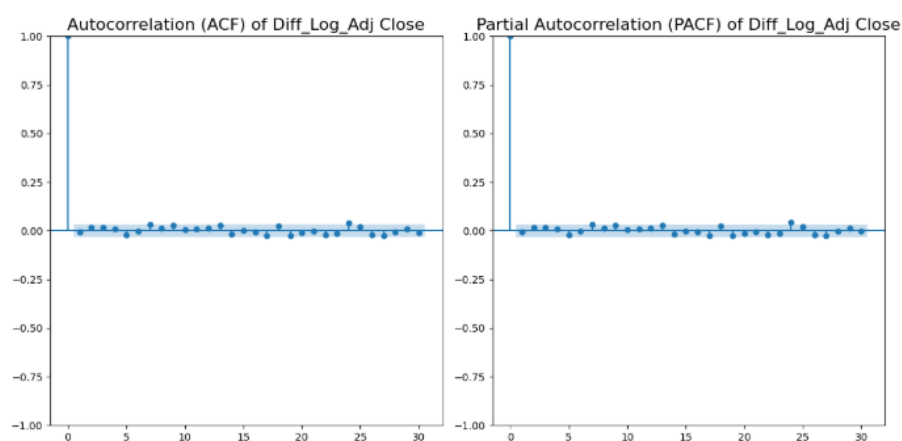
### 3.5. Differencing



Gambar 5. Grafik Differencing

Gambar 5 Grafik Differencing & Standard Deviation over Time menggambarkan hasil transformasi diferensiasi logaritmik harga penutupan yang disesuaikan ( $\text{Diff\_log\_Adj Close}$ ) serta standar deviasi logaritmik dalam jendela bergerak 30 periode ( $\text{Std\_log\_Adj Close}$ ). Diferensiasi digunakan untuk menghilangkan tren dalam data, membuatnya lebih stasioner agar lebih sesuai untuk pemodelan deret waktu, sementara standar deviasi menunjukkan volatilitas harga dalam periode tertentu. Dalam grafik ini, garis biru menunjukkan perubahan harian dalam log harga penutupan setelah diferensiasi, sementara garis merah putus-putus merepresentasikan standar deviasi yang menunjukkan fluktuasi pasar seiring waktu. Pola pada grafik dapat membantu dalam menentukan apakah data sudah cukup stasioner untuk analisis lebih lanjut serta mengidentifikasi periode dengan volatilitas tinggi yang mungkin mempengaruhi prediksi harga.

### 3.6. Autocorrelation & Partial Autocorrelation : $\text{Diff\_Log\_Adj Close}$



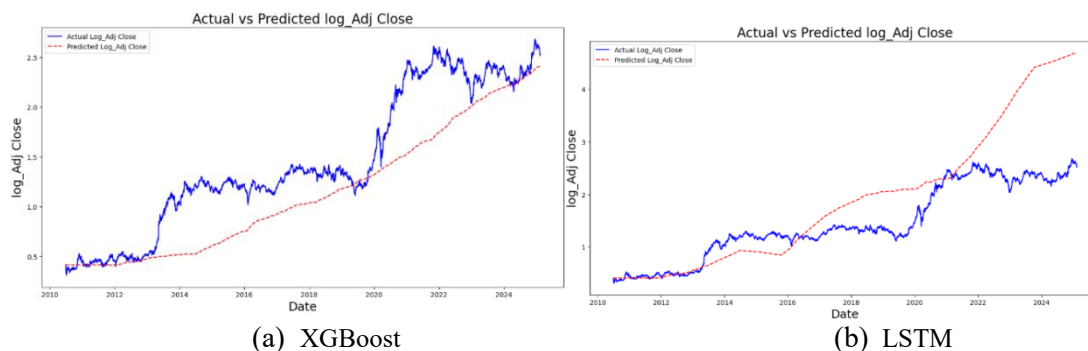
Gambar 6. Grafik Autocorrelation & Partial Autocorrelation :  $\text{Diff\_Log\_Adj Close}$

Gambar 6 Grafik Autocorrelation Function (ACF) dan Partial Autocorrelation Function (PACF) dari  $\text{Diff\_Log\_Adj Close}$  digunakan untuk menganalisis struktur ketergantungan dalam data deret waktu setelah diferensiasi logaritma harga penutupan yang disesuaikan. ACF mengukur korelasi antara nilai saat ini dengan nilai masa lalu pada berbagai lag, menunjukkan adanya pola autokorelasi dalam data. Jika ACF menunjukkan penurunan yang lambat, ini mengindikasikan bahwa data masih memiliki tren yang perlu dihilangkan. Sementara



itu, PACF mengisolasi pengaruh langsung dari lag tertentu dengan menghilangkan efek dari lag lainnya, yang berguna untuk menentukan orde model AR (AutoRegressive). Dalam grafik tersebut, lonjakan signifikan pada lag awal PACF bisa mengindikasikan bahwa model ARIMA yang sesuai mungkin memerlukan komponen autoregressive (AR)..

### 3.7. XGBoost vs LSTM



Gambar 7. Actual vs Predicted log\_Adj Close

Gambar 7.a menunjukkan perbandingan antara nilai log\_Adj Close aktual (garis biru) dengan nilai prediksi dari model XGBoost (garis merah putus-putus) pada data saham Tesla. Model XGBoost dilatih menggunakan fitur Lag\_Diff dan Std\_log\_Adj Close, dengan optimasi hyperparameter menggunakan GridSearchCV dan validasi menggunakan TimeSeriesSplit. Prediksi dilakukan secara bertahap dengan memperbarui model pada setiap iterasi validasi. Dari hasil visualisasi, garis prediksi cukup mengikuti pola data aktual, menunjukkan bahwa model mampu menangkap tren utama harga saham.

Gambar 7.b menunjukkan perbandingan antara nilai log\_Adj Close aktual (garis biru) dengan nilai prediksi dari model LSTM (garis merah putus-putus) pada data saham Tesla. Model LSTM dibangun menggunakan 50 unit neuron dengan fungsi aktivasi ReLU, dioptimasi menggunakan Adam Optimizer, dan dilatih dengan skema TimeSeriesSplit (10 fold) untuk memastikan model dapat menangkap pola time series dengan baik. Proses pelatihan menggunakan Early Stopping untuk mencegah overfitting. Prediksi dilakukan secara bertahap pada setiap split, dan hasilnya dikumulatifkan untuk mendapatkan nilai log\_Adj Close yang sesuai.

### 3.8. Hasil Evaluasi Model XGBoost vs LSTM

Tabel 2. Hasil Evaluasi Model XGBoost vs LSTM

Evaluasi Model XGBoost		Evaluasi Model LSTM	
MAE	13.7196	MAE	13.4788
RMSE	20.4188	RMSE	17.8713
R2	0.9048	R2	0.9271

Berdasarkan hasil evaluasi, model LSTM menunjukkan performa yang lebih baik dibandingkan XGBoost dalam memprediksi harga saham Tesla selama satu dekade. Hal ini terlihat dari nilai Mean Absolute Error (MAE) dan Root Mean Squared Error (RMSE) yang lebih rendah pada LSTM (MAE: 13.4788, RMSE: 17.8713) dibandingkan dengan XGBoost (MAE: 13.7196, RMSE: 20.4188). Selain itu, nilai  $R^2$  (koefisien determinasi) LSTM sebesar 0.9271 lebih tinggi daripada XGBoost yang hanya mencapai 0.9048, menunjukkan bahwa model LSTM dapat menangkap pola data dengan lebih baik dan memberikan prediksi yang lebih akurat. Keunggulan LSTM kemungkinan besar berasal dari kemampuannya dalam menangani data time series dengan dependensi jangka panjang, sementara XGBoost lebih efektif dalam data berbasis fitur yang tidak memiliki keterkaitan urutan waktu yang kuat.

- Mengapa LSTM lebih Baik ? Keunggulan LSTM dalam memprediksi harga saham Tesla selama satu dekade dapat dijelaskan oleh kemampuannya dalam memahami dan memproses dependensi jangka panjang dalam data deret waktu (time series), yang sangat penting dalam konteks pergerakan harga saham yang bersifat temporal dan seringkali dipengaruhi oleh tren historis.
- Apakah hasil ini konsisten dengan penelitian sebelumnya ? Hasil penelitian ini konsisten dengan sejumlah studi sebelumnya yang menunjukkan bahwa model LSTM umumnya lebih unggul dalam memprediksi data

time series finansial dibandingkan algoritma tradisional seperti XGBoost. Beberapa penelitian terdahulu, seperti yang dilakukan oleh Fischer dan Krauss (2018) maupun Nelson et al. (2017), juga menemukan bahwa LSTM mampu menangkap pola temporal dan ketergantungan jangka panjang dalam data pasar saham secara lebih efektif, menghasilkan akurasi prediksi yang lebih tinggi daripada metode non-rekursif.

- Bagaimana potensi penerapan dalam skenario nyata? Potensi penerapan hasil penelitian ini dalam skenario nyata sangat besar, khususnya dalam industri keuangan dan investasi yang sangat bergantung pada akurasi prediksi pasar. Model LSTM yang terbukti mampu menangkap pola jangka panjang dan fluktuasi harga saham secara lebih akurat dapat diintegrasikan ke dalam sistem rekomendasi investasi, algoritma perdagangan otomatis (automated trading), atau platform manajemen risiko untuk membantu investor institusional maupun individu dalam membuat keputusan yang lebih informasional dan tepat waktu. Dengan mengandalkan model prediktif berbasis LSTM, perusahaan sekuritas atau manajer portofolio dapat memantau tren pasar secara real-time, mengantisipasi pergerakan harga yang ekstrem, serta merespons dinamika pasar dengan strategi yang lebih adaptif. Selain itu, model ini juga dapat dikombinasikan dengan indikator ekonomi makro atau sentimen pasar dari media sosial untuk membentuk sistem prediksi hibrida yang lebih kuat dan responsif dalam menghadapi kompleksitas pasar saham di dunia nyata.

#### 4. KESIMPULAN

Penelitian ini membandingkan kinerja algoritma XGBoost dan LSTM dalam memprediksi harga saham Tesla menggunakan data tahun 2025. Berdasarkan hasil analisis, LSTM menunjukkan keunggulan dalam menangkap pola jangka panjang dan hubungan kompleks dalam data time series, menghasilkan prediksi yang lebih akurat terutama dalam tren harga yang fluktuatif. Sementara itu, XGBoost memberikan hasil yang lebih cepat dengan interpretasi yang lebih jelas, namun memiliki keterbatasan dalam memahami dependensi temporal yang mendalam. Evaluasi menggunakan metrik seperti RMSE, MAE, dan  $R^2$  mengonfirmasi bahwa LSTM cenderung lebih unggul dalam akurasi, meskipun dengan waktu pelatihan yang lebih lama dibandingkan XGBoost. Dengan demikian, pemilihan algoritma terbaik bergantung pada kebutuhan spesifik—XGBoost lebih sesuai untuk prediksi cepat dengan interpretasi yang mudah, sementara LSTM lebih ideal untuk analisis mendalam terhadap pola pergerakan harga saham.

#### DAFTAR PUSTAKA

- [1] H. Oukhouya, H. Kadiri, K. El Himdi, and R. Guerbaz, "Forecasting International Stock Market Trends: XGBoost, LSTM, LSTM-XGBoost, And Backtesting XGBoost Models," *Statistics, Optimization and Information Computing*, vol. 12, no. 1, pp. 200–209, 2024, doi: 10.19139/soic-2310-5070-1822.
- [2] R. Zhang, Y. Wang, J. Zhang, L. Zhang, and J. Qu, "Three Machine Learning Predictions of U.S. Stock Prices," *International Conference on Economics, Education and Social Research*, vol. 5, pp. 283–289, 2024, doi: 10.25236/iceesr.2024.046.
- [3] Z. Xu, W. Zhang, Y. Sun, and Z. Lin, "Multi-Source Data-Driven LSTM Framework for Enhanced Stock Price Prediction and Volatility Analysis," 2024.
- [4] E. Hadjaidji, M. C. A. Korba, and K. Khelil, "COVID-19 Detection from Cough Sounds Using XGBoost and LSTM Networks," *Traitement du Signal*, vol. 42, no. 02, pp. 939–947, Apr. 2024, doi: 10.18280/ts.410234.
- [5] Y. Zheng, S. Guan, K. Guo, Y. Zhao, and L. Ye, "Technical indicator enhanced ultra-short-term wind power forecasting based on long short-term memory network combined XGBoost algorithm," *IET Renewable Power Generation*, 2024, doi: 10.1049/rpg2.12952.
- [6] I. Gede, J. Kurniarwan, C. Dewi, and M. A. Rahman, "PENERAPAN MACHINE LEARNING EXTREME GRADIENT BOOSTING DALAM KLASIFIKASI POTENSI TSUNAMI BERDASARKAN DATA GEMPA BUMI," 2025. [Online]. Available: <http://j-ptiik.ub.ac.id>
- [7] A. A. Saputra, B. N. Sari, C. Rozikin, U. Singaperbangsa, and K. Abstrak, "Penerapan Algoritma Extreme Gradient Boosting (Xgboost) Untuk Analisis Risiko Kredit," *Jurnal Ilmiah Wahana Pendidikan*, vol. 10, no. 7, pp. 27–36, 2024, doi: 10.5281/zenodo.10960080.
- [8] P. Pathak, "Stock Market Prediction Using Machine Learning," 2024. [Online]. Available: [www.ijfmr.com](http://www.ijfmr.com)

- 
- [9] M. T. Hidayat and M. Sulistiyono, "Analisis Performa Algoritma XGBoost, GRU, dan Prophet dalam Peramalan Penjualan Obat untuk Optimasi Rantai Pasok Farmasi," *Jurnal Pendidikan dan Teknologi Indonesia*, vol. 5, no. 1, pp. 65–73, Jan. 2025, doi: 10.52436/1.jpti.562.
- [10] Khalis Sofi, Aswan Supriyadi Sunge, Sasmitoh Rahmad Riady, and Antika Zahrotul Kamalia, "PERBANDINGAN ALGORITMA LINEAR REGRESSION, LSTM, DAN GRU DALAM MEMPREDIKSI HARGA SAHAM DENGAN MODEL TIME SERIES," *SEMINASTIKA*, vol. 3, no. 1, pp. 39–46, Nov. 2021, doi: 10.47002/seminastika.v3i1.275.
- [11] B. Pratama and L. Yuniar Banowosari, "COMPARISON OF EXTREME GRADIENT BOOSTING (XGBOOST) METHOD WITH LONG SHORT-TERM MEMORY (LSTM) FOR STOCK PREDICTION PT. BANK MANDIRI TBK. (BMRI) PERBANDINGAN METODE EXTREME GRADIENT BOOSTING (XGBOOST) DENGAN LONG SHORT-TERM MEMORY (LSTM) UNTUK PREDIKSI SAHAM PT. BANK MANDIRI TBK. (BMRI)," *Journal of Economic, Business and Accounting*, vol. 7, no. 3, pp. 5631–5636, 2024.
- [12] A. Haris Prayoga et al., "PENGEMBANGAN APLIKASI BANK ACCOUNT FRAUD DETECTION DENGAN MENGGUNAKAN ALGORITMA XGBOOST," 2024.
- [13] R. Saputra et al., "ANALISIS PREDIKSI SAHAM TESLA MENGGUNAKAN ALGORITMA LONG SHORT TERM MEMORY (LSTM)," *Journal of Computer Science and Information Technology (JCSIT)*, vol. 2, no. 1, 2024.
- [14] J. Pasaribu, N. Yudistira, and W. F. Mahmudy, "Tabular Data Classification and Regression : XGBoost or Deep Learning with Retrieval-Augmented Generation," *IEEE Access*, 2024, doi: 10.1109/ACCESS.2024.3518205.
- [15] Z. Chen, Z. W. Li, J. Huang, S. Z. Liu, and H. X. Long, "An effective method for anomaly detection in industrial Internet of Things using XGBoost and LSTM," *Sci Rep*, vol. 14, no. 1, p. 23969, Dec. 2024, doi: 10.1038/s41598-024-74822-6.
- [16] K. Fu and Y. Zhang, "Incorporating Multi-Source Market Sentiment and Price Data for Stock Price Prediction," *Mathematics*, vol. 12, no. 10, May 2024, doi: 10.3390/math12101572.
- [17] A. Frifra, M. Maanan, M. Maanan, and H. Rhinane, "Harnessing LSTM and XGBoost algorithms for storm prediction," *Sci Rep*, vol. 14, no. 1, Dec. 2024, doi: 10.1038/s41598-024-62182-0.
- [18] T. Sugihartono, B. Wijaya, Marini, A. F. Alkayes, and H. A. Anugrah, "Optimizing Stunting Detection through SMOTE and Machine Learning: a Comparative Study of XGBoost, Random Forest, SVM, and k-NN," *Journal of Applied Data Sciences*, vol. 6, no. 1, pp. 667–682, Jan. 2025, doi: 10.47738/jads.v6i1.494.
- [19] I. Rifky Hendrawan, "'Jurnal TRANSFORMASI (Informasi & Pengembangan Iptek)' (STMIK BINA PATRIA) PERBANDINGAN ALGORITMA NAÏVE BAYES, SVM DAN XGBOOST DALAM KLASIFIKASI TEKS SENTIMEN MASYARAKAT TERHADAP PRODUK LOKAL DI INDONESIA," 2022.
- [20] Intan Permata and Esther Sorta Mauli Nababan, "Application Of Game Theory In Determining Optimum Marketing Strategy In Marketplace," *JURNAL RISET RUMPUN MATEMATIKA DAN ILMU PENGETAHUAN ALAM*, vol. 2, no. 2, pp. 65–71, Jul. 2023, doi: 10.55606/jurrimipa.v2i2.1336.
- [21] T. Abdupatah and C. Rozikin, "Prediksi Harga Saham Menggunakan Machine Learning Dengan Metode Long Short Therm Memory (LSTM) Study Kasus : Saham Tesla," *Jurnal Informatika MULTI*, vol. 3, no. 1, 2025, [Online]. Available: <https://jurnal.publikasitecno.id/index.php/jim>
- [22] D. Kusuma Ningrum and A. Maytsa Ismawardi, "EFEKTIVITAS ALGORITMA KECERDASAN BUATAN DALAM IMPLEMENTASI KESEHATAN MENTAL : SYSTEMATIC LITERATURE REVIEW," 2025.