

KOMPARASI METODE LSTM DAN ARIMA DALAM PREDIKSI HARGA SAHAM NVIDIA PERIODE 2012-2025

Sopiyan Apandi, Andre Effendi, Fairus Priyogi, Rajwaa Nazir Yatim

Teknik Informatika, Universitas Pamulang

Jl. Puspitek, Buaran, Kec. Pamulang, Kota Tangerang Selatan, Banten

dosen02601@unpam.ac.id

ABSTRAK

Pergerakan harga saham yang fluktuatif dan dinamis menjadikan proses prediksi harga saham sebagai salah satu topik penting dalam analisis time series finansial. Namun, pemilihan metode prediksi yang mampu menghasilkan akurasi tinggi masih menjadi tantangan karena karakteristik data saham yang kompleks dan berubah-ubah. Penelitian ini bertujuan untuk membandingkan performa metode *Autoregressive Integrated Moving Average* (ARIMA) dan *Long Short-Term Memory* (LSTM) dalam memprediksi harga saham NVIDIA Corporation menggunakan data historis periode 2012-2025. Dataset yang digunakan terdiri atas 3.415 data *adjusted closing price* yang dibagi menjadi data training dan testing dengan rasio 80:20. Model yang diuji meliputi ARIMA(5,1,0), LSTM-20, LSTM-50, dan LSTM-ES dengan mekanisme *EarlyStopping*. Evaluasi dilakukan menggunakan *Mean Absolute Error* (MAE), *Root Mean Squared Error* (RMSE), dan *Mean Absolute Percentage Error* (MAPE). Hasil penelitian menunjukkan bahwa ARIMA menghasilkan performa terbaik dengan nilai MAE sebesar 1.9576, RMSE sebesar 3.0646, dan MAPE sebesar 2.3933%, sedangkan LSTM-ES menjadi model terbaik pada kelompok LSTM dengan nilai MAPE sebesar 3.4719%. Temuan ini menunjukkan bahwa ARIMA masih mampu memberikan prediksi yang lebih akurat pada data saham NVIDIA yang memiliki pola autoregresif dominan.

Kata kunci : ARIMA, Forecasting, Harga Saham, LSTM, NVIDIA, Time Series.

1. PENDAHULUAN

Pasar modal global dalam beberapa tahun terakhir berkembang menjadi ekosistem yang sangat dinamis, di mana fluktuasi harga saham sektor teknologi menjadi perhatian utama investor, analis, dan akademisi. Perubahan valuasi yang cepat menunjukkan bahwa pasar semakin sensitif terhadap inovasi, sentimen industri, serta prospek pertumbuhan jangka panjang. Dalam konteks ini, kemampuan memprediksi harga saham menjadi kebutuhan strategis untuk memaksimalkan keuntungan sekaligus meminimalkan risiko investasi di tengah ketidakpastian ekonomi global. Salah satu emiten yang menunjukkan pertumbuhan paling signifikan adalah NVIDIA Corporation. Sebagai pemimpin pasar dalam penyediaan unit pemroses grafis (GPU) dan infrastruktur *Artificial Intelligence* (AI), NVIDIA mengalami lonjakan nilai perusahaan sepanjang periode 2012–2025 [1], [2]. Namun demikian, pergerakan harga saham tersebut juga diwarnai volatilitas tinggi, sehingga prediksi harga saham NVIDIA menjadi tantangan penting dalam analisis deret waktu (*time series*).

Secara teoritis, metode prediksi harga saham dapat dikelompokkan ke dalam dua paradigma utama, yaitu model statistik klasik dan model berbasis kecerdasan buatan. Dalam paradigma statistik, *Autoregressive Integrated Moving Average* (ARIMA) telah lama digunakan sebagai acuan (*benchmark*) karena memiliki dasar matematis yang kuat, interpretasi yang jelas, serta efektif dalam memodelkan pola linear [3], [4]. Akan tetapi,

kemampuan ARIMA sering kali terbatas ketika dihadapkan pada data yang non-linear, *non-stationary*, dan mengandung *noise* tinggi sebagaimana karakteristik pasar saham modern [5], [6]. Di sisi lain, perkembangan *deep learning* menghadirkan *Long Short-Term Memory* (LSTM), yaitu varian *Recurrent Neural Network* (RNN) yang dirancang untuk menangkap ketergantungan jangka panjang dan pola kompleks pada data sekuensial [7], [8]. Dengan mekanisme memori internal, LSTM dinilai lebih adaptif dalam mengenali perubahan tren dan dinamika pasar yang cepat [9], [10].

Berbagai penelitian terdahulu menunjukkan bahwa LSTM dan turunannya memiliki performa yang baik dalam memprediksi aset berisiko tinggi, termasuk saham teknologi dan cryptocurrency [10]. Sebagian penelitian bahkan mengembangkan model *hybrid* seperti GARCH-LSTM atau optimasi metaheuristik guna meningkatkan akurasi prediksi [4]. Namun, fokus yang terlalu besar pada model kompleks menyebabkan studi komparasi langsung (*head-to-head*) antara ARIMA sebagai representasi model statistik klasik dan LSTM sebagai representasi *deep learning* masih relatif terbatas, khususnya pada saham spesifik NVIDIA [1], [2]. Selain itu, sebagian besar penelitian sebelumnya menggunakan dataset yang hanya mencakup periode hingga tahun 2023 atau awal 2024 [8], [6], sehingga belum sepenuhnya merepresentasikan fase lonjakan harga NVIDIA pada puncak adopsi AI global tahun 2025. Oleh karena itu, masih diperlukan penelitian yang mengevaluasi dan membandingkan secara langsung performa metode ARIMA dan LSTM pada data harga saham NVIDIA

yang mencakup periode terbaru untuk mengetahui metode yang mampu menghasilkan tingkat akurasi prediksi yang lebih baik.

Berdasarkan permasalahan tersebut, penelitian ini bertujuan untuk membandingkan performa metode ARIMA dan LSTM dalam memprediksi harga saham NVIDIA menggunakan data historis periode 2012-2025. Rentang data yang lebih panjang memungkinkan analisis terhadap berbagai fase pasar, mulai dari kondisi stabil, pertumbuhan moderat, hingga lonjakan eksponensial. Penelitian ini juga menggunakan tiga metrik evaluasi sekaligus, yaitu *Root Mean Squared Error* (RMSE), *Mean Absolute Error* (MAE), dan *Mean Absolute Percentage Error* (MAPE), sehingga penilaian performa model menjadi lebih komprehensif dibandingkan penelitian yang hanya menggunakan satu indikator [7], [6].

Untuk mencapai tujuan penelitian, dilakukan serangkaian tahapan yang meliputi prapemrosesan data, pembangunan model ARIMA dan LSTM, serta evaluasi hasil prediksi menggunakan metrik MAE, RMSE, dan MAPE. Seluruh proses penelitian dilaksanakan menggunakan platform Google Colaboratory untuk memastikan konsistensi pengolahan dan pengujian model. Kebaruan penelitian ini terletak pada penggunaan data historis harga saham NVIDIA periode 2012-2025, komparasi langsung antara pendekatan statistik klasik dan *deep learning*, serta penerapan evaluasi menggunakan beberapa metrik performa secara simultan. Hasil penelitian diharapkan dapat memberikan kontribusi terhadap pengembangan studi *forecasting* keuangan sekaligus menjadi referensi dalam pemilihan metode prediksi yang sesuai untuk saham sektor teknologi dengan tingkat volatilitas yang tinggi.

2. TINJAUAN PUSTAKA

Dalam penelitian ini, teori yang digunakan mencakup konsep prediksi harga saham, metode *Autoregressive Integrated Moving Average* (ARIMA), metode *Long Short-Term Memory* (LSTM), serta metrik evaluasi model prediksi. Selain itu, kajian terhadap penelitian terdahulu dilakukan untuk mengetahui perkembangan penelitian di bidang prediksi harga saham, membandingkan hasil penelitian sebelumnya, serta mengidentifikasi *research gap* yang menjadi dasar dilakukannya penelitian ini.

2.1. Prediksi harga saham

Prediksi harga saham merupakan proses memperkirakan nilai saham pada periode mendatang berdasarkan data historis dan pola pergerakan pasar sebelumnya. Prediksi ini menjadi salah satu aspek penting dalam analisis pasar modal karena dapat membantu investor dalam menentukan strategi investasi, meminimalkan risiko, serta mengoptimalkan keuntungan. Data harga saham termasuk ke dalam data deret waktu (*time series*) yang memiliki karakteristik dinamis, fluktuatif, dan dipengaruhi oleh berbagai

faktor ekonomi maupun sentimen pasar. Oleh karena itu, proses prediksi harga saham memerlukan metode yang mampu menangkap pola perubahan data secara akurat.

Dalam perkembangannya, metode prediksi harga saham dapat dibedakan menjadi dua pendekatan utama, yaitu metode statistik klasik dan metode berbasis kecerdasan buatan (*Artificial Intelligence*). Metode statistik seperti ARIMA banyak digunakan karena memiliki kemampuan dalam memodelkan hubungan linear pada data historis serta interpretasi matematis yang jelas. Namun, metode ini memiliki keterbatasan ketika dihadapkan pada data yang bersifat non-linear dan memiliki volatilitas tinggi. Sebaliknya, perkembangan *machine learning* dan *deep learning* menghadirkan metode seperti LSTM yang dinilai lebih adaptif dalam menangkap pola kompleks dan ketergantungan jangka panjang pada data saham.

Perkembangan teknologi kecerdasan buatan juga mendorong meningkatnya penggunaan model prediksi berbasis *deep learning* pada sektor keuangan. Model seperti LSTM mampu mempelajari pola historis secara lebih mendalam melalui mekanisme memori internal, sehingga sering digunakan dalam prediksi saham, *cryptocurrency*, nilai tukar, dan berbagai data finansial lainnya. Meskipun demikian, efektivitas model prediksi tetap bergantung pada karakteristik data, periode pengamatan, metode evaluasi, serta pendekatan *forecasting* yang digunakan. Oleh sebab itu, penelitian mengenai komparasi performa model prediksi masih menjadi topik yang relevan dalam pengembangan analisis pasar saham.

2.2. Metode ARIMA

Autoregressive Integrated Moving Average (ARIMA) merupakan salah satu metode statistik yang paling banyak digunakan dalam analisis dan prediksi data deret waktu (*time series*). Model ARIMA dikembangkan untuk memprediksi nilai masa depan berdasarkan pola historis dari data sebelumnya. ARIMA bekerja dengan memanfaatkan hubungan autoregresif antar data, proses *differencing* untuk mencapai kondisi stasioner, serta komponen *moving average* untuk memodelkan kesalahan residual pada data.

Model ARIMA dinyatakan dalam bentuk $ARIMA(p,d,q)$, di mana parameter p menunjukkan jumlah *lag autoregresif*, parameter d menunjukkan tingkat *differencing* yang digunakan untuk membuat data menjadi stasioner, dan parameter q menunjukkan jumlah komponen *moving average*. Salah satu asumsi utama dalam penggunaan ARIMA adalah data harus berada dalam kondisi stasioner, yaitu memiliki rata-rata dan variansi yang relatif konstan sepanjang waktu. Oleh karena itu, pengujian stasioneritas seperti *Augmented Dickey-Fuller* (ADF) sering digunakan sebelum proses pemodelan dilakukan.

ARIMA memiliki beberapa keunggulan, di antaranya interpretasi model yang jelas, efisiensi komputasi yang baik, serta kemampuan memodelkan

pola linear pada data historis. Metode ini juga sering digunakan sebagai model *benchmark* dalam penelitian prediksi saham karena relatif stabil dan mudah diimplementasikan. Namun demikian, ARIMA memiliki keterbatasan dalam menangkap hubungan non-linear dan perubahan pola yang kompleks, terutama pada data saham dengan volatilitas tinggi. Oleh sebab itu, performa ARIMA sangat dipengaruhi oleh karakteristik data yang digunakan.

2.3. Metode LSTM

Long Short-Term Memory (LSTM) merupakan salah satu varian dari *Recurrent Neural Network* (RNN) yang dirancang khusus untuk memproses data berurutan (*sequential data*) seperti data deret waktu. LSTM dikembangkan untuk mengatasi permasalahan *vanishing gradient* yang sering terjadi pada RNN konvensional ketika mempelajari ketergantungan data jangka panjang. Dengan mekanisme memori internal, LSTM mampu menyimpan dan mempelajari informasi historis dalam periode yang lebih panjang dibandingkan metode *neural network* biasa.

Arsitektur LSTM terdiri dari beberapa komponen utama, yaitu *forget gate*, *input gate*, dan *output gate*. Ketiga komponen tersebut bekerja untuk menentukan informasi mana yang perlu dipertahankan, diperbarui, maupun dihapus dari memori model. Mekanisme ini memungkinkan LSTM mengenali pola non-linear, tren jangka panjang, serta perubahan dinamis pada data saham yang kompleks dan volatil. Oleh karena itu, LSTM banyak digunakan dalam berbagai penelitian prediksi harga saham, *cryptocurrency*, nilai tukar mata uang, hingga permintaan pasar.

Dibandingkan metode statistik klasik, LSTM memiliki kemampuan yang lebih baik dalam menangkap pola kompleks dan hubungan non-linear antar data. Namun demikian, model LSTM membutuhkan proses pelatihan yang lebih lama, sumber daya komputasi yang lebih besar, serta pengaturan parameter yang lebih kompleks. Selain itu, performa LSTM sangat dipengaruhi oleh jumlah data, teknik *preprocessing*, dan konfigurasi arsitektur model yang digunakan selama proses pelatihan.

2.4. Metrik evaluasi

Evaluasi model prediksi dilakukan untuk mengukur tingkat akurasi hasil peramalan terhadap data aktual. Dalam penelitian prediksi harga saham, metrik evaluasi digunakan untuk mengetahui seberapa baik model mampu meminimalkan kesalahan prediksi. Penelitian ini menggunakan tiga metrik evaluasi, yaitu *Mean Absolute Error* (MAE), *Root Mean Squared Error* (RMSE), dan *Mean Absolute Percentage Error* (MAPE).

MAE digunakan untuk mengukur rata-rata selisih absolut antara nilai prediksi dan nilai aktual tanpa memperhatikan arah kesalahan. Nilai MAE yang semakin kecil menunjukkan bahwa hasil prediksi model semakin mendekati data sebenarnya. RMSE digunakan untuk mengukur akar rata-rata kuadrat

kesalahan prediksi. Metrik ini memberikan penalti yang lebih besar terhadap kesalahan dengan nilai tinggi, sehingga sensitif terhadap *error* ekstrem pada data saham. Sementara itu, MAPE digunakan untuk mengukur tingkat kesalahan prediksi dalam bentuk persentase sehingga lebih mudah diinterpretasikan.

Penggunaan ketiga metrik tersebut secara bersamaan bertujuan untuk memberikan evaluasi yang lebih komprehensif terhadap performa model prediksi. Dengan kombinasi MAE, RMSE, dan MAPE, penelitian dapat menilai tingkat kesalahan prediksi dari sisi deviasi absolut, sensitivitas terhadap *error* besar, serta persentase kesalahan relatif terhadap data aktual.

2.5. Penelitian terdahulu

Penelitian terdahulu mengenai prediksi harga saham menunjukkan bahwa metode *Autoregressive Integrated Moving Average* (ARIMA) dan *Long Short-Term Memory* (LSTM) merupakan dua pendekatan yang paling banyak digunakan dalam analisis *time series*. Model ARIMA dikenal efektif dalam memodelkan data stasioner dan pola linear melalui komponen *autoregressive*, *differencing*, dan *moving average* [5]. Sementara itu, LSTM sebagai pengembangan dari *Recurrent Neural Network* (RNN) memiliki kemampuan dalam menangkap ketergantungan jangka panjang dan pola non-linear melalui mekanisme *forget gate*, *input gate*, dan *output gate*, sehingga dinilai lebih adaptif terhadap karakteristik data saham yang kompleks dan volatil [6], [8]. Perbedaan karakteristik kedua metode tersebut menyebabkan banyak penelitian mencoba membandingkan performa ARIMA dan LSTM dalam berbagai konteks pasar keuangan untuk menentukan model prediksi yang paling akurat.

Penelitian oleh Milniadi dan Adiwijaya membandingkan model ARIMA dan LSTM pada enam kategori saham berdasarkan klasifikasi Peter Lynch menggunakan data historis periode 2018–2022 [5]. Penelitian tersebut menggunakan pembagian data *training* dan *testing* serta evaluasi berbasis RMSE dan MAPE untuk menentukan performa model terbaik. Hasil penelitian menunjukkan bahwa performa ARIMA maupun LSTM berbeda pada tiap kategori saham, sehingga efektivitas model sangat dipengaruhi oleh karakteristik data yang digunakan. Penelitian lain oleh Zhang pada indeks CSI 300 di pasar saham Tiongkok juga membandingkan kemampuan prediksi ARIMA dan LSTM menggunakan data *adjusted closing price* periode 2022–2023 [6]. Hasil penelitian menunjukkan bahwa LSTM memiliki kemampuan prediksi lebih kuat dibandingkan ARIMA karena mampu mempelajari hubungan non-linear dan memori jangka panjang, namun model tersebut mengalami *lag* dalam menangkap perubahan tren ekstrem jangka pendek. Penelitian tersebut juga menegaskan bahwa data saham memiliki sifat *autocorrelation* dan *long-term memory* yang memungkinkan prediksi dilakukan berdasarkan pola historis.

Selain itu, penelitian Wang mengenai prediksi harga saham Tesla menggunakan ARIMA, LSTM, *neural network*, dan *decomposition linear model* menunjukkan bahwa metode *deep learning* tidak selalu menjadi model terbaik dalam seluruh kondisi data. Penelitian tersebut menemukan bahwa nilai MAE model ARIMA sebesar 7.2400, sedangkan LSTM menghasilkan MAE sebesar 6.8390, namun *decomposition linear model* justru memberikan performa terbaik secara keseluruhan. Temuan ini memperlihatkan bahwa peningkatan kompleksitas model tidak selalu menghasilkan peningkatan akurasi secara signifikan. Bahkan, penelitian tersebut juga menyoroti adanya keraguan terhadap efektivitas model *deep learning* yang terlalu kompleks dalam *forecasting* saham [8].

Berbeda dengan penelitian terdahulu, penelitian ini memfokuskan analisis pada saham NVIDIA Corporation periode 2012–2025 yang merepresentasikan fase transformasi besar industri kecerdasan buatan (*Artificial Intelligence*). Dataset yang digunakan lebih panjang dan mencakup masa stabil hingga lonjakan harga ekstrem akibat meningkatnya adopsi AI global. Selain itu, penelitian ini menerapkan metode *rolling forecast* pada model ARIMA (5,1,0), yaitu pendekatan yang memperbarui data aktual setiap hari sebelum melakukan prediksi berikutnya sehingga lebih mendekati kondisi pasar nyata. Penelitian ini juga menggunakan evaluasi yang lebih komprehensif melalui tiga metrik sekaligus, yaitu RMSE, MAE, dan MAPE.

3. METODE PENELITIAN

Penelitian ini dilaksanakan melalui serangkaian tahapan sistematis yang mencakup pengumpulan data historis harga saham NVIDIA Corporation, *preprocessing* data, pembentukan model prediksi, hingga evaluasi performa model menggunakan pendekatan statistik dan *deep learning*. Proses *forecasting* dilakukan menggunakan metode *Autoregressive Integrated Moving Average* (ARIMA) dan *Long Short-Term Memory* (LSTM) untuk menganalisis kemampuan masing-masing model dalam memprediksi pergerakan harga saham pada data *time series* dengan tingkat volatilitas tinggi. Seluruh proses komputasi dan pengolahan data dilakukan menggunakan platform Google Colaboratory berbasis bahasa pemrograman Python.

3.1. Desain penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan metode eksperimen komputasional pada data deret waktu (*time series*) untuk menganalisis dan membandingkan performa metode statistik klasik *Autoregressive Integrated Moving Average* (ARIMA) dan metode *deep learning Long Short-Term Memory* (LSTM) dalam memprediksi harga saham NVIDIA Corporation (NVDA). Pendekatan komparatif digunakan untuk mengevaluasi kemampuan masing-

masing model dalam menangkap pola historis harga saham yang bersifat dinamis, volatil, dan non-linear.

Penelitian difokuskan pada analisis akurasi prediksi berdasarkan tiga metrik evaluasi utama, yaitu *Mean Absolute Error* (MAE), *Root Mean Squared Error* (RMSE), dan *Mean Absolute Percentage Error* (MAPE). Seluruh tahapan penelitian, mulai dari *preprocessing* data, pemodelan, proses pelatihan, hingga evaluasi model dilakukan menggunakan platform Google Colaboratory berbasis bahasa pemrograman Python.

3.2. Dataset Penelitian

Dataset yang digunakan dalam penelitian ini merupakan data historis harga saham NVIDIA Corporation (NVDA) yang diperoleh dari platform Kaggle dalam bentuk *adjusted closing price*. Penggunaan *adjusted closing price* bertujuan untuk mempertahankan konsistensi historis data terhadap peristiwa stock split, dividen, dan aksi korporasi lainnya sehingga data lebih representatif untuk analisis *time series* jangka Panjang.

Dataset mencakup periode perdagangan dari 3 Januari 2012 hingga 1 Agustus 2025 dengan total sebanyak 3.415 data harian perdagangan saham. Penelitian hanya menggunakan dua atribut utama, yaitu tanggal perdagangan (*date*) dan harga penutupan saham (*close*). Harga minimum pada dataset berada pada kisaran \$0.32 per saham pada awal periode pengamatan tahun 2012, sedangkan harga maksimum mencapai lebih dari \$176 per saham pada tahun 2025. Nilai harga yang relatif kecil pada awal periode bukan merupakan anomali data, melainkan konsekuensi dari proses penyesuaian historis akibat *stock split* NVIDIA.

Karakteristik dataset menunjukkan adanya fluktuasi harga yang tinggi, terutama pada periode 2023–2025 ketika peningkatan adopsi teknologi *Artificial Intelligence* (AI) mendorong kenaikan valuasi NVIDIA secara signifikan. Kondisi tersebut menjadikan dataset penelitian memiliki kompleksitas tinggi dan relevan untuk menguji kemampuan model statistik maupun *deep learning* dalam melakukan *forecasting* harga saham.

3.3. Pembagian Data

Pembagian dataset dilakukan secara kronologis menggunakan rasio 80:20, di mana sebanyak 2.732 data digunakan sebagai data training dan 683 data digunakan sebagai data testing. Pendekatan kronologis dipilih untuk mempertahankan urutan temporal data sehingga model tidak memperoleh informasi masa depan pada saat proses pelatihan berlangsung.

Metode pembagian ini berbeda dengan *random split* yang umum digunakan pada data *non-time-series*. Dalam penelitian *forecasting*, penggunaan *random split* dapat menyebabkan data *leakage temporal* yang berpotensi menghasilkan evaluasi model yang bias dan tidak realistis. Oleh karena itu, penelitian ini menggunakan data historis periode awal sebagai data pelatihan dan data periode terbaru sebagai data

pengujian agar proses evaluasi lebih mendekati kondisi prediksi di dunia nyata.

Tabel 1. Pembagian Dataset

Dataset	Jumlah Data	Persentase
Data Training	2.732	80%
Data Testing	683	20%
Total	3.415	100%

3.4. Implementasi Model ARIMA

Autoregressive Integrated Moving Average (ARIMA) digunakan sebagai representasi model statistik klasik dalam penelitian ini. Model ARIMA dipilih karena memiliki kemampuan yang baik dalam memodelkan pola linear dan hubungan *autoregresif* pada data *time series*.

Model ARIMA dinyatakan dalam bentuk $ARIMA(p,d,q)$, dengan parameter:

- p menunjukkan jumlah *lag autoregressive*.
- d menunjukkan tingkat *differencing*.
- q menunjukkan jumlah komponen *moving average*.

Penelitian ini menggunakan konfigurasi ARIMA (5,1,0), yang berarti model memanfaatkan lima *lag* historis, satu kali *differencing*, dan tanpa komponen *moving average*. Pemilihan parameter dilakukan berdasarkan karakteristik data saham yang menunjukkan pola *autoregresif* cukup kuat.

Proses prediksi dilakukan menggunakan pendekatan *rolling forecast*. Pada metode ini, model dilatih menggunakan seluruh data training, kemudian melakukan prediksi secara bertahap terhadap data *testing*. Setelah satu nilai diprediksi, data aktual hari tersebut dimasukkan kembali ke dalam *history* sebelum model melakukan prediksi berikutnya. Proses tersebut diulang hingga seluruh data *testing* selesai diprediksi.

Pendekatan *rolling forecast* dipilih karena mampu mensimulasikan kondisi prediksi pasar saham secara realistis, di mana informasi harga aktual hari sebelumnya telah diketahui sebelum melakukan prediksi harga hari berikutnya. Selain itu, pendekatan ini memungkinkan model ARIMA menyesuaikan pola prediksi secara dinamis terhadap perubahan data terbaru.

3.5. Implementasi Model LSTM

Long Short-Term Memory (LSTM) digunakan sebagai representasi model *deep learning* dalam penelitian ini. LSTM dipilih karena memiliki kemampuan untuk mempelajari hubungan non-linear serta ketergantungan jangka panjang pada data sekuensial melalui mekanisme memori internal.

Sebelum proses pelatihan dilakukan, data harga saham dinormalisasi menggunakan *MinMaxScaler* dengan rentang nilai 0 hingga 1. Normalisasi dilakukan untuk mempercepat proses konvergensi model dan menjaga stabilitas pelatihan jaringan saraf. Setelah proses prediksi selesai, hasil prediksi

dikembalikan ke skala harga asli menggunakan *inverse transformation*. Arsitektur model LSTM terdiri atas:

- Input sequence* dengan *look-back window* selama 60 hari.
- Dua *hidden layer* LSTM masing-masing sebanyak 50 unit.
- Satu *dense layer* sebanyak 25 unit.
- Satu *output layer* dengan satu *neuron* untuk menghasilkan prediksi harga saham.

Optimizer yang digunakan adalah Adam dengan fungsi *loss Mean Squared Error* (MSE). Penelitian ini mengembangkan tiga strategi pelatihan model LSTM, yaitu:

- LSTM-20 dengan *fixed epoch* sebanyak 20.
- LSTM-50 dengan *fixed epoch* sebanyak 50.
- LSTM-ES menggunakan mekanisme *EarlyStopping* dengan maksimum 200 *epoch* dan *patience* sebesar 15 *epoch*.

Untuk menjaga validitas komparasi antar model, seluruh varian LSTM menggunakan *shared initial weights* sehingga setiap model memiliki bobot awal yang identik sebelum proses pelatihan dilakukan. Pendekatan ini bertujuan untuk memastikan bahwa perbedaan performa model benar-benar berasal dari strategi pelatihan dan jumlah *epoch*, bukan akibat variasi inisialisasi bobot secara acak.

Pada model LSTM-ES, mekanisme *EarlyStopping* digunakan untuk menghentikan pelatihan secara otomatis ketika *validation loss* tidak menunjukkan peningkatan selama 15 *epoch* berturut-turut. Model terbaik diperoleh pada *epoch* ke-42 dengan *validation loss* sebesar 3.76×10^{-5} , sedangkan proses pelatihan berhenti secara otomatis pada *epoch* ke-57.

3.6. Tahapan Penelitian

Tahapan penelitian dilakukan secara sistematis melalui beberapa proses utama sebagai berikut:

- Pengumpulan dataset harga saham NVIDIA Corporation.
- Seleksi atribut dan *preprocessing* data.
- Normalisasi data menggunakan *MinMaxScaler* untuk model LSTM.
- Pembagian data *training* dan *testing* secara kronologis.
- Pemodelan menggunakan ARIMA (5,1,0).
- Pelatihan tiga varian model LSTM.
- Proses *forecasting* terhadap data *testing*.
- Evaluasi performa model menggunakan MAE, RMSE, dan MAPE.
- Analisis komparatif terhadap hasil prediksi seluruh model.

3.7. Teknik Evaluasi Model

Evaluasi performa model dilakukan untuk mengukur tingkat akurasi prediksi harga saham yang dihasilkan oleh metode ARIMA dan LSTM terhadap data aktual. Pengukuran performa dilakukan menggunakan tiga metrik evaluasi, yaitu *Mean Absolute Error* (MAE), *Root Mean Squared Error*

(RMSE), dan *Mean Absolute Percentage Error* (MAPE).

MAE digunakan untuk mengukur rata-rata selisih absolut antara nilai prediksi dan nilai aktual tanpa mempertimbangkan arah kesalahan. RMSE digunakan untuk mengukur sensitivitas model terhadap *error* bernilai besar melalui proses kuadrat *error*. Sementara itu, MAPE digunakan untuk mengukur tingkat kesalahan prediksi dalam bentuk persentase sehingga mempermudah interpretasi tingkat akurasi model *forecasting*.

Penggunaan ketiga metrik secara bersamaan bertujuan untuk menghasilkan evaluasi performa yang lebih komprehensif, baik dari sisi deviasi absolut, sensitivitas terhadap *error* ekstrem, maupun persentase kesalahan relatif terhadap data aktual. Semakin kecil nilai MAE, RMSE, dan MAPE yang dihasilkan, maka semakin baik kemampuan model dalam melakukan prediksi harga saham.

4. HASIL DAN PEMBAHASAN

Hasil penelitian diperoleh melalui proses pengujian model ARIMA dan LSTM terhadap data testing harga saham NVIDIA Corporation periode 2022–2025. Evaluasi dilakukan menggunakan metrik *Mean Absolute Error* (MAE), *Root Mean Squared Error* (RMSE), dan *Mean Absolute Percentage Error* (MAPE) untuk mengukur tingkat akurasi prediksi masing-masing model. Selain menyajikan hasil evaluasi kuantitatif, pembahasan juga difokuskan pada analisis karakteristik prediksi, perilaku proses pelatihan model *deep learning*, serta interpretasi performa model berdasarkan pola data *time series* yang digunakan.

4.1. Hasil Evaluasi Model

Evaluasi performa model dilakukan terhadap empat model prediksi, yaitu ARIMA, LSTM-20, LSTM-50, dan LSTM-ES menggunakan data testing sebanyak 683 data perdagangan saham NVIDIA periode 2022–2025.

Tabel 2. Hasil Evaluasi Performa Model

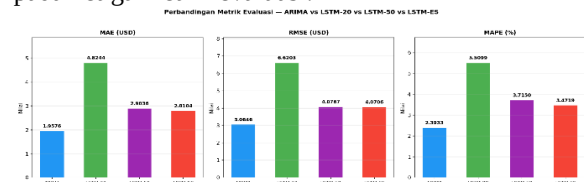
Model	MAE	RMSE	MAPE
ARIMA	1.9576	3.0646	2.3933%
LSTM-20	4.8244	6.6203	5.5099%
LSTM-50	2.9036	4.0767	3.7150%
LSTM-ES	2.8104	4.0706	3.4719%

Berdasarkan hasil evaluasi pada Tabel 4.1, model ARIMA menghasilkan performa terbaik pada seluruh metrik evaluasi dibandingkan seluruh varian LSTM. Nilai MAE sebesar 1.9576 menunjukkan bahwa rata-rata kesalahan prediksi model ARIMA hanya berada pada kisaran \$1.96 terhadap harga aktual saham NVIDIA. Selain itu, nilai MAPE sebesar 2.3933% menunjukkan bahwa model ARIMA memiliki tingkat kesalahan prediksi yang sangat rendah dan termasuk dalam kategori sangat akurat.

Di sisi lain, model LSTM-ES menghasilkan performa terbaik di antara seluruh model LSTM

dengan nilai MAPE sebesar 3.4719%, diikuti oleh LSTM-50 sebesar 3.7150%, dan LSTM-20 sebesar 5.5099%.

Visualisasi perbandingan nilai MAE, RMSE, dan MAPE dari masing-masing model ditunjukkan pada Gambar 1. Grafik memperlihatkan bahwa model ARIMA secara konsisten menghasilkan nilai *error* paling rendah dibandingkan seluruh model LSTM pada ketiga metrik evaluasi.



Gambar 1. Perbandingan Metrik Evaluasi

4.2. Analisis Performa Model ARIMA

Model ARIMA menunjukkan performa terbaik dibandingkan seluruh model lainnya pada penelitian ini. Keunggulan tersebut mengindikasikan bahwa pola harga saham NVIDIA pada periode pengamatan masih memiliki karakteristik *autoregresif* yang cukup kuat sehingga dapat dimodelkan secara efektif menggunakan pendekatan statistik klasik.

Penggunaan metode *rolling forecast* memberikan kontribusi signifikan terhadap stabilitas prediksi model ARIMA. Pada pendekatan ini, model selalu memperbarui *history* menggunakan data aktual terbaru sebelum melakukan prediksi berikutnya. Mekanisme tersebut memungkinkan model menyesuaikan pola prediksi secara dinamis terhadap perubahan harga saham dari waktu ke waktu.

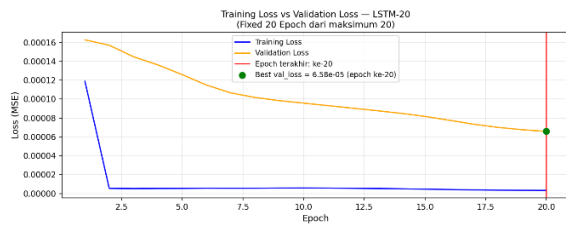
Selain itu, rendahnya nilai RMSE pada model ARIMA menunjukkan bahwa model mampu mempertahankan stabilitas prediksi dan meminimalkan *error* ekstrem selama periode volatilitas tinggi saham NVIDIA, khususnya pada fase peningkatan harga akibat pertumbuhan industri *Artificial Intelligence* pada periode 2023–2025.

Hasil penelitian menunjukkan bahwa pendekatan statistik klasik masih memiliki relevansi tinggi dalam *forecasting* harga saham, terutama pada data *time series* yang masih didominasi pola linear dan dependensi temporal jangka pendek.

4.3. Analisis Performa Model LSTM

Pada model LSTM-20 menghasilkan performa terendah dibandingkan model lainnya dengan nilai MAPE sebesar 5.5099%. Kondisi tersebut mengindikasikan terjadinya *underfitting*, yaitu model belum mampu mempelajari pola data secara optimal akibat jumlah *epoch* pelatihan yang relatif terbatas.

Kurva *training loss* dan *validation loss* pada model LSTM-20 ditunjukkan pada Gambar 2. Grafik menunjukkan bahwa *validation loss* masih mengalami penurunan hingga *epoch* terakhir sehingga proses pelatihan dihentikan sebelum model mencapai kondisi konvergensi optimal.

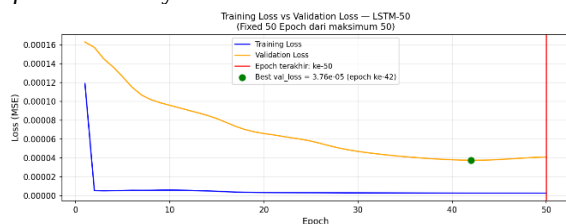


Gambar 2. Training Loss dan Validation Loss LSTM-20

Selisih yang cukup besar antara performa LSTM-20 dan model lainnya menunjukkan bahwa jumlah *epoch* memiliki pengaruh signifikan terhadap kemampuan model *deep learning* dalam mempelajari pola data *time series*.

Peningkatan jumlah *epoch* pada model LSTM-50 menghasilkan performa yang lebih baik dibandingkan LSTM-20. Model ini memperoleh nilai MAPE sebesar 3.7150% dengan penurunan *error* yang cukup signifikan pada seluruh metrik evaluasi.

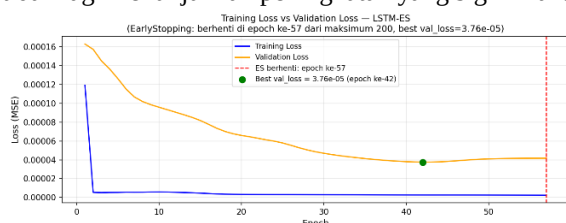
Kurva *training loss* dan *validation loss* pada model LSTM-50 ditunjukkan pada Gambar 3. *Validation loss* mengalami penurunan yang lebih stabil dan mencapai nilai terbaik pada *epoch* ke-42 sebelum mengalami kecenderungan stagnasi pada *epoch* berikutnya.



Gambar 3. Training Loss dan Validation Loss LSTM-50

Hasil tersebut menunjukkan bahwa peningkatan jumlah *epoch* memungkinkan model mempelajari pola temporal data secara lebih optimal sehingga menghasilkan prediksi yang lebih akurat dibandingkan LSTM-20.

Pada model LSTM-ES menghasilkan performa terbaik di antara seluruh varian LSTM dengan nilai MAPE sebesar 3.4719%. Penggunaan mekanisme *EarlyStopping* memungkinkan model menghentikan proses pelatihan secara otomatis ketika *validation loss* tidak lagi menunjukkan peningkatan yang signifikan.



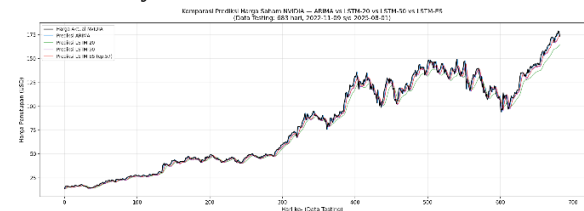
Gambar 4. Training Loss dan Validation Loss LSTM-ES

Kurva *training loss* dan *validation loss* pada model LSTM-ES ditunjukkan pada Gambar 4. Proses pelatihan berhenti secara otomatis pada *epoch* ke-57 dengan nilai *validation loss* terbaik sebesar 3.76×10^{-5} yang diperoleh pada *epoch* ke-42.

Hasil tersebut menunjukkan bahwa mekanisme *EarlyStopping* mampu membantu model mencapai performa optimal sekaligus meminimalkan risiko *overfitting* akibat proses pelatihan yang terlalu panjang.

4.4. Komparasi Hasil Prediksi Model

Komparasi hasil prediksi terhadap data aktual ditunjukkan pada Gambar 5. Grafik memperlihatkan bahwa model ARIMA menghasilkan kurva prediksi yang paling mendekati pola harga aktual dibandingkan model lainnya.



Gambar 5. Komparasi Prediksi Harga Saham Dengan Seluruh Model

Secara visual, model ARIMA mampu mengikuti pola fluktuasi harga saham dengan lebih stabil, terutama pada periode volatilitas tinggi. Sementara itu, model LSTM-20 menunjukkan deviasi yang lebih besar terhadap data aktual dibandingkan model lainnya.

Model LSTM-50 dan LSTM-ES menghasilkan pola prediksi yang relatif mendekati data aktual, meskipun masih menunjukkan *smoothing effect* pada beberapa titik perubahan harga yang tajam. Kondisi tersebut menunjukkan bahwa model *deep learning* cenderung menghasilkan prediksi yang lebih halus dibandingkan model statistik klasik.

4.5. Pembahasan Hasil Penelitian

Hasil penelitian menunjukkan bahwa kompleksitas model tidak selalu berbanding lurus dengan peningkatan akurasi prediksi. Meskipun LSTM memiliki kemampuan dalam mempelajari pola non-linear dan dependensi jangka panjang, model ARIMA justru menghasilkan performa lebih unggul pada dataset penelitian ini.

Dominasi performa ARIMA mengindikasikan bahwa dinamika harga saham NVIDIA pada periode observasi masih didominasi dependensi linear jangka pendek sehingga pendekatan *autoregressive* mampu memodelkan struktur temporal data secara lebih stabil dibandingkan pendekatan *deep learning* yang memerlukan kompleksitas pola non-linear lebih tinggi.

Selain itu, hasil penelitian juga menunjukkan bahwa strategi pelatihan memiliki pengaruh signifikan terhadap performa model *deep learning*. Jumlah *epoch* yang terlalu sedikit menyebabkan model mengalami *underfitting* sebagaimana terlihat pada model LSTM-20. Sebaliknya, peningkatan *epoch* dan penggunaan *EarlyStopping* mampu meningkatkan kemampuan generalisasi model sehingga menghasilkan performa yang lebih baik.

Temuan penelitian ini sejalan dengan beberapa penelitian terdahulu yang menyatakan bahwa model statistik klasik masih dapat menghasilkan performa kompetitif bahkan lebih unggul dibandingkan model *deep learning* pada data *time series* tertentu, khususnya ketika pola *autoregresif* masih mendominasi struktur data.

Secara keseluruhan, tiga model yaitu ARIMA, LSTM-50, dan LSTM-ES termasuk dalam kategori sangat akurat karena memiliki nilai MAPE di bawah 5%, sedangkan model LSTM-20 masih berada dalam kategori akurat dengan nilai MAPE sebesar 5.5099%.

Penelitian ini masih memiliki beberapa keterbatasan yang dapat menjadi pengembangan pada penelitian selanjutnya. Dataset yang digunakan hanya memanfaatkan *historical adjusted closing price* tanpa mempertimbangkan variabel eksternal seperti volume perdagangan, sentimen berita, kondisi makroekonomi, maupun indikator teknikal lainnya yang berpotensi memengaruhi pergerakan harga saham.

Selain itu, penelitian hanya menggunakan pendekatan *univariate forecasting* sehingga hubungan antar variabel eksternal belum dianalisis secara lebih mendalam. Penelitian selanjutnya dapat mengembangkan model *hybrid* atau *multivariate forecasting* untuk meningkatkan kemampuan prediksi terhadap pola pasar saham yang lebih kompleks.

5. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian yang telah dilakukan, metode *Autoregressive Integrated Moving Average* (ARIMA) dan *Long Short-Term Memory* (LSTM) mampu digunakan untuk melakukan prediksi harga saham NVIDIA Corporation periode 2012–2025 dengan tingkat akurasi yang baik. Hasil evaluasi menunjukkan bahwa model ARIMA(5,1,0) menghasilkan performa terbaik dibandingkan seluruh varian model LSTM dengan nilai MAE sebesar 1.9576, RMSE sebesar 3.0646, dan MAPE sebesar 2.3933%, sedangkan model LSTM-ES menjadi model *deep learning* terbaik dengan nilai MAPE sebesar 3.4719%. Temuan penelitian menunjukkan bahwa karakteristik data saham NVIDIA pada periode pengamatan masih didominasi pola *autoregresif* dan dependensi temporal jangka pendek sehingga model statistik klasik mampu menghasilkan prediksi yang lebih stabil dibandingkan model *deep learning* yang lebih kompleks. Selain itu, strategi pelatihan terbukti memengaruhi performa model LSTM, di mana peningkatan jumlah *epoch* dan penerapan *EarlyStopping* mampu meningkatkan kemampuan generalisasi model dan menurunkan nilai *error* prediksi. Penelitian selanjutnya disarankan untuk menggunakan pendekatan *multivariate forecasting* dengan menambahkan variabel eksternal seperti indikator teknikal, volume perdagangan, dan sentimen pasar, serta mengembangkan model *hybrid* untuk meningkatkan performa prediksi pada data *time series* dengan volatilitas tinggi.

DAFTAR PUSTAKA

- [1] Z. Yang and Z. Wang, "The Research of NVIDIA Stock Price Prediction Based on LSTM And ARIMA Model," *Highlights in Business, Economics and Management*, vol. 24, pp. 896–902, Jan. 2024, doi: 10.54097/dndydw34.
- [2] H. Evanthe, F. Fhadillah, R. Fauzan, N. Farezi, and F. Ramadhani, "Analisis Pola Dan Pergerakan Harga Saham Nvidia Dengan Klasterisasi K-Means Dan Prediksi Time-Series," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 9, no. 4, pp. 6043–6048, May 2025, doi: 10.36040/jati.v9i4.13992.
- [3] S. Kuang, "A Comparison of Linear Regression, LSTM model and ARIMA model in Predicting Stock Price A Case Study: HSBC's Stock Price," *BCP Business & Management*, vol. 44, pp. 478–488, Apr. 2023, doi: 10.54691/bcpbm.v44i.4858.
- [4] B. Zhang, "The Stock Price Forecasting Based on Time Series Model and Neural Network," *BCP Business & Management*, vol. 38, pp. 3423–3428, Mar. 2023, doi: 10.54691/bcpbm.v38i.4319.
- [5] A. D. Milniadi and N. O. Adiwijaya, "Analisis Perbandingan Model Arima Dan Lstm Dalam Peramalan Harga Penutupan Saham (Studi Kasus : 6 Kriteria Kategori Saham Menurut Peter Lynch)," *Sibatik Journal: Jurnal Ilmiah Bidang Sosial, Ekonomi, Budaya, Teknologi, dan Pendidikan*, vol. 2, no. 6, pp. 1683–1692, May 2023, doi: 10.54443/sibatik.v2i6.798.
- [6] Y. Zhang, "Stock Price Prediction Using ARIMA and LSTM Models: An Application to CSI 300 Closing Prices," *Highlights in Science, Engineering and Technology*, vol. 88, pp. 286–293, Mar. 2024, doi: 10.54097/7dgy1c14.
- [7] H. Wu, S. Chen, and Y. Ding, "Comparison of ARIMA and LSTM for Stock Price Prediction," *Financial Engineering and Risk Management*, vol. 6, no. 1, 2023, doi: 10.23977/ferm.2023.060101.
- [8] C. Wang, "Stock price forecasting by ARIMA, linear regression, LSTM and decomposition linear models," *Applied and Computational Engineering*, vol. 55, no. 1, pp. 61–68, Apr. 2024, doi: 10.54254/2755-2721/55/20241572.
- [9] L. Setiawan, D. Susanti, and R. Riaman, "Analisis Perbandingan Hasil Peramalan Harga Saham Menggunakan Model Autoregressive Integrated Moving Average dan Long Short Term Memory," *Jurnal Matematika Integratif*, vol. 19, no. 2, p. 223, Dec. 2023, doi: 10.24198/jmi.v19.n2.42164.223-234.
- [10] Moch Farryz Rizkilloh and Sri Widiyanesti, "Prediksi Harga Cryptocurrency Menggunakan Algoritma Long Short Term Memory (LSTM)," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 6, no. 1, pp. 25–31, Feb. 2022, doi: 10.29207/resti.v6i1.3630