



Evaluasi Model ARIMA, Random Forest dan LSTM untuk Peramalan Harga Minyak Dunia

Hendri Mahmud Nawawi^{1*}, Agung Baitul Hikmah², Ali Mustopa³, Fitra Septia Nugraha⁴

^{1,4} Informatika, Fakultas Teknologi Informasi, Universitas Nusa Mandiri, Indonesia

² Sistem Informasi Kampus Kota Tasikmalaya, Fakultas Teknik dan Informatika, Universitas Bina Sarana Informatika, Indonesia

³ Informatika Kampus Kota Pontianak, Fakultas Teknik dan Informatika, Universitas Bina Sarana Informatika, Indonesia

Info Artikel

Riwayat Artikel:

Diterima : **26 Februari 2026**

Direvisi : **05 Maret 2026**

Diterbitkan : **01 Juni 2026**

Kata Kunci:

ARIMA,
Minyak Mentah,
LSTM,
Peramalan,
Random Forest

Keywords:

ARIMA,
Crude Oil,
Forecasting,
LSTM,
Random Forest

ABSTRAK

Harga minyak dunia merupakan indikator ekonomi global yang sangat volatil dan dipengaruhi oleh berbagai faktor dinamis, sehingga diperlukan metode peramalan yang mampu menangkap pola linier maupun non-linier pada data deret waktu. Penelitian ini bertujuan mengevaluasi performa model ARIMA, Random Forest, dan LSTM dalam memprediksi harga minyak mentah West Texas Intermediate (WTI). Dataset yang digunakan berupa 1.316 data penutupan harian periode Januari 2020 hingga Desember 2024 yang diperoleh dari Investing.com. Data dibagi menggunakan pendekatan time-based split dengan rasio 80% untuk pelatihan dan 20% untuk pengujian guna menjaga karakteristik temporal. Evaluasi model dilakukan menggunakan MAE, MSE, RMSE, MAPE, dan R-squared. Hasil penelitian menunjukkan bahwa LSTM menghasilkan performa terbaik dengan MAE 2,914, RMSE 3,66, MAPE 0,854%, dan R^2 0,8514. Sebaliknya, ARIMA dan Random Forest menghasilkan R^2 negatif, menunjukkan keterbatasan dalam menangkap dinamika non-linier harga minyak. Temuan ini menegaskan efektivitas pendekatan deep learning dalam peramalan harga minyak.

ABSTRACT

Global crude oil prices are a highly volatile economic indicator influenced by various dynamic factors, requiring forecasting methods capable of capturing both linear and non-linear patterns in time series data. This study aims to evaluate the performance of ARIMA, Random Forest, and LSTM models in predicting West Texas Intermediate (WTI) crude oil prices. The dataset consists of 1,316 daily closing price records from January 2020 to December 2024 obtained from Investing.com. The data were divided using a time-based split approach with a ratio of 80% for training and 20% for testing to preserve temporal characteristics. Model evaluation was conducted using MAE, MSE, RMSE, MAPE, and R-squared. The results indicate that LSTM achieved the best performance with an MAE of 2.914, RMSE of 3.66, MAPE of 0.854%, and an R^2 of 0.8514. In contrast, ARIMA and Random Forest produced negative R^2 values, indicating limitations in capturing the non-linear dynamics of oil prices. These findings confirm the effectiveness of deep learning approaches for crude oil price forecasting.

Penulis Korespondensi:

Hendri Mahmud Nawawi,
Program Studi Informatika,
Universitas Nusa Mandiri
Email: hendri.hiw@nusamandiri.ac.id

This is an open access article under the [CC BY-SA](#) license



1. PENDAHULUAN

Harga minyak mentah merupakan salah satu indikator ekonomi global yang sangat berpengaruh, tidak hanya terhadap sektor energi, tetapi juga terhadap sektor lain seperti transportasi, manufaktur dan keuangan [1]. Fluktuasi harga minyak mentah dapat dipengaruhi oleh berbagai faktor, termasuk geopolitik, permintaan dan penawaran global, serta perubahan kebijakan energi [2]. Oleh karena itu, kemampuan untuk memprediksi harga minyak mentah menjadi sangat penting bagi para pengambil kebijakan, pelaku industri, serta para investor untuk mengantisipasi risiko dan merencanakan strategi yang efektif [3].

Peramalan harga minyak mentah telah menjadi topik penelitian yang menarik selama beberapa dekade [3]. Metode peramalan tradisional seperti *Autoregressive Integrated Moving Average* (ARIMA) telah lama digunakan untuk memodelkan data *time series* dalam berbagai bidang, termasuk peramalan harga komoditas [4]. ARIMA dikenal karena kemampuannya dalam menangkap pola linier dalam data historis, namun sering kali terbatas dalam menghadapi pola yang lebih kompleks atau non-linier [5].

Seiring dengan perkembangan teknologi dan *data science*, metode-metode peramalan berbasis *machine learning*, seperti Random Forest, mulai diterapkan dalam peramalan harga komoditas [4]. Random Forest, yang merupakan *ensemble learning method*, memiliki kemampuan untuk menangani kompleksitas data yang tinggi dengan menggabungkan hasil prediksi dari berbagai pohon keputusan (*decision trees*) [6]. Kemampuannya dalam menangani hubungan non-linier dan interaksi variabel model *random forest* menjadi algoritma yang efektif untuk peramalan harga minyak mentah [7].

Selain metode *machine learning*, model *deep learning* khususnya *Long Short-Term Memory* (LSTM), juga semakin banyak diterapkan dalam peramalan data *time series* [8]. LSTM merupakan pengembangan dari *Recurrent Neural Network* (RNN) yang dirancang untuk mampu menangkap ketergantungan jangka panjang (*long-term dependencies*) pada data deret waktu [9]. Keunggulan LSTM dalam memodelkan pola non-linier dan dinamika temporal yang kompleks menjadikannya sangat cocok untuk peramalan harga komoditas dengan fluktuasi tinggi seperti minyak mentah [10]. Beberapa penelitian terdahulu menunjukkan bahwa LSTM mampu menghasilkan performa peramalan yang lebih baik dibandingkan model statistik dan *machine learning* pada berbagai studi kasus tentang keuangan dan energi [11].

Penelitian tentang peramalan harga minyak menggunakan metode *Fuzzy Time Series* (FTS) dengan logika Singh untuk meramalkan harga minyak mentah dunia, yang menghasilkan nilai MAPE sangat rendah (0,30%), menunjukkan tingkat akurasi yang sangat tinggi. Pola peramalan yang dihasilkan juga konsisten dengan fluktuasi data aktual, sehingga metode ini dianggap sangat efektif untuk peramalan harga minyak mentah [3].

Penelitian kedua membandingkan model ARIMA, *Neural Network* dan *Neural Network* dengan Bootstrap dalam peramalan produksi minyak mentah di Indonesia. Hasilnya menunjukkan bahwa *Neural Network*, khususnya yang menggunakan pendekatan Bootstrap memberikan hasil peramalan yang lebih baik dalam menangkap pola non-linier dibandingkan ARIMA [1]. Sebaliknya, penelitian ketiga yang fokus pada penggunaan model ARIMA untuk meramalkan harga minyak mentah menemukan bahwa meskipun ARIMA sederhana dan efektif untuk data linier, model ini memiliki keterbatasan dalam menangani fluktuasi harga yang kompleks dan non-linier [12].

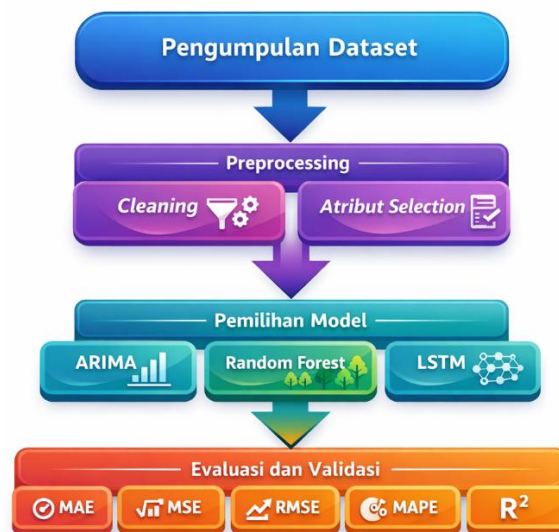
Penelitian lainnya dilakukan oleh [9] membandingkan model ARIMA dan *Long Short-Term Memory* (LSTM) dalam peramalan harga penutupan saham menggunakan enam kategori saham menurut Peter Lynch. Hasil penelitian menunjukkan bahwa performa model tidak hanya ditentukan oleh kompleksitas algoritma, tetapi juga oleh karakteristik data dan tujuan evaluasi yang digunakan. Meskipun LSTM dirancang untuk menangani pola non-linier dan dinamika temporal yang kompleks, hasil evaluasi berbasis RMSE, MAPE, dan waktu komputasi menunjukkan bahwa ARIMA lebih unggul secara rata-rata pada data saham yang diuji.

Meskipun berbagai metode peramalan harga minyak telah dikembangkan, sebagian besar penelitian masih berfokus pada satu atau dua pendekatan tertentu serta menggunakan periode data yang terbatas. Belum banyak studi yang mengevaluasi secara terpadu tiga paradigma pemodelan, yaitu statistik, *machine learning* dan *deep learning*, pada periode 2020–2024 yang ditandai volatilitas tinggi akibat pandemi COVID-19 dan dinamika geopolitik global. Kondisi tersebut menimbulkan kebutuhan akan evaluasi empiris yang mampu menguji konsistensi performa model pada karakteristik data yang fluktuatif dan non-linier.

Berdasarkan hal tersebut, penelitian ini bertujuan mengevaluasi performa model ARIMA, Random Forest, dan LSTM dalam peramalan harga minyak mentah *West Texas Intermediate* (WTI) menggunakan 1.316 data harian dari Investing.com periode Januari 2020 hingga Desember 2024. Novelty penelitian ini terletak pada evaluasi empiris tiga paradigma pemodelan dalam satu kerangka eksperimen yang seragam berbasis time-based split pada periode volatilitas tinggi, sehingga memungkinkan analisis komparatif yang lebih objektif terhadap kemampuan masing-masing pendekatan dalam menangkap dinamika harga minyak dunia. Hasil penelitian diharapkan memberikan kontribusi praktis bagi pengambilan keputusan sektor energi serta menjadi dasar pengembangan model prediksi yang lebih adaptif.

2. METODE PENELITIAN

Untuk menggambarkan alur atau metodologi penelitian supaya penelitian ini lebih terarah maka perlu perlu sebuah metode Penelitian, Metode penelitian ini menguraikan metode dan langkah-langkah yang digunakan dalam penelitian [13]. Berikut Gambar 1 menunjukkan alur metode penelitian yang digunakan.



Gambar 1. Metode Penelitian

2.1. Dataset Penelitian

Tahap awal adalah pengumpulan data yang relevan untuk mendukung penelitian ini. Dataset yang digunakan adalah data yang bersumber dari data publik pada halaman website Investing.com dengan jumlah data sebanyak 1316 data yang merupakan data histori harga minyak mentah dari *West Texas Intermediate* (WTI) selama periode 5 tahun terakhir dari 1 Januari 2020 hingga 31 Desember 2024. Data ini merupakan rentetan harga timeseries yang mencakup harga pembukaan, harga paling tinggi, harga paling rendah, harga penutupan harian dan volume penjualan minyak mentah. Atribut yang digunakan sebagai dasar dalam peramalan *time series* mengacu pada atribut tanggal dan harga penutupan.

2.2. Preprocessing

Setelah data dikumpulkan, data tersebut perlu diproses atau dibersihkan. *Preprocessing* mencakup berbagai kegiatan seperti penanganan data yang hilang, normalisasi, pengkodean variabel, dan transformasi data lainnya yang diperlukan untuk memastikan bahwa data siap untuk analisis model [14]. Tahap ini penting untuk meningkatkan kualitas data dan memastikan akurasi hasil analisis [15]

2.3. Pemilihan Model

Pada tahap ini, model analisis atau algoritma yang paling sesuai dipilih untuk menganalisis data. Penelitian ini mengusulkan model regresi statistik ARIMA, *Machine Learning Random Forest* dan *Deep Learning* (LSTM) untuk melakukan evaluasi terhadap dataset yang diteliti.

1. *AutoRegressive Integrated Moving Average* (ARIMA) model adalah salah satu teknik peramalan time series yang digunakan untuk memodelkan data yang menunjukkan hubungan linear antara nilai masa lalu (lag) dari suatu variabel dan *error terms* [16].
2. Random Forest, Algoritma *Random Forest* merupakan algoritma *ensemble learning* berbasis pohon keputusan yang bekerja dengan membangun banyak *decision tree* secara acak dan menggabungkan hasil prediksinya. Pendekatan ini mampu mengurangi *overfitting* serta menangkap hubungan non-linier dan interaksi kompleks dalam data, sehingga efektif digunakan untuk peramalan nilai kontinu pada permasalahan regresi, termasuk peramalan data deret waktu [17].
3. *Long Short-Term Memory* (LSTM)
LSTM merupakan pengembangan dari *Recurrent Neural Network* (RNN) yang dirancang untuk memodelkan data deret waktu dengan kemampuan menangkap ketergantungan jangka pendek dan jangka panjang [18]. LSTM menggunakan mekanisme *gating* (*input gate*, *forget gate*, dan *output gate*) untuk mengontrol aliran informasi, sehingga mampu mengatasi permasalahan *vanishing gradient* dan efektif dalam mempelajari pola non-linier pada data time series yang kompleks [19].

2.4. Evaluasi dan Validasi

Setelah model dipilih dan diterapkan pada data, hasil model tersebut dievaluasi dan divalidasi. Evaluasi ini dilakukan menggunakan berbagai metrik.

1. *Mean Absolute Error* (MAE), menghitung rata-rata dari selisih absolut antara nilai aktual dan nilai prediksi [20]. Nilai MAE yang lebih kecil menunjukkan performa model yang lebih baik [2].
Persamaan MAE:

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

2. *Mean Squared Error* (MSE), mengukur rata-rata kuadrat selisih antara nilai yang diprediksi dan nilai sebenarnya [21]. Berikut rumus MSE pada persamaan 2.

$$MSE = \frac{\sum (y_i - \hat{y}_i)^2}{n} \quad (2)$$

3. *Root Mean Square Error* (RMSE) adalah aturan penilaian kuadrat yang juga mengukur besarnya rata-rata kesalahan [22]. Persamaan RMSE:

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

4. *Mean Absolute Percentage Error* (MAPE), digunakan untuk mengukur tingkat kesalahan prediksi dalam skala persentase [12]. Persamaan MAPE:

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{\hat{y}_i} \right| \quad (4)$$

5. R-squared (R^2), digunakan untuk mengukur kemampuan model dalam menjelaskan variasi data aktual, di mana nilai R^2 yang mendekati 1 menunjukkan performa model yang baik, sedangkan nilai mendekati 0 atau negatif menunjukkan kemampuan prediksi yang rendah [23]. Persamaan R-Squared:

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

Dimana:

y_i : nilai aktual

$\hat{y}_i$: nilai prediksi

$\bar{y}$: nilai rata-rata dari data aktual

n : jumlah data

$|\cdot|$: nilai absolut.

3. HASIL DAN ANALISIS

3.1. Analisis Dataset Penelitian

Bagian ini menyajikan hasil penelitian mengenai evaluasi performa model Random Forest dan ARIMA dalam peramalan harga minyak dunia. Pembahasan dilakukan berdasarkan tahapan penelitian yang meliputi pengumpulan dataset, preprocessing, pemilihan model, serta proses evaluasi dan validasi untuk menilai akurasi prediksi masing-masing metode.

Dataset yang digunakan merupakan data historis harga minyak mentah West Texas Intermediate (WTI) periode 2020–2024 dengan interval harian, yang diperoleh dari sumber data sekunder [24]. Total data yang digunakan sebanyak 1.316 observasi, dengan variabel utama berupa harga penutupan (closing price) sebagai target peramalan, karena dianggap merepresentasikan nilai harga minyak pada setiap periode perdagangan. Berikut Tabel 1 Dataset penelitian yang sudah di dapatkan berdasarkan histori harian dari harga minyak mentah WTI.

Tabel 1. Data Histori Harga Minyak Mentah WTI

	<i>Date</i>	<i>Close</i>	<i>Open</i>	<i>Hight</i>	<i>Low</i>	<i>Vol.</i>	<i>Change</i>
0	12/31/2024	71.25	70.69	71.52	70.43	63,63K	1,02%
1	12/30/2024	70.53	70	71.05	69.75	93,10K	0,50%
2	12/27/2024	70.18	69.27	70.32	69.06	79,13K	1,37%
3	12/26/2024	69.23	69.72	70.25	68.94	69,64K	-0,59%
4	12/24/2024	69.64	69.04	69.94	68.95	40,39K	1,21%
...	...	...	...	...	...	...	...
1311	1/7/2020	62.51	62.67	62.92	61.96	173,85K	-0,84%
1312	1/6/2020	63.04	63.43	64.45	62.43	176,99K	0,35%
1313	1/3/2020	62.82	60.96	63.8	60.91	236,06K	3,07%
1314	1/2/2020	60.95	61.29	61.32	60.42	101,52K	-0,62%
1315	1/1/2020	61.33	61.52	61.54	61.33	NaN	0,92%

Penjelasan dari masing-masing atribut dijelaskan pada Tabel 2

Tabel 2. Deskripsi Atribut Dataset Penelitian

Kolom	Penjelasan
<i>Date</i>	Tanggal pencatatan harga minyak (format YYYY-MM-DD)
<i>Close</i>	Harga penutupan minyak WTI pada hari tersebut (USD)
<i>Open</i>	Harga pembukaan pada awal perdagangan
<i>Hight</i>	Harga tertinggi yang tercapai pada hari tersebut
<i>Low</i>	Harga terendah yang tercapai pada hari tersebut
<i>Vol.</i>	Volume transaksi harian dalam format teks (K/M)
<i>Change</i>	Persentase perubahan harga dari hari sebelumnya

Setelah dataset di deskripsikan langkah selanjutnya adalah melakukan *preprocessing*. Tahap *preprocessing* dilakukan untuk memastikan data yang digunakan layak dan sesuai dengan kebutuhan pemodelan. Karena pokok pembahasan pada penelitian ini menggunakan analisis time series maka atribut yang digunakan adalah tanggal dan harga *close*/penutupan. Pada tahap preprocessing diantaranya:

1. Data Cleaning

Pada tahap ini dilakukan membersihkan data atau *data cleaning* dengan meliputi pembersihan data sebagai berikut:

- Konversi kolom tanggal menjadi format *datetime* agar dapat diproses sebagai data deret waktu (*time series*).
- Konversi data numerik yang masih bertipe string menjadi tipe numerik (*float*), termasuk penanganan tanda pemisah ribuan dan desimal.
- Penghapusan data yang memiliki nilai kosong (*missing value*) pada variabel harga penutupan.
- Pengurutan data berdasarkan urutan waktu dari yang paling lama ke yang terbaru.

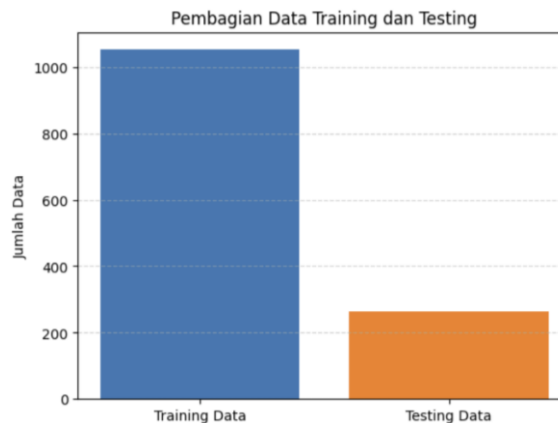
Tahap ini bertujuan untuk menghilangkan potensi kesalahan perhitungan dan memastikan konsistensi data sebelum masuk ke proses pemodelan.

2. Seleksi Atribut

Pada penelitian ini digunakan pendekatan univariat, yaitu hanya memanfaatkan variabel harga penutupan sebagai atribut utama. Pemilihan ini dilakukan untuk:

- Menyederhanakan proses pemodelan
- Memfokuskan penelitian pada evaluasi performa metode,
- Menjaga kesetaraan antara model ARIMA dan Random Forest dalam menerima input data.

Selanjutnya adalah memodelkan dataset untuk mendapatkan hasil paling optimal dari model-model yang diuji, namun sebelum pemodelan dataset terlebih dahulu dibagi dua menjadi data training dan data testing dengan pembagian 80:20.



Gambar 3. Pembagian Dataset

Pada gambar 3 Sebanyak 80% data digunakan sebagai data training untuk membangun dan melatih model peramalan, sedangkan 20% data sisanya digunakan sebagai data testing untuk menguji performa model terhadap data yang tidak dilibatkan dalam proses pelatihan. Pembagian data dilakukan berdasarkan urutan waktu (*time-based split*) untuk menjaga karakteristik deret waktu dan menghindari kebocoran informasi (*data leakage*).

Tabel 3. Sample Dataset Penelitian

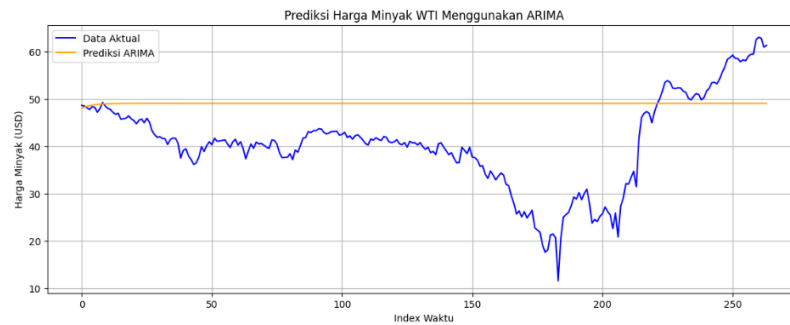
Date	Close
12/31/2024	71.25
12/30/2024	70.53
12/27/2024	70.18
12/26/2024	69.23
12/24/2024	69.64
...	...
1/7/2020	62.51
1/6/2020	63.04
1/3/2020	62.82
1/2/2020	60.95
1/1/2020	61.33

Setelah dataset dibagi ke dalam data latih dan data uji, setiap model peramalan diuji untuk menilai kemampuannya dalam memprediksi harga minyak dunia berdasarkan data historis.

3.2. Analisis Performa Model

Evaluasi performa dilakukan sebagai langkah penting untuk membandingkan efektivitas pendekatan statistik, *machine learning* dan *deep learning* yang digunakan dalam penelitian ini. Beberapa metrik evaluasi yang digunakan meliputi MAE, MSE, RMSE, MAPE, dan R-squared, yang masing-masing mencerminkan tingkat kesalahan prediksi dan kemampuan model dalam menjelaskan variasi data aktual. Hasil pengujian performa ketiga model adalah sebagai berikut:

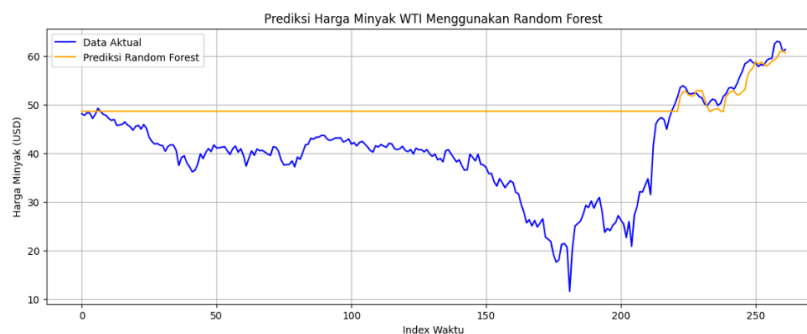
1. Hasil prediksi model ARIMA



Gambar 4. Prediksi Harga Minyak dengan ARIMA

Pada gambar 4 grafik prediksi harga minyak WTI menggunakan model ARIMA menunjukkan bahwa hasil prediksi cenderung membentuk pola yang relatif konstan dan hanya mengikuti nilai rata-rata tren data. Meskipun model mampu menangkap kecenderungan umum pergerakan harga, terlihat adanya jarak yang cukup besar antara nilai prediksi dan data aktual pada periode dengan fluktuasi harga yang tinggi. Hal ini mengindikasikan keterbatasan model ARIMA dalam menangkap perubahan harga yang bersifat non-linier dan volatilitas tajam pada data harga minyak dunia.

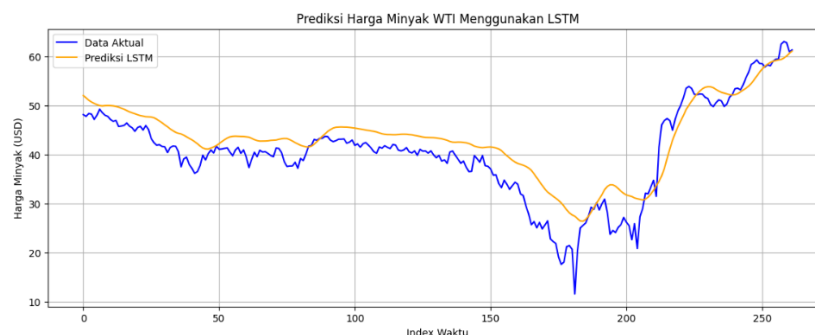
2. Hasil prediksi model Random Forest



Gambar 5. Prediksi Harga Minyak dengan Random Forest

Pada Gambar 5 grafik prediksi harga minyak WTI menggunakan model *Random Forest* menunjukkan bahwa hasil prediksi cenderung mengikuti nilai rata-rata pergerakan harga dan menghasilkan pola yang lebih halus dibandingkan data aktual. Meskipun pada beberapa periode prediksi mendekati nilai aktual, terlihat bahwa model kurang responsif terhadap fluktuasi harga yang tajam. Hal ini menunjukkan bahwa Random Forest, pada pendekatan univariat yang digunakan, belum sepenuhnya mampu menangkap dinamika temporal dan volatilitas tinggi pada pergerakan harga minyak dunia.

3. Hasil prediksi model LSTM



Gambar 6. Prediksi Harga Minyak dengan LSTM

Gambar 6 prediksi harga minyak WTI menggunakan model LSTM menunjukkan bahwa hasil prediksi mampu mengikuti pola pergerakan data aktual dengan lebih baik dibandingkan model lainnya. Garis prediksi

LSTM terlihat lebih responsif terhadap perubahan tren dan mampu menyesuaikan diri pada periode fluktuasi harga yang tinggi, sehingga selisih antara data aktual dan hasil prediksi relatif lebih kecil. Hal ini menunjukkan keunggulan LSTM dalam menangkap hubungan non-linier dan ketergantungan waktu jangka panjang pada data harga minyak dunia

Evaluasi performa ketiga model pada nilai yang dijadikan pembandingan secara detail pada tabel 3.

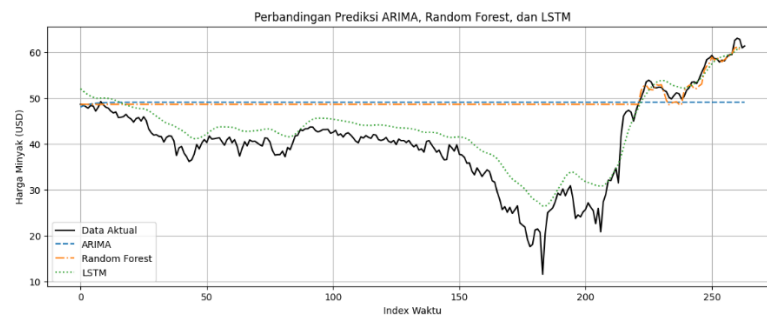
Tabel 3. Evaluasi Performa Model

Model	MAE	MSE	RMSE	MAPE (%)	R_squared
ARIMA	10.441574	162.999643	12.767131	33.033178	-0.808061
Random Forest	9.429672	147.937138	12.162941	31.004307	-0.637469
LSTM	2.916466	13.423989	3.663876	8.854504	0.851414

Berdasarkan hasil evaluasi performa pada Tabel 3, terlihat perbedaan yang jelas pada kinerja model ARIMA, Random Forest, dan Long Short-Term Memory (LSTM) dalam meramalkan harga minyak mentah WTI. Model LSTM menghasilkan nilai MAE, MSE, RMSE, dan MAPE yang paling rendah serta nilai R-squared (R^2) sebesar 0,8514, yang menunjukkan kemampuannya dalam menjelaskan sebagian besar variasi data harga minyak dunia. Nilai R^2 yang positif dan tinggi mengindikasikan bahwa LSTM lebih efektif dalam menangkap pola pergerakan harga yang bersifat non-linier dan memiliki ketergantungan waktu jangka panjang.

Sebaliknya, model ARIMA dan Random Forest menghasilkan nilai R-squared yang negatif, yang menunjukkan keterbatasan dalam merepresentasikan dinamika harga minyak dunia secara optimal pada pendekatan yang digunakan. Random Forest menunjukkan performa yang lebih baik dibandingkan ARIMA dengan tingkat kesalahan prediksi yang lebih rendah, namun masih belum mampu menangkap ketergantungan temporal secara memadai. Sementara itu, ARIMA sebagai model statistik linier menunjukkan performa terendah, terutama pada data harga minyak yang memiliki volatilitas tinggi, sehingga hasil ini menegaskan keunggulan pendekatan deep learning dibandingkan model statistik dan machine learning konvensional dalam peramalan harga minyak dunia.

Jika digambarkan secara keseluruhan perbandingan ketiga model adalah sebagai berikut



Gambar 7. Perbandingan Model ARIMA, Random Forest dan LSTM

Pada Gambar 7 grafik perbandingan prediksi menunjukkan perbedaan kemampuan yang jelas antara ketiga model dalam merepresentasikan pergerakan harga minyak dunia WTI. Model ARIMA dan Random Forest menghasilkan prediksi yang relatif datar dan cenderung mengikuti nilai rata-rata, sehingga kurang mampu menyesuaikan diri terhadap fluktuasi harga yang tajam. Sebaliknya, model LSTM menunjukkan kemampuan yang lebih baik dalam mengikuti pola data aktual, terutama pada periode volatilitas tinggi dan perubahan tren yang signifikan. Hasil ini menjelaskan bahwa LSTM lebih efektif dalam menangkap hubungan non-linier dan ketergantungan temporal jangka panjang dibandingkan model statistik dan *machine learning* konvensional, serta sejalan dengan hasil evaluasi metrik yang menunjukkan keunggulan LSTM secara kuantitatif

4. KESIMPULAN

Berdasarkan hasil analisis dan evaluasi performa, penelitian ini menyimpulkan bahwa peramalan harga minyak mentah WTI tidak dapat direpresentasikan secara optimal menggunakan pendekatan linier maupun machine learning konvensional secara univariat. Model *Long Short-Term Memory* (LSTM) terbukti memiliki performa paling baik dibandingkan ARIMA dan Random Forest, yang ditunjukkan oleh nilai MAE, MSE, RMSE, dan MAPE yang lebih rendah, serta nilai R-squared sebesar 0.8514. Hasil ini menunjukkan bahwa LSTM mampu menangkap pola non-linier dan ketergantungan temporal jangka panjang yang kuat pada data harga minyak dunia, sehingga secara efektif menjawab permasalahan penelitian yang diungkapkan pada bagian pendahuluan.

Secara umum, temuan penelitian ini menguatkan hasil penelitian terdahulu yang menyatakan bahwa pendekatan deep learning lebih sesuai untuk memodelkan data deret waktu dengan volatilitas tinggi dibandingkan model statistik dan *machine learning*. Keterbatasan yang ditunjukkan oleh model ARIMA dan Random Forest mengindikasikan perlunya pendekatan yang mampu mengakomodasi dinamika temporal yang kompleks. Oleh karena itu, penelitian ini merekomendasikan pengembangan lanjutan menggunakan pendekatan multivariat dengan penambahan variabel eksternal, seperti faktor makroekonomi dan geopolitik, atau penerapan model hybrid guna meningkatkan akurasi peramalan dan mendukung pengambilan keputusan di sektor energi dan ekonomi global.

REFERENSI

- [1] R. Erison, B. R. AMS, P. Ayu, and N. A. Sufiawan, "Pengaruh Harga Emas Dunia, Harga Minyak Mentah Dunia Dan Pasar Saham Di Indonesia," *J. Ekon. Pembang. STIE Muhammadiyah Palopo*, vol. 9, no. 1, p. 44, 2023, doi: 10.35906/jep.v9i1.1388.
- [2] D. Fakhriyana and I. I. Brilliant, "Penerapan Metode Fuzzy Time Series (FTS) Cheng dan Markov-Chain untuk Peramalan Indonesia Crude Oil Price (ICP)," *Indones. J. Appl. Stat.*, vol. 6, no. 1, p. 44, 2024, doi: 10.13057/ijas.v6i1.79907.
- [3] S. Hasibuan, Y. Asdi, and A. Nazra, "Peramalan Harga Minyak Mentah Dunia Menggunakan Metode Fuzzy Time Series Logika Singh," *J. Mat. UNAND*, vol. 13, no. 1, p. 66, 2024, doi: 10.25077/jmua.13.1.66-74.2024.
- [4] P. Kinerja *et al.*, "Comparison of the performance of the arima , multi-layer perceptron , and random forest methods in forecasting precious metal futures prices that contain outliers," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 11, no. 2, 2024, doi: 10.25126/jtiik.20241127392.
- [5] S. Wijaya and Fauziah, "a Comperative Analysis of Linear Regression and Random Forest Methods for Predicting Palm Oil Price Sales," *J. Ilm. Teknol. dan Rekayasa*, vol. 28, no. 3, pp. 265–274, 2023, doi: 10.35760/tr.2023.v28i3.10111.
- [6] A. Ubaidillah, E. M. S. Rochman, D. A. Fatah, and A. Rachmad, "Classification of Corn Diseases using Random Forest, Neural Network, and Naive Bayes Methods," *J. Phys. Conf. Ser.*, vol. 2406, no. 1, pp. 0–10, 2022, doi: 10.1088/1742-6596/2406/1/012023.
- [7] L. M. Cendani and A. Wibowo, "Perbandingan Metode Ensemble Learning pada Klasifikasi Penyakit Diabetes," *J. Masy. Inform.*, vol. 13, no. 1, pp. 33–44, 2022.
- [8] N. Sirimevan, I. G. U. H. Malmagaha, C. Jayasekara, Y. S. Mayuran, and C. Jayawardena, "Stock Market Prediction Using Machine Learning Techniques," *2019 Int. Conf. Adv. Comput. ICAC 2019*, pp. 192–197, 2019, doi: 10.1109/ICAC49085.2019.9103381.
- [9] A. D. Milniadi and N. O. Adiwijaya, "ANALISIS PERBANDINGAN MODEL ARIMA DAN LSTM DALAM PERAMALAN HARGA PENUTUPAN SAHAM (STUDI KASUS: 6," *SIBATIK J.*, vol. 2, no. 6, 2023, doi: <https://doi.org/10.54443/sibatik.v2i6.798>.
- [10] C. Zhang, Z. Chen, and J. Zhou, "Research on Short-Term Load Forecasting Using K-means Clustering and CatBoost Integrating Time Series Features," *Chinese Control Conf. CCC*, vol. 2020-July, no. July 2016, pp. 6099–6104, 2020, doi: 10.23919/CCC50068.2020.9188856.

- [11] F. Liantoni and O. Simanjuntak, "Peramalan Nilai Saham BBKA Melalui Pendekatan Time Series Menggunakan Teknik Exponential Smoothing," *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 10, no. 3, pp. 259–266, 2025.
- [12] S. Nurlela, A. Fanani, and Hani Khaulasari, "Prediksi Harga Minyak Mentah WTI Menggunakan Metode Fuzzy Time Series Markov Chain," *J. Fourier*, vol. 12, no. 1, pp. 10–19, 2023, doi: 10.14421/fourier.2023.121.10-19.
- [13] K. T. Hidayah, B. Arifitama, and S. D. H. Permana, "Klasifikasi Penyakit Kanker Serviks Berdasarkan Kebiasaan dan Rekam Medis dengan Metode C4.5," *J. Nas. Teknol. dan Sist. Inf.*, vol. 10, no. 1, pp. 36–44, 2024, doi: 10.25077/teknosi.v10i1.2024.36-44.
- [14] R. Novia, S. S. Prasetyowati, and Y. Sibaroni, "Identify User Behavior Based on The Type of Tweet on Twitter Platform Using Gaussian Mixture Model Clustering," *J. Comput. Syst. Informatics*, vol. 3, no. 4, pp. 502–506, 2022, doi: 10.47065/josyc.v3i4.2208.
- [15] S. Khotimatul Wildah, A. Latif, S. Agustiani, A. Mustopa, and S. Suharyanto, "Color Histogram Dan Support Vectore Machine Untuk Mengklasifikasikan Biji Kopi Berdasarkan Tingkat Pemanggang," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 8, no. 1, pp. 580–586, 2024, doi: 10.36040/jati.v8i1.8724.
- [16] A. Rizki, "Aplikasi Model ARIMA dalam Peramalan Data Harga Emas Dunia Tahun 2010-2022," *J. Stat. dan Apl.*, vol. 7, no. 1, pp. 84–92, 2023, doi: 10.21009/jsa.07108.
- [17] N. Nurussakinah, M. Faisal, and I. B. Santoso, "Algoritma Random Forest dan Synthetic Minority Oversampling Technique (SMOTE) untuk Deteksi Diabetes," *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 10, no. 2, pp. 221–234, 2025, doi: 10.14421/jiska.2025.10.2.221-234.
- [18] Z. Wang and L. Jia, "Short-Term Photovoltaic Power Generation Prediction Based on LightGBM-LSTM Model," *2020 5th Int. Conf. Power Renew. Energy, ICPRE 2020*, pp. 543–547, 2020, doi: 10.1109/ICPRE51194.2020.9233298.
- [19] H. Harumy, M. Zarlis, M. S. Lydia, and S. Effendi, "Pengembangan model protis neural network untuk prediksi dan klasifikasi data timeseries dan image," *Pros. Semin. Nas. Ris. Dan Inf. Sci. (*, vol. 4, pp. 1–7, 2022.
- [20] R. Komansilan, V. Tarigan, and A. Yusupa, "Analisis Perbandingan Metode Trend Moment dan Regresi Linear Untuk Meramal Harga Saham Bank BRI," *J-SISKO TECH (Jurnal Teknol. Sist. Inf. dan Sist. Komput. TGD)*, vol. 7, no. 1, p. 24, 2024, doi: 10.53513/jsk.v7i1.9474.
- [21] T. Aravind, B. S. Reddy, S. Avinash, and G. Jeyakumar, "Information of Under Graduate Students," *Proc. Third Int. Conf. I-SMAC*, pp. 542–546, 2019.
- [22] R. Julian and M. R. Pribadi, "Peramalan Harga Saham Pertambangan Pada Bursa Efek Indonesia (BEI) Menggunakan Long Short Term Memory (LSTM)," *J. Tek. Inform. dan Sist. Inf.*, vol. 8, no. 3, pp. 1570–1580, 2021.
- [23] S. B. Koduri, L. Guniseti, C. R. Ramesh, K. Mutyalu, and D. Ganesh, "Prediction of crop production using adaboost regression method Prediction of crop production using adaboost regression method," *J. Phys. Conf. Ser.*, 2019, doi: 10.1088/1742-6596/1228/1/012005.
- [24] Investing.com, "Minyak Mentah WTI Berjangka - Jan '26 (OIL)," Online. [Online]. Available: <https://id.investing.com/commodities/crude-oil-historical-data>