

PENERAPAN METODE *LONG SHORT-TERM MEMORY* UNTUK PREDIKSI PRODUKSI MINYAK BUMI BERDASARKAN DATA *TIME SERIES* HARIAN

Wardah Shifa Maharani, Dinna Yunika Hardiyanti

Sistem Informasi, Universitas Sriwijaya
Jalan Srijaya Negara, Bukit Lama, Kota Palembang, Indonesia
09031282227111@student.unsri.ac.id

ABSTRAK

Produksi minyak bumi merupakan komponen strategis dalam sektor energi sehingga membutuhkan sistem prediksi yang akurat untuk mendukung perencanaan dan pengambilan keputusan. Seiring waktu, produksi minyak menunjukkan pola fluktuatif dan memiliki ketergantungan jangka panjang yang menjadikan proses prediksi sebagai tantangan tersendiri. Kondisi tersebut menimbulkan permasalahan dalam menentukan metode yang mampu memodelkan pola *time series* data produksi secara efektif dengan tingkat kesalahan yang rendah. Oleh karena itu, penelitian ini bertujuan untuk menerapkan metode *Long Short-Term Memory* (LSTM) sebagai metode dalam memprediksi produksi minyak bumi berdasarkan data *time series* harian. Proses penelitian diawali dengan tahapan pra-pemrosesan data yang meliputi pembersihan data, normalisasi, serta pembagian data. Selanjutnya, model LSTM dilatih menggunakan data produksi periode 1 Januari 2022 hingga 31 Juli 2025 dan dievaluasi menggunakan metrik *Root Mean Square Error* (RMSE) dan *Mean Absolute Percentage Error* (MAPE). Hasil evaluasi menunjukkan nilai RMSE sebesar 100,49 dan MAPE sebesar 1,72% yang mengindikasikan tingkat kesalahan prediksi rendah dan akurasi yang sangat baik. Dengan demikian, metode LSTM mampu mempertahankan pola fluktuasi produksi serta menghasilkan prediksi yang selaras dengan tren historis sehingga layak digunakan sebagai pendukung pemantauan dan perencanaan produksi minyak bumi.

Kata kunci : LSTM, MAPE, Minyak Bumi, Prediksi, RMSE, Time Series

1. PENDAHULUAN

Minyak bumi hingga saat ini masih memegang peranan penting sebagai sumber energi dunia, meskipun secara global upaya transisi menuju energi terbarukan terus mengalami peningkatan. Selain berfungsi sebagai sumber bahan bakar, minyak bumi juga merupakan komoditas strategis yang digunakan sebagai bahan baku berbagai sektor industri. Oleh karena itu, perubahan tingkat produksi minyak bumi memiliki implikasi langsung terhadap ketahanan energi, stabilitas ekonomi, dan dinamika geopolitik suatu negara [1], [2]. Dalam konteks tersebut, ketersediaan informasi prediksi produksi yang akurat menjadi faktor penting dalam mendukung perencanaan operasional perusahaan migas serta perumusan kebijakan energi jangka menengah dan panjang [3], [4].

Seiring dengan meningkatnya kompleksitas pengelolaan lapangan minyak, karakteristik produksi minyak bumi juga menjadi semakin dinamis. Produksi minyak dipengaruhi oleh berbagai faktor, seperti laju deplesi reservoir, perubahan kondisi geologi, serta fluktuasi pasar energi global [5]. Kombinasi faktor-faktor tersebut menyebabkan data produksi minyak bumi bersifat fluktuatif dan non-linier. Kondisi ini menimbulkan tantangan tersendiri bagi metode prediksi konvensional, seperti *Autoregressive Integrated Moving Average* (ARIMA) dan *Decline Curve Analysis* (DCA) yang cenderung kurang mampu menangkap pola non-linier dan dinamika kompleks pada data produksi [6], [7].

Selain keterbatasan metode, tantangan prediksi produksi minyak bumi juga berkaitan dengan jenis

data yang digunakan. Sebagian besar penelitian terdahulu masih memanfaatkan data produksi dalam skala bulanan atau tahunan. Padahal, fluktuasi produksi harian sering kali merepresentasikan kondisi lapangan yang krusial bagi pengambilan keputusan operasional. Penggunaan data yang teragregasi berpotensi mengaburkan variasi jangka pendek yang penting dalam strategi pengelolaan produksi [4], [8]. Pemanfaatan data historis harian untuk prediksi produksi minyak bumi, khususnya di Indonesia, hingga saat ini masih relatif terbatas sehingga menunjukkan adanya kesenjangan penelitian yang perlu dikaji lebih lanjut.

Keterbatasan tersebut juga tercermin dalam praktik operasional di PT P, di mana strategi produksi masih banyak mengandalkan keahlian praktisi serta evaluasi data historis secara umum. Tidak tersedianya sistem prediksi jangka panjang berbasis data harian menyebabkan perencanaan operasional lapangan dihadapkan pada tingkat ketidakpastian yang tinggi, terutama dalam menghadapi kondisi lapangan yang dinamis dan kompleks. Akibatnya, efektivitas dan efisiensi pengelolaan produksi minyak bumi berpotensi mengalami penurunan [1], [3].

Berdasarkan kondisi tersebut, diperlukan pendekatan prediksi yang mampu menangani karakteristik data produksi minyak bumi yang fluktuatif, non-linier, serta memiliki ketergantungan jangka panjang. Salah satu metode yang relevan untuk menjawab kebutuhan tersebut adalah *Long Short-Term Memory* (LSTM). LSTM merupakan arsitektur *deep learning* yang dirancang untuk memodelkan data *time series* dengan kemampuan mengenali pola non-linier

dan mengatasi permasalahan *long-term dependency* [9], [10], [11]. Keunggulan ini menjadikan LSTM lebih sesuai untuk diterapkan pada data produksi minyak bumi berbasis data *time series* harian [10], [11].

Dalam penelitian ini, metode LSTM diterapkan pada data *time series* produksi harian minyak bumi periode 1 Januari 2022 hingga 31 Juli 2025. Proses penelitian diawali dengan tahap pra-pemrosesan data yang meliputi pembersihan data, normalisasi, serta pembagian data. Selanjutnya, model LSTM dilatih dan dievaluasi menggunakan metrik *Root Mean Squared Error* (RMSE) dan *Mean Absolute Percentage Error* (MAPE). RMSE digunakan untuk mengukur besarnya kesalahan prediksi dalam satuan asli data, sedangkan MAPE memberikan gambaran kesalahan relatif dalam bentuk persentase. Penggunaan kedua metrik tersebut secara bersamaan memungkinkan evaluasi kinerja model yang lebih komprehensif karena masing-masing memberikan perspektif berbeda terhadap kesalahan prediksi [12], [13], [14]. Melalui metode ini, penelitian diharapkan mampu menghasilkan prediksi produksi minyak bumi yang akurat serta memberikan kontribusi praktis berupa proyeksi produksi jangka panjang yang dapat mendukung perencanaan dan pengambilan keputusan operasional perusahaan.

2. TINJAUAN PUSTAKA

2.1. Penelitian Terdahulu

Penelitian prediksi produksi minyak bumi telah berkembang dari metode klasik seperti ARIMA dan Decline Curve Analysis yang mudah diinterpretasikan namun kurang mampu menangani pola non-linier dan fluktuasi tinggi, sebagaimana ditunjukkan oleh temuan [7], [12]. Pendekatan machine learning kemudian diperkenalkan untuk meningkatkan akurasi, dan studi oleh [3], [4] menunjukkan bahwa model regresi lebih adaptif, meskipun belum optimal dalam menangkap ketergantungan jangka panjang. Kemajuan signifikan muncul melalui deep learning, terutama LSTM, yang konsisten terbukti lebih akurat dibandingkan ARIMA, ANN, maupun Prophet [3], [10], [11].

Varian lanjutan seperti SA-LSTM [15] juga memperlihatkan peningkatan performa, walaupun LSTM murni tetap kompetitif. Di Indonesia, penelitian terkini menunjukkan dominasi LSTM dibandingkan metode lain dalam memprediksi data energi [9], [16], namun sebagian besar masih terbatas pada data bulanan atau tahunan. Kondisi ini menegaskan adanya celah penelitian untuk penggunaan LSTM pada data harian produksi minyak bumi, terutama di tingkat lapangan, guna menghasilkan prediksi yang lebih sensitif terhadap dinamika operasional dan mendukung perencanaan produksi jangka panjang.

2.2. Minyak Bumi

Minyak bumi tetap menjadi sumber energi global yang penting meskipun transisi menuju energi terbarukan terus meningkat [2], [17]. Selain menjadi bahan baku utama industri petrokimia, minyak bumi

juga memasok kebutuhan transportasi dan pembangkit listrik sehingga memiliki peran strategis dalam geopolitik, ketahanan energi, dan stabilitas ekonomi [3], [4]. Di banyak negara penghasil minyak dan gas, fluktuasi produksinya berpengaruh langsung terhadap pendapatan negara, inflasi, serta harga barang dan jasa [12]. Perubahan kecil dalam volume produksi dapat memicu pergerakan harga global sehingga pengelolaan produksinya mencakup pertimbangan teknis dan strategi kebijakan energi [2].

Dinamika produksi dipengaruhi oleh kondisi reservoir, faktor geologi, metode eksploitasi, serta pasar energi dunia, sehingga tren produksinya cenderung fluktuatif dan sulit diprediksi [5], [8]. Di Indonesia sendiri, cadangan minyak terus menurun sehingga menjaga keberlanjutan produksi menjadi tantangan penting bagi kebutuhan energi domestik dan ketahanan energi nasional [1], [18]. Ketidakpastian harga minyak global turut memperkuat urgensi perencanaan produksi yang adaptif dan berbasis prediksi yang akurat [2], [3].

2.3. Data Historis dan Time Series

Data historis merupakan catatan masa lalu yang menggambarkan perubahan suatu variabel dalam periode tertentu dan sering digunakan untuk memahami pola fluktuasi produksi minyak bumi, baik harian, bulanan, maupun tahunan [4], [18], [19]. Data *time series* sendiri adalah data yang tersusun berdasarkan urutan waktu sehingga setiap nilai memiliki keterhubungan dengan periode sebelumnya atau sesudahnya, serta memuat pola seperti tren, musiman, autokorelasi, dan fluktuasi acak [2], [3], [12].

Dalam konteks produksi minyak bumi, data *time series* membantu menangkap variasi jangka pendek maupun panjang untuk kebutuhan analisis dan prediksi [19]. Hubungan keduanya terletak pada fungsi analisis, yaitu data historis menyediakan rekaman masa lalu, sedangkan data *time series* memanfaatkan rekaman tersebut untuk mengidentifikasi pola dan memproyeksikan kondisi mendatang [3]. Oleh karena itu, setiap *time series* merupakan data historis, tetapi tidak semua data historis bersifat *time series* dan analisis *time series* terbukti efektif untuk memahami perubahan output serta menghasilkan prediksi produksi yang lebih akurat [2], [20].

2.4. Deep Learning

Deep Learning (DL) merupakan cabang kecerdasan buatan yang berkembang dari *Artificial Neural Network* dengan struktur multilayer yang mampu mengekstraksi fitur pada berbagai tingkat abstraksi [21]. Perkembangannya didorong oleh ketersediaan big data, kemampuan komputasi tinggi, dan kemajuan algoritma optimasi [22]. Berbeda dari metode statistik tradisional yang bergantung pada asumsi distribusi, DL mempelajari pola langsung dari data mentah tanpa rekayasa fitur manual [23]. Arsitektur DL dirancang sesuai karakteristik data,

seperti CNN untuk citra dan data spasial, RNN untuk data sekuensial, serta transformer yang unggul menangani dependensi jangka panjang pada bahasa alami maupun deret waktu [24].

Keunggulan utama DL terletak pada feature learning otomatis yang mampu menangkap pola non-linier kompleks dan memberikan generalisasi lebih baik pada data baru [22], [24], [25]. Penerapannya meluas ke bidang kesehatan, transportasi, dan keuangan, mulai dari diagnosis medis hingga prediksi harga saham dan deteksi anomali [9], [21], [23], [24]. Dalam sektor energi, khususnya analisis deret waktu, DL terbukti efektif menangani pola dinamis dan fluktuatif sehingga menjadi landasan penting dalam penelitian prediksi modern [22], [24], [25].

2.5. Prediksi (Forecasting)

Prediksi atau forecasting merupakan proses memperkirakan nilai masa depan berdasarkan pola yang terdapat pada data historis [14]. Prediksi berfungsi menyediakan informasi kuantitatif untuk perencanaan, pengendalian, dan pengambilan keputusan strategis, termasuk dalam aktivitas produksi, lifting, investasi, dan pengelolaan reservoir pada industri minyak bumi [4], [26]. Berdasarkan jangka waktunya, prediksi dibagi menjadi jangka pendek, menengah, dan panjang [3]. Prediksi jangka pendek digunakan untuk kebutuhan operasional harian [27], prediksi jangka menengah untuk perencanaan tahunan dan penyusunan anggaran [28], sedangkan prediksi jangka panjang mendukung kebijakan energi dan strategi investasi [29].

Keberhasilan prediksi *time series* bergantung pada pemahaman pola seperti tren, musiman, autokorelasi, dan fluktuasi acak [14]. Dalam produksi minyak bumi, variasi akibat *workover*, *shut-in*, atau perubahan metode eksploitasi sering menimbulkan pola non-linier sehingga metode prediksi harus disesuaikan dengan karakteristik data agar hasilnya akurat dan dapat dimanfaatkan secara optimal [3], [23].

2.6. Root Mean Squared Error (RMSE)

Root Mean Squared Error (RMSE) merupakan metrik evaluasi yang umum digunakan dalam prediksi data *time series* untuk mengukur tingkat akurasi model [30]. RMSE diperoleh dengan mengakarkan rata-rata kuadrat selisih antara nilai aktual dan prediksi, sehingga menunjukkan seberapa besar penyimpangan hasil prediksi dari kenyataan [14]. Rumus RMSE adalah [6]:

$$\text{RMSE} = \sqrt{\frac{\sum_{t=1}^n (y_t - \hat{y}_t)^2}{n}} \quad (1)$$

RMSE bersifat sensitif terhadap error besar karena adanya proses pengkuadratan, sehingga mampu memberikan penalti tinggi pada kesalahan ekstrem [29]. Karakteristik ini menjadikan RMSE efektif untuk mendeteksi ketidakakuratan signifikan,

terutama pada data yang berfluktuasi atau mengandung *outlier*. Keunggulannya adalah nilai RMSE berada pada satuan yang sama dengan data sehingga mudah diinterpretasikan [13]. Namun, sensitivitas terhadap outlier dapat menyebabkan nilai RMSE terlihat lebih tinggi daripada rata-rata kesalahan sebenarnya [14].

2.7. Mean Absolute Percentage Error (MAPE)

Mean Absolute Percentage Error (MAPE) adalah metrik evaluasi yang menyajikan tingkat kesalahan dalam bentuk persentase sehingga mudah dipahami oleh praktisi maupun pengambil keputusan [30]. MAPE dihitung dari rata-rata kesalahan absolut relatif terhadap nilai aktual, sehingga dapat digunakan untuk membandingkan performa model pada dataset dengan skala berbeda [29]. Secara matematis, rumus MAPE sebagai berikut [31]:

$$\text{MAPE} = \frac{1}{n} \sum_{t=1}^n \left| \frac{y_t - \hat{y}_t}{y_t} \right| \times 100\% \quad (2)$$

Nilai MAPE memiliki indikator [31]:

Tabel 1. Indikator MAPE

Nilai MAPE	Kategori
<10%	Sangat baik
$10\% \leq \text{MAPE} < 20\%$	Baik
$20\% \leq \text{MAPE} < 50\%$	Wajar
$\geq 50\%$	Buruk

Tabel 1 menunjukkan kategori untuk menentukan keakuratan prediksi melalui nilai MAPE. Keunggulan utama MAPE adalah sifatnya yang *scale-independent* sehingga hasil evaluasi tetap konsisten meskipun skala data berubah [14]. Interpretasinya juga mudah, karena persentase menunjukkan besarnya rata-rata kesalahan relatif; misalnya MAPE 5% berarti rata-rata kesalahan prediksi hanya 5% dari nilai aktual [30]. Namun, MAPE kurang andal ketika nilai aktual mendekati nol karena menghasilkan error sangat besar yang dapat menyesatkan [29].

2.8. Long Short-Term Memory (LSTM)

Long Short-Term Memory (LSTM) adalah pengembangan *Recurrent Neural Network* (RNN) yang dirancang untuk mengatasi *vanishing* dan *exploding gradients*, sehingga mampu mempelajari ketergantungan jangka panjang pada data *time series* [22]. Dengan mekanisme gerbang yang mengatur aliran informasi, LSTM dapat menangkap pola temporal yang kompleks dan non-linier.

Model ini banyak digunakan dalam pemrosesan bahasa alami, pengenalan suara, rekomendasi sistem, serta prediksi deret waktu di bidang ekonomi, keuangan, energi, dan cuaca. Berbagai studi menunjukkan bahwa LSTM mampu melampaui performa metode statistik tradisional seperti ARIMA pada data yang fluktuatif dan dipengaruhi banyak faktor [24], [26].

Meski unggul dalam akurasi, LSTM memiliki keterbatasan seperti kebutuhan data besar, waktu komputasi tinggi, risiko *overfitting*, dan interpretabilitas yang rendah [25], [32]. Namun demikian, fleksibilitas dan kemampuannya mengelola pola jangka panjang menjadikan LSTM tetap relevan sebagai salah satu arsitektur penting dalam *deep learning* modern.

3. METODE PENELITIAN

Tahapan penelitian ini disusun secara sistematis agar proses penerapan metode prediksi produksi minyak bumi dapat berjalan runtun. Berikut tahapan utama dalam penelitian:

3.1. Pengumpulan Data

Data yang digunakan dalam penelitian ini merupakan data sekunder berupa data historis harian produksi minyak bumi harian. Data ini berbentuk data *time series*, yaitu data yang dicatat secara kronologis dengan interval waktu harian. Periode data yang digunakan adalah 1 Januari 2023 sampai 31 Juli 2025 dengan cakupan penelitian di PT P. Sumber data penelitian diperoleh langsung dari bagian terkait yang mengelola catatan produksi harian.

3.2. Pra-pemrosesan Data

Data historis produksi minyak bumi harian perlu melalui tahap preprocessing sebelum digunakan sebagai input pemodelan. Proses ini mencakup pembersihan data, normalisasi, dan pembagian data untuk memastikan model menerima data yang bersih, terstandarisasi, dan siap digunakan dalam pelatihan maupun evaluasi.

a. Pembersihan Data (*Data Cleaning*)

Data diperiksa untuk memastikan tidak terdapat *missing values*, duplikasi, maupun anomali. Nilai kosong diatasi menggunakan interpolasi agar kontinuitas data tetap terjaga, sementara duplikasi dihapus guna mencegah bias pada proses pelatihan. Anomali yang menyimpang jauh dari tren normal diperiksa kembali dan dapat disesuaikan atau dihapus apabila tidak valid.

b. Normalisasi Menggunakan *Min-Max Normalization*

Tahap berikutnya adalah normalisasi min-max agar seluruh nilai berada pada rentang 0–1. Normalisasi dilakukan menggunakan persamaan:

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (3)$$

Metode ini dipilih karena sesuai untuk data produksi minyak yang bersifat non-negatif dan tidak memiliki distribusi ekstrem, sehingga membantu menstabilkan pelatihan LSTM dan mempercepat proses konvergensi [9], [12], [29].

c. Pembagian Data (*Train-Test Split*)

Data yang telah dinormalisasi dibagi menjadi data latih dan data uji dengan proporsi 80% dan 20%.

Data latih digunakan untuk membantu model mempelajari pola historis, sementara data uji berfungsi untuk menilai akurasi prediksi terhadap data yang belum pernah dilihat sebelumnya sehingga evaluasi performa model lebih objektif.

3.3. Pelatihan Metode LSTM

Setelah data melewati tahap pra-pemrosesan, langkah selanjutnya adalah pelatihan metode *Long Short-Term Memory* (LSTM) untuk membangun model prediksi. Pada fase ini, data yang telah dibersihkan, dinormalisasi, dibagi menjadi data latih dan data uji, serta ditransformasikan ke bentuk sekuensial digunakan sebagai input utama ke dalam arsitektur LSTM.

Arsitektur LSTM disusun secara bertingkat, dengan lapisan LSTM yang bertugas menangkap pola temporal dari data produksi minyak harian. Lapisan ini dapat dilengkapi dengan dropout untuk mengurangi risiko *overfitting*, kemudian diakhiri dengan lapisan output (*dense layer*) yang menghasilkan prediksi nilai produksi. Penentuan hyperparameter, seperti *window size*, jumlah lapisan dan unit LSTM, *batch size*, *epoch*, serta algoritma optimisasi, dilakukan melalui pengujian awal untuk memperoleh konfigurasi yang memberikan hasil paling stabil.

Model dilatih menggunakan data latih hingga mencapai konvergensi sesuai fungsi kerugian regresi yang digunakan. Setelah proses pelatihan selesai, model diaplikasikan pada data uji untuk menghasilkan nilai prediksi, yang kemudian didenormalisasi agar dapat dibandingkan secara langsung dengan nilai aktual produksi minyak bumi.

3.4. Evaluasi Metode

Evaluasi dilakukan dengan membandingkan hasil prediksi LSTM terhadap data aktual pada bagian data uji. Dua metrik digunakan, yaitu *Root Mean Square Error* (RMSE) untuk mengukur besar deviasi prediksi dalam satuan asli data, serta *Mean Absolute Percentage Error* (MAPE) untuk melihat tingkat kesalahan dalam bentuk persentase. Penggunaan kedua metrik ini memberikan gambaran yang lebih lengkap mengenai akurasi dan konsistensi performa metode.

3.5. Prediksi Produksi Minyak Bumi

Tahap ini menghasilkan prediksi produksi minyak bumi harian untuk tiga tahun ke depan menggunakan metode LSTM yang telah dilatih pada data historis. Data yang sudah dinormalisasi diubah menjadi urutan (*sequence*) berdasarkan window size tertentu sehingga model dapat mempelajari pola temporal produksi.

Setelah pelatihan, dilakukan dua jenis prediksi, yaitu *in-sample prediction* untuk menilai kemampuan model merepresentasikan pola historis dan *future forecasting* untuk meramalkan nilai produksi secara iteratif menggunakan teknik *sliding window*. Model juga diberi elemen *adaptive noise* berdasarkan

volatilitas 90 hari terakhir agar prediksi lebih realistis. Hasil prediksi kemudian dikembalikan ke skala asli melalui *inverse MinMaxScaler*.

4. HASIL DAN PEMBAHASAN

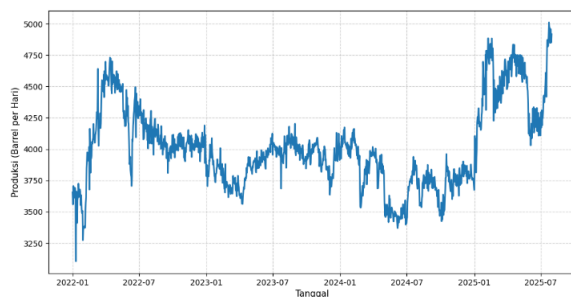
4.1. Hasil Pengumpulan Data

Dataset yang digunakan merupakan data *time series* yang merepresentasikan produksi minyak bumi harian. Data terdiri atas tiga kolom, yaitu *Date*, *Oil (bo) Target*, dan *Oil (bo) Real*. Kolom *Oil (bo) Real* dipilih sebagai variabel prediksi karena mencerminkan output aktual dan fluktuasi nyata di lapangan. Jumlah data yang tersedia sebanyak 1.308 data harian sehingga cukup untuk membentuk pola temporal yang diperlukan dalam pemodelan LSTM.

Tabel 2. Contoh data yang digunakan

No.	Date	Oil (bo)	
		Target	Real
1.	01/01/2022	3956	3655
2.	02/01/2022	3956	3558
3.	03/01/2022	3956	3705
4.	04/01/2022	3956	3643
...			
1305.	28/07/2025	4608	4848
1306.	29/07/2025	4608	4958
1307.	30/07/2025	4608	4851
1308.	31/07/2025	4608	4917

Tabel 2 menampilkan sebagian data produksi minyak harian yang tersusun secara kronologis berdasarkan tanggal. Tren fluktuasi produksi harian dapat diamati secara visual pada Gambar 1 yang menampilkan perubahan volume produksi selama periode observasi.



Gambar 1. Tren produksi minyak bumi harian

Gambar di atas memperlihatkan tren produksi minyak bumi harian selama periode penelitian yang menunjukkan fluktuasi signifikan dari waktu ke waktu. Terdapat kecenderungan peningkatan pada awal 2022 dan awal 2025, sementara penurunan terlihat pada pertengahan 2023 hingga pertengahan 2024. Pola yang tidak stabil ini menegaskan bahwa karakteristik data bersifat non-linier dan dinamis yang dipengaruhi oleh variasi operasional serta kondisi lapangan.

4.2. Hasil Pra-Pemrosesan Data

a. Hasil Pembersihan Data

Proses pembersihan data mencakup pemeriksaan nilai kosong dan duplikasi berdasarkan kolom *Date* guna memastikan kualitas dan keandalan data sebelum tahap pelatihan. Hasil pemeriksaan dan pembersihan data ditampilkan pada gambar berikut untuk memastikan tidak terdapat nilai kosong maupun duplikasi.

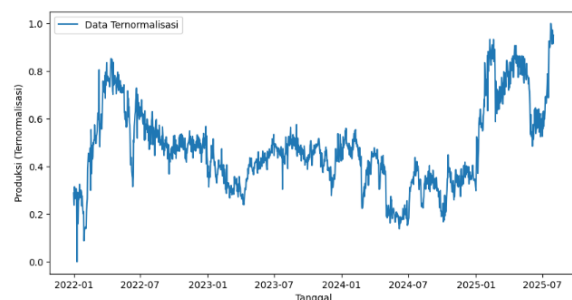
```
Jumlah nilai kosong tiap kolom:
Date      0
Target    0
Real      0
dtype: int64
Data duplikat:
Empty DataFrame
Columns: [Date, Target, Real]
Index: []
Jumlah data setelah dibersihkan: 1308
Periode data: 2022-01-01 00:00:00 sampai 2025-07-31 00:00:00
```

Gambar 2. Hasil pembersihan Data

Gambar 2 menunjukkan eluruh kolom tidak mengandung nilai kosong (*missing value*) dan tidak ditemukan duplikasi pada kolom *Date*. Karena tidak ditemukan masalah pada data, dataset siap untuk masuk ke tahap berikutnya.

b. Hasil Normalisasi Data

Setelah proses pembersihan, data produksi minyak bumi dinormalisasi menggunakan metode *min-max* agar seluruh nilai berada dalam rentang 0–1 sehingga skala data menjadi seragam dan tidak menimbulkan bias dalam proses pelatihan model. Langkah ini dilakukan untuk menjaga stabilitas proses pembelajaran dan memastikan pola perubahan produksi dapat dipelajari secara konsisten oleh metode. Perubahan nilai setelah normalisasi dapat dilihat pada gambar berikut:

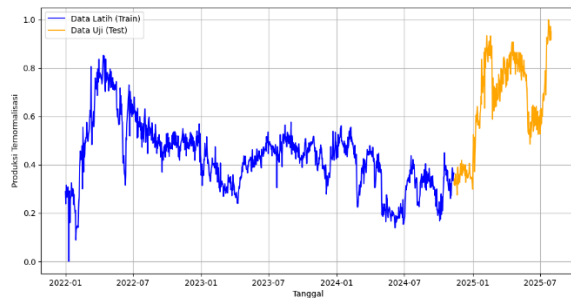


Gambar 3. Grafik data setelah normalisasi

Gambar 3 menunjukkan penyesuaian skala data ke rentang 0-1. Data yang telah dinormalisasi tetap mempertahankan pola tren aslinya, hanya skala nilainya yang diperkecil. Ini menegaskan bahwa normalisasi berjalan baik tanpa mengubah pola temporal data dan data siap digunakan untuk tahap selanjutnya.

c. Hasil Pembagian Data

Setelah dinormalisasi, dataset dipisahkan menjadi data latih dan data uji dengan komposisi 80% dan 20%. Data latih berfungsi membentuk model LSTM berdasarkan pola temporal, sedangkan data uji digunakan untuk mengukur kinerja model pada data baru. Pembagian kedua subset ditunjukkan pada gambar berikut:



Gambar 4. Hasil Pembagian Data

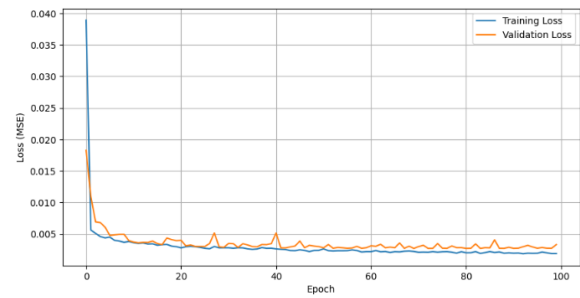
Berdasarkan hasil pembagian, dari total 1.308 data harian, sebanyak 1.046 observasi digunakan sebagai data latih dan 262 sebagai data uji. Visualisasi proporsi pembagian kedua subset tersebut ditampilkan pada Gambar 4, di mana data latih ditampilkan dengan warna biru dan data uji dengan warna oranye. Pembagian ini memastikan metode mendapatkan proporsi data yang seimbang antara proses pelatihan dan pengujian.

4.3. Hasil Pelatihan Metode

Pelatihan metode LSTM dilakukan untuk membangun model prediksi produksi minyak bumi berdasarkan data historis harian yang telah diproses sebelumnya. Pada tahap ini, data diubah ke dalam bentuk berurutan (*sequence*), kemudian digunakan untuk melatih jaringan agar mampu mengenali pola temporal produksi dari waktu ke waktu.

Model dirancang menggunakan arsitektur LSTM dengan 50 unit neuron, *dropout* 0,2 untuk mengurangi *overfitting*, dan satu *dense layer* sebagai keluaran. Model dikompilasi menggunakan Adam dengan fungsi *loss* MSE, lalu dilatih selama 100 *epoch* dengan *batch size* 32. Selama proses pelatihan, nilai *loss* dan *validation loss* menunjukkan tren penurunan yang konsisten hingga mencapai sekitar 0,0019 dan 0,0028. Hal ini menandakan proses pembelajaran berjalan stabil dan model berhasil mencapai konvergensi.

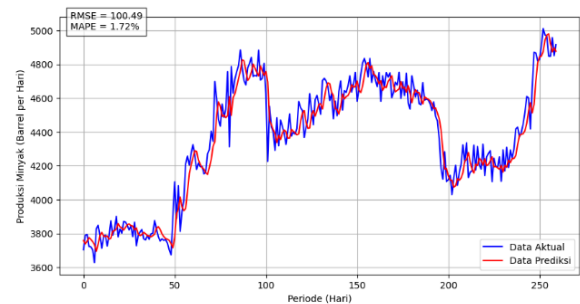
Gambar 5 menunjukkan Nilai *loss* dan *validation loss* sama-sama mengalami penurunan signifikan pada *epoch* awal dan kemudian mencapai kestabilan di bawah nilai 0.01 setelah sekitar *epoch* ke-20. Pola kedua garis yang saling berdekatan menunjukkan bahwa metode tidak mengalami *overfitting*, karena performa pada data latih dan data uji relatif seimbang. Hal ini menunjukkan bahwa arsitektur dan parameter pelatihan yang digunakan sudah sesuai untuk data produksi minyak bumi yang bersifat *time series*.



Gambar 5. Grafik perbandingan nilai *loss*

4.4. Hasil Evaluasi Metode

Evaluasi dilakukan untuk menilai kemampuan metode LSTM dalam memprediksi produksi minyak bumi berdasarkan data historis harian. Kinerja model dinilai menggunakan dua metrik utama, yaitu RMSE dan MAPE.



Gambar 6. Perbandingan nilai aktual dan prediksi

Visualisasi pada Gambar 6 memperlihatkan bahwa kurva prediksi mengikuti pola kurva aktual dengan baik, baik pada perubahan tren jangka panjang maupun fluktuasi jangka pendek. Hal ini menunjukkan bahwa LSTM mampu menangkap struktur temporal data secara efektif dan menghasilkan prediksi yang stabil dan presisi.

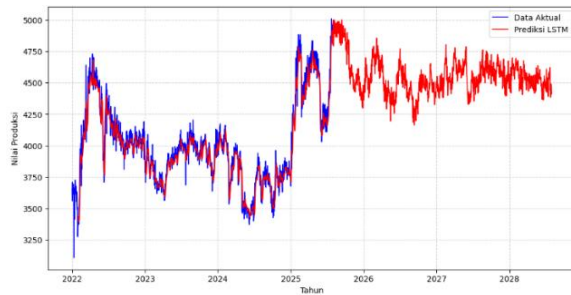
Hasil evaluasi menunjukkan bahwa metode LSTM memiliki performa prediksi yang sangat baik. Nilai RMSE sebesar 100,49 mengindikasikan bahwa deviasi rata-rata prediksi terhadap data aktual hanya sekitar 100 barel per hari. Jika dibandingkan dengan rata-rata produksi harian sebesar 4.012 barel, tingkat kesalahan absolut ini hanya sekitar 2,49%, yang termasuk sangat kecil dalam konteks operasional.

Nilai MAPE sebesar 1,72% semakin memperkuat bahwa kesalahan relatif model berada di bawah 2%. Berdasarkan standar umum evaluasi, MAPE di bawah 10% tergolong akurasi sangat tinggi, sehingga hasil ini menegaskan konsistensi dan keandalan model dalam menangani pola data yang dinamis.

4.5. Analisis Hasil Prediksi

Pada tahap ini dilakukan proses peramalan produksi minyak bumi untuk tiga tahun ke depan setelah model LSTM selesai dilatih menggunakan data historis 2022–2025. Prediksi dilakukan melalui serangkaian prosedur komputasional untuk

menghasilkan proyeksi yang stabil dan realistis sekaligus tetap mempertahankan karakter fluktuatif pada pola produksi.



Gambar 7. Hasil prediksi produksi minyak bumi

Grafik pada Gambar 7 memperlihatkan perbandingan antara data aktual produksi minyak bumi (garis biru) dan hasil prediksi metode LSTM (garis merah). Secara visual, model menunjukkan kemampuan yang baik dalam mengikuti pola historis, terlihat dari pergerakan kurva prediksi yang sejajar dengan perubahan nilai aktual, baik saat tren meningkat maupun menurun. Konsistensi ini menunjukkan bahwa LSTM mampu mempelajari pola temporal dan dinamika deret waktu secara efektif.

Pada bagian setelah data aktual berakhir, garis merah menggambarkan hasil proyeksi produksi untuk tiga tahun ke depan. Prediksi menunjukkan pola fluktuatif yang tetap stabil, dengan rentang nilai sekitar 4.400 hingga 4.800 barel per hari. Rentang ini mencerminkan proyeksi yang realistis dan tidak berlebihan, sekaligus mempertahankan variasi alami yang lazim terjadi pada produksi minyak. Fluktuasi tersebut mengindikasikan adanya dinamika operasional dan faktor eksternal yang secara tipikal memengaruhi perubahan nilai produksi dari waktu ke waktu.

5. KESIMPULAN DAN SARAN

Penelitian ini menunjukkan bahwa metode *Long Short-Term Memory* (LSTM) mampu memberikan prediksi produksi minyak bumi harian dengan akurasi tinggi, dibuktikan oleh nilai RMSE sebesar 100,49 dan MAPE sebesar 1,72%, serta kemampuan model dalam mengikuti pola fluktuasi data baik pada tahap validasi maupun proyeksi tiga tahun ke depan. Meskipun begitu, penelitian ini masih memiliki keterbatasan berupa penggunaan satu variabel tunggal dan rentang data historis yang relatif pendek. Oleh karena itu, disarankan agar penelitian selanjutnya menambah variabel teknis dan operasional yang relevan, memperluas rentang data historis, serta melakukan perbandingan dengan metode lain seperti GRU, *Bidirectional-LSTM*, atau *hybrid CNN-LSTM* untuk mendapatkan model prediksi yang lebih komprehensif, robust, dan adaptif terhadap dinamika kondisi lapangan.

DAFTAR PUSTAKA

- [1] H. Wahyudi, "PERAMALAN KONSUMSI DAN PRODUKSI MINYAK MENGGUNAKAN MODEL ARIMA DI INDONESIA," *Jurnal Ekonomi Pembangunan*, vol. 12, no. 1, pp. 12–26, Apr. 2023, doi: 10.23960/JEP.V12I1.1090.
- [2] Z. Xu and G. Yu, "A Time Series Forecasting Approach Based on Meta-Learning for Petroleum Production under Few-Shot Samples," *Energies* 2024, Vol. 17, Page 1947, vol. 17, no. 8, p. 1947, Apr. 2024, doi: 10.3390/EN17081947.
- [3] Y. Ning, H. Kazemi, and P. Tahmasebi, "A comparative machine learning study for time series oil production forecasting: ARIMA, LSTM, and Prophet," *Comput Geosci*, vol. 164, p. 105126, Jul. 2022, doi: 10.1016/J.CAGEO.2022.105126.
- [4] X. Song *et al.*, "Time-series well performance prediction based on Long Short-Term Memory (LSTM) neural network model," *J Pet Sci Eng*, vol. 186, p. 106682, Mar. 2020, doi: 10.1016/J.PETROL.2019.106682.
- [5] E. Kozhevnikov, E. Riabokon, and M. Turbakov, "A Model of Reservoir Permeability Evolution during Oil Production," *Energies* 2021, Vol. 14, Page 2695, vol. 14, no. 9, p. 2695, May 2021, doi: 10.3390/EN14092695.
- [6] E. Kartawijaya, M. Munawar, G. Firmansyah, and B. Tjahjono, "Comparative Analysis of Time Series Methods LSTM and ARIMA for Predicting Inventory Availability (Case Study: PT XYZ)," *CCIT (Creative Communication and Innovative Technology) Journal*, vol. 18, no. 1, pp. 9–19, Aug. 2025, doi: 10.33050/CCIT.V18I1.3443.
- [7] A. Tadjer, A. Hong, and R. B. Bratvold, "Machine learning based decline curve analysis for short-term oil production forecast," *Energy Exploration & Exploitation*, vol. 39, no. 5, pp. 1747–1769, Sep. 2021, doi: 10.1177/01445987211011784.
- [8] H. Yan, M. Liu, B. Yang, Y. Yang, H. Ni, and H. Wang, "Short-term forecasting approach of single well production based on multi-intelligent agent hybrid model," *PLoS One*, vol. 19, no. 4, p. e0301349, Apr. 2024, doi: 10.1371/JOURNAL.PONE.0301349.
- [9] F. Fathoni, A. Ibrahim, R. Septiana, A. R. Putri, T. Ispahan, and W. S. Maharani, "ANALISIS PREDIKSI HARGA SAHAM PADA PERUSAHAAN T MENGGUNAKAN KOMBINASI CNN-LSTM," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 9, no. 4, pp. 6669–6675, May 2025, doi: 10.36040/JATI.V9I4.14104.
- [10] S. Hosseini and T. Akilan, "Advanced Deep Regression Models for Forecasting Time Series Oil Production," Aug. 2023, Accessed: Sep. 19, 2025. [Online]. Available: <https://arxiv.org/pdf/2308.16105>

- [11] J. B. Correia, M. Pivetta, G. S. do Nascimento, and K. Becker, "Comparing ARIMA and LSTM models to predict time series in the oil industry," *Symposium on Knowledge Discovery, Mining and Learning (KDMiLe)*, pp. 129–136, Oct. 2021, doi: 10.5753/KDMiLe.2021.17470.
- [12] K. Zhang and M. Hong, "Forecasting crude oil price using LSTM neural networks," *Data Science in Finance and Economics* 2022 3:163, vol. 2, no. 3, pp. 163–180, 2022, doi: 10.3934/DSFE.2022008.
- [13] K. Alsalem, "A hybrid time series forecasting approach integrating fuzzy clustering and machine learning for enhanced power consumption prediction," *Sci Rep*, vol. 15, no. 1, pp. 1–16, Dec. 2025, doi: 10.1038/S41598-025-91123-8;SUBJMETA.
- [14] R. J. Hyndman and G. Athanasopoulos, "Forecasting: Principles and Practice (3rd ed)." Accessed: Sep. 19, 2025. [Online]. Available: <https://otexts.com/fpp3/>
- [15] D. He *et al.*, "Oil Production Rate Forecasting by SA-LSTM Model in Tight Reservoirs," *Lithosphere*, vol. 2024, no. 1, Jan. 2024, doi: 10.2113/2024/LITHOSPHERE_2023_197.
- [16] H. K. Chavan and R. K. Sinharay, "Prophet modeling for oil production forecasting in an enhanced oil recovery field," *Physics of Fluids*, vol. 36, no. 9, Sep. 2024, doi: 10.1063/5.0224299.
- [17] M. Ahsan, "Tantangan dan Peluang Pembangunan Proyek Pembangkit Listrik Energi Baru Terbarukan (EBT) di Indonesia," *SUTET*, vol. 11, no. 2, pp. 81–93, Dec. 2021, doi: 10.33322/SUTET.V11I2.1575.
- [18] D. Y. Situmorang, H. B. Sulistyarso, and D. L. Wijayanti, "Examining the Relation between Reserves Replacement Ratio, Production, Proven Reserves and Upstream Oil & Gas Investment in Indonesia (2010-2022)," *Journal of Energy Research and Reviews*, vol. 17, no. 3, pp. 42–56, Mar. 2025, doi: 10.9734/JENRR/2025/V17I3401.
- [19] C. Mgbemena and E. Chinwuko, "FORECAST OF CRUDE OIL PRODUCTION OUTPUT IN AN OIL FIELD IN THE NIGER DELTA REGION OF NIGERIA," *International Journal of Industrial Engineering & Production Research*, vol. 31, no. 1, pp. 161–170, 2020, doi: 10.22068/IJIEPR.31.1.161.
- [20] Y. Li, Y. Li, and Y. Liu, "Research on used car price prediction based on random forest and LightGBM," *2022 IEEE 2nd International Conference on Data Science and Computer Application, ICDSCA 2022*, pp. 539–543, 2022, doi: 10.1109/ICDSCA56264.2022.9988116.
- [21] K. Sharifani and M. Amini, "Machine Learning and Deep Learning: A Review of Methods and Applications," 2023. Accessed: Sep. 19, 2025. [Online]. Available: <https://papers.ssrn.com/abstract=4458723>
- [22] B. Lim and S. Zohren, "Time-series forecasting with deep learning: a survey," *Philosophical Transactions of the Royal Society A*, vol. 379, no. 2194, Apr. 2021, doi: 10.1098/RSTA.2020.0209.
- [23] W. Li and K. L. E. Law, "Deep Learning Models for Time Series Forecasting: A Review," *IEEE Access*, vol. 12, pp. 92306–92327, 2024, doi: 10.1109/ACCESS.2024.3422528.
- [24] A. Casolaro, V. Capone, G. Iannuzzo, and F. Camastra, "Deep Learning for Time Series Forecasting: Advances and Open Problems," *Information* 2023, Vol. 14, Page 598, vol. 14, no. 11, p. 598, Nov. 2023, doi: 10.3390/INFO14110598.
- [25] I. Malashin, V. Tynchenko, A. Gantimurov, V. Nelyub, and A. Borodulin, "Applications of Long Short-Term Memory (LSTM) Networks in Polymeric Sciences: A Review," *Polymers* 2024, Vol. 16, Page 2607, vol. 16, no. 18, p. 2607, Sep. 2024, doi: 10.3390/POLYM16182607.
- [26] H. Y. Tang *et al.*, "Production decline curve analysis of shale oil wells: A case study of Bakken, Eagle Ford and Permian," *Pet Sci*, vol. 21, no. 6, pp. 4262–4277, Dec. 2024, doi: 10.1016/J.PETSCI.2024.07.029.
- [27] K. A. Rijal, A. V. Vitianingsih, Y. Kristyawan, A. L. Maukar, and S. F. A. Wati, "Forecasting Model of Indonesia's Oil & Gas and Non-Oil & Gas Export Value using Var and LSTM Methods," *Jurnal Teknologi dan Manajemen Informatika*, vol. 10, no. 1, pp. 59–69, Jun. 2024, doi: 10.26905/JTMI.V10I1.13127.
- [28] H. Rahmanifard and I. Gates, "A Comprehensive review of data-driven approaches for forecasting production from unconventional reservoirs: best practices and future directions," *Artif Intell Rev*, vol. 57, no. 8, pp. 1–40, Aug. 2024, doi: 10.1007/S10462-024-10865-5/FIGURES/6.
- [29] J. Kim, H. Kim, H. G. Kim, D. Lee, and S. Yoon, "A comprehensive survey of deep learning for time series forecasting: architectural diversity and open challenges," *Artif Intell Rev*, vol. 58, no. 7, Jul. 2025, doi: 10.1007/S10462-025-11223-9.
- [30] S. Makridakis, E. Spiliotis, and V. Assimakopoulos, "The M4 Competition: 100,000 time series and 61 forecasting methods," *Int J Forecast*, vol. 36, no. 1, pp. 54–74, Jan. 2020, doi: 10.1016/J.IJFORECAST.2019.04.014.
- [31] H. Hewamalage, K. Ackermann, and C. Bergmeir, "Forecast evaluation for data scientists: common pitfalls and best practices," *Data Min Knowl Discov*, vol. 37, no. 2, p. 788, Mar. 2022, doi: 10.1007/S10618-022-00894-5.
- [32] H. Kuswanto *et al.*, "Prediksi Nilai Ekspor Migas Indonesia Menggunakan Metode SARIMA dan LSTM," *SATESI: Jurnal Sains Teknologi dan Sistem Informasi*, vol. 5, no. 1, pp. 69–79, Apr. 2025, doi: 10.54259/SATESI.V5I1.4103