

Pendekatan *Transformer Deep Learning* dalam Meramalkan Harga Minyak Sumatran Light Crude

Ni Luh Putu Ika Candraweni^{1*}, Yadhusani Dewi Amritha¹, Md. Wira Putra Dananjaya¹

¹*Universitas Pendidikan Nasional, Denpasar, Bali, Indonesia*

ikacandraweni@undiknas.ac.id^{*}

Received: 06/12/2025 | Revised: 15/12/2025 | Accepted: 25/12/2025

Copyright©2025 by authors. Authors agree that this article remains permanently open access under the terms of the Creative Commons

Abstrak

Peramalan deret waktu memegang peranan penting dalam memahami dinamika data yang bersifat volatil dan bergantung pada pola historis jangka panjang, seperti harga minyak mentah. Pendekatan statistik parametrik sering menghadapi keterbatasan akibat asumsi yang ketat, sehingga metode *deep learning* nonparametrik menjadi alternatif yang lebih fleksibel. Penelitian ini mengusulkan penerapan model *Transformer-based deep learning* untuk memprediksi harga Minyak Mentah Ringan Sumatra (*Sumatran Light Crude Oil – SLC*), dengan memanfaatkan mekanisme *self-attention* guna menangkap ketergantungan jangka panjang dalam data deret waktu. Eksperimen dilakukan dengan mengevaluasi berbagai konfigurasi jumlah *multi-head attention* dan jumlah lapisan (*layers*), sementara dimensi model serta jendela input–output dijaga tetap konsisten. Hasil penelitian menunjukkan bahwa konfigurasi Transformer dengan 16 *heads* dan 4 *layers* memberikan kinerja terbaik dengan nilai *Root Mean Square Error* (RMSE) sebesar 8,19818. Temuan ini mengindikasikan bahwa Transformer mampu memodelkan tren jangka panjang harga SLC secara efektif, meskipun sensitivitas terhadap fluktuasi jangka pendek masih terbatas. Kontribusi utama penelitian ini terletak pada pemanfaatan Transformer sebagai pendekatan alternatif untuk peramalan harga minyak mentah di Indonesia, yang sebelumnya masih didominasi oleh metode statistik dan model rekuren. Secara praktis, hasil penelitian ini memberikan dasar bagi pengembangan sistem peramalan harga minyak yang lebih adaptif dalam mendukung analisis energi dan pengambilan keputusan berbasis data.

Kata kunci: Deep Learning, Forecasting, SLC, *Transformer*

Abstract

Time series forecasting plays an important role in understanding the dynamics of volatile data that depends on long-term historical patterns, such as crude oil prices. Parametric statistical approaches often face limitations due to strict assumptions, making nonparametric deep learning methods a more flexible alternative. This study proposes the application of a Transformer-based deep learning model to predict the price of Sumatran Light Crude Oil (SLC), utilizing a self-attention mechanism to

capture long-term dependencies in time series data. Experiments were conducted by evaluating various configurations of multi-head attention and number of layers, while keeping the model dimensions and input-output windows consistent. The results show that the Transformer configuration with 16 heads and 4 layers provides the best performance with a Root Mean Square Error (RMSE) value of 8.19818. These findings indicate that Transformer is capable of effectively modeling long-term trends in SLC prices, although its sensitivity to short-term fluctuations is still limited. The main contribution of this research lies in the use of Transformer as an alternative approach to forecasting crude oil prices in Indonesia, which was previously dominated by statistical methods and recurrent models. In practical terms, the results of this study provide a basis for the development of a more adaptive oil price forecasting system to support energy analysis and data-driven decision making.

Keywords: Deep-learning, Forecasting, SLC, Transformer

Pendahuluan

Beberapa tahun terakhir, *machine learning* dan *deep learning* banyak digunakan untuk memodelkan dan meramalkan data *time-series*, khususnya pada data yang bersifat dinamis dan volatil seperti harga minyak. Salah satu model yang umum digunakan adalah *Recurrent Neural Network* (RNN). Namun, RNN memiliki keterbatasan dalam menangkap ketergantungan jangka panjang karena masalah *vanishing gradient* (Harrou et al., 2020). Untuk mengatasi kelemahan tersebut, dikembangkan *Long Short-Term Memory* (LSTM) yang mampu mempertahankan informasi historis lebih lama melalui mekanisme gerbang (Hochreiter & Schmidhuber, 1997).

Meskipun LSTM telah berhasil diterapkan pada berbagai kasus *time-series* seperti pemantauan kondisi mesin dan deteksi anomali (Zhao et al., 2017; Pariaman et al., 2021), model ini masih menghadapi keterbatasan ketika dihadapkan pada pola jangka panjang yang kompleks dan fluktuatif, terutama pada data dengan rentang waktu yang panjang. Ketergantungan pada mekanisme rekuren membuat LSTM kurang efisien dalam memanfaatkan seluruh informasi historis secara simultan.

Model berbasis *Transformer* Sebagai alternatif dikembangkan dengan mekanisme *self-attention* yang memungkinkan model mengakses seluruh urutan data historis secara langsung tanpa ketergantungan rekuren (Vaswani et al., 2017). Pendekatan ini terbukti efektif dalam berbagai aplikasi *time-series*, termasuk peramalan konsumsi listrik dan lalu lintas (Farsani & Pazouki, 2021) deteksi anomali, serta peramalan spasial-temporal (Wen et al., 2022). Berbagai pengembangan arsitektur Transformer, seperti LogTrans (Li et al., 2019) dan Informer (Zhou et al., 2021), menunjukkan bahwa Transformer unggul dalam peramalan jangka panjang dibandingkan model rekuren konvensional.

Di Indonesia, Minyak Mentah Ringan Sumatra (*Sumatran Light Crude* / SLC) merupakan komoditas strategis yang digunakan sebagai dasar indeksasi harga bagi berbagai jenis minyak mentah nasional. Harga SLC menunjukkan volatilitas tinggi dan dipengaruhi oleh dinamika global serta kebijakan energi, sehingga memerlukan metode peramalan yang mampu menangkap pola nonlinier dan ketergantungan jangka panjang. Penelitian sebelumnya terkait peramalan harga SLC masih didominasi oleh pendekatan statistik, seperti model campuran autoregresif yang

digunakan oleh Sholikhah (Sholikhah, 2022), yang meskipun menunjukkan kinerja baik, memiliki keterbatasan dalam menangkap pola kompleks pada data dengan fluktuasi tinggi.

Berdasarkan tinjauan tersebut, masih terdapat *research gap* berupa minimnya kajian yang mengimplementasikan dan mengevaluasi model *Transformer-based Deep Learning* secara khusus untuk peramalan harga SLC. Padahal, karakteristik data harga SLC yang panjang, volatil, dan dipengaruhi banyak faktor menjadikannya sangat relevan untuk dimodelkan menggunakan Transformer. Oleh karena itu, penelitian ini berfokus pada penerapan dan evaluasi kinerja model Transformer dalam memprediksi harga Minyak Mentah Ringan Sumatra, guna memberikan alternatif metode peramalan yang lebih adaptif dan komprehensif dibandingkan pendekatan konvensional maupun model rekuren.

Metodologi Penelitian

Dataset Penelitian

Penelitian ini menggunakan data sekunder harga SLC dari Januari 1989 hingga Desember 2021. Data ini dapat ditemukan pada publikasi Kementerian Energi dan Sumber Daya Mineral Indonesia (ESDM, 2022). Variabel yang digunakan dalam penelitian ini adalah harga SLC dengan struktur data seperti yang ditunjukkan pada Tabel 1.

Tabel 1 Struktur data penelitian

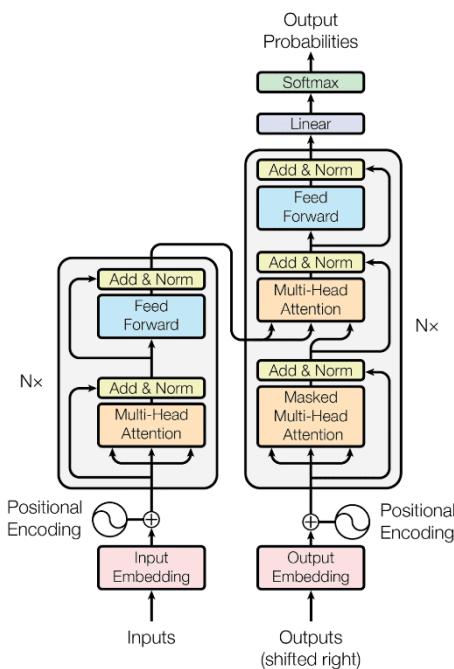
Date	SLC Price
t_1	x_1
t_2	x_2
⋮	⋮
t_t	x_t
⋮	⋮
t_n	x_n

Arsitektur Model *Transformer* yang diusulkan

Penelitian ini menggunakan model *Transformer-based deep learning* untuk memprediksi harga Minyak Sumatran Light Crude Oil (SLC) berbasis data *time-series*. Transformer dipilih karena kemampuannya menangkap ketergantungan jangka panjang melalui mekanisme *self-attention* tanpa ketergantungan rekuren, sehingga lebih sesuai untuk data harga minyak yang bersifat volatil dan memiliki rentang historis panjang.

Arsitektur Transformer yang digunakan terdiri atas dua komponen utama, yaitu *encoder* dan *decoder*. Pada penelitian ini, bagian *encoder* berfungsi untuk mengekstraksi representasi fitur dari rangkaian data harga historis, sedangkan *decoder* digunakan untuk menghasilkan nilai prediksi harga pada periode selanjutnya secara autoregresif. Struktur umum arsitektur mengikuti Transformer standar sebagaimana diperkenalkan oleh Vaswani et al. (2017), dengan penyesuaian pada input data *time-series*. Arsitektur *Transformer* dapat dilihat pada Gambar 1. Pada *encoder*,

terdapat beberapa sub-*encoder* yang memiliki parameter dan lapisan, di mana setiap lapisan memiliki dua sub-lapisan, yaitu multi-head attention dan lapisan feed forward. Sementara pada *decoder*, jumlah lapisan sama dengan jumlah lapisan yang digunakan pada *encoder*. Setiap lapisan memiliki parameternya sendiri dan tidak terkait dengan lapisan lain. Pada *decoder*, ditambahkan tiga sub-lapisan untuk melakukan *multi-head attention* pada outputnya.

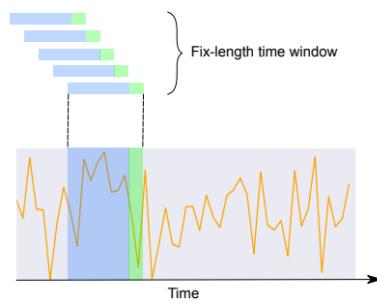


Gambar 1. Ilustrasi arsitektur *transformer*

Model “*Transformer*” menggunakan data dari seri sebelumnya sebagai input untuk memprediksi data ke-t, sehingga input ini akan mengandung banyak informasi. Hal ini sering terjadi dalam peramalan dengan seri yang jauh. Input seri yang jauh akan membuat model kesulitan menentukan informasi mana dalam seri yang paling penting. *Self-attention* adalah algoritma yang dapat digunakan untuk mengatasi masalah ini. Blok *self-attention* mengkode setiap input baru dengan memperhatikan semua input sebelumnya dan menempatkan penekanan berdasarkan perhitungan relevansi token saat ini (Pandey & Jain, 2021). Langkah-langkah untuk menghasilkan model “*Transformer*” dijelaskan sebagai berikut.

Pembentukan Data Input dengan Sliding Window

Data harga SLC dibentuk menjadi rangkaian input menggunakan pendekatan *sliding window*. Metode ini membagi data historis ke dalam segmen berdurasi tetap, sehingga setiap sampel input merepresentasikan pola harga pada beberapa periode sebelumnya untuk memprediksi harga pada periode berikutnya. Jika periode data lebih pendek dari panjang data, akan terjadi tumpang tindih (Pal et al., 2018) . Ilustrasi windows *sliding* dapat dilihat pada Gambar 2.



Gambar 2 Ilustrasi windows sliding

Pembentukan Positional Encoding

Untuk mempertahankan informasi urutan waktu, *positional encoding* ditambahkan ke setiap vektor input. Pendekatan ini memungkinkan model membedakan posisi temporal antar observasi dalam satu rangkaian data. Skema *sinusoidal positional encoding* digunakan sebagaimana pada Transformer standar, tanpa modifikasi bentuk matematis. dinotasikan sebagai $\mathbf{x}_i \in \mathbb{R}^d$ untuk memberikan *encoding* untuk posisi *ke-I*, sehingga vector $\mathbf{PE}_i \in \mathbb{R}^d$ dapat dihitung menggunakan Persamaan (1) dan Persamaan (2) (Vaswani et al., 2017).

$$\mathbf{PE}_i(2j) = \sin(i / 10000^{2j/p}), \quad (1)$$

$$\mathbf{PE}_i(2j+1) = \cos(i / 10000^{2j/p}), \quad (2)$$

Mekanisme Self-Attention dan Multi-Head Attention

Pada setiap lapisan *encoder* dan *decoder*, mekanisme *self-attention* digunakan untuk menghitung tingkat kepentingan setiap titik waktu terhadap titik waktu lainnya dalam satu rangkaian. Mekanisme ini memungkinkan model memfokuskan perhatian pada periode historis yang paling relevan dalam membentuk prediksi harga SLC. Diketahui data, query, and key dibentuk dalam matriks $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_n] \in \mathbb{R}^{d \times n}$, $\mathbf{Q} = [\mathbf{q}_1, \dots, \mathbf{q}_n] \in \mathbb{R}^{p \times n}$, $\mathbf{K} = [\mathbf{k}_1, \dots, \mathbf{k}_n] \in \mathbb{R}^{p \times n}$, and $\mathbf{V} = [\mathbf{v}_1, \dots, \mathbf{v}_n] \in \mathbb{R}^{r \times n}$. Untuk mendapatkan skor *attention* yang telah dinormalisasi, dapat ditunjukkan dalam Persamaan **Error! Reference source not found..**

$$\mathbb{R}^{r \times n} \ni \mathbf{Z} = \text{attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{soft max} \left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{p}} \right) \mathbf{V}, \quad (3)$$

dimana, $\mathbf{Z} = [z_1, \dots, z_n]$ adalah skor *attention* untuk semua observasi yang menunjukkan seberapa besar pengaruh observasi tersebut terhadap observasi berikutnya, dan (=8).

Skor *attention* yang dihasilkan oleh scaled dot-product *attention* kemudian digabungkan menjadi multi-head *attention*. Multi-head *attention* adalah *attention* yang dilakukan sebanyak *h* kali. Multi-head *attention* memungkinkan model untuk merekam informasi dari representasi spasial yang berbeda dari setiap posisi secara bersamaan. Multi-head *attention* juga membantu “Transformers” mengatasi kelemahan pelatihan non-paralel dan panjang kapasitas memori dalam proses ketergantungan jangka panjang pada LSTM atau GRU (Yi et al., 2021). Fungsi multi-head *attention* dapat ditunjukkan dalam Persamaan (8) dan Persamaan (9).

$$\mathbf{Z}_t = \text{concat}(\text{head}_1, \dots, \text{head}_h) \mathbf{W}^O, \quad (4)$$

$$\text{head}_i = \text{Attention}(\mathbf{QW}_i^Q, \mathbf{KW}_i^K, \mathbf{VW}_i^V), \quad (5)$$

di mana hasilnya adalah matriks dengan parameter $\mathbf{W}_i^Q \in \mathbb{R}^{d \times p}$, $\mathbf{W}_i^K \in \mathbb{R}^{d \times p}$, $\mathbf{W}_i^V \in \mathbb{R}^{d \times r}$ dan $\mathbf{W}^O \in \mathbb{R}^{hr \times d}$ di mana setiap *head* akan diinisialisasi secara acak.

Decoder dan Masked Attention

Bagian *decoder* menggunakan *masked multi-head attention* untuk mencegah model mengakses informasi masa depan pada saat proses prediksi. Masking ini memastikan bahwa prediksi harga pada waktu ke-t hanya bergantung pada data hingga waktu ke-(t-1). Output dari *decoder* kemudian diteruskan ke lapisan *feed-forward* untuk menghasilkan nilai prediksi akhir.

Optimisasi Model dengan Adam Optimizer

Proses pelatihan model Transformer dilakukan menggunakan *Adam optimizer*. Nama Adam berasal dari *adaptive moment estimation*. *Adam optimizer* adalah salah satu algoritma optimasi yang didasarkan pada gradien orde pertama dari fungsi tujuan stokastik berdasarkan estimasi momen orde rendah. Metode ini sangat efisien dalam perhitungan karena membutuhkan sedikit memori, tidak terpengaruh oleh gradien yang diskalakan, dan sangat cocok untuk data berskala besar atau memiliki banyak parameter (Kingma & Ba, 2014). Variasi laju pembelajaran selama proses pelatihan ditunjukkan dalam Persamaan .

$$lrate = d_{\text{model}}^{-0.5} \cdot \min(step_num^{-0.5}, step_num \cdot warmup_steps^{-1.5}), \quad (6)$$

dengan $warmup_steps = 5000$ and $step_num$ adalah jumlah epoch yang dijalankan.

Loss Function

Loss function adalah metode yang digunakan untuk mengevaluasi seberapa baik algoritma yang digunakan untuk memodelkan data. *Loss function* diperoleh dari perbandingan hasil prediksi dari lapisan output dengan hasil prediksi. Dalam studi ini, MSE digunakan untuk menghitung *loss function*. Rumus *loss function* MSE dapat ditunjukkan dalam Persamaan (18) (Goodfellow et al., 2016).

$$\text{loss}(\hat{x}, x) = \frac{1}{N} \sum_{i=1}^N (\hat{x}_i - x_i)^2 \quad (7)$$

dimana, N adalah panjang data, \hat{x}_i adalah hasil prediksi output, and x_i adalah data observasi.

Evaluasi Model

Root Mean Square Error (RMSE) digunakan untuk mengukur kesalahan atau ketidakakuratan pengukuran berdasarkan dua nilai, yaitu nilai prediksi dan nilai aktual. RMSE juga dapat digunakan untuk mengukur kualitas model setelah mendapatkan model. Semakin kecil RMSE, semakin baik hasil peramalan (Goodfellow et al., 2016). Rumus RMSE dapat ditunjukkan pada Persamaan (19).

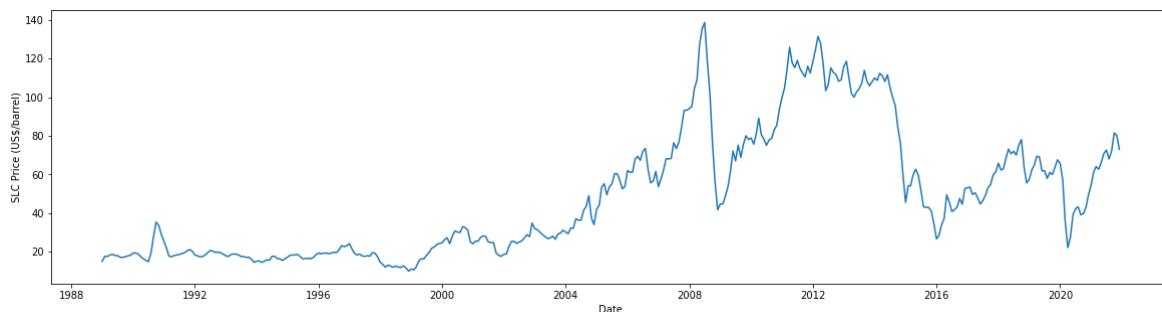
$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{x}_i - x_i)^2} \quad (8)$$

dimana, N adalah panjang data, \hat{x}_i adalah hasil prediksi output, and x_i adalah data observasi.

Hasil dan Pembahasan

Tren Harga Minyak SLC

Sumatran Light Crude Oil (SLC), yang merupakan salah satu jenis minyak yang digunakan sebagai dasar pengindeksan beberapa jenis minyak di Indonesia, mengalami perubahan dari tahun ke tahun. Harga SLC selalu berubah setiap bulan mengikuti harga minyak mentah di Indonesia. Gambar 4 menunjukkan pergerakan harga SLC bulanan selama periode 1989 hingga 2021.



Gambar 3 Tren harga minyak SLC periode 1989-2021

Gambar 4 menunjukkan harga SLC dari Januari 1989 hingga Desember 1998 yang stabil di sekitar US\$ 17,51 per barel, kemudian cenderung meningkat hingga mencapai US\$ 138,73 per barel pada Juli 2008. Pada Agustus hingga Desember 2007, harganya tiba-tiba turun sebesar 69,96% hingga mencapai US\$ 41,69 per barel. Pada periode berikutnya, harga naik kembali, namun pada Juni 2014 hingga Januari 2015, harga turun lagi hingga mencapai US\$ 26,63 per barel. Fenomena ini disebabkan oleh faktor pasokan dan permintaan minyak mentah, namun faktor pasokan menjadi yang paling berpengaruh. Selain itu, harga dapat turun drastis akibat kebijakan OPEC pada November sebelumnya tentang tidak membatasi produksi minyak[21] . Kenaikan harga SLC juga terjadi pada Juni 2021 dari sebelumnya hanya US\$ 66,25 per barel menjadi US\$ 70,82 per barel.

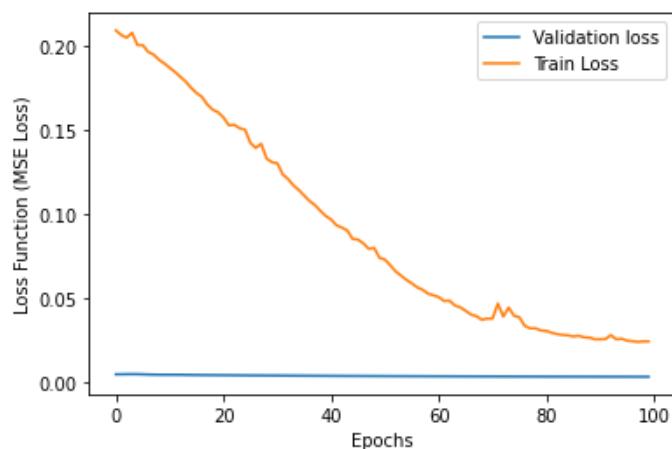
Hasil Performa Model *Transformer*

Dalam studi ini, harga SLC akan diprediksi menggunakan model *transformer*. Akan dilakukan percobaan untuk setiap model menggunakan jumlah *head* dan layer yang berbeda untuk mendapatkan model terbaik yang memiliki kinerja baik dalam memprediksi harga SLC. Adapun setup dari model *Transformer* yang digunakan adalah menggunakan $h = 8$ dan $h = 16$ serta $d_{model} = 512$. Dimensi d_{model} ini mengikuti makalah asli *Transformer*[5] . Jumlah lapisan juga diuji dengan variasi 3, 4, dan 5 lapisan. Tingkat drop-out yang digunakan dalam studi ini adalah 0,1. Tabel 2 menunjukkan RMSE untuk setiap model yang dilatih selama 100 epoch. Tabel 2 menunjukkan bahwa model dengan 4 *layer* memiliki kinerja yang lebih baik dalam memprediksi harga SLC. Model dengan multi-*head attention* 8 dan 16 pada 4 *layer* memiliki nilai RMSE yang lebih kecil dibandingkan model yang menggunakan 5 atau 3 lapisan *attention*. Model terbaik adalah model *transformer* dengan multihead *attention* 16 pada 4 lapisan yang mencapai RMSE sekitar 8.19818. Oleh karena itu, model terbaik ini akan digunakan untuk meramalkan harga SLC.

Tabel 2 Hasil performa model *transformer*

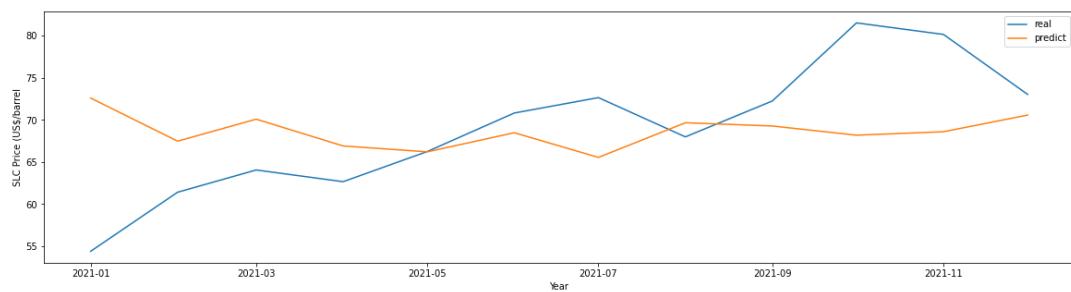
<i>d_{model}</i>	<i>head</i>	<i>n layer</i>	RMSSE
512	8	5	12.35000
512	8	4	12.00347
512	8	3	16.49219
512	16	4	8.19818
512	16	5	11.86072
512	16	3	21.08466

Gambar 5 mengilustrasikan tren penurunan pada training loss seiring dengan bertambahnya jumlah epoch. Meskipun terdapat sedikit fluktuasi di beberapa titik, tren secara umum menunjukkan penurunan hingga mencapai titik konvergensi (stabil). Menariknya, validation loss menunjukkan nilai yang lebih rendah dibandingkan training loss dan bergerak stabil menuju konvergensi. Cela (gap) yang cukup signifikan antara kedua kurva ini mengindikasikan bahwa model dapat menangkap pola pada data validasi dengan lebih mudah dibandingkan pada data latih.



Gambar 4 Loss function model *transformer*

Selanjutnya, Gambar 6 memvisualisasikan hasil peramalan pada data uji. Terlihat bahwa kurva hasil prediksi mampu mengikuti pola data aktual (observasi) dengan cukup akurat.



Gambar 5 Hasil actual vs prediksi model *transformer*

Hasil pada Gambar 6 menunjukkan bahwa model *Transformer* mampu mengikuti arah tren umum dari data aktual, namun prediksinya cenderung lebih rata dan kurang responsif terhadap fluktuasi tajam. Model terlihat melakukan *regression to the mean*, sehingga kenaikan dan penurunan besar pada data real tidak tertangkap dengan baik, meskipun pola besarnya tetap terbaca. Secara keseluruhan, model menangkap struktur dasar deret waktu, tetapi belum akurat dalam memodelkan variasi lokal, sehingga presisi prediksi perlu ditingkatkan melalui penyesuaian parameter, penambahan data, atau perluasan window input.

Pembahasan Hasil Penelitian

Berdasarkan hasil analisis, pergerakan harga Minyak Sumatran Light Crude Oil (SLC) selama periode pengamatan menunjukkan dinamika yang sangat fluktuatif, dengan fase stabil, lonjakan tajam, hingga penurunan drastis yang dipengaruhi oleh kebijakan produksi dan kondisi pasar minyak global. Pola tersebut menegaskan bahwa harga SLC tidak hanya dipengaruhi oleh perubahan jangka pendek, tetapi juga oleh ketergantungan historis yang panjang dan kompleks. Dalam konteks ini, penerapan model *Transformer-based deep learning* terbukti mampu menangkap struktur dasar deret waktu harga SLC, khususnya dalam merepresentasikan tren jangka panjang. Hasil pengujian berbagai konfigurasi model menunjukkan bahwa peningkatan kompleksitas arsitektur hingga titik tertentu berkontribusi positif terhadap akurasi, di mana model dengan 16 *multi-head attention* dan 4 lapisan menghasilkan performa terbaik dengan nilai RMSE sebesar 8,19818. Selama proses pelatihan, model menunjukkan konvergensi yang stabil dan kemampuan generalisasi yang baik, sebagaimana tercermin dari perilaku *training loss* dan *validation loss* yang konsisten. Meskipun demikian, hasil peramalan pada data uji mengindikasikan bahwa prediksi cenderung lebih halus dibandingkan data aktual, sehingga fluktuasi ekstrem belum sepenuhnya terakomodasi. Temuan ini menunjukkan bahwa *Transformer* efektif dalam membaca arah dan struktur utama pergerakan harga SLC, namun sensitivitas terhadap variasi lokal masih menjadi tantangan yang perlu diperhatikan dalam pengembangan model selanjutnya.

Kesimpulan

Penelitian ini menunjukkan bahwa pendekatan *Transformer-based deep learning* merupakan alternatif yang relevan dan menjanjikan untuk peramalan harga Minyak Mentah Ringan Sumatra (SLC), khususnya dalam menangkap pola umum dan tren jangka panjang pada data harga minyak yang bersifat volatil. Kontribusi utama penelitian ini terletak pada penerapan dan evaluasi model *Transformer* pada konteks harga minyak mentah Indonesia, yang selama ini masih relatif jarang dikaji dibandingkan pendekatan statistik konvensional atau model rekuren. Hasil penelitian mengindikasikan bahwa arsitektur *Transformer* mampu melakukan generalisasi dengan baik terhadap data historis dan memberikan gambaran tren pergerakan harga secara konsisten.

Meskipun demikian, penelitian ini memiliki beberapa keterbatasan. Model yang dibangun masih bergantung pada satu variabel utama berupa harga historis, sehingga belum sepenuhnya mempertimbangkan pengaruh faktor eksternal seperti kondisi makroekonomi, kebijakan energi, maupun dinamika geopolitik global. Selain itu, penggunaan data bulanan membatasi kemampuan model dalam merepresentasikan fluktuasi jangka pendek dan perubahan harga yang bersifat ekstrem.

Oleh karena itu, penelitian selanjutnya disarankan untuk mengintegrasikan variabel eksogen yang relevan serta menggunakan resolusi data yang lebih tinggi guna meningkatkan sensitivitas model terhadap variasi lokal. Pengembangan lebih lanjut juga dapat dilakukan dengan mengeksplorasi varian Transformer khusus untuk *time-series forecasting* atau pendekatan hibrida untuk memperoleh keseimbangan antara akurasi tren jangka panjang dan respons terhadap fluktuasi jangka pendek.

Daftar Pustaka

- Farsani, R. M., & Pazouki, E. (2021). A *transformer self-attention* model for time series forecasting. *Journal of Electrical and Computer Engineering Innovations (JECEI)*, 9(1), 1–10.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT press.
- Harrou, F., Sun, Y., Hering, A. S., & Madakyaru, M. (2020). *Statistical process monitoring using advanced data-driven and deep learning approaches: theory and practical applications*. Elsevier.
- Hochreiter, S., & Schmidhuber, J. (1997). *Long Short-Term Memory*. *Neural Computation*, 9(8), 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *ArXiv Preprint ArXiv:1412.6980*.
- Li, S., Jin, X., Xuan, Y., Zhou, X., Chen, W., Wang, Y.-X., & Yan, X. (2019). Enhancing the locality and breaking the memory bottleneck of *transformer* on time series forecasting. *Advances in Neural Information Processing Systems*, 32.
- Lim, B., Arik, S. O., Loeff, N., & Pfister, T. (2019). Temporal fusion *transformers* for interpretable multi-horizon time series forecasting. *ArXiv Preprint ArXiv:1912.09363*.
- Pal, G., Li, G., & Atkinson, K. (2018). *Big Data Real Time Ingestion and Machine Learning*. <https://doi.org/10.1109/DSMP.2018.8478598>
- Pandey, S., & Jain, M. (2021). *Time Series Visualization using Transformer for Prediction of Natural Catastrophe*.
- Pariaman, H., Luciana, G. M., Wisyaldin, M. K., & Hisjam, M. (2021). *Anomaly Detection Using LSTM-Autoencoder to Predict Coal Pulverizer Condition on Coal-Fired Power Plant*.
- Penetapan Harga Minyak Mentah Januari 2022, Pub. L. No. 11.K/MG.03/DJM/2022 (2022).
- Sholikhah, A. (2022). *Pengembangan Model Mixture Autoregressive (STUDI KASUS: Peramalan Harga Sumatran Light Crude)*. Institut Teknologi Sepuluh Nopember.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 30.
- Wen, Q., Zhou, T., Zhang, C., Chen, W., Ma, Z., Yan, J., & Sun, L. (2022). *Transformers in Time Series: A Survey*. *ArXiv Preprint ArXiv:2202.07125*.
- Wu, N., Green, B., Ben, X., & O'Banion, S. (2020). Deep *transformer* models for time series

- forecasting: The influenza prevalence case. *ArXiv Preprint ArXiv:2001.08317*.
- Wu, S., Xiao, X., Ding, Q., Zhao, P., Wei, Y., & Huang, J. (2020). Adversarial sparse *transformer* for time series forecasting. *Advances in Neural Information Processing Systems*, 33, 17105–17115.
- Yi, S., Chen, X., & Tang, C. (2021). Tsformer: Time series *Transformer* for tourism demand forecasting. *ArXiv Preprint ArXiv:2107.10977*.
- Zhao, R., Yan, R., Wang, J., & Mao, K. (2017). Learning to monitor machine health with convolutional bi-directional LSTM networks. *Sensors*, 17(2), 273.
- Zhou, H., Zhang, S., Peng, J., Zhang, S., Li, J., Xiong, H., & Zhang, W. (2021). Informer: Beyond efficient *transformer* for long sequence time-series forecasting. *Proceedings of AAAI*.