Jurnal Pendidikan dan Teknologi Indonesia (JPTI)
Vol. 5, No. 9, September 2025, Hal. 2866-2877
DOI: https://doi.org/10.52436/1.jpti.1091
p-ISSN: 2775-4227
e-ISSN: 2775-4219

Prediksi Harga Bitcoin Menggunakan Model Hibrida Transformer Dan Long Short Term Memory

Rio Irawan[1], Ema Utami[2]

[1,2]Magister of Informatics, Amikom University, Yogyakarta, Indonesia
Email: [1]rio.blt@students.amikom.ac.id, [2]ema.u@amikom.ac.id

Abstrak

Peramalan harga Bitcoin merupakan tantangan signifikan akibat fluktuasi pasar yang tinggi, di mana model deret waktu konvensional seringkali gagal menangkap dependensi data jangka pendek dan panjang secara bersamaan. Untuk mengatasi masalah tersebut, penelitian ini bertujuan mengembangkan dan mengevaluasi model *deep learning* hibrida guna meningkatkan presisi peramalan harga Bitcoin. Model yang diusulkan mengintegrasikan arsitektur *Long Short-Term Memory* (LSTM), yang andal dalam memproses data sekuensial, dengan mekanisme atensi pada *Transformer* yang mampu mengidentifikasi hubungan data yang relevan. Metode penelitian mencakup penggunaan data harga historis Bitcoin yang telah melalui tahap pra-pemrosesan dan normalisasi. Arsitektur model terdiri dari dua lapisan LSTM (30 unit), diikuti oleh lapisan *Transformer* dengan *Multi-Head Attention* (satu *head*, dimensi kunci empat), dan diakhiri dengan lapisan *Global Average Pooling*. Pelatihan model dilakukan menggunakan konfigurasi optimal selama 10 *epoch* dengan ukuran *batch* 64. Hasil evaluasi menunjukkan kinerja prediksi yang sangat akurat, dibuktikan dengan nilai *Mean Absolute Percentage Error* (MAPE) sebesar 0,000695 dan *R-squared* (R^2) 0,99999. Temuan ini menegaskan bahwa pendekatan hibrida efektif dalam menangkap pola kompleks pada data harga Bitcoin, sehingga menawarkan alat yang lebih kuat untuk analisis keuangan di pasar *cryptocurrency*.

Kata kunci: *Deep Learning, LSTM, Model Hibrida, Peramalan Harga Bitcoin, Transformer.*

Bitcoin Price Prediction Using a Hybrid Transformer and Long Short-Term Memory Model

Abstract

Forecasting Bitcoin prices is a significant challenge due to high market volatility, where conventional time series models often fail to capture both short-term and long-term data dependencies. To address this issue, this study aims to develop and evaluate a hybrid deep learning model to improve the precision of Bitcoin price forecasting. The proposed model integrates the Long Short-Term Memory (LSTM) architecture, which excels at processing sequential data, with the attention mechanism of the Transformer, capable of identifying relevant data relationships. The methodology involves using historical Bitcoin price data that was preprocessed and normalized. The model architecture consists of two LSTM layers (30 units), a Transformer layer with Multi-Head Attention (one head, a key dimension of four), and a Global Average Pooling layer. The model was trained using an optimal configuration of 10 epochs and a batch size of 64. The evaluation results demonstrate highly accurate predictive performance, evidenced by a Mean Absolute Percentage Error (MAPE) of 0.000695 and an R-squared (R^2) score of 0.99999. These findings confirm that the hybrid approach effectively captures complex patterns in Bitcoin price data, offering a more robust tool for financial analysis in the cryptocurrency market.

Keywords: *Bitcoin Price Forecasting, Deep Learning, Hybrid Model, LSTM, Transformer.*

1. PENDAHULUAN

Bitcoin merupakan aset digital pertama dan paling dominan dalam ekosistem mata uang kripto, yang diperkenalkan sebagai sistem keuangan alternatif berbasis teknologi blockchain pada tahun 2009 [1]. Sejak saat itu, Bitcoin telah mengalami perkembangan nilai yang sangat signifikan disertai volatilitas harga yang ekstrem [2]. Fluktuasi tajam ini, yang terjadi dalam waktu yang sangat singkat, menjadikan Bitcoin tidak hanya menarik sebagai objek investasi [3], tetapi juga menimbulkan tantangan besar dalam hal peramalan harga. Ketidakpastian harga ini

p-ISSN: 2775-4227
e-ISSN: 2775-4219

dapat menimbulkan risiko tinggi bagi investor, pelaku pasar, serta pengembang sistem perdagangan algoritmik yang mengandalkan ketepatan dalam memprediksi harga untuk mengambil keputusan secara otomatis.

Harga Bitcoin dipengaruhi oleh berbagai faktor multidimensional, termasuk dinamika pasar global, kebijakan regulasi, sentimen media sosial [4], aktivitas spekulatif, dan kondisi ekonomi makro[5]. Kompleksitas ini menghasilkan data deret waktu yang bersifat non-linier, tidak stasioner, dan dinamis. Dalam kondisi tersebut, pendekatan konvensional berbasis model statistik klasik seperti Autoregressive Integrated Moving Average (ARIMA) [6] menjadi kurang efektif karena model-model tersebut cenderung bergantung pada asumsi linieritas dan kestasioneran data, yang tidak sesuai dengan karakteristik pasar kripto.

Seiring berkembangnya teknologi komputasi dan kecerdasan buatan, pendekatan berbasis machine learning [7] dan deep learning [8] telah muncul sebagai solusi yang lebih adaptif dan fleksibel dalam menangani prediksi deret waktu kompleks. Model deep learning memiliki kemampuan untuk menangkap hubungan non-linier, menemukan pola tersembunyi dalam data, serta melakukan pemodelan jangka panjang berdasarkan data historis. Khususnya, model Recurrent Neural Network (RNN) [9] dan turunannya seperti Long Short-Term Memory (LSTM) [10] telah banyak digunakan dalam berbagai aplikasi deret waktu, termasuk prediksi harga saham [11], cuaca , dan harga cryptocurrency.

Beberapa penelitian telah mencoba meningkatkan akurasi prediksi menggunakan model-model tersebut. Salah satunya adalah pendekatan gabungan *Wavelet Transform* dan *Bi-Directional LSTM* (Bi-LSTM) oleh Parameswaran et al. [12] untuk memprediksi harga SHIBA INU (SHIB). Meskipun menunjukkan hasil yang baik, pendekatan tersebut masih memiliki keterbatasan dalam menangkap dependensi kontekstual jangka panjang, karena struktur sekuensial LSTM memproses informasi secara linier.

Untuk mengatasi keterbatasan tersebut, penelitian ini mengusulkan pengembangan model prediksi harga Bitcoin berbasis arsitektur hibrida yang menggabungkan Long Short-Term Memory (LSTM) dan Transformer. Arsitektur LSTM dirancang untuk menangani masalah jangka pendek dan menengah dalam deret waktu, dengan kemampuan mempertahankan informasi penting melalui mekanisme sel memori internalnya. Sementara itu, Transformer yang diperkenalkan pertama kali dalam bidang pemrosesan bahasa alami (Natural Language Processing) menggunakan mekanisme self-attention untuk mengidentifikasi dan mempelajari hubungan kontekstual jangka panjang secara efisien tanpa mengandalkan urutan waktu secara eksplisit [13].

Dengan menggabungkan kedua arsitektur tersebut, diharapkan model prediksi yang dihasilkan mampu memanfaatkan keunggulan dari masing-masing pendekatan: LSTM dalam hal dependensi temporal dan Transformer dalam hal perhatian global terhadap pola antar waktu. Kombinasi ini tidak hanya memberikan ketahanan terhadap noise dan fluktuasi mendadak dalam data kripto, tetapi juga mampu menangkap struktur makro dan mikro dari pola harga historis Bitcoin.

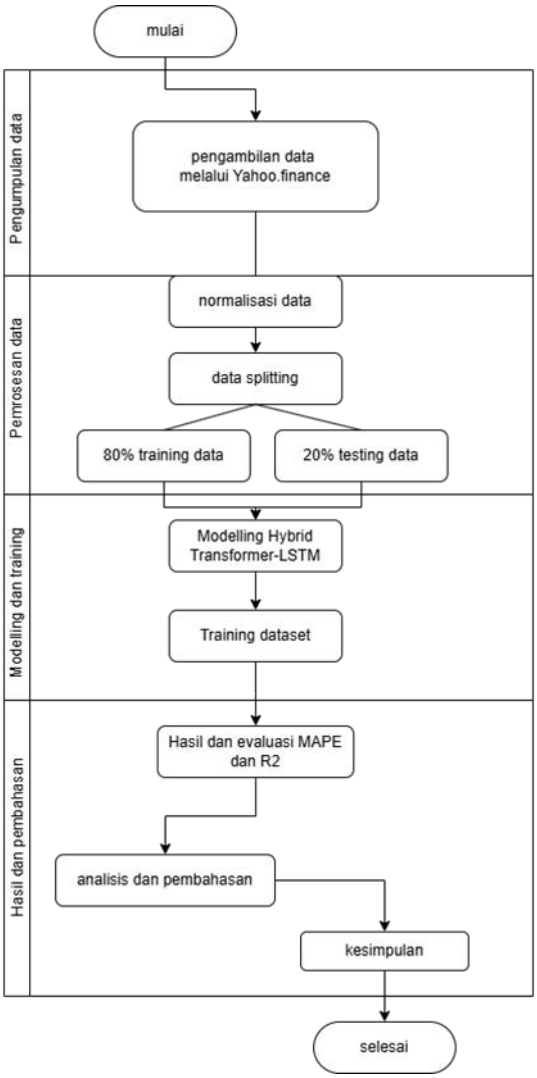
Tujuan utama dari penelitian ini adalah untuk mengembangkan, menerapkan, dan mengevaluasi model prediksi harga Bitcoin berbasis pendekatan hibrida LSTM-Transformer. Ruang lingkup penelitian mencakup tahapan pengumpulan data historis harga Bitcoin dari sumber terpercaya, pemrosesan awal data melalui normalisasi dan segmentasi, perancangan arsitektur model yang terdiri dari dua lapisan LSTM dan satu lapisan Multi-Head Attention dari Transformer, serta pelatihan model menggunakan konfigurasi hiperparameter optimal, yakni batch size 64 dan epoch sebanyak 10. Evaluasi performa dilakukan menggunakan metrik evaluasi standar seperti Mean Absolute Percentage Error (MAPE) dan koefisien determinasi (R-squared, R^2) [14], untuk menilai akurasi dan presisi model.

Kontribusi utama penelitian ini adalah perancangan dan validasi model hibrida inovatif yang menggabungkan keunggulan LSTM dan Transformer untuk prediksi harga Bitcoin. Model ini tidak hanya berpotensi meningkatkan akurasi peramalan, tetapi juga dapat menjadi acuan untuk pengembangan sistem perdagangan otomatis yang lebih adaptif. Dengan demikian, penelitian ini menawarkan solusi untuk mengatasi tantangan prediktabilitas di pasar kripto yang dinamis serta memperkaya literatur di bidang peramalan keuangan berbasis kecerdasan buatan.

2. METODE PENELITIAN

Untuk memperjelas metode yang digunakan dalam penelitian ini, Gambar 1 di bawah ini menggambarkan alur penelitian [15]. Alur penelitian dimulai dari tahap pengumpulan data historis harga Bitcoin, kemudian dilanjutkan dengan proses normalisasi data untuk memastikan keseragaman skala. Setelah itu, data dibagi menjadi dua bagian melalui proses *data splitting*, yakni data pelatihan dan data pengujian. Tahap selanjutnya adalah proses *modelling* menggunakan arsitektur hibrida Transformer dan LSTM. Setelah model dilatih, dilakukan evaluasi terhadap hasil prediksi menggunakan metrik yang relevan. Tahap akhir berupa pembahasan terhadap hasil evaluasi guna menarik kesimpulan dan memberikan interpretasi terhadap performa model yang dibangun.

p-ISSN: 2775-4227
e-ISSN: 2775-4219



Gambar 1. Alur Penelitian

2.1. Dataset

Dataset yang digunakan dalam penelitian ini diunduh dari platform Kaggle melalui tautan https://www.kaggle.com/datasets/mczielinski/bitcoin-historical-data [16]. Dataset ini merupakan data historis harga Bitcoin dengan resolusi satu menit (1-minute interval), yang mencakup berbagai atribut seperti waktu (timestamp), harga pembukaan (open), harga tertinggi (high), harga terendah (low), harga penutupan (close), dan volume transaksi (volume). Secara keseluruhan, dataset ini terdiri dari lebih dari 6,7 juta baris data yang merepresentasikan pergerakan harga Bitcoin secara rinci dari waktu ke waktu. Gambar 2 memperlihatkan tampilan struktur awal dataset, sebelum dilakukan proses transformasi lebih lanjut. Dataset ini bersifat deret waktu (time series) dan digunakan sebagai dasar untuk pelatihan serta pengujian model prediksi harga.

p-ISSN: 2775-4227
e-ISSN: 2775-4219

	Timestamp	Open	High	Low	Close	Volume
0	2012-01-01 10:01:00	4.58	4.58	4.58	4.58	0.0
1	2012-01-01 10:02:00	4.58	4.58	4.58	4.58	0.0
2	2012-01-01 10:03:00	4.58	4.58	4.58	4.58	0.0
3	2012-01-01 10:04:00	4.58	4.58	4.58	4.58	0.0
4	2012-01-01 10:05:00	4.58	4.58	4.58	4.58	0.0

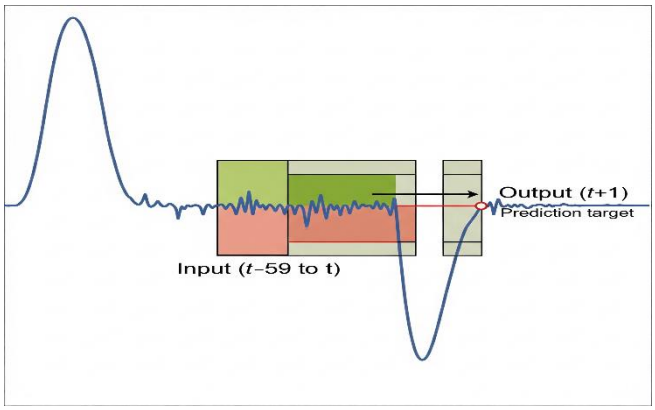
	Timestamp	Open	High	Low	Close	Volume
6742275	2025-01-10 23:56:00	94725.0	94725.0	94690.0	94691.0	0.901410
6742276	2025-01-10 23:57:00	94691.0	94691.0	94672.0	94680.0	1.067289
6742277	2025-01-10 23:58:00	94680.0	94713.0	94680.0	94710.0	2.215471
6742278	2025-01-10 23:59:00	94710.0	94710.0	94700.0	94702.0	1.324348
6742279	2025-01-11 00:00:00	94702.0	94702.0	94678.0	94685.0	3.407287

Gambar 2. Data Historis Bitcoin

2.2. Preprocessing Data

Sebelum digunakan dalam proses pemodelan, data mentah dari dataset harga historis Bitcoin terlebih dahulu melalui beberapa tahapan preprocessing. Tahap awal yang dilakukan adalah konversi kolom *timestamp* dari format UNIX time menjadi format *datetime* yang lebih mudah dibaca dan diolah, khususnya untuk keperluan visualisasi dan analisis deret waktu. Selanjutnya, dilakukan pengecekan terhadap nilai-nilai pada kolom numerik seperti *Open*, *High*, *Low*, *Close*, dan *Volume*. Semua kolom ini dikonversi ke tipe data numerik, dan entri yang mengandung nilai kosong atau tidak valid dihapus untuk menjaga integritas data. Dari seluruh atribut yang tersedia, hanya kolom *Close* yang dipilih sebagai target utama dalam prediksi harga, karena kolom ini secara umum digunakan sebagai representasi harga akhir pada suatu waktu tertentu dalam konteks perdagangan aset digital. Nilai-nilai dalam kolom *Close* [17] kemudian dinormalisasi ke dalam rentang [0, 1] menggunakan teknik *Min-Max Scaling*. Penggunaan *Min-Max digunakan karena* kemampuannya untuk menjaga hubungan proporsional antar nilai dalam data asli dan efektif untuk algoritma *deep learning* yang menggunakan fungsi aktivasi seperti Sigmoid atau Tanh, yang sensitif terhadap skala input. Teknik ini memastikan proses pelatihan berjalan lebih stabil dan konvergen lebih cepat tanpa mengubah bentuk distribusi data secara signifikan.

Setelah normalisasi, data dibentuk menjadi struktur sekuens menggunakan pendekatan *sliding window*. (Sebuah jendela (*window*) dengan panjang 60 langkah waktu (*timesteps*) digunakan sebagai input (fitur) untuk memprediksi harga pada langkah waktu ke-61 (target). Proses ini digambarkan pada Gambar 3. Pendekatan ini memungkinkan model untuk belajar dari pola historis dalam durasi waktu tertentu.) Terakhir, data dibagi menjadi 80% data pelatihan (5.393.775 entri) dan 20% data pengujian (1.348.444 entri) untuk evaluasi objektif.



Gambar 3. Sliding Window

p-ISSN: 2775-4227
e-ISSN: 2775-4219

2.3. Transformer

Transformer adalah arsitektur yang mengandalkan mekanisme *self-attention* untuk memproses seluruh urutan input secara paralel[18]. Dalam konteks deret waktu, *Transformer* mampu menangkap dependensi jangka panjang secara efisien. Komponen utamanya adalah *Multi-Head Attention*, *Feed Forward Network*, dan *Positional Encoding*. Mekanisme atensi dihitung menggunakan *query* (Q), *key* (K), dan *value* (V) sesuai Persamaan (1) .

$$Attention\ Q,K,V = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (1)$$

Di mana d_k adalah dimensi dari *key*. Rumus ini menghitung skor atensi untuk memberikan bobot pada setiap bagian dari input berdasarkan relevansinya. Dalam penelitian ini, key_dim diatur ke nilai 4. Pemilihan dimensi kunci yang relatif kecil ini disengaja untuk mengurangi kompleksitas komputasi dan mencegah *overfitting*, mengingat model hanya memproses satu fitur univariat (harga penutupan), sehingga mekanisme atensi yang terlalu kompleks tidak diperlukan.

Dalam model hybrid yang digunakan pada penelitian ini, Transformer berperan sebagai komponen yang mengidentifikasi keterkaitan global antar elemen waktu secara paralel. Setelah data diproses oleh LSTM, hasilnya dipadukan dengan output dari blok Transformer. Proses ini memungkinkan pemodelan konteks jangka panjang dengan pemahaman global yang lebih menyeluruh.

2.4. Long Short Term Memory

Long Short-Term Memory [19] adalah varian dari RNN yang dirancang untuk mengatasi masalah *vanishing gradient* dan menangkap dependensi jangka panjang melalui sel memori dan mekanisme gerbang (*gate*). Arsitekturnya terdiri dari tiga gerbang utama: *input gate* (i_t), *forget gate* (f_t), dan *output gate* (o_t).

Gerbang Lupa (Forget Gate): Menentukan informasi yang akan dibuang dari sel memori.

$$f_t = \sigma\left(W_f . [h_{t-1}, x_t] + b_f\right) \quad (2)$$

Gerbang Masukan (Input Gate): Menentukan informasi baru yang akan disimpan.

$$i_t = \sigma\left(W_i . [h_{t-1}, x_t] + b_i\right) \quad (3)$$

$$\tilde{c}_t = \tanh\left(W_c . [h_{t-1}, x_t] + b_c\right) \quad (4)$$

Pembaruan Sel Memori (Cell State Update): Menggabungkan informasi lama dan baru.

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{c}_t \quad (5)$$

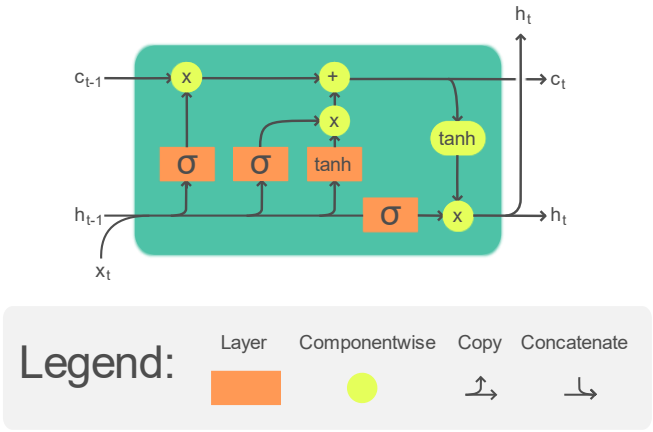
Gerbang Keluaran (Output Gate): Menentukan output berdasarkan sel memori.

$$o_t = \sigma\left(W_o . [h_{t-1}, x_t] + b_o\right) \quad (6)$$

$$h_t = o_t \odot \tanh\left(c_t\right) \quad (7)$$

Di mana xt adalah input, ht−1 adalah *hidden state* sebelumnya, Ct−1 adalah *cell state* sebelumnya, W dan b adalah bobot dan bias, σ adalah fungsi sigmoid, dan ⊙ adalah perkalian Hadamard. Dalam penelitian ini, setiap lapisan LSTM memiliki 30 unit. Jumlah unit ini dipilih sebagai kompromi antara kapasitas model untuk menangkap pola kompleks dan risiko *overfitting*. Jumlah yang moderat memastikan model cukup ekspresif tanpa menjadi terlalu berat secara komputasi. Gambar 4 merupakan bentuk dari arsitektur Long Short Term Memory.

p-ISSN: 2775-4227
e-ISSN: 2775-4219



Gambar 4. Arsitektur Long Short Term Memory

2.5. Hyperparameter tunning

Pelatihan model dilakukan selama 10 *epoch*. Dengan dataset yang sangat besar (lebih dari 5 juta data latih), jumlah *epoch* yang relatif sedikit sudah cukup untuk memungkinkan model mencapai konvergensi tanpa risiko *overfitting*. Pengamatan selama eksperimen menunjukkan bahwa penambahan *epoch* tidak memberikan peningkatan performa yang signifikan namun secara drastis meningkatkan waktu komputasi. Ukuran *batch* (*batch size*) yang diuji adalah 32, 64, dan 128 untuk menemukan keseimbangan optimal antara kecepatan pelatihan dan stabilitas gradien.

2.6. Model Evaluasi

Untuk mengevaluasi performa model prediksi harga Bitcoin, digunakan dua metrik utama, yaitu *Mean Absolute Percentage Error* (MAPE) dan *Coefficient of Determination* (R^2) [14]. MAPE digunakan untuk mengukur tingkat kesalahan prediksi dalam bentuk persentase terhadap nilai aktual. Metrik ini banyak digunakan karena memberikan hasil dalam satuan persen, sehingga memudahkan interpretasi dan memungkinkan perbandingan performa antar model. Secara matematis, MAPE didefinisikan sebagai berikut:

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{\hat{y}_i}\right| \times 100 \qquad (8)$$

Sementara itu, Koefisien Determinasi atau R-Squared (R^2) mengukur proporsi variabilitas dalam data target yang dapat dijelaskan oleh model prediksi. Nilai R^2 berada dalam rentang antara 0 hingga 1, di mana nilai yang mendekati 1 menunjukkan bahwa model memiliki kemampuan yang baik dalam menjelaskan variansi data. Rumus R^2 dinyatakan sebagai:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2} \qquad (9)$$

3. HASIL DAN PEMBAHASAN

Bab ini membahas hasil eksperimen yang dilakukan untuk mengevaluasi performa model Hybrid [20] Transformer dan Long Short-Term Memory dalam dua skenario berbeda. Skenario pertama menggunakan arsitektur LSTM murni tanpa komponen Transformer, sementara skenario kedua menggunakan model hibrida yang menggabungkan LSTM dan Transformer.

Setiap skenario diuji dengan beberapa konfigurasi ukuran batch (batch size), yaitu 32, 64, dan 128, dengan jumlah epoch yang tetap sebanyak 10. Evaluasi performa model dilakukan menggunakan sejumlah metrik utama, seperti Mean Absolute Percentage Error (MAPE) dan koefisien determinasi (R^2), untuk mengukur tingkat akurasi dan presisi prediksi harga Bitcoin. Selain itu, dilakukan analisis visual terhadap hasil prediksi untuk menilai sejauh mana model mampu mengikuti pola aktual pergerakan harga. Seluruh eksperimen dirancang untuk menguji

p-ISSN: 2775-4227
e-ISSN: 2775-4219

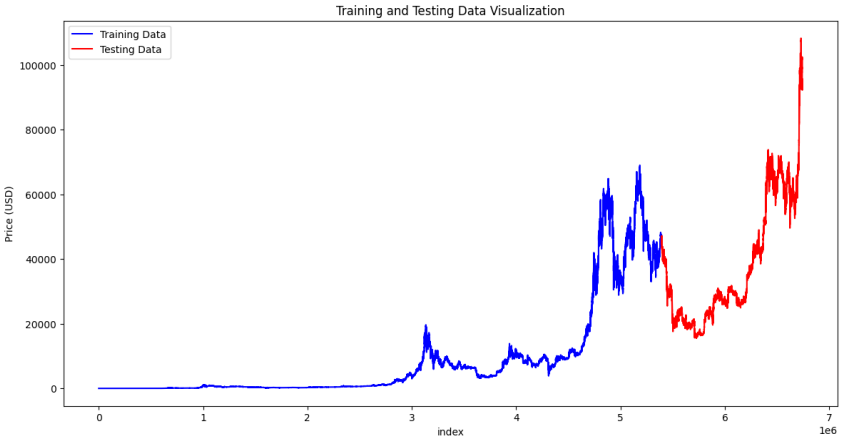
efektivitas kombinasi arsitektur LSTM dan Transformer dalam memodelkan data deret waktu kripto yang bersifat non-linier dan sangat fluktuatif.

3.1. Dataset

Dataset yang digunakan dalam penelitian ini bersifat deret waktu kontinu dengan resolusi sangat tinggi, yakni per satu menit. Data diambil dari repositori Kaggle dan mencakup lebih dari 6,7 juta baris, menjadikannya salah satu dataset kripto paling rinci yang tersedia secara publik. Atribut yang tersedia dalam dataset antara lain: timestamp, open, high, low, close, dan volume. Setiap baris merepresentasikan aktivitas harga dalam jendela waktu satu menit, memungkinkan pengamatan sangat granular terhadap dinamika pasar Bitcoin.

Pemilihan resolusi 1-menit memiliki implikasi penting. Sebagian besar penelitian sebelumnya cenderung menggunakan resolusi harian atau per jam, yang meskipun lebih mudah diolah, sering kali kehilangan banyak informasi penting terkait lonjakan harga jangka pendek dan pola volatilitas mikro. Dalam pasar mata uang kripto yang beroperasi 24/7 dan sangat dipengaruhi oleh sentimen pasar sesaat, menggunakan resolusi tinggi seperti ini memungkinkan model untuk mengenali pola mikro-tren yang sering luput dalam resolusi rendah. Hal ini penting terutama untuk strategi perdagangan frekuensi tinggi (high-frequency trading).

Dataset juga mencerminkan berbagai kondisi pasar, seperti lonjakan tajam akibat berita global, periode stabil karena rendahnya volume perdagangan, serta efek akhir pekan atau hari libur tertentu. Oleh karena itu, model yang dilatih dengan data ini diharapkan lebih adaptif terhadap berbagai situasi pasar. Namun, besarnya ukuran dataset ini juga menghadirkan tantangan tersendiri. Dataset yang terdiri dari jutaan entri membutuhkan optimasi pemrosesan data yang baik dan pengelolaan memori yang efisien, terutama saat diolah dengan framework deep learning seperti TensorFlow atau PyTorch. Oleh karena itu, dilakukan pembagian dataset menjadi 80% data latih (5.393.775 baris) dan 20% data uji (1.348.444 baris), serta implementasi pendekatan sliding window [21] untuk membentuk struktur input-output yang representatif dari pola temporal. Visualisasi pembagian dataset dapat dilihat pada gambar 5.



Gambar 5. Visualisasi pembagian data

3.2. Konfigurasi Eksperimen

Untuk memastikan pengujian yang adil dan komprehensif, arsitektur dasar model ditetapkan seperti yang dirangkum pada Tabel 1. Konfigurasi ini menjadi fondasi untuk model hibrida yang diuji. Eksperimen kemudian difokuskan pada penyesuaian *hyperparameter* ukuran *batch* (*batch size*), dengan tiga variasi (32, 64, dan 128) untuk setiap skenario model. Jumlah *epoch* dijaga konstan sebanyak 10 pada semua pengujian.

Tabel 1. Hasil Evaluasi Model LSTM tanpa Transformer

Komponen	Konfigurasi
Input Layer	Shape = (60, 1) dari sliding window
LSTM Layer 1	30 unit, return_sequences=True
LSTM Layer 2	30 unit
Transformer Layer	MultiHeadAttention (num_heads=1, key_dim=4)

p-ISSN: 2775-4227
e-ISSN: 2775-4219

3.3. Skenario 1

Skenario pertama melibatkan penggunaan model LSTM tanpa integrasi komponen Transformer. Model ini mengandalkan kekuatan memori jangka panjang dari LSTM untuk mempelajari pola historis data harga Bitcoin. Tiga konfigurasi batch size diuji: 32, 64, dan 128. Hasil evaluasi model ditampilkan pada Tabel 2 berikut:

Tabel 2. Hasil Evaluasi Model LSTM tanpa Transformer

No.	Epoch	Batch Size	MAPE	R^2
1	10	32	0.0011279	0.999935
2	10	64	0.0006479	0.999961
3	10	128	0.0009631	0.999983

Hasil dari Tabel 2 menunjukkan bahwa batch size 64 menghasilkan performa prediksi terbaik. Studi oleh Zhong [19]menegaskan bahwa konfigurasi batch size menengah memberikan trade-off terbaik antara stabilitas model dan efisiensi pelatihan. Selain itu, hasil ini juga didukung oleh penelitian Saud dan Shakya [9] yang menunjukkan bahwa LSTM cukup efektif dalam mempelajari pola temporal jangka pendek, namun memiliki keterbatasan dalam mengenali pola jangka panjang dan kompleks.

3.4. Skenario 2

Skenario kedua menggunakan model hybrid yang menggabungkan kekuatan LSTM dan Transformer. LSTM menangani dependensi sekuensial, sedangkan Transformer menyoroti hubungan temporal secara global melalui mekanisme perhatian (attention). Tiga konfigurasi batch size juga diuji.

Tabel 3. Hasil Evaluasi Model Hybrid LSTM – Transformer

No.	Epoch	Batch Size	MAPE	R^2
1	10	32	0.0008982	0.999978
2	10	64	0.0006954	0.999994
3	10	128	0.0014954	0.999571

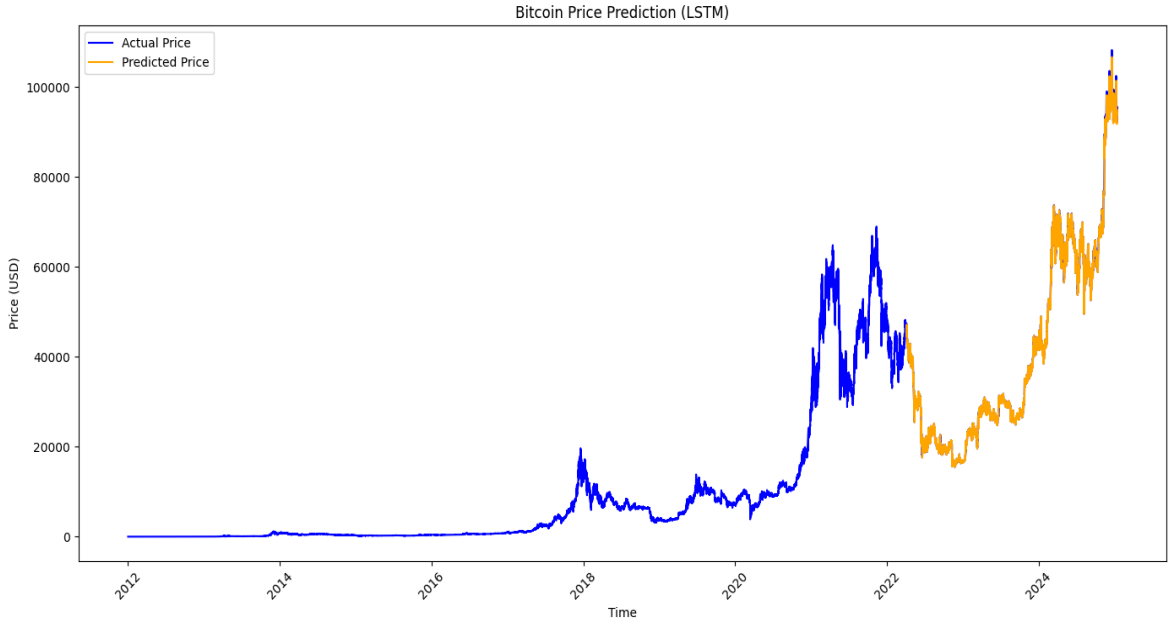
Konfigurasi batch size 64 kembali menunjukkan hasil terbaik. Penelitian oleh Al-Ali et al menunjukkan bahwa integrasi Transformer ke dalam arsitektur LSTM memberikan peningkatan signifikan dalam menangani deret waktu kompleks seperti produksi energi dan harga pasar. Kinerja model yang tinggi dengan R^2 mencapai 0.999994 memperlihatkan bahwa model mampu menjelaskan hampir seluruh variabilitas dalam data, yang sejalan dengan studi Dong et al [21] mengenai efektivitas pendekatan sliding window dan LSTM untuk prediksi berbasis waktu.

3.5. Visualisasi hasil prediksi

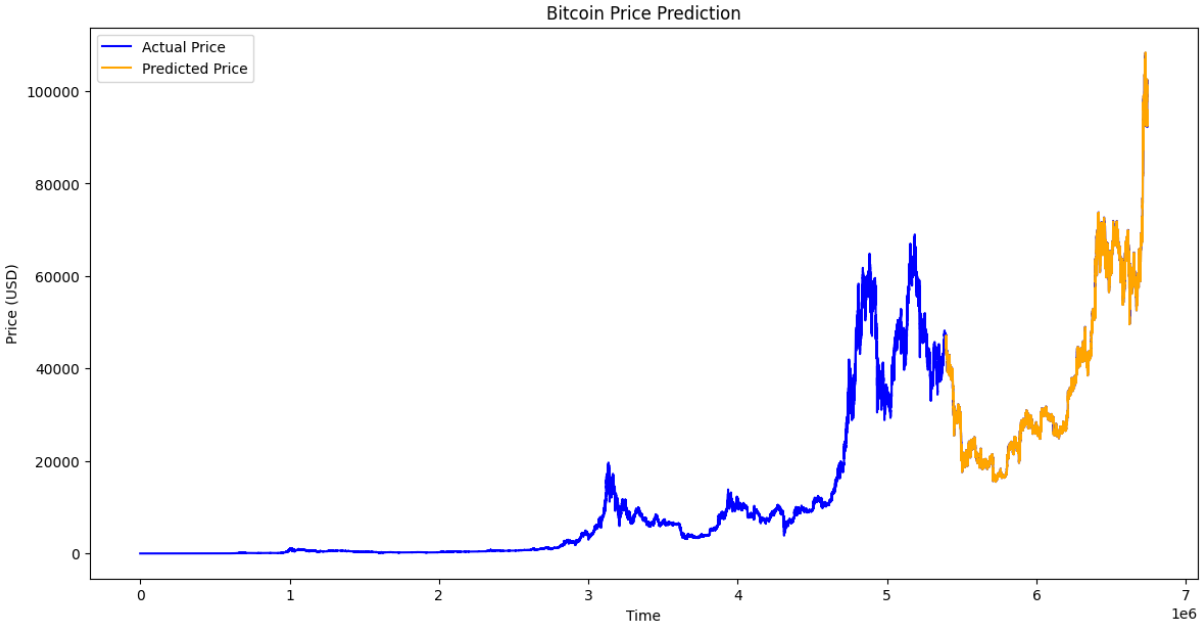
Visualisasi hasil prediksi sangat penting dalam menilai keberhasilan model secara intuitif dan mendalam. Pada penelitian ini, digunakan grafik perbandingan antara harga aktual dan hasil prediksi pada subset data uji untuk menilai akurasi temporal model.

Gambar 6 menunjukkan bahwa model LSTM-only mengalami keterlambatan respons terhadap perubahan harga ekstrem. Hal ini mengindikasikan keterbatasan LSTM dalam menangkap hubungan global antar waktu, sebagaimana dikemukakan oleh Rothman [18] dalam bukunya yang dikutip secara luas di riset NLP dan time series.

p-ISSN: 2775-4227
e-ISSN: 2775-4219



Gambar 6. Hasil prediksi LSTM



Gambar 7. Hasil prediksi hybrid LSTM-Transformer

Sebaliknya, Gambar 7 menampilkan kurva prediksi dari model hybrid yang nyaris tumpang tindih dengan harga aktual. Ini memperlihatkan bahwa model mampu menangkap fluktuasi ekstrem dengan lebih baik. Penelitian oleh Parameswaran et al [12] mendukung klaim ini, menunjukkan bahwa model yang menggabungkan pemrosesan lokal (seperti LSTM) dengan perhatian global (seperti Transformer) memberikan prediksi lebih stabil dalam konteks kripto.

Visualisasi prediksi memperjelas bahwa pendekatan hybrid mampu memahami dinamika jangka pendek dan jangka panjang secara bersamaan. Hasil ini penting terutama dalam konteks kripto yang sangat fluktuatif seperti dibuktikan oleh Wu dan Ding [5].

3.6. Diskusi

Hasil eksperimen menunjukkan bahwa model hibrida LSTM-*Transformer* memberikan keunggulan performa dibandingkan model LSTM murni. Meskipun model LSTM mampu menangani pola historis secara sekuensial, ia

p-ISSN: 2775-4227
e-ISSN: 2775-4219

menunjukkan kelemahan dalam merespons anomali atau lonjakan harga yang mendadak, sebuah temuan yang sejalan dengan studi oleh Saud dan Shakya [9]. Keterlambatan respons yang terlihat pada Gambar 5 mengindikasikan keterbatasan LSTM dalam menangkap hubungan kontekstual yang non-linier dan bersifat global.

Integrasi *Transformer* terbukti efektif mengatasi keterbatasan ini. Dengan mekanisme *self-attention*, model mampu memberikan bobot pada data historis yang paling relevan tanpa terikat oleh urutan waktu, sehingga dapat belajar secara global. Hasil ini mendukung temuan dari Al-Ali et al [20] dan Dong et al [21] mengenai peran signifikan *Transformer* dalam meningkatkan akurasi prediksi pada data deret waktu yang kompleks. Kemampuan model hibrida untuk mencapai R^2 sebesar 0.999994, seperti terlihat pada Tabel 3, menunjukkan bahwa model mampu menjelaskan hampir seluruh variabilitas dalam data.

Jika dibandingkan dengan penelitian terkait oleh Risna Sari et al. [17] yang menghasilkan R^2 sebesar 0.9962 pada prediksi harga Solana, model hibrida dalam penelitian ini menunjukkan kinerja yang lebih unggul. Keunggulan ini dicapai dengan jumlah *epoch* yang jauh lebih sedikit (10 *epoch*), yang mengindikasikan efisiensi arsitektur hibrida yang diusulkan. Temuan ini memperkuat argumen bahwa kombinasi LSTM dan *Transformer* menawarkan solusi prediktif yang tidak hanya akurat tetapi juga efisien untuk pasar *cryptocurrency* yang sangat dinamis.

Namun, model ini juga memiliki kelemahan. Kompleksitas arsitektur *Transformer* dengan mekanisme *self-attention* secara inheren membutuhkan sumber daya komputasi yang lebih besar dibandingkan model LSTM murni. Meskipun efisien dalam hal *epoch*, proses pelatihan untuk setiap *epoch* bisa lebih lambat. Selain itu, interpretasi terhadap bobot atensi pada data deret waktu finansial masih menjadi tantangan, sehingga sulit untuk mengetahui secara pasti mengapa model memberikan perhatian lebih pada titik data tertentu.

Secara akademik, penelitian ini memberikan bukti empiris yang kuat mengenai efektivitas arsitektur hibrida LSTM-*Transformer* untuk data *cryptocurrency* yang sangat volatil. Temuan ini memperkaya literatur di bidang peramalan keuangan dengan menyajikan sebuah arsitektur yang dapat menjadi tolok ukur (*benchmark*) untuk penelitian selanjutnya. Studi ini juga mendorong pergeseran dari model sekuensial murni ke pendekatan berbasis atensi yang lebih dinamis untuk analisis deret waktu.

Dari sisi praktis, hasil penelitian ini memiliki implikasi signifikan. Model dengan akurasi tinggi ini dapat diintegrasikan langsung ke dalam sistem perdagangan otomatis (*trading bots*) untuk meningkatkan profitabilitas dan mengurangi risiko. Selain itu, model ini dapat digunakan sebagai alat bantu dalam manajemen portofolio dan analisis risiko, memberikan proyeksi harga yang lebih andal bagi investor dan analis keuangan untuk pengambilan keputusan yang lebih baik.

Meskipun memberikan hasil yang positif, penelitian ini memiliki beberapa keterbatasan yang perlu diakui. Pertama, model ini hanya menggunakan data univariat, yaitu harga penutupan (*close price*). Padahal, pergerakan harga Bitcoin juga dipengaruhi oleh faktor lain seperti volume perdagangan, sentimen media sosial, dan indikator makroekonomi. Pengabaian fitur-fitur ini dapat membatasi kemampuan prediksi model dalam skenario pasar yang lebih kompleks. Kedua, penelitian ini hanya menguji model pada satu aset kripto, yaitu Bitcoin. Generalisasi performa model pada aset lain dengan karakteristik volatilitas yang berbeda masih belum teruji. Terakhir, eksplorasi *hyperparameter* terbatas pada beberapa konfigurasi. Penyelidikan yang lebih mendalam terhadap arsitektur, seperti jumlah *head* pada *Multi-Head Attention* atau jumlah lapisan, berpotensi menemukan konfigurasi yang lebih optimal.

4. KESIMPULAN

Penelitian ini bertujuan untuk menjawab permasalahan utama dalam prediksi harga Bitcoin yang bersifat non-linier dan tidak stasioner, yaitu bagaimana merancang model yang mampu menangkap pola jangka pendek dan panjang secara efektif dalam data deret waktu kripto. Hasil penelitian menunjukkan bahwa model hybrid yang menggabungkan arsitektur Long Short-Term Memory (LSTM) dan Transformer mampu menjawab permasalahan tersebut secara terukur. Model ini berhasil mencapai nilai Mean Absolute Percentage Error (MAPE) sebesar 0.000695 dan koefisien determinasi (R^2) sebesar 0.999994 dengan konfigurasi optimal (batch size 64 dan 10 epoch). Hal ini membuktikan bahwa model mampu melakukan prediksi harga Bitcoin secara sangat akurat dan stabil, bahkan dalam kondisi pasar yang sangat fluktuatif.

Keunggulan model ini terletak pada integrasi dua pendekatan: LSTM yang efektif menangkap pola temporal jangka pendek, dan Transformer yang unggul dalam memahami hubungan jangka panjang melalui mekanisme self-attention. Temuan ini tidak hanya meningkatkan akurasi prediksi, tetapi juga menunjukkan bahwa pendekatan hybrid lebih efisien dibandingkan pendekatan tradisional yang memerlukan jumlah epoch lebih besar untuk mencapai akurasi serupa.

Berdasarkan hasil eksperimen yang telah dilakukan, saran untuk penelitian selanjutnya adalah agar fokus pada optimalisasi waktu komputasi dengan menerapkan teknik seperti pruning, distillation, atau penggunaan arsitektur Transformer yang lebih ringan seperti Linformer dan Performer. Selain itu, penelitian lanjutan

p-ISSN: 2775-4227
e-ISSN: 2775-4219

disarankan untuk menguji generalisasi model pada aset kripto lain seperti Ethereum atau Solana, serta pada berbagai resolusi temporal (per jam, harian, atau real-time) guna mengevaluasi adaptabilitas model terhadap berbagai kondisi pasar. Penambahan fitur eksternal seperti sentimen media sosial dan indikator makroekonomi juga dapat meningkatkan kemampuan prediktif model.

DAFTAR PUSTAKA

[1] M. Conti, K. E. Sandeep, C. Lal, and S. Ruj, "A survey on security and privacy issues of bitcoin," *IEEE Communications Surveys and Tutorials*, vol. 20, no. 4, 2018, doi: 10.1109/COMST.2018.2842460.

[2] L. Ø. Bergsli, A. F. Lind, P. Molnár, and M. Polasik, "Forecasting volatility of Bitcoin," *Res Int Bus Finance*, vol. 59, 2022, doi: 10.1016/j.ribaf.2021.101540.

[3] X. Song, "Predicting cryptocurrency investment suitability using machine learning techniques," *Applied and Computational Engineering*, vol. 29, no. 1, 2023, doi: 10.54254/2755-2721/29/20230726.

[4] L. Afuan and N. Hidayat, "Sentiment Analysis of the Kampus Merdeka Program on Twitter Using Support Vector Machine and a Feature Extraction Comparison: TF-IDF vs. FastText," *Journal of Applied Data Sciences*, vol. 5, no. 4, pp. 1738–1753, Dec. 2024, doi: 10.47738/jads.v5i4.436.

[5] X. Wu and S. Ding, "The impact of the Bitcoin price on carbon neutrality: Evidence from futures markets," *Financ Res Lett*, vol. 56, 2023, doi: 10.1016/j.frl.2023.104128.

[6] M. D. Angelo, I. Fadhiilrahman, and Y. Purnama, "Comparative Analysis of ARIMA and Prophet Algorithms in Bitcoin Price Forecasting," in *Procedia Computer Science*, Elsevier B.V., 2023, pp. 490–499. doi: 10.1016/j.procs.2023.10.550.

[7] Ardiansyah Risko Anwari and S. Sukirman, "RECOMMENDATION SYSTEM TO SELECT A MAJOR OF VOCATIONAL SCHOOL USING DECISION TREE," *Jurnal Teknik Informatika (Jutif)*, vol. 5, no. 2, pp. 589–598, Apr. 2024, doi: 10.52436/1.jutif.2024.5.2.1327.

[8] S. Kashyap and M. Singh, "A Deep Learning Approach for Crypto Price Prediction," *International Research Journal of Engineering and Technology*, vol. 3, no. 3, 2022.

[9] A. S. Saud and S. Shakya, "Analysis of look back period for stock price prediction with RNN variants: A case study on banking sector of NEPSE," in *Procedia Computer Science*, Elsevier B.V., 2020, pp. 788–798. doi: 10.1016/j.procs.2020.03.419.

[10] S. Samsul, I. Tahyudin, and F. Setyo utomo, "Perbandingan Model CNN, LSTM, dan FNN dalam Klasifikasi Kulit Penderita Diabetes," *Jurnal Pendidikan dan Teknologi Indonesia*, vol. 5, no. 3, pp. 571–583, Mar. 2025, doi: 10.52436/1.jpti.536.

[11] N. Afrianto, D. H. Fudholi, and S. Rani, "Prediksi Harga Saham Menggunakan BiLSTM dengan Faktor Sentimen Publik," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 6, no. 1, pp. 41–46, Feb. 2022, doi: 10.29207/resti.v6i1.3676.

[12] S. E. Parameswaran, V. Ramachandran, and S. Shukla, "Crypto Trend Prediction Based on Wavelet Transform and Deep Learning Algorithm," in *Procedia Computer Science*, Elsevier B.V., 2024, pp. 1179–1189. doi: 10.1016/j.procs.2024.04.112.

[13] D. Rothman, *Transformers for Natural Language Processing Build innovative deep neural network architectures for NLP & Python, PyTorch, TensorFlow, BERT, RoBERTa & more*. 2021.

[14] D. Chicco, M. J. Warrens, and G. Jurman, "The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation," *PeerJ Comput Sci*, vol. 7, pp. 1–24, 2021, doi: 10.7717/PEERJ-CS.623.

[15] S. Mounika, "Crypto-Currency Price Prediction using CNN and LSTM Models," *Int J Res Appl Sci Eng Technol*, vol. 9, no. 3, 2021, doi: 10.22214/ijraset.2021.33191.

[16] S. Singh, A. Pise, and B. Yoon, "Prediction of bitcoin stock price using feature subset optimization," *Heliyon*, vol. 10, no. 7, Apr. 2024, doi: 10.1016/j.heliyon.2024.e28415.

[17] R. Sari, K. Kusrini, T. Hidayat, and T. Orphanoudakis, "Improved LSTM Method of Predicting Cryptocurrency Price Using Short-Term Data," *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, vol. 17, no. 1, p. 33, Feb. 2023, doi: 10.22146/ijccs.80776.

[18] D. Rothman, *Transformers for natural language processing : build, train, and fine-tuning deep neural network architectures for NLP with Python, PyTorch, TensorFlow, BERT, and GPT-3*. 2022.

p-ISSN: 2775-4227
e-ISSN: 2775-4219

[19] B. Zhong, "Deep learning integration optimization of electric energy load forecasting and market price based on the ANN–LSTM–transformer method," *Front Energy Res*, vol. 11, 2023, doi: 10.3389/fenrg.2023.1292204.

[20] E. M. Al-Ali *et al.*, "Solar Energy Production Forecasting Based on a Hybrid CNN-LSTM-Transformer Model," *Mathematics*, vol. 11, no. 3, 2023, doi: 10.3390/math11030676.

[21] L. Dong *et al.*, "Prediction of streamflow based on dynamic sliding window lstm," *Water (Switzerland)*, vol. 12, no. 11, 2020, doi: 10.3390/w12113032.