

Pemodelan *Dataset On-chain* pada BiLSTM untuk Prediksi Harga *Bitcoin*

Gamar Ramadhani Malik¹, Rizky Parluka², Kartini³

^{1,2,3}Informatika, Fakultas Ilmu Komputer, UPN "Veteran" Jawa Timur

¹gamarramadhani16@gmail.com, ²rizkyparlika.if@upnjatim.ac.id *, ³kartini.if@upnjatim.ac.id

Abstract

Bitcoin is a crypto asset for investment. It can give high profit, but it also has high risk because the price changes very fast and is not stable. To reduce the risk of loss, we need a prediction system that can read price changes well. This research aims to model and predict the closing price of Bitcoin using network activity data (on-chain metrics). The method used is Deep Learning with the BiLSTM algorithm. This method is chosen because it can process data in two directions (forward and backward), so it can learn patterns better than standard LSTM. The dataset is taken from the public Blockchain network using BigQuery, from August 18, 2011, to February 6, 2026, with 5,287 daily data. The model uses the main input active_spending_addresses and two volatility indicators: Percent of Top Range (PTR) and Percent Low Range (PLR). Before modeling, the data is processed using a sliding window of 60 days, with 90% training data and 10% testing data. The results show that the BiLSTM model has high accuracy, with MAE 2.958, RMSE 3.905, and MAPE 3.22%. The comparison shows that BiLSTM is better than other models. LSTM has MAPE 29.06%, and MLP has MAPE 4.01%. In conclusion, BiLSTM can handle extreme crypto market changes very well, so it gives stable and accurate Bitcoin price predictions.

Keywords: bitcoin, price prediction, on-chain data, deep learning, BiLSTM

Abstrak

Bitcoin merupakan instrumen investasi aset kripto berpotensi keuntungan tinggi, namun berisiko besar karena pergerakan harganya sangat fluktuatif dan bersifat non-Gaussian. Guna meminimalkan risiko kerugian bagi investor, diperlukan sistem prediksi komputasi yang konsisten dalam membaca fluktuasi harga tersebut. Penelitian ini bertujuan memodelkan dan memprediksi harga penutupan (close price) Bitcoin berbasis aktivitas jaringan menggunakan metrik on-chain. Metode yang diusulkan adalah Deep Learning dengan algoritma BiLSTM. Metode ini dipilih karena terbukti mampu mengatasi keterbatasan pemrosesan satu arah pada LSTM standar, sehingga memungkinkan model menangkap pola informasi secara lebih optimal dan simultan melalui lapisan maju (forward) dan mundur (backward). Dataset historis diekstrak dari jaringan Blockchain publik melalui BigQuery mencakup periode 18 Agustus 2011 hingga 6 Februari 2026, berisikan 5.287 baris data harian. Model menggunakan input utama active_spending_addresses yang dikombinasikan dengan dua indikator volatilitas hasil rekayasa fitur, yaitu Percent of Top Range (PTR) dan Percent Low Range (PLR). Sebelum pemodelan, dataset diproses menggunakan sekuens sliding window sepanjang 60 hari dengan rasio pembagian data latih dan uji sebesar 90:10. Hasil pengujian menunjukkan bahwa arsitektur BiLSTM menghasilkan akurasi tinggi dengan nilai MAE 2,958, RMSE 3,905, dan MAPE 3,22%. Evaluasi komparasi membuktikan BiLSTM jauh mengungguli performa algoritma pembandingnya, di mana model LSTM menghasilkan MAPE 29,06% dan model Multi-Layer Perceptron (MLP) sebesar 4,01%. Kesimpulannya, pemrosesan dua arah BiLSTM terbukti paling andal dalam menangani volatilitas ekstrem pasar kripto, sehingga menghasilkan kerangka kerja prediksi harga Bitcoin yang stabil, presisi, dan dapat diandalkan.

Kata kunci: bitcoin, prediksi harga, on-chain data, deep learning, BiLSTM

©This work is licensed under a Creative Commons Attribution - ShareAlike 4.0 International License

1. Pendahuluan

Perubahan besar-besaran yang terjadi dalam perkembangan teknologi pada saat ini, memberikan dampak yang signifikan terutama di bidang teknologi dan investasi. Sebagai wujud nyata dari perpaduan kedua bidang tersebut, Bitcoin muncul sebagai mata uang kripto yang paling populer di dunia. Posisinya sebagai aset digital yang paling banyak diperdagangkan menjadikannya instrumen investasi utama yang menyita perhatian besar dari berbagai kalangan investor [1]. Transaksi Bitcoin dapat dilakukan secara langsung antarpihak tanpa harus melibatkan perantara seperti lembaga keuangan atau otoritas pusat. Hal ini dimungkinkan oleh pemanfaatan jaringan peer-to-peer yang dibangun di atas teknologi Blockchain yang bersifat open-source [2].

Bitcoin sendiri dapat menjadi aset yang memiliki resiko tinggi, karena karakteristik dari harganya yang sangat fluktuatif. Nilai Bitcoin dapat berubah drastis dalam waktu singkat sehingga memiliki potensi untuk menghasilkan return yang tidak sebanding terutama jika dibandingkan dengan instrument investasi lain seperti saham dan obligasi. Meskipun dengan resiko yang tinggi Bitcoin tetap menjadi aset dari cryptocurrency dengan kapitalisasi pasar terbesar yang meningkat dari \$128 miliar pada 30 Mei 2018 menjadi \$1.156 miliar pada 1 Maret 2021[3] Setelah mengalami kenaikannya meningkat terus sampai tahun 2021, pada 1 Juni 2022 kapitalisasi pasar Bitcoin turun hingga menyentuh angka \$380 miliar. Kondisi ini sering disebut dengan distribusi yang miring dan tidak mengikuti distribusi normal atau singkatnya non-

Gaussian dan lebih mudah terjadi perubahan harga yang sangat ekstrim baik waktu kenaikan maupun penurunan bisa dalam waktu yang cukup singkat [4].

Bitcoin memiliki sistem yang dapat menjamin untuk melakukan transaksi itu akan aman dan transparan tanpa harus bergantung kepada pihak lainnya atau lembaga hukum[5]. Bitcoin sendiri sudah memiliki sebuah infrastruktur keamanan yang canggih seperti teknologi Blockchain, karena setiap transaksi dari Bitcoin bersifat public dan divalidasi oleh sebuah *node* jaringan[6]. Setiap Bitcoin terjadi transaksi yang baru saja selesai akan diverifikasi ke dalam jaringan Bitcoin dan dimasukkan ke Blockchain dalam bentuk "*block*", sebuah unit untuk menyimpan data yang mempunyai sifat permanen dan tidak bisa mengubah catatan data dengan menggunakan fungsi *hashing*[7].

Penjelasan sebelumnya menegaskan bahwa Bitcoin terbilang sangat aman dan terdesentralisasi dan berkat itu juga perhatian para investor dapat mendorong momentum yang salah karena ekspektasi mereka yang tinggi dan keliru tentang harga Bitcoin di masa depan. Sebagai contoh, penelitian yang dilakukan oleh Brauneis dkk. Menunjukkan terjadinya peningkatan biaya transaksi dan volume transaksi *on-chain* berhubungan dengan peningkatan likuiditas[8]. Penelitian tersebut menemukan bahwa analisis terkait pasar Bitcoin dapat diperluas dengan memasukkan faktor-faktor berdasarkan *on-chain* seperti biaya transaksi, volume transaksi dan *activity addresses* ikut andil dalam menguatkan harga Bitcoin. Peran metrik *on-chain activity addresses* dikuatkan pada studi literatur oleh [9] membuktikan bahwa jaringan Blockchain pada Bitcoin mematuhi prinsip hukum *Metcalf* berdasarkan hukum komputasi ini, nilai dari kapitalisasi dari sebuah jaringan digital khususnya harga Bitcoin memiliki korelasi langsung dengan jumlah *activity addresses* di dalamnya. Dalam analisis *on-chain* metrik *activity addresses* direpresentasikan secara kuantitatif melalui jumlah dari *active spending addresses* setiap harinya.

Berdasarkan hal yang sudah dijelaskan di atas, khususnya mengenai korelasi kuat antara valuasi jaringan Bitcoin dengan *activity spending addresses* memunculkan ide guna memprediksi harga dari Bitcoin berdasarkan data *on-chain* tersebut. Prediksi yang akurat memiliki kegunaan guna meminimalkan dari kerugian dan tentu meningkatkan keuntungan untuk para investor yang investasi di Bitcoin. Hal ini menjadi sangat penting karena pergerakan dari harga Bitcoin dipengaruhi oleh banyak faktor yang sangat kompleks dan sulit diprediksi. Deep Learning merupakan salah satu penerapan teknologi dalam melakukan prediksi harga Bitcoin dengan menggunakan *input* dari data *on-chain* sehingga menjadi pilihan untuk melakukan prediksi sebuah harga Bitcoin yang memiliki sifat volatile. Salah satu banyaknya algoritma dari Deep Learning yang efektif untuk memodelkan harga Bitcoin dan data tipe *time series* adalah *Long Short-Term Memory* (LSTM). LSTM merupakan hasil

pengembangan dari *Recurrent Neural Network* (RNN) yang dirancang untuk menangani data historis dan mampu mengingat informasi jangka Panjang dalam proses prediksi harga Bitcoin[10].

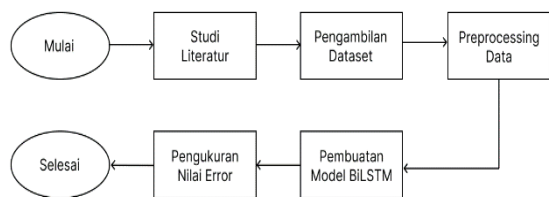
LSTM memiliki keunggulan terletak pada kemampuannya untuk mengatasi permasalahan *vanishing gradient* yang umum terjadi pada RNN, sehingga dapat mengenali pola dan tren jangka panjang pada data *time series*[11]. Keterbatasan dari LSTM hanya memproses data dalam satu arah. Keterbatasan ini mendorong pengembangan *Bidirectional LSTM* (BiLSTM) yang terdiri dari dua jaringan LSTM, yaitu *forward layer* untuk memproses informasi dari arah maju dan *backward layer* untuk memproses informasi dari arah sebaliknya[12]. Dengan memanfaatkan informasi dari kedua arah secara bersamaan, BiLSTM mampu mengenali pola perubahan harga Bitcoin yang bersifat fluktuatif dengan lebih akurat. Pergerakan harga Bitcoin yang sangat volatile dipengaruhi oleh berbagai faktor termasuk *on-chain* seperti *activity spending addresses*, transaksi volume BTC, dan lain-lainnya. Sehingga penggunaan BiLSTM menjadi lebih relevan dalam pemodelan harga *Bitcoin*.

Berbagai penelitian sebelumnya telah dilakukan dengan topik yang serupa. Penelitian yang dilakukan oleh Puteri (2023) menunjukkan bahwa penggunaan model BiLSTM yang memproses data dua arah berhasil memprediksi harga tiga saham syariah (ANTM.JK, UNVR.JK, dan INDF.JK) periode 2018-2023 dengan akurat menghasilkan nilai MAPE masing-masing sebesar 2.59%, 1.77%, dan 1.05%[13]. Penelitian berikutnya yang dilakukan oleh Yang dan Wang (2022) menunjukkan bahwa penggunaan model BiLSTM yang memproses data dua arah berhasil mengatasi masalah rendahnya akurasi pada data *time series* yang bersifat *non-linear* dan volatil. Pengujian pada data harga *close price* indeks CSI 300 membuktikan bahwa model ini mampu menghasilkan prediksi yang akurat dan stabil dengan perolehan nilai RMSE sebesar 0.96 dan MAPE sebesar 0.79%[14].

Oleh karena itu, algoritma BiLSTM lebih cocok dijadikan sebuah hipotesa untuk melakukan kemampuan dalam menangani volatilitas dan data yang *non-linear* yang biasa terjadi di pasar Bitcoin. Dengan menggunakan BiLSTM, prediksi harga Bitcoin menggunakan dataset *on-chain* yang fokusnya pada *activity spending addresses* dan serta dua indikator hasil dari rekayasa fitur volatilitas harga yaitu *Percent of Top Range* (PTR) dan *Percent Low Range* (PLR) sebagai input dalam model dan *Close* sebagai target dari model. Sehingga dapat dilakukan dengan lebih akurat dan mampu untuk menangkap pola dari pergerakan harga dalam jangka panjang.

2. Metode Penelitian

Penelitian ini dilakukan dengan menggunakan beberapa rangkaian yang sistematis sehingga menjadi lebih teratur dan dapat dijabarkan pada Gambar 1 sebagai berikut.



Gambar 1. Alur Metode Penelitian

Setiap dari langkah pada Gambar 1 mempunyai peran masing-masing yang sama pentingnya di penelitian ini dan memastikan juga dapat berjalan dengan sistematis. Penjelasan lebih lanjut mengenai masing-masing tahapan dari awal sampai tahap evaluasi akan dijelaskan pada di bawah ini.

2.1 Studi Literatur

Tahapan pertama kali yang dilakukan adalah studi literatur dimana peneliti melakukan kajian-kajian di beberapa referensi ataupun sumber ilmiah yang relevan dengan topik pada penelitian ini, khususnya pada analisis data terhadap data *on-chain*, prediksi harga, Bitcoin, dan penerapan algoritma BiLSTM pada tipe data *time series*.

2.2 Pengambilan Dataset

Tahap kedua ialah pengambilan dataset dengan menggunakan metode *scrapping* yang dilakukan dengan memperoleh data *on-chain* Bitcoin dari *BigQuery* dengan mengambil fitur *activity spending addresses* yang nantinya akan digunakan sebagai *input*. Target dari model ini ada pada harga *Close* Bitcoin yang nantinya akan diprediksi. Adapun jumlah data yang digunakan berjumlah 5287 data, pengambilan data dimulai pada tanggal 18 Agustus 2011 sampai dengan 06 Februari 2026 dengan *timeframe* satu jam dan dapat dilihat pada Tabel 1.

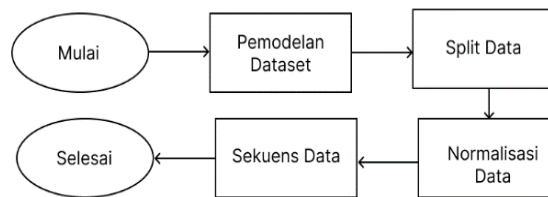
Tabel 1. Dataset On-Chain

date	activity spending addresses	Close
18/08/2011	10,461	10,9
19/08/2011	10,825	11,69
20/08/2011	9,487	11,7
21/08/2011	10,400	11,7
...
06/02/2026	409,686	67,005

Langkah pertama dan terakhir dalam pengambilan dataset adalah melakukan mengambil data dari *BigQuery*. Dapat dilihat pada Tabel 1, dataset ini akan mengambil fitur *on-chain* yaitu *active_spending_addresses*, Data ini nantinya akan memberikan gambaran mengenai hasil dari transaksi yang dicatat pada saat terjadinya transaksi.

2.3 Preprocessing Data

Pada tahap ini dilakukannya *preprocessing* data yang memiliki tujuan untuk memproses data sebelum memasuki tahapan dari *training model*, proses ini guna menyiapkan dataset nantinya akan siap pakai, sesuai dengan model dan bersih untuk masuk ke tahap pemodelan dapat dilihat pada Gambar 2.



Gambar 2. Flowchart Preprocessing Data

Selesai melakukan pengambilan data, langkah berikutnya ialah melakukan *preprocessing data* seperti yang sudah ditunjukkan pada Gambar 2 menunjukkan alur dari proses *preprocessing data*. Diawali dengan tahapan pemodelan dataset, kemudian dilanjutkan dengan melakukan *split data* menjadi 2 bagian *training data dan test data*, setelah itu dilakukan normalisasi data agar skala dari variabel menjadi seimbang, lalu memasuki tahap *sekuensi data* yang mengubah bentuknya data agar sesuai dengan yang dibutuhkan oleh model di tahap *reshape data*. Adapun penjelasan lebih lanjut terkait tahapan-tahapan yang ada pada proses *preprocessing data* ini akan dijelaskan di bawah ini.

a. Pemodelan Dataset

Pada tahap ini, data dari historis pasar yang masih mentah akan ditransformasikan menggunakan pendekatan *Patterned Dataset Model* untuk memodelkan dan mengukur volatilitas dari harga Bitcoin. Model ini memproses data histori yang meliputi harga tertinggi harian (*High/H*), harga terendah harian (*Low/L*), dan harga saat ini/penutupan (*Current/Close/C*). Dari variabel-variabel tersebut dilakukan ekstraksi fitur matematika untuk mendapatkan nilai *Total Range (R)*, *Top Range (TR)*, *Lower Range (LR)*, *Percentage Top Range (PTR)*, dan *Percentage Lower Range (PLR)*[15]. Rumus yang digunakan dalam pembentukan dataset ini adalah sebagai berikut:

$$R = H - L \quad (1)$$

$$TR = H - C \quad (2)$$

$$LR = C - L \quad (3)$$

$$PTR = \frac{TR}{R} \times 100\% \quad (4)$$

$$PLR = \frac{LR}{R} \times 100\% \quad (5)$$

Melalui pemodelan *dataset* dan persamaan diatas nantinya akan menghasilkan *dataset* baru yang mengelompokkan pergerakan aset ke dalam kondisi pasar tertentu. Jika pasar cenderung menunjukkan nilai $PTR = 100$ dan $PLR = 0$ maka data tersebut akan diklasifikasikan ke dalam penurunan harga atau *dip*. Sebaliknya jika nilai $PLR = 100$ dan $PTR = 0$ data diklasifikasikan ke dalam kondisi kenaikan harga atau *rocket*. Hasil dari pemodelan ini menghasilkan kolom baru yaitu *PTR* dan *PLR* bisa dilihat pada Tabel 2.

Tabel 2. Transformasi ke Pemodelan Dataset

Date	High	Low	Close	PTR(%)	PLR(%)
18/08/2011	10,9	10,7	10,9	0	100
19/08/2011	11,85	10,9	11,69	16,84	83,16
20/08/2011	11,7	11,69	11,7	0	100
21/08/2011	11,71	11,7	11,7	100	0

...
06/02/2026	67630	59930	67005	8,12	91,88

Berdasarkan pada Tabel 2 memvisualisasikan proses tranformasi dari data harga historis mentah menjadi indikator persentase dari sentiment pasar yang berupa PTR dan PLR. Hasil dari perhitungan matematis pada fitur tersebut, pergerakan aset dapat diklasifikasikan ke dalam beberapa kondisi pasar yang spesifik; contohnya pada tanggal 18 dan 20 Agustus 2011 harga di penutupan menyentuh nilai tertinggi harian sehingga bisa menghasilkan nilai PTR 0% dan PLR 100% yang mengindikasikan kondisi *rocket* (kenaikan dominan), sedangkan pada tanggal 21 Agustus 2011 harga dari penutupan anjlok ke titik terendah dengan PTR 100% dan PLR 0% yang menandakan kondisi *dip* (penurunan dominan). Adapun hari-hari dengan pergerakan fluktuatif seperti pada 19 Agustus 2011 dan 6 Februari 2026, nilai dari harga penutupan di antara rentang harga *High* dan *Low* sehingga kondisinya disebut dengan fluktuatif atau *sideway*. Kedua fitur nantinya akan dikombinasikan dengan aktivitas jaringan *on-chain* yaitu *active spending addresses* digunakan untuk memperkaya konteks data *time series* yang akan dipelajari oleh model BiLSTM

b. Split Data

Tahap *split data* merupakan proses penting dalam memisahkan *dataset* yang telah melalui tahapan sebelumnya. Prosesnya memisahkan *dataset* menjadi 2 bagian yaitu *data training* dan *test*. *data training* digunakan dalam membangun dan melatih model sedangkan *data test* berfungsi untuk mengevaluasi kemampuan model. Untuk pembagian data bisa dilihat pada Tabel 3.

Tabel 3. Jumlah Data Training dan Test

Rasio(%)	Jumlah Data	
	Training	Test
90:10	4,704	523

c. Normalisasi Data

Tahap selanjutnya yaitu normalisasi data yang bertujuan agar skala nilai dari setiap fitur dalam rentang 0-1. Dataset yang digunakan memiliki perbedaan nilai skala yang cenderung jauh antar fitur dan model BiLSTM sensitif dengan perbedaan skala tersebut. Dengan normalisasi proses dari *training model* akan menjadi lebih stabil, akurat, dan menghindari dominan dari fitur nilai numerik yang besar. Untuk visualisasi dari normalisasi dapat dilihat pada Tabel 4.

Tabel 4. Data Sebelum dan Sesudah Normalisasi

<i>active spending addresses</i>	
Sebelum	Sesudah
10,461	0,73
10,825	1
9,487	0
10,400	0,68

Berdasarkan pada Tabel 4 dapat dilihat perbedaan dari fitur *tx_count* diubah dalam rentang skala nilai 0 hingga 1 dengan menggunakan *Min-Max*. Adapun rumus dari *Min-Max* yang selanjutnya akan diimplementasikan

pada keseluruhan di dataset sehingga skalanya akan menjadi seimbang. Berikut persamaan matematis untuk melakukan normalisasi di bawah ini.

$$x_n = \frac{x_0 - x_{min}}{x_{max} - x_{min}} \quad (6)$$

Keterangan:

- x_n : Nilai hasil normalisasi
- x_0 : Nilai data aktual
- x_{min} : Nilai minimal dari data aktual
- x_{max} : Nilai maksimal dari data aktual

d. Sekuens Data

Tahap terakhir dari *preprocessing data* yaitu pembentukan sekuens data yang bertujuan mengubah *data time series* menjadi bentuk yang dapat diproses oleh model BiLSTM. Pada penelitian ini menerapkan metode *sliding window* dengan panjang sekuens sebanyak 60 hari. Pada tabel 5 di bawah akan ditampilkan hasil dari sekuens data dan hanya menampilkan fitur dari *active spending addresses* fitur lainnya akan diproses juga.

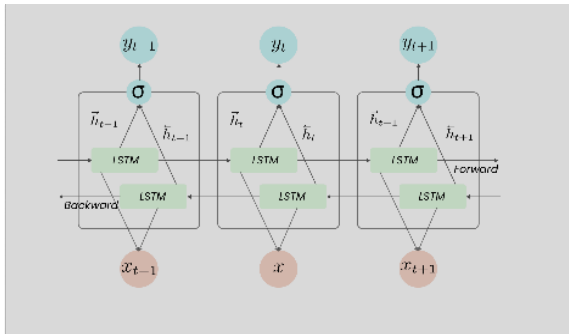
Tabel 5. Sekuens Data

No	Hari ke-1	Hari ke-2	...	Hari ke-60	Prediksi Hari ke-61
1.	10,461	10,825	...	7,400	8,628
2	10,825	9,487	...	8,628	7,125
3	9,487	10,400	...	7,125	7,522

Berdasarkan Tabel 5 setiap baris data merepresentasikan satu sampel sekuens yang dibentuk menggunakan metode *sliding window*. Pada baris pertama, nilai dari *active spending addresses* pada hari ke-1 hingga hari ke-60 digunakan sebagai data *input* untuk memprediksi nilai dari fitur *tx_count* pada hari ke-61. Selanjutnya pada baris ke-2 dan seterusnya hingga hari ke-60 digunakan untuk memprediksi nilai pada hari ke-61.

2.4 Pembuatan Model BiLSTM

Proses pembuatan model prediksi menggunakan metode BiLSTM dan menggunakan bahasa python. *Bidirectional Long Short-Term Memory* (BiLSTM) merupakan hasil dari pengembangan LSTM yang menggabungkan dua jaringan LSTM dan dilakukan secara paralel, jaringan yang pertama membaca dan mempelajari data dalam urutan awal hingga akhir (*forward direction*) sedangkan jaringan kedua memproses data dari arah sebaliknya memproses data dari akhir menuju awal (*backward direction*). Jaringan *forward LSTM* dan *backward LSTM* memiliki kesamaan dari komposisi dan prinsip kerja yang membedakan ialah urutan *input* yang berlawanan[16]. Hasil dari kedua pemrosesan tersebut digabung pada setiap urutan waktu sehingga model mampu menangkap informasi masa dari masa lalu dan memanfaatkan pola yang berasal dari masa selanjutnya[17]. Arsitektur BiLSTM dapat dilihat pada Gambar 3.



Gambar 3. Arsitektur Model BiLSTM

Pada Gambar 3 menunjukkan arsitektur dari model BiLSTM serta proses kerja dari model. Menjelaskan proses arah forward merupakan keluaran dari waktu ke-t dapat dihitung dengan persamaan dalam matematika sebagai $\vec{h}_t = LSTM(x_t, h_{t-1})$ yang memanfaatkan input sekarang dan *hidden state* dari *timestamp* sebelumnya. Sementara itu, untuk arah *backward* perhitungan bisa dilakukan dengan persamaan $\vec{h}_t = LSTM(x_t, h_{t+1})$ Model akan memproses hasil urutan data dari akhir ke awal. Hasil akhir BiLSTM penggabungan dari kedua *hidden state* tersebut yang biasa simbolkan dalam persamaan $h_t = [\vec{h}_t, \vec{h}_t]$ [18]

$$y_t = U_y \vec{h}_t + W_y \vec{h}_t + b_y \quad (7)$$

Keterangan:

y_t = Output jaringan pada waktu t

U_y = Nilai dari bobot *output gate* pada \vec{h}_t

\vec{h}_t = Nilai dari *output* pada LSTM *forward*

W_y = Nilai dari bobot *output gate* pada \vec{h}_t

\vec{h}_t = Nilai dari *output* pada LSTM *backward*

b_y = Vektor bias

Pada tahap perancangan arsitektur model BiLSTM ditetapkan terlebih dahulu nilai dari parameter pendukung yang dapat ditunjukkan pada Tabel 6.

Tabel 6. Parameter Model

Parameter	Fungsi
<i>Optimizer</i>	Meningkatkan kinerja model dengan mengubah bobot dan bias, yang akan digunakan ialah Adam. Dikarenakan dapat memberikan kecepatan untuk konvergensi yang lebih tinggi(Liu, 2021)
<i>Loss Function</i>	Menghitung hasil pada tingkat kesalahan dari hasil prediksi dengan data asli. Peneliti ini menggunakan MSE.
<i>Learning Rate</i>	Menentukan kecepatan dalam mempelajari data di <i>train model</i> . Rentang nilainya 0.0001 dalam nilai tersebut bisa menghasilkan hasil yang stabil dan tidak terlalu lambat(Peng, 2024)

2.5 Pengukuran Nilai Error

Prediksi juga tidak bisa memberikan nilai dengan seratus persen akurat namun bisa dimasukkan dalam mempertimbangkan dalam mengambil keputusan.

Karena hasil dari prediksi tidak selalu sama dengan nilai actual sehingga diperlukan pengukuran guna melihat seberapa besar nilai dari perbedaannya. Cara menganalisa perbedaan tersebut menggunakan 3 metrik yaitu MSE, RMSE, dan MAPE. Semakin kecil nilai dari perbedaannya semakin akurat hasil dari model, 3 metrik diatas memiliki persamaannya masing-masing dapat dilihat pada di bawah ini.

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_1 - \hat{y}_1)^2 \quad (8)$$

$$RMSE = \sqrt{\sum_{i=1}^n \frac{(y_1 - \hat{y}_1)^2}{n}} \quad (9)$$

$$MAPE = \frac{1}{n} \sum_{t=1}^n \frac{|y_1 - \hat{y}_1|}{y_1} \times 100 \quad (10)$$

Keterangan:

y_1 = Nilai data actual

\hat{y}_1 = Nilai data prediksi

n = Banyaknya data

3. Hasil dan Pembahasan

3.1 Preprocessing Data

Dalam proses ini *dataset* diolah dengan beberapa proses; pemodelan *dataset*, *split data*, normalisasi, dan sekuens data. Pada proses pemodelan *dataset* sehingga total fitur dalam dataset terdapat 4 data.

	date	active_spending_addresses	Close	PTR(%)	PLR(%)
0	18/08/2011	10,461	10,9	0	100
1	19/08/2011	10,825	11,69	16,84	83,16
2	20/08/2011	9,487	11,7	0	100
3	21/08/2011	10,400	11,7	100	0
...
5286	06/02/2026	409,686	67005	8,12	91,88

Gambar 4. Dataset Setelah Preprocessing

Dari Gambar 4 bisa didapatkan setelah melakukan pemodelan *dataset* yang menciptakan 2 fitur tambahan sebagai *input*. Sehingga terdiri dari 4 fitur dengan 1 fitur yang diprediksi yaitu harga "Close" dari *Bitcoin*, setelah melewati rangkaian tersebut akan dilanjutkan pada proses pelatihan Model.

3.2 Pelatihan Model

Pelatihan model dilakukan secara bertahap guna memastikan model BiLSTM memiliki kinerja yang optimal dalam memprediksi harga Bitcoin. Proses pelatihan diikuti dengan penetapan konfigurasi beberapa *hyperparameter*; *learning rate* 0.0001, Dalam BiLSTM menggunakan 2 lapis *hidden layer* dengan konfigurasi 16 *neuron* pada lapisan pertama dan 32 *neuron* pada lapisan kedua yang dimana nantinya pada *hidden layer* ini kesepuluh fitur tersebut akan diproses selanjutnya pada proses lapisan akhir *output layer* model dirancang hanya mengeluarkan satu nilai tunggal yaitu harga penutupan. Kesepuluh fitur lainnya tidak hilang melainkan telah dilebur menjadi bobot informasi di dalam model untuk menghasilkan satu nilai prediksi. Sehingga rangkuman *hyperparameter* bisa dilihat pada Tabel 7.

Tabel 7. *Hyperparameter*

Model	Fungsi Aktivasi	Optimizer	Loss
BiLSTM	ReLU	Adam	MSE

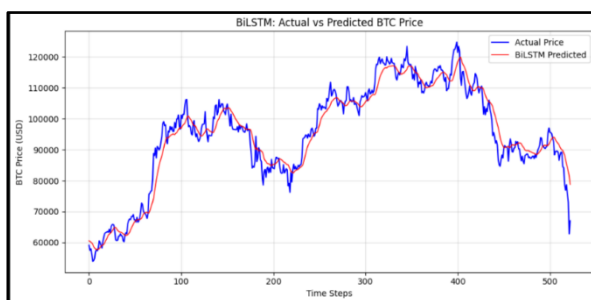
Berdasarkan penjelasan di atas dan pengaturan pada Tabel 7, arsitektur model nantinya dibangun secara sekuensial untuk memproses dataset dengan tipe *time series*.

Sesuai dengan konfigurasi, *layer* BiLSTM yang pertama menggunakan 16 *neuron* sehingga menghasilkan *output shape* sebesar 32 hasil dari penggabungan arah *forward* dan *backward*. Lapisan BiLSTM kedua menggunakan 32 *neuron* yang menghasilkan *output shape* sebesar 63. Untuk mencegah terjadinya *overfitting* selama proses *training*, *Dropout layer* ditambahkan tepat setelah masing-masing *layer* BiLSTM. Pada tahap akhir *Dense layer* dengan 1 *neuron* digunakan untuk menyatukan hasil dari ekstraksi fitur menjadi satu nilai prediksi akhir yaitu harga “Close” Bitcoin.

3.3 Evaluasi Model

Hasil evaluasi dari model BiLSTM yang menggunakan *split data* 90:10, konfigurasi *hyperparameter* yang terdapat pada Tabel 7 di atas menghasilkan MSE sebesar 15,25, RMSE sebesar 3,905 dan MAPE sebesar 3,22%.

Dari hasil tersebut, dapat dilakukan pemetaan *visual* atau *plotting* untuk membandingkan harga actual Bitcoin dengan harga hasil dari prediksi. Visualisasi ini bertujuan melihat secara langsung seberapa akurat model dalam menangkap pola dan mengikuti tren pergerakan harga pada data pengujian (*testing data*). Grafik perbandingan hasil *plotting* prediksi dari model BiLSTM dapat dilihat pada Gambar 4. di bawah ini.



Gambar 4. Grafik Prediksi dari Model BiLSTM

Uji kemampuan dan keunggulan dari model BiLSTM yang telah dibangun akan dilakukan perbandingan kinerja dengan beberapa algoritma dari Deep Learning dari penelitian terdahulu yang memiliki fokus serupa pada pemodelan fluktuasi dari harga mata uang kripto. Berdasarkan hasil dari komparasi metrik tingkat kesalahan seperti; MAE, RMSE, dan MAPE terbukti bahwa model BiLSTM yang diusulkan di penelitian ini merupakan arsitektur yang paling akurat di antara kedua model pembanding tersebut. Kemampuan dari model BiLSTM dalam mengolah data dengan memanfaatkan dua arah sekaligus membantu untuk

menekan angka error secara signifikan. Rangkuman hasil perbandingan evaluasi kinerja dari ketiga model tersebut dapat dilihat lebih jelas pada Tabel 8 di bawah ini.

Tabel 8. Hasil Komparasi Metrik

Metode	MAE	RMSE	MAPE(%)	Author
BiLSTM	2,958	3,905	3,22%	Penulis
LSTM	19,873	24,321	29,06%	[19].
MLP	0,0418	0,06921	4,01%	[20]

Berdasarkan Tabel 8 model BiLSTM yang diusulkan pada penelitian ini terbukti paling rendah nilai dari MAPE yakni 1.41%. Sebagai perbandingan model MLP pada penelitian Jagannath et al. menghasilkan tingkat kesalahan MAPE 4.01% meskipun nilai MAE dan RMSE pada penelitian MLP tersebut lebih kecil [19]. Sementara itu penelitian yang dilakukan Jay et al. Menggunakan model LSTM mencatatkan nilai MAPE yang cukup tinggi yaitu sebesar 29.06% [20]. Hasil ini membuktikan bahwa arsitektur pemrosesan dua arah pada BiLSTM lebih unggul dalam menangani volatilitas ekstrem yang terjadi di pasar kripto.

4. Kesimpulan

Penelitian ini berhasil merancang dan mengimplementasikan arsitektur dari Deep Learning BiLSTM guna memprediksi arah dari pergerakan harga Bitcoin secara akurat dengan memanfaatkan masukan dari data *on-chain*. Pemodelan BiLSTM terbukti memiliki kemampuan komputasi yang sangat baik dalam menangani karakteristik dari data harga Bitcoin yang sangat fluktuatif dan *non-linear*. Berkat kemampuannya memproses urutan informasi dari dua arah (kiri ke kanan dan dari kanan ke kiri). Berdasarkan tahapan evaluasi menggunakan skenario rasio data *training* dan *testing* sebesar 90:10 model ini menunjukkan performa prediksi yang presisi dengan nilai MAE sebesar 2,958, RMSE sebesar 3,905 dan MAPE yang sangat minim yaitu sebesar 3.22%. Jika dibandingkan dengan performa model prediktif dari dua penelitian sebelumnya seperti menggunakan LSTM dan MLP, arsitektur BiLSTM secara meyakinkan terbukti menjadi metode yang konsisten dan akurat.

Saran untuk tahapan riset lanjutan atau *future works*. Mengembangkan arsitektur model ini dengan menggabungkan beberapa variabel data eksternal seperti analisis sentimen publik dan indeks berita makroekonomi untuk fitur *input* dan menyesuaikan prediksi terhadap anomaly dari lonjakan sentiment pasar yang bisa terjadi secara tiba-tiba

Daftar Rujukan

- [1] P. Zhu, X. Zhang, Y. Wu, H. Zheng, and Y. Zhang, “Investor attention and cryptocurrency: Evidence from the Bitcoin market,” *PLoS One*, vol. 16, no. 2, p. e0246331, 2021.
- [2] S. Nakamoto, “Bitcoin: A peer-to-peer electronic cash system,” Available at SSRN 3440802, 2008.
- [3] D. O. Milando, R. Rahim, and F. Adrianto, “Analisis Pengaruh World Commodity Price terhadap Harga Bitcoin dengan Indeks Dolar sebagai Variabel Moderasi,” *Jurnal Informatika Ekonomi Bisnis*, pp. 1107–1114, 2023.

- [4] P. H. Padhila, I. Cholissodin, and P. P. Adikara, "Prediksi Harga Bitcoin berdasarkan Data Historis Harian dan Google Trend Index menggunakan Algoritme Extreme Learning Machine," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 6, no. 7, pp. 3515–3524, 2022.
- [5] S. Singh, A. Pise, and B. Yoon, "Prediction of bitcoin stock price using feature subset optimization," *Heliyon*, vol. 10, no. 7, p. e28415, 2024, doi: <https://doi.org/10.1016/j.heliyon.2024.e28415>.
- [6] S. M. Raju and A. M. Tarif, "Real-time prediction of BITCOIN price using machine learning techniques and public sentiment analysis," *arXiv preprint arXiv:2006.14473*, 2020.
- [7] R. Dubey and D. Enke, "Bitcoin price direction prediction using on-chain data and feature selection," *Machine Learning with Applications*, vol. 20, p. 100674, 2025, doi: <https://doi.org/10.1016/j.mlwa.2025.100674>.
- [8] A. Brauneis, R. Mestel, R. Riordan, and E. Theissen, "Bitcoin unchained: Determinants of cryptocurrency exchange liquidity," *J. Empir. Finance*, vol. 69, pp. 106–122, 2022, doi: <https://doi.org/10.1016/j.jempfin.2022.08.004>.
- [9] K. Alabi, "Digital blockchain networks appear to be following Metcalfe's Law," *Electron. Commer. Res. Appl.*, vol. 24, pp. 23–29, 2017.
- [10] H. Utama, "Pendekatan Deep Learning Menggunakan Metode Lstm Untuk Prediksi Harga Bitcoin," *The Indonesian Journal of Computer Science Research*, vol. 2, no. 2, pp. 43–50, 2023.
- [11] F. Febriansyah, A. Sujjada, and F. Sembiring, "Prediksi Harga Bitcoin Menggunakan Algoritma Long Short Term Memory (LSTM)," *INOVTEK Polbeng-Seri Informatika*, vol. 9, no. 1, 2024.
- [12] A. F. Hanif, T. B. Sasongko, and A. D. Laksito, "Perbandingan Kinerja LSTM, Bi-LSTM, dan GRU pada Klasifikasi Judul Berita Clickbait," *The Indonesian Journal of Computer Science*, vol. 12, no. 4, 2023.
- [13] D. I. Puteri, "Implementasi Long Short Term Memory (LSTM) dan Bidirectional Long Short Term Memory (BiLSTM) Dalam Prediksi Harga Saham Syariah," *Euler: Jurnal Ilmiah Matematika, Sains dan Teknologi*, vol. 11, no. 1, pp. 35–43, 2023.
- [14] M. Yang and J. Wang, "Adaptability of Financial Time Series Prediction Based on BiLSTM," *Procedia Comput. Sci.*, vol. 199, pp. 18–25, 2022, doi: <https://doi.org/10.1016/j.procs.2022.01.003>.
- [15] R. Parlita, R. R. Isnanto, and B. Rahmat, "Prediction of ROI achievements and potential maximum profit on spot Bitcoin Rupiah trading using K-means clustering and patterned dataset model," *JOIV: International Journal on Informatics Visualization*, vol. 8, no. 3–2, pp. 1987–2001, 2024.
- [16] I. B. N. A. P. Wiryawan, "Perbandingan Time Step pada Model Prediksi State of Health Baterai Lithium-ion Berbasis BiLSTM," in *Seminar Nasional Riset Inovatif*, 2024.
- [17] E. N. Waroi, A. Arief, and K. Khusnawi, "Prediksi harga laptop menggunakan algoritma GRU dan BiLSTM," *Jurnal Sosial Teknologi*, vol. 4, no. 7, pp. 408–424, 2024.
- [18] Z. Hameed and B. Garcia-Zapirain, "Sentiment classification using a single-layered BiLSTM model," *Ieee Access*, vol. 8, pp. 73992–74001, 2020.
- [19] N. Jagannath *et al.*, "An on-chain analysis-based approach to predict ethereum prices," *IEEE Access*, vol. 9, pp. 167972–167989, 2021.
- [20] P. Jay, V. Kalariya, P. Parmar, S. Tanwar, N. Kumar, and M. Alazab, "Stochastic neural networks for cryptocurrency price prediction," *Ieee access*, vol. 8, pp. 82804–82818, 2020.