

**Analisis dan Perbandingan Algoritma SVR, XGBOOST, dan Lightgbm
dalam Prediksi *Cryptocurrency* Ethereum**

**Ryan Anthony¹, Jechenthia Maria Taso², Stephen Yohanes Christopher³,
Ridhwan Ardiyansyah⁴**

¹²³⁴Universitas Tarumanagara, Indonesia

ryan.535220074@stu.untar.ac.id¹, jechenthia.535220081@stu.untar.ac.id², stephen.
535220105@stu.untar.ac.id³, ridhwan.535220074@stu.untar.ac.id⁴

ABSTRACT

This study aims to analyze and compare the performance of three algorithms, namely Support Vector Regression (SVR) with a linear kernel, XGBoost, and LightGBM, in predicting the Price of Ethereum cryptocurrency based on daily historical data. The study uses Ethereum Price data in USD for the last five years obtained from the investing.com website. The variables used are Close, Open, High, and Low Prices. The study uses two data splitting scenarios: 80% training data and 20% testing data, and 70% training data and 30% testing data. This study also uses time step variations to test the effect of time dependency on algorithm performance. The results indicate that the LightGBM algorithm has the best performance compared to the other two algorithms with an average MAE value for High Price of 75.486, SVR has a value of 115.590, and XGBoost has a value of 77.314 in the 80% training data and 20% testing data split. In the 70% training data and 30% testing data split, the LightGBM algorithm still excels with an average MAE value for High Price of 78.228, SVR of 104.356, and XGBoost of 83.573. Other evaluations such as RMSE and R2 also show the superiority of the LightGBM algorithm. For the required computation time, the SVR algorithm outperforms the other two algorithms.

Keywords: Ethereum, LightGBM, SVR, XGBoost

ABSTRAK

Penelitian bertujuan untuk menganalisis dan membandingkan kinerja tiga algoritma yaitu SVR dengan kernel linear, XGBoost, dan LightGBM dalam memprediksi harga *cryptocurrency* Ethereum berdasarkan data historis harian. Penelitian menggunakan *dataset* harga Ethereum dalam satuan USD selama lima tahun terakhir yang diperoleh dari situs investing.com. Variabel yang digunakan adalah harga *Close*, *Open*, *High*, dan *Low*. Penelitian menggunakan dua skenario pembagian data yaitu 80% data latih dan 20% data uji serta 70% data latih dan 30% data uji. Penelitian ini juga menggunakan variasi *time step* untuk menguji pengaruh ketergantungan waktu terhadap kerja algoritma. Hasil penelitian menunjukkan bahwa algoritma LightGBM memiliki kinerja terbaik dibandingkan dua algoritma lainnya dengan nilai rata-rata MAE untuk *High Price* yaitu 75.486, SVR memiliki nilai 115.590, dan XGBoost memiliki nilai 77.314 pada pembagian 80% data latih dan 20% data uji. Pada pembagian 70% data latih dan 30% data uji algoritma LightGBM masih unggul dengan nilai rata-rata MAE untuk *High Price* yaitu 78.228, SVR sebesar 104.356, dan XGBoost sebesar 83.573. Evaluasi lainnya seperti RMSE dan R2 juga menunjukkan keunggulan algoritma LightGBM. Untuk waktu komputasi yang diperlukan algoritma SVR unggul dua algoritma lainnya.

Kata kunci: Ethereum, LightGBM, SVR, XGBoost

PENDAHULUAN

Perkembangan teknologi digital terjadi sangat cepat dan masif di seluruh dunia. Perkembangan ini mendorong berkembangnya inovasi teknologi di berbagai sektor kehidupan. Salah satu inovasi yang berkembang pada sektor keuangan yaitu *cryptocurrency* (Shin & Rice, 2022). *Cryptocurrency* merupakan inovasi pada sistem perekonomian berbasis digital yang terdesentralisasi (Marzo et al., 2022). *Cryptocurrency* dibangun di atas teknologi *blockchain* sehingga mendukung penggunaanya untuk melakukan pembayaran secara mudah dan cepat tanpa melibatkan lembaga keuangan (Robusti et al., 2025).

Selama beberapa dekade terakhir *cryptocurrency* ramai diperbincangkan karena keefisienan dan keamanannya sebagai mata uang virtual (Cunha et al., 2021). *Cryptocurrency* dapat digunakan untuk berbagai tujuan seperti melakukan pembelian barang atau jasa, melakukan penambangan (*mining*), serta dalam investasi (Astuti et al., 2022). Salah satu *cryptocurrency* yang cukup populer adalah Ethereum. Ethereum merupakan salah satu dari *cryptocurrency* yang berkembang dan memiliki komunitas yang kuat selain Bitcoin (De Vries, 2023). Berbeda dengan Bitcoin yang difokuskan sebagai media pertukaran dan penyimpanan nilai, Ethereum mendukung *smart contract* dan platform terdesentralisasi dalam pengembangan aplikasi berbasis *blockchain* (Robusti et al., 2025).

Dalam bursa perdagangan *cryptocurrency*, Ethereum menempati posisi kedua setelah Bitcoin sebagai *cryptocurrency* dengan kapitalisasi pasar tertinggi. Tercatat pada tahun 2025, dominasi Ethereum menunjukkan kenaikan signifikan hingga melonjak ke angka 14.65%. Angka ini merupakan kenaikan kapitalisasi pasar tertinggi yang terjadi pada tahun 2025 (UMAIR, 2025). Hal ini semakin memperkuat posisi Ethereum dalam pasar *cryptocurrency* global yang membawa fungsi lebih dari sekadar alat tukar tetapi sebagai pelopor lahirnya ekosistem *smart contract*, aplikasi terdesentralisasi serta DeFi (*Decentralized Finance*) (Sriman & Kumar, 2022).

Beberapa penelitian terdahulu telah dilakukan dalam melakukan prediksi terhadap harga *cryptocurrency*. Penelitian menggunakan metode *Long Short-Term Memory* (LSTM) dengan data historis Ethereum periode Januari 2018 hingga Januari 2022 menghasilkan nilai rata-rata MSE sebesar 0.0121 (Sianturi et al., 2023). Adapun penelitian yang mencoba membandingkan beberapa algoritma seperti *Linear Regression*, *Random Forest*, *Support Vector Machine* (SVM), dan *Long Short-Term Memory* (LSTM) menggunakan data dari 3 jenis *cryptocurrency* yang berbeda. Diperoleh hasil terbaik pada algoritma LSTM untuk Litecoin (LTC) dengan nilai MAE 6.32 dan nilai MAE 42.15 pada Ethereum (Khoiri et al., 2024). Selanjutnya penelitian menggunakan metode *Vector Autoregressive* pada data Ethereum memperoleh hasil evaluasi RMSE sebesar 978.37 pada harga penutupan (Pradana Ananda Raharja, 2021). Kemudian studi menggunakan algoritma *Gated Recurrent Unit* (GRU) dan XGBoost untuk *cryptocurrency* Ethereum memperoleh hasil GRU lebih unggul dengan nilai RMSE 101.37 dan R2 sebesar 0.9718 (Reza & Ghufon, 2025). Di sisi lain penelitian menggunakan metode *Support Vector Regression* (SVR) dan *Extreme Learning Machine* (ELM) dalam prediksi harga Bitcoin memperoleh hasil bahwa algoritma ELM lebih baik dengan R2 sebesar 0.987 pada *split* 80/20 lalu R2 sebesar 0.989 pada *split* 60/40 (Ferdinand et al., 2025).

Berbagai penelitian menunjukkan bahwa tingkat akurasi dalam prediksi harga *cryptocurrency* sangat beragam dan dipengaruhi oleh metode yang digunakan. Oleh karena itu penelitian ini berfokus pada penggunaan algoritma *Support Vector Regression* (SVR), XGBoost, dan LightGBM. Algoritma SVR digunakan karena algoritma ini mampu memodelkan hubungan non-linear antar variabelnya menggunakan fungsi kernel. SVR juga tidak memerlukan banyak parameter dibanding dengan algoritma seperti *deep learning* dan bisa memberikan hasil yang stabil dan mudah dipahami. XGBoost dipilih karena algoritma ini dapat mencegah *overfitting* sehingga dapat mempelajari pola harga yang mudah berubah pada pasar *cryptocurrency*. Algoritma LightGBM dipilih karena algoritma ini cepat dan lebih efisien dalam proses pelatihan model khususnya pada data non-linear. Penelitian ini diharapkan mampu memberikan perbandingan dan analisis mendalam untuk melakukan perbandingan performa komputasi dan akurasi yang diperoleh dari setiap algoritma yang digunakan.

Pergerakan harga pasar *cryptocurrency* khususnya Ethereum sangat *Volatil* sehingga memiliki risiko yang lebih tinggi dibanding investasi pada saham (Zhang & Mani, 2021). Meski demikian peluang keuntungan yang didapatkan dari *cryptocurrency* berkali-kali lipat jika dibandingkan dengan instrumen investasi lainnya. Fluktuasi harga yang tinggi tersebut disebabkan oleh faktor geopolitik dan sosioekonomi yang terjadi pada seluruh dunia (Theiri et al., 2023). Dengan mempertimbangkan sifat harga Ethereum yang berisiko tinggi dan sangat *Volatil*, maka diperlukan eksperimen dan analisis menggunakan algoritma *machine learning* dan membandingkan algoritma mana yang paling optimal.

TINJAUAN LITERATUR

Support Vector Regression (SVR) adalah algoritma *supervised learning* yang digunakan dalam memprediksi nilai variabel kontinu (Gananta et al., 2023). Algoritma ini menggunakan prinsip yang sama dengan *Support Vector Machine* (SVM). Tujuan dari algoritma ini adalah untuk mencari fungsi $f(x)$ yang memiliki nilai *error* kurang dari nilai ambang ϵ untuk seluruh data pelatihan. SVR dengan kernel linear dirumuskan sebagai berikut (Ferdinand et al., 2025):

$$f(x) = \omega^T x + b \quad (1)$$

$$\min \left(\frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \right) \quad (2)$$

Keterangan:

ω = vektor bobot kemiringan fungsi regresi

x = vektor fitur input prediksi

b = intersep fungsi regresi

ξ_i, ξ_i^* = variabel *slack* pelanggaran margin di atas/bawah batas ϵ

C = parameter regulasi yang mengatur keseimbangan kompleksitas model dan penalti *error*

Selain SVR, penelitian ini juga mengimplementasikan algoritma berbasis pohon keputusan. *Extreme Gradient boosting* (XGBoost) merupakan algoritma *ensemble learning* berbasis *gradient boosting decision tree*. Konsep utama algoritma ini adalah setiap *tree* mencoba memperbaiki kesalahan prediksi dari *tree* sebelumnya (Asnawi et al., 2024). Hasil akhirnya algoritma ini bekerja dengan menggabungkan beberapa model keputusan sederhana untuk membentuk model yang lebih kuat. Fungsi prediksi XGBoost dirumuskan sebagai berikut (Astutiningsih et al., 2023):

$$\hat{y}_i = \sum_{k=1}^K f_k(x_i) \quad (3)$$

Keterangan:

$\hat{y}_i$ = hasil prediksi untuk data ke-i

$f_k(x_i)$ = pohon keputusan ke-k

K = jumlah total pohon yang dibangun

x_i = vektor fitur input

Sebagai pembanding untuk menguji peningkatan efisiensi komputasi pada keluarga algoritma yang sama, penelitian ini turut menganalisis LightGBM. Algoritma *Light Gradient boosting Machine* (LightGBM) adalah model berbasis *gradient boosting* yang dikembangkan oleh Microsoft pada tahun 2017 (Septiana Rizky et al., 2022). Algoritma ini merupakan pengembangan dari algoritma XGBoost dengan optimisasi pada kecepatan komputasi dan penggunaan memori yang lebih efisien. Kecepatan dan keefisienan tersebut diperoleh karena algoritma ini menggunakan pendekatan *histogram-based algorithm*. Berbeda dengan XGBoost yang membuat *tree* secara *level-wise*, algoritma LightGBM membuat *tree* secara *leaf-wise* (berdasarkan *loss* terbesar) (Syam et al., 2025). Algoritma LightGBM untuk prediksi dapat dirumuskan sebagai berikut (Immanuel & Handhayani, 2025):

$$\hat{y}_i = \sum_{t=1}^T f_t(x_i) \quad (4)$$

Keterangan:

$\hat{y}_i$ = hasil prediksi untuk data ke-i

$f_t(x_i)$ = pohon keputusan ke-t

T = jumlah total pohon yang dibangun

x_i = vektor fitur input

Untuk mengukur sejauh mana kinerja dan tingkat akurasi dari ketiga algoritma tersebut, diperlukan pengujian yang terstruktur. Dalam menggunakan algoritma prediksi perlu adanya metrik evaluasi yang digunakan untuk menilai performa dari tiap algoritma yang digunakan. Dalam penelitian ini metrik evaluasi yang digunakan untuk menilai algoritma yang digunakan yaitu Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), dan Koefisien Determinasi (R2). MAE

menunjukkan rata-rata absolut selisih nilai hasil prediksi dengan nilai aktual. RMSE mengukur akar rata-rata kuadrat selisih nilai prediksi dan nilai aktual. Koefisien Determinasi (R^2) menunjukkan seberapa besar proporsi variansi data aktual dapat dijelaskan oleh model prediksi. Berikut ini persamaan yang digunakan untuk menghitung nilai MAE, RMSE, dan R^2 secara berurutan (Tatachar, 2021):

$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_i - \hat{Y}_i| \quad (5)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2} \quad (6)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (7)$$

Keterangan:

n = jumlah total data uji

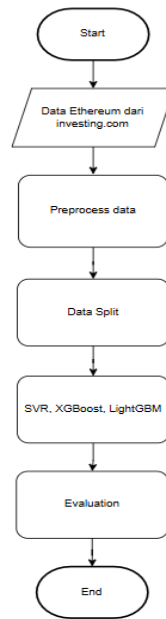
y_i = nilai aktual untuk data ke- i

$\hat{y}_i$ = nilai prediksi untuk data ke- i

$\bar{y}$ = rata-rata dari nilai aktual

METODE PENELITIAN

Pada penelitian ini metode yang digunakan adalah perbandingan antara *Support Vector Regression* (SVR) dengan kernel linear, XGBoost, serta LightGBM yang akan digunakan untuk melakukan prediksi dari harga Ethereum. Data yang digunakan dalam penelitian ini berisi data harga Ethereum dalam Dollar Amerika Serikat (USD) yang diperoleh dari situs investing.com. Berikut ini metodologi penelitian yang digunakan yang bisa dilihat pada Gambar 1. Langkah pertama dalam tahapan metodologi ini adalah pengumpulan data mentah yang akan diobservasi. *Dataset* yang digunakan dalam penelitian ini diperoleh dari situs investing.com, yang menyediakan data historis harga Ethereum selama 5 tahun. Harga Ethereum dinyatakan dalam dollar Amerika Serikat (USD). Data yang digunakan adalah data *time-series* harian dengan variabel seperti *Price (Close)*, *Open*, *High*, *Low*, *Vol*, dan *Change %*. Secara berurutan setiap variabel menjelaskan harga penutupan, harga pembukaan, harga tertinggi dalam satu hari, harga terendah dalam satu hari, *Volume* transaksi perdagangan, dan persentase perubahan harga harian.

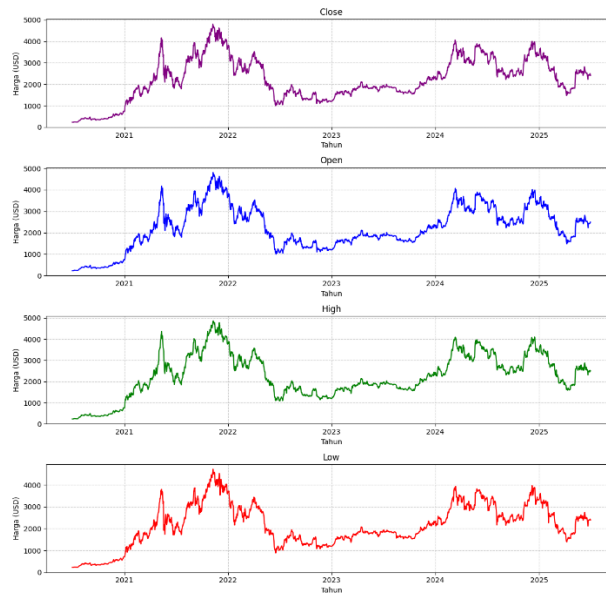


Gambar 1 FLOWchart Penelitian

Setelah data mentah berhasil dikumpulkan, tahap selanjutnya adalah melakukan prapemrosesan data (*preprocessing*) untuk membuang variabel yang tidak relevan. *Dataset* yang digunakan diunduh dari situs investing.com dalam format .csv, berisi 1.827 baris data historis harian harga Ethereum. *Dataset* tidak memiliki *missing value*. Periode data yang digunakan adalah data historis harga Ethereum selama 5 tahun (1 Juli 2020 hingga 1 Juli 2025). Variabel *Vol* dan *Change %* dihapus sedangkan variabel lain seperti *Price (Close)*, *Open*, *High*, dan *Low* digunakan untuk melakukan prediksi harga Ethereum. Berikut ini sampel dari *dataset* yang digunakan yang dapat dilihat pada Tabel 1. Untuk persebaran data yang digunakan dapat dilihat pada Gambar 2.

Date	Close	Open	High	Low
2020-07-01	230.95	225.61	232.47	224.33
2020-07-02	226.42	230.96	232.04	223.53
2020-07-03	225.12	226.42	228.14	224.77
2020-07-04	229.15	225.09	230.40	224.80
2020-07-05	227.98	229.09	229.80	223.79

Tabel 1 Sampel 5 data pertama dari *dataset*



Gambar 2 Grafik dari data *Close*, *Open*, *High*, dan *Low*

Setelah data dibersihkan dan variabel target ditetapkan, data tersebut kemudian dibagi ke dalam proporsi tertentu untuk keperluan pelatihan dan pengujian model. Pada penelitian ini digunakan 2 skenario dalam pembagian data. Skenario pertama menggunakan pembagian 70% data latih dan 30% data uji. Skenario kedua menggunakan pembagian 80% data latih dan 20% data uji. Pembagian data tersebut dimaksudkan untuk melakukan analisis terhadap dua skenario pembagian yang telah ditentukan untuk melihat bagaimana performa model bekerja. Pada setiap skenario pembagian data akan menggunakan tujuh variasi *time step* (3,5,7,10,14,21,28). Variasi tersebut digunakan untuk mengatur seberapa banyak hari yang digunakan untuk melakukan prediksi untuk hari berikutnya. Pengujian dengan berbagai nilai *time step* dilakukan untuk menganalisis sejauh mana ketergantungan terhadap data historis dapat memengaruhi performa model dalam melakukan prediksi.

HASIL DAN PEMBAHASAN

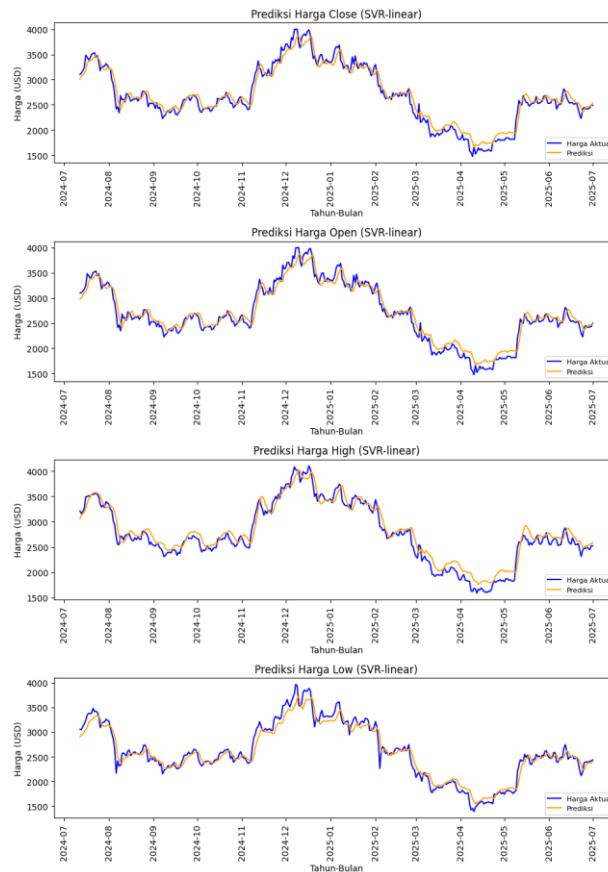
Untuk mengevaluasi kinerja masing-masing model, eksperimen pertama dilakukan menggunakan algoritma *Support Vector Regression* (SVR). Penelitian dilakukan menggunakan algoritma SVR dengan kernel linear, dengan skenario *split* data 80/20 untuk 7 eksperimen. Algoritma SVR menunjukkan hasil evaluasi yang menjanjikan. Nilai Mean Absolute Error (MAE) pada *Close Price* sebesar 105.383 USD menunjukkan bahwa rata-rata perbedaan nilai prediksi harga Ethereum dan nilai aktualnya adalah sebesar 105.383 USD. Untuk Root Mean Squared Error (RMSE) pada seluruh eksperimen yang dilakukan untuk *Close Price* adalah 133.108 USD, nilai ini menunjukkan evaluasi yang bagus dari sisi akurasi. Kemudian untuk rata-rata koefisien determinasinya sebesar 0.947, hal ini menandakan bahwa model yang terbentuk mampu melakukan prediksi yang baik sehingga mendekati nilai aktualnya. Rata-rata waktu komputasi yang diperlukan dalam algoritma SVR dalam membentuk

model prediksi juga terbilang sangat cepat yaitu 0.005 detik. Rata-rata hasil evaluasi dengan pembagian 80% data latih dan 20% data uji dapat terlihat pada Tabel 2.

Tabel 2 Rerata Evaluasi Algoritma SVR
(80% data latih dan 20% data uji)

<i>Price</i>	MAE	RMSE	R2	Runtime
<i>Close</i>	105.383	133.108	0.947	0.005
<i>Open</i>	105.832	133.458	0.947	0.004
<i>High</i>	115.590	140.194	0.944	0.003
<i>Low</i>	99.501	134.126	0.944	0.005

Selain dari *Close Price*, hasil evaluasi untuk harga lainnya terlihat optimal meskipun pada *High Price* model yang terbentuk memiliki nilai MAE terbesar yaitu di angka 115.590 USD dan nilai RMSE terbesar yaitu 140.194 USD. Nilai koefisien determinasi (R2) untuk *High Price* juga salah satu yang terkecil jika dibandingkan dengan *Close Price*. Hal ini menandakan bahwa model paling sulit memprediksi *High Price* harian karena fluktuasi dari *High Price* yang *Volatil* dan sulit diprediksi dengan pola linear. Tingginya nilai *error* yang dihasilkan pada *High Price* juga bisa disebabkan oleh penggunaan *time step* = 10, yang dimana nilai MAE, RMSE, dan R2 berturut-turut adalah 120.144, 145.323, dan 0.94. Meskipun demikian rata-rata waktu komputasi yang diperlukan pada *High Price* adalah yang paling cepat yaitu 0.003 detik. Gambar berikut akan memperlihatkan grafik prediksi dari setiap harga dengan *time step* = 10.



Gambar 3 Grafik harga *Close*, *Open*, *High*, dan *Low* (80% data latih, 20% data uji, dan *time step* =10)

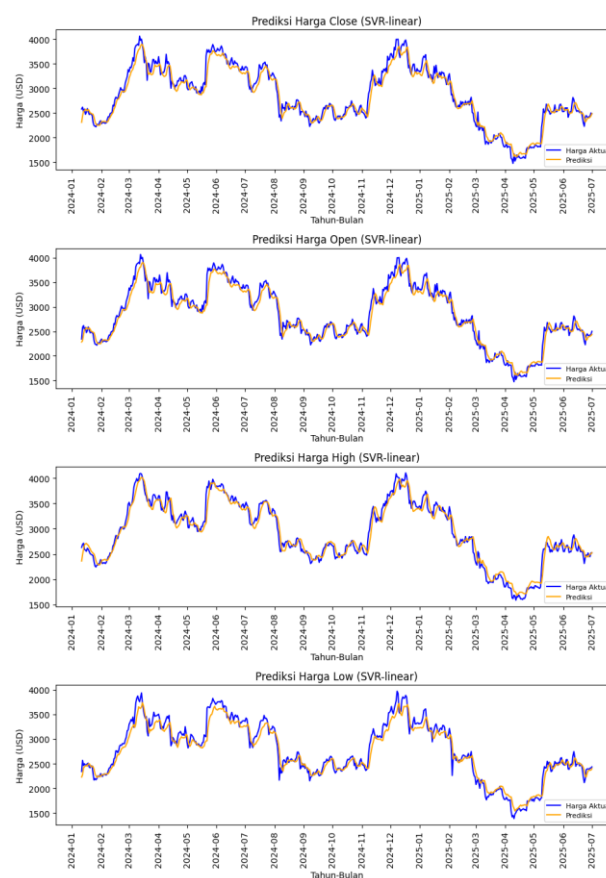
Selain pengujian dengan rasio tersebut, performa SVR juga diuji menggunakan proporsi data latih yang lebih sedikit. Dengan algoritma yang sama yaitu SVR dengan data latih 70% dan data uji 30% diperoleh hasil evaluasi model menunjukkan peningkatan performa pada *Close*, *Open*, dan *High Price* sedangkan pada *Low Price* mengalami penurunan. Sebagai contoh pada *Close Price*, model menghasilkan nilai MAE sebesar 98.697 USD menandakan rata-rata perbedaan nilai prediksi dan nilai aktualnya sebesar 98.697 USD. Nilai tersebut menunjukkan peningkatan dibandingkan dengan pengujian pada pembagian data 80/20. Untuk nilai koefisien determinasi (R^2) pada *Close Price* juga mengalami peningkatan sebesar 0.005 dibanding sebelumnya. Untuk waktu komputasi yang diperlukan secara keseluruhan tidak ada yang berubah secara signifikan. Berikut tabel rata-rata nilai evaluasi setiap harga:

Tabel 3 Rerata Evaluasi Algoritma SVR (70% data latih dan 30% data uji)

Price	MAE	RMSE	R2	Runtime
Close	98.697	130.621	0.952	0.005
Open	98.461	130.371	0.953	0.005
High	104.356	118.1123	0.953	0.002

<i>Low</i>	107.238	141.4513	0.942	0.003
------------	---------	----------	-------	-------

Dapat terlihat pada Tabel 3 bahwa *High Price* mengalami peningkatan yang paling besar dibandingkan harga lainnya, sedangkan pada *Low Price* evaluasi model mengalami penurunan. Hal ini bisa terjadi karena pada pembagian data 70/30 model memiliki variasi data yang lebih baik khususnya untuk *High Price* sehingga prediksinya lebih baik jika dibandingkan dengan *Low Price* yang kemungkinan lebih sensitif terhadap *Volatilitas* harga. Sementara untuk waktu komputasi yang diperlukan untuk masing-masing harga masih sangat cepat karena algoritma SVR memiliki komputasi yang ringan. Berikut ini gambar grafik perbandingan setiap harga pada pembagian data latih 70% dan data uji 30% serta *time step* = 10.



Gambar 4 Grafik harga *Close*, *Open*, *High*, dan *Low* (70% data latih, 30% data uji, dan *time step* = 10)

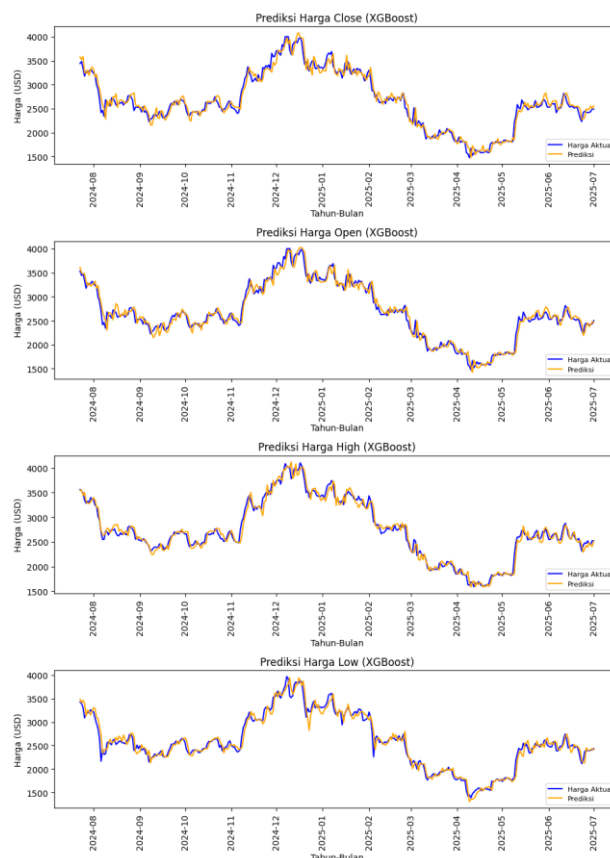
Setelah mengevaluasi SVR, eksperimen dilanjutkan dengan menggunakan algoritma *ensemble*, yaitu XGBoost. Penelitian dengan algoritma XGBoost dilakukan tanpa melakukan *tuning parameter*. Hasil evaluasi yang dihasilkan dengan algoritma XGBoost menunjukkan performa yang sangat baik. Rata-rata nilai MAE yang dihasilkan oleh algoritma ini berada di bawah angka 90 USD. Untuk nilai rata-rata RMSE yang diperoleh juga sudah cukup baik dengan kisaran diantara 102-125 USD. Untuk nilai koefisien determinasinya (R^2) juga sangat memuaskan berkisar antara 0.95 – 0.97 dengan nilai tertinggi dihasilkan oleh *High Price* sebesar 0.97, artinya

model dengan algoritma XGBoost mampu menjelaskan lebih dari 95% variasi harga Ethereum yang fluktuatif.

Tabel 4 Rerata Evaluasi Algoritma XGBoost
(80% data latih dan 20% data uji)

<i>Price</i>	MAE	RMSE	R2	Runtime
<i>Close</i>	87.252	117.7463	0.959	0.119
<i>Open</i>	88.686	119.5509	0.958	0.103
<i>High</i>	77.314	102.6709	0.970	0.112
<i>Low</i>	88.381	124.9911	0.951	0.108

Tingginya performa model karena algoritma XGBoost lebih unggul dalam menangkap pola non-linear pada data harga historis. Waktu komputasi yang diperlukan pada algoritma XGBoost juga tergolong sangat cepat yaitu di kisaran rata-rata 0.103 – 0.119 detik, meskipun kecepatan tersebut masih lebih lambat dibanding algoritma SVR. Nilai terbaik pada evaluasi algoritma XGBoost terjadi pada *time step* = 21 pada *High Price* dengan nilai MAE, RMSE, dan R2 secara berturut – turut adalah 74.638, 99.207, dan 0.972. Hal ini menandakan bahwa model XGBoost mampu menangkap pola historis besar dan kompleks dengan stabil. Grafik perbandingan setiap harga dengan *time step* = 21 akan terlihat pada Gambar 5.

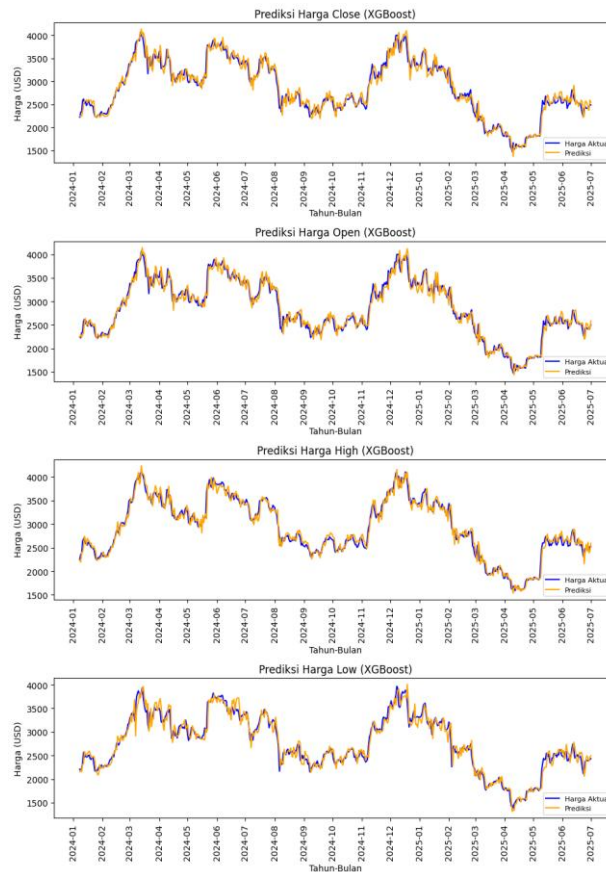


Gambar 5 Grafik harga *Close*, *Open*, *High*, dan *Low*
(80% data latih, 20% data uji, dan *time step* =21)

Sama halnya dengan SVR, performa XGBoost juga dievaluasi pada skenario pembagian data yang lebih kecil. Pada skenario pembagian data 70% data latih dan 30% data uji diperoleh hasil model mengalami penurunan. Penurunan performa terjadi secara jelas terlihat untuk nilai *error* yaitu MAE dan RMSE pada *Close*, *Open*, dan *Low Price*. Sementara untuk *High Price* penurunan yang terjadi tidak terlalu signifikan. Secara garis besar pembagian data 70/30 menghasilkan performa yang sedikit lebih rendah dibanding 80/20, namun model XGBoost tetap mampu mengenali pola data dengan baik dan tidak bergantung pada banyaknya jumlah data latih yang ditentukan dengan rata-rata koefisien determinasi (R^2) yang masih berada pada kisaran 0.941 – 0.967. Selain itu waktu komputasi yang diperlukan juga relatif stabil pada dua skenario pembagian data yaitu rata-rata 0.1 detik, yang berarti model XGBoost tetap efisien secara komputasi. Contoh grafik perbandingan setiap harga prediksi dan aktualnya dapat dilihat pada Gambar 6.

Tabel 5 Rerata Evaluasi Algoritma XGBoost
(70% data latih dan 30% data uji)

<i>Price</i>	MAE	RMSE	R^2	Runtime
<i>Close</i>	98.086	130.8543	0.952	0.103
<i>Open</i>	97.711	130.6773	0.953	0.098
<i>High</i>	83.573	110.2787	0.967	0.097
<i>Low</i>	101.373	141.9163	0.941	0.100



Gambar 6 Grafik harga *Close*, *Open*, *High*, dan *Low* (70% data latih, 30% data uji, dan *time step* =7)

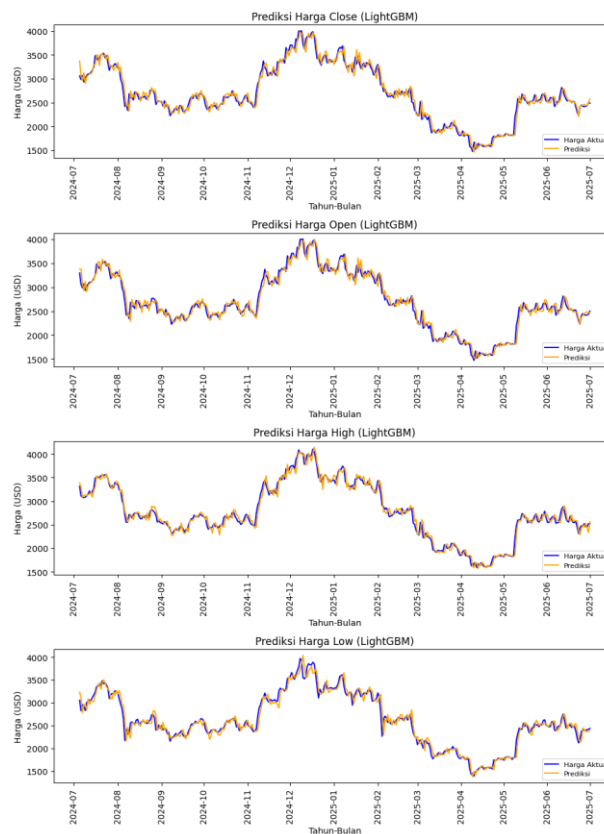
Untuk menguji efisiensi algoritma boosting yang lebih modern, penelitian ini kemudian mengevaluasi algoritma LightGBM. Penelitian dengan algoritma LightGBM dilakukan tanpa melakukan *tuning parameter* tertentu untuk menjaga objektivitas perbandingan dengan model lainnya. Pada skenario pembagian data 80% data latih dan 20% data uji, algoritma ini memiliki performa yang paling unggul yang dibandingkan algoritma SVR dan XGBoost. Nilai MAE yang diperoleh berada di bawah 90 USD, lalu untuk nilai RMSE yang dihasilkan oleh model berada di bawah 120 USD. Hal ini berarti algoritma ini mampu melakukan generalisasi dengan sangat baik dan kemampuan dalam mengenali pola kompleks antar variabel juga efisien. Dengan nilai koefisien determinasi (R^2) berada pada kisaran 0.959 – 0.972, model ini mampu menjelaskan variasi harga Ethereum dengan baik dan menunjukkan kecocokan prediksi yang tinggi terhadap data sebenarnya.

Tabel 6 Rerata Evaluasi Algoritma LightGBM (80% data latih dan 20% data uji)

Price	MAE	RMSE	R2	Runtime
Close	85.553	116.306	0.959	0.088
Open	86.966	117.263	0.959	0.055
High	75.486	99.601	0.972	0.054

Low	82.525	114.160	0.959	0.052
-----	--------	---------	-------	-------

Pada Tabel 6 dapat dilihat bahwa variabel *High Price* menunjukkan hasil evaluasi model terbaik jika dibandingkan dengan variabel lainnya. Ini menandakan bahwa model memiliki kemampuan yang kuat dalam mengenali pola pergerakan harga maksimum harian Ethereum. Nilai koefisien determinasi (R^2) yang tertinggi pada angka 0.972 ini menandakan bahwa algoritma LightGBM menjadi algoritma yang cocok untuk digunakan pada prediksi harga Ethereum berdasarkan data historisnya. Selain itu, kinerja komputasi algoritma LightGBM memiliki waktu eksekusi yang berada pada kisaran 0.05 - 0.09 detik, menjadikan algoritma ini cepat sekaligus akurat. Contoh grafik perbandingan harga prediksi dan aktual untuk setiap variabel dapat dilihat pada Gambar 7.



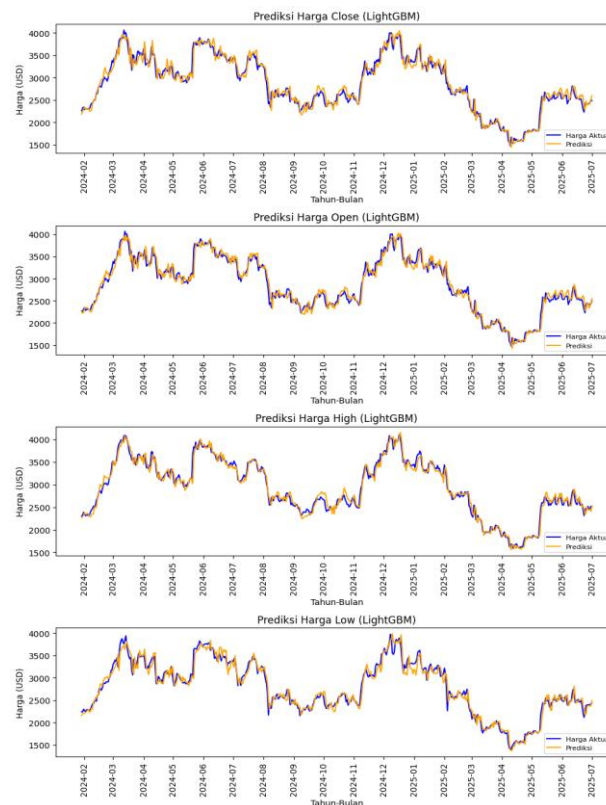
Gambar 7 Grafik harga *Close*, *Open*, *High*, dan *Low* (80% data latih, 20% data uji, dan *time step* = 3)

Konsistensi algoritma ini kembali diuji pada skenario pembagian data yang kedua. Pada pembagian data 70% data latih dan 30% data uji model mengalami penurunan yang sangat kecil. Nilai *error* yaitu MAE dan RMSE mengalami kenaikan tipis pada seluruh variabel. Sementara nilai koefisien determinasi (R^2) masih tetap tinggi yaitu diatas 0.95 pada semua variabel. Hal ini menunjukkan bahwa algoritma LightGBM tetap dapat bekerja secara stabil dan konsisten meskipun proporsi data latih berkurang secara signifikan. Selain itu waktu komputasi yang diperlukan juga relatif stabil dan bahkan pada beberapa skenario mengalami penurunan.

**Tabel 7 Rerata Evaluasi Algoritma LightGBM
(70% data latih dan 30% data uji)**

<i>Price</i>	MAE	RMSE	R2	Runtime
<i>Close</i>	93.987	125.562	0.956	0.076
<i>Open</i>	92.869	124.381	0.957	0.059
<i>High</i>	78.228	103.487	0.971	0.053
<i>Low</i>	91.438	124.741	0.955	0.053

Terlihat pada Tabel 7, variabel *High Price* masih menjadi variabel yang paling akurat dan stabil. Nilai MAE meningkat hanya sekitar 2.7 USD dan untuk nilai koefisien determinasinya (R2) hampir tidak menunjukkan perubahan. Ini menunjukkan bahwa fluktuasi harga tertinggi harian Ethereum masih dapat diprediksi dengan baik oleh model ini bahkan ketika proporsi data latih dikurangi. Berikut in grafik perbandingan harga prediksi dan aktualnya:



**Gambar 8 Grafik harga *Close*, *Open*, *High*, dan *Low*
(70% data latih, 30% data uji, dan *time step* =28)**

Berdasarkan serangkaian eksperimen yang telah dilakukan untuk masing-masing algoritma, dapat ditarik sebuah perbandingan kinerja secara komprehensif. Pada eksperimen yang telah dilakukan untuk masing-masing algoritma diperoleh hasil evaluasi pada pembagian data 80% data latih dan 20% uji bahwa secara keseluruhan ketiga algoritma yang digunakan menunjukkan kemampuan prediksi yang baik terhadap data harga Ethereum, namun dengan tingkat performa yang

berbeda. Algoritma SVR memberikan hasil yang paling rendah dibanding dua algoritma lainnya, namun kecepatan eksekusi algoritma SVR adalah yang tercepat. Di sisi lain algoritma XGBoost memiliki hasil yang lebih baik dibanding algoritma SVR meskipun kecepatan eksekusinya lebih lambat. Secara keseluruhan algoritma LightGBM memiliki hasil yang paling menjanjikan, hal ini dibuktikan dari hasil yang diberikan paling tinggi diantara algoritma lainnya. Untuk performanya algoritma LightGBM memiliki kecepatan eksekusi yang sangat cepat meskipun lebih lambat beberapa milidetik dari SVR, tapi hasil evaluasinya jauh lebih baik. Pada skenario pembagian data 70% data latih dan 30% data uji, masing-masing algoritma menunjukkan penurunan performa, namun algoritma LightGBM masih menunjukkan keunggulannya. Meskipun proporsi data latih dikurangi algoritma LightGBM masih dapat memprediksi harga Ethereum dengan sangat baik.

KESIMPULAN DAN SARAN

Cryptocurrency merupakan instrumen investasi yang beresiko sangat tinggi. Fluktuasi harga *cryptocurrency* khususnya Ethereum bisa disebabkan oleh beberapa faktor seperti faktor sosioekonomi dan geopolitik. *Volatilitas* harga *cryptocurrency* tersebut membuat penelitian untuk melakukan prediksi harga dengan metode *machine learning* diperlukan. Evaluasi yang dilakukan menggunakan beberapa algoritma dalam penelitian ini dilakukan untuk melihat algoritma mana yang lebih optimal dalam melakukan prediksi harga Ethereum yang fluktuatif. Hasil yang didapat pada penelitian ini bahwa algoritma LightGBM memiliki performa yang paling optimal dibanding dua algoritma lainnya yaitu SVR dan XGBoost. Performa tersebut dinilai berdasarkan dua skenario pembagian data yang berbeda yaitu 80% data latih dan 20% data uji serta 70% data latih dan 30% data uji. Sebagai contoh pada pembagian data 80% data latih dan 20% data uji rata-rata nilai MAE pada variabel *High Price* dengan algoritma SVR adalah 115.590, pada algoritma XGBoost nilainya adalah 77.314, sedangkan pada LightGBM nilainya adalah 75.486. Hal ini terjadi juga pada pembagian data 70% data latih dan 30% data uji, LightGBM masih menunjukkan superioritasnya dengan nilai rata-rata MAE untuk SVR, XGBoost, dan LightGBM adalah 104.356, 83.573, dan 78.228. Berdasarkan hasil tersebut diperoleh bahwa algoritma LightGBM bisa dipertimbangkan digunakan untuk prediksi harga Ethereum. Untuk penelitian selanjutnya, saran yang bisa diberikan adalah melakukan optimisasi parameter khususnya pada algoritma XGBoost dan LightGBM.

DAFTAR PUSTAKA

- Asnawi, M. F., Bisono, H. H., Megantara, M. A., & Kusrini, K. (2024). Aplikasi Prediksi Banjir Menggunakan Algoritma XGBoost Berbasis Website. *Journal of Economic, Management, Accounting and Technology*, 7(2), 379–389. <https://doi.org/10.32500/JEMATECH.V7I2.7644>
- Astuti, I. D., Rajab, S., & Setiyoudi, D. (2022). *Cryptocurrency Blockchain Technology in the Digital ReVolution Era*. *Aptisi Transactions on Technopreneurship (ATT)*, 4(1), 9–15. <https://doi.org/10.34306/ATT.V4I1.216>
- Astutiningsih, T., Saputro, D. R. S., & Sutanto. (2023). Optimasi Algoritme Xtreme Gradient boosting (XGBoost) pada Harga Saham PT United Tractors Tbk. *SPECTA*

- Journal of Technology*, 7(3), 632–641.
<https://doi.org/10.35718/specta.v7i3.1031>
- Cunha, P. R., Melo, P., & Sebastião, H. (2021). From bitcoin to central bank digital currencies: Making sense of the digital money reVolution. *Future Internet*, 13(7).
<https://doi.org/10.3390/FI13070165>
- De Vries, A. (2023). Cryptocurrencies on the road to sustainability: Ethereum paving the way for Bitcoin. *Patterns*, 4(1), 100633.
<https://doi.org/10.1016/j.patter.2022.100633>
- Ferdinand, F., Anthony, R., Winata, T. J., Sutanto, J., Souwiko, R., & Fernando, C. (2025). Comparison of *Support Vector Regression* and *Extreme Learning Machine* Methods for Predicting Bitcoin Prices. *Journal of Applied Informatics and Computing*, 9(5), 2121–2130. <https://doi.org/10.30871/JAIC.V9I5.10983>
- Gananta, I. M., Purnama, I. N., & Fredlina, K. Q. (2023). OPTIMASI PREDIKSI HARGA EMAS DENGAN METODE *SUPPORT VECTOR REGRESSION* (SVR) MENGGUNAKAN ALGORITMA GRID SEARCH. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(6), 3160–3165. <https://doi.org/10.36040/JATI.V7I6.8000>
- Immanuel, J. P., & Handhayani, T. (2025). PREDIKSI HARGA PANGAN JAYAPURA MENGGUNAKAN ELM, LSTM, LIGHTGBM, DAN GB. *Jurnal Ilmu Komputer Dan Sistem Informasi*, 13(2). <https://doi.org/10.24912/JIKSI.V13I2.35122>
- Khoiri, S. A., Wahid, A., & Koresponden, P. (2024). Analisis Kinerja Algoritma *Machine learning* dalam Prediksi Harga *Cryptocurrency*. *JUSTINDO (Jurnal Sistem Dan Teknologi Informasi Indonesia)*, 9(2), 133–141.
<https://doi.org/10.32528/JUSTINDO.V9I2.1965>
- Marzo, G. De, Pandolfelli, F., & Servedio, V. D. P. (2022). Modeling innovation in the *cryptocurrency* ecosystem. *Scientific Reports*, 12(1), 1–12.
<https://doi.org/10.1038/S41598-022-16924-7;SUBJMETA>
- Pradana Ananda Raharja, P. (2021). Prediksi Harga Ethereum Menggunakan Metode *Vector Autoregressive*. *Journal of Informatics Information System Software Engineering and Applications (INISTA)*, 3(2), 71–79.
<https://doi.org/10.20895/INISTA.V3I2.285>
- Reza, M. F., & Ghufro. (2025). IMPLEMENTATION OF THE *GATED RECURRENT UNIT* (GRU) OR *EXTREME GRADIENT BOOSTING* (XGBOOST) MODEL FOR ETHEREUM *CRYPTOCURRENCY PRICE* PREDICTION. *Jurnal Rekayasa Sistem Informasi Dan Teknologi*, 3(1), 209–222. <https://doi.org/10.70248/JRSIT.V3I1.3031>
- Robusti, C. da S., Avelar, A. B. A., Farina, M. C., & Gananca, C. A. (2025). *Blockchain* and *smart contracts*: transforming digital entrepreneurial finance and venture funding. *Journal of Small Business and Enterprise Development*, 32(3), 739–761.
<https://doi.org/10.1108/JSBED-12-2023-0584>
- Septiana Rizky, P., Haiban Hirzi, R., Hidayaturrohmah, U., Hamzanwadi Selong Jl TGKH Muhammad Zainuddin Abdul Madjid Pancor, U., & Timur, L. (2022). Perbandingan Metode LightGBM dan XGBoost dalam Menangani Data dengan Kelas Tidak Seimbang. *J Statistika: Jurnal Ilmiah Teori Dan Aplikasi Statistika*, 15(2), 228–236. <https://doi.org/10.36456/JSTAT.VOL15.NO2.A5548>
- Shin, D., & Rice, J. (2022). *Cryptocurrency*: A panacea for economic growth and sustainability? A critical review of crypto innovation. *Telematics and Informatics*,

71, 101830. <https://doi.org/10.1016/J.TELE.2022.101830>

Sianturi, T. B., Cholissodin, I., & Yudistira, N. (2023). Penerapan Algoritma *Long Short-Term Memory* (LSTM) berbasis Multi Fungsi Aktivasi Terbobot dalam Prediksi Harga Ethereum. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 7(3), 1101–1107.

Sriman, B., & Kumar, S. G. (2022). *Decentralized Finance* (DeFi): The Future of Finance and Defi Application for Ethereum *blockchain* based Finance Market. Proceedings - IEEE International Conference on Advances in Computing, Communication and Applied Informatics, ACCAI 2022. <https://doi.org/10.1109/ACCAI53970.2022.9752657>

Syam, U. A., Irdyanti, I., Magfirrah, I., Sartono, B., & Firdawanti, A. R. (2025). Evaluasi Kinerja Model *Random Forest* dan LightGBM untuk Klasifikasi Status Imunisasi Hepatitis B (HB-0) pada Balita. *Euler: Jurnal Ilmiah Matematika, Sains Dan Teknologi*, 13(1), 1–8. <https://doi.org/10.37905/EULER.V13I1.29762>

Tatachar, A. V. (2021). Comparative Assessment of Regression Models Based On Model Evaluation Metrics. *International Research Journal of Engineering and Technology*, 8(9), 853–860.

Theiri, S., Nekhili, R., & Sultan, J. (2023). *Cryptocurrency* liquidity during the Russia–Ukraine war: the case of Bitcoin and Ethereum. *The Journal of Risk Finance*, 24(1), 59–71. <https://doi.org/10.1108/JRF-05-2022-0103>

UMAIR. (2025, August). \$ETH Market Dominance Reaches 14.65%, *Highest Point In 2025*.

Zhang, S., & Mani, G. (2021). Popular cryptoassets (Bitcoin, Ethereum, and Dogecoin), Gold, and their relationships: *Volatility and correlation modeling*. *Data Science and Management*, 4, 30–39. <https://doi.org/10.1016/J.DSM.2021.11.001>