

PERBANDINGAN ALGORITMA REGRESI DALAM MEMPREDIKSI HARGA CRYPTOCURRENCY DOGE COIN

Jason Sutanto¹⁾ Tanjaya Jason Winata²⁾ Olivia Clarabella Khotiera³⁾ Wahyu Pamungkas⁴⁾

^{1) 2) 3) 4)} Program Studi Teknik Informatika, Fakultas Teknologi Informasi, Universitas Tarumanagara
Jln. Letjen S. Parman No. 1, Jakarta, 11440, Indonesia

Email : ¹⁾ jason.535220052@stu.untar.ac.id, ²⁾ tanjaya.535220041@stu.untar.ac.id, ³⁾ olivia.535220050@stu.untar.ac.id,
⁴⁾ wahyu.535220037@stu.untar.ac.id

ABSTRAK

Dogecoin merupakan salah satu aset *cryptocurrency* yang awalnya dibuat sebagai parodi namun berkembang menjadi aset digital dengan kapitalisasi pasar yang besar. Transaksi dan investasi menggunakan Dogecoin memiliki tingkat volatilitas yang tinggi sehingga harga sering berubah secara signifikan dalam waktu singkat. Oleh karena itu, diperlukan prediksi harga yang akurat untuk membantu pengambilan keputusan investasi. Penelitian ini bertujuan untuk membandingkan tiga algoritma *machine learning* linear, yaitu *Support Vector Regression* (SVR) dengan *Kernel Linear*, *Regresi Linear*, dan *AdaBoost* dalam memprediksi harga Dogecoin untuk satu hari ke depan. Dataset yang digunakan merupakan data historis harga harian Dogecoin dari tahun 2019 hingga 2025 yang diperoleh dari *Investing.com*, meliputi harga *Open*, *High*, *Low*, dan *Close*. Proses prapemrosesan data dilakukan menggunakan *Z-Score Normalization*, sedangkan pembagian data dilakukan dalam dua skenario, yaitu 80% pelatihan 20% pengujian dan 60% pelatihan 40% pengujian. Hasil penelitian menunjukkan bahwa model SVR *Kernel Linear* memiliki performa terbaik dengan *MAE* 0.00471 USD, *RMSE* 0.01355 USD, dan *R*² 0.98032, mengungguli *Regresi Linear* dan jauh lebih baik dibandingkan *AdaBoost*. Berdasarkan hasil evaluasi, SVR *Kernel Linear* terbukti sebagai model paling akurat dan stabil dalam memprediksi harga Dogecoin karena mampu menangkap pola hubungan linear antar variabel secara efektif.

Key words

AdaBoost, *Cryptocurrency*, *Dogecoin*, *Regresi Linear*, *SVR*

1. Pendahuluan

Era revolusi industri 4.0 telah membawa perubahan yang besar bagi seluruh aspek kehidupan manusia, termasuk sektor keuangan. Saat ini, kemajuan teknologi tidak hanya mengubah cara manusia berinteraksi dan

bekerja, tetapi juga memicu perubahan dalam sistem transaksi, investasi, dan pengelolaan data keuangan yang semakin cepat, efektif, dan berbasis digital [1]. Salah satu kemajuan teknologi dalam sektor keuangan adalah munculnya *cryptocurrency* sebagai instrumen keuangan baru yang menggabungkan inovasi dan teknologi.

Cryptocurrency merupakan aset digital yang beroperasi secara terdesentralisasi dengan memanfaatkan kriptografi sebagai sistem keamanannya. *Cryptocurrency* dirancang untuk bekerja melalui jaringan komputer tanpa bergantung pada otoritas pusat seperti pemerintah atau lembaga perbankan [2]. Karakteristik tersebut menjadikan *cryptocurrency* berbeda dari mata uang konvensional, karena transaksi dapat dilakukan secara transparan dan tidak mudah untuk dimanipulasi [3].

Popularitas *cryptocurrency* telah bertumbuh secara pesat dalam satu dekade terakhir. Hal tersebut dibuktikan dengan jumlah pengguna aset *cryptocurrency* diseluruh dunia pada akhir tahun 2024 diperkirakan telah mencapai lebih dari 560 juta pengguna [4] serta volume perdagangan global yang telah mencapai \$108 triliun [5]. Pasar *cryptocurrency* sendiri terdiri dari ribuan aset digital dengan karakteristik yang berbeda, mulai dari Bitcoin (BTC), Ethereum (ETH), USD Coin (USDC), dan *cryptocurrency* lain yang terus bermunculan. Salah satu diantaranya merupakan DogeCoin, yang pada awalnya diciptakan sebagai parodi, telah berkembang menjadi aset digital dengan komunitas yang besar dan kapitalisasi pasar yang signifikan [6].

Semua aset *cryptocurrency*, memiliki permasalahan utama yang besar yaitu volatilitas harga pasar. Harga dari sebuah *cryptocurrency* dapat berubah secara signifikan dalam waktu yang sangat dekat. Harga dari *cryptocurrency* dipengaruhi oleh berbagai jenis faktor seperti regulasi pemerintahan, sentimen publik hingga pemberitaan di media sosial. Tingkat volatilitas yang tinggi membuat *cryptocurrency* dianggap sebagai instrumen investasi yang beresiko tinggi dibandingkan aset keuangan konvensional [7]. Kondisi menyebabkan para investor menghadapi kesulitan dalam melakukan

analisis teknikal maupun memprediksi pergerakan harga dari *cryptocurrency* secara akurat.

Untuk mengatasi permasalahan tersebut, salah satu pendekatan yang mulai banyak digunakan adalah *machine learning*. Beberapa penelitian terkait sebelumnya telah menerapkan algoritma *machine learning* dalam memprediksi pergerakan harga *cryptocurrency*. Diantaranya penelitian yang berjudul “*Cryptocurrency Price Prediction Using Support Vector Regression*” menggunakan algoritma SVR kernel polynomial untuk memprediksi harga *open*, *high*, *low* dan *close* dari berbagai *cryptocurrency* diantaranya Dogecoin. Hasil dari penelitian menunjukkan rata-rata nilai R^2 sebesar 44,92% dan RMSE sebesar 11,3% berdasarkan 30 *cryptocurrency* yang diprediksi [8]. Penelitian lain yang berjudul “*Perbandingan Akurasi Algoritma XGBoost Dan SVR Dalam Prediksi Harga Cryptocurrency*” membandingkan algoritma XGBoost dengan SVR kernel RBF untuk prediksi harga 10 macam *cryptocurrency* diantaranya DogeCoin. Hasil dari penelitian menunjukkan bahwa algoritma *XGBoost* memiliki nilai MAPE yang lebih rendah sebesar 1,1600 sedangkan algoritma SVR kernel RBF sebesar 38,70 untuk DogeCoin [9]. Selanjutnya penelitian yang berjudul “*Averaged bars for cryptocurrency price forecasting across different horizons*” membandingkan algoritma *AdaBoost*, *LightGBM*, *Random Forest* dan *KNN* untuk memprediksi harga *close* menggunakan data harga *open*, *high*, dan *low* dari 10 jenis *cryptocurrency*, termasuk Dogecoin. Hasil menunjukkan bahwa AdaBoost memiliki nilai R^2 yang konsisten dan tinggi dengan nilai R^2 tertinggi yaitu 32% untuk Dogecoin [10]. Kemudian penelitian yang berjudul “*Comparison of Support Vector Regression and Extreme Learning Machine Methods for Predicting Bitcoin Prices*” membandingkan algoritma SVR Kernel Linear dengan *Extreme Learning Machine* yang memprediksi harga Bitcoin untuk satu hari kedepan. Hasil penelitian menunjukkan bahwa algoritma *Extreme Learning Machine* menghasilkan nilai R^2 tertinggi sebesar 0,987 [11]. Lalu, penelitian yang berjudul “*Cryptocurrency Price Prediction Using Supervised Machine Learning Algorithms*” membandingkan algoritma SVR, Regresi Linear, dan Decision Tree untuk memprediksi harga 4 *cryptocurrency* termasuk Dogecoin untuk 5 hari kedepan. Hasil penelitian menunjukkan R^2 regresi linear sebesar 92,50% untuk Dogecoin [12].

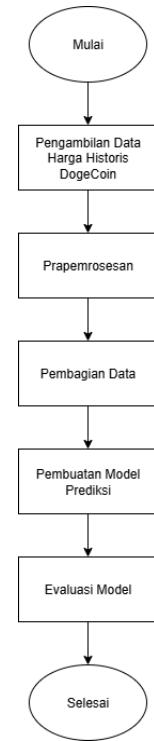
Algoritma yang dipakai dalam penelitian ini adalah *Support Vector Regression* (SVR) dengan kernel linear, AdaBoost dan Regresi Linear. Algoritma SVR adalah pengembangan dari algoritma *Support Vector Machine* (SVM) yang dirancang untuk kasus regresi untuk data linear maupun non-linear. SVR memiliki berbagai jenis kernel diantaranya linear, polinomial, RBF, dan Sigmoid [13]. Selain SVR, algoritma AdaBoost merupakan salah satu algoritma *ensemble learning* yang menggabungkan banyak model sederhana menjadi satu model prediksi yang lebih kuat [14]. Sementara itu, algoritma Regresi Linear merupakan metode paling sederhana dalam regresi yang digunakan untuk memodelkan hubungan antara

variabel independen dan variabel dependen dengan menarik garis lurus terbaik untuk meminimalkan selisih antara nilai prediksi dan nilai aktual [15].

Tujuan serta *novelty* dari penelitian ini adalah untuk membandingkan algoritma linear dari *machine learning* yaitu algoritma SVR dengan kernel linear, AdaBoost, dan Regresi Linear untuk memprediksi harga *open*, *high*, *low* dan *close* dari *cryptocurrency* Dogecoin. Keluaran yang dihasilkan dari perbandingan ketiga algoritma yang digunakan di penelitian ini adalah prediksi harga *open*, *high*, *low* dan *close* untuk satu hari kedepan.

2. Metode Penelitian

Penelitian ini akan membandingkan performa tiga algoritma *Machine Learning* dalam memprediksi harga *Cryptocurrency* Dogecoin, yaitu Regresi Linear, *AdaBoost*, dan *Support Vector Regression* (SVR) dengan kernel linear. Dataset yang akan digunakan berisi harga historis *Cryptocurrency* Dogecoin dalam Dolar Amerika Serikat (USD) yang diambil dari platform *Investing.com* [16]. Ilustrasi metodologi dalam penelitian ini dapat dilihat melalui *flowchart* pada Gambar 1.



Gambar 1. *Flowchart* Penelitian

2.1 Dataset

Data yang digunakan dalam penelitian ini merupakan data deret waktu (*time series*) harga Dogecoin yang bersumber dari platform *Investing.com*. Data disajikan secara harian dengan fitur-fitur utama seperti *Date*, *Open*, *High*, *Low*, *Close* (*Price*), *Volume* (VoL), serta *Change*

%. Fitur *Date* menunjukkan tanggal perdagangan, sedangkan *Open*, *High*, *Low*, dan *Close* masing-masing menggambarkan harga pembukaan, harga tertinggi, harga terendah, serta harga penutupan pada suatu hari yang tercatat dalam Dolar Amerika Serikat (USD). Fitur *VoL* mewakili jumlah Dogecoin yang diperdagangkan dalam satu hari, sementara fitur *Change %* menunjukkan persentase perubahan harga dibandingkan dengan hari sebelumnya.

2.2 Prapemrosesan

Dataset harga historis Dogecoin diunduh langsung dari platform Investing.com dalam format .csv. Dataset ini memuat catatan harga harian Dogecoin dalam periode 1 Januari 2019 hingga 1 Januari 2025, dengan total 2.193 baris data. Seluruh data ini tercatat dengan lengkap tanpa adanya nilai yang hilang/kosong (*missing value*), sehingga dapat langsung digunakan dalam tahap pengolahan.

Pada tahap awal prapemrosesan, dilakukan penghapusan dua fitur yang kurang relevan dengan kebutuhan penelitian, yaitu *VoL* dan *Change %*. Dengan demikian, fitur yang digunakan untuk pembuatan model adalah *Date*, *Open*, *High*, *Low*, dan *Price (Close)*. Selanjutnya, dilakukan proses standarisasi data menggunakan metode *Z-Score Normalization*. Metode ini mentransformasi setiap nilai data ke dalam bentuk skor standar. Proses ini menghasilkan distribusi dengan rata-rata 0 dan standar deviasi 1. Standarisasi ini dipilih karena sesuai untuk data dengan rentang nilai yang bervariasi dan tidak dibatasi pada interval tertentu. Dalam konteks prediksi harga Dogecoin, ini sangat membantu menyeimbangkan skala antar fitur sehingga algoritma *Machine Learning* yang digunakan dapat mempelajari pola pergerakan harga secara lebih stabil tanpa bias terhadap fitur dengan rentang nilai yang lebih besar. Tabel 1 menunjukkan sampel 10 baris pertama pada dataset.

Tabel 1 Sampel 10 Baris Pertama Dataset

Date	Price	Open	High	Low
01/01/2025	0,32501	0,31595	0,32722	0,31188
	4	6	6	2
12/31/2024	0,31596	0,31385	0,32851	0,30955
	4	4	3	6
12/30/2024	0,31386	0,31428	0,32295	
	4	8	1	1
12/29/2024	0,31426	0,32423	0,32944	0,31197
	4	9	8	1
12/28/2024	0,32422	0,31118	0,32698	0,31145
	4	6	6	7
12/27/2024	0,31118	0,31273		0,30902
	4	9	0,32465	3
12/26/2024	0,31275	0,33425	0,33783	0,30861
	4	5	6	1
12/25/2024	0,33424	0,33701	0,34193	0,32673
	4	5	1	1
12/24/2024	0,33701	0,32485	0,34028	0,31623
	4	2		4

12/23/2024	0,32484	0,31269	0,32848	0,30314
4	7	6	2	3

2.3 Pembagian Data

Tahap berikutnya adalah pembagian dataset ke dalam set pelatihan dan pengujian. Dikarenakan data yang digunakan merupakan deret waktu (*time series*), maka proses pembagian data tidak dilakukan secara acak agar urutan kronologis tetap terjaga dan pola temporal tidak hilang. Data dibagi secara berurutan, di mana bagian awal digunakan sebagai data latih dan bagian akhir digunakan sebagai data uji. Data kemudian diubah ke dalam bentuk pasangan input dan target dengan pendekatan *sliding window*. Metode ini memungkinkan pembuatan model dengan memanfaatkan beberapa nilai harga sebelumnya sebagai prediktor untuk memprediksi harga pada hari berikutnya. Pendekatan ini dipilih karena sesuai dengan karakteristik data *time series* yang memiliki ketergantungan antarperiode.

Dalam penelitian ini diterapkan dua skenario/eksperimen, yaitu pembagian data dengan perbandingan 60% untuk pelatihan dan 40% untuk pengujian, serta perbandingan 80% untuk pelatihan dan 20% untuk pengujian. Selain itu, ditetapkan juga panjang *time step* sebesar 3, artinya tiga hari berturut-turut harga Dogecoin akan dijadikan input untuk memprediksi harga pada hari keempat. Pemilihan panjang *time step* ini didasarkan pada pertimbangan bahwa interval tiga hari cukup merepresentasikan pola jangka pendek tanpa membuat dimensi input menjadi terlalu besar. Jika *time step* terlalu kecil, maka informasi penting dari pergerakan harga dapat terabaikan, sedangkan *time step* yang terlalu besar justru meningkatkan kompleksitas perhitungan dan berpotensi memasukkan data historis yang kurang relevan.

2.4 Regresi Linear

Regresi linear merupakan salah satu metode dalam *Machine Learning* yang digunakan untuk menganalisis hubungan antara satu variabel dependen (target) dengan satu atau lebih variabel independen (prediktor). Tujuan utama regresi linear adalah membentuk persamaan garis lurus yang dapat digunakan untuk memperkirakan nilai target berdasarkan data historis [17]. Regresi linear bekerja dengan mencari koefisien terbaik yang meminimalkan kesalahan prediksi melalui pendekatan *Ordinary Least Square* (OLS), yaitu dengan meminimalkan jumlah kuadrat selisih antara nilai aktual dan prediksi sehingga terbentuk garis lurus terbaik (*best fit line*) untuk mendekati pola data. Kelebihan regresi linear terletak pada kesederhanaannya, baik dari sisi interpretasi maupun kebutuhan komputasi, sehingga sering dijadikan baseline dalam perbandingan model [18].

Dalam penelitian ini, regresi linear digunakan dengan pendekatan time series regression, di mana setiap model

memanfaatkan tiga nilai sebelumnya (*timestep* = 3) untuk memprediksi nilai pada periode berikutnya. Dengan kata lain, input bagi model bukan hanya nilai pada satu waktu, tetapi juga nilai pada beberapa waktu sebelumnya, sehingga penyelesaian dapat menggunakan persamaan (1) [19].

$$\hat{y} = \alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 \quad (1)$$

Keterangan:

\hat{y} = nilai prediksi

α = koefisien intercept

$\beta_1, \beta_2, \beta_3$ = koefisien regresi untuk masing-masing waktu

x_1, x_2, x_3 = nilai variabel pada masing-masing waktu

2.5 AdaBoost Regressor

AdaBoost Regressor merupakan metode *ensemble learning* yang menggabungkan sejumlah model lemah (*weak learners*) secara bertahap untuk membentuk model prediksi yang lebih akurat [20]. Algoritma ini memiliki beberapa kelebihan, seperti meningkatkan akurasi, mengurangi bias dan varians, serta fleksibel dalam menangani data nonlinier. Prinsip utamanya adalah memberikan bobot lebih besar pada data yang mengalami kesalahan prediksi tinggi pada iterasi sebelumnya, sehingga model berikutnya lebih fokus memperbaiki kelemahan tersebut [21]. Pendekatan ini efektif karena kombinasi seluruh model dapat menghasilkan prediksi yang lebih akurat dibandingkan hanya menggunakan satu model saja.

Proses kerja *AdaBoost* dimulai dengan pemberian bobot sama pada seluruh data, kemudian melatih *weak learner* pertama. Setelah itu, bobot diperbarui berdasarkan besar kecilnya kesalahan prediksi pertama, dan *weak learner* baru dilatih kembali dengan bobot yang sudah disesuaikan. Proses ini diulang terus menerus hingga jumlah estimator M tercapai, dan hasil akhirnya merupakan kombinasi berbobot dari seluruh *weak learners*, yang bisa dirumuskan pada persamaan (2) [20].

$$\hat{y}(x) = \sum_{m=1}^M \alpha_m h_m(x) \quad (2)$$

Keterangan:

M = jumlah estimator

$h_m(x)$ = prediksi dari *weak learner* ke- m

α_m = bobot tingkat kontribusi model terhadap hasil akhir

2.6 Support Vector Regression

Support Vector Regression (SVR) merupakan sebuah algoritma ekstensi dari algoritma *Support Vector Machine* (SVM) yang diperuntukkan untuk tugas regresi. Berbeda dengan algoritma SVM yang berfungsi untuk

menyelesaikan masalah klasifikasi berdasarkan biner. Algoritma SVR digunakan untuk memprediksi nilai yang bersifat kontinu [13].

Cara kerja algoritma SVR, serupa dengan cara kerja SVM, yaitu dengan membentuk *regression hyperplane* beserta garis batas *margin* (*epsilon insensitive margin*) berdasarkan sebaran data, di mana titik data yang berada di luar margin dianggap sebagai *slack variables* dan titik di dekat margin disebut sebagai *support vectors* [22].

Untuk memperoleh fungsi regresi yang sederhana dan mudah dinterpretasikan, maka pada penelitian ini digunakanlah algoritma SVR kernel linear. Dengan kernel linear, model SVR dapat membentuk hubungan langsung antar variabel masukan dan keluaran dalam bentuk garis regresi yang meminimalkan kesalahan prediksi terhadap data. Persamaan regresi dari SVR kernel linear dirumuskan oleh persamaan (3). Agar fungsi regresi yang dibentuk memiliki tingkat kesalahan sekecil mungkin, SVR melakukan optimasi terhadap bobot dan biasnya. Persamaan dari fungsi objektif yang digunakan dalam SVR kernel linear dapat dilihat pada persamaan (4) [23].

$$f(x) = w\phi(x) + b \quad (3)$$

$$\min \left(\frac{1}{2} \|w\|^2 + c \sum_{i=1}^n (\xi_i + \xi_i^*) \right) \quad (4)$$

Keterangan:

w = bobot

b = bias

c = koefisien

ξ_i, ξ_i^* = *slack* variabel yang mewakili deviasi titik data dari batas margin

2.7 Evaluasi Model

Untuk menilai kinerja model prediksi harga Dogecoin, penelitian ini menggunakan tiga metrik, yaitu *Mean Absolute Error* (MAE), *Root Mean Square Error* (RMSE), dan Koefisien Determinasi (R^2). MAE mengukur rata-rata kesalahan absolut antara nilai prediksi dan nilai aktual, sehingga memberikan gambaran seberapa jauh prediksi secara rata-rata meleset tanpa memperhitungkan kesalahan. Semakin kecil nilai MAE, maka semakin baik model dalam menghasilkan prediksi yang mendekati nilai aktual. RMSE adalah akar kuadrat dari rata-rata kuadrat kesalahan, yang memberi penalti lebih besar terhadap error besar karena kesalahan dikuadratkan. Nilai RMSE yang kecil menunjukkan prediksi lebih akurat, sedangkan nilai yang tinggi menunjukkan adanya kesalahan besar yang signifikan [24].

Sementara itu, R^2 atau koefisien determinasi, menunjukkan proporsi varians atau variasi dari variabel target yang dapat dijelaskan oleh model. Nilai R^2 yang mendekati 1 menandakan model berhasil menjelaskan variasi data dengan baik, sedangkan R^2 mendekati 0 (atau bahkan negatif) menunjukkan model gagal menjelaskan pola data. Cara untuk menghitung MAE, RMSE, dan R^2

dituliskan pada persamaan (5), (6), dan (7) berturut-turut [25].

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

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

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

Keterangan:

n = jumlah sampel data

y_i = nilai aktual data ke- i

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

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

3. Hasil Percobaan

3.1 Regresi Linear

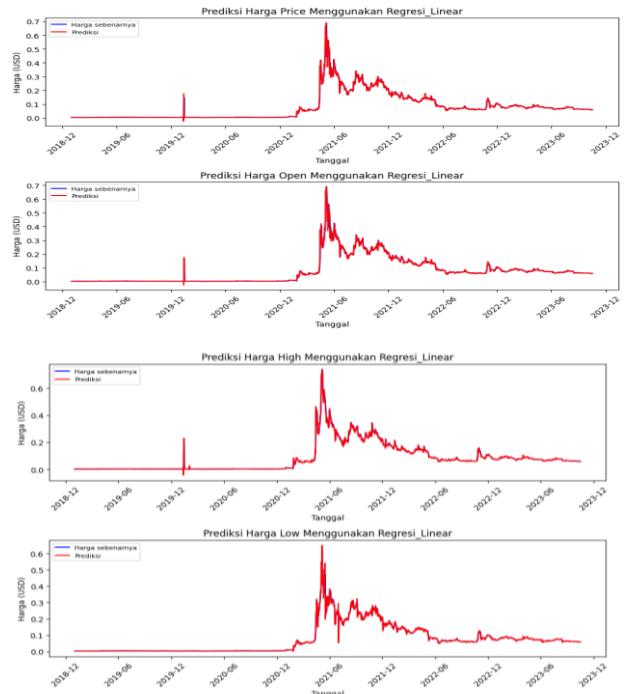
Hasil percobaan pada model Regresi Linear ini dilakukan dengan dua skenario pembagian data yaitu dengan 80% untuk pelatihan dan 20% untuk pengujian serta 60% untuk pelatihan dan 40% untuk pengujian. Eksperimen pembagian data 80% pelatihan dan 20% pengujian untuk model Regresi Linear menunjukkan MAE sebesar 0,00457, RMSE sebesar 0,01414 dan R^2 sebesar 0,97858 untuk harga *price (close)*. Nilai MAE menunjukkan bahwa rata-rata kesalahan absolut antara harga aktual dan hasil prediksi sekitar 0,00457 USD, sedangkan RMSE menunjukkan bahwa rata-rata besarnya penyimpangan sekitar 0,01414 USD. Nilai R^2 sebesar 0,97858 menunjukkan bahwa model Regresi Linear mampu menangkap pola hubungan antara variabel input dan harga *close* dengan sangat baik. Keseluruhan evaluasi model Regresi Linear untuk eksperimen pembagian data 80% pelatihan dan 20% pengujian ditunjukkan oleh Tabel 2.

Tabel 2 Hasil Evaluasi Model Regresi Linear (80% latih dan 20% uji)

Harga	MAE	RMSE	R ²
Price (Close)	0,00457	0,01414	0,97858
Open	0,00457	0,01411	0,97867
High	0,00441	0,01362	0,98277
Low	0,00417	0,01411	0,97479

Berdasarkan tabel 2, prediksi harga *high* memiliki performa terbaik dibandingkan dengan prediksi harga *price (close)*, *open*, dan *low*. Dengan nilai MAE dan RMSE terkecil yaitu sebesar 0,00441 USD (MAE) dan 0,01362 USD (RMSE) serta nilai R^2 sebesar 0,98277. Secara keseluruhan, keempat prediksi harga menunjukkan nilai R^2 di atas 0,95, yang menandakan

bahwa model Regresi Linear memiliki kemampuan prediksi yang tinggi dan stabil untuk seluruh jenis harga, dengan performa terbaik pada harga *high*. Grafik dari perbandingan nilai aktual dan prediksi model Regresi Linear untuk eksperimen pembagian data 80% pelatihan dan 20% pengujian ditunjukkan oleh Gambar 1.



Gambar 1 Grafik Perbandingan Harga Aktual Dan Harga Prediksi Regresi Linear (80% latih dan 20% uji)

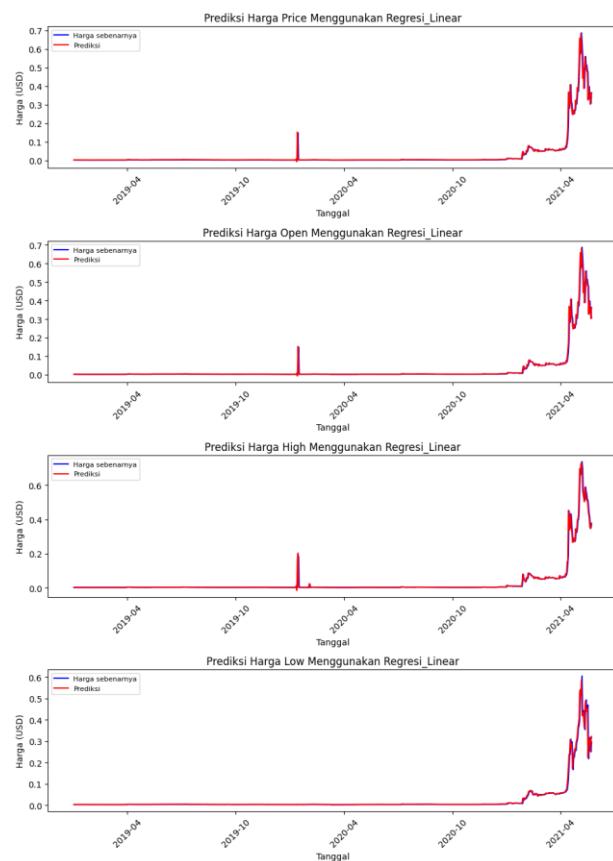
Berdasarkan Gambar 1, terdapat dua garis yaitu garis berwarna merah (prediksi) memiliki pola yang sangat mendekati garis berwarna biru (aktual) untuk keempat jenis harga, *price (close)*, *open*, *high* dan *low*. Pada keempat grafik, terutama pada pertengahan tahun 2021, model Regresi Linear dapat mengikuti pola lonjakan harga yang signifikan dan penurunan harga setelahnya dengan arah tren yang serupa dengan data aktual. Model juga mampu memprediksi tren penurunan yang berlangsung secara konsisten hingga akhir tahun 2023.

Sedangkan eksperimen pembagian data 60% pelatihan dan 40% pengujian untuk model Regresi Linear mengalami sedikit penurunan terhadap tiap jenis harga dibandingkan dengan eksperimen sebelumnya. Dengan performa terbaik untuk harga *high* yaitu MAE sebesar 0,00380 USD, RMSE sebesar 0,01664 USD dan R^2 sebesar 0,97132. Keseluruhan evaluasi model Regresi Linear untuk eksperimen pembagian data 60% pelatihan dan 40% pengujian ditunjukkan oleh Tabel 3.

Tabel 3 Hasil Evaluasi Model Regresi Linear (60% latih dan 40% uji)

Harga	MAE	RMSE	R ²
Price (Close)	0,00422	0,01704	0,96243
Open	0,00418	0,01700	0,96199
High	0,00380	0,01664	0,97132
Low	0,00339	0,01460	0,96219

Berdasarkan tabel 3, terlihat bahwa pembagian data pelatihan yang lebih kecil (60%) menyebabkan performa model sedikit menurun dibandingkan pembagian 80% pelatihan dan 20% pengujian. Namun demikian, nilai R^2 yang tetap berada di atas 0,95 menunjukkan bahwa model Regresi Linear masih mampu melakukan prediksi dengan tingkat akurasi yang baik. Grafik dari perbandingan nilai aktual dan prediksi model Regresi Linear untuk eksperimen pembagian data 60% pelatihan dan 40% pengujian ditunjukkan oleh Gambar 2.



Gambar 2 Grafik Perbandingan Harga Aktual Dan Harga Prediksi Regresi Linear (60% latih dan 40% uji)

3.2 AdaBoost Regressor

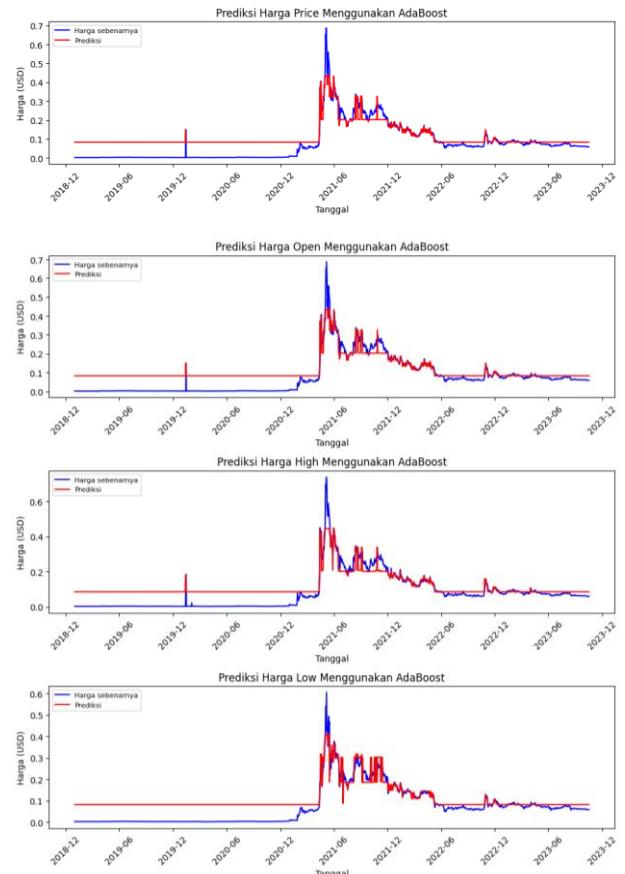
Eksperimen pembagian data 80% pelatihan dan 20% pengujian untuk model *AdaBoost* menunjukkan MAE sebesar 0,04628, RMSE sebesar 0,05814 dan R^2 sebesar 0,63778 untuk harga *price (close)*. Nilai MAE menunjukkan bahwa rata-rata kesalahan absolut antara harga aktual dan hasil prediksi yang cukup besar yaitu sekitar 0,04628 USD, sedangkan RMSE menunjukkan bahwa rata-rata besarnya penyimpangan sekitar 0,05814 USD. Nilai R^2 sebesar 0,63778 menunjukkan bahwa model *AdaBoost* kurang mampu menangkap pola hubungan antara variabel input dan harga *close* jika dibandingkan dengan model Regresi Linear. Keseluruhan evaluasi model *AdaBoost* untuk eksperimen pembagian

data 80% pelatihan dan 20% pengujian ditunjukkan oleh Tabel 4.

Tabel 4 Hasil Evaluasi Model *AdaBoost* (80% latih dan 20% uji)

Harga	MAE	RMSE	R^2
Price (Close)	0,04628	0,05814	0,63778
Open	0,04584	0,05771	0,64320
High	0,04824	0,06083	0,65632
Low	0,04432	0,05548	0,61034

Berdasarkan tabel 4, prediksi harga *high* tetap memiliki performa terbaik dibandingkan dengan prediksi harga *price (close)*, *open*, dan *low*. Dengan nilai MAE dan RMSE terkecil yaitu sebesar 0,04824 USD (MAE) dan 0,06083 USD (RMSE) serta nilai R^2 sebesar 0,65632. Secara keseluruhan, keempat prediksi harga menunjukkan nilai R^2 di atas 0,60, yang menandakan bahwa model *AdaBoost* kurang mampu untuk memprediksi seluruh jenis harga jika dibandingkan dengan model Regresi Linear. Grafik dari perbandingan nilai aktual dan prediksi model *AdaBoost* untuk eksperimen pembagian data 80% pelatihan dan 20% pengujian ditunjukkan oleh Gambar 3.



Gambar 3 Grafik Perbandingan Harga Aktual Dan Harga Prediksi *AdaBoost* (80% latih dan 20% uji)

Berdasarkan Gambar 3, model *AdaBoost* mampu mengikuti pola umum pergerakan harga pada keempat jenis harga, *price (close)*, *open*, *high* dan *low*. Keempat

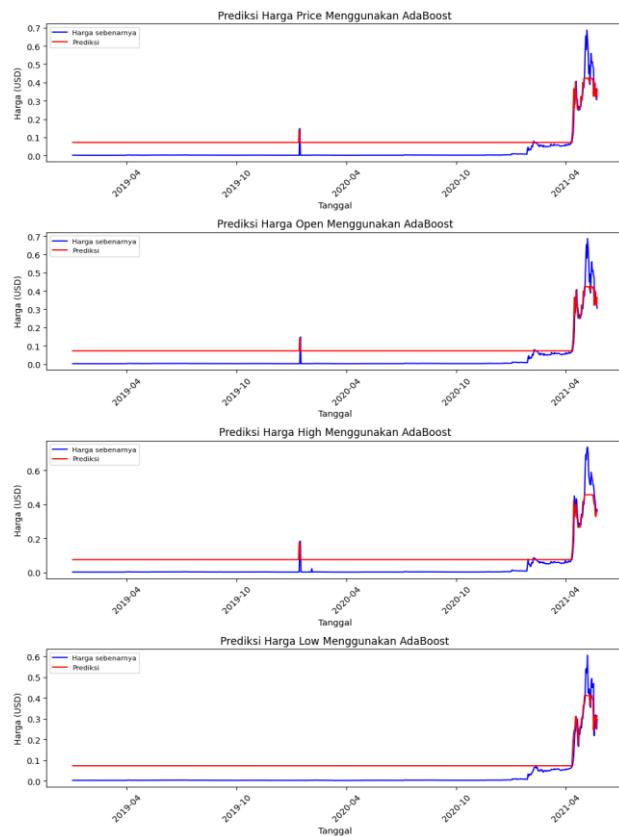
grafik menunjukkan bahwa model *AdaBoost* kurang responsif terhadap perubahan harga yang cepat, hal tersebut ditunjukkan oleh *gap* yang cukup besar antara nilai prediksi dan nilai aktual pada beberapa periode, terutama ketika harga Dogecoin mengalami pelonjakan harga secara tajam di tahun 2021. Pada periode tersebut, nilai yang diprediksi tidak mampu mencapai puncak harga sebenarnya.

Sedangkan eksperimen pembagian data 60% pelatihan dan 40% pengujian untuk model *AdaBoost* mengalami penurunan yang drastis terhadap tiap jenis harga dibandingkan dengan eksperimen sebelumnya. Dengan performa terbaik untuk harga *high* yaitu MAE sebesar 0,06826 USD, RMSE sebesar 0,07216 USD dan R^2 sebesar 0,46064. Keseluruhan evaluasi model *AdaBoost* untuk eksperimen pembagian data 60% pelatihan dan 40% pengujian ditunjukkan oleh Tabel 5.

Tabel 5 Hasil Evaluasi Model *AdaBoost* (60% latih dan 40% uji)

Harga	MAE	RMSE	R^2
Price (Close)	0,06539	0,06859	0,39160
Open	0,06536	0,06856	0,38177
High	0,06826	0,07216	0,46064
Low	0,06492	0,06726	0,19694

Berdasarkan tabel 5, terlihat bahwa pembagian data pelatihan yang lebih kecil (60%) menyebabkan performa model menurun secara drastis dibandingkan pembagian 80% pelatihan dan 20% pengujian. Nilai MAE, RMSE, dan R^2 untuk seluruh jenis harga menunjukkan perbedaan yang signifikan, di mana kesalahan prediksi meningkat, dan kemampuan model untuk menjelaskan variasi data menurun secara drastis. Jika dibandingkan dengan eksperimen sebelumnya, di mana prediksi harga *price (close)* menghasilkan MAE sebesar 0,04628 USD, RMSE sebesar 0,05814 USD dan R^2 sebesar 0,63778, eksperimen dengan data latih sebesar 60% menyebabkan nilai MAE dan RMSE meningkat, sedangkan nilai R^2 menurun hampir 40% dari performa eksperimen sebelumnya. Hal ini menunjukkan bahwa jumlah data pelatihan yang lebih sedikit berdampak langsung terhadap kemampuan generalisasi model *AdaBoost*. Grafik dari perbandingan nilai aktual dan prediksi model *AdaBoost* untuk eksperimen pembagian data 60% pelatihan dan 40% pengujian ditunjukkan oleh Gambar 4.



Gambar 4 Grafik Perbandingan Harga Aktual Dan Harga Prediksi *AdaBoost* (60% latih dan 40% uji)

3.3 SVR Kernel Linear

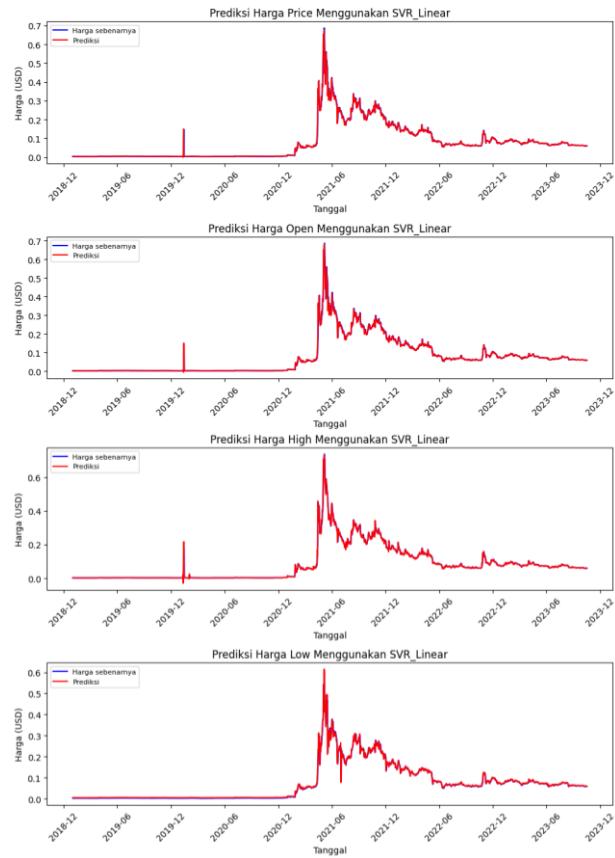
Eksperimen pembagian data 80% pelatihan dan 20% pengujian untuk model SVR Kernel Linear menunjukkan MAE sebesar 0,00471, RMSE sebesar 0,01355 dan R^2 sebesar 0,98032 untuk harga *price (close)*. Nilai MAE menunjukkan bahwa rata-rata kesalahan absolut antara harga aktual dan hasil prediksi yang sedikit yaitu sekitar 0,00471 USD, sedangkan RMSE menunjukkan bahwa rata-rata besarnya penyimpangan sekitar 0,01355 USD. Nilai R^2 sebesar 0,98032 menunjukkan bahwa model SVR Kernel Linear sangat mampu menangkap pola hubungan antara variabel input dan harga *close* serta lebih unggul jika dibandingkan dengan model Regresi Linear dan *AdaBoost*. Keseluruhan evaluasi model *AdaBoost* untuk eksperimen pembagian data 80% pelatihan dan 20% pengujian ditunjukkan oleh Tabel 6.

Tabel 6 Hasil Evaluasi Model SVR Kernel Linear (80% latih dan 20% uji)

Harga	MAE	RMSE	R^2
Price (Close)	0,00471	0,01355	0,98032
Open	0,00451	0,01357	0,98027
High	0,00452	0,01354	0,98298
Low	0,00501	0,01340	0,97725

Berdasarkan tabel 6, prediksi harga *high* tetap memiliki performa terbaik dibandingkan dengan prediksi harga *price (close)*, *open*, dan *low*. Dengan nilai MAE dan

RMSE terkecil yaitu sebesar 0,00452 USD (MAE) dan 0,01354 USD (RMSE) serta nilai R^2 sebesar 0,98298. Secara keseluruhan, keempat prediksi harga menunjukkan nilai R^2 di atas 0,97, yang menandakan bahwa model SVR Kernel Linear mampu dan unggul dalam memprediksi seluruh jenis harga jika dibandingkan dengan model Regresi Linear dan *AdaBoost*. Grafik dari perbandingan nilai aktual dan prediksi model SVR Kernel Linear untuk eksperimen pembagian data 80% pelatihan dan 20% pengujian ditunjukkan oleh Gambar 5.



Gambar 5 Grafik Perbandingan Harga Aktual Dan Harga Prediksi SVR Kernel Linear (80% latih dan 20% uji)

Berdasarkan Gambar 5, model SVR Kernel Linear menunjukkan hasil yang serupa dengan model Regresi Linear, dimana tingkat kesesuaianya sangat baik terhadap nilai aktual dari keempat jenis harga *price (close)*, *open*, *high* dan *low*. Hal tersebut terlihat jelas pada grafik, khususnya pada periode sekitar tahun 2021. Pada periode tersebut model mampu mengikuti pola kenaikan tajam dan penurunan harga yang terjadi tanpa perbedaan signifikan antara nilai prediksi dan nilai aktual.

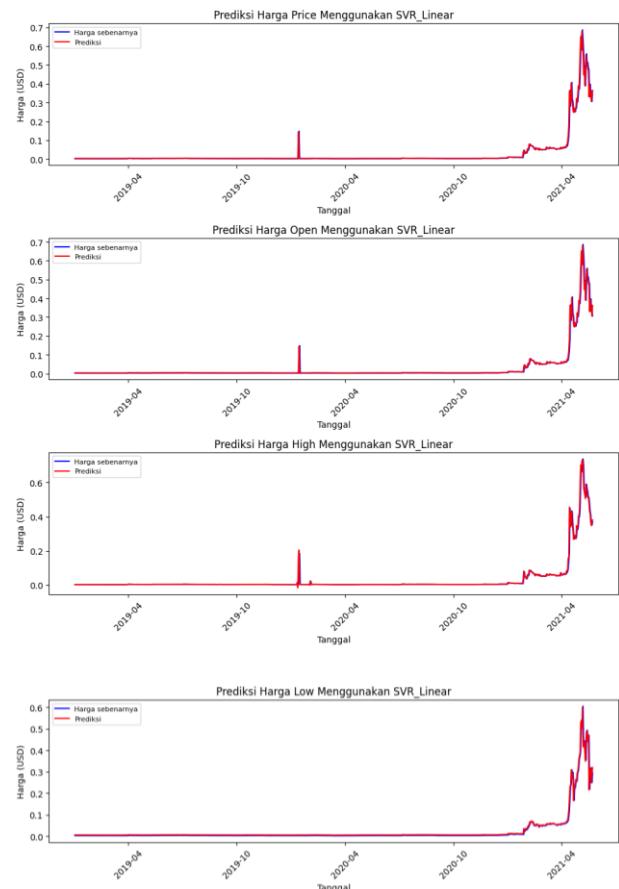
Sedangkan eksperimen pembagian data 60% pelatihan dan 40% pengujian untuk model SVR Kernel Linear mengalami sedikit penurunan terhadap tiap jenis harga dibandingkan dengan eksperimen sebelumnya. Dengan performa terbaik untuk harga *high* yaitu MAE sebesar 0,00345 USD, RMSE sebesar 0,01666 USD dan R^2 sebesar 0,97124. Keseluruhan evaluasi model SVR

Kernel Linear untuk eksperimen pembagian data 60% pelatihan dan 40% pengujian ditunjukkan oleh Tabel 7.

Tabel 7 Hasil Evaluasi Model SVR Kernel Linear (60% latih dan 40% uji)

Harga	MAE	RMSE	R^2
Price (Close)	0,00440	0,01678	0,96359
Open	0,00435	0,01675	0,96310
High	0,00345	0,01666	0,97124
Low	0,00517	0,01504	0,95985

Berdasarkan tabel 7, terlihat bahwa pembagian data pelatihan yang lebih kecil (60%) menyebabkan performa model menurun walaupun tidak signifikan, dibandingkan pembagian 80% pelatihan dan 20% pengujian. Nilai MAE, RMSE, dan R^2 untuk seluruh jenis harga menunjukkan perbedaan, di mana kesalahan prediksi meningkat, dan kemampuan model untuk menjelaskan variasi data menurun secara drastis. Grafik dari perbandingan nilai aktual dan prediksi model *AdaBoost* untuk eksperimen pembagian data 60% pelatihan dan 40% pengujian ditunjukkan oleh Gambar 6.



Gambar 6 Grafik Perbandingan Harga Aktual Dan Harga Prediksi SVR Kernel Linear (60% latih dan 40% uji)

4. Kesimpulan

Berdasarkan hasil pengujian pada seluruh jenis harga (*open*, *high*, *low* dan *close*) serta pada kedua skenario pembagian data, model SVR Kernel Linear secara konsisten memberikan hasil terbaik dibandingkan dua algoritma lainnya. Pada skenario pembagian data 80% pelatihan dan 20% pengujian, model SVR Kernel Linear memperoleh MAE sebesar 0,00471 USD, RMSE sebesar 0,01355 USD dan R^2 sebesar 0,98032 untuk harga *price (close)*. Sebagai perbandingan, model Regresi Linear menghasilkan MAE sebesar 0,00457 USD, RMSE sebesar 0,01414 USD, dan R^2 sebesar 0,97858, sedangkan model *AdaBoost* menunjukkan performa yang jauh lebih rendah dengan MAE sebesar 0,04628 USD, RMSE sebesar 0,05814 USD dan R^2 sebesar 0,63778. Nilai kesalahan yang sangat kecil pada SVR Kernel Linear menunjukkan bahwa prediksi model ini sangat akurat dengan deviasi rata-rata hanya sekitar 0,0047 USD dari harga aktual, jauh lebih kecil dibandingkan dengan *AdaBoost* dan lebih unggul daripada Regresi Linear. Ketika dilakukan pembagian data 60% pelatihan dan 40% pengujian, model SVR Kernel Linear tetap menunjukkan performa yang stabil dengan MAE sebesar 0,00440 USD, RMSE sebesar 0,01678 USD dan R^2 sebesar 0,96359, menandakan penurunan akurasi yang tidak signifikan. Jika dibandingkan dengan model *AdaBoost*, SVR Kernel Linear mampu mengurangi MAE hingga sekitar 90% dan RMSE lebih dari 75%. Sementara itu, dibandingkan dengan Regresi Linear, SVR Kernel Linear memiliki nilai error yang sedikit lebih rendah dan nilai R^2 yang lebih tinggi. Secara keseluruhan, nilai R^2 di atas 0,97 untuk semua jenis harga membuktikan bahwa algoritma SVR Kernel Linear mampu menangkap pola pergerakan harga Dogecoin dengan sangat baik. Berdasarkan percobaan yang telah dilakukan, algoritma SVR Kernel Linear menjadi algoritma yang unggul dan akurat dalam memprediksi harga Dogecoin dibandingkan algoritma linear lainnya seperti Regresi Linear dan *AdaBoost*, karena mampu memberikan keseimbangan terbaik antara akurasi dan stabilitas pada kedua skenario.

REFERENSI

- [1] B. Machkour and A. Abriane, "Industry 4.0 and its Implications for the Financial Sector," *Procedia Comput Sci*, vol. 177, pp. 496–502, 2020, doi: 10.1016/j.procs.2020.10.068.
- [2] M. El Hajj and I. Farran, "The Cryptocurrencies in Emerging Markets: Enhancing Financial Inclusion and Economic Empowerment," *Journal of Risk and Financial Management*, vol. 17, no. 10, p. 467, Oct. 2024, doi: 10.3390/jrfm17100467.
- [3] Z. Qaroush, S. Zakerneh, and A. Dawabsheh, "Cryptocurrencies Advantages and Disadvantages: A Review," *International Journal of Applied Sciences and Smart Technologies*, vol. 4, no. 1, pp. 1–20, Jun. 2022, doi: 10.24071/ijassst.v4i1.4610.
- [4] D. Bianchi, M. Babiak, and A. Dickerson, "Trading volume and liquidity provision in cryptocurrency markets," *J Bank Financ*, vol. 142, p. 106547, Sep. 2022, doi: 10.1016/j.jbankfin.2022.106547.
- [5] F. Fang *et al.*, "Cryptocurrency Trading: A Comprehensive Survey," Jan. 2022.
- [6] M. Ben Osman, C. Urom, K. Guesmi, and R. Benkraiem, "Economic sentiment and the cryptocurrency market in the post-COVID-19 era," *International Review of Financial Analysis*, vol. 91, p. 102962, Jan. 2024, doi: 10.1016/j.irfa.2023.102962.
- [7] M. S. Ahmed, A. A. El-Masry, A. I. Al-Maghyereh, and S. Kumar, "Cryptocurrency volatility: A review, synthesis, and research agenda," *Res Int Bus Finance*, vol. 71, p. 102472, Aug. 2024, doi: 10.1016/j.ribaf.2024.102472.
- [8] T. Stephen, "CRYPTOCURRENCY PRICE PREDICTION USING SUPPORT VECTOR REGRESSION," *International Journal of Application on Sciences*, vol. 1, no. 4, 2023, doi: 10.24912/ijaste.v1.i4.1657-1667.
- [9] Nazwa Fadhil, "PERBANDINGAN AKURASI ALGORITMA XGBOOST DAN SVR DALAM PREDIKSI HARGA CRYPTOCURRENCY," *Jurnal Ilmu Komputer dan Sistem Informasi*, vol. 13, no. 1, Jan. 2025, doi: 10.24912/jiksi.v13i1.32865.
- [10] A. El Youssefi, A. Hessane, I. Zeroual, and Y. Farhaoui, "Averaged bars for cryptocurrency price forecasting across different horizons," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 14, no. 3, p. 1910, Jun. 2025, doi: 10.11591/ijai.v14.i3.pp1910-1918.
- [11] F. Ferdinand, R. Anthony, T. Jason Winata, J. Sutanto, R. Souwiko, and C. Fernando, "Comparison of Support Vector Regression and Extreme Learning Machine Methods for Predicting Bitcoin Prices," *Journal of Applied Informatics and Computing (JAIC)*, vol. 9, no. 5, p. 2121, 2025, [Online]. Available: <http://jurnal.polibatam.ac.id/index.php/JAIC>
- [12] D. Chaudhary and S. K. Saroj, "Cryptocurrency Price Prediction Using Supervised Machine Learning Algorithms," *ADCAIJ: Advances in Distributed Computing and Artificial Intelligence Journal*, vol. 12, p. e31490, Dec. 2023, doi: 10.14201/adcaij.31490.
- [13] Y. Jia, S. Zhou, Y. Wang, F. Lin, and Z. Gao, "A quadratic v-support vector regression approach for load forecasting," *Complex & Intelligent Systems*, vol. 11, no. 1, p. 123, Jan. 2025, doi: 10.1007/s40747-024-01730-7.
- [14] E. M. Ferrouhi and I. Bouabdallaoui, "A comparative study of ensemble learning algorithms for high-frequency trading," *Sci Afr*, vol. 24, p. e02161, Jun. 2024, doi: 10.1016/j.sciaf.2024.e02161.
- [15] S. Syed, S. M. Talha, A. Iqbal, N. Ahmad, and M. A. Alshara, "Seeing Beyond Noise: Improving Cryptocurrency Forecasting with Linear Bias Correction," *AI*, vol. 5, no. 4, pp. 2829–2851, Dec. 2024, doi: 10.3390/ai5040136.
- [16] Investing.com, "Dogecoin Historical Data 2019 - 2025." Accessed: Oct. 07, 2025. [Online]. Available: <https://www.investing.com/indices/investing.com-doge-usd-historical-data>
- [17] N. Roustaei, "Application and interpretation of linear-regression analysis," *Medical hypothesis discovery and innovation in ophthalmology*, vol. 13, no. 3, pp. 151–159, Oct. 2024, doi: 10.51329/mehdiophthal1506.
- [18] J. Lu, "A rigorous introduction to linear models," May 2025, [Online]. Available: <http://arxiv.org/abs/2105.04240>
- [19] F. Ding, "The Research on the Prediction of Cryptocurrency Based on Linear Regression and LSTM,"

- 2023, pp. 1472–1481. doi: 10.2991/978-94-6463-102-9_154.
- [20] P. Beja-Battais, “Overview of AdaBoost : Reconciling its views to better understand its dynamics.” Oct. 2023, [Online]. Available: <http://arxiv.org/abs/2310.18323>
 - [21] K. Abegaz and İ. Etikan, “Boosting the Performance of Artificial Intelligence-Driven Models in Predicting COVID-19 Mortality in Ethiopia,” *Diagnostics*, vol. 13, no. 4, p. 658, Feb. 2023, doi: 10.3390/diagnostics13040658.
 - [22] A. G. Usman *et al.*, “Optimized SVR with nature-inspired algorithms for environmental modelling of mycotoxins in food virtual-water samples,” *Sci Rep*, vol. 15, no. 1, p. 16569, May 2025, doi: 10.1038/s41598-025-99908-7.
 - [23] K.-L. Du, B. Jiang, J. Lu, J. Hua, and M. N. S. Swamy, “Exploring Kernel Machines and Support Vector Machines: Principles, Techniques, and Future Directions,” *Mathematics*, vol. 12, no. 24, p. 3935, Dec. 2024, doi: 10.3390/math12243935.
 - [24] T. O. Hodson, “Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not,” *Geosci Model Dev*, vol. 15, no. 14, pp. 5481–5487, Jul. 2022, doi: 10.5194/gmd-15-5481-2022.
 - [25] 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.