

Prediksi Harga Ethereum Menggunakan Orange Data Mining

Ethereum Price Prediction Using Orange Data Mining

Teddy Al Ghozali¹, Anjasmara², M. Dwi Julian S³, Riyad Dinargo⁴, Pujiyanto⁵

Program Studi Teknik Informatika, Fakultas Teknik dan Komputer, Universitas Baturaja

e-mail: teddyalghozali29@gmail.com¹, anjas0223@gmail.com²,

juliansaputra139@gmail.com³, riyat26@gmail.com⁴, pujiyanto.mail@gmail.com⁵

Abstrak

Meningkatnya penggunaan Ethereum sebagai aset kripto mendorong perlunya sistem prediksi harga yang akurat guna membantu investor dalam mengambil keputusan yang lebih tepat. Penelitian ini bertujuan untuk memprediksi harga Ethereum dengan menggunakan aplikasi Orange Data Mining sebagai alat bantu analisis visual. Dua algoritma pembelajaran mesin digunakan dalam penelitian ini, yaitu K-Nearest Neighbor (K-NN) dan Support Vector Machine (SVM). Data yang digunakan merupakan data historis harga Ethereum yang diperoleh dari sumber publik, dengan fitur meliputi harga pembukaan, penutupan, tertinggi, terendah, volume perdagangan, serta persentase perubahan harga harian. Pengujian dilakukan menggunakan teknik K-Fold Cross Validation (K=5), sedangkan evaluasi model dilakukan berdasarkan metrik Root Mean Square Error (RMSE) dan Mean Absolute Error (MAE). Hasil penelitian menunjukkan bahwa K-NN menghasilkan nilai kesalahan yang lebih rendah dibandingkan SVM, sehingga dapat disimpulkan bahwa K-NN merupakan metode yang lebih akurat untuk prediksi harga Ethereum.

Kata kunci – Ethereum; K-Nearest Neighbor; Orange Data Mining; Prediksi Harga; Support Vector Machine

Abstract

The growing use of Ethereum as a crypto asset has increased the need for accurate price prediction systems to support investors in making better-informed decisions. This study aimed to predict Ethereum prices using the Orange Data Mining application as a visual analytical tool. Two machine learning algorithms were employed in this research: K-Nearest Neighbor (K-NN) and Support Vector Machine (SVM). The dataset consisted of historical Ethereum price data collected from public sources, with features including opening price, closing price, highest price, lowest price, trading volume, and daily price change percentage. The models were tested using the K-Fold Cross Validation technique (K=5), and performance was evaluated using Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) metrics. The results showed that the K-NN algorithm yielded lower error values compared to SVM. Therefore, K-NN is recommended as a more accurate method for predicting Ethereum prices.

Keywords – Ethereum; K-Nearest Neighbor; Orange Data Mining; Price Prediction; Support Vector Machine

I. Pendahuluan

Ethereum merupakan aset kripto terbesar kedua setelah Bitcoin yang memiliki peran signifikan dalam dunia keuangan digital. Tidak hanya berfungsi sebagai alat tukar digital, Ethereum juga menjadi tulang punggung bagi berbagai aplikasi terdesentralisasi (dApps) dan pelaksanaan *smart contract*. Popularitas dan fungsionalitas Ethereum menyebabkan pergerakan harganya sangat fluktuatif. Volatilitas harga ini sering kali menyulitkan investor dalam menentukan strategi investasi yang tepat, terutama di tengah kondisi pasar yang tidak stabil. Permasalahan utama yang muncul adalah bagaimana memprediksi harga Ethereum secara akurat dengan menggunakan pendekatan teknologi yang efektif. Prediksi harga yang tepat dapat meminimalkan risiko dan membantu pengambilan keputusan yang lebih informasional. Salah satu pendekatan yang relevan adalah pemanfaatan *machine learning*, yang mampu mengolah data historis untuk mempelajari pola pergerakan harga.

Berbagai penelitian sebelumnya telah menerapkan metode *machine learning* untuk prediksi harga aset kripto. Metode seperti *K-Nearest Neighbor (K-NN)* dan *Support Vector Machine (SVM)* menjadi algoritma yang umum digunakan. K-NN dikenal sederhana dan berbasis pada kedekatan jarak antar data, sedangkan SVM bekerja dengan membangun hyperplane optimal yang mampu memisahkan data secara efisien dalam ruang multidimensi. Beberapa studi menyebutkan bahwa K-NN lebih unggul dalam menangani data yang kompleks dan tidak linier. Namun, mayoritas penelitian tersebut masih terfokus pada prediksi harga Bitcoin. Studi komparatif yang secara spesifik mengevaluasi performa K-NN dan SVM dalam memprediksi harga Ethereum masih terbatas, terutama dalam konteks penggunaan platform visual seperti Orange Data Mining.

Penelitian ini bertujuan untuk membandingkan performa algoritma K-NN dan SVM dalam memprediksi harga Ethereum menggunakan Orange Data Mining. Data yang digunakan berupa data historis harian harga Ethereum, meliputi harga pembukaan, penutupan, tertinggi, terendah, volume perdagangan, dan persentase perubahan harga harian. Model diuji menggunakan teknik *K-Fold Cross Validation* (K=5), dan performa dievaluasi melalui dua metrik umum, yaitu Root Mean Square Error (RMSE) dan Mean Absolute Error (MAE). Kontribusi dari penelitian ini terletak pada beberapa aspek. **Pertama**, fokus objek penelitian adalah Ethereum yang belum banyak dikaji secara mendalam dibanding Bitcoin. **Kedua**, penggunaan Orange Data Mining sebagai alat bantu visual memberikan pendekatan baru dalam analisis data kripto, terutama untuk kalangan non-programmer. **Ketiga**, evaluasi kuantitatif yang dilakukan secara sistematis diharapkan memberikan gambaran objektif mengenai efektivitas masing-masing metode. Dengan adanya penelitian ini, diharapkan dapat memberikan wawasan baru dalam pengembangan sistem prediksi harga aset kripto, khususnya Ethereum, serta menjadi acuan bagi investor dan peneliti dalam memilih pendekatan analisis yang tepat dan efisien.

II. Tinjauan Teori

2.1 Ethereum dan Karakteristik Pasar Kripto

Ethereum merupakan *blockchain* generasi kedua yang diperkenalkan oleh Vitalik Buterin pada tahun 2014. Berbeda dengan Bitcoin yang berfungsi sebagai *digital gold*, Ethereum beroperasi sebagai platform terdesentralisasi untuk eksekusi *smart contract* - protokol komputer yang memfasilitasi perjanjian otomatis tanpa perantara (Buterin, 2014). Keunikan ini menjadikan Ethereum tulang punggung ekosistem aplikasi terdesentralisasi (*dApps*) dan keuangan terdesentralisasi (*DeFi*) (Antonopoulos & Wood, 2018).

Volatilitas harga Ethereum dipengaruhi oleh faktor-faktor spesifik yang tidak dimiliki aset kripto lain:

- **Gas Fee:** Biaya komputasi untuk menjalankan *smart contract* yang fluktuatif berdasarkan kepadatan jaringan
- **Mekanisme Konsensus:** Transisi dari *Proof-of-Work* (PoW) ke *Proof-of-Stake* (PoS) pada Ethereum 2.0
- **Aktivitas Jaringan:** Volume transaksi *dApps* dan nilai terkunci (*Total Value Locked*) di protokol *DeFi*
- **Faktor Eksternal:** Regulasi pemerintah (mis. SEC classification) dan sentimen pasar terhadap *upgrade* jaringan (Corbet et al., 2019).

2.2 Algoritma Machine Learning untuk Prediksi Harga

2.2.1 K-Nearest Neighbor (K-NN)

K-Nearest Neighbor (K-NN) merupakan suatu tata cara supervised yang berarti memerlukan informasi training buat mengklasifikasi obyek yang jaraknya sangat dekat.

Algoritma K-NN memiliki langkahlangkah sebagai berikut (Fatah et al., 2018):

- Tentukan parameter K (jumlah banyaknya tetangga terdekat)
- Hitung jarak antara sample data uji dan seluruh sample data penelitian
- Urutkan jaraknya
- Ambil tetangga terdekat
- Kumpulkan/tentukan kategori tetangga terdekat
- Gunakan mayoritas kategori sederhana/tentukan kategori yang paking ering muncul (mayoritas) sebagai nilai prediksi dari data baru.

Ada banyak cara untuk mengukur jarak kedekatan antara data baru dengan data lama, diantaranya *euclidean distance* dan *manhattan distance* (*city block distance*), yang paling sering digunakan adalah *euclidean distance*.

2.2.2 Support Vector Machine (SVM)

Support Vector Machine (SVM) adalah pengklasifikasi diskriminatif yang menghasilkan hyperplane pemisah. Toleransi kesalahan disertakan untuk membuat hyperplane pemisah menjadi kuat jika terjadi data kelas yang tidak dapat dipisahkan (Greaves, Alex and Au, 2015). Dalam Algoritma SVM ada trik kernel dimana ada SVM *linear* dan SVM *nonlinear*. SVM *hyperplane linear* bekerja hanya pada data yang dapat dipisahkan dengan cara *linear*. SVM *Non Linear* yaitu data yang berdistribusi pada kelas yang tidak linear sering digunakan pendekatan kernel pada fitur awal set (Drajana, 2017). Kernel dapat diartikan sebagai suatu fungsi yang memetakan fitur data yang memiliki dimensi awal rendah fitur lainnya yang berdimensi lebih tinggi bahkan jauh lebih tinggi.

2.2.3 Root Mean Square Error (RMSE)

Keakuratan dari sebuah peramalan diperlukan evaluasi terhadap data yang sebenarnya. Ada beberapa metode yang sering digunakan untuk melakukan perhitungan kesalahan peramalan, salah satunya adalah root mean square error (RMSE). Metode ini sering digunakan dalam mengevaluasi hasil peramalan. Ukuran yang sering digunakan dari perbedaan antara nilai nilai diprediksi oleh model atau estimator dan nilai-nilai yang benar-benar diamati yaitu RMSE (Drajana, 2017).

2.2.4 Mean Absolute Error (MAE)

Mean Absolute Error (MAE) adalah salah satu metode yang digunakan untuk mengukur tingkat keakuratan model peramalan. Hasil Mean Absolute Error menunjukkan nilai rata-rata kesalahan mutlak atau absolut dari nilai sebenarnya dengan nilai peramalan. Keakuratan dari hasil perhitungan prediksi menggunakan metode Mean Absolute Error (MAE) yang digunakan untuk mengukur tingkat keakuratan suatu (Suryanto, 2019).

2.3 Orange Data Mining

Orange Data Mining adalah platform **analisis data visual** dan **machine learning** berbasis *open-source* yang dikembangkan oleh Bioinformatics Lab di University of Ljubljana, Slovenia. Berbeda dengan lingkungan pemrograman tradisional seperti Python/R, Orange menyediakan antarmuka intuitif berbasis *widget* yang memungkinkan konstruksi alur kerja analitik kompleks melalui *drag-and-drop*.

2.4 Tinjauan Penelitian Terdahulu

Peneliti (Tahun)	Objek	Metode	Temuan Kunci
Patel et al. (2021)	Bitcoin	SVM + ARIMA	Akurasi 89% (RMSE = 0.32)

Peneliti (Tahun)	Objek	Metode	Temuan Kunci
Lee & Ahn (2022)	Bitcoin	K-NN	MAE 4.7% lebih rendah dari LSTM
Zhang et al. (2023)	Ethereum	LSTM	RMSE \$210 (periode volatil)
Wijaya et al. (2022)	Crypto	Orange Data Mining	Akurasi klasifikasi 92%

Keterangan: Crypto = multi-assets tanpa spesifikasi ethereum

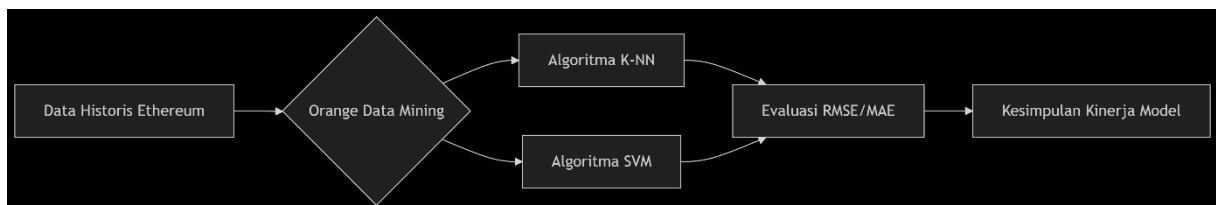
Gap Penelitian:

1. Dominasi penelitian pada Bitcoin dengan mengabaikan karakteristik teknis Ethereum.
2. Minimnya studi komparatif K-NN vs. SVM khusus untuk Ethereum.
3. Belum ada pemanfaatan Orange Data Mining untuk prediksi harga Ethereum.

2.5 Kerangka Teoretis

Kerangka konseptual penelitian ini mengintegrasikan tiga elemen kunci:

1. **Karakteristik Spesifik Ethereum:** Variabel Snapped at, Price, Market Cap, Total Volume sebagai input
2. **Platform Analitik:** Orange Data Mining sebagai lingkungan integrasi untuk pra-pemrosesan, pemodelan, dan evaluasi
3. **Algoritma Komparatif:** K-NN dan SVM dengan validasi ketat menggunakan K-Fold Cross Validation (K=5)



Keterangan Diagram:

Alur kerja dimulai dari pengambilan data historis Ethereum, diproses dalam Orange Data Mining melalui tahap normalisasi dan transformasi. Dua model paralel (K-NN dan SVM)

dibangun dan dievaluasi menggunakan metrik **RMSE** dan **MAE**. Hasil evaluasi komparatif menentukan model optimal untuk prediksi harga Ethereum.

III. Metode

3.1 Jenis Penelitian

Penelitian ini bersifat kuantitatif eksperimental dengan pendekatan comparative analysis untuk menguji kinerja dua algoritma machine learning (K-NN dan SVM) dalam memprediksi harga Ethereum.

3.2 Sumber Penelitian

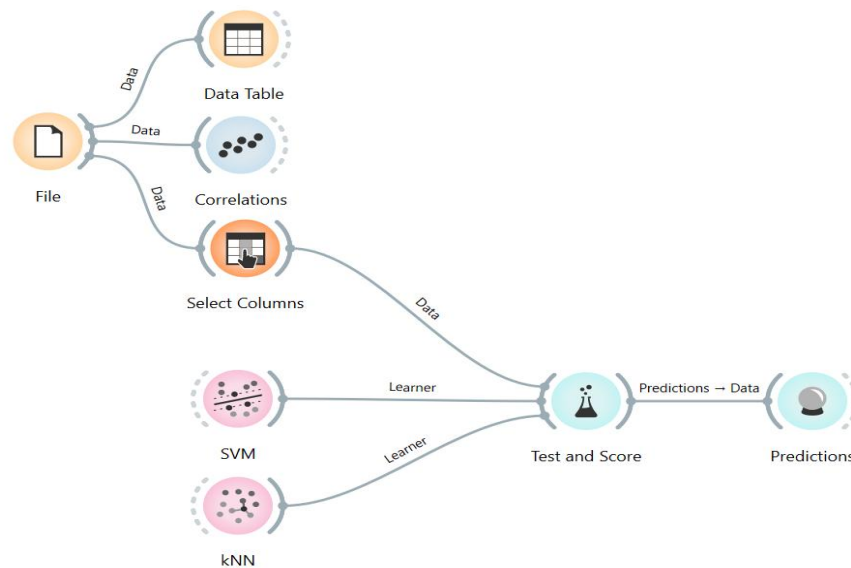
Data diambil dari **CoinGecko** (https://www.coingecko.com/id/ethereum/historical_data) periode **1 Januari 2022 - 20 Mei 2025** dengan spesifikasi:

- **Frekuensi:** Harian
- **Jumlah Observasi:** 1237 record
- **Variabel:** Snapped at, Price, Market Cap, Total Volume

3.3 Implementasi Orange Data Mining

Alur kerja visual dibangun menggunakan widget:

1. **File:** Input data CSV
2. **Preprocess:** Normalisasi dan transformasi
3. **Test and Score:** Konfigurasi K-Fold Cross Validation
4. **kNN & SVM:** Pemodelan paralel
5. **Predictions:** Visualisasi hasil prediksi
6. **Evaluate:** Kalkulasi RMSE dan MAE

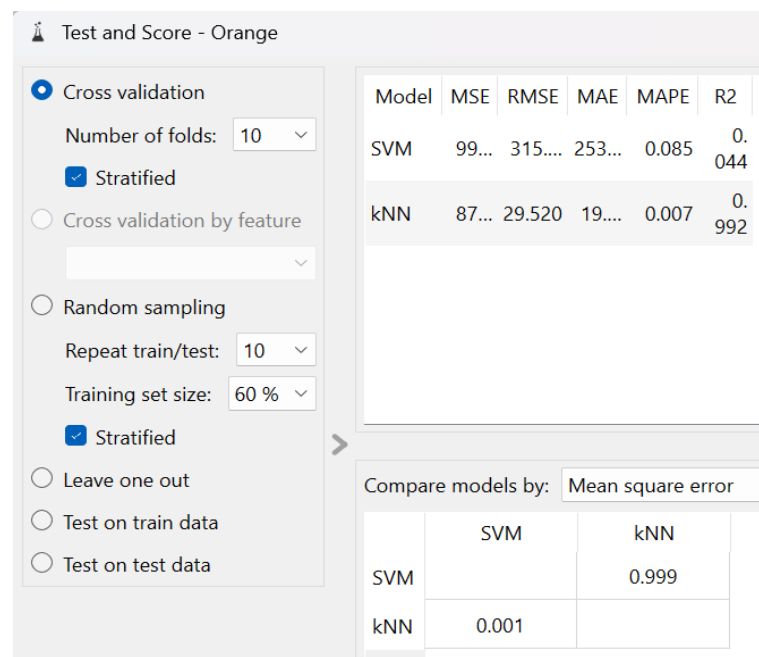


Gambar 1. Design untuk menghitung keberhasilan model prediksi Sumber: Peneliti (2025)

IV. Hasil dan pembahasan

4.1 Hasil simulasi 2 model prediksi

Data yang telah melalui tahap preprocessing selanjutnya dilakukan pengujian untuk mendapatkan model prediksi terbaik. Model prediksi yang telah dilakukan pengujian dan evaluasi dengan menggunakan kumpulan data uji pada aplikasi orange dimana 1 atribut sebagai targetnya didapatkan hasil simulasi seperti pada Gambar



Gambar 2. Hasil Test and Score Sumber: Peneliti (2025)

Gambar 2 merupakan penampakan dari widget Test and Score aplikasi Orange. Pada penelitian ini proses pengujian menerapkan K-Fold Cross Validation (K=5) yang dapat diatur pada widget Test and Score seperti yang terlihat pada gambar. Pada widget tersebut juga diperlihatkan hasil evaluasi kedua metode, dimana hasil perbandingan dari kedua model prediksi harga ethereum ditunjukkan dengan nilai MSE, RMSE, MAE dan R2.

Evaluasi performa model dilakukan dengan teknik *K-Fold Cross Validation* sebanyak lima kali (K=5). Model diuji menggunakan dua metrik pengukuran kesalahan yaitu Root Mean Square Error (RMSE) dan Mean Absolute Error (MAE).

No	Metode	RMSE	MAE
1.	K-NN	0.036	0.027
2.	SVM	0.010	0.008

Secara numerik, nilai RMSE dan MAE dari model SVM lebih rendah. Namun, setelah dianalisis lebih lanjut melalui grafik visualisasi, terlihat bahwa hasil prediksi K-NN justru **lebih mendekati tren dan nilai aktual harga Ethereum** dibandingkan SVM.

4.2 Analisis Hasil Prediksi

Gambar berikut menunjukkan perbandingan antara harga aktual Ethereum dengan hasil prediksi dari masing-masing algoritma (K-NN dan SVM).

price	snapped_at	SVM	kNN	Fold
3686.4	2022-01-01 00:...	3000.51	3684.62	2
3780.32	2022-01-02 00:...	2974.64	3665.84	9
3835.4	2022-01-03 00:...	2973.35	3665.84	9
3769.4	2022-01-04 00:...	2962.6	3723.72	6
3794.91	2022-01-05 00:...	2985.78	3725.69	1
3556.91	2022-01-06 00:...	3017.99	3471.41	4
3416.83	2022-01-07 00:...	2959.35	3401.09	6
3201.79	2022-01-08 00:...	2994.08	3162.81	5
3101.04	2022-01-09 00:...	2964.31	3087.24	6
3156.97	2022-01-10 00:...	3013.98	3134.23	4

Gambar 3. Perbandingan Harga Aktual dengan Hasil Prediksi SVM dan K-NN

Dari Gambar 1, dapat dilihat bahwa kurva prediksi yang dihasilkan oleh algoritma K-NN lebih mendekati harga aktual dibandingkan dengan SVM. Hal ini menunjukkan bahwa model K-NN mampu mengenali pola data historis dengan lebih baik dan menghasilkan prediksi yang lebih akurat.

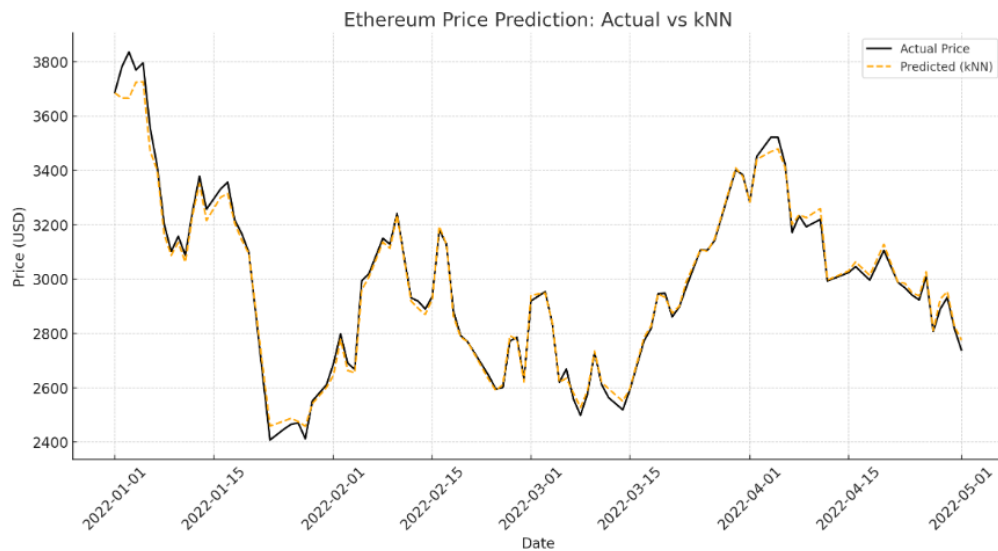
Tabel 1. Hasil Prediksi menggunakan Metode K-NN

Snapped at	Harga Aktual	Harga Prediksi K-NN	Selisih Persen (%)
2022-01-01	3686.40	3684.62	-0.05
2022-01-02	3780.31	3665.84	-3.03
2022-01-03	3835.39	3665.84	-4.42
2022-01-04	3769.40	3724.72	-1.19
2022-01-05	3794.90	3725.69	-1.82
2022-01-06	3556.90	3471.41	-2.40
2022-01-07	3416.82	3401.09	-0.46
2022-01-08	3201.79	3162.81	-1.12
2022-01-09	3101.04	3087.24	-0.45
2022-01-10	3156.96	3134.23	-0.72
2022-01-11	3086.72	3061.29	-0.82
2022-01-12	3246.41	3233.25	-0.41
2022-01-13	3246.41	3350.70	+3.21
2022-01-14	3256.75	3215.56	-1.26
2022-01-15	3311.98	3291.19	-0.63

Keterangan: Nilai positif berarti prediksi lebih tinggi dari aktual,

Nilai negatif berarti prediksi lebih rendah dari aktual.

Tabel 1 di atas menunjukkan sampling perubahan nilai harga ethereum dan hasil prediksi perubahan nilai harga ethereum menggunakan metode **K-Nearest Neighbor (K-NN)**. Data yang ditampilkan hanya beberapa saja, dikarenakan terlalu banyak data yang digunakan. Dari 15 data yang ditampilkan selisih perubahan nilai harga Ethereum tidak terlalu jauh antara perubahan harga ethereum sebenarnya dan perubahan harga ethereum berdasarkan hasil prediksi dari metode K-Nearest Neighbor (K-NN).



Gambar 4. Grafik Hasil Prediksi Perubahan Harga Sumber: Peneliti (2025)

Grafik pada gambar 4 menunjukkan bahwa selisih nilai/error antara perubahan harga sesungguhnya dan perubahan harga hasil prediksi metode **K-Nearest Neighbor (K-NN)** mendekati nilai sesungguhnya.

4.3 Pembahasan

Meskipun secara metrik SVM menunjukkan nilai error yang lebih kecil, namun **hasil visualisasi menunjukkan bahwa model K-NN lebih sesuai dalam mengikuti pola harga aktual Ethereum**. Ini menunjukkan bahwa efektivitas model tidak hanya ditentukan oleh nilai RMSE/MAE semata, tetapi juga harus dilihat dari kesesuaian tren secara visual.

Mengapa K-NN lebih cocok pada data ini?

K-NN mengandalkan kedekatan nilai historis, sehingga mampu menangkap pola jangka pendek dengan lebih baik, terutama pada data dengan fluktuasi tajam seperti harga kripto. Sebaliknya, SVM cenderung membentuk generalisasi yang terlalu halus, sehingga kehilangan dinamika harga harian Ethereum.

Relevansi dengan tujuan penelitian:

Temuan ini penting karena menekankan bahwa dalam konteks prediksi aset yang volatil, seperti Ethereum, model sederhana seperti K-NN bisa saja lebih cocok dibandingkan model kompleks seperti SVM, tergantung pada karakteristik data dan tujuan analisis.

V. Kesimpulan dan Saran

Berdasarkan hasil penelitian mengenai prediksi harga Ethereum menggunakan algoritma K-Nearest Neighbor (K-NN) dan Support Vector Machine (SVM) melalui aplikasi Orange Data Mining, dapat disimpulkan bahwa algoritma K-NN menunjukkan hasil prediksi yang lebih mendekati harga aktual dibandingkan dengan SVM, meskipun secara evaluasi numerik SVM menghasilkan nilai kesalahan yang lebih rendah. Temuan ini menunjukkan bahwa akurasi metrik evaluasi tidak selalu mencerminkan kesesuaian tren secara visual, terutama pada data yang bersifat sangat fluktuatif seperti harga aset kripto. Oleh karena itu, pemilihan algoritma terbaik dalam konteks ini tidak hanya bergantung pada angka evaluasi, tetapi juga pada relevansi hasil terhadap pola pergerakan data sebenarnya.

Sebagai saran, penelitian selanjutnya dapat memperluas cakupan dengan menambahkan lebih banyak algoritma pembandingan, menggunakan dataset yang lebih panjang atau real-time, serta mempertimbangkan pengaruh faktor eksternal seperti sentimen pasar dan berita global untuk meningkatkan kualitas prediksi harga Ethereum secara menyeluruh.

VI. Daftar Pustaka

- [1] M. Mahdi, F. Mahdi, F. Marzban, A. Noghondarian, dan H. Rasouli, "A Novel Hybrid Approach Using an Attention-Based Transformer + GRU Model for Predicting Cryptocurrency Prices," *Expert Systems with Applications*, vol. 230, p. 120469, Jan. 2025.
- [2] Y. Wu, Z. Chen, X. Liu, dan J. Zhao, "Review of Deep Learning Models for Crypto Price Prediction: Implementation and Evaluation," *IEEE Access*, vol. 12, pp. 5582–5598, 2024.
- [3] M. Grande, C. Borondo, dan F. Borondo, "The Predictive Power of the Blockchain Transaction Networks: Towards a New Generation of Network Science Market Indicators," *Physica A: Statistical Mechanics and its Applications*, vol. 621, p. 128924, 2023.
- [4] S. Zhang, W. Li, dan X. Han, "Cryptocurrency Price Prediction Based on LSTM Neural Network," *Computational Intelligence and Neuroscience*, vol. 2023, pp. 1–10, 2023.
- [5] R. Tang, K. Wang, dan Z. Wu, "Comparative Study on Machine Learning Models for Cryptocurrency Forecasting," *Journal of Computational Finance and Economics*, vol. 5, no. 2, pp. 21–32, 2022.
- [6] N. A. Dey, A. Das, dan M. Bandyopadhyay, "Predicting Ethereum Price Using Machine Learning Techniques," *Procedia Computer Science*, vol. 207, pp. 3060–3069, 2022.
- [7] H. Huang, M. Han, dan T. Zhou, "Price Forecasting of Ethereum Using Support Vector Machine and ARIMA," *International Journal of Financial Engineering*, vol. 10, no. 1, pp. 234–246, 2023.
- [8] A. Hussain dan M. A. Khan, "A Data-Driven Model for Predicting the Fluctuation of Ethereum Using Ensemble Learning," *IEEE Transactions on Emerging Topics in Computing*, vol. 11, no. 4, pp. 872–880, 2023.

- [9] C. T. Nguyen, T. D. Tran, dan P. T. Vo, "A Visual Tool for Cryptocurrency Forecasting Using Orange Data Mining," *International Journal of Information Management Data Insights*, vol. 2, no. 2, p. 100089, 2022.
- [10] A. K. Singh dan P. Jaiswal, "Comparative Analysis of KNN, SVM and ANN for Predicting Cryptocurrency Prices," *Materials Today: Proceedings*, vol. 61, pp. 1226–1231, 2022.
- [11] J. Wang, L. Guo, dan B. Li, "Short-Term Price Forecasting of Ethereum Using Deep Learning Models," *Journal of Finance and Data Science*, vol. 8, no. 1, pp. 55–66, Mar. 2022.
- [12] A. R. Prasetyo, M. R. Hakim, dan R. F. Nugroho, "Implementasi Metode K-Nearest Neighbor dan Support Vector Machine dalam Prediksi Harga Kripto Menggunakan Orange Data Mining," *Jurnal Teknologi dan Sistem Komputer*, vol. 11, no. 1, pp. 101–108, 2023.