

Prediksi Harga Emas Berdasarkan Finansial Global Dengan Menggunakan Algoritma Gradien Boosting Regression

Yennimar¹, Gabriella Br. Surbakti²

^{1,2}Universitas Prima Indonesia

yennimar@unprimdn.ac.id¹ gabriellabrsurbakti@gmail.com²

Abstrak. Penelitian ini bertujuan untuk mengembangkan model prediksi harga emas yang akurat dan dapat diandalkan dengan memanfaatkan algoritma Gradient Boosting Regression. Model ini dilatih menggunakan data historis harian dari tahun 2010 hingga 2024, yang mencakup beragam variabel ekonomi makro dan mikro yang secara empiris diketahui memiliki pengaruh signifikan terhadap pergerakan harga emas, seperti indeks pasar saham utama, harga komoditas energi dan logam mulia lainnya, tingkat inflasi, suku bunga acuan, dan nilai tukar mata uang terhadap dolar AS. Melalui proses pra-pemrosesan data yang cermat, pemilihan fitur yang relevan, dan optimasi parameter model, diperoleh model prediksi yang mampu menangkap pola kompleks dalam data dan menghasilkan perkiraan harga emas dengan tingkat akurasi yang tinggi. Hasil evaluasi model menunjukkan bahwa Gradient Boosting Regression mampu mengungguli metode prediksi tradisional, dengan nilai Root Mean Squared Error (RMSE) sebesar 0,323 dan Mean Absolute Percentage Error (MAPE) sebesar 0,002 sehingga model ini dapat mencapai akurasi 99,96%. Analisis sensitivitas terhadap perubahan parameter model juga dilakukan untuk menguji robustitas model. Model prediksi yang dikembangkan dalam penelitian ini memiliki potensi besar untuk dimanfaatkan oleh investor, analis pasar, dan pembuat kebijakan sebagai alat bantu dalam pengambilan keputusan investasi yang lebih baik. Namun, perlu diingat bahwa model ini memiliki keterbatasan, seperti asumsi stasioneritas data dan potensi perubahan struktur pasar yang dapat mempengaruhi kinerja model di masa depan. Oleh karena itu, perlu dilakukan pembaruan model secara berkala untuk memastikan keakuratan prediksi. Penelitian ini juga menyarankan perlunya eksplorasi lebih lanjut mengenai pengaruh peristiwa global yang tidak terduga, seperti pandemi atau konflik geopolitik, terhadap pergerakan harga emas.

Keyword: Prediksi, harga emas, gradien boosting regression

1. Pendahuluan

Emas, sebagai aset safe-haven yang telah teruji waktu, telah menjadi pilar penting dalam portofolio investasi banyak investor. Pasokan dan nilai intrinsiknya yang terbatas membuatnya kurang rentan terhadap fluktuasi yang berdampak pada mata uang fiat dan aset lainnya. Hal ini dipandang sebagai lindung nilai terhadap ketidakpastian, memberikan stabilitas ketika aset tradisional (seperti saham dan obligasi) berkinerja buruk atau berisiko[1].

Namun, memprediksi pergerakan harga emas bukanlah tugas yang mudah. Fluktuasi harga emas dipengaruhi oleh berbagai faktor kompleks yang saling terkait, mulai dari kebijakan moneter bank sentral hingga sentimen investor. Dalam kondisi pasar yang semakin volatil, kemampuan untuk memprediksi pergerakan harga emas menjadi semakin krusial bagi investor untuk membuat keputusan

investasi yang lebih baik dan mengelola risiko dengan lebih efektif[2]. Penelitian ini bertujuan untuk mengatasi tantangan ini dengan mengembangkan model prediksi yang lebih akurat dan handal.

Bank-bank sentral memiliki cadangan emas dalam jumlah besar sebagai bagian dari cadangan devisa mereka[3], yang mencerminkan pentingnya logam ini dalam sistem keuangan global sejak lama. Memahami dan memprediksi pergerakan harga emas menjadi sangat krusial dalam pengambilan keputusan investasi. Investor, pelaku bisnis, dan pembuat kebijakan membutuhkan alat analisis yang handal untuk mengantisipasi perubahan pasar dan merumuskan strategi yang efektif[4]. Penelitian ini hadir dengan tujuan untuk menjawab tantangan tersebut dengan menawarkan pendekatan yang lebih inovatif dalam memprediksi harga emas.

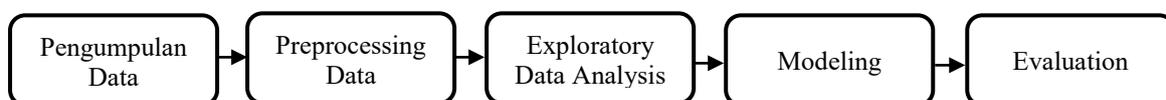
Permasalahan yang dialami penelitian sebelumnya mulai dari kompleksitas data dan variable yang besar, interaksi antar variable yang tidak dapat ditangkap dengan baik dan juga data yang bervariasi. Kualitas data yang noise dan terdapat banyak missing values. Permasalahan lain yang ditemukan juga terdapat di hasil pengolahan data yang dilakukan dimana terdapat overfitting[2], [5]–[9]. Dalam mengatasi permasalahan tersebut pada penelitian ini, peneliti melakukan pra-pemrosesan data yang cermat, melakukan regulasi dalam mencegah overfitting, dan menggabungkan beberapa model machine learning untuk meningkatkan akurasi.

Dalam era big data dan kecerdasan buatan, ilmu data telah membuka peluang baru dalam analisis data keuangan. Penelitian ini memanfaatkan kemajuan dalam bidang ilmu data dengan mengadopsi algoritma gradient boosting regression. Algoritma ini telah terbukti unggul dalam menangani masalah prediksi non-linear yang kompleks, seperti yang sering terjadi pada data pasar keuangan. Dengan menggabungkan kekuatan algoritma gradient boosting regression dengan data historis harga emas dan berbagai indikator ekonomi global, diharapkan model prediksi yang dihasilkan dapat memberikan akurasi yang lebih tinggi dibandingkan dengan model-model konvensional.

2. Metodologi Penelitian

2.1. Tahapan Penelitian

Studi ini bertujuan untuk membangun model prediksi yang tidak hanya mampu menangkap tren jangka panjang, tetapi juga merespons perubahan pasar secara real-time. melakukan analisis terhadap pengaruh algoritma yang digunakan dalam meningkatkan hasil akurasi yang didapatkan dengan mengidentifikasi variabel-variabel kunci yang memiliki dampak paling signifikan terhadap pergerakan harga emas, sehingga dapat memberikan pemahaman yang lebih mendalam mengenai dinamika pasar emas. Penelitian ini juga bertujuan untuk menambahkan wawasan baru pada bidang penelitian prediksi harga komoditas dengan mengaplikasikan pendekatan machine learning yang mutakhir.



Gambar 1. Flowchart Metode Penelitian

2.1.1. Pengumpulan Data

Dataset yang digunakan dalam penelitian ini di ambil dari data Gold Price Regression yang diperoleh dari Kaggle. Data tersebut merupakan kumpulan data deret waktu dengan informasi keuangan untuk beberapa indeks pasar, komoditas, indikator ekonomi, dan nilai tukar mata uang. Indeks pasar dan komoditas diwakili melalui dana yang diperdagangkan di bursa masing-masing. Ini mencakup nilai-nilai dari 14 Oktober 2010 hingga 23 Oktober 2024[10]. Kemudian data tersebut dipisahkan dan dikelompokkan berdasarkan data emas untuk digunakan dalam penelitian ini.

2.1.2. Preprocessing Data

Pada tahapan ini, dilakukan preprocessing data yang bertujuan untuk mengorganisir, mengubah, dan memperbaiki data yang hilang. Hal ini bertujuan untuk mempermudah proses pengolahan data yang dilakukan pada tahap berikutnya.

date	gold open	gold high	gold low	gold volume	gold close	gold open	gold high	gold low	gold volume	gold close	gold open	gold high	gold low	gold volume	gold close
2020-01-14	114.00	114.50	114.00	114.00	114.00	114.00	114.50	114.00	114.00	114.00	114.00	114.50	114.00	114.00	114.00
2020-01-15	114.75	114.84	114.26	114.84	114.75	114.75	114.84	114.26	114.84	114.75	114.75	114.84	114.26	114.84	114.75
2020-01-16	114.00	114.00	114.00	114.00	114.00	114.00	114.00	114.00	114.00	114.00	114.00	114.00	114.00	114.00	114.00
2020-01-17	113.00	113.00	113.00	113.00	113.00	113.00	113.00	113.00	113.00	113.00	113.00	113.00	113.00	113.00	113.00
2020-01-20	114.25	114.25	114.25	114.25	114.25	114.25	114.25	114.25	114.25	114.25	114.25	114.25	114.25	114.25	114.25

Gambar 2. Showing Data

date	gold open	gold high	gold low	gold volume	gold close	date	gold open	gold high	gold low	gold volume	gold close
2020-01-14	114.00	114.50	114.00	114.00	114.00	2020-01-14	114.00	114.50	114.00	114.00	114.00
2020-01-15	114.75	114.84	114.26	114.84	114.75	2020-01-15	114.75	114.84	114.26	114.84	114.75
2020-01-16	114.00	114.00	114.00	114.00	114.00	2020-01-16	114.00	114.00	114.00	114.00	114.00
2020-01-17	113.00	113.00	113.00	113.00	113.00	2020-01-17	113.00	113.00	113.00	113.00	113.00
2020-01-20	114.25	114.25	114.25	114.25	114.25	2020-01-20	114.25	114.25	114.25	114.25	114.25

Gambar 3. Data setelah di Cleaning

Preprocessing data adalah langkah krusial dalam analisis data yang bertujuan untuk mempersiapkan data mentah agar dapat diolah dan dianalisis secara efektif. Data mentah seringkali mengandung ketidakkonsistenan, kesalahan, atau format yang tidak sesuai untuk langsung diproses oleh algoritma. Melalui preprocessing, data dapat dibersihkan, diubah, dan diintegrasikan sehingga menghasilkan informasi yang lebih akurat dan relevan.

Pada gambar 2 dilakukan pembersihan data. Kualitas model machine learning sangat bergantung pada kualitas data. Oleh karena itu, data cleaning adalah langkah krusial sebelum memulai pelatihan model. Dengan membersihkan data, kita dapat meningkatkan kinerja model machine learning dan mengurangi risiko mendapatkan hasil yang bias atau tidak akurat.

2.1.3. Exploratory Data Analysis

Melakukan Exploratory Data Analysis (EDA) menjadi tahap awal yang sangat krusial dalam melakukan analisis data. Hal ini dilakukan untuk melihat bagaimana data disusun, mengidentifikasi data dan melakukan pengelompokan data. Pada penelitian ini, data yang banyak tersebut dipisahkan dan diambil data gold (gold open, gold close, gold high, gold low, gold volume).

	gold open	gold low	gold high	gold close	gold volume
0	111.51	110.79	112.37	112.03	18305238.0
1	111.35	110.38	112.01	110.86	18000724.0
2	NaN	NaN	NaN	NaN	NaN
3	110.95	110.83	111.75	111.52	10467927.0
4	109.97	108.46	110.05	108.94	17534231.0
...
3899	247.75	247.62	249.06	248.63	5176170.0
3900	250.00	249.90	251.37	251.27	7833614.0
3901	252.74	250.73	253.14	251.22	9258590.0
3902	253.06	252.52	253.94	253.93	5756321.0
3903	253.08	250.20	253.18	250.87	7899995.0

3904 rows x 5 columns

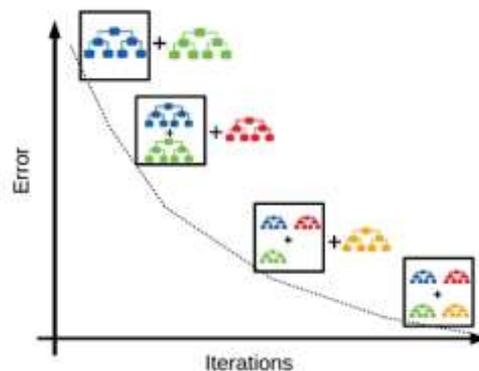
	gold open	gold low	gold high	gold close	gold volume
count	3719.000000	3719.000000	3719.000000	3719.000000	3.719000e+03
mean	145.454975	144.761329	146.101477	145.453861	9.658138e+06
std	29.604259	29.422517	29.754711	29.611348	6.182342e+06
min	100.920000	100.232600	100.990000	100.500000	1.436508e+06
25%	120.555000	120.170000	121.030000	120.590000	5.795310e+06
50%	137.620000	136.990000	138.140000	137.710000	8.087993e+06
75%	167.770000	167.055000	168.412500	167.835000	1.156729e+07
max	253.080000	252.520000	253.940000	253.930000	9.369811e+07

Gambar 4. Pengelompokan data emas

2.1.4. Modeling

Pada penelitian ini, peneliti menggunakan algoritma Gradient Boosting Regression untuk melakukan prediksi harga emas berdasarkan data yang dimiliki. Keunggulan utama dari algoritma ini adalah dapat menggabungkan banyak model sederhana untuk menciptakan model yang lebih kompleks dan akurat[11].

Model ini memiliki akurasi tinggi dalam menganalisis pola harga emas yang memiliki pola non-linear yang kompleks. Kemampuannya yang dapat menggabungkan banyak model membuatnya bisa mengurangi overfitting dan underfitting, sehingga dapat menghasilkan model yang lebih baik.



2.1.5. Model Testing dan Evaluation

Pada tahap ini dilakukan uji dari model yang dibuat menggunakan algoritma Gradient Boosting yang dilatih. Dengan pengujian ini, kita dapat melakukan evaluasi dari hasil pengujian model yang

didapatkan. Kemudian model tersebut di uji coba dengan mencoba memprediksi harga emas selama 30 hari kedepan terhitung dari setelah tanggal terakhir pada rangkuman model tersebut.

2.2. Mean Absolute Percentage Error (MAPE)

Mean Absolute Percentage Error (MAPE) merupakan metrik evaluasi yang lazim digunakan untuk mengukur tingkat ketepatan suatu model prediksi. Berbeda dengan Mean Absolute Error (MAE) yang mengevaluasi kesalahan prediksi dalam satuan yang identik dengan data observasi, MAPE mengekspresikan kesalahan prediksi sebagai persentase dari nilai aktual. Hal ini memungkinkan interpretasi yang lebih intuitif, khususnya ketika membandingkan kinerja model pada berbagai dataset dengan skala yang beragam[12].

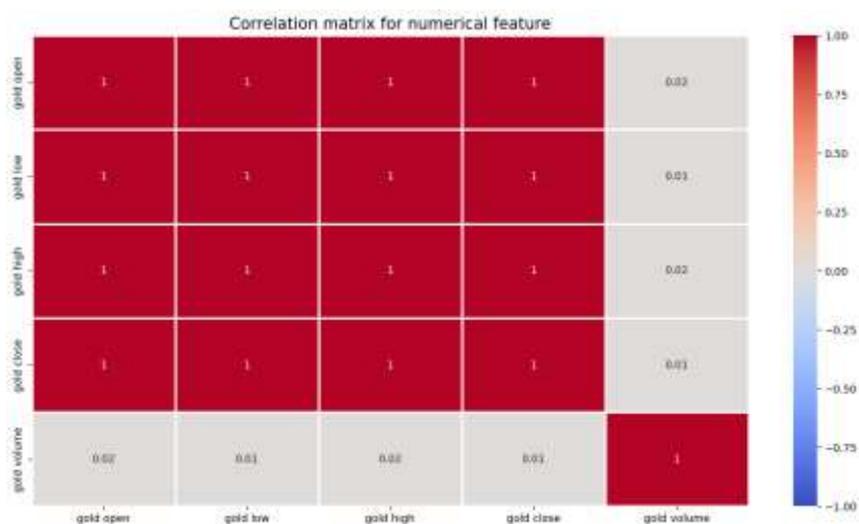
2.3. Mean Squared Error (MSE)

Mean Squared Error (MSE) merupakan suatu metrik statistik yang umum digunakan untuk mengevaluasi tingkat ketepatan suatu model prediksi. MSE menghitung rata-rata kuadrat dari deviasi antara nilai prediksi yang dihasilkan oleh model dengan nilai aktual yang telah diketahui. Dengan demikian, MSE memberikan bobot yang lebih signifikan terhadap kesalahan prediksi yang besar. Nilai MSE yang lebih rendah mengindikasikan kinerja model yang lebih baik dalam membuat prediksi. MSE secara luas diaplikasikan dalam berbagai disiplin ilmu seperti statistika, pembelajaran mesin, dan ilmu data untuk menilai performa model[13].

3. Hasil dan Pembahasan

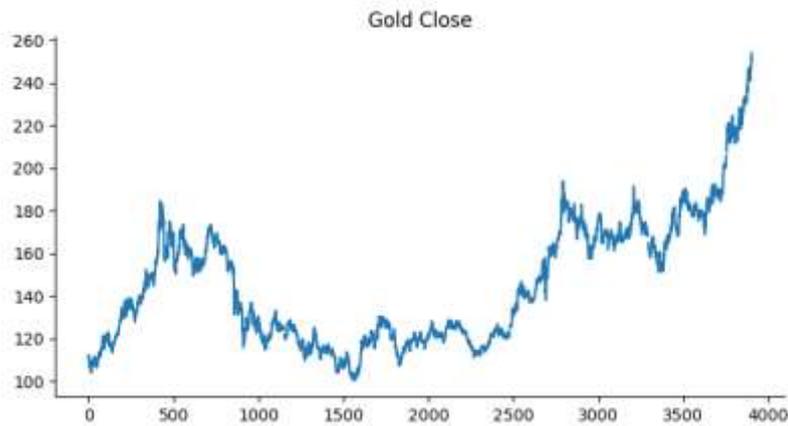
3.1. Hasil Analisis

Setelah dilakukannya pengelompokkan data emas yang dipisahkan dari data utama yang digunakan. Kemudian kita melakukan cek korelasi untuk memahami nilai dan hubungan antar indikator. Selanjutnya melakukan inisiasi prediksi dan menghitung serta memprediksi nilai residual.



Gambar 5. Korelasi antar indikator

Setelah memahami hubungan antar indikator, kemudian melakukan prediki terhadap nilai gold close sebagai faktor y dan gold open, gold low, gold high, dan gold volume sebagai faktor x. hal ini dikarenakan, gold close merupakan nilai akhir emas tiap periode.



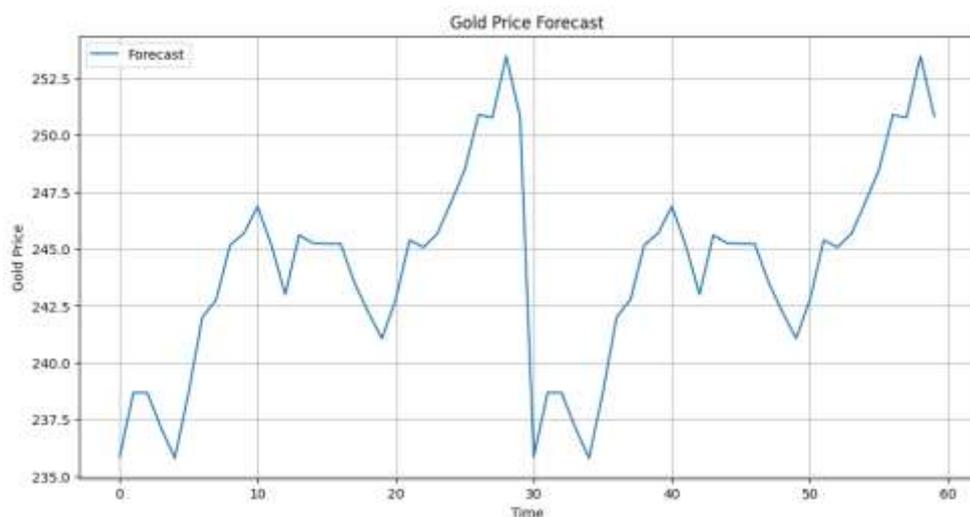
Gambar 6. Gold Close

Kemudian dilakukan splitting data dan lalu hasilnya dimasukkan dalam train model dimana menggunakan algoritma gradien boosting. Hasil yang didapatkan menunjukkan gradien boosting mendapatkan score sebesar 99,96%. Setelahnya mencari nilai MAPE dan MSE dan mendapatkan 0,323 untuk MAPE dan 0,002 untuk MSE. Dari hasil tersebut, menunjukkan model yang dilatih sangat baik untuk melakukan prediksi harga emas.

3.2. Uji Coba Model

Uji coba model dilakukan dengan cara meramalkan harga emas dalam 60 hari kedepan dengan menggunakan model yang sudah dilatih. Pada hari ke 50, prediksi harga emas jatuh pada 242,76 ETF. Pada hari ke 57, harga emas melonjak naik ke angka 250,76 ETF.

Forecast	
50	242.767973
51	245.377344
52	245.070225
53	245.658207
54	247.023457
55	248.474906
56	250.889653
57	250.766141
58	253.467484
59	250.827645



Gambar 7. Prediksi harga emas

Daftar Pustaka

- [1] F. Allen, X. Gu, C. W. Li, J. “QJ” Qian, and Y. Qian, “Implicit guarantees and the rise of shadow banking: The case of trust products,” *J. financ. econ.*, vol. 149, no. 2, pp. 115–141, 2023, doi: 10.1016/j.jfineco.2023.04.012.

- [2] R. K. Akmal, "Tinjauan Sistematis Untuk Merekomendasi Prediksi Harga Emas," *Maret*, vol. 2022, no. 7, pp. 18–24, 2022.
- [3] Bareksa, "Bank Sentral Terus Borong Emas Meskipun Harga Naik, Meningkatkan 2 Kali Lipat Jadi 37 Ton pada Juli," 2024. <https://www.bareksa.com/berita/emas/2024-09-09/bank-sentral-terus-borong-emas-meskipun-harga-naik-meningkat-2-kali-lipat-jadi-37-ton-pada-juli>
- [4] M. Adiningsih, D. Yulianti, and A. Wibowo, "Analisis Penerapan Forecasting dalam Menentukan Kebutuhan Storage pada Pusat Data Kementerian Keuangan," *JRST (Jurnal Ris. Sains dan Teknol.)*, vol. 7, no. 2, p. 229, 2023, doi: 10.30595/jrst.v7i2.17579.
- [5] A. Aulia, B. Aprianti, Y. Supriyanto, and C. Rozikin, "Prediksi Harga Emas dengan Menggunakan Algoritma Support Vector Regression (Svr) dan Linear Regression," *J. Ilm. Wahana Pendidik.*, vol. 8, no. 5, pp. 84–88, 2022, doi: 10.5281/zenodo.6408864.
- [6] M. Hadi, A. M., Witanti, W., Melina, "Hal. 486-496 Vol. 6; No. 3," *Prediksi Pergerakan Harga Emas Menggunakan Metod. Genet. Support Vector Regres.*, vol. 6, pp. 486–496, 2024.
- [7] Hafid Akbar Fikri, "Prediksi Harga Emas Dengan Algoritma Backpropagation," *J. Sains Komput. Inform.*, vol. 7, no. 1, pp. 182–189, 2023.
- [8] M. Muharrom, "Analisis Komparasi Algoritma Data Mining Naive Bayes, K-Nearest Neighbors dan Regresi Linier Dalam Prediksi Harga Emas," *Bull. Inf. Technol.*, vol. 4, no. 4, pp. 430–438, 2023, doi: 10.47065/bit.v4i4.986.
- [9] A. Syahputra, "Implementasi Metode Double Exponential Smoothing Pada Aplikasi Prediksi Harga Emas Dunia," *J. Teknol. Terap. Sains 4.0*, vol. 1, no. 1, 2020, doi: 10.29103/tts.v1i1.3250.
- [10] Franciscogcc, "Gold Price Regression," *Kaggle*, 2024. <https://www.kaggle.com/datasets/franciscogcc/financial-data/data> (accessed Nov. 20, 2024).
- [11] KANTINIT, "Gradient Boosting: Pengertian, Cara Kerja dan Contoh Skripsi," 2023. <https://kantinit.com/kecerdasan-buatan/gradient-boosting-pengertian-cara-kerja-dan-contoh-skripsi/>
- [12] I. Nabillah and I. Ranggadara, "Mean Absolute Percentage Error untuk Evaluasi Hasil Prediksi Komoditas Laut," *JOINS (Journal Inf. Syst.)*, vol. 5, no. 2, pp. 250–255, 2020, doi: 10.33633/joins.v5i2.3900.
- [13] H. Nuha, "Mean Squared Error (MSE) dan Penggunaannya," *Papers.Ssrn.Com*, vol. 52, pp. 1–1, 2023, [Online]. Available: <https://ssrn.com/abstract=4420880>