

PENERAPAN METODE RANDOM FOREST REGRESSION UNTUK PREDIKSI HARGA EMAS DI INDONESIA

Willy Calvin Candra Lay, Putu Manik Prihatini, Made Pasek Agus Ariawan

Teknologi Informasi, Politeknik Negeri Bali
Kampus Bukit, Jimbaran, Kuta Selatan, Badung, Indonesia
willycandra850@gmail.com

ABSTRAK

Harga emas bersifat *non-linear* dan dipengaruhi oleh beberapa faktor eksternal, sehingga menyulitkan masyarakat dalam menentukan waktu yang tepat untuk berinvestasi. Tanpa alat prediksi yang memadai, keputusan investasi berisiko untuk diambil pada waktu yang tidak tepat dan berpotensi menimbulkan kerugian finansial. Penelitian ini bertujuan mengevaluasi kinerja model *Random Forest Regression* dalam memprediksi harga emas di Indonesia menggunakan data historis harga emas harian dalam satuan Rupiah per gram. Model tidak memprediksi harga secara langsung, melainkan memprediksi persentase perubahan harga (*pct_change*) untuk mengatasi *distribusi shift* akibat tren harga yang kuat pada data *time-series*. Prediksi akan dilakukan tanpa ada tambahan variabel eksternal. Optimasi *hyperparameter* dilakukan menggunakan *GridSearch* dengan skema *Weighted TimeSeriesSplit* 7-fold terhadap 72 kombinasi parameter, menghasilkan parameter terbaik pada penelitian ini yaitu *n_estimators* = 200, *max_depth* = 5, *max_features* = *sqrt* dan *min_sample_leaf* = 8. Prediksi jangka menengah dilakukan menggunakan metode *Recursive Forecasting*. Hasil evaluasi pada data uji menunjukkan MAE sebesar Rp20.105, RMSE Rp34.988, dan MAPE sebesar 0,9081%, membuktikan bahwa pola historis harga emas sudah cukup memadai sebagai dasar prediksi yang akurat tanpa variabel eksternal.

Kata kunci : *random forest regression, prediksi harga emas, recursive forecasting, machine learning, data mining, non-linear*

1. PENDAHULUAN

Investasi merupakan suatu bentuk penanaman modal yang dijadikan tabungan dari masa sekarang hingga masa yang akan datang untuk menjaga stabilitas ekonomi [1]. Diantara berbagai pilihan instrumen yang tersedia, emas menempati posisi yang istimewa. Emas telah dipercaya sebagai *safe haven*, aset yang nilainya cenderung bertahan bahkan menguat justru ketika kondisi ekonomi sedang tidak menentu. Karena hal ini, banyak masyarakat memilih emas untuk meredam risiko portofolio saat pasar bergejolak [2].

Salah satu alasan kenapa emas menjadi pilihan favorit adalah ketahanannya terhadap inflasi [3]. Namun, harga emas tidak bersifat statis. Nilai tukar mata uang, tingkat suku bunga, serta kondisi ekonomi dan geopolitik global semuanya berkontribusi pada fluktuasi harga emas yang dapat terjadi secara signifikan baik dalam jangka pendek maupun jangka panjang [4].

Fluktuasi yang tidak menentu ini menjadi permasalahan nyata bagi masyarakat, karena tanpa informasi prediktif yang memadai, keputusan investasi berisiko diambil pada waktu yang tidak tepat sehingga berpotensi menimbulkan kerugian finansial [5]. Persoalan ini semakin kompleks karena pergerakan harga emas bersifat *non-linear*, sehingga pendekatan statistik konvensional seperti regresi *linear* yang mengasumsikan hubungan *linear* antar variabel dinilai kurang andal untuk menangkap pola tersebut [6].

Beberapa penelitian sebelumnya telah dilakukan untuk menjawab tantangan tersebut. Penelitian yang dilakukan oleh Agusmawati dkk. menguji *Long Short*

Term Memory (LSTM) dan *Gated Recurrent Unit* (GRU) menggunakan data historis harga emas mendapatkan hasil *Mean Absolute Percentage Error* (MAPE) 5.2047% pada LSTM, namun kedua model ini membutuhkan sumber daya komputasi yang besar [7]. Gananta dkk. menerapkan *Support Vector Regression* (SVR) dengan optimasi *GridSearch* untuk prediksi harga emas, dimana tercatat hasil *Mean Absolute Percentage Error* (MAPE) sebesar 3,73%, namun optimasi parameter SVR terbukti membutuhkan proses yang kompleks dan sensitif terhadap pemilihan nilai parameter [5]. Dari kajian tersebut, meskipun metode yang ada mampu menghasilkan prediksi yang akurat, masing-masing memiliki keterbatasan tersendiri. Di sisi lain, *Random Forest Regression* belum banyak dievaluasi secara spesifik untuk prediksi harga emas di Indonesia, padahal metode ini dinilai mampu menangani pola *non-linear* dan lebih *robust* terhadap *overfitting* [8].

Berdasarkan permasalahannya dan *research gap* tersebut, penelitian ini bertujuan mengevaluasi kinerja model *Random Forest Regression* sebagai alternatif yang lebih efisien dalam memprediksi harga emas harian di Indonesia dalam satuan Rupiah (IDR) per gram. Untuk mencapai tujuan tersebut, penelitian ini menggunakan data historis harga emas, pendekatan prediksi berbasis persentase perubahan harga (*pct_change*) guna mengatasi ketidakstabilan distribusi pada data *time-series* yang memiliki tren naik kuat. Proses optimasi model dilakukan menggunakan *GridSearch* dengan skema *Weighted TimeSeriesSplit* 7-fold, dan prediksi jangka menengah dihasilkan melalui metode *Recursive Forecasting*.

Kinerja model diukur menggunakan tiga metrik yaitu *Mean Absolute Error* (MAE), *Root Mean Square Error* (RMSE), dan *Mean Absolute Percentage Error* (MAPE). Dengan harapan penelitian ini dapat menjadi referensi bagian pengembangan sistem pendukung keputusan investasi emas selanjutnya.

2. TINJAUAN PUSTAKA

2.1. Penelitian Pendahuluan

Beberapa penelitian terdahulu telah dilakukan terkait prediksi harga emas menggunakan berbagai pendekatan *machine learning*. Aulia dkk. membandingkan SVR dan *Linear Regression* menggunakan data harian [Investing.com](https://www.investing.com) dan menentukan *Linear Regression* memiliki MSE lebih rendah 4.04444791059 dibandingkan SVR 7.524505784357 [1].

Agusmawati dkk. menguji *Long Short Term Memory* (LSTM) dan *Gated Recurrent Unit* (GRU) pada data historis harga emas, di mana LSTM menghasilkan MAPE sebesar 5.2047% sedangkan GRU menghasilkan MAPE 6.0688% [7].

Gananta dkk. mengoptimasi SVR menggunakan *GridSearch* dan mencapai MAPE 3,73%, meski proses optimasi parameternya terbukti kompleks [5].

Afifa dkk. membandingkan ARIMA dan *Random Forest*, dengan *Random Forest* unggul pada MAPE 2,25% dibanding ARIMA 4.26% [9]. Ringkasan penelitian pendahuluan dapat dilihat pada Tabel 1 berikut.

Tabel 1. Ringkasan penelitian pendahuluan

No	Peneliti	Metode	Hasil
1	Aulia dkk. (2022)	SVR, LR	SVR MSE: 7.524505784357 LR MSE: 4.04444791059
2	Agusmawati dkk. (2023)	LSTM, GRU	MAPE GRU: 6.0688% LSTM: 5.2047%
3	Gananta dkk. (2023)	SVR + Grid Search	MAPE 3,73%
4	Afifa dkk. (2024)	ARIMA, RF	MAPE ARIMA: 4.26% RF: 2,25%

Berdasarkan Tabel 1, terlihat bahwa penelitian terdahulu menggunakan berbagai metode dengan hasil yang bervariasi. Metode *deep learning* seperti LSTM menghasilkan akurasi yang cukup baik namun dengan kebutuhan komputasi yang tinggi, sedangkan SVR memerlukan optimasi parameter yang kompleks. *Random Forest* pada penelitian Afifa dkk. menunjukkan performa yang kompetitif dengan MAPE 2.25% meski menggunakan variabel eksternal. Penelitian ini melengkapi *gap* tersebut dengan mengevaluasi *Random Forest Regression* secara khusus menggunakan data historis harga emas tanpa variabel eksternal.

2.2. Emas

Emas merupakan jenis logam mulia yang memiliki nilai intrinsik yang relatif stabil dan tidak mengalami penurunan nilai secara signifikan dari waktu ke waktu. Emas dipandang sebagai aset *safe haven* daripada aset kripto yang dapat memberikan perlindungan terhadap ketidakpastian ekonomi [10].

2.3. Data Mining

Data mining sering diasosiasikan dengan istilah *Knowledge Discovery in Database* (KDD), yakni keseluruhan rangkaian proses sistematis untuk menentukan pola baru dan sah dari data yang tersedia. *Knowledge Discovery in Database* (KDD) mencakup beberapa tahap, mulai dari pemilihan data, pengolahan data, transformasi, pemodelan, hingga evaluasi hasil yang diperoleh [11].

2.4. Machine Learning

Machine learning merupakan salah satu cabang kecerdasan buatan (*artificial intelligence*) yang berfokus pada pengembangan metode dan algoritma yang memungkinkan sistem mempelajari pola-pola atau hubungan dalam sebuah data secara otomatis. Dengan demikian, sistem *machine learning* dapat meningkatkan performanya seiring dengan banyaknya data tanpa intervensi pemrograman manual secara langsung [12].

2.5. Pengolahan Data

Pengolahan data adalah berbagai langkah atau proses yang dilakukan pada data, sebelum digunakan dalam analisis atau pemodelan. Tujuan dari pengolahan data ini adalah untuk meningkatkan kualitas output dengan melakukan *cleaning* pada data, seperti memperbaiki atau menghapus data yang rusak atau tidak relevan. Tahap ini dilakukan untuk memastikan data dapat diolah dengan maksimal oleh model [13].

2.6. Train-Test Split

Teknik *train-test split* digunakan untuk memisahkan sebagian data yang digunakan untuk pelatihan model dan sebagian digunakan untuk melakukan evaluasi performa model. Tujuan dilakukan pembagian dataset menjadi dua, adalah untuk menghindari fenomena *overfitting*, yaitu kondisi dimana model memiliki performa yang baik pada data latih tetapi buruk pada data baru [14].

2.7. Random Forest Regression

Random Forest Regression adalah algoritma yang digunakan untuk melakukan prediksi nilai kontinu (numerik) dengan membangun dan menggabungkan banyak pohon keputusan (*decision tree*). Setiap pohon akan memberikan hasil prediksi berdasarkan subset data yang berbeda, kemudian semua hasil prediksi tersebut dirata-ratakan untuk memperoleh satu nilai estimasi akhir. Pendekatan ini membantu mengurangi risiko kesalahan yang mungkin

muncul jika hanya menggunakan satu pohon keputusan, sehingga hasil akhir menjadi lebih stabil, akurat, dan tahan terhadap *overfitting* [8].

2.8. Grid Search

GridSearch adalah metode optimasi *hyperparameter* yang mencoba setiap kombinasi nilai parameter secara sistematis dan memilih kombinasi terbaik berdasarkan metrik evaluasi tertentu [5]. Pada penelitian ini, *GridSearch* digunakan untuk menemukan *hyperparameter* terbaik pada model yang akan dibangun melalui skema *TimeSeriesSplit* 7-fold dengan metrik *Mean Absolute Error* (MAE) dan *Root Mean Square Error* (RMSE)

2.9. Metode Evaluasi

a. Mean Absolute Error (MAE)

Mean Absolute Error (MAE) mengukur akurasi model prediksi dengan menghitung selisih persentase antara nilai aktual dan prediksi [15].

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

Dimana:

n = jumlah total baris data

y_i = nilai actual

$\hat{y}_i$ = nilai prediksi

b. Root Mean Square Error (RMSE)

Root Mean Square Error (RMSE) mengukur rata-rata perbedaan kuadrat antara nilai prediksi dan nilai aktual dengan menghitung akar kuadrat dari rata-rata selisih kuadrat [15].

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

Dimana:

n = jumlah baris data.

y_i = nilai aktual

$\hat{y}_i$ = nilai prediksi

c. Mean Absolute Percentage Error (MAPE)

Mean Absolute Percentage Error (MAPE) mengukur akurasi model prediksi dengan menghitung selisih persentase antara nilai aktual dan prediksi. [15].

$$MAPE = \frac{\sum \frac{|Y - Y'|}{Y} \times 100\%}{n} \quad (3)$$

Dimana:

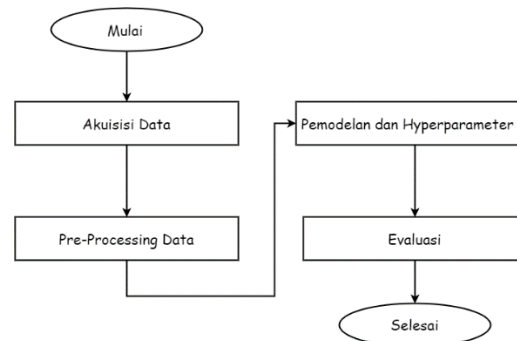
n = jumlah baris data

Y = nilai aktual

Y' = nilai prediksi

3. METODE PENELITIAN

Penelitian ini memiliki beberapa tahap, berikut merupakan alur penelitian secara keseluruhan digambarkan pada Gambar 1.



Gambar 1. Bagan alir penelitian

Pada Gambar 1 diatas, penelitian ini terdiri dari beberapa tahap yang meliputi akuisisi data, *preprocessing* data, pemodelan dan *hyperparameter*, serta evaluasi.

3.1. Akuisisi data

Data yang digunakan pada penelitian ini bersumber dari *Investing.com* dengan rentang waktu 18 September 2020 hingga 16 April 2026. Total data yang diperoleh sebanyak 1.644 data harga harian dalam satuan Rupiah (IDR) per gram. Karakteristik dataset yang digunakan disajikan pada Tabel 2.

Tabel 2. Karakteristik dataset

Karakteristik	Keterangan
Sumber data	Investing.com
Rentang waktu	18 September 2020 – 16 April 2026
Jumlah data	1.644
Satuan	Rupiah (IDR) per gram
Frekuensi	Harian

Berdasarkan Tabel 2, dataset yang digunakan merupakan data deret waktu (*time-series*) harga emas harian dalam satuan Rupiah (IDR) per gram.

3.2. Data Preprocessing

Pre-premrosesan merupakan tahap yang cukup krusial untuk meningkatkan kualitas data sebelum digunakan pada model. Langkah-langkah yang dilakukan meliputi :

- Pembersihan Data: menghapus beberapa kolom yang tidak diperlukan selain kolom tanggal (*date*) dan harga penutup (*close price*).
- Tranformasi Data: melakukan konversi tipe data setiap kolom.
- Penyeragaman Data: melakukan penyeragaman data menggunakan metode *forward fill* (*last observation carried forward*), yaitu menggunakan nilai harga penutup (*close price*) pada tanggal sebelumnya. Metode ini digunakan untuk menjaga kontinuitas data dan menghindari kekosongan data, karena dataset harga emas harian memiliki kekosongan pada hari-hari tertentu seperti hari libur dan akhir pekan.
- Finalisasi Data: memastikan bahwa data sudah sesuai dengan kriteria yang diperlukan oleh model.

3.3. Pemodelan dan Hyperparameter

Model yang digunakan pada penelitian ini adalah *Random Forest Regression*, algoritma *ensemble learning* yang bekerja dengan membangun banyak pohon keputusan secara paralel dan merata-ratakan hasilnya untuk menghasilkan prediksi yang stabil [16].

Sebelum pelatihan model, dataset akan dibagi menjadi dua data dengan perbandingan 80:20 secara sekuensial sesuai urutan waktu. 80% dari data digunakan sebagai data latih untuk melatih model, sedangkan 20% dari data digunakan sebagai data uji untuk melakukan evaluasi model. Tujuan pembagian dataset tersebut untuk mencegah kebocoran informasi masa depan (*feature leakage*) yang terjadi pada data *time-series*.

Agar model bekerja pada kondisi terbaik, dilakukan *hyperparameter tuning* menggunakan *Grid Search* dengan skema *TimeSeriesSlipt* 7-fold. *Grid Search* menguji setiap kombinasi parameter secara sistematis dan memilih kombinasi yang menghasilkan nilai MAE + RMSE terendah pada data validasi. Parameter yang diuji ditampilkan pada Tabel 3.

Tabel 3. *Hyperparameter Tuning*

Parameter	Nilai Uji
<i>n_estimators</i>	50, 100, 200
<i>max_depth</i>	5, 10, 15, 20
<i>max_features</i>	sqrt, log2
<i>min_sample_leaf</i>	2, 4, 8

Dari Tabel 3, total kombinasi yang akan diuji adalah 72 kombinasi (3x4x2x3). Dua parameter tambahan ditetapkan secara tetap yaitu *max_sample* = 0.8 dan *min_samples_split* = 10 sebagai mekanisme *anti-overfitting*.

Untuk prediksi jangka menengah, model akan menerapkan metode *Recursive Forecasting*, di mana hasil prediksi sebelumnya digunakan kembali sebagai input untuk menghasilkan estimasi harga pada periode berikutnya.

3.4. Evaluasi

Kinerja model diukur pada data uji menggunakan tiga metrik. Ketiganya dipilih karena masing-masing memberikan sudut pandang yang berbeda.

a. Mean Absolute Error (MAE)

MAE menghitung rata-rata selisih absolut antara harga emas aktual dan hasil prediksi model dalam satuan rupiah (IDR) per gram [15].

b. Root Mean Square Error (RMSE)

RMSE menghitung akar kuadrat dari rata-rata selisih kuadrat antara harga aktual dan harga prediksi model. Berbeda dengan MAE, RMSE memberikan penalti lebih besar pada kesalahan prediksi yang ekstrem. Ini berguna untuk mendeteksi apakah model sesekali melakukan kesalahan besar yang tidak terlihat dari MAE saja [15].

c. Mean Absolute Percentage Error (MAPE)

MAPE mengukur rata-rata persentase kesalahan prediksi terhadap harga aktual. Metriks ini paling mudah diinterpretasikan dalam konteks investasi karena hasilnya dinyatakan dalam bentuk persentase [15].

4. HASIL DAN PEMBAHASAN

Bagian ini menyajikan hasil setiap tahapan penelitian, mulai dari hasil *preprocessing* data, optimasi *hyperparameter*, evaluasi model, hingga hasil prediksi. Setiap tahap dibahas secara kontekstual untuk memberikan pemahaman menyeluruh terhadap proses dan capaian yang diperoleh.

4.1. Hasil Preprocessing Data

Dataset awal yang diperoleh dari *Investing.com* terdiri dari enam kolom yaitu *Date*, *Price*, *Open*, *High*, *Low*, *Vol*, dan *Change %*. Setelah tahap pembersihan, hanya dua kolom yang dipertahankan yaitu *Date* dan *Close Price*. Selanjutnya, penelitian ini melakukan pengisian kekosongan data pada hari libur dan akhir pekan menggunakan metode *forward fill*, sehingga seluruh tanggal dalam rentang 18 September 2020 hingga 16 April 2026 terisi penuh tanpa nilai kosong maupun urutan *time series* yang terputus. Sepuluh data teratas hasil *preprocessing* disajikan pada Tabel 4.

Tabel 4. Hasil *preprocessing* data

No	Date	Close Price (Rp)
1	2020-09-18	923.929
2	2020-09-19	923.929
3	2020-09-20	923.015
4	2020-09-21	906.079
5	2020-09-22	905.271
6	2020-09-23	883.681
7	2020-09-24	894.130
8	2020-09-25	890.016
9	2020-09-26	890.016
10	2020-09-27	890.346

Pada Tabel 4 diatas, merupakan sepuluh data teratas hasil *preprocessing*. Data yang ditampilkan merupakan sepuluh data pertama secara kronologis, yaitu dimulai dari tanggal 18 September 2020 sebagai titik awal dataset.

4.2. Optimasi Hyperparameter

Pada tahap awal pengembangan model, penelitian ini mencoba memprediksi harga emas secara langsung dalam satuan Rupiah (harga absolut). Namun, hasil yang diperoleh menunjukkan performa yang tidak memuaskan, di mana model gagal menggeneralisasi dengan baik pada data uji meski menunjukkan performa tinggi pada data latih. Setelah ditelusuri, ditemukan bahwa penyebab utamanya adalah data harga emas pada penelitian ini memiliki tren naik yang kuat, dari sekitar Rp900.000 pada awal periode pelatihan hingga di atas Rp2.600.000 pada periode uji, sehingga menyebabkan distribusi *shift* antar fold pada *TimeSeriesSplit*.

Berdasarkan kondisi tersebut, penelitian ini mengubah pendekatan prediksi dari harga *absolut* menjadi persentase perubahan harga (*pct_change*). Berbeda dengan harga *absolut* yang memiliki tren naik kuat, *pct_change* memiliki distribusi yang stabil sepanjang waktu sehingga model dapat mempelajari pola yang lebih konsisten. Harga absolut kemudian diperoleh kembali dengan mengalikan harga hari sebelumnya dengan hasil konversi *pct_change*.

Dengan pendekatan *pct_change* tersebut, penelitian ini melakukan optimasi hyperparameter menggunakan *GridSearch* dengan skema *Weighted TimeSeriesSplit* 7-fold, di mana bobot lebih tinggi diberikan pada fold terbaru untuk menangkap tren pasar yang lebih relevan. Dua parameter ditetapkan secara tetap sebagai langkah preventif terhadap *overfitting*, yaitu *max_samples*=0.8 dan *min_samples_split*=10. Seleksi parameter terbaik dilakukan berdasarkan nilai MAE + RMSE terendah pada data validasi. Hasil empat kombinasi teratas dari 72 kombinasi yang diuji disajikan pada Tabel 5.

Tabel 5. Empat hasil *GridSearch*

Rank	1	2	3	4
<i>n_estimators</i>	200	200	200	200
<i>max_depth</i>	5	5	5	5
<i>max_features</i>	sqrt	log2	log2	sqrt
<i>min_leaf</i>	8	8	4	4
val mae	0,004896	0,004896	0,004904	0,004904
val rmse	0,007387	0,007387	0,007390	0,007390
mae + rmse	0,012283	0,012283	0,012294	0,012294
overfit rasio	4,98%	4,98%	5,71%	5,71%

Berdasarkan Tabel 5, seluruh empat kombinasi terbaik secara konsisten menghasilkan *max_depth* = 5, yang mengindikasikan bahwa kedalaman pohon lebih dangkal untuk data ini. Dari hasil tersebut, dipilih parameter dengan rangking pertama yaitu *n_estimators* = 200, *max_depth* = 5, *max_features* = sqrt dan *min_sample_leaf* = 8.

4.3. Hasil Evaluasi Model

Parameter terbaik hasil *GridSearch* digunakan untuk melatih model pada data latih (80%) dan di evaluasi pada data uji (20%). Hasil evaluasi pada ketiga subset data disajikan pada Tabel 6.

Tabel 6. Hasil evaluasi model

Data	MAE (Rp)	RMSE (Rp)	MAPE (%)
Training	4.606	7.187	0,4691%
Holdout	6.520	10.018	0,4962%
Test	20.105	34.988	0,9081%

Berdasarkan Tabel 6, model menghasilkan MAE sebesar Rp20.105, RMSE sebesar Rp34.988, dan MAPE sebesar 0,9081% pada data uji. Terdapat selisih yang cukup besar antara nilai MAE dan RMSE pada data training dan data uji dalam satuan Rupiah. Ini terjadi karena konsekuensi langsung dari perbedaan skala harga gram, di mana data latih harganya kisaran

Rp900.000 – Rp2.500.000 per gram sementara data uji diatas Rp2.500.000 per gram. Dengan perbedaan skala tersebut, kesalahan absolut dalam Rupiah memang ikut besar secara proporsional meski persentase kesalahannya tetap kecil. Hal ini dikonfirmasi oleh nilai MAPE sebesar 0.9081% yang tergolong masih baik.

4.4. Hasil Prediksi

Model menghasilkan prediksi emas melalui pendekatan *Recursive Forecasting*, di mana hasil prediksi *pct_change* digunakan kembali sebagai input untuk menghasilkan estimasi hari berikutnya secara berurutan. Perbandingan harga aktual dan harga prediksi pada data uji disajikan pada Gambar 2 berikut

Gambar 2. Grafik harga *actual* vs harga prediksi

Dari Gambar 2, terlihat bahwa, secara umum, model mampu mengikuti tren pergerakan harga emas dengan baik pada data uji. Kesalahan prediksi cenderung membesar pada periode dengan volatilitas tinggi, yang merupakan karakteristik umum dari model berbasis data historis tanpa variabel eksternal. Meski demikian, nilai MAPE sebesar 0,9081% menunjukkan bahwa secara rata-rata prediksi model berada dalam rentang yang sangat dapat diterima untuk keperluan estimasi harga jangka menengah.

5. KESIMPULAN DAN SARAN

Penelitian ini menunjukkan bahwa model *Random Forest Regression* mampu memprediksi harga emas di Indonesia dengan baik melalui pendekatan prediksi persentase perubahan harga (*pct_change*). Melalui optimasi *GridSearch* dengan *Weighted TimeSeriesSplit* 7-fold terhadap 72 kombinasi, diperoleh parameter terbaik *n_estimators* = 200, *max_depth* = 5, *max_features* = sqrt dan *min_sample_leaf* = 8 dengan *overfit ratio* 4,98%, yang menghasilkan MAPE sebesar 0,9081% pada data uji yang membuktikan bahwa pola historis harga emas sudah memadai sebagai dasar prediksi jangka menengah tanpa memerlukan variabel eksternal.

DAFTAR PUSTAKA

- [1] A. Aulia, B. Aprianti, Y. Supriyanto, and C. Rozikin, "Prediksi Harga Emas dengan Menggunakan Algoritma Support Vector Regression (Svr) dan Linear Regression (LR)," *Jurnal Ilmiah Wahana Pendidikan*, vol. 8, no. 5, Apr. 2022, doi: 10.5281/zenodo.6408864.

- [2] M. A. Pandu W, R. E. Saputro, P. Purwadi, and U. A. Rohmah, "Analisis Tingkat Akurasi Metode Naive Bayes dan Random Forest dalam Prediksi Penjualan Emas," *Jurnal Pendidikan dan Teknologi Indonesia*, vol. 5, no. 7, pp. 1809–1821, Jul. 2025, doi: 10.52436/1.jpti.732.
- [3] G. W. Gusnindar and Syafri, "FAKTOR PENGARUH HARGA EMAS DENGAN METODE ERROR CORRECTION MODEL (ECM)," *Jurnal Ekonomi Trisakti*, vol. 3, no. 2, pp. 2359–2366, Jul. 2023, doi: 10.25105/jet.v3i2.16556.
- [4] G. Taneva-Angelova, S. Raychev, and G. Ilieva, "A Framework for Gold Price Prediction Combining Classical and Intelligent Methods with Financial, Economic, and Sentiment Data Fusion," *International Journal of Financial Studies*, vol. 13, no. 2, Jun. 2025, doi: 10.3390/ijfs13020102.
- [5] I. M. Gananta, I. N. Purnama, and K. Q. Fredlina, "OPTIMASI PREDIKSI HARGA EMAS DENGAN METODE SUPPORT VECTOR REGRESSION (SVR) MENGGUNAKAN ALGORITMA GRID SEARCH," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 7, no. 6, pp. 3160–3165, Dec. 2023, doi: 10.36040/jati.v7i6.8000.
- [6] S. Wahyuningsih, Kusri, and Kusri, "Studi Literatur Mengenai Prediksi Harga Emas Menggunakan Machine Learning," *Jurnal Dielektrika – Department of Electrical Engineering University of Mataram*, vol. 10, no. 2, pp. 112–117, Aug. 2023.
- [7] N. K. Agusmawati, F. Khoiriyah, and A. Tholib, "PREDIKSI HARGA EMAS MENGGUNAKAN METODE LSTM DAN GRU," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 11, no. 3, Aug. 2023, doi: 10.23960/jitet.v11i3.3250.
- [8] A. Agustino Maulana, K. Latifah, and N. D. Saputro, "IMPLEMENTASI ALGORITMA RANDOM FOREST REGRESSION UNTUK ESTIMASI HARGA MOBIL BEKAS MEREK TOYOTA," *Jurnal Informatika Teknologi dan Sains*, vol. 7, no. 3, pp. 1294–1303, Aug. 2025, doi: 10.51401/jinteks.v7i3.6439.
- [9] Afifa, R. Ardiansyah, C. Lamasitudju, R. Laila, and D. S. Angreni, "PERBANDINGAN METODE ARIMA DAN RANDOM FOREST DALAM MEMPREDIKSI HARGA EMAS BERDASARKAN PERGERAKAN MATA UANG DAN SUKU BUNGA," *JUPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 10, no. 4, pp. 3266–3279, Jun. 2024, doi: <https://doi.org/10.29100/jupi.v10i4.6346>.
- [10] S. Nawatmi, A. B. Santosa, and A. Nusantara, "THE ROLE OF GOLD AS SAFE HAVEN OR DIVERSIFIER INVESTMENT IN INDONESIA," *Business and Accounting Research (IJEBA) Peer Reviewed-International Journal*, vol. 8, no. 2, May 2024, doi: <https://doi.org/10.29040/ijebar.v8i2.13955>.
- [11] Yendrizal and Y. Alfero, "Data Mining Penjualan Bibit Tanaman Menggunakan Algoritma Apriori," *Jurnal Nasional Komputasi dan Teknologi Informasi (JNKTI)*, vol. 7, no. 6, pp. 1908–1917, Dec. 2024, doi: <https://doi.org/10.32672/jnkti.v7i6.8346>.
- [12] A. Putra Argadianata, D. Abdul Fatah, and H. Sukri, "KLASIFIKASI KUALITAS BUAH APEL MENGGUNAKAN METODE RANDOM FOREST," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 9, no. 2, pp. 2016–2022, Apr. 2025, doi: <https://doi.org/10.36040/jati.v9i2.12854>.
- [13] R. Hidayat *et al.*, "Implementasi Algoritma Random Forest Regression Untuk Memprediksi Penjualan Produksi di Supermarket," *Jurnal Sistem Informasi dan Sistem Komputer*, vol. 10, no. 1, pp. 101–109, Jan. 2025, doi: 10.51717/simkom.v10i1.703.
- [14] M. Sivakumar, S. Parthasarathy, and T. Padmapriya, "Trade-off between training and testing ratio in machine learning for medical image processing," *PeerJ Comput. Sci.*, vol. 10, Sep. 2024, doi: 10.7717/PEERJ-CS.2245.
- [15] Y. Handoko, E. Satria, and A. Nasrulloh, "Algoritma Regresi Linier Dalam Prediksi Jumlah Pendaftar Program Pendidikan Di Lingkungan Pesantren," *Jurnal Algoritma*, vol. 22, no. 1, pp. 557–568, May 2025, doi: 10.33364/algoritma/v.22-1.1802.
- [16] J. Ganapa, S. Choudari, and M. Rao K, "Gold Price Prediction Using Random Forest Regression," *Educational Administration Theory and Practices*, vol. 30, no. 1, pp. 1052–1055, Jan. 2024, doi: 10.53555/kuey.v30i1.5928