

Prediction of World Gold Price Using k -Nearest Neighbor Method

Muhamad Rayhan Nanda P, Zamahsary Martha*, Dodi Vionanda, Admi Salma

Departemen Statistika, Universitas Negeri Padang, Padang, Indonesia

*Corresponding author : zamahsarymartha@fmipaunp.ac.id

Submitted : 18 Oktober 2024

Revised : 11 November 2024

Accepted : 12 November 2024

ABSTRACT

This research aims to predict world gold prices using k -nearest neighbor (KNN) method with secondary data from the London Bullion Market Association (LBMA) in monthly time series data from January 2019 to December 2023. The analysis process divides the data into two parts: 80% for training data from January 2019 to December 2022 and 20% for testing data from January to December 2023. The analysis results show that the Mean Absolute Percentage Error (MAPE) value of the KNN method is 4.5%, which indicates a very good level of accuracy. With a MAPE below 10%, the KNN model is proven to be able to accurately predict world gold prices. Gold price predictions for January to December 2024 period show a consistent upward trend, which is influenced by factors such as global economic fluctuations, increased gold demand, and geopolitical uncertainty. These results show that the KNN model is reliable as a tool for forecasting future world gold prices.

Keywords: *Global Gold Prices, k -Nearest Neighbor (KNN), Price Forecasting*



This is an open access article under the Creative Commons 4.0 Attribution License, which permits unrestricted use, distribution, and reproduction in medium, provided the original work is properly cited. ©2022 by author and Universitas Negeri Padang.

I. PENDAHULUAN

Investasi merupakan konsep ekonomi yang berarti penanaman modal dalam jangka panjang dengan tujuan meraih keuntungan. Saat ini, ada banyak instrumen investasi yang tersedia, seperti deposito, saham, valuta asing, properti, emas, bitcoin, dan lainnya. Salah satu bentuk investasi yang paling populer adalah emas. Emas adalah komoditas tradisional yang digunakan sejak zaman dahulu dan diterima oleh seluruh dunia (Marwondo dan Hidayah, 2023). Emas sebagai barang investasi dapat berbentuk perhiasan, batangan, maupun koin.

Memantau perkembangan harga jual emas didunia sangat penting sebab emas merupakan salah satu aset yang paling stabil dan dianggap sebagai tempat berlindung nilai (*safe haven*) oleh banyak investor. Fluktuasi harga emas dunia dapat memberikan indikasi tentang kondisi ekonomi global, tingkat inflasi, dan ketidakpastian geopolitik. Selama periode ketidakstabilan ekonomi atau politik, harga emas relatif meningkat karena investor mencari aset yang lebih aman. Dengan demikian, analisis harga emas bisa menjadi alat yang berguna untuk meramalkan tren ekonomi dan perilaku pasar keuangan (Chai, 2021).

Harga emas yang ditetapkan oleh *London Bullion Market Association* (LBMA) menjadi acuan harga emas global melalui sistem *London Gold Fixing*. LBMA menetapkan perubahan harga emas dua kali sehari pada hari kerja, yaitu pukul 10.30 (Gold A.M) dan pukul 15.00 (Gold P.M), dengan satuan dalam Dolar Amerika Serikat per troy ounce (31,103 gram). Penetapan harga ini dilakukan oleh lima anggota LBMA, yaitu *Bank of Nova Scotia*, *Barclays Capital*, *Deutsche Bank*, *HSBC*, dan *Societe Generale*. *Harga Gold P.M* menjadi acuan harga emas bagi semua negara di dunia (Izati dkk., 2019).

Sebagai salah satu logam yang berharga, emas menunjukkan kinerja yang sangat baik, mampu bertahan saat terjadi krisis ekonomi global, sementara banyak komoditas keuangan lainnya mengalami penurunan. Meskipun demikian, harga emas di pasar internasional juga mengalami fluktuasi yang tidak terduga. Berbagai faktor memengaruhi harga emas, seperti kurs dolar, inflasi, permintaan dan penawaran, serta kebijakan moneter. Mengingat fluktuasi harga emas yang terus berubah seiring waktu, penting untuk melakukan ramalan harga suatu emas di masa depan agar dapat jadi acuan bagi investor untuk membuat keputusan investasi yang tepat. Dalam berinvestasi emas diperlukannya strategi khusus, salah satunya yaitu mengetahui prediksi harga emas dunia (Thalib dkk., 2023). Prediksi harga emas dunia sangat dibutuhkan karena dari ramalan harga emas tersebut para investor bisa mendapatk keuntungan sesuai dengan perencanaan atau untuk meminimalkan resiko. Prediksi harga emas dunia sangat diperlukan agar investor dapat mencapai keuntungan sesuai dengan rencana dan mengurangi

risiko. Prediksi yang tepat membantu dalam pengambilan keputusan yang lebih baik. Setiap jenis investasi, tanpa terkecuali, mengandung risiko kerugian. Namun, risiko ini dapat diminimalkan dengan prediksi yang akurat.

Pada umumnya, peramalan menggunakan runtun waktu atau *time series* ialah sebuah metode yang digunakan untuk meramalkan nilai di masa yang akan datang dengan cara mengamati data berdasarkan urutan waktu (Nanja, 2015). Dalam banyak kasus, data deret waktu yang memenuhi asumsi yang diperlukan sangat jarang ditemukan. Salah satu alasannya adalah terdapat hubungan non-linear antar variabel, sehingga dibutuhkan metode yang lebih efisien. Berbagai metode telah dikembangkan untuk memprediksi data runtun waktu dengan ukuran besar dan banyak variabel penjelas, salah satunya adalah metode *k-Nearest Neighbor* (kNN). Metode kNN dapat diterapkan pada data yang tidak memenuhi asumsi klasik dan memiliki karakteristik *non-linear* (Sinta, 2015). Awalnya, kNN adalah metode yang digunakan untuk analisis klasifikasi, tetapi dalam beberapa tahun terakhir, metode ini juga diterapkan untuk prediksi (Alfin, 2022). Algoritma kNN adalah algoritma dengan cara perhitungan yang sederhana dan mudah, tetapi memiliki tingkat keakuratan yang cukup tinggi, sehingga menjadi suatu metode paling populer dalam analisis data mining. Kelebihan dari metode kNN adalah kemampuannya dalam mengatasi data yang tidak linear dan tidak homogen, serta tidak memerlukan asumsi tertentu mengenai distribusi data (Virdaus, 2020).

II. METODOLOGI PENELITIAN

A. Jenis Penelitian dan Sumber Data

Pada penelitian ini menggunakan penelitian terapan karena penelitian ini mencari solusi dari masalah investasi emas dunia, sehingga dapat dilakukan penanganan yang tepat dari masalah yang ada. Dalam hal ini eksperimen dilakukan melalui proses pengambilan data harga emas dunia kemudian dilakukan peramalan menggunakan metode kNN. Dalam penelitian ini, digunakan jenis data sekunder yang didapatkan dari situs resmi yaitu LBMA (*London Bullion Market Association*). Penelitian ini menggunakan data runtun waktu atau *time series* harga emas perbulan di dunia. Terdapat 60 observasi dari bulan Januari 2019 - Desember 2023. Pada penelitian ini terdapat variabel univariat yang mana variabelnya adalah harga emas dunia (USD).

B. Teknik Analisis Data

Untuk menganalisis data pada penelitian ini memakai bantuan aplikasi *RStudio*. Berikut tahapan analisis yang akan dilakukan

1. Mengumpulkan data dari situs LBMA (*London Bullion Market Association*) dari tahun 2019 - 2023 dengan periode bulanan.
2. Membagi dataset menjadi data *training* dan data *testing*. Adapun rasio yang dipakai pada pembagiannya adalah 80% : 20% (Rogel, 2018). Data dari bulan Januari 2019 - Desember 2022 sebagai data *training* dan data bulan Januari - Desember 2023 sebagai data *testing*.
3. Mengkonversi dari univariat ke multivariat

Pada metode kNN, setiap data terdiri dari variabel fitur dan variabel target. Menurut Martinez dkk (2017) untuk menerapkan kNN pada pengaturan prediksi *time series* data univariat terlebih dahulu harus dikonversi ke bentuk data multivariat agar bisa digunakan. Mengkonversi data dari univariat menjadi multivariat merupakan mengubah data yang awalnya hanya memiliki satu variabel menjadi data yang memiliki dua variabel atau lebih. Konversi data ini dapat diperoleh dengan melakukan penambahan fitur baru pada dataset. Fitur yang baru dibuat berdasarkan nilai waktu sebelumnya, yang sering disebut sebagai lag (nilai kelambanan), dapat disusun menjadi kolom-kolom baru.

Dalam data *time series* jika panjang periode musiman adalah m , maka akan digunakan lag $1:m$. Contohnya untuk data bulanan maka digunakan lag $1:12$. Secara sederhana konversi data dapat digambarkan sebagai berikut. Misalkan terdapat data *time series* bulanan selama 5 tahun, yaitu $t = \{x_1, x_2, x_3, x_4, x_5, \dots, x_{60}\}$. Maka digunakan lag $1:12$ variabel fitur. Hasil konversi data dapat dilihat pada Tabel 1.

Tabel 1. Struktur Data Setelah Dikonversi

No	Fitur	Target
1	$x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}, x_{12}$	x_{13}
2	$x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}, x_{12}, x_{13}$	x_{14}
⋮	⋮	⋮
36	$x_{36}, x_{37}, x_{38}, x_{39}, x_{40}, x_{41}, x_{42}, x_{43}, x_{44}, x_{45}, x_{46}, x_{47}$	x_{48}

4. Melakukan Transformasi

Transformasi data pada *time series* mampu mengoptimalkan kinerja model dan akurasi peramalan. Karena alasan ini, transformasi memiliki peranan penting dalam model peramalan. Menurut Martinez dkk., (2022). Salah satu transformasi yang bisa digunakan ialah transformasi additive ditunjukkan pada Persamaan

$$f'_{x,i} = f_{x,i} - \frac{1}{n} \sum_{i=1}^n f_{x,i} \quad (1)$$

Keterangan:

$f_{x,i}$: Nilai fitur ke- i baris ke- x .

$f'_{x,i}$: Hasil tranformasi nilai fitur ke- i dan data ke- x .

n : Banyaknya fitur.

$$t'_x = t_x - \frac{1}{n} \sum_{i=1}^n f_{x,i} \quad (2)$$

Keterangan:

t_x : Nilai target data ke- x .

n : Banyaknya fitur.

$f_{x,i}$: Nilai fitur ke- i dan data ke- x .

t'_x : Hasil tranformasi nilai target data ke- x .

5. Memilih beberapa nilai k untuk model k NN

Pada penelitian terdapat 36 nilai k yang akan digunakan sebagai model, yaitu $k = 1$ sampai $k = 36$. Alasan menggunakan seluruh nilai k karena ingin melihat performa seluruh nilai k .

6. Menghitung jarak *euclidean*

Jarak yang paling banyak digunakan untuk data numerik adalah jarak *euclidean* rumus pada Persamaan

$$d(f^i, q) = \sqrt{\sum_{x=1}^n (f_x^i - q_x)^2} \quad (3)$$

Keterangan:

$d(f^i, q)$: Jarak *euclidean* antara data training ke- i dengan data yang akan di prediksi.

f_x^i : Nilai fitur ke- x baris ke- i dari data *training*.

q_x : Nilai fitur ke- x dari data yang akan di prediksi

n : Banyak data *training*.

7. Menghitung nilai prediksi model k NN

Data time series merupakan data yang memperhatikan urutan waktu, maka data training yang terpilih sebagai k tetangga terdekat dapat diberi bobot. Tujuannya adalah untuk memberikan bobot lebih kepada tetangga yang lebih dekat. Pemberian bobot menggunakan Persamaan

$$\hat{y} = \frac{\sum_{i=1}^k w_i g_i}{\sum_{i=1}^k g_i} \quad (4)$$

Keterangan:

\hat{y} : Nilai Prediksi.

w_i : Nilai pembobot tetangga terdekat ke- i .

g_i : Nilai target tetangga terdekat ke- i .

k : Banyak nya tetangga terdekat.

Untuk memboboti nilai target yang terpilih menjadi tetangga terdekat digunakan nilai pembobot. Untuk menghitung nilai pembobot digunakan Persamaan

$$w_i = \frac{1}{d_i^2} \quad (5)$$

Keterangan:

w_i : Nilai pembobot untuk tetangga terdekat ke- k yang terpilih.

d_i : Jarak *euclidean* antara data yang akan diprediksi dengan tetangga ke- i .

8. Tranformasi hasil prediksi ke bentuk semula

Untuk mentransformasi hasil prediksi ke bentuk awal, rata-rata vektor fitur dari data yang akan diprediksi ditambahkan kedalam hasil prediksi. Secara matematis dapat dilihat pada Persamaan

$$\hat{y}' = \hat{y} + \frac{1}{n} \sum_{i=1}^n q_x \quad (6)$$

Keterangan:

- \hat{y} : Hasil prediksi (dalam bentuk tranformasi additive)
- n : Banyaknya fitur.
- q_x : Nilai fitur ke- x dari data yang akan di prediksi.
- \hat{y}' : Hasil prediksi setelah ditranformasi ke bentuk awal.

9. Evaluasi menggunakan MAPE

Ada beberapa indikator yang umum atau sering digunakan untuk evaluasi ini, salah satunya bisa dengan menggunakan MAPE, semakin kecil nilai dari indikator evaluasi ini maka akan semakin akurat hasil prediksi

$$MAPE = \frac{100}{n} \sum_{t=1}^n \left| \frac{x_t - f_t}{x_t} \right| \tag{7}$$

Keterangan:

- x_t : Nilai aktual ke- t .
- f_t : Nilai prediksi dari model ke- t .
- n : Banyaknya fitur.

Menurut Chang dkk., (2007) tabel nilai dari hasil prediksi yang didapatkan dari nilai MAPE dapat dilihat Tabel 2.

Tabel 2. Kategori Prediksi MAPE

MAPE	Kategori
<10%	Sangat baik
10-20%	Baik
20-50%	Layak
>50%	Buruk

III. HASIL DAN PEMBAHASAN

Analisis data harga emas dunia rata-rata bulanan pada LBMA (*London Bullion Market Association*) yang di prediksi dengan memakai metode k-nearest neighbor dilakukan bantuan *software R Studio*. Langkah awal dalam analisis ini adalah melakukan training (80%) dan testing (20%), yaitu data dari bulan Januari 2019 - Desember 2022 sebagai data training dan bulan Januari - Desember 2023 sebagai data testing. Setelah dataset tadi dibagi langkah selanjutnya yaitu mengubah data training menjadi bentuk multivariat yang dapat dilihat pada Tabel 3.

Tabel 3. Struktur Data *Training* Setelah Konversi

No.	Fitur	Target
1	1291.75, 1320.07, 1300.90, 1286.45, 1283.95, 1359.04, 1412.98, 1498.80, 1511.31, 1494.80, 1470.02, 1476.04	1560.67
2	1320.07, 1300.90, 1286.45, 1283.95, 1359.04, 1412.98, 1498.80, 1511.31, 1494.80, 1470.02, 1476.04, 1560.67	1597.10
⋮	⋮	⋮
36	1786.65, 1816.77, 1856.30, 1947.83, 1933.90, 1848.30, 1833.83, 1736.37, 1765.65, 1682.97, 1664.45, 1726.45	1798.74

Tabel 3 menunjukkan bahwa struktur data training setelah konversi diperoleh 36 baris data yang terdiri atas fitur dan target. Setelah data dikonversi maka selanjutnya adalah melakukan transformasi data seperti pada Tabel 4.

Tabel 4. Struktur Data *Training* Hasil Transformasi

No.	Fitur	Target
1	-100.43, -72.11, -91.28, -105.73, -108.23, -33.14, 20.80, 106.62, 119.13, 102.62, 77.84, 83,86	168,49
2	-94.52, -113.69, -128.14, -130.64, -55.55, -1.61, 84.21, 96.72, 80.21, 55.43, 61.45, 146.08	182,51
⋮	⋮	⋮
36	-13.31, 16.81, 56.34, 147.87, 133.94, 48.34, 33.87, -63.59, -34.31, -116.99, -135.51, -73,51	-1,22

Berdasarkan Tabel 4 terlihat bahwa hasil transformasi data menghasilkan nilai positif dan negatif. Transformasi data ini dilakukan dengan tujuan untuk meningkatkan akurasi peramalan.

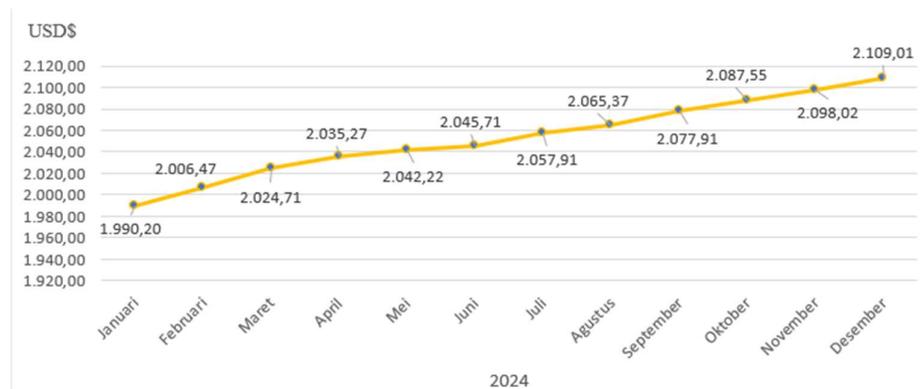
Tahapan selanjutnya yaitu memilih nilai k optimal untuk model KNN. Untuk itu nilai k yang akan digunakan adalah k=1 sampai k=36. Lalu melakukan prediksi untuk periode selanjutnya. Hasil prediksi diperoleh dapat dilihat pada Tabel 5.

Tabel 5. Nilai MAPE kNN

K	MAPE
1	9,001656
2	8,53701
3	8,855471
4	8,710897
5	8,524028
⋮	⋮
30	5,283718
31	5,206646
32	5,217486
33	5,139397
34	4,830893
35	4,624083
36	4,488219

Tabel 5 menunjukkan bahwa nilai MAPE dengan k yang berbeda memberikan hasil yang berbeda, dengan kategori prediksi pada semua k menunjukkan kemampuan peramalan sangat baik berada di bawah 10%. Diperoleh bahwa k=36 memberikan MAPE terkecil dengan nilai 4,488219 \approx 4,5%, sehingga kNN dengan k=36 dipilih jadi model terbaik.

Tahapan selanjutnya adalah melakukan prediksi harga emas dunia dari bulan Januari – Desember 2024 menggunakan kNN dengan k=36. Hasil prediksi dapat dilihat pada Gambar 1.



Gambar 1. Hasil Prediksi Harga Emas Dunia pada data LBMA Tahun 2024

Berdasarkan hasil prediksi pada Tabel 6, dapat disimpulkan bahwa harga emas dunia pada periode Januari hingga Desember 2024 diperkirakan mengalami kenaikan yang konsisten. Tren kenaikan ini tampak terjadi secara berkala, hal ini mungkin dipengaruhi oleh berbagai faktor seperti fluktuasi ekonomi global, permintaan emas yang terus meningkat, serta ketidakpastian etidakpastian geopolitik yang mendorong para investor untuk beralih ke aset aman seperti emas. Secara keseluruhan, pergerakan harga emas diperkirakan akan menunjukkan peningkatan yang stabil sepanjang tahun 2024.

IV. KESIMPULAN

Berdasarkan hasil analisis yang telah dilakukan dengan menggunakan algoritma K-Nearest Neighbor (KNN) dalam prediksi harga emas menunjukkan hasil yang cukup baik dengan nilai MAPE yang rendah dimana nilai k=36 dengan MAPE 4.5% menandakan tingkat akurasi yang tinggi dalam peramalan harga emas. Hasil prediksi harga emas dunia tahun 2024 meningkat seriap bulannya dimana harga terendah terjadi pada bulan Januari sebesar 1990.20 USD dan harga emas tertinggi pada bulan Desember sebesar 2109.01 USD dengan rata-rata harga emas dunia tahun 2024 sebesar 2053.36 USD. Manfaat utama dari hasil prediksi ini bagi investor adalah kemampuannya

untuk merencanakan strategi investasi yang lebih efektif, mengurangi risiko kerugian, serta meningkatkan potensi keuntungan melalui keputusan yang didasarkan pada prediksi harga emas yang lebih akurat.

DAFTAR PUSTAKA

- Alfin, N. F. (2022). Penerapan Metode Clustering Untuk Prediksi Produksi Bawang Merah (Ensemble K-Nearest Neighbor). Vol 3, 30-37.
- Chai, J. (2021) Structural Analysis and Forecast of Gold Price Return. *Journal of Management Science and Engineering* vol. 6 135 – 145
- Chang, P. C., Wang, Y. W., dan Liu, C.H. (2007). The Development Of A Weighted Evolving Fuzzy Neural Network For PCB Sales Forecasting. *Expert Systems With Applications*, 32(1), 86-96.
- Izati, N. A., Warsito, B., dan Widiharih, T. (2019). Prediksi Harga Emas Menggunakan Feed Forward Neural Network dengan Metode Extreme Learning Machine. *Jurnal Gaussian*, 8(2), 171-183.
- Martínez, F., Frías, M. P., Pérez, M. D., & Rivera, A. J. (2017). A Methodology For Applying k-Nearest Neighbor To Time Series Forecasting. *Artificial Intelligence Review*, 52(3), 2019–2037.
- Martínez, F., Charte, F., Frías, M. P., dan Martínez-Rodríguez, A. M. (2022). Strategies for Time Series Forecasting with Generalized Regression Neural Networks. *Neurocomputing Journal*, 491, 509–521.
- Nanja, M. (2015). Metode K-Nearest Neighbor Berbasis Forward Selection untuk Prediksi Harga Komoditi Lada. *Jurnal Pseudocode, Volume 2*.
- Rogel, S. J. (2018). Data Science and Analytics with Python. Chapman and Hall/CRC.
- Sinta, D. (2015). Metode Ensemble K-Nearest Neighbor Untuk Prediksi Harga Beras Di Indonesia. Bogor: Insitut Pertanian Bogor.
- Tholib, A., Agusmawati, N.K., dan Khoiriyah, F. (2023). Prediksi Harga Emas Menggunakan Metode LSTM dan GRU. *Jurnal Informatika dan Teknik Elektro Terapan*. vol. 12 no. 1
- Virdaus, D. (2020). Penerapan Data Mining untuk Memprediksi Harga Bawang Merah di Yogyakarta Menggunakan Metode K-Nearest Neighbor. *Jurnal Multimedia & Artificial Intelligence, Fakultas Teknologi Informasi*. no. 84, pp. 1–8.