

PENERAPAN ALGORITMA K-NEAREST NEIGHBOR (KNN) DALAM KLASIFIKASI TREND HARGA EMAS BERBASIS WEB

Rahmat¹, Helmi Fauzi Siregar.S.T., M.Kom.²

Universitas Asahan, Asahan

e-mail: ¹rahmat7359@gmail.com, ²fauzi.helmi.hf@gmail.com

Abstract: Gold is widely used as a store of value and a safe-haven investment, yet its daily price fluctuations make it difficult for investors and the public to determine the direction of price movement. This study applies the K-Nearest Neighbor (KNN) algorithm to classify gold price trends into three classes (Up, Down, and Stable) based on the monthly percentage change in the average gold price. The data are public historical gold prices from January 2022 to August 2026, in which each month consists of four period values that are processed into the highest, lowest, and average price. The percentage change in the average price relative to the previous month is used as the single numerical feature; it is labeled with a $\pm 1\%$ threshold and normalized using Min-Max Normalization. The data are split chronologically into 47 training data (February 2022–December 2025) and 8 testing data (January–August 2026). Classification is carried out by calculating the Euclidean Distance and applying majority voting among $K = 5$ nearest neighbors. The model is implemented in a web-based application built with PHP and MySQL. Testing on the eight test data yields an accuracy of 100%, with precision, recall, and F1-score of 100% for every class. However, the small test set and the imbalanced class distribution (only four Down training records and one Stable test record) indicate that the result needs further validation on longer data periods.

Keywords: K-Nearest Neighbor, classification, gold price, trend, data mining.

Abstrak: Emas banyak digunakan sebagai penyimpan nilai dan investasi aman (safe haven), namun harganya yang berubah setiap hari menyulitkan investor dan masyarakat dalam menentukan arah pergerakan harga. Penelitian ini menerapkan algoritma K-Nearest Neighbor (KNN) untuk mengklasifikasikan tren harga emas ke dalam tiga kelas (Naik, Turun, dan Stabil) berdasarkan persentase perubahan harga rata-rata bulanan. Data yang digunakan adalah data historis harga emas dari sumber publik periode Januari 2022–Agustus 2026, yang setiap bulannya terdiri atas empat nilai Periode dan diolah menjadi Harga Tertinggi, Harga Terendah, dan Harga Rata-rata. Persentase perubahan harga rata-rata terhadap bulan sebelumnya digunakan sebagai fitur numerik tunggal, diberi label dengan ambang batas $\pm 1\%$, dan dinormalisasi menggunakan Min-Max Normalization. Data dibagi berdasarkan urutan waktu menjadi 47 data latih (Februari 2022–Desember 2025) dan 8 data uji (Januari–Agustus 2026). Klasifikasi dilakukan dengan menghitung Euclidean Distance dan voting mayoritas dari $K = 5$ tetangga terdekat. Model diimplementasikan ke dalam aplikasi berbasis web yang dibangun menggunakan PHP dan MySQL. Pengujian terhadap delapan data uji menghasilkan akurasi 100%, dengan precision, recall, dan F1-score 100% pada setiap kelas. Namun, jumlah data uji yang kecil dan distribusi kelas yang tidak seimbang (data latih kelas Turun hanya empat dan data uji kelas Stabil hanya satu) menunjukkan bahwa hasil ini perlu divalidasi pada periode data yang lebih panjang.

Kata kunci: K-Nearest Neighbor, klasifikasi, harga emas, tren, data mining.

PENDAHULUAN

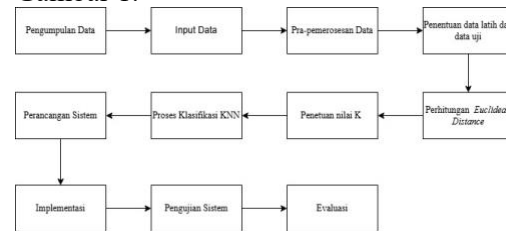
Emas banyak digunakan masyarakat tidak hanya sebagai instrumen investasi, tetapi juga sebagai sarana penyimpanan nilai kekayaan untuk menghadapi ketidakpastian ekonomi. Namun, harga emas yang berubah setiap hari menyulitkan investor maupun masyarakat umum dalam menentukan waktu yang tepat untuk membeli atau menjual emas. Kesalahan dalam memahami tren harga dapat menyebabkan kerugian atau hilangnya peluang investasi, sehingga diperlukan pendekatan yang mampu mengidentifikasi pola pergerakan harga emas berdasarkan data historis.

Fenomena tersebut terlihat dari kenaikan harga emas selama tahun 2024–2025. Harga emas Antam beberapa kali mencetak rekor tertinggi dan pada 19 Maret 2025 mencapai Rp1.759.000 per gram, didorong oleh meningkatnya permintaan terhadap aset aman (safe haven) di tengah ketidakpastian ekonomi global, inflasi, dan kondisi geopolitik (Kompas Keuangan, 2025). Sebagian besar masyarakat masih menganalisis harga hanya dengan melihat kenaikan atau penurunan pada hari sebelumnya, padahal data harga emas yang tersimpan dalam jumlah besar dapat dimanfaatkan untuk menemukan pola yang menjadi dasar pengambilan keputusan (Ferdinandus et al., 2023).

Salah satu metode klasifikasi yang dapat digunakan adalah K-Nearest Neighbor (KNN), yaitu metode yang mengklasifikasikan data baru berdasarkan mayoritas kelas dari sejumlah data terdekatnya. KNN dikenal sederhana, mudah diterapkan, dan tidak memerlukan pembentukan model yang kompleks (Nuraeni et al., 2024). Penelitian sebelumnya menunjukkan bahwa KNN mampu menganalisis dan memprediksi harga emas dunia dengan nilai Mean Absolute Percentage Error (MAPE) sebesar 4,5% (Nanda dkk., 2024), serta dapat mempelajari pola pergerakan harga emas harian dari data historis jangka

panjang saat dibandingkan dengan Regresi Linear (Martin dkk., 2025).

Berdasarkan kondisi tersebut, penelitian ini menerapkan algoritma KNN untuk mengklasifikasikan tren pergerakan harga emas ke dalam tiga kelas, yaitu Naik, Turun, dan Stabil, menggunakan data historis Januari 2022 sampai Agustus 2026. Penelitian dibatasi pada klasifikasi tren dan tidak melakukan prediksi harga secara langsung. Hasil klasifikasi diimplementasikan ke dalam aplikasi berbasis web agar proses perhitungannya dapat ditelusuri secara transparan. Tahapan penelitian ditunjukkan pada Gambar 1.



Gambar 1. Kerangka Kerja Penelitian

Objek dan Sumber Data

Objek penelitian adalah data historis harga emas bulanan periode Januari 2022 sampai Agustus 2026 yang diperoleh dari sumber data publik yang dapat diakses secara daring. Setiap bulan terdiri atas empat nilai Periode (Periode_I sampai Periode_IV) yang diolah menjadi atribut Harga Tertinggi, Harga Terendah, dan Harga Rata-rata. Data Januari 2022 tidak memiliki bulan pembanding sehingga tidak diberi kelas, dan 55 data sisanya dibagi menjadi 47 data latih (Februari 2022–Desember 2025) dan 8 data uji (Januari–Agustus 2026).

Penerapan Algoritma K-Nearest Neighbor

Tahap pertama adalah menghitung atribut turunan dari empat nilai Periode tiap bulan, yaitu Harga Tertinggi = MAX, Harga Terendah = MIN, dan Harga Rata-rata = AVERAGE. Fitur yang digunakan pada perhitungan jarak adalah Persentase Perubahan Harga Rata-rata terhadap bulan sebelumnya:

$$\Delta\% = \frac{Rata_t - Rata_t-1}{Rata_t-1} \times 100\%$$

Persentase perubahan dipilih sebagai fitur tunggal karena tren naik atau turun lebih tepat dinyatakan sebagai laju perubahan relatif antarbulan, bukan nilai harga absolut yang terus meningkat sepanjang waktu. Kelas ditentukan dengan ambang batas $\pm 1\%$, yaitu Naik jika $\Delta\% > 1\%$, Turun jika $\Delta\% < -1\%$, dan Stabil jika $-1\% \leq \Delta\% \leq 1\%$. Nilai $\Delta\%$ kemudian dinormalisasi dengan Min-Max Normalization menggunakan nilai minimum dan maksimum dari data latih, yang juga diterapkan pada data uji:

$$X' = (X - X_{min}) / (X_{max} - X_{min})$$

Pembagian data dilakukan berdasarkan urutan waktu, bukan secara acak, sehingga data latih sepenuhnya berasal dari periode sebelum data uji. Karena fitur hanya satu dimensi, jarak Euclidean antara data uji dan data latih menyederhanakan menjadi selisih mutlak:

$$d(x, y) = \sqrt{(x_i - y_i)^2} = |x_i - y_i|$$

Jarak diurutkan dari yang terkecil, kemudian diambil K data latih terdekat. Kelas data uji ditentukan melalui voting mayoritas:

$$Kelas\ Prediksi = \operatorname{argmax}(V_{Naik}, V_{Turun}, V_{Stabil})$$

dengan V_{Naik} , V_{Turun} , dan V_{Stabil} adalah jumlah suara tiap kelas dari K tetangga terdekat.

Alat dan Evaluasi

Sistem dibangun berbasis web menggunakan bahasa pemrograman PHP, basis data MySQL, dan XAMPP sebagai web server lokal, dengan Visual Studio Code sebagai code editor dan Draw.io untuk perancangan diagram UML (Ramadhan et al., 2022; Sukardi et al., 2022). Evaluasi dilakukan dengan Confusion Matrix untuk memperoleh akurasi, precision, recall, dan F1-score. Akurasi dihitung dengan:

$$Akurasi = \frac{Jumlah\ Prediksi\ Benar}{Jumlah\ Seluruh\ Data} \times 100\%$$

Sistem web sengaja tidak menyertakan halaman evaluasi dan tidak menyimpan kelas aktual data uji; perhitungan akurasi dilakukan terpisah pada lembar kerja Microsoft Excel.

HASIL DAN PEMBAHASAN

Analisa Data

Atribut dan Kriteria Kelas

Atribut penelitian terdiri atas Harga Tertinggi, Harga Terendah, Harga Rata-rata, dan Persentase Perubahan. Hanya Persentase Perubahan yang ternormalisasi digunakan sebagai fitur KNN. Kriteria kelas keluaran disajikan pada Tabel 1.

Tabel 1 Kriteria Penentuan Kelas

Kondisi $\Delta\%$	Kelas
$\Delta\% > 1\%$	Naik
$-1\% \leq \Delta\% \leq 1\%$	Stabil
$\Delta\% < -1\%$	Turun

Data Training

Data latih berjumlah 47 data (Februari 2022–Desember 2025). Sebagai contoh, pada Februari 2022 diperoleh Harga Rata-rata Rp819.435,5 dan rata-rata Januari 2022 sebesar Rp833.258,25, sehingga $\Delta\% = -1,66\%$ dan bulan tersebut berkelas Turun dengan nilai normalisasi 0,0069. Nilai $\Delta\%$ minimum pada data latih adalah $-1,7456\%$ (Desember 2025) dan maksimum $10,7655\%$ (November 2025); keduanya menjadi X_{min} dan X_{max} pada normalisasi. Distribusi kelas data latih disajikan pada Tabel 2: kelas Naik 48,9%, Stabil 42,6%, dan Turun hanya 8,5%.

Tabel 2. Distribusi Kelas Data Training

Tahun	Data	Naik	Stabil	Turun
2022	11	3	5	3
2023	12	4	8	0
2024	12	8	4	0
2025	12	8	3	1
Total	47	23	20	4

Data Testing

Data uji berjumlah 8 data, yaitu Januari sampai Agustus 2026 (Tabel 3). Kelas aktual tidak disertakan pada data yang diproses sistem. Nilai X' pada Januari 2026 melebihi 1 (1,9193) karena $\Delta\%$ sebesar 22,27% berada di luar

rentang data latih; normalisasi tetap memakai batas data latih agar skala konsisten.

Tabel 3 Data Testing (Januari–Agustus 2026)

Bulan 2026	Rata-rata (Rp)	$\Delta\%$	X'
Januari	2.380.159	22,27	1,9193
Februari	2.583.484	8,54	0,8223
Maret	2.603.998	0,79	0,2030
April	2.550.777	-2,04	-0,0238
Mei	2.430.795	-4,70	-0,2364
Juni	2.474.436	1,80	0,2830
Juli	2.386.526	-3,55	-0,1444
Agustus	2.342.373	-1,85	-0,0084

Perhitungan Klasifikasi KNN

Rumus $K = \sqrt{n}$ dengan $n = 47$ menghasilkan $\sqrt{47} = 6,86 \approx 7$. Nilai tersebut dibandingkan dengan uji coba $K = 3, 5, 7$, dan 9 terhadap kelas aktual data uji (Tabel 4). $K = 3, 5$, dan 7 sama-sama menghasilkan akurasi 100%, sedangkan $K = 9$ turun menjadi 50% karena kelas Turun hanya memiliki empat data latih sehingga tidak dapat memenangkan voting dari sembilan tetangga. Penelitian ini menetapkan $K = 5$, sama dengan nilai K yang digunakan pada aplikasi, dengan opsi bagi pengguna untuk mengubahnya.

Tabel 4 Hasil Uji Coba Nilai K

Nilai K	Prediksi Benar	Akurasi
3	8 dari 8	100%
5	8 dari 8	100%
7	8 dari 8	100%
9	4 dari 8	50%

Sebagai contoh, data uji Januari 2026 ($\Delta\% = 22,27\%$; $X' = 1,9193$) memiliki jarak 0,9193 terhadap November 2025 ($X' = 1,0000$), yaitu tetangga terdekatnya. Seluruh lima tetangga terdekatnya berkelas Naik (Naik = 5, Turun = 0, Stabil = 0) sehingga hasilnya Naik. Contoh kedua adalah April

2026 yang kelima tetangga terdekatnya ditunjukkan pada Tabel 5.

Tabel 5 Lima Tetangga Terdekat Data Uji April 2026

Peringkat	Data Training	Kelas	Jarak
1	Desember 2025	Turun	0,0238
2	Februari 2022	Turun	0,0308
3	Agustus 2022	Turun	0,0742
4	Juni 2022	Turun	0,0825
5	Agustus 2025	Stabil	0,1025

Empat tetangga berkelas Turun dan satu berkelas Stabil (Turun = 4, Stabil = 1), sehingga April 2026 diklasifikasikan Turun meskipun tidak seluruh tetangga sekelas. Proses yang sama diterapkan pada seluruh data uji dengan rekapitulasi pada Tabel 6.

Tabel 6 Rekapitulasi Voting KNN (K = 5)

Bulan 2026	Naik	Turun	Stabil	Hasil
Januari	5	0	0	Naik
Februari	5	0	0	Naik
Maret	0	0	5	Stabil
April	0	4	1	Turun
Mei	0	4	1	Turun
Juni	5	0	0	Naik
Juli	0	4	1	Turun
Agustus	0	4	1	Turun

Implementasi Sistem

Hasil klasifikasi diimplementasikan ke dalam aplikasi berbasis web (PHP dan MySQL) dengan alur kerja: Upload Data Latih, Pengolahan dan Normalisasi, Grafik Data Latih, Penentuan Nilai K, Data Uji, Proses KNN, serta Hasil dan Grafik Klasifikasi. Hasil dapat dicetak sebagai PDF dan diekspor ke CSV. Seluruh halaman, kecuali Login, hanya dapat diakses setelah pengguna berhasil masuk.

Halaman Login

Halaman Login merupakan gerbang masuk sistem. Pengguna memasukkan username dan password yang diperiksa terhadap tabel pengguna pada basis data.



Gambar 2 Halaman Login

Halaman Beranda

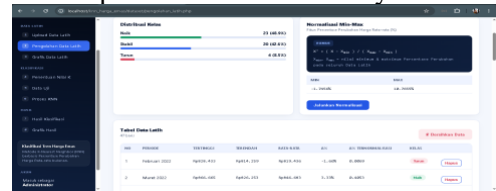
Beranda menampilkan ringkasan 47 data latih, harga rata-rata terakhir Rp1.946.690 (Desember 2025), status alur klasifikasi, grafik tren harga rata-rata, dan distribusi kelas.



Gambar 3 Halaman Beranda

Halaman Pengolahan Data Latih

Halaman ini menampilkan 47 data latih beserta distribusi kelasnya dan menyediakan tombol untuk menjalankan normalisasi Min-Max. Setelah dijalankan, sistem menampilkan nilai minimum -1,7456% dan maksimum 10,7655% serta nilai $\Delta\%$ ternormalisasi pada setiap data, sesuai Tabel 2 dan perhitungan manual pada sub-bab sebelumnya.



Gambar 4 Halaman Pengolahan Data Latih

Halaman Upload Data Latih menerima berkas CSV dalam format siap pakai atau format data mentah empat Periode. Halaman Data Uji menerima satu data secara langsung atau banyak data melalui CSV. Halaman Proses KNN menampilkan perhitungan normalisasi, jarak ke seluruh data latih, lima tetangga

terdekat, dan hasil voting untuk setiap data uji.

Pengujian Sistem

Pengujian membandingkan hasil prediksi sistem ($K = 5$) dengan kelas aktual pada 8 data uji Januari–Agustus 2026. Kelas aktual ditentukan dengan kriteria ambang batas $\Delta\%$ yang sama seperti pada data latih (Tabel 1). Hasil perbandingan disajikan pada Tabel 7.

Tabel 7 Perbandingan Prediksi dan Kelas Aktual

Bulan	$\Delta\%$	Aktual	Prediksi	Ket.
Januari	22,27	Naik	Naik	Sesuai
Februari	8,54	Naik	Naik	Sesuai
Maret	0,79	Stabil	Stabil	Sesuai
April	-2,04	Turun	Turun	Sesuai
Mei	-4,70	Turun	Turun	Sesuai
Juni	1,80	Naik	Naik	Sesuai
Juli	-3,55	Turun	Turun	Sesuai
Agustus	-1,85	Turun	Turun	Sesuai

Seluruh 8 data uji diklasifikasikan dengan tepat tanpa kesalahan. Confusion Matrix pada Tabel 8 membentuk diagonal penuh: 3 data Naik, 4 data Turun, dan 1 data Stabil seluruhnya benar.

Tabel 8 Confusion Matrix

$$\text{Akurasi} = (8 / 8) \times 100\% = 100\%$$

Aktual \ Prediksi	Naik	Turun	Stabil
Naik	3	0	0
Turun	0	4	0
Stabil	0	0	1

Tabel 8 Confusion Matrix
 $\text{Akurasi} = (8 / 8) \times 100\% = 100\%$

Tabel 9 Precision, Recall, dan F1-Score per kelas

Kelas	Precision	Recall	F1-Score
Naik	100,00%	100,00%	100,00%
Turun	100,00%	100,00%	100,00%
Stabil	100,00%	100,00%	100,00%

Analisis Hasil Pengujian

Sistem dengan $K = 5$ memperoleh akurasi 100%, dengan precision, recall, dan F1-score 100% pada ketiga kelas. Hasil ini menunjukkan bahwa persentase perubahan harga rata-rata yang ternormalisasi cukup representatif untuk mengenali tren naik, turun, dan stabil pada data penelitian. Meskipun demikian, hasil ini perlu ditafsirkan hati-hati. Data uji hanya berjumlah 8 dengan kelas Stabil diwakili satu data, dan data latih tidak seimbang (Turun hanya 4 dari 47 data). Sensitivitas terhadap ketidakseimbangan tersebut terlihat pada $K = 9$, ketika akurasi turun menjadi 50%. Selain itu, kelas aktual dan label data latih dibentuk dengan aturan ambang batas yang sama, sehingga akurasi ini mengukur kesesuaian KNN terhadap aturan tersebut dan bukan kemampuan memprediksi pergerakan harga di masa depan.

SIMPULAN

Penelitian ini berhasil menerapkan algoritma K-Nearest Neighbor untuk mengklasifikasikan tren harga emas bulanan ke dalam kelas Naik, Turun, dan Stabil berdasarkan persentase perubahan harga rata-rata. Tahapannya meliputi pengolahan empat nilai Periode menjadi atribut harga, normalisasi Min-Max, penentuan $K = 5$, perhitungan Euclidean

Distance, dan voting mayoritas pada 8 data uji Januari–Agustus 2026. Sistem memperoleh akurasi 100% dengan precision, recall, dan F1-score 100% pada setiap kelas.

Model telah diimplementasikan pada aplikasi web berbasis PHP dan MySQL yang mendukung pengelolaan data latih, normalisasi, penentuan K , input data uji, proses KNN, serta penyajian hasil dan grafik. Karena data uji kecil dan distribusi kelas tidak merata, penelitian selanjutnya disarankan menggunakan periode data uji yang lebih panjang, variasi K yang lebih luas, tambahan fitur seperti kurs rupiah, inflasi, dan suku bunga, serta modul evaluasi otomatis pada sistem.

DAFTAR PUSTAKA

- Akbar, et al. (2022). Penerapan Unified Modeling Language (UML) dalam Perancangan Sistem Informasi.
- Nuraeni, et al. (2024). Penerapan Machine Learning Berbasis K-Nearest Neighbor untuk Klasifikasi Data.
- Ramadhan, et al. (2022). Penerapan Bahasa Pemrograman PHP dalam Pengembangan Sistem Informasi Berbasis Web.
- Sukardi, et al. (2022). Penerapan Basis Data dalam Pengelolaan Sistem Informasi Berbasis Web