

Penerapan *Data Mining* pada Algoritma *Multiple Linear Regression* dalam Peramalan Harga Emas

Intan Rachma Dina^{a,*}, Mula Agung Barata^b, Pelangi Eka Yuwita^c

^{a,b}Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Nahdlatul Ulama Sunan Giri, Bojonegoro

^cTeknik Mesin, Fakultas Sains dan Teknologi, Universitas Nahdlatul Ulama Sunan Giri, Bojonegoro

*correspondence email : intanrachmadina06@gmail.com

Abstract—Gold is a highly valued precious metal, recognized for its ease of maintenance and minimal risk of loss, which makes it a popular choice for investment. However, gold prices are subject to fluctuations influenced by various factors, including the dollar exchange rate, market demand and supply, and monetary crises. Gaining a thorough understanding of these fluctuations is essential for investors looking to minimize losses and maximize profits. The dataset, obtained from Investing.Com, covers the period from January 2019 to December 2024 and consists of 1,548 records with five attributes. We assessed the error rate using Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE). This study aims to forecast gold prices using the Multiple Linear Regression algorithm, implementing K-Fold Cross Validation to enhance model accuracy. The findings indicate RMSE and MAPE values of 695.7909 and 0.27%, respectively, demonstrating that the Multiple Linear Regression algorithm effectively predicts gold prices.

Index Terms—Data mining, multiple linear regression, forecasting, gold, K-fold cross validation

Abstrak—Emas merupakan logam mulia yang sangat dihargai, dikenal karena kemudahan perawatannya dan risiko kehilangan yang minimal, sehingga menjadikannya pilihan yang populer untuk investasi. Namun, harga emas dapat berfluktuasi yang dipengaruhi oleh berbagai faktor, termasuk nilai tukar dolar, permintaan dan penawaran pasar, serta krisis moneter. Memahami fluktuasi ini secara menyeluruh sangat penting bagi investor yang ingin meminimalkan kerugian dan memaksimalkan keuntungan. Dataset yang digunakan dalam penelitian ini diperoleh dari Investing.Com dan mencakup periode dari Januari 2019 hingga Desember 2024, terdiri dari 1.548 catatan dengan lima atribut. Kami menilai tingkat kesalahan menggunakan Root Mean Square Error (RMSE) dan Mean Absolute Percentage Error (MAPE). Penelitian ini bertujuan untuk meramalkan harga emas dengan menggunakan algoritma *Multiple Linear Regression*, serta menerapkan K-Fold Cross Validation untuk meningkatkan akurasi model. Temuan penelitian menunjukkan nilai RMSE dan MAPE masing-masing sebesar 695,7909 dan 0,27%, yang menunjukkan bahwa algoritma *Multiple Linear Regression* secara efektif dapat memprediksi harga emas.

Kata Kunci—Data mining, multiple linear regression, peramalan, emas, K-fold cross validation

I. PENDAHULUAN

Emas, sebagai salah satu logam mulia, memiliki banyak peminat karena nilai investasinya yang stabil dan risiko kerugian yang relatif rendah[1]. Hal ini disebabkan oleh fakta bahwa emas merupakan aset yang menguntungkan dan biasa digunakan untuk investasi, dengan risiko kerugian yang relatif kecil[2]. Berbeda dengan saham, yang nilainya berubah sesuai dengan kondisi pasar[3], emas terbukti mampu bertahan selama krisis ekonomi global, seperti pada tahun 2008 dan 2020, di mana harga emas meningkat secara signifikan sementara aset keuangan lainnya mengalami penurunan[4]. Ketersediaan emas yang terbatas serta tingginya permintaan menjadikannya sebagai sarana investasi untuk menjaga aset kekayaan[5]. Fluktuasi harga emas dipengaruhi oleh beberapa faktor, yaitu nilai tukar dolar, inflasi, tingkat permintaan dan penawaran pasar, serta kebijakan moneter[4].

Investasi merupakan cara untuk memperoleh keuntungan di masa mendatang. Umumnya, investasi dilakukan melalui pembelian aset keuangan seperti obligasi, saham, asuransi, dan emas. Investasi emas merupakan bentuk investasi jangka panjang yang dilakukan dalam periode bulanan bahkan tahunan. Faktanya, emas dapat disimpan di mana saja dan kapan saja karena kadar, bentuk, dan ukuran emas tidak berubah atau mengalami kerusakan. Oleh karena itu, investasi emas menjadi sangat populer[6]. Berdasarkan beberapa faktor tersebut, investasi emas banyak digemari tidak hanya di kalangan pengusaha, tetapi juga oleh masyarakat umum. Namun, ketidakmampuan untuk memahami strategi

investasi yang tepat menyebabkan pengambilan keputusan investasi yang kurang optimal, yang pada akhirnya dapat mengakibatkan kerugian[7]. Oleh karena itu, diperlukan metode untuk memprediksi harga emas di masa mendatang dengan menggunakan teknik *data mining*.

Data mining menjadi salah satu solusi untuk mengatasi permasalahan yang ada, karena *data mining* melakukan proses pencarian pola-pola tersembunyi untuk menggabungkan informasi yang sebelumnya tidak diketahui dari sekumpulan data yang tersedia[8]. Teknik dalam *data mining* terdiri dari lima jenis, yaitu klasifikasi, pengelompokan (*clustering*), asosiasi, peramalan (*forecasting*), dan estimasi. Peramalan diperlukan dalam proses pengambilan keputusan, karena teknik ini mampu meramalkan nilai di masa mendatang berdasarkan data historis[9]. Prediksi atau peramalan dapat dilakukan dengan menggunakan metode *exponential smoothing*, *fuzzy time series*, *moving average*, dan *multiple linear regression*.

Berdasarkan beberapa algoritma *data mining* yang digunakan dalam peramalan, peneliti memilih algoritma *Multiple Linear Regression* untuk prediksi harga emas. Hal ini disebabkan oleh hasil penelitian sebelumnya yang menunjukkan bahwa algoritma tersebut menghasilkan prediksi yang tinggi dengan error yang relatif kecil. Dalam sebuah penelitian yang membandingkan algoritma *Multiple Linear Regression* dan *Support Vector Machine* (SVM) untuk prediksi harga emas, *Multiple Linear Regression* terbukti lebih akurat dengan nilai MAPE sebesar 0,27%, dibandingkan dengan SVM yang memiliki nilai MAPE lebih tinggi. Selain itu, analisis komparatif antara algoritma *data mining Naive Bayes*, *K-Nearest Neighbors*, dan regresi linier menunjukkan bahwa nilai RMSE dan MAE untuk K-NN masing-masing sebesar 0,007 dan 0,006, sedangkan untuk SVM nilainya sebesar 0,006 dan 0,005, dan regresi linier masing-masing sebesar 0,004 dan 0,003[11]. Temuan ini membuktikan bahwa regresi linier lebih akurat dibandingkan dengan K-NN dan SVM.

Penelitian yang membandingkan metode regresi linier berganda dan ARIMA untuk prediksi harga emas menunjukkan bahwa regresi linier berganda lebih akurat[6], dengan nilai RMSE sebesar 4.902.782,346, dibandingkan dengan ARIMA yang memiliki RMSE lebih tinggi, yaitu 5.876.287,332. Selain itu, prediksi harga emas juga pernah diteliti menggunakan metode regresi linier dengan data dari *investing.com*, [12] yang menghasilkan MAE sebesar 4.341,140, lebih akurat dibandingkan dengan RMSE yang mencapai 4.893,132. Penelitian lain mengenai prediksi harga emas menggunakan regresi linier menunjukkan bahwa data yang digunakan bersumber dari laman website rumah123.com, [13] dengan hasil akurasi dan kinerja yang ditunjukkan oleh mean squared error (MSE) sebesar 464,87397, R-squared (R^2) sebesar 0,7818811, dan Mean Absolute Error (MAE) sebesar 313,00083.

Penelitian ini menerapkan algoritma *multiple linear regression* dan menggabungkannya dengan metode evaluasi *K-Fold Cross Validation*. *K-Fold Cross Validation* digunakan untuk mengevaluasi kinerja suatu metode algoritma dengan membagi data sampel secara acak, kemudian mengelompokkan data tersebut sebanyak nilai K[14]. Dalam sebuah penelitian[15], *K-Fold Cross Validation* terbukti efektif untuk memprediksi kinerja dengan hasil fold ke-4 menunjukkan akurasi tertinggi, yaitu 97,29% pada algoritma model naïve bayes. Penelitian lain mengenai *cross validation* juga menunjukkan bahwa *K-Fold Cross Validation* dapat menurunkan hasil sebesar 0,1 terhadap setiap model yang digunakan. Hal ini terjadi karena adanya *overfitting* atau *underfitting* pada setiap model, sehingga permasalahan tersebut dapat diatasi dengan menggunakan *K-Fold Cross Validation*[16]. Penelitian ini dapat menjadi referensi bagi pelaku investasi untuk memprediksi harga emas dengan akurasi tinggi, sehingga mereka dapat mengambil keputusan investasi yang lebih tepat dan meminimalkan risiko kerugian. Selain itu, disarankan untuk mempertimbangkan faktor eksternal seperti inflasi dan nilai tukar mata uang dalam analisis lebih lanjut.

II. METODE

Data mining adalah proses mengidentifikasi pola dan tren dalam data menggunakan teknik statistik dan matematika[17]. Proses ini memungkinkan pelaku bisnis untuk memproses data menjadi informasi penting, sehingga mereka dapat membuat keputusan dengan cepat dan tepat. Beberapa metode *data mining* yang digunakan untuk menemukan pola dan pengetahuan dari informasi antara lain:

1. Classification.

Metode *supervised learning* ini digunakan untuk mengkategorikan data ke dalam kelompok atau jenis yang berbeda berdasarkan data latih atau atribut tertentu. Model klasifikasi memanfaatkan data historis untuk mengkategorikan data baru ke dalam kelas yang sesuai. Contoh metode klasifikasi termasuk *Decision Tree*, jaringan saraf tiruan, dan *Naive Bayes*.

2. Clustering

Metode *unsupervised learning* ini dimanfaatkan untuk mengelompokkan data ke dalam kategori atau kluster berdasarkan kesamaan karakteristik atau fitur tertentu. Data yang serupa akan

dikelompokkan dalam satu kluster. Algoritma *K-Means* dan *Hierarchical Clustering* adalah dua contoh metode *clustering*.

3. *Association*

Dalam *data mining*, metode asosiasi diterapkan untuk mengukur hubungan antara dua atau lebih atribut data[18]. Metode ini digunakan untuk mengidentifikasi hubungan antara variabel atau pola tersembunyi antara item data dalam dataset yang besar. Aturan asosiasi dalam data transaksi dicari dengan algoritma terkenal seperti Apriori, *FP-Growth*, dan *Eclat*.

4. *Prediction*

Prediksi atau *forecasting* adalah teknik *data mining* yang digunakan untuk memperkirakan hasil nilai masa depan berdasarkan data historis. Algoritma yang digunakan dalam prediksi antara lain *Support Vector Machine*, *Random Forest*, dan *Linear Regression*.

5. *Estimation*

Metode estimasi mirip dengan prediksi, namun lebih fokus pada nilai numerik yang diberikan. Estimasi sering digunakan untuk memprediksi hasil atau nilai yang belum diketahui menggunakan data yang sudah ada. Contoh algoritma yang digunakan adalah *Linear Regression*, *Support Vector Machine*, dan *Neural Network*.

Peramalan (*forecasting*) merupakan unsur penting dari perencanaan yang sukses. Peramalan adalah dugaan yang mengacu pada data historis. Teknik ini dapat diterapkan di berbagai bidang, termasuk bisnis, industri, ekonomi, ilmu lingkungan, dan keuangan. Menurut Kamus Besar Bahasa Indonesia (KBBI), prediksi merupakan hasil dari kegiatan meramal atau memperkirakan yang dapat berupa metode ilmiah maupun subjektif[19].

Peramalan telah dimanfaatkan sebagai alat dalam pengambilan keputusan, terutama pada sektor ekonomi dan bisnis, di mana upaya dilakukan untuk meminimalkan kerugian serta memaksimalkan keuntungan[20]. Hasil dari peramalan tidak selalu tepat karena berkaitan dengan kejadian yang tidak pasti di masa depan. Namun, apabila persentase error yang dihasilkan kecil, maka hasil peramalan tersebut dapat dianggap akurat dan mendekati kondisi yang sebenarnya[21].

Algoritma *Multiple Linear Regression* merupakan perluasan dari *Linear Regression* yang melibatkan lebih dari satu variabel bebas. Teknik *Multiple Linear Regression* digunakan untuk mengetahui apakah terdapat pengaruh signifikan dari dua atau lebih variabel bebas ($X_1, X_2, X_3, \dots, X_k$) terhadap variabel terikat (Y). Rumus *Multiple Linear Regression* dapat dinyatakan sebagai berikut:

$$y = a + b_1 X_1 + b_2 X_2 + \dots + b_n X_n \tag{1}$$

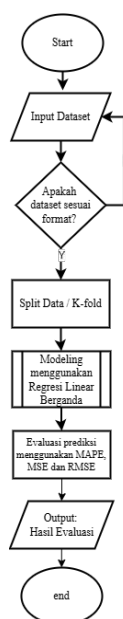
Keterangan:

Y : Variabel Terikat

X_1, X_2 : Variabel bebas

a : Konstanta

b_1, b_2 : Koefisien Regresi



Gambar 1. Alur algoritma *Multiple Linear Regression*

Cross Validation adalah metode validasi yang bertujuan untuk mengevaluasi hasil analisis statistik guna menggeneralisasi kumpulan data independen. *K-fold cross-validation* merupakan teknik validasi yang membagi dataset menjadi k subset dengan ukuran yang sama. Setiap subset secara bergantian digunakan sebagai data uji, sementara k-1 subset lainnya digunakan sebagai data latih. Dalam penelitian ini, *k-fold cross-validation* diterapkan untuk meningkatkan kemampuan generalisasi dalam evaluasi data. Metode ini membagi *dataset* menjadi beberapa bagian, di mana setiap bagian akan secara bergantian berfungsi sebagai data uji, dan evaluasi dilakukan secara berulang[22].

Table 1 Proses 5-Fold Cross Validation

K- Fold	Cross Validation				
1	Validation	Train	Train	Train	Train
2	Train	Validation	Train	Train	Train
3	Train	Train	Validation	Train	Train
4	Train	Train	Train	Validation	Train
5	Train	Train	Train	Train	Validation

Root Mean Square Error (RMSE) merupakan salah satu metrik statistik yang berguna untuk menilai performa suatu model dalam melakukan prediksi pada data numerik. RMSE menghitung rata-rata akar kuadrat dari kesalahan kuadrat antara nilai aktual (observasi sebenarnya) dan nilai prediksi. Metrik ini memberikan gambaran mengenai seberapa besar perbedaan antara prediksi model dan data sebenarnya, dalam satuan yang sama dengan variabel target.

$$RMSE = \sqrt{\frac{1}{n} \sum (Y_t - Y'_t)^2}$$

Mean Absolute Percentage Error (MAPE) merupakan rata-rata perbedaan absolut antara nilai prediksi dan nilai aktual. Nilai ini direpresentasikan dalam bentuk persentase dari nilai aktual dan digunakan untuk mengevaluasi hasil peramalan guna menentukan tingkat akurasi antara nilai prediksi dan nilai aktual[23]. Nilai MAPE dapat dihitung sebagai berikut:

$$MAPE = \frac{1}{n} \sum_{i=0}^n \left| \frac{y' - yi}{yi} \right| * 100 \% X_n$$

Keterangan:

- $y' i$: Hasil Prediksi
 y_i : Data aktual yang diprediksi
 n : Total data yang diprediksi

Tabel 2. Deskripsi Kriteria MAPE

Persentase MAPE	Kriteria
<10%	Sangat Baik
10 > 20%	Baik
20 > 50%	Cukup
> 50%	Buruk

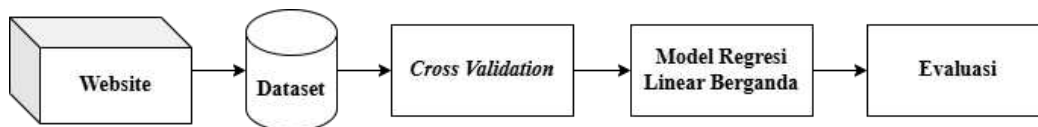
III. HASIL DAN PEMBAHASAN

Dataset yang digunakan adalah data historis harga emas berjangka dari *website investing.com* (<https://www.investing.com/commodities/gold-futures-historical-data>) dengan rentang waktu dari 1 Januari 2019 – 31 Desember 2024 dengan total sebanyak 1549 data. Pada dataset ini terdapat atribut yaitu, *Date*, *Close*, *Open*, *High*, dan *Low*. Berikut dataset yang dapat dilihat pada tabel 1:

Tabel 3. Dataset Harga Emas Berjangka

No	Date	Close	Open	High	Low
1	1/2/2019	1.284100	1.285000	1.291000	1.280600
2	1/3/2019	1.294800	1.288500	1.296900	1.286400
3	1/4/2019	1.285800	1.296500	1.300400	1.278100
4	1/7/2019	1.289900	1.287000	1.297000	1.284100
5	1/8/2019	1.285900	1.289900	1.291400	1.280200
6	1/9/2019	1.292000	1.286100	1.295000	1.280900
7	1/10/2019	1.287400	1.294700	1.298000	1.286700
8	1/11/2019	1.289500	1.287600	1.295700	1.287000
9	1/14/2019	1.291300	1.288000	1.296600	1.287700
10	1/15/2019	1.288400	1.292400	1.294800	1.286500
...
1548	12/31/2024	2.641000	2.620000	2.642000	2.614300

Berdasarkan metode yang diusulkan, berikut skema metode pada Gambar 4 dengan menggunakan algoritma *Multiple Linear Regression*.



Gambar 2. Alur Skenario Penelitian

Pengujian dataset pada algoritma *Multiple Linear Regression* dilakukan dengan evaluasi 5-fold cross validation, yang menghasilkan akurasi dari lima iterasi. Dalam *5-fold cross validation*, dataset dibagi menjadi data latih dan data uji. Proses ini diulang sebanyak lima kali sesuai dengan yang telah ditentukan. Pada tahap validasi ini, dataset dibagi menjadi lima bagian melalui pengacakan data. Model akan diuji sebanyak lima kali, di mana setiap iterasi menggunakan empat subset sebagai data latih dan satu subset sebagai data uji, dengan data yang bergantian berfungsi sebagai data uji. Hasil akurasi dari *5-fold cross validation* dapat dilihat pada Tabel 3 di bawah:

Tabel 4. Hasil Pengujian *Multiple Linear Regression* dengan *5 fold Cross Validation*

K- Fold	MAPE	RMSE
1	0.30%	757.8690
2	0.30%	767.6348
3	0.25%	614.1426
4	0.26%	675.2802
5	0.26%	664.0279
Rata- rata	0.27%	695.7909

Grafik data prediksi dan data aktual dari algoritma *Multiple Linear Regression* dengan *5-fold cross validation* dapat dilihat pada Gambar 5. Grafik tersebut memperlihatkan perbandingan antara nilai aktual dan nilai prediksi harga emas menggunakan metode validasi *K-Fold Cross Validation*. Grafik menunjukkan bahwa garis prediksi (merah) hampir mendekati garis aktual (biru), yang mengindikasikan bahwa model *Multiple Linear Regression* memiliki performa yang baik. Tingkat kesalahan yang kecil, dengan RMSE sebesar 695.7909 dan MAPE sebesar 0.27%, membuktikan akurasi model ini dalam memprediksi harga emas.



Gambar 3. Grafik dengan *5-Fold Cross Validation*

Berikut adalah hasil akurasi metrix evaluasi yang diperoleh menggunakan MAPE terhadap model yang telah di bangun.

```

# Menampilkan hasil rata-rata dari semua folds
print("Hasil Rata-rata K-Fold:")
print(f" RMSE Rata-rata: {np.mean(rmse_scores):.4f}")
print(f" MAPE Rata-rata: {np.mean(mape_scores):.4f}")

Hasil Rata-rata K-Fold:
RMSE Rata-rata: 695.7909
MAPE Rata-rata: 0.2743
  
```

Gambar 4. RMSE dan MAPE dengan 5-fold

IV. KESIMPULAN

Berdasarkan hasil penelitian di atas, penerapan algoritma *Multiple Linear Regression* dengan memanfaatkan metode validasi *5-Fold Cross Validation* untuk peramalan harga emas menghasilkan RMSE dan MAPE masing-masing sebesar 695.7909 dan 0.27%. Hal ini disebabkan oleh metode *5-Fold Cross Validation* yang melakukan *cross check* terhadap dataset secara menyeluruh. Pada penelitian selanjutnya, disarankan untuk membandingkan algoritma *Multiple Linear Regression* dengan algoritma peramalan lainnya, seperti ARIMA dan *Exponential Smoothing*, untuk melihat perbandingan akurasi antar algoritma.

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," *J. Ilm. Wahana Pendidik.*, vol. 8, no. 5, pp. 84–88, 2022, doi: 10.5281/zenodo.6408864.
- [2] F. Ristiano, N. Nurmalasari, and A. Yoraeni, "Implementasi Metode Naive Bayes Untuk Prediksi Harga Emas," *Comput. Sci.*, vol. 1, no. 1, pp. 62–71, 2021, doi: 10.31294/coscience.v1i1.201.
- [3] Hafid Akbar Fikri, "Prediksi Harga Emas Dengan Algoritma Backpropagation," *J. Sains Komput. Inform.*, vol. 7, no. 1, pp. 182–189, 2023.
- [4] N. K. Agusmawati, F. Khoiriyah, and A. Tholib, "prediksi harga emas menggunakan metode lstm dan gru," *JITEKH*, vol. 8, no. 1, pp. 32–36, 2020, doi: 10.35447/jitekh.v8i1.194.
- [5] A. P. Rahma and C. C. Canggih, "Analisis Faktor-Faktor Yang Mempengaruhi Minat Masyarakat Terhadap Investasi Emas," *J. Ekon. dan Bisnis Islam*, vol. 4, no. 2, pp. 98–108, 2021, doi: 10.26740/jekobi.v4n2.p98-108.
- [6] Y. F. Wijaya and A. Triayudi, "Penerapan Data Mining Pada Prediksi Harga Emas dengan Menggunakan Algoritma Regresi Linear Berganda dan ARIMA," *J. Comput. Syst. Informatics*, vol. 5, no. 1, pp. 73–81, 2023, doi: 10.47065/josyc.v5i1.4615.
- [7] M. Fauzi, M. J. Vikri, and S. Wahyudhi, "Sistem Pendukung Keputusan Pemilihan Jenis Investasi Menggunakan Metode Analytical Hierarchy Process (AHP)," ... *Appl. Quantum ...*, pp. 8–13, 2024, [Online]. Available: <https://journal.unugiri.ac.id/index.php/almantiq/article/view/2183%0Ahttps://journal.unugiri.ac.id/index.php/almantiq/article/download/2183/1297>
- [8] R. Takdirillah, "Penerapan Data Mining Menggunakan Algoritma Apriori Terhadap Data Transaksi Sebagai Pendukung Informasi Strategi Penjualan," *Edumatic J. Pendidik. Inform.*, vol. 4, no. 1, pp. 37–46, 2020, doi: 10.29408/edumatic.v4i1.2081.
- [9] R. Apriandi, M. B. Insan, F. Rizmawan, H. A. Haq, K. A. Azizi, and D. D. Priyono, "Perancangan Aplikasi Prediksi Harga Emas, Perak, Dolar, Menggunakan Algoritma Regression Berbasis Web," vol. 10, no. 3, 2022.
- [10] R. S. Sinambela, M. Ula, and A. F. Ulva, "Prediksi Harga Emas Menggunakan Algoritma Regresi Linear Berganda dan Support Vector Machine (SVM)," *J. Sist. dan Teknol. Inf.*, vol. 12, no. 2, p. 253, 2024, doi: 10.26418/justin.v12i2.73386.
- [11] 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.
- [12] W. Andriani, Gunawan, and A. E. Prayoga, "Prediksi Nilai Emas Menggunakan Algoritma Regresi Linear," *J. Ilm. Inform. Komput.*, vol. 28, no. 1, pp. 27–35, 2023, doi: 10.35760/ik.2023.v28i1.8096.
- [13] N. Nuris, "Analisis Prediksi Harga Rumah Pada Machine Learning Metode Regresi Linear," *Explore*, vol. 14, no. 2, pp. 108–112, 2024, doi: 10.35200/ex.v14i2.123.
- [14] R. R. R. Arisandi, B. Warsito, and A. R. Hakim, "Aplikasi Naïve Bayes Classifier (Nbc) Pada Klasifikasi Status Gizi Balita Stunting Dengan Pengujian K-Fold Cross Validation," *J. Gaussian*, vol. 11, no. 1, pp. 130–139, 2022, doi: 10.14710/j.gauss.v11i1.33991.
- [15] W. Wijiyanto, A. I. Pradana, S. Sopingi, and V. Atina, "Teknik K-Fold Cross Validation untuk Mengevaluasi Kinerja Mahasiswa," *J. Algoritma*, vol. 21, no. 1, pp. 239–248, 2024, doi: 10.33364/algoritma/v.21-1.1618.
- [16] W. A. Firmansyah, U. Hayati, and Y. Arie Wijaya, "Analisa Terjadinya Overfitting Dan Underfitting Pada Algoritma Naive Bayes Dan Decision Tree Dengan Teknik Cross Validation," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 7, no. 1, pp. 262–269, 2023, doi: 10.36040/jati.v7i1.6329.
- [17] M. A. Ghofur, A. R. Zayn, and R. N. Faila, "Implementasi Metode K-Means Clustering untuk Menentukan Persediaan Stok Barang pada Toko At-Thullab Tuban," pp. 1–6, 2024.
- [18] M. A. Barata et al., "Perancangan Sistem Electronic Nose Berbasis," pp. 117–126, 2024.
- [19] A. Riyandi, A. Aripin, I. N. Ardiansyah, R. Dany, and Y. Yusrizal, "Analisis Data Mining untuk Prediksi Harga Saham: Perbandingan Metode Regresi Linier dan Pola Historis," *J. Teknol. Sist. Inf.*, vol. 4, no. 2, pp. 278–288, 2023, doi: 10.35957/jtsi.v4i2.5158.
- [20] L. H. Hasibuan and S. Musthofa, "Penerapan Metode Regresi Linear Sederhana Untuk Prediksi Harga Beras di Kota Padang," *JOSTECH J. Sci. Technol.*, vol. 2, no. 1, pp. 85–95, 2022, doi: 10.15548/jostech.v2i1.3802.

- [21] D. R. Deni, M. A. Barata, and Sahri, "Forecasting Metode Single Exponential Smoothing Dalam Meramalkan Penjualan Barang," *J. Inform. Polinema*, vol. 9, no. 4, pp. 435–444, 2023, doi: 10.33795/jip.v9i4.1405.
- [22] S. D. Qirani and I. Sukarsih, "Penerapan Metode K-Nearest Neighbor untuk Prediksi Harga Gas Alam Menggunakan Python," pp. 57–64, 2024.
- [23] 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.

Intan Rachma Dina, Mahasiswa Program Studi Teknik Informatika Fakultas Sains dan Teknologi Universitas Nahdlatul Ulama Sunan Giri Bojonegoro.

Mula Agung Barata, Meraih gelar Sarjana Sains Terapan pada tahun 2017 dari Politeknik Negeri Malang. Kemudian meraih gelar Magister Komputer dari Universitas Dian Nuswantoro pada tahun 2022. Saat ini Penulis menjadi dosen Program Studi Teknik Informatika Universitas Nahdlatul Ulama Sunan Giri Bojonegoro.

Pelanggi Eka Yuwita, Meraih gelar Sarjana Sains pada tahun 2013 dari Universitas Negeri Malang. Kemudian meraih gelar Magister Sains dari Institut Teknologi Sepuluh Nopember pada tahun 2015. Saat ini Penulis menjadi dosen Program Studi Teknik Mesin Universitas Nahdlatul Ulama Sunan Giri Bojonegoro.