



Application of Machine Learning for Premium Rice Price Increase Prediction Using Linear Regression Algorithm

Penerapan *Machine Learning* untuk Prediksi Kenaikan Harga Beras Premium Menggunakan Algoritma Regresi Linier

Arif Tri Widiyatmoko^{1*}, Sufajar Butsianto², Agung Nugroho³

^{1,2,3}Teknik Informatika, Universitas Pelita Bangsa, Indonesia

E-Mail: ¹arif.tri@pelitabangsa.ac.id, ²sufajar@pelitabangsa.ac.id,
³agung@pelitabangsa.ac.id

Received Jun 29th 2025; Revised Jul 13th 2025; Accepted Jul 30th 2025; Available Online Jul 31th 2025, Published Aug 15th 2025

Corresponding Author: Arif Tri Widiyatmoko

Copyright © 2025 by Authors, Published by Institut Riset dan Publikasi Indonesia (IRPI)

Abstract

The volatility of premium rice prices as a staple food commodity requires an accurate prediction solution to aid economic planning. This research applies a Machine Learning algorithm, namely Linear Regression, to predict premium rice price increases. The model was trained using historical price data and evaluated for its performance with MAE (0.244), MSE (0.092), and R-squared (0.893) metrics, showing a fairly good level of accuracy in predicting prices. Furthermore, the successfully developed model was implemented into a Streamlit-based interactive web application. The application allows users to input dates and directly gets predictions of premium rice prices. The results show that Linear Regression is effective in predicting premium rice prices, and the implementation into the Streamlit application successfully provides an easily accessible prediction tool. However, further research could focus on improving the accuracy of the model and exploring other Machine Learning algorithms for commodity price prediction.

Keyword: Linear Regression, Mean Absolute Error, Rice Price Prediction, Streamlit

Abstrak

Ketidakstabilan harga beras premium sebagai komoditas pangan pokok memerlukan solusi prediksi yang akurat untuk membantu perencanaan ekonomi. Penelitian ini menerapkan algoritma Machine Learning, yaitu Regresi Linier, untuk memprediksi kenaikan harga beras premium. Model dilatih menggunakan data historis harga dan dievaluasi kinerjanya dengan metrik MAE (0.244), MSE (0.092), dan R-squared (0.893), menunjukkan tingkat akurasi yang cukup baik dalam memprediksi harga. Selanjutnya, model yang berhasil dikembangkan diimplementasikan ke dalam aplikasi web interaktif berbasis Streamlit. Aplikasi ini memungkinkan pengguna untuk memasukkan tanggal dan secara langsung mendapatkan prediksi harga beras premium. Hasil penelitian menunjukkan bahwa Regresi Linier efektif dalam memprediksi harga beras premium, dan implementasi ke dalam aplikasi Streamlit berhasil menyediakan alat prediksi yang mudah diakses. Meskipun demikian, penelitian lanjutan dapat berfokus pada peningkatan akurasi model dan eksplorasi algoritma Machine Learning lainnya untuk prediksi harga komoditas.

Kata Kunci: Mean Absolute Error, Prediksi Harga Beras, Regresi Linier, Streamlit

1. PENDAHULUAN

Harga beras merupakan salah satu indikator penting dalam ekonomi agrikultural, terutama di negara-negara dengan konsumsi beras yang tinggi seperti Indonesia. Fluktuasi harga beras dapat berdampak signifikan pada stabilitas ekonomi dan kesejahteraan masyarakat [1]. Studi menunjukkan bahwa harga beras pada tingkat produsen dan konsumen memiliki pola transmisi yang berbeda, dengan harga di tingkat produsen lebih rentan terhadap perubahan yang cepat dibandingkan harga di tingkat konsumen [2]. Ketidakpastian ini berdampak pada perencanaan pasar dan kesejahteraan petani serta konsumen, sehingga diperlukan model prediksi yang akurat untuk mengantisipasi lonjakan harga. Prediksi kenaikan harga beras premium dapat membantu pemerintah dan pemangku kepentingan lainnya dalam mengambil langkah-langkah strategis untuk menjaga stabilitas harga dan ketahanan pangan.

Dalam beberapa dekade terakhir, perkembangan teknologi dan ilmu data telah memungkinkan penggunaan algoritma machine learning untuk berbagai tujuan prediktif, termasuk prediksi harga komoditas [3]. Machine learning adalah cara bagi komputer untuk belajar dari data dan membuat prediksi, seperti membedakan spam dari email biasa, tanpa perlu diprogram secara spesifik. Beberapa penelitian telah mengeksplorasi berbagai metode dalam memprediksi harga beras, termasuk model stokastik dan analisis berbasis data historis [4]. Salah satu metode yang sering digunakan dalam analisis prediktif adalah regresi linier, yang terkenal karena kesederhanaan dan interpretabilitasnya [5]. Regresi linier mampu memodelkan hubungan linier antara variabel-variabel prediktor dan variabel target, sehingga dapat digunakan untuk memprediksi harga beras berdasarkan data historis [6].

Regresi linier adalah metode statistika yang paling dasar dan umum digunakan dalam analisis prediktif. Model ini berusaha menemukan garis lurus terbaik yang meminimalkan jarak total antara titik data aktual dan garis prediksi. Menurut Montgomery, regresi linier sederhana dapat digunakan untuk memodelkan hubungan antara satu variabel independen dan satu variabel dependen, sementara regresi linier berganda melibatkan lebih dari satu variabel independen [7]. Dalam konteks prediksi harga beras, regresi linier dapat digunakan untuk memodelkan hubungan antara waktu (hari, bulan, tahun) dan harga beras premium.

Beberapa studi terbaru menunjukkan efektivitas regresi linier dalam prediksi harga komoditas. Misalnya, studi oleh Zhao (2025) menemukan bahwa regresi linier mampu mencerminkan tren harga saham berdasarkan data historis, meskipun akurasi masih dapat ditingkatkan dengan model yang lebih kompleks [8]. Selain itu, penelitian oleh Ijsrem Journal (2024) mengaplikasikan regresi linier untuk memprediksi harga komoditas di sektor pertanian dan menunjukkan bahwa pendekatan ini dapat digunakan secara efektif untuk memodelkan korelasi antar variabel harga [9]. Studi lain oleh Bi et al. (2023) menggunakan regresi linier dalam kombinasi dengan metode lain untuk memprediksi harga minyak bumi dan menunjukkan hasil prediksi yang dapat diterima dalam skenario volatilitas tinggi [10]. Meskipun regresi linier memiliki keterbatasan dalam menangani hubungan non-linier, kesederhanaan dan kecepatan komputasinya menjadikannya pilihan yang baik untuk analisis prediktif awal [11].

Penelitian mengenai prediksi harga dan konsumsi beras di Indonesia telah dilakukan dengan berbagai pendekatan, khususnya metode regresi linear dan algoritma pembelajaran mesin. Pramono et al. [12] menerapkan regresi linear untuk meramalkan harga beras di Kota Bengkulu tahun 2024, dengan memanfaatkan data deret waktu untuk enam jenis beras, seperti IR 64 Lampung sebesar Rp12.202 dan Bulog/Dolog sebesar Rp10.900. Penelitian ini menekankan pentingnya metode peramalan harga yang akurat guna mengatasi fluktuasi harga yang berpengaruh terhadap stabilitas masyarakat. Hasibuan dan Musthofa [13] juga menggunakan regresi linear sederhana untuk memprediksi harga beras di Kota Padang, dengan hasil prediksi yang memiliki tingkat akurasi tinggi, ditunjukkan oleh nilai *Root Mean Square Error* (RMSE) yang sangat rendah.

Sementara itu, Muchtar dan Ratyadi [14] membandingkan metode regresi linear dan Random Forest untuk memprediksi harga beras premium di Jawa Barat. Hasil penelitian menunjukkan bahwa regresi linear mencapai akurasi 95,08%, namun Random Forest memiliki kinerja lebih baik dengan akurasi 98,69%, sehingga dinilai lebih efektif dalam menangani fluktuasi harga beras. Penelitian lain oleh Hanif et al. [15] fokus pada prediksi konsumsi beras dalam sistem kotak beras pintar (smart rice box) berbasis IoT. Dengan regresi linear, model prediksi memiliki nilai *Mean Squared Error* (MSE) sebesar 0,2588 untuk jendela prediksi 43 hari, dan mampu memprediksi kehabisan beras hingga 24 hari sebelumnya.

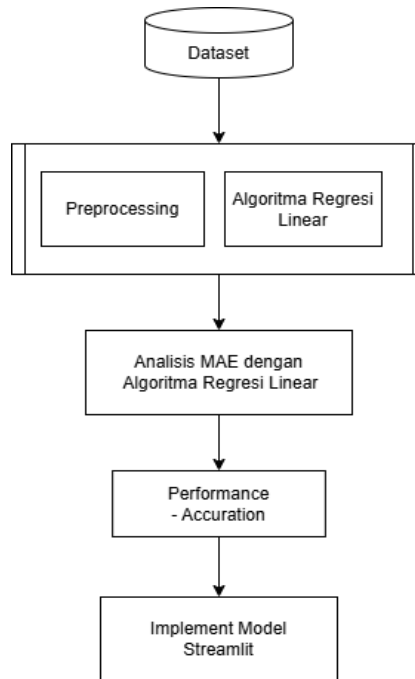
Tujuan dari penelitian ini adalah untuk mengembangkan model prediksi harga beras premium menggunakan algoritma regresi linier, yang diharapkan mampu mengurangi ketidakpastian dalam perencanaan pasar. Dengan menggunakan teknik ini, penelitian ini akan mengevaluasi efektivitas prediksi harga berdasarkan data historis dan faktor-faktor yang mempengaruhi harga beras [16]. Aplikasi prediksi diimplementasikan menggunakan Streamlit, sebuah framework yang memungkinkan pembuatan aplikasi web interaktif dengan mudah.

2. MATERIALS AND METHOD

Metode penelitian yang digunakan adalah penerapan algoritma regresi linier pada prediksi harga beras premium. Gambar 1 menjelaskan tahapan penelitian.

Metode yang digunakan dalam penelitian ini adalah regresi linear, sebuah pendekatan statistik yang digunakan untuk memodelkan hubungan linier antara variabel independen, dalam hal ini waktu, dan variabel dependen, yaitu harga beras. Pendekatan ini dipilih untuk memberikan pemahaman yang lebih baik terhadap tren harga beras seiring berjalannya waktu.

Tahap awal dalam penelitian ini adalah pengumpulan atau penggunaan sebuah dataset. Dataset ini berisi data yang akan dianalisis dan digunakan untuk membangun model regresi linear. Dataset yang digunakan pada penelitian ini adalah data harga beras premium yang di dapatkan dari situs <https://panelharga.badanpangan.go.id> website tersebut adalah webste daftar harga dari badan pangan Indonesia yang khusus nya data disini adalah wilayah Jawa Barat, adapun rentang data yang digunakan dari rentang waktu 1 May 2022 sampai dengan 26 Desember 2023.



Gambar 1. Tahapan Penelitian

Setelah dataset diperoleh, langkah selanjutnya adalah melakukan *preprocessing*. Tahap ini melibatkan berbagai teknik untuk mempersiapkan data agar lebih sesuai untuk pemodelan. Langkah awal dimulai dari pembersihan data. Tujuan utama dari pembersihan data adalah untuk memastikan bahwa data yang akan diproses selanjutnya akurat, relevan, dan bebas dari kesalahan yang signifikan. Kemudian dilanjutkan dengan mengubah kolom tanggal dan mengekstrak fitur waktu menjadi hari, bulan dan tahun. Pemisahan komponen tanggal ini memungkinkan model untuk menangkap pola atau tren yang mungkin terkait dengan hari, bulan, atau tahun tertentu.

Preprocessing selanjutnya adalah menghilangkan *missing value*. *Missing value* adalah masalah umum dalam dataset dan dapat mengganggu proses pelatihan model. Teknik interpolasi digunakan untuk mengisi nilai-nilai yang hilang. Interpolasi adalah metode untuk memperkirakan nilai yang hilang berdasarkan nilai-nilai lain dalam urutan data [17]. Salah satu teknik interpolasi dalam time series dataset adalah interpolasi linier waktu, yaitu dengan melakukan interpolasi linear antara titik-titik data yang valid, tetapi dengan mempertimbangkan jarak waktu antara titik-titik tersebut [18].

$$\hat{y}_{missing} = y_1 + (y_2 - y_1) \times \frac{t_{missing} - t_1}{t_2 - t_1} \quad (1)$$

Dimana $\frac{t_{missing} - t_1}{t_2 - t_1}$ adalah proporsi waktu antara t_1 dan t_2 di mana $t_{missing}$ berada. Interpolasi berdasarkan waktu memberikan perkiraan yang lebih realistis dibandingkan interpolasi linear berbasis indeks, terutama jika kejadian antar waktu tidak seragam [19].

Setelah data dibersihkan, fitur waktu diekstrak, dan *missing value* ditangani, langkah selanjutnya adalah membagi dataset menjadi dua bagian utama, yaitu data *train* dan data *test*. Pembagian data menjadi *train* dan *test* sangat penting untuk menghindari *overfitting* [20], [21], yaitu kondisi di mana model bekerja sangat baik pada data latih tetapi buruk pada data baru. Rasio pembagian dataset yang digunakan adalah 80 data *train* dan 20 data *test*.

Tahapan berikutnya, algoritma regresi linear diterapkan pada data pelatihan yang telah diproses. Regresi linier adalah metode statistik yang digunakan untuk memodelkan hubungan antara variabel independen dan variabel dependen dengan pendekatan garis lurus [22][23]. Dalam konteks prediksi harga beras, regresi linier banyak digunakan untuk menganalisis faktor-faktor yang mempengaruhi harga, seperti produksi, konsumsi, dan faktor eksternal lainnya [24]. Model ini menjadi dasar bagi metode *machine learning* yang lebih kompleks dan sering digunakan sebagai baseline dalam penelitian prediksi harga komoditas pangan [25].

Model regresi linear dapat dinyatakan secara matematis pada Persamaan 2.

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n \quad (2)$$

Dimana $\hat{y}$ adalah nilai prediksi dari variabel dependen. β_0 adalah *intercept* (titik potong garis dengan sumbu y). $\beta_1, \beta_2, \dots, \beta_n$ adalah koefisien regresi untuk masing-masing variabel independen. Sedangkan $x_1, x_2, \dots, x_n$ adalah nilai dari variabel-variabel independen. Algoritma regresi linear akan mencari nilai-nilai koefisien ($\beta_0, \beta_1, \dots, \beta_n$) yang meminimalkan perbedaan antara nilai prediksi ($\hat{y}$) dan nilai sebenarnya dari variabel dependen pada data pelatihan.

Setelah model regresi linear dilatih menggunakan data pelatihan, model tersebut digunakan untuk membuat prediksi pada data pengujian. Penelitian ini juga menitikberatkan pada evaluasi hasil prediksi menggunakan *Mean Absolute Error* (MAE), sebuah metrik yang umum digunakan dalam penelitian prediktif. MAE memberikan gambaran tentang sejauh mana perbedaan antara nilai prediksi dan nilai sebenarnya, yang menjadi indikator tingkat akurasi model.

$$MAE = \frac{1}{m} \sum_{i=1}^m |\hat{y}_i - y_i| \quad (3)$$

Dimana m adalah jumlah total data pengujian, $\hat{y}_i$ adalah nilai prediksi untuk data ke- i , dan y_i adalah nilai sebenarnya untuk data ke- i .

MAE adalah salah satu metrik evaluasi yang digunakan untuk mengukur akurasi model prediksi dengan menghitung rata-rata selisih absolut antara nilai prediksi dan nilai aktual. MAE sering digunakan dalam penelitian prediksi harga beras untuk menilai performa model, di mana nilai MAE yang lebih rendah menunjukkan tingkat kesalahan prediksi yang lebih kecil [26].

Selain MAE, pengujian model regresi linear juga dilakukan dengan metrik evaluasi lainnya yaitu MSE dan R-Square (Koefisien Determinasi). MSE mengukur rata-rata dari kuadrat selisih antara nilai prediksi ($\hat{y}_i$) dan nilai sebenarnya (y_i) dari data pengujian [27]. Secara matematis, MSE dihitung berdasarkan persamaan (4).

$$MSE = \frac{1}{m} \sum_{i=1}^m (\hat{y}_i - y_i)^2 \quad (4)$$

Dimana m adalah jumlah total data pengujian, $\hat{y}_i$ adalah nilai prediksi untuk data ke- i , dan y_i adalah nilai sebenarnya untuk data ke- i .

Nilai MSE yang lebih rendah menunjukkan bahwa model memiliki kesalahan prediksi yang lebih kecil secara rata-rata. MSE sangat sensitif terhadap varians kesalahan [28][29]. Model dengan varians kesalahan yang tinggi akan memiliki nilai MSE yang lebih tinggi, meskipun kesalahan rata-ratanya mungkin tidak terlalu besar. Sedangkan *R-Square* mengukur proporsi varians dalam variabel dependen yang dapat dijelaskan oleh model regresi linear [27]. Nilainya berkisar antara 0 hingga 1. Secara matematis, *R-Square* dihitung berdasarkan persamaan (5).

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} = 1 - \frac{\sum_{i=1}^m (y_i - \hat{y}_i)^2}{\sum_{i=1}^m (y_i - \bar{y})^2} \quad (5)$$

Dimana SS_{res} (*Sum of Squared Residuals*) adalah jumlah kuadrat selisih antara nilai sebenarnya dan nilai prediksi (sama dengan $m \times MSE$). SS_{tot} (*Total Sum of Squares*) adalah jumlah kuadrat selisih antara nilai sebenarnya dan rata-rata nilai sebenarnya ($\bar{y}$).

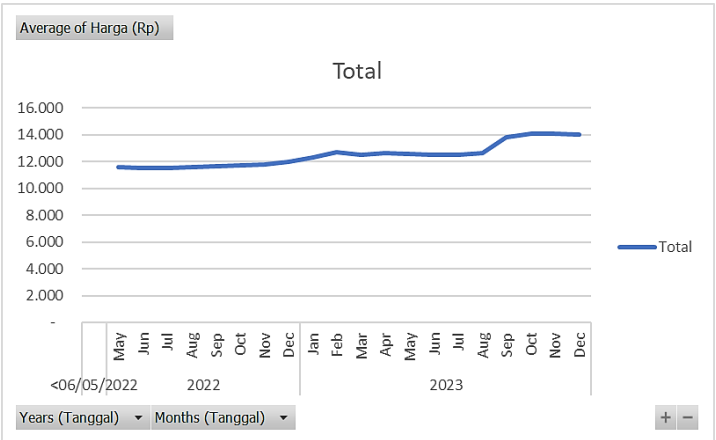
R-Square, atau koefisien determinasi, mengukur proporsi varians variabel dependen yang dapat dijelaskan oleh model regresi linear. Nilainya berkisar antara 0 hingga 1. *R-Square* mendekati 1 menandakan model mampu menjelaskan sebagian besar variasi data, menunjukkan kesesuaian model yang baik. Sebaliknya, nilai mendekati 0 berarti model tidak banyak menjelaskan varians dan mungkin tidak lebih baik dari rata-rata. *R-Square* memberikan indikasi seberapa baik model sesuai dengan data secara keseluruhan.

Tahap terakhir adalah evaluasi performa model secara keseluruhan. Dalam konteks regresi, "akurasi" seringkali diinterpretasikan berdasarkan nilai kesalahan (seperti MAE) dan metrik evaluasi regresi lainnya seperti MSE, *Root Mean Squared Error* (RMSE), atau koefisien determinasi (R^2). Nilai-nilai metrik ini digunakan untuk menilai seberapa tepat dan andal model regresi linear dalam memprediksi nilai variabel dependen berdasarkan variabel-variabel independen.

Selanjutnya model prediksi harga beras premium diimplementasikan pada aplikasi web berbasis Streamlit. Streamlit adalah sebuah framework yang memungkinkan pembuatan aplikasi web interaktif dengan mudah. Streamlit adalah framework berbasis Python yang digunakan untuk membangun aplikasi web interaktif secara cepat dan efisien [30]. Dalam konteks prediksi harga beras, Streamlit sering digunakan untuk membuat dashboard analitik yang memungkinkan pengguna untuk memvisualisasikan data serta menjalankan model prediksi dengan antarmuka yang ramah pengguna [31][32].

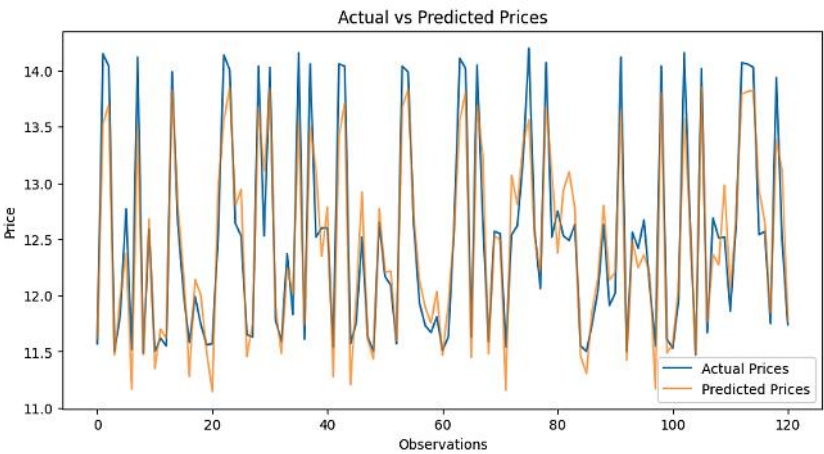
3. RESULTS AND DISCUSSION

Penelitian ini menggunakan data harga beras premium, yang merupakan kualitas beras dengan harga tertinggi di pasaran. Data tersebut mencakup harga pertanggal beras premium dari bulan mei 2022 hingga bulan desember 2023. Untuk memvisualisasikan tren harga, data diolah menjadi grafik seperti yang ditunjukkan pada Gambar 2. Pada grafik ditampilkan tren kenaikan harga rata-rata perbulannya.



Gambar 2. Grafik Harga Rata-rata perbulan

Dari data tersebut kemudian dilakukan analisa terhadap model regresi linier. Analisis grafik perbandingan antara harga aktual dan prediksi menunjukkan bahwa model regresi linier mampu menangkap tren harga beras premium dengan baik, seperti yang terlihat pada Gambar 3. Meskipun terdapat beberapa penyimpangan, secara keseluruhan model telah memberikan hasil yang cukup akurat. Hal ini mengindikasikan potensi model untuk memprediksi harga di masa mendatang, namun perlu adanya penyempurnaan untuk meningkatkan akurasi pada titik-titik tertentu.



Gambar 3. Grafik perbandingan harga aktual dan harga prediksi

Untuk mengevaluasi kinerja model secara lebih kuantitatif, diperlukan metrik evaluasi yang terukur seperti MSE atau MAE. Metrik-metrik ini memberikan gambaran numerik mengenai seberapa besar kesalahan prediksi model dibandingkan dengan nilai sebenarnya. Semakin kecil nilai MSE atau MAE, semakin baik kinerja model dalam melakukan prediksi. Hasil pengukuran kinerja model menggunakan metrik-metrik tersebut dapat dilihat secara rinci pada Tabel 1, yang menyajikan nilai-nilai kuantitatif untuk analisis lebih lanjut. Analisis terhadap nilai-nilai ini memungkinkan pemahaman yang lebih mendalam tentang keakuratan dan keandalan model.

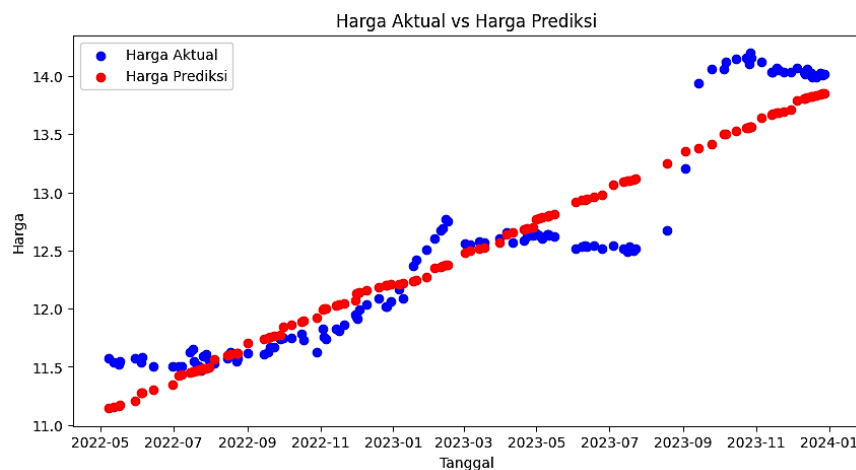
Tabel 1. Matrix Evaluasi

Matrix	Nila
MAE	0.2440308020161823
MSE	0.09207803742428074
R-squared	0.8931928096783356

Dengan nilai MAE sebesar 0.244, rata-rata prediksi model meleset sekitar 0.244 unit dari harga sebenarnya. MSE sebesar 0.092 menunjukkan rata-rata kuadrat kesalahan prediksi. Sementara itu, R-squared sebesar 0.893 atau 89.3% mengindikasikan bahwa model mampu menjelaskan sekitar 89.3% variasi harga aktual. Secara keseluruhan, model ini menunjukkan kinerja yang cukup baik dalam memprediksi harga, dengan sebagian besar variasi harga dapat diprediksi dan rata-rata kesalahan yang relatif kecil.

Berdasarkan metrik evaluasi ini, model prediksi menunjukkan kinerja yang cukup baik. *R-squared* yang tinggi menunjukkan bahwa model memiliki kemampuan yang baik dalam menjelaskan variabilitas harga. Namun, nilai MAE dan MSE yang tidak nol mengindikasikan bahwa masih ada ruang untuk perbaikan dalam mengurangi kesalahan prediksi rata-rata.

Secara keseluruhan, model prediksi yang digunakan mampu menangkap tren umum pergerakan harga aktual, terutama dalam jangka waktu yang lebih pendek. Namun, model tersebut kurang akurat dalam memprediksi fluktuasi harga yang lebih tajam atau perubahan tren yang tiba-tiba, seperti lonjakan harga di sekitar September 2023, seperti yang terlihat pada Gambar 4. Evaluasi lebih lanjut terhadap model prediksi diperlukan untuk meningkatkan akurasinya, terutama dalam mengantisipasi perubahan pasar yang signifikan.



Gambar 4. Grafik perbandingan harga aktual dan prediksi

Setelah melalui tahapan pengembangan dan pengujian yang ketat, model prediksi harga beras premium kini terintegrasi dalam antarmuka pengguna interaktif. Streamlit dipilih sebagai platform pengembangan aplikasi web ini karena kemampuannya yang unggul dalam membangun aplikasi analitik berbasis Python dengan cepat dan efisien. Integrasi ini memungkinkan pengguna untuk berinteraksi langsung dengan model, memasukkan parameter yang relevan, dan memperoleh prediksi harga beras premium secara visual dan intuitif. Kemudahan penggunaan Streamlit mempercepat penyebaran model kepada pengguna akhir, memfasilitasi pemahaman yang lebih baik terhadap dinamika harga beras.

Gambar 5. Visualisasi Program Menggunakan Streamlit

Gambar 5 menampilkan aplikasi Streamlit untuk akses real-time prediksi harga beras premium. Aplikasi ini menyajikan grafik interaktif dan fitur input data baru. Pengguna dapat melihat proyeksi harga berdasarkan model regresi linier yang terintegrasi. Tujuannya adalah mempermudah pengguna dalam memantau dan memahami tren harga beras premium secara dinamis melalui visualisasi dan kemampuan prediksi instan berdasarkan data yang dimasukkan. Dengan adanya Aplikasi ini memberikan kemudahan akses informasi harga beras secara aktual.

4. CONCLUSION

Melalui proses pelatihan dan pengujian, model yang dihasilkan menunjukkan kemampuan yang cukup baik dalam memprediksi harga, tercermin dari metrik evaluasi seperti MAE sebesar 0.244, MSE sebesar 0.092, dan koefisien determinasi (*R-squared*) sebesar 0.893. Nilai *R-squared* yang tinggi mengindikasikan bahwa model mampu menjelaskan sebagian besar variabilitas dalam data harga beras premium. Model prediksi ini berhasil diimplementasikan ke dalam sebuah aplikasi web interaktif menggunakan platform Streamlit. Aplikasi ini menyediakan antarmuka yang sederhana dan mudah digunakan, memungkinkan pengguna untuk memasukkan tanggal spesifik dan secara langsung mendapatkan prediksi harga beras premium untuk tanggal tersebut dalam mata uang Rupiah. Meskipun hasil evaluasi menunjukkan kinerja yang memuaskan, penelitian dan pengembangan lebih lanjut dapat difokuskan pada peningkatan akurasi model, eksplorasi algoritma machine learning yang lebih kompleks, serta integrasi fitur-fitur tambahan yang dapat memperkaya informasi yang disajikan kepada pengguna.

REFERENCES

- [1] Ugi Sugiharto, "Factors Affecting Rice Prices in Indonesia (Production, Consumption, Imports, International Prices, Crop Damage)," *International Journal of Economics and Management Sciences*, vol. 1, no. 3, pp. 329–337, Aug. 2024, doi: 10.61132/ijems.v1i3.183.
- [2] L. R. Shaffitri, E. A. Suryana, and J. F. Sinuraya, "Market integration and rice price transmission in Indonesia," *BIO Web Conf*, vol. 119, p. 02007, Jul. 2024, doi: 10.1051/bioconf/202411902007.
- [3] M. K. Afkar, M. Wali, and Imilda, "Aplikasi Prediksi Produksi Cabai dengan Algoritma C.45 untuk Dinas Pertanian Provinsi Aceh Berbasis Web," *Jurnal Ilmu Komputer dan Teknologi Informasi*, vol. 1, no. 1, pp. 1–13, Mar. 2024, doi: 10.35870/jikti.v1i1.732.
- [4] Respatiwulan, D. Prabandari, Y. Susanti, S. S. Handayani, and Hartatik, "The stochastic model of rice price fluctuation in Indonesia," *J Phys Conf Ser*, vol. 1217, no. 1, p. 012107, May 2019, doi: 10.1088/1742-6596/1217/1/012107.
- [5] D. Maulud and A. M. Abdulazeez, "A Review on Linear Regression Comprehensive in Machine Learning," *Journal of Applied Science and Technology Trends*, vol. 1, no. 2, pp. 140–147, Dec. 2020, doi: 10.38094/jastt1457.
- [6] K. F. Nimon and F. L. Oswald, "Understanding the Results of Multiple Linear Regression," *Organ Res Methods*, vol. 16, no. 4, pp. 650–674, Oct. 2013, doi: 10.1177/1094428113493929.
- [7] D. C. Montgomery, E. A. Peck, and G. G. Vining, *Introduction to linear regression analysis*. John Wiley & Sons, 2021.
- [8] Y. Zhao, "Stock price prediction based on linear regression," *Theoretical and Natural Science*, vol. 84, no. 1, pp. 85–90, Feb. 2025, doi: 10.54254/2753-8818/2025.21203.
- [9] I. Journal, "Forecasting Commodity Prices," *Interantional Journal Of Scientific Research In Engineering And Management*, vol. 08, no. 01, pp. 1–13, Jan. 2024, doi: 10.55041/ijrsrem28035.
- [10] J. Bi, E. Li, and Y. Luo, "Petroleum Price Prediction Based on the Linear Regression and Random Forest," *Applied and Computational Engineering*, vol. 8, no. 1, pp. 292–296, Aug. 2023, doi: 10.54254/2755-2721/8/20230170.
- [11] S. D. Saputra and A. D. Widianoro, "BBCA Stock Price Prediction Using Linear Regression Method," *International Journal of Artificial Intelligence and Science*, vol. 1, no. 1, pp. 25–36, Sep. 2024, doi: 10.63158/IJAIS.v1.i1.7.
- [12] Azhar Dyo Pramono, Herlina Latipa Sari, and Ila Yati Beti, "Penerapan Regresi Linear dalam Perkiraan Harga Beras di Kota Bengkulu," *VISA: Journal of Vision and Ideas*, vol. 4, no. 2, Jul. 2024, doi: 10.47467/visa.v4i2.3662.
- [13] L. H. Hasibuan and S. Musthofa, "Penerapan Metode Regresi Linear Sederhana Untuk Prediksi Harga Beras di Kota Padang," *JOSTECH: Journal of Science and Technology*, vol. 2, no. 1, pp. 85–95, Mar. 2022, doi: 10.15548/jostech.v2i1.3802.
- [14] I. R. Muchtar and A. Afiyati, "Comparison of Linear Regression and Random Forest Algorithms for Premium Rice Price Prediction (Case Study: West Java)," *Jurnal Indonesia Sosial Teknologi*, vol. 5, no. 7, pp. 3122–3132, Jul. 2024, doi: 10.59141/jist.v5i7.1184.
- [15] M. Hanif, M. Abdurrohman, and A. G. Putrada, "Rice consumption prediction using linear regression method for smart rice box system," *Jurnal Teknologi dan Sistem Komputer*, vol. 8, no. 4, pp. 284–288, Oct. 2020, doi: 10.14710/jtsiskom.2020.13353.
- [16] Rahmat Hidayat and Irawan Wibisonya, "Rice Price Prediction with Long Short-Term Memory (LSTM) Neural Network," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 8, no. 5, pp. 658–664, Oct. 2024, doi: 10.29207/resti.v8i5.6041.
- [17] M. Benckekroun, B. Chevallier, V. Zalc, D. Istrate, D. Lenne, and N. Vera, "The Impact of Missing Data on Heart Rate Variability Features: A Comparative Study of Interpolation Methods for Ambulatory Health Monitoring," *IRBM*, vol. 44, no. 4, p. 100776, Aug. 2023, doi: 10.1016/j.irbm.2023.100776.

- [18] T. Decorte *et al.*, “Missing Value Imputation of Wireless Sensor Data for Environmental Monitoring,” *Sensors*, vol. 24, no. 8, p. 2416, Apr. 2024, doi: 10.3390/s24082416.
- [19] A. M. Alonso, A. E. Sipols, and S. Quintas, “A single-index model procedure for interpolation intervals in time series,” *Comput Stat*, vol. 28, no. 4, pp. 1463–1484, Aug. 2013, doi: 10.1007/s00180-012-0355-8.
- [20] 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, p. e2245, Sep. 2024, doi: 10.7717/peerj-cs.2245.
- [21] V. Kholiev and O. Barkovska, “Analysis of the of training and test data distribution for audio series classification,” *Information and control systems at railway transport*, vol. 28, no. 1, pp. 38–43, Mar. 2023, doi: 10.18664/ikszt.v28i1.276343.
- [22] X. Su, X. Yan, and C. Tsai, “Linear regression,” *WIREs Computational Statistics*, vol. 4, no. 3, pp. 275–294, May 2012, doi: 10.1002/wics.1198.
- [23] D. Maulud and A. M. Abdulazeez, “A Review on Linear Regression Comprehensive in Machine Learning,” *Journal of Applied Science and Technology Trends*, vol. 1, no. 2, pp. 140–147, Dec. 2020, doi: 10.38094/jastt1457.
- [24] I. R. Muchtar and A. Afiyati, “Comparison of Linear Regression and Random Forest Algorithms for Premium Rice Price Prediction (Case Study: West Java),” *Jurnal Indonesia Sosial Teknologi*, vol. 5, no. 7, pp. 3122–3132, Jul. 2024, doi: 10.59141/jist.v5i7.1184.
- [25] M. L. Zakaria, S. A. Wibowo, and I. Kurniawan, “Implementation of Temporal Fusion Transformer Optimized by Grey Wolf Optimizer In Predicting Rice Price In Bandung Regency,” in *2024 8th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE)*, IEEE, Aug. 2024, pp. 127–132. doi: 10.1109/ICITISEE63424.2024.10730751.
- [26] I. N. G. A. M. Wardhiana *et al.*, “Comparative Study of Statistical, Machine Learning, and Deep Learning for Rice Retail Price Forecasting in West Java,” in *2024 2nd International Conference on Technology Innovation and Its Applications (ICTIIA)*, IEEE, Sep. 2024, pp. 1–6. doi: 10.1109/ICTIIA61827.2024.10761587.
- [27] D. Chicco, M. J. Warrens, and G. Jurman, “The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation,” *PeerJ Comput Sci*, vol. 7, p. e623, Jul. 2021, doi: 10.7717/peerj-cs.623.
- [28] E. Fradinata, S. Suthummanon, and W. Suntiamorntut, “Forecasting Determinant of Cement Demand in Indonesia with Artificial Neural Network,” *Journal of Asian Scientific Research*, vol. 5, no. 7, pp. 373–384, 2015, doi: 10.18488/journal.2/2015.5.7/2.7.373.384.
- [29] C. Wang *et al.*, “Predicting Plant Growth and Development Using Time-Series Images,” *Agronomy*, vol. 12, no. 9, p. 2213, Sep. 2022, doi: 10.3390/agronomy12092213.
- [30] J. M. Nápoles-Duarte, A. Biswas, M. I. Parker, J. P. Palomares-Baez, M. A. Chávez-Rojó, and L. M. Rodríguez-Valdez, “Stmol: A component for building interactive molecular visualizations within streamlit web-applications,” *Front Mol Biosci*, vol. 9, Sep. 2022, doi: 10.3389/fmolb.2022.990846.
- [31] L. Setivani, H. H. Handayani, and W. A. Geraldine, “Rice Price Forecasting Using GridSearchCV and LSTM,” in *2023 International Conference on Modeling & E-Information Research, Artificial Learning and Digital Applications (ICMERALDA)*, IEEE, Nov. 2023, pp. 127–131. doi: 10.1109/ICMERALDA60125.2023.10458179.
- [32] M. P. Keerthi, G. S. Reddy, V. S. Raghava, and K. B. Reddy, “Streamlit Interface for Multiple Disease Diagnosis,” *Int J Res Appl Sci Eng Technol*, vol. 11, no. 2, pp. 1159–1164, Feb. 2023, doi: 10.22214/ijraset.2023.49166.