

Rice Production Prediction System Using Multiple Linear Regression for Food Security Optimization in Minahasa Regency

¹Oliver Simon Hardianto Raharusun, ²Alfiansyah Hasibuan, ³Gladly Caren Rorimpandey
^{1, 2, 3}Prodi Teknik Informatika, Fakultas Teknik, Universitas Negeri Manado
¹oliverraharusun@gmail.com, ²alfiansyahhasibuan@unima.ac.id,
³gladlyrorimpandey@unima.ac.id

Submit : 14 Sept 25 | Diterima : 11 Okt 2025 | Terbit : 19 Okt 2025

ABSTRAK

Sistem prediksi produksi beras menggunakan algoritma *Multiple Linear Regression* dirancang untuk membantu pemerintah daerah Kabupaten Minahasa dalam mengoptimalkan ketahanan pangan melalui perencanaan yang lebih akurat. Sebanyak 19 baris dan 7 kolom data histori produksi beras dikumpulkan, kemudian dilakukan tahap *preprocessing* sebelum dimodelkan menggunakan algoritma *Multiple Linear Regression*. Model dievaluasi dengan mengukur nilai koefisien determinasi (R^2) dan *Mean Absolute Percentage Error* (MAPE) untuk memastikan tingkat akurasi prediksi. Implementasi sistem diwujudkan dalam bentuk aplikasi berbasis web menggunakan *Streamlit*, yang memungkinkan pengambil kebijakan memperoleh prediksi produksi beras secara *real-time*.

Kata Kunci: Sistem Prediksi, Produksi Beras, Regresi Linear Berganda, Ketahanan Pangan, *Streamlit*

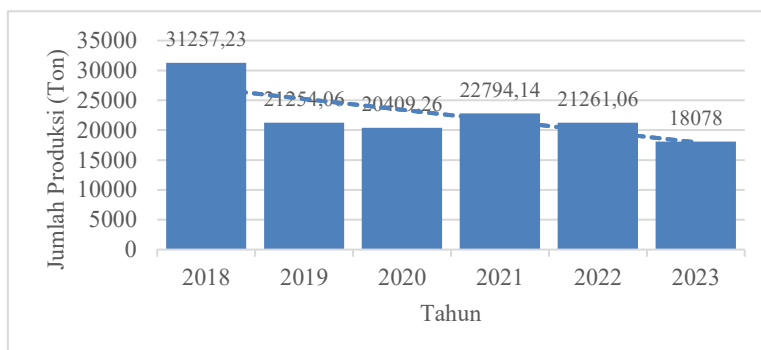
PENDAHULUAN

Ketahanan pangan merupakan isu strategis global yang semakin krusial akibat pertumbuhan populasi, perubahan iklim, dan keterbatasan sumber daya pertanian. Organisasi Pangan dan Pertanian (FAO), badan khusus di bawah Perserikatan Bangsa-Bangsa (PBB), memperkirakan bahwa produksi pangan harus meningkat sekitar 60%–70% pada tahun 2050 untuk memenuhi kebutuhan pangan penduduk dunia yang diperkirakan mencapai 9,7 miliar jiwa (www.detik.com, 2023). Tantangan ini semakin kompleks dengan adanya ancaman perubahan iklim, yang berdampak pada produktivitas pertanian, ketersediaan lahan, serta efisiensi sistem distribusi pangan. Perubahan iklim menyebabkan cuaca ekstrem, seperti kekeringan dan banjir, yang mengganggu siklus tanam dan panen, sehingga mengancam stabilitas pasokan pangan global (Ajjaz dkk., 2025).

Ketahanan pangan merupakan salah satu isu strategis global yang terus mendapatkan perhatian besar, terutama di negara berkembang seperti Indonesia. Di Indonesia, sektor pertanian, khususnya produksi beras, memainkan peran vital dalam memenuhi kebutuhan pangan lebih dari 280 juta penduduknya. Beras sebagai makanan pokok utama, sangat mempengaruhi stabilitas sosial dan ekonomi. Namun, dalam beberapa tahun terakhir, sektor pertanian Indonesia mengalami tantangan besar, termasuk ketidakstabilan produksi beras yang disebabkan oleh berbagai faktor.

Kabupaten Minahasa merupakan salah satu wilayah di Sulawesi Utara, Indonesia yang memiliki ketergantungan tinggi terhadap produksi beras sebagai kebutuhan pangan utama masyarakat. Dalam beberapa tahun terakhir, terjadi ketidakstabilan pada tingkat produksi di daerah ini bahkan tren produksi beras cenderung menurun seperti pada Gambar 1.1. Kondisi ini berdampak pada kenaikan harga beras, yang kemudian memicu masalah ketersediaan pangan bagi masyarakat. Seperti yang dilaporkan oleh Ibu Sandra Karlos Arsjad, Kepala Bidang Dinas Pangan Kabupaten Minahasa dalam wawancaranya pada Agustus 2024, mengatakan bahwa fluktuasi harga beras menjadi salah satu indikator permasalahan stok dan produksi yang belum sepenuhnya stabil. Semakin tinggi stok yang tersedia, maka harga beras cenderung lebih stabil dan rendah, namun

ketika stok berkurang, harga beras meningkat secara signifikan. Permasalahan ini mengindikasikan adanya ketergantungan pada manajemen produksi dan distribusi yang lebih baik agar ketahanan



pangan dapat terjaga.

Gambar 1 Diagram Produksi Beras Kab. Minahasa

Untuk dalam upaya menjaga ketahanan pangan di daerah Minahasa, peneliti merancang dan membangun sistem prediksi produksi beras menggunakan algoritma *Multiple Linear Regression*. Algoritma ini dipilih karena keunggulannya dalam prediksi, terbukti dengan nilai *R-squared* yang tinggi dalam berbagai penelitian sebelumnya. Misalnya *Multiple Linear Regression* mencapai nilai *R-squared* 0.99 atau 99% dalam prediksi hasil panen wortel (S dkk., 2024) dan 1 atau 100% dalam memprediksi hasil panen kacang kedelai di Kecamatan Surade (Raehan dkk., 2024).

Implementasi sistem ini diharapkan mampu memberikan estimasi produksi beras yang akurat sehingga diharapkan mampu menjadi acuan dalam pengambilan keputusan strategis terkait ketahanan pangan. Sistem ini tidak hanya diharapkan dapat menjaga ketahanan pangan di Kabupaten Minahasa, tetapi juga diharapkan meningkatkan stabilitas harga beras.

TINJAUAN PUSTAKA

Sistem

Menurut Azhar Susanto, Sistem adalah kumpulan dari komponen apapun fisik ataupun bukan fisik yang terhubung satu dan lainnya dan bersama-sama agar mencapai tujuan tertentu (Susanto, 2017). Berdasarkan dari pengertian diatas bisa disimpulkan sistem merupakan suatu bagian atau unsur yang saling tersinkron menjadi satu-kesatuan bersama-sama dan saling berkaitan untuk mencapai suatu sasaran/tujuan tertentu yang sama.

Streamlit

Streamlit adalah sebuah *framework open-source* yang membantu pengembang membangun aplikasi web yang dapat diakses menggunakan bahasa pemrograman *Python*. Dengan menggunakan *framework* ini, pengembang dapat membangun aplikasi dengan mudah menggunakan skrip *Python* yang sederhana tanpa perlu memiliki pengetahuan mendalam tentang pengembangan *web frontend*. Dengan menggunakan Streamlit, Anda dapat membuat antarmuka pengguna (UI) responsif dengan cepat yang menampilkan visualisasi data, tabel, grafik, dan elemen interaktif lainnya. Selain itu, fitur-fiturnya mempermudah penyebaran aplikasi sehingga pengguna lain dapat dengan mudah mengaksesnya secara online. Karena itu, pengembang *data science* dan *machine learning* telah memilih Streamlit untuk mempercepat pembuatan dan distribusi aplikasi berbasis web (Azhar dkk., 2024).

Multiple Linear Regression

Linear Regression atau Regresi Linear adalah salah satu metode regresi yang paling sederhana, namun cukup handal dalam performanya. *Linear Regression* digunakan untuk memprediksi suatu variabel berdasarkan variabel lainnya. Variabel yang akan diprediksi disebut

variabel dependen (target), sedangkan variabel yang digunakan untuk memprediksi disebut variabel independen (prediktor)(Sarno dkk., 2023). Bila variabel yang akan memprediksi atau variabel independennya berjumlah lebih satu, maka digunakan *Multiple Linear Regression*.

Koefisien Determinasi (*R-squared*)

Koefisien determinasi (*R-squared*) merupakan suatu ukuran metrik yang digunakan untuk mengevaluasi sejauh mana suatu model mampu menjelaskan proporsi variasi yang terjadi pada variabel dependen. Nilai R^2 berkisar antara 0 hingga 1, di mana nilai yang semakin mendekati 1 mengindikasikan bahwa variabel-variabel independen dalam model memiliki kemampuan yang tinggi dalam menjelaskan variasi yang terdapat pada variabel dependen (Zaka dkk., 2017).

Mean Absolute Percentage Error (MAPE)

Untuk menilai tingkat akurasi hasil prediksi, digunakan metode *Mean Absolute Percentage Error* (MAPE), yang merupakan rata-rata dari persentase kesalahan absolut antara nilai aktual dan nilai hasil prediksi. MAPE digunakan sebagai indikator evaluasi dalam analisis peramalan, khususnya pada data runtun waktu, dan dinyatakan dalam bentuk persentase. Semakin kecil nilai MAPE yang dihasilkan, maka semakin tinggi tingkat kedekatan hasil prediksi terhadap data aktual (Sumari dkk., 2020).

Tabel 1 Interpretasi MAPE

MAPE (%)	Interpretasi
< 10	Prediksi sangat akurat
10-20	Prediksi yang baik
20-50	Prediksi yang layak
> 50	Prediksi yang tidak akurat

METODE PENELITIAN

Lokasi dan Waktu Penelitian

Lokasi penelitian di Dinas Pangan Kabupaten Minahasa yang berlokasi di Kelurahan Kembuan, Kecamatan Tondano Utara, Kabupaten Minahasa, Sulawesi Utara. Penelitian ini dimulai dari tanggal 24 Agustus 2024 sampai dengan 28 Desember 2024.

Teknik Pengumpulan Data

Data dikumpulkan menggunakan 3 metode, yaitu:

1. Studi Literatur, yaitu proses ini melibatkan pengumpulan referensi dari berbagai sumber seperti artikel online, buku/jurnal ilmiah, dan materi pembelajaran yang tersedia di internet.
2. Studi Dokumentasi, yaitu mengumpulkan data histori produksi beras dari website resmi Badan Pusat Statistik Sulawesi Utara dan Badan Pusat Statistik Kabupaten Minahasa.
3. Wawancara, yaitu untuk mendapatkan informasi terperinci mengenai informasi produksi dan stok beras di Kabupaten Minahasa.

Algoritma *Multiple Linear Regression*

Pada penelitian ini, metode yang digunakan yaitu metode regresi menggunakan algoritma *Multiple Linear Regression*. MLR (*Multiple Linear Regression*) adalah Regresi Linear Berganda adalah regresi yang memiliki satu variabel dependen (tidak bebas) dan lebih dari satu variabel independen (bebas) (Adiguno dkk., 2022). Secara umum, model persamaan *multiple linear regression* dapat dinyatakan sebagai berikut.

$$Y = a + b_1X_1 + b_2X_2 + \dots + b_nX_n$$

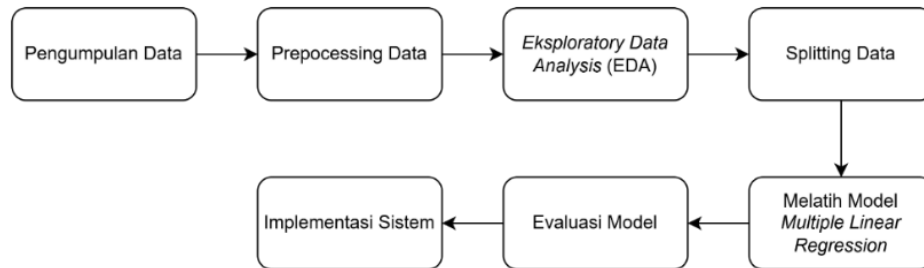
Dimana:

- Y : Variabel dependen (variabel terikat)
X : Variabel independen (variabel bebas)
a : Nilai konstanta
b : Nilai koefisien regresi
X₁ : Variabel independen pertama

X_2 : Variabel independen kedua
 X_n : Variabel independen ke-n

Tahapan Penelitian

Penelitian ini dilakukan melalui beberapa tahap untuk merancang dan membangun sistem prediksi produksi beras menggunakan algoritma *Multiple Linear Regression* di Kabupaten Minahasa. Tahapan-tahapan dalam metode penelitian yang akan dilaksanakan adalah sebagai berikut.



Gambar 2 Tahapan Penelitian

1. **Pengumpulan Data**
Dataset yang digunakan untuk pelatihan model adalah dataset histori produksi beras sebanyak 19 baris dan 7 kolom data yang diperoleh dari *website* resmi Badan Pusat Statistik (BPS) Kabupaten Minahasa dan Provinsi Sulawesi Utara.
2. **Preprocessing Data**
Pada tahap ini dilakukan pembersihan data yaitu dengan memeriksa *duplicate data*, *missing value*, dan *outlier*. Setelah itu, dilakukan *treatment* yang sesuai.
3. **Exploratory Data Analysis (EDA)**
Pada tahap ini dilakukan analisis awal pada dataset guna memahami pola, distribusi, dan korelasi antar variabel (*corellation heatmap*).
4. **Splitting Data**
Pada tahap ini, dataset dibagi menjadi dua yaitu *data training* dan *data testing* dengan masing-masing rasio 80:20.
5. **Pelatihan Model**
Pada tahap ini, *data training* digunakan untuk melatih model *multiple linear regression* menggunakan *library python scikit-learn*.
6. **Evaluasi Model**
Setelah model dilatih, kemudian model diuji menggunakan metrik-metrik evaluasi seperti *R-squared* dan *Mean Absolute Percentage Error (MAPE)*.
7. **Implementasi Sistem**
Setelah model dievaluasi, kemudian model tersebut diintegrasikan ke dalam sebuah sistem berbasis web menggunakan *streamlit*.

HASIL DAN PEMBAHASAN

Hasil *Preprocessing Data*

Setelah melalui proses pengumpulan, pemilihan dan pembersihan data, peneliti memperoleh 19 baris dan 6 kolom data yang akan dibagi menjadi data *training* dan data *testing*. Hasil data yang telah dibersihkan adalah seperti di bawah ini:

Tabel 2 Hasil *Preprocessing Data*

Produksi Beras (Ton)	Luas Panen (Ha)	Jumlah Curah Hujan (mm)	Jumlah Hari Hujan (Hari)	Rata-rata suhu (Celcius)	Rata-rata Kelembapan Udara (%)
68322	15560	2177.868	200.15	23.372	84.815
51531	10998	2177.868	200.15	23.372	84.815
64292	13050	2177.868	200.15	23.372	84.815
57859	11800	2177.868	200.15	23.372	84.815
65941	13500	2177.868	200.15	23.372	84.815
70482	13793	2177.868	200.15	23.372	84.815
82448	16473	2177.868	200.15	23.372	84.815
82475	16535	2207.757	200.15	23.372	84.815
86456	16162	2706.422	200.15	23.372	84.815
89455	18032	2237.252	200.15	23.372	84.815
86004	17511	1459.2	200.15	23.372	84.815
86004	17511	2345.835	200.15	23.372	84.815
86004	17511	2874.26	200.15	23.372	84.815
31257.23	11662	1858	189	22.9	86.64
21254.06	8764.31	1610.862	149.75	22.8	86.04
20409.26	8024.89	2136	232	24.79	80.67
22794.14	8466.15	2306	245	23.1	84.815
21261.06	7889.73	2595.824	200.15	23.372	84.815
18078	7053	1797	185	23.27	85.91

Hasil Evaluasi Model

Setelah model *multiple linear regression* dilatih, kemudian performa model tersebut dievaluasi menggunakan metrik-metrik evaluasi seperti *R-squared* dan *Mean Absolute Percentage Error (MAPE)*. Berikut hasil evaluasi model *multiple linear regression*:

Tabel 3 Hasil Evaluasi Model

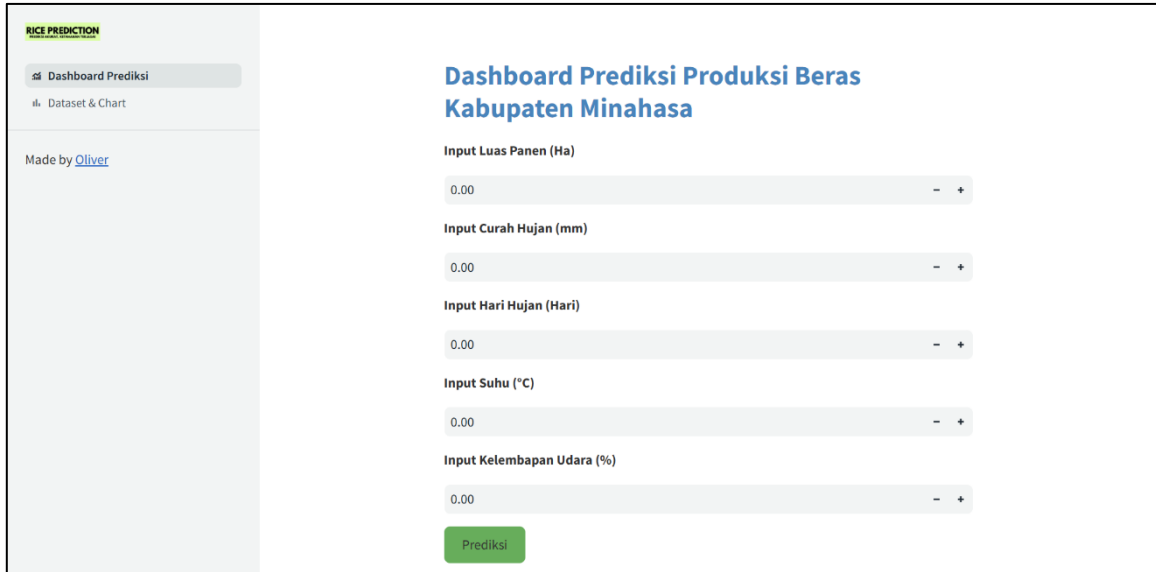
	<i>R-squared (R²)</i>	MAPE
<i>Multiple Linear Regression</i>	0.667	10.61%

Dalam *R-squared*, semakin nilai mendekati 1 maka model semakin bagus. Begitupun sebaliknya, jika model menjadi nilai 1 atau mendekati 0 maka model semakin buruk. Dengan nilai *R-squared* sebesar 0.667, model *multiple linear regression* mampu menjelaskan 66,7% variasi dalam data produksi beras. Hal ini menunjukkan model cukup baik dalam melakukan prediksi.

Dengan *Mean Absolute Percentage Error (MAPE)* sebesar 10.61% menunjukkan bahwa rata-rata kesalahan prediksi hanya sekitar 10.61% dari nilai aktual. Nilai MAPE yang berada diantara 10%-20% menunjukkan bahwa model memiliki kemampuan prediksi yang baik (Sumari dkk., 2020).

Implementasi Sistem

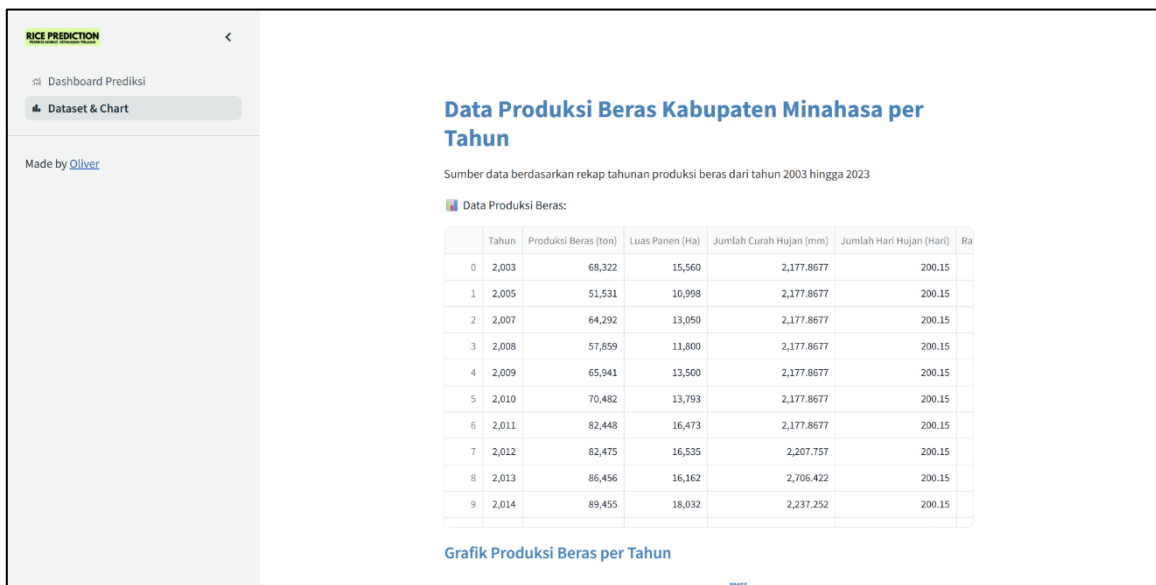
Pada halaman dashboard prediksi sistem ini, pengguna/*user* dapat langsung melakukan prediksi dengan menginput beberapa variabel prediktor seperti jumlah luas panen (Ha), curah hujan (mm), hari hujan, suhu (°C), dan kelembapan udara (%) seperti pada Gambar 3.



The screenshot shows a web interface titled "RICE PREDICTION". On the left is a sidebar with "Dashboard Prediksi" and "Dataset & Chart". The main content area is titled "Dashboard Prediksi Produksi Beras Kabupaten Minahasa". It contains five input fields, each with a value of 0.00 and a minus/plus toggle: "Input Luas Panen (Ha)", "Input Curah Hujan (mm)", "Input Hari Hujan (Hari)", "Input Suhu (°C)", and "Input Kelembapan Udara (%)". A green "Prediksi" button is at the bottom.

Gambar 3 Tampilan Halaman Dashboard Prediksi

Pada halaman Dataset & Chart sistem ini pengguna/user bisa melihat dataset histori produksi beras dari tahun 2003-2023 (Gambar 4).

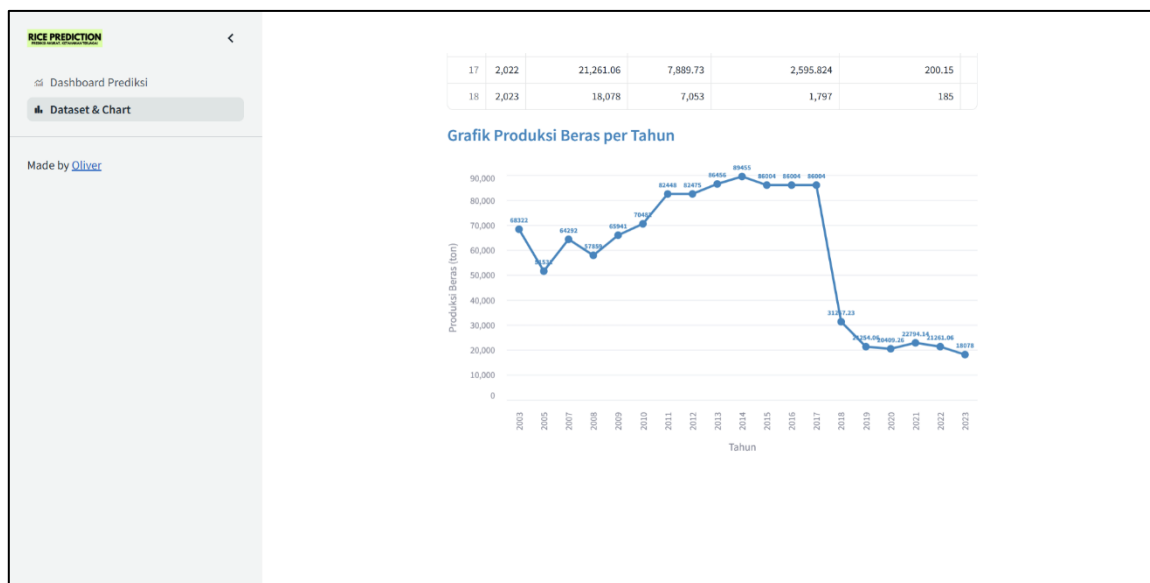


The screenshot shows the "Dataset & Chart" page. It features a table titled "Data Produksi Beras Kabupaten Minahasa per Tahun". Below the title, it states "Sumber data berdasarkan rekap tahunan produksi beras dari tahun 2003 hingga 2023". The table has 6 columns: Tahun, Produksi Beras (ton), Luas Panen (Ha), Jumlah Curah Hujan (mm), Jumlah Hari Hujan (Hari), and Ra. The data rows are numbered 0 to 9, representing years from 2003 to 2014.

Tahun	Produksi Beras (ton)	Luas Panen (Ha)	Jumlah Curah Hujan (mm)	Jumlah Hari Hujan (Hari)	Ra
0	2,003	68,322	15,560	2,177.8677	200.15
1	2,005	51,531	10,998	2,177.8677	200.15
2	2,007	64,292	13,050	2,177.8677	200.15
3	2,008	57,859	11,800	2,177.8677	200.15
4	2,009	65,941	13,500	2,177.8677	200.15
5	2,010	70,482	13,793	2,177.8677	200.15
6	2,011	82,448	16,473	2,177.8677	200.15
7	2,012	82,475	16,535	2,207.757	200.15
8	2,013	86,456	16,162	2,706.422	200.15
9	2,014	89,455	18,032	2,237.252	200.15

Gambar 4 Tampilan Halaman Dataset & Chart (Dataset)

Pada halaman Dataset & Chart sistem ini pengguna/user bisa melihat tren produksi beras dari tahun ke tahun, mulai dari 2003-2023 (Gambar 5).



Gambar 5 Tampilan Halaman Dataset & Chart (Chart)

KESIMPULAN

Hasil penelitian menunjukkan bahwa model regresi linier berganda yang dibangun memiliki nilai *R-Squared* sebesar 0,667. Artinya, 66,7% variasi produksi beras dapat dijelaskan oleh faktor luas panen, curah hujan, hari hujan, suhu, dan kelembapan, sementara 33,3% sisanya dipengaruhi oleh faktor lain di luar penelitian. Nilai *Mean Absolute Percentage Error* (MAPE) sebesar 10,61% menegaskan bahwa model ini cukup andal dalam melakukan prediksi. Model yang dihasilkan juga telah diimplementasikan pada sistem berbasis *web* menggunakan *framework Streamlit*, sehingga diharapkan mendukung pemerintah dan pemangku kepentingan dalam perencanaan kebijakan ketahanan pangan berbasis data.

REFERENSI

- Adiguno, S., Syahra, Y., & Yetri, M. (2022). Prediksi Peningkatan Omset Penjualan Menggunakan Metode Regresi Linier Berganda. *Jurnal Sistem Informasi Triguna Dharma (JURSI TGD)*, 1(4), 275. <https://doi.org/10.53513/jursi.v1i4.5331>
- Aijaz, N., Lan, H., Raza, T., Yaqub, M., Iqbal, R., & Salman, M. (2025). Artificial intelligence in agriculture : Advancing crop productivity and sustainability. *Journal of Agriculture and Food Research journal*, 20. <https://doi.org/https://doi.org/10.1016/j.jafr.2025.101762>
- Alitawan, A. A. I., & Sutrisna, K. (2017). Faktor-Faktor yang Mempengaruhi Pendapatan Petani Jeruk pada Desa Gunung Bau Kecamatan Kintamani Kabupaten Bangli. *E-Jurnal Ekonomi Pembangunan Universitas Udayana*, 6(5), 796–826.
- Azhar, P. A., Pratama, M. A., & Fitriani, R. (2024). Prediksi Harga Mobil Audi Bekas Menggunakan Model Regresi Linear dengan Framework Streamlit. *Journal of Technology and Informatics (JoTI)*, 6(1), 22–28. <https://doi.org/10.37802/joti.v6i1.763>
- FAO. (2019). *The State of Food and Agriculture: Moving Forward On Food Loss and Waste Reduction 2019*. Rome.
- Raehan, M. F. I., Kusdinar, A. B., & Indrayana, D. (2024). Penerapan Regresi Linier Berganda Untuk Memprediksi Hasil Panen Kacang Kedelai (Studi Kasus: Kecamatan Surade). *Jurnal Mahasiswa Teknik Informatika (JATI)*, 8(5), 10572–10579.
- Rohmawati, F., Rohman, G., & Mujilawati, S. (2017). Sistem Prediksi Jumlah Pengunjung Wisata Wego Kec.Sugio Kab.Lamongan Menggunakan Metode Fuzzy Time Series. *JOUTICLA*, 3(2). <https://doi.org/10.30736/jti.v2i2.66>
- S, N. A., Az-zahra, F. F., & Prajoko. (2024). Prediksi Hasil Panen Wortel Menggunakan Algoritma Regresi Linear Berganda. *Jurnal Mahasiswa Teknik Informatika (JATI)*, 8(5), 10255–10262.
- Sarno, R., Sabilla, S. I., Malikkah, Purbawa, D. S., & Ardani, M. S. hanif. (2023). *Machine*

- Learning dan Deep Learning - Konsep dan Pemrograman Python (I)*. Penerbit Andi. Diambil dari <https://play.google.com/books/reader?id=byWFEAAAQBAJ&pg=GBS.PA145&hl=en>
- Sumari, A. D. W., Musthafa, M. B., Ngatmari, & Putra, disman R. H. (2020). Perbandingan Kinerja Metode-Metode Prediksi pada Transaksi Dompot Digital di Masa Pandemi. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 4(4), 642–647. <https://doi.org/10.29207/resti.v4i4.2024>
- Susanto, A. (2017). *Sistem Informasi Akuntansi* (1 ed.). Bandung: Lingga Jaya.
- Susilo, A. A. T., Kurniawan, R., & Julyansa, G. P. (2024). Penerapan Model Regresi Linier Berganda Pada Sistem Prediksi Produksi Hasil Pertanian(Padi) Di Kota Lubuklinggau Dengan Bahasa Pemrograman Phyton. *Jurnal Informatika dan Komputer (JIK)*, 15(1), 24–30.
- www.detik.com. (2023). Membangun Masa Depan Pangan dari Perkotaan. Diambil 28 Februari 2025, dari detikJabar website: <https://www.detik.com/jabar/berita/d-7021032/membangun-masa-depan-pangan-dari-perkotaan>
- Zaka, A. R., & Sutopo. (2017). Analisis Faktor - Faktor Yang Mempengaruhi Kepuasan Pelanggan Pada LBB Antologi Semarang. *Diponegoro Journal Of Manajement*, 6(1), 1–13.