

# ***Prediksi Produktivitas Padi Berdasarkan Luas Panen dan Produksi padi Menggunakan Metode Regresi Linear Berganda***

Alya Yayan Apriyandi, Baiq Fiky Renita, Phyta Rahima

Universitas Bumigora, Mataram, Indonesia

Correspondence : e-mail: [alyayanapryndi@gmail.com](mailto:alyayanapryndi@gmail.com)

## ***Abstrak***

*Fluktuasi produktivitas padi di Nusa Tenggara Barat selama tahun 2019 hingga 2023 terjadi meskipun luas panen relatif konsisten, yang menimbulkan kesulitan dalam perencanaan produksi dan ketahanan pangan di wilayah tersebut. Untuk menghadapi keadaan ini, sebuah model prediktif dikembangkan yang bertujuan untuk mengidentifikasi pengaruh simultan dari luas panen dan volume produksi terhadap produktivitas padi. Proses pemodelan dilakukan dengan menggunakan pendekatan regresi linear berganda berdasarkan data sekunder tahunan yang mencakup variabel luas panen, produksi, dan produktivitas. Hasil dari analisis menunjukkan bahwa kedua variabel independen, yaitu luas panen dan volume produksi, memiliki pengaruh yang signifikan terhadap produktivitas padi. Model yang dirancang ini menunjukkan tingkat akurasi yang baik, yang terlihat dari nilai koefisien determinasi yang tinggi. Temuan ini menunjukkan bahwa regresi linear berganda dapat berfungsi sebagai alat analitis awal dalam merumuskan kebijakan pertanian yang berbasis pada data. Peluang untuk pengembangan lebih lanjut sangat terbuka dengan penambahan variabel lain untuk meningkatkan ketepatan prediksi di masa yang akan datang.*

***Kata kunci:*** produktivitas padi, regresi linear berganda, luas panen, produksi padi, prediksi pertanian

## ***Abstract***

*The Fluctuations in rice productivity in West Nusa Tenggara from 2019 to 2023 occurred even though the harvested area was relatively consistent, which caused difficulties in production planning and food security in the region. To deal with this situation, a predictive model was developed that aims to identify the simultaneous influence of harvested area and production volume on rice productivity. The modeling process was carried out using a multiple linear regression approach based on annual secondary data that includes the variables of harvested area, production, and productivity. The results of the analysis show that both independent variables, namely harvested area and production volume, have a significant influence on rice productivity. The designed model shows a good level of accuracy, as seen from the high coefficient of determination value. These findings indicate that multiple linear regression can function as an initial analysis tool in formulating data-based agricultural policies. Opportunities for further development are wide open with the addition of other variables to increase the precision of predictions in the future.*

***Keywords:*** rice productivity, multiple linear regression, harvested area, rice production, agricultural prediction

## **1. Pendahuluan**

Sektor pertanian memegang peranan penting dalam struktur sosial dan ekonomi di Provinsi Nusa Tenggara Barat (NTB), dengan sebagian besar penduduk menggantungkan pendapatan dari aktivitas budidaya tanaman pangan, terutama padi[1]. Padi tidak hanya menjadi komoditas utama konsumsi masyarakat, tetapi juga berkontribusi terhadap stabilitas ekonomi dan ketahanan pangan wilayah[2]. Sayangnya, data produktivitas padi di NTB menunjukkan pola yang tidak stabil antar tahun, yang dapat mengganggu keberlangsungan sistem produksi dan distribusi pangan[3]. Berdasarkan data dari Badan Pusat Statistik, periode 2019 hingga 2023 menunjukkan fluktuasi yang cukup mencolok dalam produksi padi, walaupun luas panen relatif tidak banyak berubah. Misalnya, pada tahun 2019 produksi tercatat sebesar

1.192.036 ton, namun mengalami penurunan menjadi 1.140.446 ton pada 2020. Ketidaksesuaian antara luas panen dan hasil produksi ini mengindikasikan bahwa variabel lain turut memengaruhi hasil panen, seperti pergeseran musim tanam, variabilitas iklim, dan efisiensi penggunaan input pertanian[4].

Fenomena ini memperlihatkan pentingnya pengembangan model prediksi berbasis data yang dapat menjelaskan keterkaitan antara berbagai faktor produksi terhadap tingkat produktivitas padi secara kuantitatif[5]. Salah satu pendekatan yang relevan adalah metode regresi linear berganda, yang mampu menganalisis pengaruh lebih dari satu variabel independen terhadap satu variabel dependen[6]. Dalam konteks ini, luas panen dan produksi padi dijadikan variabel bebas yang digunakan untuk memprediksi tingkat produktivitas sebagai variabel terikat[7]. Studi-studi terdahulu telah banyak menggunakan regresi linear berganda untuk menganalisis data pertanian, termasuk dalam memprediksi hasil padi[8]. Namun, sebagian besar penelitian masih terbatas pada wilayah tertentu atau belum mengintegrasikan faktor-faktor lokal yang relevan[9]. Oleh karena itu, penerapan regresi linear berganda pada konteks wilayah NTB menjadi upaya yang diperlukan untuk menghasilkan model yang lebih akurat dan aplikatif dalam mendukung pengambilan keputusan berbasis bukti[10].

Ketidakpastian dalam produktivitas padi yang tidak tertangani dengan baik berpotensi menimbulkan ketidakseimbangan antara kebutuhan konsumsi dan ketersediaan hasil pertanian[11]. Kondisi ini dapat berdampak pada ketahanan pangan dan menurunkan kesejahteraan petani lokal. Maka dari itu, dibutuhkan pendekatan prediktif yang mampu menyederhanakan kompleksitas sistem pertanian menjadi model analitis yang dapat digunakan untuk perencanaan yang lebih efektif dan terukur[5]. Penggunaan regresi linear berganda dipandang tepat karena mampu menjelaskan kontribusi masing-masing variabel independen secara simultan terhadap variabel dependen[12]. Selain menghasilkan nilai prediktif yang terukur, metode ini juga memungkinkan pengujian signifikansi hubungan antar variabel menggunakan uji statistik. Dengan bantuan perangkat lunak seperti SPSS, regresi linear berganda dapat diterapkan secara efisien untuk menghasilkan model yang valid dan dapat dipertanggungjawabkan secara ilmiah[13].

Penelitian ini bertujuan untuk membangun model prediksi produktivitas padi di Provinsi NTB menggunakan metode regresi linear berganda, dengan memanfaatkan data historis luas panen dan produksi selama periode 2019–2023. Hasil dari model ini diharapkan dapat menjadi dasar dalam penyusunan strategi pertanian yang lebih adaptif, serta mendukung kebijakan pengelolaan pangan yang berkelanjutan dan berbasis data.

## 2. Metode Penelitian

Pendekatan kuantitatif diterapkan dengan menggunakan metode regresi linear berganda untuk memprediksi produktivitas padi di Provinsi Nusa Tenggara Barat (NTB) berdasarkan data historis selama periode 2019 hingga 2023. Metode ini dipilih karena mampu mengkaji pengaruh simultan dua variabel bebas, yaitu luas panen dan produksi padi, terhadap satu variabel terikat, yakni produktivitas (ton per hektar). Regresi linear berganda memungkinkan analisis lebih dari satu variabel independen secara bersamaan terhadap variabel dependen[14].

Persamaan matematis yang digunakan mengacu pada bentuk dasar regresi linear berganda dan dirumuskan sebagai berikut:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon [15]$$

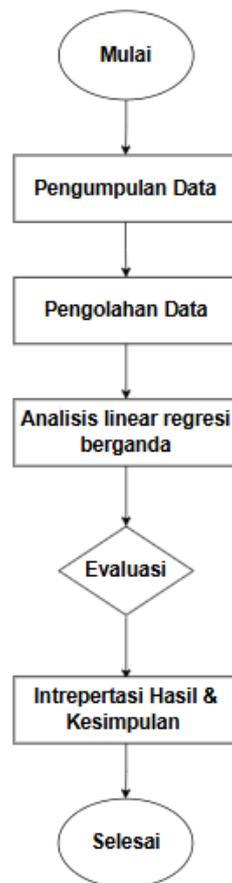
- $Y$  = **Produktivitas padi** (ton/ha)
- $X_1$  = **Luas panen** (hektar)
- $X_2$  = **Produksi padi** (ton)
- $\beta_0$  = **Intercept** atau konstanta (nilai  $Y$  saat  $X_1=0$  dan  $X_2=0$ )
- $\beta_1$  = **Koefisien regresi** untuk variabel  $X_1$  (luas panen)
- $\beta_2$  = **Koefisien regresi** untuk variabel  $X_2$  (produksi padi)
- $\varepsilon$  = **Error term** atau galat, yaitu komponen residual dari model (mewakili faktor lain yang tidak dijelaskan oleh  $X_1$  dan  $X_2$ )

Data yang digunakan bersumber dari publikasi Badan Pusat Statistik (BPS) Provinsi NTB dan terdiri dari 50 observasi. Variabel yang dianalisis meliputi tahun, luas panen, dan produksi padi. Produktivitas dihitung dari rasio antara produksi terhadap luas panen. Proses analisis dilakukan menggunakan perangkat lunak statistik seperti SPSS atau Microsoft Excel.

Sebelum analisis regresi dilakukan, data terlebih dahulu dianalisis secara deskriptif dan diuji melalui serangkaian pengujian asumsi klasik. Uji tersebut meliputi normalitas residual, multikolinearitas antar

variabel independen, serta heteroskedastisitas. Tahapan ini bertujuan untuk memastikan bahwa model regresi linear berganda yang dibangun memenuhi syarat validitas dan dapat diinterpretasikan secara ilmiah.

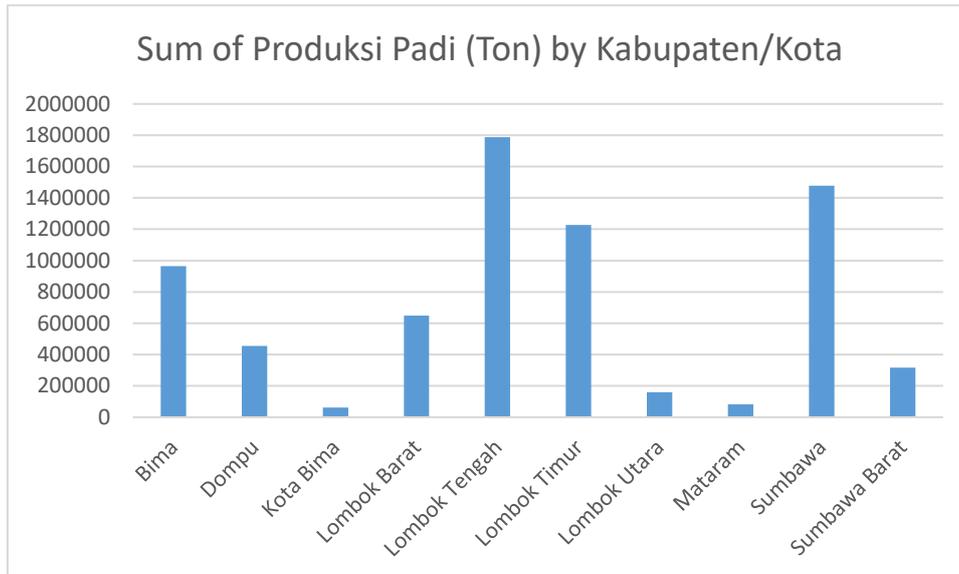
Hasil yang diperoleh diharapkan mampu memberikan pemahaman yang lebih baik tentang kontribusi variabel luas panen dan produksi terhadap tingkat produktivitas padi, serta menyajikan model prediksi yang relevan bagi pengambilan keputusan di bidang pertanian pada tingkat regional[15]. Kendati regresi linear memiliki keterbatasan dalam menangkap dinamika non-linier dan pengaruh variabel eksternal, pendekatan ini tetap dianggap representatif dalam menyusun analisis awal terhadap fluktuasi produktivitas padi di NTB. Alur analisis data dalam studi ini ditunjukkan pada gambar berikut:



Gambar 1. *Flowchart* Proses Analisis Linear Berganda

### 3. Hasil dan Pembahasan

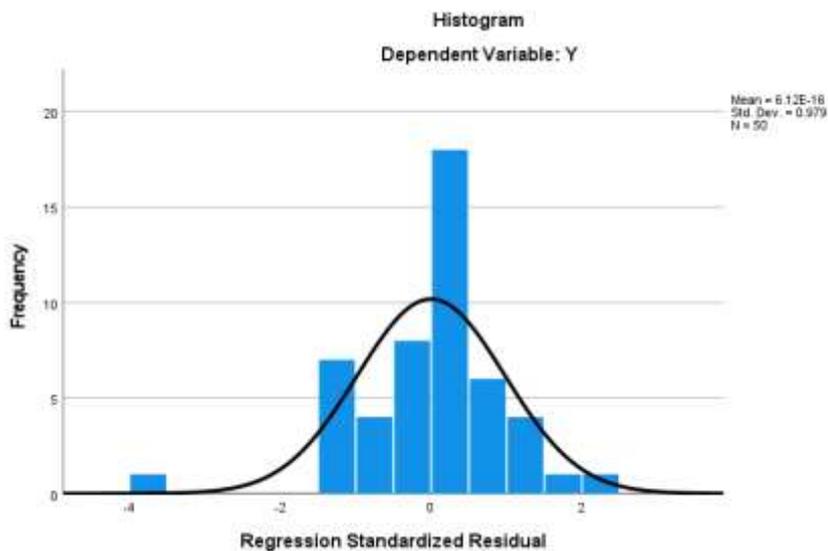
Penelitian ini mengandalkan data sekunder dengan total 50 observasi yang mencakup variabel luas panen ( $X_1$ ), produksi padi ( $X_2$ ), dan produktivitas padi ( $Y$ ) di Provinsi Nusa Tenggara Barat (NTB) selama periode 2019 hingga 2023. Data diperoleh dari sumber resmi yang sudah melalui proses validasi. Secara keseluruhan, nilai produktivitas padi berfluktuasi antara 43,96 kuintal/ha sampai 67,74 kuintal/ha, dengan rata-ratanya adalah 52,91 kuintal/ha. Luas panen bervariasi cukup signifikan, dari 2.316 ha hingga 72.414 ha, sedangkan jumlah produksi padi berkisar antara 10.903 ton sampai 380.812 ton.



Gambar 2. Jumlah Total Produksi Padi

Proses analisis regresi linear berganda diawali dengan verifikasi terhadap asumsi-asumsi mendasar guna memastikan bahwa model yang dibangun memiliki validitas yang baik. Salah satu asumsi penting yang diuji adalah normalitas residual. Pemeriksaan awal dilakukan melalui tampilan visual berupa histogram residual dan normal P-P plot. Hasil histogram menunjukkan bahwa penyebaran residual yang belum distandarasi membentuk pola menyerupai distribusi normal, ditandai dengan kemiringan yang relatif simetris terhadap titik nol. Arah horizontal menggambarkan nilai residual, sedangkan sumbu vertikal mencerminkan jumlah kemunculan. Adanya kurva normal yang melapisi histogram menunjukkan bahwa pola distribusi residual berada dalam kisaran yang wajar, meskipun terdapat sedikit penyimpangan pada bagian ujung distribusi.

Pada grafik normal P-P plot, titik-titik residual terlihat sejajar mengikuti garis diagonal dari arah kiri bawah menuju kanan atas. Kesesuaian ini menunjukkan bahwa residual menyebar dengan pola yang mendekati distribusi normal. Secara keseluruhan, interpretasi visual ini memberikan dukungan terhadap terpenuhinya asumsi normalitas, meskipun diperlukan pengujian statistik tambahan untuk memperkuat kesimpulan secara kuantitatif.



Gambar 3. Histogram Residual Regresi Linear Berganda

Pengujian normalitas residual dilakukan melalui pendekatan visual menggunakan histogram dan dilengkapi dengan uji statistik Kolmogorov-Smirnov serta Shapiro-Wilk untuk menilai apakah residual mengikuti distribusi normal. Berdasarkan hasil pengujian yang tercantum pada Tabel X, diperoleh nilai signifikansi sebesar 0,031 untuk Kolmogorov-Smirnov dan 0,002 untuk Shapiro-Wilk. Kedua nilai tersebut berada di bawah tingkat signifikansi 0,05, yang menunjukkan bahwa residual dalam model tidak terdistribusi normal secara statistic. Pelanggaran terhadap asumsi normalitas ini tidak mengurangi kelayakan model, mengingat jumlah sampel yang digunakan dalam penelitian mencukupi ( $n = 50$ ) dan asumsi klasik lainnya seperti multikolinearitas dan homoskedastisitas telah terpenuhi. Model regresi linear berganda menunjukkan kinerja yang sangat baik dengan nilai koefisien determinasi ( $R^2$ ) sebesar 0,992 serta signifikansi tinggi pada uji F dan uji t. Hasil ini mengindikasikan bahwa model mampu menjelaskan hampir seluruh variasi pada variabel dependen secara signifikan.

Tabel 1. Tabel Test Normality

<b>Tests of Normality</b>								
	Kolmogorov-Smirnov <sup>a</sup>				Shapiro-Wilk			
	tistic	Sta	df	g.	tistic	Sta	df	g.
Unstandardized Residual	.13		50	.03	.91		50	.00
	1			1	7		2	

Hasil pengujian terhadap statistik residual yang disajikan dalam tabel menunjukkan bahwa nilai residual yang telah distandarisasi berada dalam rentang  $-3,871$  hingga  $2,112$ , dengan nilai rata-rata mendekati nol. Rentang ini mengindikasikan tidak adanya pencilan (outlier) yang ekstrem pada data. Temuan ini juga mendukung terpenuhinya asumsi homoskedastisitas, yang ditandai dengan penyebaran residual yang relatif konstan di seluruh rentang prediksi. Berdasarkan indikator tersebut, dapat disimpulkan bahwa distribusi residual dalam model memenuhi kriteria normalitas dan kestabilan varians, sehingga model regresi layak digunakan untuk analisis lebih lanjut.

Tabel 2. Tabel Residual Statistik  
**Residuals Statistics<sup>a</sup>**

	Minimu m	Maxim um	Mean	Std. Deviation	N
Predicted Value	1542.88 57	37400 0.0000	14260 8.2600	117791. 59711	50
Residual	- 42987.21094	23459. 60742	.00000	10876.9 2221	50
Std. Predicted Value	-1.198	1.964	.000	1.000	50
Std. Residual	-3.871	2.112	.000	.979	50

Hasil analisis regresi linear berganda menunjukkan bahwa model memiliki nilai koefisien determinasi (R Square) sebesar 0,992, yang berarti 99,2% variasi pada produktivitas padi dapat dijelaskan oleh variabel luas panen dan produksi padi. Nilai Adjusted R Square sebesar 0,991 menunjukkan bahwa model tetap stabil meskipun menggunakan dua variabel bebas. Hal ini menunjukkan bahwa model memiliki tingkat akurasi yang sangat tinggi dan sangat baik digunakan untuk prediksi.

Hasil uji F yang tercantum dalam tabel ANOVA menghasilkan nilai F hitung sebesar 2756,033 dengan signifikansi  $< 0,001$ , sehingga dapat disimpulkan bahwa model regresi secara simultan signifikan. Artinya, variabel luas panen dan produksi padi secara bersama-sama berpengaruh signifikan terhadap produktivitas padi.

Tabel 3. Tabel Anova  
**ANOVA<sup>a</sup>**

Model	Sum of Squares	df	Mean Square	F	Sig
Regression	6798681	2	3399340	275	.00
Total	57139.972		78569.986	6.033	0 <sup>b</sup>

Residual	5797064	47	1233417		
al	405.648		95.865		
Total	6856652	49			
	21545.620				

Uji t dilakukan untuk mengetahui pengaruh masing-masing variabel bebas secara parsial. Hasilnya menunjukkan bahwa luas panen (X1) memiliki nilai t sebesar 4,703 ( $p < 0,001$ ), dan produksi padi (X2) memiliki nilai t sebesar 70,992 ( $p < 0,001$ ), yang berarti keduanya berpengaruh signifikan terhadap produktivitas padi. Nilai koefisien regresi tak tersatukan (B) menunjukkan bahwa setiap peningkatan satu satuan pada luas panen akan meningkatkan produktivitas sebesar 1.465,126 satuan, dan setiap peningkatan satu satuan pada produksi akan meningkatkan produktivitas sebesar 5,194 satuan, dengan asumsi variabel lain tetap.

Tabel 4. Tabel *coefficients*

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error			
(Constant)	-79200.331	17375.403		-4.558	.000
X1	1465.126	311.550	.067	4.703	.000
X2	5.194	.073	1.018	70.992	.000

a. Dependent Variable: Y

Hasil analisis menunjukkan bahwa di antara kedua variabel independen, produksi padi (X2) memberikan kontribusi paling dominan terhadap produktivitas padi. Hal ini tercermin dari nilai koefisien beta sebesar 1,018 serta nilai t-hitung yang secara signifikan lebih tinggi dibandingkan dengan variabel luas panen (X1). Temuan ini mengindikasikan bahwa peningkatan produktivitas padi di Provinsi Nusa Tenggara Barat (NTB) lebih efektif dicapai melalui optimalisasi hasil produksi. Strategi yang dapat diimplementasikan meliputi penerapan inovasi teknologi pertanian, peningkatan efektivitas penggunaan pupuk, serta penguatan sistem irigasi yang berkelanjutan.

Secara keseluruhan, model regresi linear berganda yang diterapkan dalam penelitian ini menunjukkan tingkat signifikansi dan validitas yang tinggi, serta memiliki kemampuan yang sangat baik dalam menjelaskan variasi produktivitas padi berdasarkan data yang dianalisis. Dengan demikian, model ini dapat dijadikan sebagai dasar pertimbangan dalam perumusan kebijakan pembangunan pertanian berbasis data, khususnya dalam menyusun strategi peningkatan produksi dan produktivitas secara terarah di wilayah Provinsi Nusa Tenggara Barat.

#### 4. Kesimpulan

Hasil penelitian mengonfirmasi bahwa model regresi linear berganda merupakan pendekatan yang efektif dalam mengkaji serta memproyeksikan produktivitas padi di Provinsi Nusa Tenggara Barat dengan menggunakan variabel luas panen dan volume produksi sebagai prediktor utama. Analisis statistik menunjukkan bahwa kedua variabel independen tersebut memiliki pengaruh yang signifikan terhadap produktivitas, baik secara simultan maupun individual, dengan dukungan nilai koefisien determinasi ( $R^2$ ) sebesar 0,992. Temuan juga mengindikasikan bahwa produksi padi memiliki kontribusi paling dominan dalam menjelaskan variabilitas produktivitas. Model yang dihasilkan dinyatakan valid dan signifikan secara statistik, serta dapat dimanfaatkan sebagai acuan dalam penyusunan kebijakan pertanian berbasis data. Untuk meningkatkan akurasi prediksi di masa mendatang, direkomendasikan agar model dikembangkan lebih lanjut dengan memasukkan variabel tambahan seperti intensitas curah hujan, penggunaan pupuk, tingkat adopsi teknologi, dan efisiensi sistem irigasi.

## Daftar Pustaka

- [1] D. Setiadi, Sasmita, and M. Yolanda, "Penerapan Algoritma Regresi Linier Berganda Untuk Memprediksi Hasil panen Padi Di Kota Pagar Alam," *Pagar Alam*, Apr. 2024. doi: pp.438-447.
- [2] E. A. Jiya, U. Illiyasu, and M. Akinyemi, "Rice Yield Forecasting: A Comparative Analysis of Multiple Machine Learning Algorithms," *Journal of Information Systems and Informatics*, vol. 5, no. 2, pp. 785–799, Jun. 2023, doi: 10.51519/journalisi.v5i2.506.
- [3] D. Nuraini *et al.*, "Prediksi Hasil Panen Padi dengan Metode Multiple Linear Regression dan Particle Swarm Optimization untuk Meningkatkan Produksi Padi di Madura," vol. 4, no. 1, pp. 1–8, 2025, doi: 10.70609/jusifor.v4i1.5857.
- [4] S. Aprizkiyandari and T. Palupi, "Rice production forecasting model in West Kalimantan with factors are rainfall and harvest area," in *IOP Conference Series: Earth and Environmental Science*, Institute of Physics, 2023. doi: 10.1088/1755-1315/1177/1/012015.
- [5] J. A. Ingio, A. S. Nsang, and A. Iorliam, "Optimizing Rice Production Forecasting Through Integrating Multiple Linear Regression with Recursive Feature Elimination," *Journal of Future Artificial Intelligence and Technologies*, vol. 1, no. 2, pp. 96–108, Aug. 2024, doi: 10.62411/faith.2024-17.
- [6] A. Prasetyo and J. Teknologi Informasi dan Komputer Politeknik Negeri Lhokseumawe, "Prediksi Produksi Kelapa Sawit Menggunakan Metode Regresi Linier Berganda," *Multimedia & Jaringan*, vol. 6, no. 2, 2021.
- [7] T. N. Padilah and R. I. Adam, "Analisis Regresi Linier Berganda Dalam Estimasi Produktivitas Tanaman Padi Di Kabupaten Karawang," *FIBONACCI: Jurnal Pendidikan Matematika dan Matematika*, vol. 5, no. 2, p. 117, Dec. 2019, doi: 10.24853/fbc.5.2.117-128.
- [8] C. P. Dhakal, "Multiple Regression Model Fitted for Rice Production Forecasting in Nepal: A Case of Time Series Data," *Nepalese Journal of Statistics*, vol. 2, pp. 89–98, Sep. 2018, doi: 10.3126/njs.v2i0.21157.
- [9] O. Simon Hardianto Raharusun *et al.*, "Sistem Prediksi Produksi Beras Menggunakan Multiple Linear Regression untuk Optimalisasi Ketahanan Pangan di Kabupaten Minahasa," *JOINTER-JOURNAL OF INFORMATICS ENGINEERING*, vol. 05, no. 02.
- [10] M. F. Badaruzaman and O. Chee Tiong, "Crop Yield Prediction with Multiple Linear Regression in Malaysia," 2022. doi: Proc.Sci.Math.9:211-221.
- [11] E. Triyanto, H. Sismoro, and A. D. Laksito, "IMPLEMENTASI ALGORITMA REGRESI LINEAR BERGANDA UNTUK MEMPREDIKSI PRODUKSI PADI DI KABUPATEN BANTUL," *Rabit : Jurnal Teknologi dan Sistem Informasi Univrab*, vol. 4, no. 2, pp. 66–75, Jul. 2019, doi: 10.36341/rabit.v4i2.666.
- [12] A. Pasaribu and M. R. Syahputra, "Comparison of Fuzzy Logic and Multiple Linear Regression in Forecasting Rice Production in Toba District," *Jurnal Matematika dan Pendidikan Matematika*, vol. 5, pp. 121–127, Dec. 2022.
- [13] G. Erman Agusta, A. Dewanto, and N. Astriawati, "International Journal of Technology and Education Research Multiple Linear Regression Model for Analyzing the Determinants of Rice Production in Sumatra," vol. 03, no. 55284, 2025, [Online]. Available: <https://e-journal.citakonsultindo.or.id/index.php/IJETER>
- [14] R. Prasad, P. Krishnamurthy, R. M. Kumar, and B. C. Viraktamath, "Multiple linear regression models for prediction of rice production and productivity based on rainfall." [Online]. Available: <http://www.indiastat.com>
- [15] S. Pazhanivelan *et al.*, "Spatial Rice Yield Estimation Using Multiple Linear Regression Analysis, Semi-Physical Approach and Assimilating SAR Satellite Derived Products with DSSAT Crop Simulation Model," *Agronomy*, vol. 12, no. 9, Sep. 2022, doi: 10.3390/agronomy12092008.