

Analisis Data Pertanian Tanaman Pangan untuk Memprediksi Hasil Panen di Kabupaten Malaka Menggunakan Metode *Multiple Linear Regression*

M K B Seran¹, F Tedy², I P A N Samane³, P Batarius⁴, P A Nani⁵, A A J Sinlae⁶

¹⁻⁶Program Studi Ilmu Komputer, Universitas Katolik Widya Mandira

E-mail: kristinebriaseran@gmail.com¹, frengkytedy@unwira.ac.id²,
pricher_samane@unwira.ac.id³, patrisbatarius@unwira.ac.id⁴,
paskalisnani@unwira.ac.id⁵, alfry.aj@unwira.ac.id⁶

Abstrak. Kabupaten Malaka merupakan daerah otonom baru dengan pertanian sebagai salah satu sektor basisnya. Dari tahun 2012 sampai dengan 2021, hasil produksi pertanian berupa tanaman pangan padi, jagung dan kacang hijau selalu mengalami kenaikan dan penurunan. Beberapa faktor yang menyebabkan terjadinya kenaikan dan penurunan hasil produksi tanaman pangan tersebut adalah adanya alih fungsi lahan pertanian menjadi lahan bangunan, masih tingginya harga bahan pangan yang ada di pasaran dan kebutuhan yang semakin meningkat, serta jumlah populasi penduduk yang terus bertambah. Untuk itu diperlukan sebuah metode yang dapat memprediksi hasil panen tanaman pangan di Kabupaten Malaka untuk 5 tahun yang akan datang. Hal ini dapat membantu Dinas Pertanian melakukan pengambilan keputusan terkait, guna menunjang Pemerintah Daerah dalam membantu menyediakan bahan pangan yang dibutuhkan bagi masyarakat setempat. Penelitian ini menggunakan metode *multiple linear regression* untuk melakukan prediksi. Hasil uji koefisien determinasi sebesar 0.918, menunjukkan bahwa 91.8% variasi variabel dependen dapat dijelaskan oleh variabel independen dalam model. Selain itu, terdapat tiga analisis statistik yang digunakan untuk melihat performansi dari metode *multiple linear regression*, yaitu nilai MSE sebesar 69611782.304, RMSE sebesar 8343.368 dan MAE sebesar 7327.695.

Kata kunci: pertanian; tanaman pangan; prediksi; *multiple linear regression*; Kabupaten Malaka

Abstract. Malaka Regency is a new autonomous region with one of its basic sectors being agriculture. From 2012 to 2021, agricultural production in the form of rice, corn, and green beans always experienced increases and decreases. Several factors cause increases and decreases in food crop production, namely the conversion of land from agriculture to building land, the still high price of food on the market and increasing demand, as well as the population continuing to increase. For these reasons, a method is needed to predict the yield of food crops in Malaka Regency for the next 5 years. It can help the agricultural department make related decisions, to support the Regional Government in helping to provide the food needed for the local community. In this study, the multiple linear regression method was used to make predictions. The results with a coefficient of determination test of 0.918 showed that 91.8% of the variation in the dependent variable can be explained by the independent variables in the model. Apart from that, there were three statistical analyses used to see the performance of the multiple linear regression method, namely the MSE value of 69611782.304, RMSE of 8343.368, and MAE of 7327.695.

Keywords: agriculture; crops; prediction; multiple linear regression; Malaka Regency

1. Pendahuluan

Indonesia adalah salah satu negara dengan kekayaan sumber daya alam yang melimpah serta memiliki kondisi tanah dan musim yang sangat cocok bagi sektor pertanian [1]. Pada umumnya, para petani akan melakukan proses penanaman terhadap beberapa jenis tanaman yang menjadi sumber kebutuhan pokok untuk di konsumsi oleh masyarakat dengan tujuan memenuhi ketersediaan pangan yang bermutu dan bergizi, dengan harga yang wajar dan terjangkau sesuai dengan kebutuhan masyarakat [2].

Di saat ini potensi sektor pertanian memiliki peran yang signifikan karena dapat menjadi sumber pendapatan dan juga untuk memenuhi ketersediaan pangan bagi masyarakat di suatu daerah, salah satunya adalah Kabupaten Malaka [3]. Untuk menjalankan program percepatan dalam bidang pertanian, Kabupaten Malaka tentunya memiliki sebuah tantangan tersendiri, di mana letak geografis kabupaten ini berada pada kawasan yang memiliki ragam karakteristik potensi iklim dan juga kondisi lahannya. Oleh karena itu, dibutuhkan suatu gagasan yang dapat meningkatkan hasil pertanian yang ada melalui ketersediaan lahan dan penjadwalan pembangunan yang merata dengan melihat potensi yang dimiliki oleh setiap daerah [4].

Hasil produksi untuk beberapa komoditas lahan pertanian di kabupaten malaka pada tahun 2021 terdiri dari tanaman padi sebanyak 18.710 ton, Jagung 74.800 ton, Kacang Hijau 2.819 ton, Ubi Kayu 19.273 ton, Ubi Jalar 1.028 ton dan Kacang Tanah 1.776 ton. Selama kurun waktu 10 tahun terakhir, yaitu dari tahun 2012 sampai dengan 2021, komoditas tanaman pangan padi, jagung dan kacang hijau merupakan jenis komoditas pangan yang paling banyak dibutuhkan oleh rumah tangga ataupun masyarakat dan selalu mengalami kenaikan dan penurunan dalam hal produksi. Adapun data yang diperoleh dari internal Dinas Pertanian Kabupaten Malaka untuk hasil produksi pertanian dari ketiga jenis tanaman pangan tersebut dapat dilihat pada Tabel 1 berikut ini.

Tabel 1. Data Hasil Produksi Pertanian 10 Tahun Terakhir

No	Tahun	Padi (ton)	Jagung (ton)	Kacang Hijau (ton)
1.	2012	30.013	70.619	2.349
2.	2013	16.666	40.591	1.942
3.	2014	16.974	52.414	1.533
4.	2015	27.910	51.370	2.295
5.	2016	27.846	50.359	1.893
6.	2017	35.245	72.326	1.904
7.	2018	36.766	84.183	2.851
8.	2019	39.390	94.573	3.600
9.	2020	31.206	87.551	3.760
10.	2021	18.710	74.800	2.819

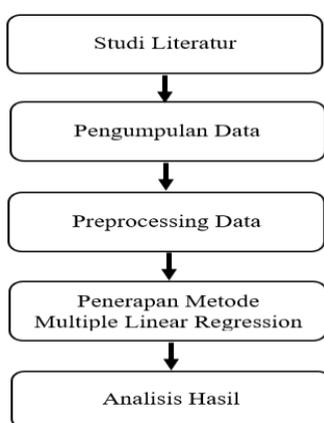
Penyebab terjadinya kenaikan dan penurunan hasil produksi pertanian untuk tanaman pangan padi, jagung dan kacang hijau di Kabupaten Malaka, dikarenakan adanya perluasan lahan bangunan di area lahan pertanian dalam beberapa tahun terakhir, yang berdampak pada kurangnya ketersediaan bahan pangan tersebut bagi kebutuhan masyarakat. Untuk itu, ketersediaan lahan sangat penting dan berpengaruh terhadap luas tanam dan luas panen yang dihasilkan dari rata-rata produksi pada lahan pertanian, selain faktor curah hujan sebagai salah satu unsur yang dapat menentukan dalam hasil akhir produksi tanaman pangan. Dengan masih tingginya harga bahan pangan yang ada di pasaran dan kebutuhan yang semakin meningkat, serta jumlah populasi penduduk yang terus bertambah, menjadikan sebagian besar masyarakat tidak mendapatkan bahan pangan yang cukup untuk kebutuhan rumah tangga di setiap tahunnya. Demi mengatasi keadaan-keadaan tersebut, perlu dilakukan dengan menggunakan sebuah metode *multiple linear regression* yang dapat memprediksi tentang hasil panen tanaman pangan di Kabupaten Malaka.

Prediksi dapat diartikan sebagai sebuah proses untuk meramalkan suatu keadaan yang paling mungkin terjadi di masa mendatang, berdasarkan pengamatan data pada masa lampau atau informasi masa lalu dan sekarang yang dimiliki [5], [6]. Untuk itu, proses prediksi dapat dilakukan dengan

menggunakan metode *multiple linear regression*. Penerapan metode *multiple linear regression* atau regresi linier berganda untuk melakukan hasil prediksi dalam bidang pertanian telah dilakukan dalam beberapa penelitian, di antaranya pada hasil panen buah pinang (*Areca Catechu*) [7], produksi jagung pada Dinas Pertanian Kabupaten Dompu [8], hasil pertanian durian [9], produksi dan hasil panen padi di Kabupaten Cirebon [10], [11] dan produksi bawang merah di Kabupaten Brebes [12]. Berdasarkan uraian yang ada, maka terbuka peluang untuk melakukan penelitian bagaimana menganalisis data pertanian tanaman pangan dengan menggunakan metode *multiple linear regression* dalam memprediksi hasil panen di Kabupaten Malaka untuk 5 tahun yang akan datang, sehingga dapat membantu dinas pertanian melakukan pengambilan keputusan terkait, guna menunjang Pemerintah Daerah dalam membantu menyediakan bahan pangan yang dibutuhkan bagi masyarakat setempat.

2. Metode

Penelitian ini menggunakan metode penelitian komparatif yang membandingkan satu variabel dengan variabel lainnya untuk mendapatkan kebenaran dari objek penelitian yang sedang diteliti [13]. Untuk metode analisisnya menggunakan metode *data mining* yaitu prediksi dengan menggunakan *tools Orange*. Adapun tahapan-tahapan yang akan dilakukan, dapat dilihat pada Gambar 1 berikut ini:



Gambar 1. Tahapan Penelitian

2.1. Studi Literatur

Pada tahapan ini dilakukan dengan mempelajari beberapa referensi dan sumber lainnya terkait penerapan metode *multiple linear regression* untuk melakukan hasil prediksi dalam bidang pertanian.

2.2. Pengumpulan Data

Pada tahapan ini pengumpulan data padi, data jagung dan data kacang hijau dari setiap kecamatan di Kabupaten Malaka, yang meliputi data kecamatan, luas tanam, luas panen, produksi dan produktivitas bersumber dari data yang ada di Dinas Pertanian Kabupaten Malaka.

2.3. Preprocessing Data

Preprocessing data dilakukan dengan proses *cleaning* data, yaitu dengan cara mereduksi atribut-atribut tertentu yang tidak diperlukan sehingga dapat mempermudah dalam melakukan analisis dan penerapan metode secara efektif [14]. Adapun dalam penelitian ini tahap *preprocessing* data yang dilakukan meliputi atribut luas tanam, luas panen, curah hujan, rata-rata produksi dan produksi.

2.4. Penerapan Metode Multiple Linear Regression

Penerapan metode *multiple linear regression* pada penelitian ini menggunakan lebih dari satu buah variabel independen (X) dan satu variabel dependen (Y) [15]. Adapun yang menjadi kategori atribut variabel independen (X), yaitu luas tanam, luas panen, produktivitas dan curah hujan, sedangkan untuk atribut produksi atau hasil panen merupakan variabel dependen (Y) sehingga dapat dinyatakan bahwa

penerapan metode ini memiliki relasi antara peubah respons (variabel dependen) dengan faktor-faktor yang mempengaruhi lebih dari satu prediktor (variabel independen) dan apabila semua kelas yang dihasilkan berupa numerik, maka dapat diekspresikan kelas tersebut sebagai kombinasi *linear* dari atribut dengan bobot yang telah ditentukan [16] dengan rumus sebagai berikut:

$$\gamma = a + b_1X_1 + b_2X_2 + \dots + b_nX_n \quad (1)$$

Di mana :

Y = Variabel tidak bebas (nilai yang diprediksikan)

X = Variabel bebas

a = Konstanta

b = Koefisien

Di mana nilai a, b₁, b₂ ... b_n dapat dihitung dengan metode kuadrat terkecil yaitu:

$$\begin{aligned} b_1 &= \frac{(\sum X_2^2)(\sum X_1 y) - (\sum X_2 y)(\sum X_1 X_2)}{(\sum X_1^2)(\sum X_2^2) - (\sum X_1 X_2)^2} \\ b_2 &= \frac{(\sum X_1^2)(\sum X_2 y) - (\sum X_1 y)(\sum X_1 X_2)}{(\sum X_1^2)(\sum X_2^2) - (\sum X_1 X_2)^2} \\ a &= \frac{\sum Y - (b_1 \sum X_1) - (b_2 \sum X_2)}{n} \end{aligned} \quad (2)$$

Di mana:

$$\begin{aligned} \sum x_1^2 &= \sum x_1^2 - \frac{(\sum x_1)^2}{n} \\ \sum x_2^2 &= \sum x_2^2 - \frac{(\sum x_2)^2}{n} \\ \sum y^2 &= \sum y^2 - \frac{(\sum y)^2}{n} \\ \sum x_1 y &= \sum x_1 y - \frac{\sum x_1 \sum y}{n} \\ \sum x_2 y &= \sum x_2 y - \frac{\sum x_2 \sum y}{n} \\ \sum x_1 x_2 &= \sum x_1 x_2 - \frac{\sum x_1 \sum x_2}{n} \end{aligned} \quad (3)$$

Untuk penggunaan matriks determinan dapat dilakukan dengan mencari nilai konstanta dan koefisien regresi dari setiap variabel bebas. Contohnya adalah ketika terdapat 3 persamaan dengan 3 variabel yang tidak diketahui nilainya, yaitu a, b₁, b₂ dan b₃, persamaan tersebut dapat dinyatakan dalam persamaan matriks sebagai berikut:

$$\begin{aligned} A &= \begin{bmatrix} N & \sum X_1 & \sum X_2 & \sum X_3 \\ \sum X_1 & \sum(X_1 \cdot X_1) & \sum(X_1 \cdot X_2) & \sum(X_1 \cdot X_3) \\ \sum X_2 & \sum(X_2 \cdot X_1) & \sum(X_2 \cdot X_2) & \sum(X_2 \cdot X_3) \\ \sum X_3 & \sum(X_3 \cdot X_1) & \sum(X_3 \cdot X_2) & \sum(X_3 \cdot X_3) \end{bmatrix} \\ A0 &= \begin{bmatrix} \sum(X \cdot Y) & \sum X_1 & \sum X_2 & \sum X_3 \\ \sum(X_1 \cdot Y) & \sum(X_1 \cdot X_1) & \sum(X_1 \cdot X_2) & \sum(X_1 \cdot X_3) \\ \sum(X_2 \cdot Y) & \sum(X_2 \cdot X_1) & \sum(X_2 \cdot X_2) & \sum(X_2 \cdot X_3) \\ \sum(X_3 \cdot Y) & \sum(X_3 \cdot X_1) & \sum(X_3 \cdot X_2) & \sum(X_3 \cdot X_3) \end{bmatrix} \\ A1 &= \begin{bmatrix} N & \sum(Y) & \sum X_2 & \sum X_3 \\ \sum X_1 & \sum(X_1 \cdot Y) & \sum(X_1 \cdot X_2) & \sum(X_1 \cdot X_3) \\ \sum X_2 & \sum(X_2 \cdot Y) & \sum(X_2 \cdot X_2) & \sum(X_2 \cdot X_3) \\ \sum X_3 & \sum(X_3 \cdot Y) & \sum(X_3 \cdot X_2) & \sum(X_3 \cdot X_3) \end{bmatrix} \end{aligned} \quad (4)$$

$$A2 = \begin{bmatrix} N & \sum X_1 & \sum(Y) & \sum X_2 \\ \sum X_1 & \sum(X_1 \cdot X_1) & \sum(X_1 \cdot Y) & \sum(X_1 \cdot X_3) \\ \sum X_2 & \sum(X_2 \cdot X_1) & \sum(X_2 \cdot Y) & \sum(X_2 \cdot X_3) \\ \sum X_3 & \sum(X_3 \cdot X_1) & \sum(X_3 \cdot Y) & \sum(X_3 \cdot X_3) \end{bmatrix}$$

$$A3 = \begin{bmatrix} N & \sum X_1 & \sum X_2 & \sum(Y) \\ \sum X_1 & \sum(X_1 \cdot X_1) & \sum(X_1 \cdot X_2) & \sum(X_1 \cdot Y) \\ \sum X_2 & \sum(X_2 \cdot X_1) & \sum(X_2 \cdot X_2) & \sum(X_2 \cdot Y) \\ \sum X_3 & \sum(X_3 \cdot X_1) & \sum(X_3 \cdot X_2) & \sum(X_3 \cdot Y) \end{bmatrix}$$

Sesudah dilakukan perhitungan untuk masing-masing matriks A, A1, A2 dan A3, maka tahap selanjutnya dilakukan perhitungan determinan dengan rumus sebagai berikut :

$$a = \frac{Det(A0)}{Det(A)}$$

$$b1 = \frac{Det(A1)}{Det(A)}$$

$$b2 = \frac{Det(A2)}{Det(A)}$$

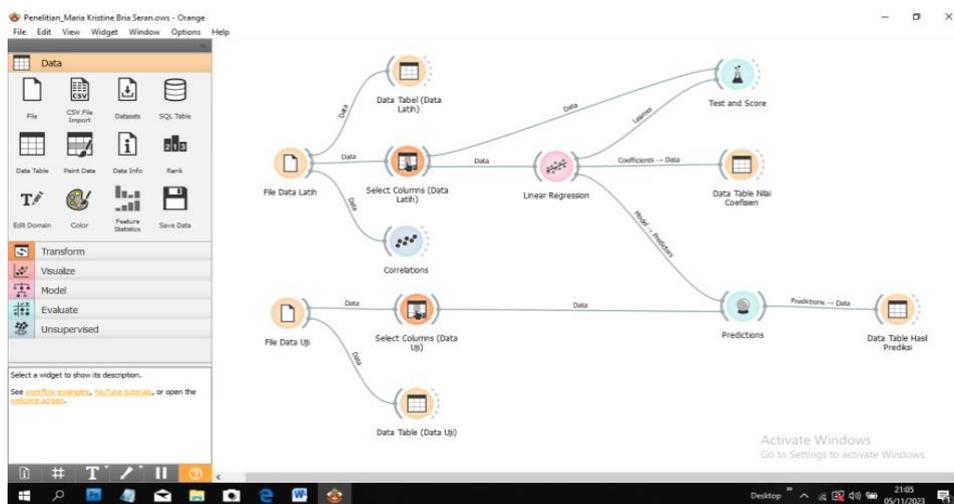
$$b3 = \frac{Det(A3)}{Det(A)}$$
(5)

2.5. Analisis Hasil

Pada tahapan analisis hasil, dilakukan prediksi terhadap hasil pertanian tanaman pangan di Kabupaten Malaka dengan menggunakan metode *multiple linear regression*, yang hasil luarannya berupa angka yang menampilkan banyaknya jumlah hasil panen tanaman pangan tersebut. Dari hasil panen yang diperoleh, hasil pertanian di masa yang akan datang dapat diprediksi untuk membantu Dinas Pertanian mengambil keputusan, guna menunjang Pemerintah Daerah dalam membantu menyediakan bahan pangan yang dibutuhkan masyarakat setempat.

3. Hasil dan Pembahasan

Dalam melakukan proses prediksi terhadap hasil panen berupa padi, jagung dan kacang hijau yang ada di Kabupaten Malaka untuk 5 tahun yang akan datang, dilakukan dengan menggunakan metode *multiple linear regression* yang terdapat pada *tools Orange*, seperti yang terlihat pada Gambar 2 berikut ini :



Gambar 2. Prediksi hasil panen menggunakan *tools Orange*

3.1. Sampel Data Latih

Sampel data latih yang digunakan dalam memprediksi hasil pertanian di Kabupaten Malaka terdiri dari beberapa atribut, yaitu tanaman pangan, tahun, luas tanam, luas panen, rata-rata produksi, curah hujan dan produksi seperti yang terlihat pada Tabel 2 berikut ini:

Tabel 2. Data Latih

Tanaman Pangan	Tahun	Luas Tanam (ha)	Luas Panen (ha)	Rata-rata Produksi (ton/ha)	Curah Hujan (mm)	Produksi (ton)
Padi	2012	8.425	8.425	35	1.355	30.013
Jagung	2012	25.944	25.944	27	1.355	70.619
Kacang Hijau	2012	1.927	1.927	12	1.355	2.349
Padi	2013	4.454	4.454	37	2.239	16.666
Jagung	2013	14.859	14.859	27	2.239	40.591
Kacang Hijau	2013	1.877	1.877	11	2.239	1.946
Padi	2014	4.437	4.437	39	609	16.974
Jagung	2014	21.821	21.821	24	609	52.414
Kacang Hijau	2014	1.874	1.874	8	609	1.533
Padi	2015	6.928	6.928	40	606	27.910
Jagung	2015	21.387	21.387	24	606	51.370
Kacang Hijau	2015	2.803	2.803	8	606	2.292
Padi	2016	7.484	7.430	37	928	27.846
Jagung	2016	19.591	18.540	27	928	50.359
Kacang Hijau	2016	2.940	2.737	7	928	1.893
Jagung	2017	24.292	22.593	32	1.442	72.326
Kacang Hijau	2017	2.563	2.563	7	1.442	1.904
Padi	2018	7.908	7.908	46	1.234	36.766
Jagung	2018	26.343	26.018	32	1.234	84.183
Kacang Hijau	2018	3.734	4.734	8	1.234	2.851
Padi	2019	8.274	8.204	48	1.161	39.390
Jagung	2019	26.979	26.955	35	1.161	94.573
Kacang Hijau	2019	4.399	4.396	8	1.161	3.600
Padi	2020	6.574	6.230	50	1.536	31.206
Jagung	2020	27.702	26.534	33	1.536	87.551
Kacang Hijau	2020	4.681	4.054	9	1.536	3.760
Padi	2021	4.963	4.060	46	1.510	18.710
Jagung	2021	25.811	23.617	32	1.510	74.800
Kacang Hijau	2021	5.133	5.118	6	1.510	2.819

3.2. Sampel Data Uji

Sampel data uji terdiri dari beberapa atribut, yaitu tanaman pangan, tahun, luas tanam, luas panen, rata-rata produksi, curah hujan dan produksi yang digunakan sebagai data yang akan diprediksi, seperti yang terlihat pada Tabel 3 berikut ini:

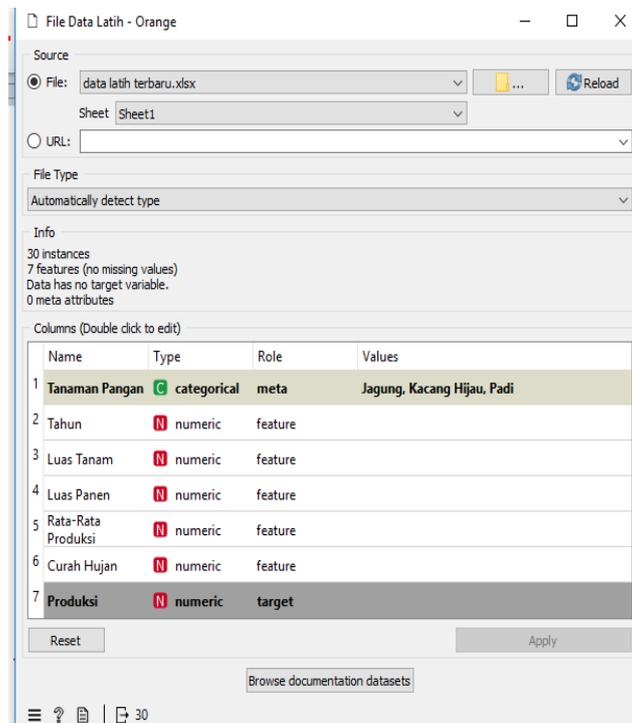
Tabel 3. Data Uji

Tanaman Pangan	Tahun	Luas Tanam (ha)	Luas Panen (ha)	Rata-rata Produksi (ton/ha)	Curah Hujan (mm)	Produksi (ton)
Padi	2022	7.710	7.710	46	1.442	?
Jagung	2022	24.292	22.593	32	1.442	?
Kacang Hijau	2022	2.563	2.563	7	1.442	?
Padi	2023	7.908	7.908	46	1.234	?
Jagung	2023	26.343	26.018	32	1.234	?
Kacang Hijau	2023	3.734	4.734	8	1.234	?
Padi	2024	8.274	8.204	48	1.161	?
Jagung	2024	26.979	26.955	35	1.161	?
Kacang Hijau	2024	4.399	4.396	8	1.161	?

Tanaman Pangan	Tahun	Luas Tanam (ha)	Luas Panen (ha)	Rata-rata Produksi (ton/ha)	Curah Hujan (mm)	Produksi (ton)
Padi	2025	6.574	6.230	50	1.536	?
Jagung	2025	27.702	26.534	33	1.536	?
Kacang Hijau	2025	4.681	4.054	9	1.536	?
Padi	2026	4.963	4.060	46	1.510	?
Jagung	2026	25.811	23.617	32	1.510	?
Kacang Hijau	2026	5.133	5.118	6	1.510	?

3.3. Proses Input Data

Pada tahap proses penginputan sampel data latih dan sampel data uji untuk hasil panen di Kabupaten Malaka, menggunakan komponen *widget file* yang terdapat dalam *tools Orange*. Komponen ini berfungsi untuk membaca sampel data beserta ukuran, jumlah dan jenis *fitur* data yang telah diinputkan ke dalam *tools Orange*. Atribut ditampilkan dalam bentuk kolom yang menjelaskan fungsi-fungsi dan nilai dari atribut yang ada, di mana terdapat nama atribut yang meliputi *type*, yang menampilkan atribut berupa *type* data *numeric* dan *categorical* serta *role* yang berfungsi untuk menetapkan atribut sebagai variabel yang menjadi *target*, *feature* dan *meta*. Tampilan *widget file* pada *tools Orange* terlihat seperti pada Gambar 3 berikut ini :



Gambar 3. Input Data Latih dan Data Uji

Tahap selanjutnya adalah menampilkan fungsi dari komponen *widget data table* yang digunakan untuk menampilkan informasi dari sampel data latih dan sampel data uji yang telah diinputkan ke dalam *tools Orange*, seperti yang terlihat pada Gambar 4 dan Gambar 5 berikut ini :

	Produksi	Tanaman Pangan	Tahun	Luas Tanam	Luas Panen	Rata-Rata Produksi	Curah Hujan
1	30013.00	Padi	2012	8425	8425	35	1355
2	70619.00	Jagung	2012	25944	25944	27	1355
3	2349.00	Kacang Hijau	2012	1927	1927	12	1355
4	16.66	Padi	2013	4454	4454	37	2239
5	40591.00	Jagung	2013	14859	14859	27	2239
6	1946.00	Kacang Hijau	2013	1877	1877	11	2239
7	16974.00	Padi	2014	4437	4437	39	609
8	52414.00	Jagung	2014	21821	21821	24	609
9	1533.00	Kacang Hijau	2014	1874	1874	8	609
10	27910.00	Padi	2015	6928	6928	40	606
11	51370.00	Jagung	2015	21387	21387	24	606
12	2292.00	Kacang Hijau	2015	364	2803	8	606
13	27846.00	Padi	2016	7484	7430	37	928
14	50359.00	Jagung	2016	19591	18540	27	928
15	1893.00	Kacang Hijau	2016	2940	2737	7	928
16	35245.00	Padi	2017	7710	7710	46	1442
17	72326.00	Jagung	2017	24292	22593	32	1442
18	1904.00	Kacang Hijau	2017	2563	2563	7	1442
19	36766.00	Padi	2018	7908	7908	46	1234
20	84183.00	Jagung	2018	26343	26018	32	1234
21	2851.00	Kacang Hijau	2018	3734	4734	8	1234
22	39390.00	Padi	2019	8274	8204	48	1161
23	94573.00	Jagung	2019	26979	26955	35	1161
24	3600.00	Kacang Hijau	2019	4399	4396	8	1161
25	31206.00	Padi	2020	6574	6230	50	1536
26	87551.00	Jagung	2020	27702	26534	33	1536
27	3760.00	Kacang Hijau	2020	4681	4054	9	1536
28	18710.00	Padi	2021	4963	4060	46	1510
29	74800.00	Jagung	2021	25811	23617	32	1510
30	2819.00	Kacang Hijau	2021	5133	5118	6	1510

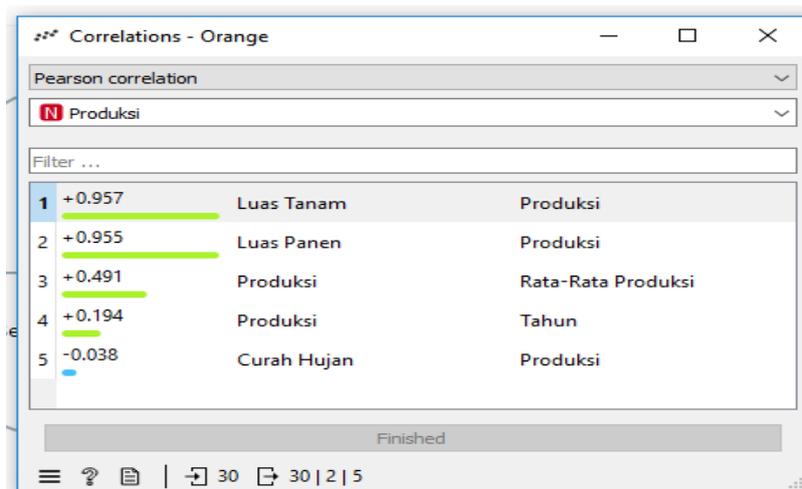
Gambar 4. Data Latih

	Tanaman Pangan	Tahun	Luas Tanam	Luas Panen	Rata-Rata Produksi	Curah Hujan	Produksi
1	Padi	2022	7710	7710	46	1442	?
2	Jagung	2022	24292	22593	32	1442	?
3	Kacang Hijau	2022	2563	2563	7	1442	?
4	Padi	2023	7908	7908	46	1234	?
5	Jagung	2023	26343	26018	32	1234	?
6	Kacang Hijau	2023	3734	4734	8	1234	?
7	Padi	2024	8274	8204	48	1161	?
8	Jagung	2024	26979	26955	35	1161	?
9	Kacang Hijau	2024	4399	4396	8	1161	?
10	Padi	2025	6574	6230	50	1536	?
11	Jagung	2025	27702	26534	33	1536	?
12	Kacang Hijau	2025	4681	4054	9	1536	?
13	Padi	2026	4963	4060	46	1510	?
14	Jagung	2026	25811	23617	32	1510	?
15	Kacang Hijau	2026	5133	5118	6	1510	?

Gambar 5. Data Uji

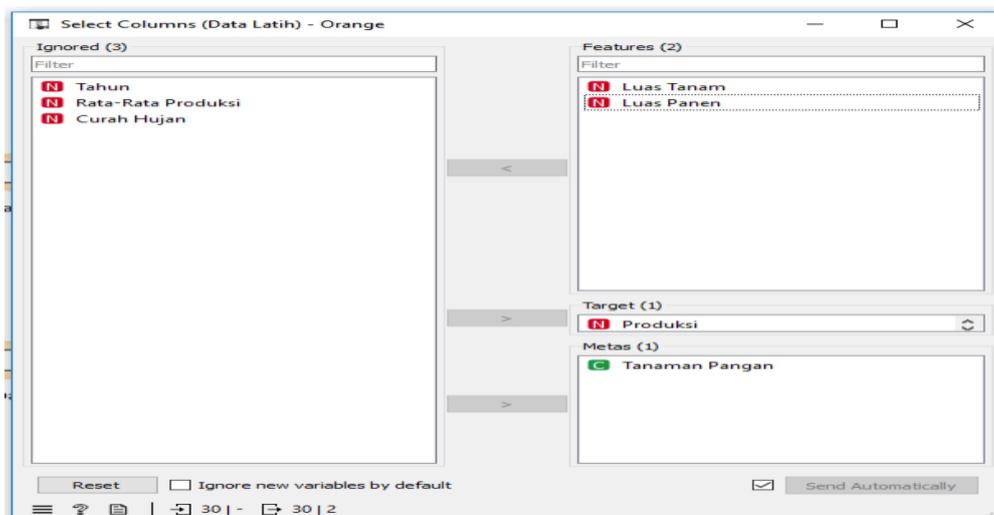
3.4. Preprocessing

Tahap *preprocessing* dilakukan dengan proses *cleaning* data yaitu dengan cara mereduksi atribut-atribut tertentu yang tidak diperlukan. Pada tahapan sebelum *preprocessing*, komponen *widget correlations* digunakan untuk melihat atribut-atribut yang saling mempengaruhi satu sama lain. Atribut yang sangat mempengaruhi hasil produksi digunakan untuk menguji *koefisien* determinasi, mengetahui seberapa besar pengaruh antara variabel independen terhadap variabel dependen, dan mengetahui kecocokan pada model *linear regression* seperti yang terlihat pada Gambar 6.



Gambar 6. Widget Correlations Data

Pada tampilan Gambar 6 terlihat bahwa luas tanam memiliki nilai 0.957 dan luas panen memiliki nilai 0.955. Hal itu menunjukkan bahwa kedua nilai tersebut mendekati 1 yang artinya kedua atribut tersebut mempengaruhi produksi. Untuk ketiga atribut lainnya seperti tahun dengan nilai 0,194, rata-rata produksi dengan nilai 0,491 dan curah hujan dengan nilai di bawah 0,038, ketiga nilai atribut tersebut tidak mempengaruhi produksi karena nilainya tidak mendekati 1. Untuk dapat memisahkan target dan atribut yang saling mempengaruhi dan yang tidak, digunakan komponen *widget select columns* pada *tools Orange* seperti pada Gambar 7 berikut ini :



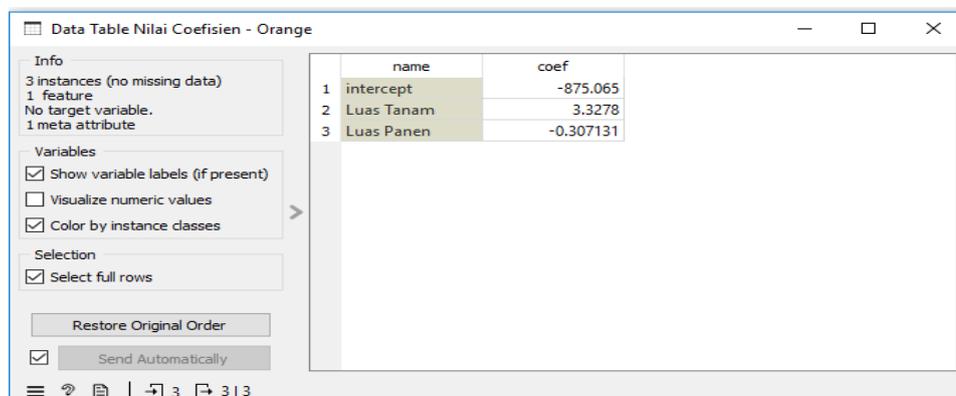
Gambar 7. Proses Pre-processing Data Hasil Pertanian (*Widget Select Columns*)

Seperti pada Gambar 7, yang menjadi target adalah produksi, yang menjadi *feature* atau atribut yang mempengaruhi produksi adalah luas tanam dan luas panen. Sedangkan tahun, rata-rata produksi dan curah hujan tidak mempengaruhi produksi dan *meta* adalah tanaman pangan yang akan diprediksi.

3.5. Algoritma Multiple Linear Regression

Penerapan algoritma *multiple linear regression* menggunakan *widget linear regression* yang telah disediakan oleh aplikasi *Orange* untuk mendapatkan data hasil prediksi. Pada tahap *widget linear regression* ini dilakukan proses perhitungan matriks determinan untuk mendapatkan nilai a yang merupakan nilai *konstanta* dan b_1, b_2 yang merupakan *koefisien regresi* dari luas tanam dan luas panen. Hasil perhitungan dari proses *widget linear regression*, didapatkan nilai *intercept konstanta* (-875.065),

nilai koefisien luas tanam (3.3278) dan nilai koefisien luas panen (-0.307131) yang akan digunakan untuk melakukan proses prediksi hasil panen tanaman pangan pada 5 tahun yang akan datang, seperti yang terlihat pada Gambar 8 berikut ini :



Gambar 8. Hasil Perhitungan Algoritma *Multiple Linear Regression*

3.6. Hasil Prediksi

Setelah melewati semua tahapan pada *tools Orange*, prediksi data hasil pertanian untuk 5 tahun yang akan datang di Kabupaten Malaka dihasilkan. Hasil terlihat pada Tabel 4 berikut :

Tabel 4. Hasil Prediksi

Tanaman Pangan	Tahun	Luas Tanam (ha)	Luas Panen (ha)	Rata-Rata Produksi (ton/ha)	Curah Hujan (mm)	Produksi (ton)
Padi	2022	7.710	7.710	46	1.442	22414
Jagung	2022	24.292	22.593	32	1.442	73025
Kacang Hijau	2022	2.563	2.563	7	1.442	6867
Padi	2023	7.908	7.908	46	1.234	23012
Jagung	2023	26.343	26.018	32	1.234	78798
Kacang Hijau	2023	3.734	4.734	8	1.234	10097
Padi	2024	8.274	8.204	48	1.161	24139
Jagung	2024	26.979	26.955	35	1.161	80627
Kacang Hijau	2024	4.399	4.396	8	1.161	12414
Padi	2025	6.574	6.230	50	1.536	19088
Jagung	2025	27.702	26.534	33	1.536	83162
Kacang Hijau	2025	4.681	4.054	9	1.536	13457
Padi	2026	4.963	4.060	46	1.510	14394
Jagung	2026	25.811	23.617	32	1.510	77765
Kacang Hijau	2026	5.133	5.118	6	1.510	14635

Berdasarkan tabel hasil prediksi, metode *multiple linear regression* dapat digunakan untuk memprediksi hasil panen dari data uji yang tersedia. Perhitungan menunjukkan hasil untuk tahun 2022 sampai dengan tahun 2026 dengan atribut luas tanam dan luas panen terhadap hasil panen produksi. Berikut ini merupakan *interpretasi* dari persamaan model *multiple linear regression*:

$$Y = a + b_1X_1 + b_2X_2:$$

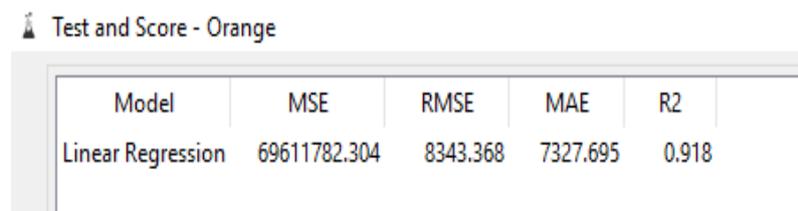
$$Y = -875.065 + 3.3278X_1 - 0.307131X_2$$

Untuk model persamaan *multiple linear regression* tersebut dapat dijelaskan sebagai berikut:

1. Nilai konstanta (a) memiliki nilai negatif sebesar -875.065. Nilai tersebut menunjukkan pengaruh negatif (berlawanan arah) terhadap variabel independen dan variabel dependen. Apabila X_1 dan X_2 bernilai 0% atau tidak mengalami perubahan, maka nilai Y adalah -875.065.
2. Nilai koefisien regresi untuk variabel luas tanam (X_1) yaitu sebesar 3.3278. Nilai tersebut menunjukkan pengaruh positif atau searah antara variabel independen dan variabel dependen. Artinya jika variabel independen (X_1) mengalami kenaikan 1% maka variabel dependen (Y) akan mengalami kenaikan 3.3278 dengan asumsi bahwa variabel lainnya tetap konstanta.
3. Nilai koefisien untuk variabel luas panen (X_2) yaitu sebesar -0.307131. Nilai tersebut menunjukkan pengaruh negatif atau berlawanan arah antara variabel independen dan variabel dependen. Artinya jika variabel independen (X_2) mengalami kenaikan 1% maka variabel dependen (Y) akan mengalami penurunan sebesar 0.307131 dengan asumsi variabel lainnya tetap konstanta.

3.7. Analisis Hasil

Berdasarkan nilai konstanta dan koefisien yang telah diperoleh dengan menggunakan metode *multiple linear regression*, dilakukan perhitungan untuk memprediksi hasil panen atau produksi menjadi keluaran angka yang dapat ditentukan berapa banyak. Pada tahap analisis hasil dilakukan proses pengujian dengan menggunakan komponen *Widget Test and Score* terhadap metode hasil nilai prediksi, yang dapat digunakan untuk mengukur tingkat akurasi dari hasil perkiraan suatu model seperti yang terlihat pada Gambar 9



Model	MSE	RMSE	MAE	R2
Linear Regression	69611782.304	8343.368	7327.695	0.918

Gambar 9. Hasil Pengujian (*Widget Test and Score*)

Dari proses pengujian hasil prediksi dengan pengujian pada data latih di Gambar 9 dengan menggunakan seluruh data untuk pelatihan dan kemudian dilakukan pengujian, didapatkan nilai MSE (*Mean Square Error*) yang digunakan untuk menghitung rata-rata dari selisih kuadrat antara nilai prediksi dan nilai aktual. RMSE (*Root Mean Square Error*) digunakan dalam menghitung rata-rata dari selisih kuadrat antara nilai prediksi dan nilai aktual, yang kemudian diambil akar kuadratnya. MAE (*Mean Absolute Error*) digunakan dalam menghitung rata-rata dari selisih absolut antara nilai prediksi dan nilai aktual. R2 (*Coefficient Determination*) digunakan untuk menunjukkan tidak adanya pengaruh antara variabel independen dengan variabel dependen. Dari nilai hasil uji keakuratan model, MSE (*Mean Square Error*) menghasilkan nilai sebesar 69611782.304 dengan rumus sebagai berikut:

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y})^2 \quad (6)$$

Ket :

y_i = Nilai aktual

\hat{y} = Nilai Prediksi

N = Banyaknya data

$$MSE = 2.088.353.469 / 30 \\ = 69611782.304$$

Pada RMSE (*Root Mean Square Error*) terlihat nilai hasil uji sebesar 8343.368, dengan menggunakan rumus sebagai berikut:

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y})^2} \quad (7)$$

Ket :

Yi = Nilai aktual
 \hat{y} = Nilai Prediksi
 N = Banyaknya data
 RMSE = $\sqrt{MSE} = \sqrt{69611782.304}$
 = 8343.368

Pada MAE (*Mean Absolute Error*) terlihat nilai hasil uji sebesar 7327.695, dengan menggunakan rumus sebagai berikut :

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}| \quad (8)$$

Ket :

yi = Nilai aktual
 \hat{y} = Nilai Prediksi
 N = Banyaknya data
 MAE = $219830,8 / 30$
 = 7327.695

Nilai hasil uji R2 sebesar 0.918 diperoleh dengan menggunakan rumus sebagai berikut:

$$r^2 = \frac{(b_1 \sum x_1 y) + (b_2 \sum x_2 y)}{\sum y^2} \quad (9)$$

$r^2 = (3,327804473 * 7669037947) + (-0,307130895 * 7417415411) / 25331294822$
 $r^2 = (25521058784 + (-2278117430)) / 25331294822$
 $r^2 = 23242941353 / 25331294822$
 $r^2 = 0,918$
 $r = 0,957892667$

Untuk itu dapat dinyatakan bahwa nilai R2 (*Coefficient Determination*) semakin mendekati nilai 1, maka semakin baik hasil dari kualitas model tersebut dan semakin kuat pengaruh variabel independen terhadap variabel dependen [17] .

4. Kesimpulan

Berdasarkan proses pengujian yang telah dilakukan untuk memprediksi hasil panen pada tanaman padi, jagung dan kacang hijau di Kabupaten Malaka untuk 5 tahun yang akan datang, terdapat dua variabel yang saling mempengaruhi yaitu luas tanam dan luas panen. Selain itu, dalam penelitian ini dihasilkan sebuah model persamaan *multiple linear regression*, yaitu $Y = -875.065 + 3.3278X_1 - 0.307131X_2$. Dari data hasil pertanian terlihat hasil uji koefisien determinasi sebesar 0.918, yang menunjukkan bahwa 91.8% variasi variabel dependen dapat dijelaskan oleh variabel independen di dalam model . Terdapat juga tiga analisis statistik yang digunakan untuk melihat *performance* dari metode *multiple linear regression*, yaitu MSE sebesar 69611782.304, RMSE sebesar 8343.368 dan MAE sebesar 7327.695.

Referensi

- [1] E. F. Laili andi H. C. Dinarto, "Pengembangan Kawasan Pertanian Berbasis Tanaman Pangan di Kecamatan Wuluhan, Kabupaten Jember," *Journal of Regional and Rural Development Planning*, vol. 2, no. 3, p. 209, Dec. 2018, doi: 10.29244/jp2wd.2018.2.3.209-217.
- [2] JDIH BPK RI, *Undang-Undang Republik Indonesia Nomor 7 Tahun 1996 Tentang Pangan*. 1996. Accessed: Apr. 14, 2024. [Online]. Available: <https://peraturan.bpk.go.id/Details/46097/uu-no-7-tahun-1996>

- [3] BP4D Kabupaten Malaka, R. Lumbantoruan, A. Nomleni, and U. Kristen Indonesia, *Rencana Tata Ruang Wilayah Kabupaten Malaka 2017-2036*. 2017. [Online]. Available: <https://journals.upi-yai.ac.id/index.php/IKRAITH-ABDIMAS/issue/archive>
- [4] D. P. Lamondjong and M. Hardjianto, “Data Mining untuk Memprediksi Jumlah Penjualan Hasil Pertanian Menggunakan Algoritma Forecasting (Studi Kasus : Dinas Pertanian Kabupaten Banggai),” *Jurnal Teknologi Informasi Respati*, vol. 6, no. 2, pp. 94–102, 2021, Accessed: Apr. 14, 2024. [Online]. Available: <https://jti.respati.ac.id/index.php/jurnaljti/article/view/405/351>
- [5] Y. S. , M. Y. Siswo Adiguno, “Prediksi Peningkatan Omset Penjualan Menggunakan Metode Regresi Linier Berganda”, Accessed: Apr. 14, 2024. [Online]. Available: <https://ojs.trigunadharma.ac.id/index.php/jsi/article/view/5331>
- [6] M. Kafil, “Penerapan Metode K-Nearest Neighbors Untuk Prediksi Penjualan Berbasis Web Pada Boutiq Dealove Bondowoso,” 2019.
- [7] M. K. H. H. M. D. Salahuddin Salahuddin, “Rancang Bangun Aplikasi Machine Learning Prediksi Hasil Panen Buah Pinang (Areca Catechu) Menggunakan Metode Regresi Linier Berganda,” in *Prosiding Seminar Nasional Politeknik Negeri Lhokseumawe* , 2022. Accessed: Apr. 14, 2024. [Online]. Available: <https://e-jurnal.pnl.ac.id/semnaspnl/article/view/3475>
- [8] M. Adha and E. Utami, “Prediksi Produksi Jagung Menggunakan Algoritma Apriori Dan Regresi Linear Berganda (Studi Kasus : Dinas Pertanian Kabupaten Dompu),” *JUPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 7, no. 3, pp. 803–820, 2022, Accessed: Apr. 14, 2024. [Online]. Available: <https://jurnal.stkipgritulungagung.ac.id/index.php/jupi/article/view/3139>
- [9] M. Firman, M. Assidiq, and S. Syarli, “Sistem Prediksi Hasil Pertanian Durian Dengan Metode Regresi Linear Berganda,” *Journal Peqquruang: Conference Series*, vol. 5, no. 1, p. 350, May 2023, doi: 10.35329/jp.v5i1.3199.
- [10] O. Muharromah, N. Suarna, and W. Prihartono, “Implementasi Algoritma Regresi Linear Berganda Untuk Prediksi Produksi Padi Di Kabupaten Cirebon,” 2023.
- [11] R. Andia, S. Eka Permana, and T. Handayani, “Peramalan Hasil Panen Padi Kabupaten Cirebon Menggunakan Algoritma Regresi Linear Berganda,” 2024.
- [12] R. Astuti and F. M. Basysyar, “Implementasi Data Mining Untuk Prediksi Produksi Bawang Merah Di Kabupaten Brebes,” 2024. [Online]. Available: <https://brebeskab.bps.go.id>.
- [13] M. Buchori Ibrahim *et al.*, *Metode Penelitian Berbagai Bidang Keilmuan (Panduan & Referensi)*, vol. 1. 2023. [Online]. Available: www.sonpedia.com
- [14] H. W. Herwanto, T. Widiyaningtyas, and P. Indriana, “Penerapan Algoritma Linear Regression untuk Prediksi Hasil Panen Tanaman Padi,” 2019.
- [15] M. Joefitra Zaqy, R. Mentari, and M. Iqbal, “Jurnal Teknologi Sistem Informasi dan Sistem Komputer TGD Implementasi Data Mining Untuk Memprediksi Penjualan Toko Roti Mawar Menggunakan Regresi Linier Berganda,” *Jurnal Teknologi Sistem Informasi dan Sistem Komputer TGD*, vol. 6, no. 2, pp. 385–392, 2023, [Online]. Available: <https://ojs.trigunadharma.ac.id/index.php/jsk/index>
- [16] 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.
- [17] S. C. Chapra and R. P. Canale, “Numerical methods for Engineers Sixth Edition,” McGraw-Hill Higher Education, 2010, pp. 454–481.