

# Implementasi Metode *Multiple Linear Regression* untuk Memprediksi Intensitas Cahaya untuk Pembungaan Buah Tomat Pada Media Kultur Jaringan

## *Implementation of Multiple Linear Regression Method To Predict Light Intensity For Tomato Flowering On Tissue Culture Media*

Vincentius Riandaru Prasetyo<sup>1</sup>, Wina Dian Savitri<sup>2</sup>, Johan Sukweenadhi<sup>3</sup>, Fenny Irawati<sup>4</sup>,  
Mohammad Farid Naufal<sup>5</sup>, Solichul Huda<sup>6</sup>

<sup>1,5</sup> Program Studi Teknik Informatika, Fakultas Teknik, Universitas Surabaya

<sup>2,3,4</sup> Program Studi Biologi, Fakultas Teknobiologi, Universitas Surabaya

<sup>6</sup> Program Studi Teknik Informatika, Universitas Dian Nuswantoro

E-mail: <sup>1</sup>vincent@staff.ubaya.ac.id, <sup>2</sup>winasavitri@staff.ubaya.ac.id, <sup>3</sup>sukwee@staff.ubaya.ac.id,  
<sup>4</sup>fenny\_i@staff.ubaya.ac.id, <sup>5</sup>faridnaufal@staff.ubaya.ac.id, <sup>6</sup>solichul.huda@dsn.dinus.ac.id

### Abstrak

Buah tomat merupakan salah satu buah yang sering dijumpai di Indonesia. Permintaan tomat juga terus meningkat dari tahun ke tahun di Indonesia. Keterbatasan lahan menjadi kendala dalam memenuhi permintaan tersebut. Oleh karena itu, tomat dapat dikembangkan melalui media kultur jaringan untuk menghasilkan buah yang sehat dan cepat. Salah satu faktor yang menentukan keberhasilan dalam budidaya tomat melalui media kultur jaringan adalah intensitas cahaya yang digunakan. Metode *multiple linear regression* diimplementasikan untuk dapat memperkirakan intensitas cahaya yang tepat. Fitur yang akan menjadi prediktor dalam penelitian ini adalah jenis media tanam dan panjang kecambah yang diharapkan. Berdasarkan evaluasi yang telah dilakukan, nilai *mean absolute error* (MAE) dan *root mean square error* (RMSE) yang dihasilkan yaitu 1,262 dan 1,427 dengan teknik *10-fold cross validation*. Hal ini menandakan bahwa sistem cukup akurat dalam melakukan prediksi.

**Kata kunci:** *multiple linear regression*, prediksi, intensitas cahaya, tomat

### Abstract

Tomato is one of the fruits that are often found in Indonesia. The demand for tomatoes also continues to increase yearly in Indonesia. Limited land is an obstacle in meeting this demand. Therefore, tomatoes can be developed through tissue culture media to produce healthy and fast fruit. One factor determining tomato cultivation's success through tissue culture media is the intensity of light used. A multiple linear regression method is implemented to estimate the right light intensity. The features that will be predictors in this study are the planting medium and the expected length of sprouts. Based on the evaluation that has been done, the resulting mean absolute error (MAE) and root mean square error (RMSE) are 1.262 dan 1.427 using 10-fold cross validation technique. This indicates that the system is entirely accurate in making predictions.

**Keywords:** *multiple linear regression*, prediction, light intensity, tomato

## 1. PENDAHULUAN

Tanaman tomat atau yang memiliki nama latin *Lycopersium esculentum* Mill., adalah tanaman pertanian yang umum dijumpai di Indonesia. Buah tomat memiliki rasa khas yang memadukan rasa manis dan asam, menjadikan tomat sebagai buah yang populer. Selain itu, permintaan pasar terhadap tomat terus meningkat setiap tahunnya. Di Indonesia, permintaan pasar tomat pada 2018 sebesar 976.772 ton, dan pada 2019 meningkat 1.020.333 ton atau 4,46 persen [1]. Akan tetapi, hal tersebut tidak berbanding lurus dengan jumlah lahan pertanian yang ada. Banyak lahan-lahan pertanian yang beralih fungsi menjadi lahan non pertanian seperti:

perumahan, gedung-gedung perkantoran, industri pabrik, dan sebagainya [2]. Selain itu, faktor lain yang dapat menjadi penghambat dalam produksi buah tomat adalah hama dan penyakit. Hama dan penyakit ini sering kali menyerang saat proses pembungaan buah tomat, di mana proses pembungaan merupakan tahapan penting suatu tanaman, sebelum berkembang menjadi buah. Kultur jaringan merupakan salah satu cara untuk mendapatkan bibit tomat yang sehat dan bebas penyakit. Teknik *in vitro* adalah metode yang digunakan untuk membudidayakan tanaman dalam kultur jaringan. Teknik *in vitro* akan memberikan keturunannya sifat yang sama seperti induknya [3]. Dengan budidaya menggunakan media kultur jaringan ini, memungkinkan produksi tomat dalam skala besar dengan waktu yang relatif cepat.

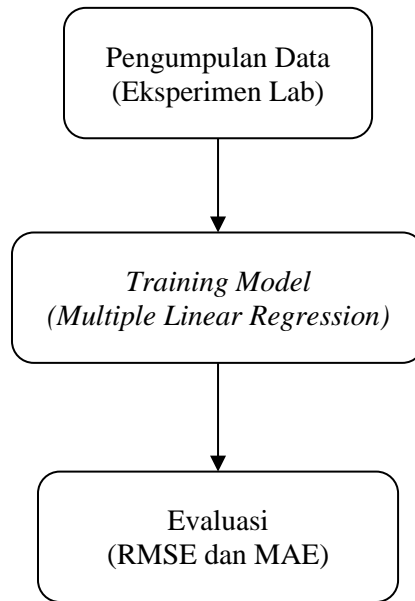
Salah satu faktor yang berpengaruh dalam keberhasilan budidaya tanaman dengan media kultur jaringan adalah jenis media yang dipakai. Jenis media kultur jaringan dapat bermacam-macam, seperti BAP (*Benzyl Amino Purine*), MS (*Murashige and Skoog*), IAA (*Indole Acetic Acid*), dan sebagainya, atau dapat merupakan kombinasi dari jenis-jenis media tersebut. Jenis media atau kombinasi dari beberapa jenis media yang dipakai akan menentukan intensitas cahaya yang harus dipakai. Intensitas cahaya akan berpengaruh pada proses fotosintesis tanaman, sehingga menentukan intensitas cahaya yang tepat akan menyebabkan pertumbuhan tanaman menjadi optimal [4]. Salah satu metode yang dapat digunakan untuk memprediksi suatu kejadian berdasarkan variabel-variabel tertentu adalah *multiple linear regression*.

*Multiple linear regression* merupakan metode yang bertujuan untuk menguji pengaruh beberapa variabel bebas terhadap variabel terikat [5]. Beberapa kelebihan dari metode *multiple linear regression*, yaitu kemampuan untuk menggeneralisasi dan mengekstrak pola data tertentu, kemampuan untuk belajar meskipun ada ketidakpastian, dan kemampuan untuk melakukan perhitungan secara bersamaan untuk mempercepat proses [6]. Herwanto et.al [7] pada penelitiannya menggunakan metode *multiple linear regression* untuk memprediksi hasil panen padi di kota Lamongan. Variabel terikat yang digunakan adalah hasil panen dan variabel bebas yang digunakan adalah luas lahan, jumlah bibit, pupuk area, dan pupuk NPK Phonska. Nilai RMSE (*Root Mean Square Error*) yang dihasilkan pada penelitian tersebut yaitu sebesar 0,432, yang menunjukkan bahwa variasi nilai dari model prediksi yang terbentuk mendekati akurat [7]. Penelitian lainnya yang mengimplementasikan metode *multiple linear regression* juga dilakukan oleh Tarigan [8]. Metode *multiple linear regression* digunakan untuk memprediksi pertumbuhan tanaman akuaponik berdasarkan parameter-parameter tertentu. Pada penelitian tersebut dihasilkan nilai akurasi tertinggi untuk model prediksi berdasarkan luas tanaman yaitu 72,4% [8]. Penelitian yang secara khusus memprediksi waktu pembungaan tanaman tomat pernah dilakukan oleh Syakur [9]. Pada penelitian tersebut, metode *artificial neural network* diimplementasikan untuk memprediksi waktu pembungaan dan matang fisiologis tanaman tomat berdasarkan kondisi iklim mikro pada suatu *green house*. Selain memprediksi waktu pembungaan, penelitian tersebut juga memprediksi satuan panas yang sesuai bagi perkembangan tanaman tomat. Hasil penelitian menunjukkan nilai koefisien korelasi (R) yang dihasilkan adalah 0,43655 dan nilai RMSE adalah 5,2314 [9].

Berbeda dengan penelitian-penelitian sebelumnya, penelitian ini bertujuan untuk menentukan atau memprediksi intensitas cahaya pada proses pembungaan buah tomat pada media kultur jaringan tertentu. Prediksi dilakukan dengan mengimplementasikan metode *multiple linear regression*, di mana variabel bebasnya adalah jenis media kultur jaringan yang digunakan dan panjang kecambah. Selain itu, evaluasi terhadap model prediksi yang dihasilkan oleh metode *multiple linear regression*, tidak hanya dengan cara mengukur nilai RMSE, tetapi juga dilakukan pengukuran nilai MAE (*Mean Absolute Error*).

## 2. METODE PENELITIAN

Metode penelitian pada penelitian ini terdiri dari tiga tahapan yaitu pengumpulan data, training model, dan evaluasi. Gambar 1 memperlihatkan alur tahapan penelitian ini.



Gambar 1 Alur Tahapan Penelitian

### 2.1 Pengumpulan Data

Pada penelitian ini, *dataset* yang digunakan untuk *training model* didapatkan melalui percobaan pembungaan buah tomat pada media kultur jaringan di laboratorium Fakultas Teknobiologi Universitas Surabaya. Media kultur jaringan yang dipakai pada percobaan ini terdiri dari lima jenis, seperti yang diperlihatkan pada Tabel 1. Sedangkan untuk intensitas cahaya yang dipakai yaitu berkisar antara 1700 hingga 3820 lux.

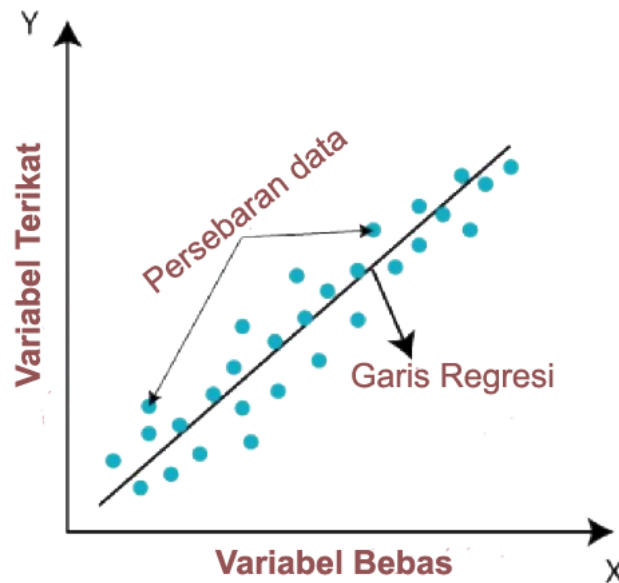
Tabel 1 Jenis Media Kultur Jaringan

Nomor	Jenis Media
1	MSO
2	MS + BAP 1
3	MS + IAA 0.5
4	MS + BAP 0.75 + 30 $\mu$ M SN
5	MS + BAP 0.75 + 45 $\mu$ M SN

Percobaan dilakukan dengan menanam benih atau biji tomat pada media tertentu dan intensitas cahaya tertentu selama dua minggu. Benih tomat yang telah mengeluarkan atau menumbuhkan kecambah akan diukur panjang kecambahnya, yang nantinya akan digunakan sebagai input untuk pembentukan model prediksi pada proses *training*.

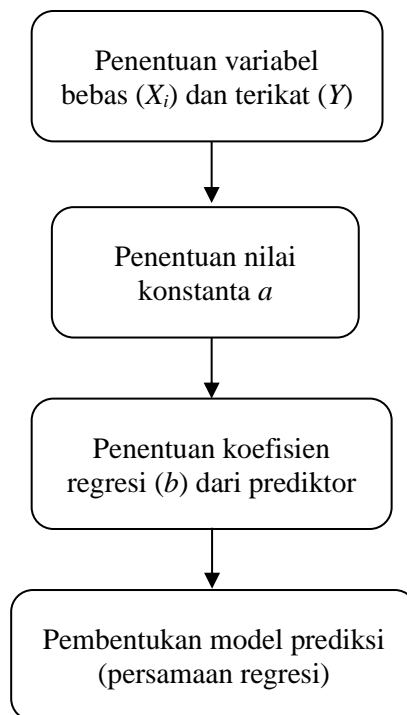
### 2.2 Training Model

Setelah *dataset* dikumpulkan, maka tahapan selanjutnya adalah melakukan proses *training* untuk membentuk model prediksi. Metode yang digunakan untuk proses training pada penelitian ini menggunakan *multiple linear regression*. Metode *multiple linear regression* merupakan pengembangan dari metode regresi linier sederhana. Pada dasarnya metode regresi linier akan memprediksi nilai numerik. Metode regresi linier berusaha mencari hubungan antara prediktor (X) dan nilai target (Y) yang dapat membuat suatu pola tertentu [10]. Gambar 2 menunjukkan ilustrasi dari metode regresi linier.



Gambar 2 Ilustrasi Metode Regresi Linier

Langkah-langkah pembentukan model prediksi dengan metode *multiple linear regression*, ditunjukkan pada Gambar 3. Pembentukan model prediksi diawali dengan menentukan atribut-atribut data yang menjadi variabel bebas dan terikat. Langkah selanjutnya adalah menentukan nilai dari konstanta  $a$  melalui persamaan (1), di mana  $Y$  adalah variabel terikat atau target dan  $X$  adalah variabel bebas atau prediktor. Setelah konstanta  $a$  didapatkan, maka akan dihitung koefisien regresi dari prediktor melalui persamaan (2). Terakhir, model prediksi dari *multiple linear regression* didapatkan melalui persamaan (3) [11].



Gambar 3 Tahapan Pembentukan Model Prediksi *Multiple Linear Regression*

$$a = \frac{(\sum_{i=1}^{i=n} y) (\sum_{i=1}^{i=n} x^2) - (\sum_{i=1}^{i=n} x) (\sum_{i=1}^{i=n} x.y)}{n (\sum_{i=1}^{i=n} x^2) - (\sum_{i=1}^{i=n} x)^2} \quad (1)$$

$$b = \frac{n (\sum_{i=1}^{i=n} x.y) - (\sum_{i=1}^{i=n} x) (\sum_{i=1}^{i=n} y)}{n (\sum_{i=1}^{i=n} x^2) - (\sum_{i=1}^{i=n} x)^2} \quad (2)$$

$$Y = a + b_1X_1 + b_2X_2 + \dots + b_nX_n \quad (3)$$

### 2.3 Evaluasi

Evaluasi dilakukan untuk menguji model prediksi yang dihasilkan dari proses *training* sebelumnya. Pada penelitian ini, evaluasi dilakukan dengan menghitung nilai *mean absolute error* (MAE). MAE mengukur besarnya rata-rata kesalahan pada serangkaian prediksi. Tujuan lain MAE adalah mengukur perbedaan antara hasil prediksi dengan pengamatan yang sesuai [12]. MAE juga telah banyak digunakan dalam mengevaluasi keakuratan sistem prediksi, di mana dihitung dengan persamaan (4) [13]. Di mana  $\tau$  berarti peringkat prediksi,  $r$  adalah nilai sebenarnya dalam pengujian *dataset*, dan  $N$  adalah jumlah pasangan data pengujian dan hasil prediksi.

$$MAE = \frac{\sum_{i=1}^{i=n} |\tau_n - r_n|}{N} \quad (4)$$

Evaluasi lainnya yang dilakukan pada penelitian ini adalah mengukur nilai *root mean square error* (RMSE). RMSE adalah akar kuadrat dari rata-rata kuadrat semua kesalahan yang terjadi saat prediksi. RMSE adalah cara umum untuk mengukur kesalahan model dalam memprediksi data kuantitatif [14]. Persamaan (5) menunjukkan cara perhitungan dari RMSE, di mana  $y'$  adalah nilai hasil prediksi,  $y$  adalah nilai sebenarnya, dan  $n$  adalah banyaknya data yang digunakan [14].

$$MAE = \sqrt{\frac{\sum_{i=1}^{i=n} (y'_i - y_i)^2}{n}} \quad (5)$$

## 3. HASIL DAN PEMBAHASAN

Berdasarkan tahapan pengumpulan data yang telah dilakukan, jumlah *dataset* yang didapatkan adalah 90 data. *Dataset* yang digunakan untuk pembentukan model prediksi terdiri dari 3 data utama yaitu: intensitas cahaya yang digunakan, media tanam yang digunakan, dan panjang kecambah yang dihasilkan. Untuk data media tanam, dikarenakan data berupa teks, maka data akan diubah terlebih dahulu dengan memberikan notasi 1 sampai 5, di mana 1 untuk MSO, 2 untuk MS + BAP 1, 3 untuk MS + IAA 0.5, 4 untuk MS + BAP 0.75 + 30uM SN, dan 5 untuk MS + BAP 0.75 + 45uM SN. Hal ini dikarenakan metode *multiple linear regression* hanya menerima data yang berupa bilangan. Selanjutnya, *dataset* yang telah siap akan dilakukan proses *training* menggunakan metode *multiple linear regression*, di mana variabel bebas atau prediktornya adalah jenis media tanam dan panjang kecambah. Sedangkan untuk variabel terikat atau targetnya adalah intensitas cahaya. Setelah proses *training* selesai dilakukan, maka sistem siap untuk melakukan prediksi.

Sistem prediksi dibangun dengan menggunakan bahasa pemrograman Python, di mana user dapat memilih jenis media tanam apa yang digunakan dan panjang kecambah yang diharapkan, seperti yang diperlihatkan pada Gambar 4. Setelah *user* menentukan jenis media tanam dan mengisi panjang kecambah yang diharapkan, maka sistem akan melakukan prediksi berdasarkan model yang sudah terbentuk sebelumnya. Contoh hasil prediksi intensitas cahaya dapat dilihat pada Gambar 5.

```

Select one (1/2): 1

====Planting Media Used====
1. MSO
2. MS + BAP 1
3. MS + IAA 0.5
4. MS + BAP 0.75 + 30uM SN
5. MS + BAP 0.75 + 45uM SN

Select one (1-5): 

Select one (1-5): 4

Expected Plantlet Length (cm): 
    
```

Gambar 4 Inputan Media Tanam dan Panjang Kecambah

```

Expected Plantlet Length (cm): 2

====Prediction Result====

Light Intensity: 2828.7 Lux
    
```

Gambar 5 Hasil Prediksi Intensitas Cahaya

Sistem prediksi yang telah terbentuk akan dievaluasi lebih lanjut untuk mengetahui apakah model prediksi yang dihasilkan sudah baik. Evaluasi dilakukan dengan mengukur nilai MAE. Apabila nilai MAE dan RMSE yang didapatkan semakin kecil atau mendekati 0, maka kemungkinan hasil prediksi semakin akurat [15]. Perhitungan nilai MAE dan RMSE juga melibatkan teknik *10-fold cross validation*. Teknik tersebut akan melakukan perhitungan nilai MAE dan RMSE sebanyak sepuluh kali, di mana masing-masing perhitungan dilakukan pada *dataset* yang terdiri dari 90% dari total *dataset* yang dipilih secara acak, yang akan digunakan sebagai data *training*, dan 10% sisanya digunakan sebagai data *testing* untuk validasi. Tabel 2 menunjukkan hasil evaluasi pertama MAE dan RMSE dengan teknik *10-fold cross validation*.

Tabel 2 Hasil Evaluasi Pertama MAE dan RMSE

Fold-	Nilai MAE	Nilai RMSE
1	1,100	1,302
2	1,025	1,275
3	0,890	1,127
4	1,100	1,192
5	1,454	1,583
6	1,154	1,335
7	1,237	1,378
8	1,509	1,581
9	1,405	1,461
10	1,747	1,882
<b>Rata-rata</b>	<b>1,262</b>	<b>1,427</b>

Dari hasil evaluasi yang dilakukan, didapatkan rata-rata nilai MAE dan RMSE dari model prediksi menggunakan metode *multiple linear regression* adalah 1,262 dan 1,427. Hal ini menandakan bahwa model prediksi yang terbentuk cukup baik dalam melakukan prediksi. Akan tetapi, nilai MAE dan RMSE ini masih dapat ditingkatkan lagi dengan menambahkan *dataset* agar model prediksi yang terbentuk lebih baik lagi. Untuk membuktikan hal tersebut, maka dilakukan percobaan kedua dengan menambahkan 30 dataset baru, sehingga total dataset yang digunakan untuk proses evaluasi berjumlah 120. Tabel 3 menunjukkan hasil evaluasi kedua MAE dan RMSE dengan teknik *10-fold cross validation*. Hasil percobaan kedua, didapatkan nilai MAE dan RMSE yang lebih kecil yaitu 1,239 dan 1,417. Hal tersebut telah membuktikan bahwa semakin banyak *dataset* yang digunakan, maka model prediksi yang terbentuk semakin baik lagi.

Tabel 3 Hasil Evaluasi Kedua MAE dan RMSE

Fold-	Nilai MAE	Nilai RMSE
1	1,130	1,332
2	1,486	1,667
3	1,129	1,239
4	1,169	1,270
5	1,203	1,361
6	0,889	1,066
7	1,426	1,598
8	1,118	1,375
9	1,404	1,583
10	1,443	1,568
<b>Rata-rata</b>	<b>1,239</b>	<b>1,417</b>

#### 4. KESIMPULAN DAN SARAN

Berdasarkan nilai MAE dan RMSE yang dihasilkan pada tahapan evaluasi sebelumnya, maka dapat disimpulkan bahwa sistem yang dibangun cukup akurat dalam melakukan prediksi intensitas cahaya pada kasus pembungaan buah tomat di media kultur jaringan. Nilai MAE dan RMSE yang dihasilkan masih dapat dikurangi lagi dengan menambahkan *dataset* di kemudian hari untuk meningkatkan hasil prediksi. Selain itu, dapat dilakukan percobaan di kemudian hari dengan membandingkan metode-metode prediksi lainnya, untuk mengetahui metode mana yang paling akurat dalam melakukan prediksi.

#### UCAPAN TERIMA KASIH

Terima kasih kepada Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi yang telah mendanai penelitian ini melalui skema PDUPT (Penelitian Dasar Unggulan Perguruan Tinggi).

#### DAFTAR PUSTAKA

- [1] E. Halid, A. Mutalib, S. Inderiati and R. D, "Pertumbuhan Dan Produksi Tanaman Tomat (*Lycopersium Esculentum* Mill.) Pada Pemberian Berbagai Dosis Bubuk Cangkang Telur," *J. Agropiantae*, vol. 10, no. 1, pp. 59-66, 2021.
- [2] A. C. Kusumastuti, L. M. Kolopaking and B. Barus, "Faktor Yang Mempengaruhi Alih Fungsi Lahan Pertanian Pangan Di Kabupaten Pandeglang," *Sodality: Jurnal Sosiologi Pedesaan*, vol. 6, no. 2, pp. 131-136, 2018.
- [3] E. Suryanti and Melissa, "Perbanyakan Tanaman Tomat Dengan Menggunakan Bap Dan Naa Secara In Vitro," *Jurnal Bioterdidik Wahana Ekspresi Ilmiah*, vol. 5, no. 6, pp. 1-7, 2017.

- [4] Lazuardi, "Pengaruh Pemberian 2,4-D (Diclorophenoxy Acetic Acid) Dan Bap (Benzyl Amino Purine) Terhadap Induksi Kalus Pada Tanaman Padi Ladang," *Jurnal Sainika*, vol. 12, no. 1, pp. 36-46, 2012.
- [5] A. A.-F. N. Wahyudin, A. Primajaya and A. S. Y. Irawan, "Penerapan Algoritma Regresi Linear Berganda Pada Estimasi Penjualan Mobil Astra Isuzu," *Techno.COM*, vol. 19, no. 4, pp. 364-374, 2020.
- [6] Amrin, "Data Mining Dengan Regresi Linier Berganda Untuk Peramalan Tingkat Inflasi," *Jurnal Techno Nusa Mandiri*, vol. XIII, no. 1, pp. 74-79, 2016.
- [7] H. W. Herwanto, T. Widiyaningtyas and P. Indriana, "Penerapan Algoritme Linear Regression untuk Prediksi Hasil Panen Tanaman Padi," *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, vol. 8, no. 4, pp. 364-370, 2019.
- [8] F. M. A. Tarigan, "Prediksi Pertumbuhan Tanaman Akuaponik Menggunakan Length Features Extraction Dan Multiple Linear Regression," Universitas Sumatera Utara, Medan, 2021.
- [9] A. Syakur, "Analisis Iklim Mikro Di Dalam Rumah Tanaman Untuk Memprediksi Waktu Pembungaan Dan Matang Fisiologis Tanaman Tomat Dengan Menggunakan Metode Artificial Neural Network," *Jurnal Agroland*, vol. 18, no. 2, pp. 97-103, 2011.
- [10] R. Seethalakshmi, "Analysis of stock market predictor variables using Linear Regression," *International Journal of Pure and Applied Mathematics*, vol. 119, no. 15, pp. 369-378, 2018.
- [11] V. R. Prasetyo, H. Lazuardi, A. A. Mulyono and C. Lauw, "Penerapan Aplikasi RapidMiner Untuk Prediksi Nilai Tukar Rupiah Terhadap US Dollar Dengan Metode Regresi Linier," *Jurnal Nasional Teknologi dan Sistem Informasi*, vol. 7, no. 1, pp. 8-17, 2021.
- [12] T. Chai and R. R. Draxler, "Root mean square error (RMSE) or mean absolute error (MAE)? Arguments against avoiding RMSE in the literature," *Geoscientific Model Development*, vol. 7, pp. 1247-1250, 2014.
- [13] W. Wang and Y. Lu, "Analysis of the Mean Absolute Error (MAE) and the Root Mean Square Error (RMSE) in Assessing Rounding Model," *IOP Conf. Series: Materials Science and Engineering*, vol. 324, pp. 1-10, 2018.
- [14] F. I. Sanjaya and D. Heksaputra, "Prediksi Rerata Harga Beras Tingkat Grosir Indonesia dengan Long Short Term Memory," *Jurnal Teknik Informatika dan Sistem Informasi*, vol. 7, no. 2, pp. 163-174, 2022.
- [15] R. Ramadania, "Peramalan Harga Beras Bulanan Di Tingkat Penggilingan Dengan Metode Weighted Moving Average," *Buletin Ilmiah Math. Stat. dan Terapannya*, vol. 7, no. 4, pp. 329-334, 2018.