

# Perbandingan Performa Algoritma SVR, LSTM, dan SARIMA dalam Peramalan Produksi Kelapa Sawit

Desvita Hendri\*, Inggih Permana, Febi Nur Salisah, M Afdal, Megawati Megawati, Eki Saputra

Fakultas Sains dan Teknologi, Sistem Informasi, Universitas Islam Negeri Sultan Syarif Kasim Riau, Pekanbaru, Indonesia

Email: <sup>1,\*</sup>12150320281@studenst-uinsuska.ac.id, <sup>2</sup>inggihpermana@uin-suska.ac.id, <sup>3</sup>febinursalisah@uin-suska.ac.id,

<sup>4</sup>m.afdal@uin-suska.ac.id, <sup>5</sup>megawati@uin-suska.ac.id, <sup>6</sup>eki.saputra@uin-suska.ac.id

Email Penulis Korespondensi: 12150320281@studenst-uinsuska.ac.id

Submitted: 03/04/2025; Accepted: 03/05/2025; Published: 01/06/2025

**Abstrak**—Produksi kelapa sawit di Indonesia mengalami fluktuasi yang signifikan akibat berbagai faktor seperti cuaca, kesuburan tanah, dan kondisi tandan buah. Perubahan ini berdampak pada stabilitas harga, pasokan, dan perencanaan industri kelapa sawit. Oleh karena itu, untuk meningkatkan pengambilan keputusan di industri ini diperlukan metode peramalan yang akurat untuk meningkatkan pengambilan keputusan yang tepat mengenai distribusi. Penelitian ini bertujuan membandingkan kinerja tiga metode peramalan berbasis machine learning, yaitu Support Vector Regression (SVR), Long Short-Term Memory (LSTM), dan Seasonal Autoregressive Integrated Moving Average (SARIMA), dalam memprediksi produksi kelapa sawit berdasarkan data historis selama 10 tahun terakhir yang diperoleh dari PTPN V Riau. Hasil evaluasi menunjukkan bahwa model SVR dengan kernel linear memberikan performa terbaik dengan nilai MSE 4.1718, RMSE 0.0020, MAE 0.0018, MAPE 0.2014% dan R<sup>2</sup> 0.9988. Model SVR memberikan hasil prediksi yang lebih unggul dibandingkan dengan LSTM dan SARIMA dalam peramalan produksi kelapa sawit. Penelitian ini diharapkan dapat memberikan kontribusi nyata dalam pengembangan sistem prediksi produksi yang lebih andal, sehingga mendukung efisiensi operasional dan stabilitas industri kelapa sawit di Indonesia.

**Kata Kunci:** Kelapa sawit; Prediksi; Support Vector Regression; Long Short-Term Memory; Seasonal Autoregressive Integrated Moving Average

**Abstract**—Oil palm production in Indonesia fluctuates significantly due to various factors such as weather, soil fertility, and fruit bunch condition. These changes have an impact on price stability, supply and planning for the palm oil industry. Therefore, to improve decision-making in this industry, an accurate forecasting method is required to improve decision-making regarding distribution. This study aims to compare the performance of three machine learning-based forecasting methods, namely Support Vector Regression (SVR), Long Short-Term Memory (LSTM), and Seasonal Autoregressive Integrated Moving Average (SARIMA), in predicting palm oil production based on historical data for the last 10 years obtained from PTPN V Riau. The evaluation results show that the SVR model with a linear kernel provides the best performance with an MSE value of 4.1718, RMSE 0.0020, MAE 0.0018, MAPE 0.2014% and R<sup>2</sup> 0.9988. The SVR model provides superior prediction results compared to LSTM and SARIMA in forecasting palm oil production. This research is expected to make a real contribution in the development of a more reliable prediction system, thus supporting operational efficiency and stability of the palm oil industry in Indonesia.

**Keywords:** Oil Palm; Prediction; Support Vector Regression; Long Short-Term Memory; Seasonal Autoregressive Integrated Moving Average

## 1. PENDAHULUAN

Kelapa sawit adalah salah satu komoditas perkebunan utama di Indonesia, yang memberikan kontribusi signifikan terhadap perekonomian nasional [1]. Komoditas ini memberikan keuntungan bagi pemerintah dan kesempatan kerja bagi masyarakat umum [2]. Tanaman ini dapat berkembang di berbagai kondisi iklim dan tanah yang terdapat di wilayah Indonesia [3]. Tahun 2018 luas perkebunan kelapa sawit di Indonesia mencapai 14.3 juta ha dan luas perkebunan kelapa sawit di provinsi Riau sampai tahun 2018 adalah 2.7 juta hektar [4]. Indonesia dan Malaysia tercatat sebagai produsen utama kelapa sawit di pasar global, di mana kedua negara ini menyumbang lebih dari 85% dari total produksi dunia [5].

Di Provinsi Riau, khususnya, kelapa sawit merupakan komoditas utama yang dibudidayakan oleh masyarakat dan badan usaha. Berdasarkan data Dinas Perkebunan Provinsi Riau, perkebunan kelapa sawit telah berkembang secara signifikan dari 1.119.798 hektar di tahun 2001 menjadi 2.372.402 hektar di tahun 2013. Perkebunan kelapa sawit milik petani mencapai 56%, dan sisanya dimiliki oleh perusahaan swasta dan BUMN [1]. Salah satu perusahaan Badan Usaha Milik Negara (BUMN) yang mengelola dan bisnis kelapa sawit di Provinsi Riau adalah PT. Perkebunan Nusantara V (PTPN V). Perusahaan PTPN V berdiri sejak tahun 1996 bergerak dibidang perkebunan kelapa sawit dan karet [6]. Perkebunan Nusantara V (PTPN-V) merupakan salah satu perusahaan perkebunan milik negara yang telah lama beroperasi dalam bidang perkebunan kelapa sawit dan karet. Saat ini, PTPN V mengelola area kebun inti kelapa sawit seluas 77. 305.20 hektar.

Produksi kelapa sawit yang dihasilkan oleh perusahaan di Riau ini sering mengalami fluktuasi produksi, baik kenaikan maupun penurunan, yang berdampak pada pendapatan dalam jangka panjang [7]. Fluktuasi ini dipengaruhi oleh faktor cuaca, curah hujan, kesuburan tanah, serta kondisi tandan buah yang belum siap panen, sehingga berpengaruh terhadap keuntungan perusahaan [7]. Fluktuasi ini dapat berdampak pada ketidakstabilan harga, pasokan, dan perencanaan industri kelapa sawit. Oleh karena itu untuk mengatasi tantangan tersebut, prediksi produksi kelapa sawit yang akurat menjadi sangat penting. Prediksi yang tepat dapat membantu para pemangku kepentingan, termasuk

petani, perusahaan perkebunan, dan pemerintah, dalam mengambil keputusan yang lebih baik terkait perencanaan produksi, pengelolaan sumber daya, dan kebijakan industri. Namun, memprediksi produksi kelapa sawit bukanlah tugas yang mudah karena kompleksitas dan non-linearitas data yang terlibat. Maka diperlukan metode peramalan yang lebih akurat dan handal guna meningkatkan pengambilan keputusan yang lebih optimal. Prediksi hasil produksi ini dilakukan setiap bulan karena evaluasi hasil produksi di PTPN V dilaksanakan secara bulanan. Sehingga prediksi ini digunakan sebagai acuan estimasi untuk pengambilan keputusan di masa depan.

Beberapa metode yang digunakan untuk melakukan peramalan produksi kelapa sawit antara lain metode Support Vector Regression (SVR), Long Short-Term Memory (LSTM), Seasonal Autoregressive Integrated Moving Average (SARIMA). Support Vector Regression merupakan metode yang didasarkan pada klasifikasi biner dalam ruang dengan properti yang berubah-ubah [8]. Jaringan LSTM adalah memastikan pembaruan informasi, status jaringan, dan keluaran informasi [9]. Metode SARIMA sifatnya kemampuan untuk mempertimbangkan perilaku musiman deret waktu stasioner atau non-stasioner [10]. Dalam konteks ini, pendekatan time series forecasting menjadi relevan karena mampu menganalisis pola data historis dan memproyeksikannya ke masa depan. Ketiga model ini menawarkan pendekatan berbeda yang berpotensi memberikan hasil prediksi yang optimal sesuai karakteristik data produksi kelapa sawit.

Penelitian tentang produksi kelapa sawit banyak dilakukan belakangan ini dapat diprediksi dengan berbagai teknik. Penelitian terbaru oleh Hastari dkk tahun 2024 menggunakan Support Vector Regression (SVR) dan Long Short-Term Memory (LSTM), menunjukkan hasil bahwa SVR lebih akurat daripada LSTM dalam memprediksi produksi minyak sawit, dengan akurasi 100% [7]. Penelitian lain yang dilakukan Ramadhan dan Fachrie (2024) menunjukkan bahwa LSTM efektif dalam memprediksi hasil panen sawit, dengan MSE loss sebesar 5.3968, membuktikan keakuratannya dalam analisis pertanian [11]. Studi lainnya oleh faris dan kawan-kawan tahun 2024 model ARIMA (1,1,3)(1,0,1)[12] dipilih sebagai model terbaik. Hasil peramalan menunjukkan adanya penurunan produksi dengan tingkat kesalahan MAPE sebesar 12,49% [12].

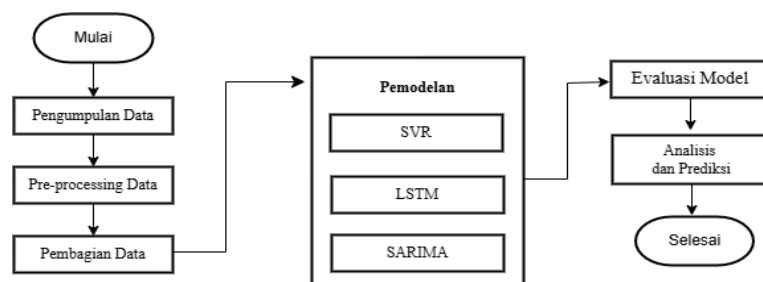
Penelitian serupa lainnya yang melakukan prediksi kelapa sawit dilakukan oleh Widiarni (2023) menunjukkan bahwa algoritma SVR dengan kernel RBF adalah yang terbaik [13]. Kajian relevan lainnya dilakukan oleh Setiawan dkk tahun 2024 membuktikan bahwa SARIMA dan LSTM dapat memprediksi produksi minyak sawit di PT X. Model SARIMA(2,1,1)(1,1,1)[12] lebih akurat dengan MAPE 1,63% dibandingkan LSTM yang mencapai 14,06% [14]. Penelitian ini melanjutkan studi sebelumnya, dari berbagai studi yang ada, terlihat bahwa masih sedikit penelitian yang secara menyeluruh membandingkan ketiga model SVR, LSTM, dan SARIMA dalam satu kerangka analisis dengan menggunakan data yang sama. Dengan demikian, penelitian ini dilakukan untuk mengisi kesenjangan tersebut sekaligus mengatasi keterbatasan yang ada dalam jumlah data dan variasi algoritma yang diterapkan pada penelitian sebelumnya.

Berdasarkan pemaparan yang telah dijelaskan dan didukung beberapa penelitian terdahulu, penelitian ini bertujuan membandingkan tiga kinerja metode peramalan dalam memprediksi produksi kelapa sawit di masa mendatang dengan tingkat kesalahan yang rendah. Selain itu, penelitian ini diharapkan dapat menjadi dasar pengembangan model yang lebih efisien di masa depan, model ini tidak hanya diharapkan memiliki kemampuan prediksi yang optimal, tetapi juga ringan dalam hal komputasi, mudah untuk diterapkan, dan mampu beradaptasi dengan perubahan data secara real-time. Dengan adanya model yang handal dan efisien, pelaku industri kelapa sawit dapat merencanakan produksi dengan lebih tepat waktu, mengelola sumber daya secara optimal, dengan harapan mampu meningkatkan efisiensi operasional serta profitabilitas industri.

## 2. METODOLOGI PENELITIAN

### 2.1 Tahapan Penelitian

Penelitian ini melakukan perbandingan beberapa algoritma pada data produksi kelapa sawit. Seluruh tahapan penelitian digambarkan melalui diagram alir yang dapat dilihat pada Gambar 1.



**Gambar 1.** Metodologi penelitian

Gambar 1 menjelaskan secara rinci semua tahapan penelitian yang dimulai dengan mengumpulkan data secara langsung dari PTPN V. Data kemudian akan dilakukan pre-processing untuk memperbaiki format data, pembagian data menggunakan teknik validasi hold-out, selanjutnya pemodelan dilakukan menggunakan tiga algoritma untuk

mendapatkan model terbaik. Kemudian dilakukan evaluasi model untuk menilai seberapa baik model yang telah dilatih. Terakhir, model terbaik dilakukan analisis perbandingan untuk mendapatkan tingkat akurasi terbaik dari model pada proses prediksi produksi kelapa sawit.

## 2.2 Pengumpulan Data

Pengumpulan data dalam penelitian ini berupa data time series produksi kelapa sawit yang diperoleh langsung dari perusahaan perkebunan PTPN V di Riau, Indonesia. Data yang digunakan mencakup periode 10 tahun terakhir, yaitu dari Januari 2014 hingga Desember 2024.

## 2.3 Pre-Processing Data

Pada penelitian ini, proses pre-processing data dilakukan untuk memastikan kesiapan data sebelum digunakan dalam pelatihan model machine learning. Tahapan ini mencakup cleaning dengan menangani data yang hilang (missing value) dan memastikan format serta tipe data konsisten agar dapat diproses dengan baik. Selain itu, dilakukan normalisasi menggunakan teknik Min-Max Normalization, yaitu mengubah nilai data ke dalam rentang 0 hingga 1. Normalisasi ini bertujuan untuk menyederhanakan struktur data, mengurangi redundansi, serta memastikan konsistensi dan integritas data, sehingga model dapat bekerja lebih optimal dalam proses analisis dan prediksi.

## 2.4 Pembagian Data

Tahap selanjutnya melakukan pembagian data untuk mengevaluasi performa model secara objektif dan memastikan model dapat bekerja dengan baik. Dataset dibagi menjadi dua bagian terpisah, yaitu data pelatihan (training set) dan data pengujian (test set), dengan proporsi 80:20 menggunakan teknik *Hold-Out*. Data pelatihan digunakan untuk melatih model, sementara data pengujian digunakan untuk mengukur performa model pada data yang belum pernah dilihat sebelumnya.

## 2.5 Pemodelan

### 2.5.1 Support Vector Regression

Support Vector Regression (SVR) adalah MLA populer yang dikembangkan berdasarkan konsep Vapnik (1995) [15]. Algoritma Support Vector Regression (SVR) menggunakan pembelajaran mesin untuk memprediksi nilai keluaran kontinu berdasarkan masukan, dengan meminimalkan kesalahan prediksi, mengelola pencilan, serta menghindari overfitting dan underfitting [16]. Tujuan dari SVR adalah untuk menemukan fungsi  $f(x)$  sebagai hyperplane yang merupakan sebuah fungsi.

$$F(X) = W^T \phi(x) + b \tag{1}$$

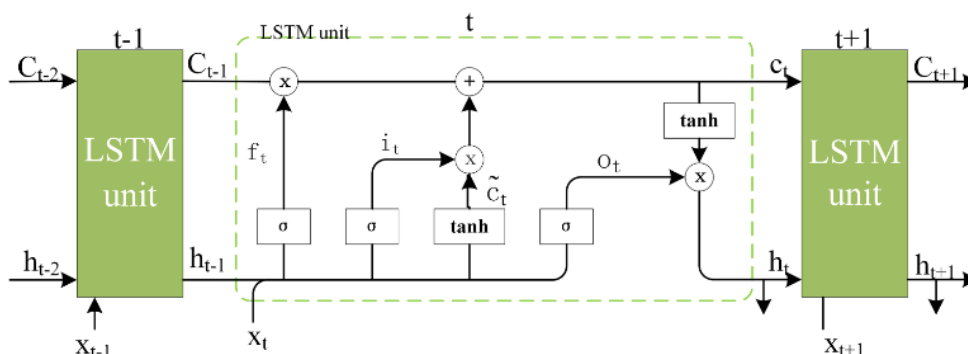
Banyak teknik data mining yang dikembangkan mengasumsikan kelinearitasan, sehingga algoritma dihasilkan terbatas untuk linier saja [17], akan tetapi SVR dapat bekerja pada dataset non-linier dengan pendekatan kernel pada data awal. Untuk SVR, fungsi kernel mengubah data menjadi ruang fitur berdimensi lebih tinggi agar memungkinkan dilakukannya pemisahan linier, SVR dianggap sebagai teknik non parametrik karena mengandalkan fungsi kernel [18].

$$G = \{(x_i, y_i): x_i, y_i, \in F\}_{i=1}^n \tag{2}$$

Di mana  $x_i$  merupakan variabel input dan  $y_i$  merupakan variabel output.

### 2.5.2 Long Short-Term Memory

Algoritma Long Short-Term Memory (LSTM) merupakan jenis Recurrent Neural Network (RNN) yang dirancang untuk mengatasi masalah error backpropagation dan meningkatkan kinerja sistem [19]. Model ini pertama kali diusulkan oleh Hochreiter dan Schmidhuber pada tahun 1997 dan telah disempurnakan selama bertahun-tahun [20].



Gambar 2. Struktur dasar model LSTM [20].

Gambar 2 menunjukkan diagram struktur sel yang terperinci dari jaringan LSTM. Langkah pertama adalah menentukan informasi mana yang perlu dibuang atau dipertahankan dari keadaan sel melalui gerbang lupa, yang diimplementasikan oleh fungsi aktivasi sigmoid. Rumusnya adalah sebagai berikut:

$$f_i = \sigma (W_f \bullet [h_{t-1}, x_t] + b_f) \quad (3)$$

di mana  $f_i$  adalah vektor hasil satuan,  $\sigma (\bullet)$  mewakili fungsi sigmoid,  $W_f$  adalah matriks bobot dari gerbang lupa, dan  $b_f$  adalah bias dari gerbang lupa. Langkah kedua adalah memutuskan informasi apa yang akan diperbarui melalui gerbang masukan. Rumusnya adalah sebagai berikut :

$$i_t = \sigma (W_i \bullet [h_{t-1}, x_t] + b_i) \quad (4)$$

$$\tilde{C}_t = \tanh (W_c \bullet [h_{t-1}, x_t] + b_c) \quad (5)$$

dimana nilai  $\tilde{C}_t$  berada dalam rentang (-1,1),  $\tanh (\bullet)$  adalah fungsi tangen hiperbolik,  $W_i$  dan  $W_c$  adalah matriks bobot dari gerbang input, dan  $b_i$  dan  $b_c$  adalah bias dari gerbang input. Langkah ketiga adalah memperbarui status sel. Rumusnya adalah sebagai berikut:

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (6)$$

Langkah keempat melibatkan pemilihan dan keluaran informasi melalui gerbang keluaran. Rumusnya adalah sebagai berikut:

$$o_t = \sigma (W_o \bullet [h_{t-1}, x_t] + b_o) \quad (7)$$

$$h_t = o_t * \tanh (C_t) \quad (8)$$

di mana  $o_t$  adalah nilai keluaran,  $W_o$  adalah matriks bobot gerbang keluaran, dan  $b_o$  adalah bias gerbang keluaran.

### 2.5.3 Seasonal Autoregressive Integrated Moving Average

SARIMA adalah model peramalan deret waktu standar yang digunakan untuk memprediksi nilai masa depan berdasarkan data historis [21]. Model Ini telah diperkenalkan oleh Box dan Jenkins pada tahun 1976 [22]. Model ini merupakan pengembangan dari model ARIMA, yang secara luas digunakan untuk analisis deret waktu.

Bentuk umum dari SARIMA dilambangkan sebagai SARIMA (p,d,q) (P,D,Q)[s], di mana: p mewakili orde bagian autoregresif, d mewakili orde differencing untuk membuat deret menjadi stasioner, q mewakili orde bagian moving average, P, D, dan Q masing-masing mewakili komponen musiman dari autoregresif, differencing, dan moving average, s adalah panjang siklus musiman [23].

### 2.6 Metriks Evaluasi

MAE, MAPE, RMSE, MSE dan  $R^2$  adalah matriks yang digunakan untuk menilai keakuratan model peramalan. MAE mengukur rata-rata perbedaan absolut antara nilai prediksi dan nilai aktual, sedangkan MAPE menyatakan kesalahan sebagai persentase dari nilai aktual. RMSE dan MSE menghitung rata-rata kesalahan kuadrat, dengan RMSE menekankan kesalahan yang lebih besar karena akar kuadratnya transformasi. Matriks ini membantu mengevaluasi model kinerja model dengan menganalisis bias dan akurasi prediksi [24]. R-Squared mengukur sejauh mana variabilitas dalam variabel target dapat dijelaskan oleh model [25]. Rumus untuk MSE, RMSE, MAE, MAPE dan  $R^2$  disajikan dalam persamaan 9-13.

$$MSE = \frac{1}{n} \sum_i^n (y_i - \hat{y}_i)^2 \quad (9)$$

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{n} \sum_i^n (y_i - \hat{y}_i)^2} \quad (10)$$

$$MAE = \frac{1}{n} \sum_i^n |y_i - \hat{y}_i| \quad (11)$$

$$MAPE = \frac{100\%}{n} \sum_i^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (12)$$

$$R^2 = 1 - \frac{SS\ Error}{SS\ Total} = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2} \quad (13)$$

Mean Squared Error (MSE) dihitung dengan menjumlahkan kuadrat selisih antara nilai aktual ( $y_i$ ) dan nilai prediksi ( $\hat{y}_i$ ) untuk setiap data ke-  $i$ , kemudian dibagi dengan jumlah total data ( $n$ ). Root Mean Squared Error (RMSE) adalah akar kuadrat dari MSE, yang mengembalikan hasil kesalahan ke satuan asli data. Mean Absolute Error (MAE) menghitung rata-rata dari nilai absolut selisih antara ( $y_i$ ) dan ( $\hat{y}_i$ ), tanpa memperhitungkan arah kesalahan. Mean Absolute Percentage Error (MAPE) menyatakan rata-rata kesalahan prediksi dalam bentuk persentase terhadap nilai aktual, dengan cara membagi selisih absolut antara ( $y_i$ ) dan ( $\hat{y}_i$ ) terhadap nilai ( $y_i$ ), lalu dirata-rata dan dikalikan 100%, Semakin rendah nilai MAPE, semakin baik kinerja prediksi dari model tersebut. Sementara itu,

koefisien determinasi ( $R^2$ ) mengukur sejauh mana model menjelaskan variasi dari data aktual, dengan membandingkan jumlah kuadrat galat prediksi  $(y_i - \hat{y}_i)^2$  terhadap jumlah kuadrat total  $(y_i - \bar{y})^2$  di mana  $(\bar{y})$  merupakan rata-rata dari seluruh nilai aktual.

### 3. HASIL DAN PEMBAHASAN

#### 3.1 Pengumpulan Data dan Pra Proses Data

Penelitian ini menggunakan sebanyak 132 data produksi kelapa sawit, Adapun data yang digunakan dalam penelitian ini dapat dilihat pada Tabel 1. Selanjutnya, langkah pengolahan data melibatkan penerapan Normalisasi Min-Maks. Data produksi kelapa sawit memiliki rentang nilai yang cukup besar antar tahun, yang dapat memengaruhi kinerja model prediksi yang sensitif terhadap skala input. Untuk mengatasi hal ini, dilakukan normalisasi dengan skala 0 hingga 1 agar pelatihan model menjadi lebih stabil dan akurat. Untuk data normalisasi dari rentang 0 hingga 1 dilihat pada Tabel 2.

**Tabel 1.** Data Awal

Periode	Hasil Produksi
Jan-14	71.261.074
Feb-14	61.795.830
Mar-14	69.893.552
...	...
Dec-24	143.705.130

**Tabel 2.** Normalisasi data

Periode	Hasil Produksi
Jan-14	0.127453
Feb-14	0.047630
Mar-14	0.115921
...	...
Dec-24	0.738394

#### 3.2 Pemodelan

Penerapan Algoritma *Support Vector Regression*, *Long Short-Term Memory*, dan *Seasonal Autoregressive Integrated Moving Average* menggunakan Python. Pada pengujian dengan algoritma Support Vector Regression (SVR), parameter yang akan dicari adalah kombinasi nilai C, epsilon ( $\epsilon$ ), dan gamma ( $\gamma$ ) yang menghasilkan nilai error paling kecil. Kernel yang digunakan mencakup RBF, linear, sigmoid, dan polynomial. Untuk menentukan nilai optimal pada algoritma SVR dilakukan optimasi menggunakan metode Grid Search. Hasil percobaan uji parameter yang dihasilkan pada produksi kelapa sawit terdapat pada Tabel 3.

**Tabel 3.** Hasil Pemodelan Algoritma SVR

Kernel	Experiment			SVR				
	C	Epsilon	Gamma	RMSE	MSE	MAPE	MAE	R <sup>2</sup>
RBF	10	0,001	0,01	0.0031	9.6963	0.3089	0.0028	0.9972
	10	0,01	0,1	0.0265	0.0007	2.7622	0.0252	0.7960
	10	0,001	0,1	0.0119	0.0001	1.1899	0.0109	0.9586
	10	0,01	0,5	0.0541	0.0029	5.4957	0.0505	0.1489
Linear	<b>2</b>	<b>0,001</b>	<b>0,01</b>	<b>0.0020</b>	<b>4.1718</b>	<b>0.2014</b>	<b>0.0018</b>	<b>0.9988</b>
	2	0,01	0,1	0.0141	0.0001	1.3712	0.0126	0.9422
	0,25	0,001	0,01	0.0022	4.8339	0.2163	0.0020	0.9986
	2	0,01	0,1	0.0141	0.0001	1.3712	0.0126	0.9422
Sigmoid	10	0,001	0,01	0.0042	1.7555	0.4134	0.0038	0.9949
	10	0,01	0,01	0.0328	0.0010	3.2386	0.0297	0.6875
	10	0,001	0,5	0.0468	0.0021	3.7470	0.0335	0.3640
	10	0,01	0,5	0.0177	0.0003	1.8586	0.0170	0.9090
Polynomial	0,25	0,01	0,5	0.1876	0.0352	1.8097	0.2502	0.2421
	0,25	0,01	0,75	0.1908	0.0387	1.8735	0.2305	0.2621
	0,1	0,001	0,5	0.1790	0.0323	1.0118	0.0504	0.8404
	0,25	0,1	0,5	0.1898	0.0360	1.9240	0.0133	0.4738

Tabel 3 memperlihatkan performa model prediksi algoritma SVR menggunakan grid search, Tabel tersebut menunjukkan bahwa epsilon ( $\epsilon$ ) = 0.001, C = 2, dan gamma ( $\gamma$ ) = 0.01 dengan kernel linear sebagai parameter terbaik. Model dengan parameter ini menunjukkan performa yang optimal berdasarkan nilai metrik evaluasi, yaitu RMSE 0.0020, MSE 4.1718, MAE 0.0018, MAPE 0.2014%, dan R<sup>2</sup> 0.9988.

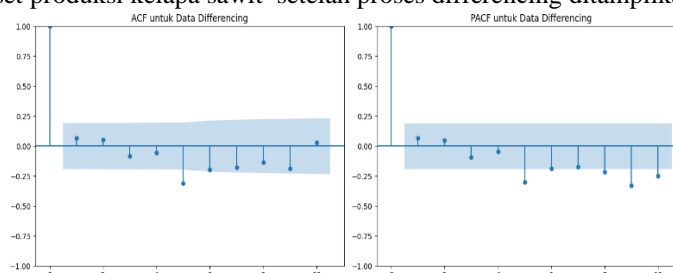
Selanjutnya pengujian dengan algoritma LSTM digunakan untuk memodelkan prediksi kelapa sawit melalui beberapa percobaan dengan tujuan menghasilkan model yang optimal. Eksperimen yang dilakukan mencakup optimasi parameter (Adam, AdamW, Adamax, RMSprop dan Nadam), batch size (8, 16, dan 32), dengan learning rate 0,001 dan 0,0001. Model dilatih selama lebih dari 100 epoch menggunakan callback (checkpoint and early stopping). Hasil percobaan uji parameter yang dihasilkan terdapat pada Tabel 4.

**Tabel 4.** Hasil Pemodelan Algoritma LSTM

Experiment			LSTM				
Optimizer	Batch Size	Learning Rate	RMSE	MSE	MAPE	MAE	R <sup>2</sup>
Adam	8	0,001	17028	28995	10.563%	13248	0.4804
Adam	8	0,0001	16170	26147	9.8949%	12437	0.5314
Adam	16	0,001	15691	24623	9.7308%	11998	0.5587
Adam	16	0,0001	16229	26339	10.256 %	13681	0.5280
Adam	32	0,001	26145	68356	17.902 %	12107	0.2250
Adam	32	0,0001	26416	69783	18.003 %	19306	0.2505
AdamW	8	0,001	18553	34424	12.134 %	12107	0.3831
AdamW	8	0,0001	26546	70469	18.130 %	19842	0.2628
<b>AdamW</b>	<b>16</b>	<b>0,001</b>	<b>15518</b>	<b>24082</b>	<b>9.9343 %</b>	<b>14356</b>	<b>0.5684</b>
AdamW	16	0,0001	26917	72457	18.287 %	14257	0.2984
AdamW	32	0,001	26632	70929	18.141 %	17654	0.2711
AdamW	32	0,0001	26956	72664	18.168 %	11489	0.3021
Adamax	8	0,001	26084	68038	17.423 %	16038	0.2193
Adamax	8	0,0001	26242	68868	17.778 %	19571	0.2341
Adamax	16	0,001	18904	35737	12.314 %	13906	0.3596
Adamax	16	0,0001	26539	70432	18.074 %	12268	0.2622
Adamax	32	0,001	25635	65719	17.634 %	18702	0.1777
Adamax	32	0,0001	26192	68603	17.913 %	26941	0.2294
RMSprop	8	0,001	24302	59060	16.151 %	23450	0.0584
RMSprop	8	0,0001	25190	63457	17.060 %	27583	0.1372
RMSprop	16	0,0001	25635	65719	17.507 %	29106	0.1777
RMSprop	32	0,0001	26083	68034	17.825 %	25849	0.2192
Nadam	8	0,001	18126	32856	12.051 %	21097	0.4112
Nadam	16	0,0001	27241	74211	18.314 %	23102	0.3299
Nadam	32	0,001	26688	71228	17.522 %	14398	0.2764
Nadam	32	0,0001	26278	69055	17.986 %	26730	0.2375

Tabel 4 menunjukkan performa model parameter menggunakan algoritma LSTM, di mana optimizer AdamW dengan batch size 16 teridentifikasi sebagai kombinasi parameter terbaik. Model dengan parameter ini menunjukkan performa yang optimal berdasarkan nilai metrik evaluasi, yaitu RMSE 15518, MSE 24082, MAE 14356, MAPE 9.9343 %, dan R<sup>2</sup> 0.5684.

Selanjutnya dilakukan pengujian pada algoritma SARIMA, identifikasi dimulai dari *pre-processing* data dengan melakukan analisis apakah data sudah stasioner apa belum. Uji Augmented Dickey-Fuller (ADF) pada dataset produksi kelapa sawit menunjukkan p-value 0.754, mengindikasikan data belum stasioner. Setelah differencing pertama, p-value menjadi 0.002, menandakan data stasioner. Setelah dataset teridentifikasi stasioner terhadap mean dan varians, langkah berikutnya adalah menentukan model terbaik dengan memilih parameter (p, d, q)(P, D, Q) berdasarkan analisis grafik Autocorrelation Function (ACF) dan Partial Autocorrelation Function (PACF). Grafik ACF dan PACF untuk dataset produksi kelapa sawit setelah proses differencing ditampilkan pada Gambar 3.



**Gambar 3.** Plot ACF dan PACF

Gambar 3 menunjukkan grafik ACF dan PACF setelah dilakukan differencing pada data produksi kelapa sawit. Grafik ACF menampilkan autokorelasi yang menurun secara signifikan, menandakan bahwa data telah lebih stasioner. Sementara itu, grafik PACF menunjukkan cutoff setelah beberapa lag pertama, yang dapat menjadi indikasi nilai p untuk model AR. Pola ini membantu dalam menentukan parameter (p, d, q) (P, D, Q, s) dalam model SARIMA, yang selanjutnya dapat dipilih berdasarkan nilai AIC untuk mendapatkan model terbaik. Hasil percobaan uji parameter yang dihasilkan pada produksi kelapa sawit terdapat pada Tabel 5.

**Tabel 5.** Hasil Pemodelan Algoritma SARIMA

Experiment	SARIMA	
Parameter	AIC	Time
ARIMA(0,1,0)(0,1,0)[12]	3269.583	0.09
ARIMA(1,1,0)(1,1,0)[12]	3268.292	0.23
ARIMA(0,1,1)(0,1,1)[12]	3267.381	0.33
ARIMA(0,1,1)(0,1,0)[12]	3267.383	0.19
ARIMA(0,1,1)(1,1,1)[12]	3259.663	1.16
ARIMA(0,1,1)(1,1,0)[12]	3268.227	0.51
ARIMA(0,1,1)(2,1,1)[12]	3258.081	1.18
ARIMA(0,1,1)(2,1,0)[12]	3265.484	0.34
ARIMA(0,1,1)(2,1,2)[12]	3258.136	1.48
ARIMA(0,1,1)(1,1,2)[12]	3259.177	0.78
ARIMA(0,1,0)(2,1,1)[12]	3259.419	0.53
ARIMA(1,1,1)(2,1,1)[12]	3257.774	2.13
ARIMA(1,1,1)(1,1,1)[12]	3261.630	0.44
ARIMA(1,1,1)(2,1,0)[12]	3267.421	0.62
ARIMA(1,1,1)(2,1,2)[12]	3257.803	2.05
ARIMA(1,1,1)(1,1,0)[12]	3270.224	0.24
ARIMA(1,1,1)(1,1,2)[12]	3258.956	2.88
ARIMA(1,1,0)(2,1,1)[12]	3257.571	1.64
ARIMA(1,1,0)(1,1,1)[12]	3259.228	0.37
ARIMA(1,1,0)(2,1,0)[12]	3265.500	0.32
<b>ARIMA(1,1,0)(2,1,2)[12]</b>	<b>3257.516</b>	<b>1.35</b>
ARIMA(1,1,0)(1,1,2)[12]	3258.710	0.63
ARIMA(0,1,0)(2,1,2)[12]	3258.773	1.30
ARIMA(2,1,0)(2,1,2)[12]	3259.113	1.96
ARIMA(2,1,1)(2,1,2)[12]	3261.245	2.68
ARIMA(1,1,0)(2,1,2)[12] intercept	3259.433	1.42

Tabel 5 menunjukkan bahwa model ARIMA(1,1,0)(2,1,2)[12] memiliki AIC terendah (3257.516), menandakan kecocokan terbaik dengan data. Dengan waktu eksekusi 1.35 detik, model ini dipilih sebagai yang paling optimal untuk peramalan produksi kelapa sawit. Model dengan parameter tersebut menunjukkan performa yang optimal berdasarkan nilai metrik evaluasi, yaitu RMSE 13901, MSE 19323, MAE 27288, MAPE 22352%, dan R<sup>2</sup> 09521.

### 3.3 Evaluasi Model

Setelah model dilatih proses selanjutnya adalah evaluasi. Evaluasi merupakan tahap yang bertujuan untuk mengukur performa serta kinerja yang dihasilkan dari model. Model dievaluasi menggunakan metrik kesalahan regresi, yaitu, MSE, RMSE, MAE, MAPE dan R<sup>2</sup>. Tabel 6 menyajikan hasil evaluasi performa model berdasarkan parameter yang telah diuji terhadap data produksi kelapa sawit.

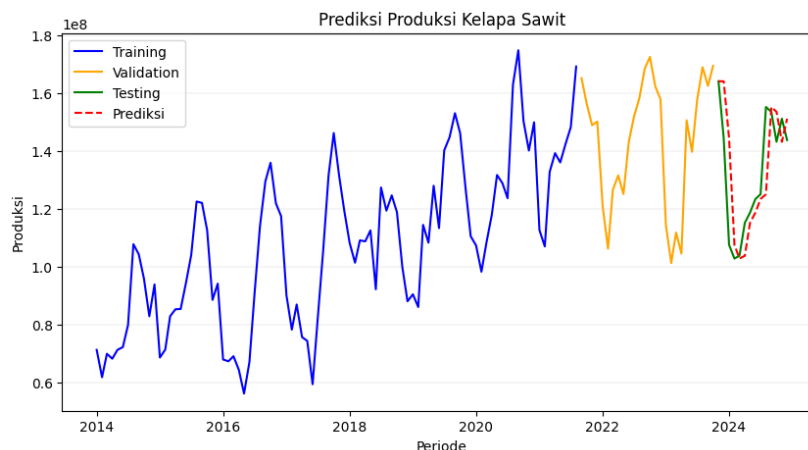
**Tabel 6.** Evaluasi Hasil Prediksi Model

Model	MSE	RMSE	MAE	MAPE	R <sup>2</sup>
SVR	41718	00020	00018	02014	09988
LSTM	24082	15518	14356	99343	05684
SARIMA	19323	13901	27288	22352	09521

### 3.4 Hasil Prediksi Terbaik

Model SVR dengan kernel linear, epsilon ( $\epsilon$ ) = 0.001, C = 2, dan gamma ( $\gamma$ ) = 0.01 adalah model yang paling efektif dari semua pengujian yang dilakukan, sesuai dengan analisis dan pengujian yang dilakukan. Berdasarkan hasil evaluasi model prediksi pada Tabel 6, algoritma Support Vector Regression (SVR) menunjukkan performa terbaik dengan nilai MSE sebesar 4.1718, RMSE sebesar 0.0020, MAE sebesar 0.0018, MAPE sebesar 0.2014% dan R<sup>2</sup>

sebesar 0.9988. Nilai error yang lebih kecil dibandingkan model LSTM dan SARIMA. Hasil visualisasi prediksi pada algoritma SVR dapat dilihat pada Gambar 4.



**Gambar 4.** Hasil Visualisasi SVR

Gambar 4 menunjukkan grafik prediksi memperlihatkan pola fluktuatif dalam produksi kelapa sawit, dengan trend yang cenderung meningkat meskipun mengalami variasi musiman. Hal ini menunjukkan bahwa model yang digunakan mampu menangkap pola produksi dengan baik dan memberikan gambaran yang cukup akurat untuk proyeksi masa depan.

## 4. KESIMPULAN

Penelitian ini membahas penerapan algoritma Support Vector Regression (SVR), Long Short-Term Memory (LSTM), dan Seasonal Autoregressive Integrated Moving Average (SARIMA) dalam memprediksi produksi kelapa sawit berdasarkan data historis selama 10 tahun terakhir. Dari analisis yang dilakukan, ditemukan bahwa setiap algoritma memiliki kelebihan dan keterbatasan masing-masing dalam menangani pola data yang kompleks. Hasil evaluasi menunjukkan bahwa SVR memiliki performa terbaik dari ketiga algoritma, dengan nilai MSE 4.1718, RMSE 0.0020, MAE 0.0018, MAPE 0.2014% dan R2 0.9988 yang lebih rendah dibandingkan dengan LSTM dan SARIMA. Penelitian ini menunjukkan bahwa karakteristik data kelapa sawit yang stabil dan memiliki pola nonlinier yang ringan lebih cocok untuk ditangani dengan pendekatan SVR. Hasil penelitian ini menegaskan bahwa pemilihan algoritma yang tepat sangat berpengaruh terhadap kinerja metode peramalan, sehingga SVR direkomendasikan sebagai metode yang lebih efektif dalam memproyeksikan produksi kelapa sawit di masa mendatang.

## REFERENCES

- [1] A. Syahza and B. Asmit, "Development of palm oil sector and future challenge in Riau Province, Indonesia," *J. Sci. Technol. Policy Manag.*, vol. 11, no. 2, pp. 149–170, 2020, doi: 10.1108/JSTPM-07-2018-0073.
- [2] F. Husaini, I. Permana, M. Afdal, and F. N. Salisah, "Penerapan Algoritma Long Short-Term Memory untuk Prediksi Produksi Kelapa Sawit," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 4, no. 2, pp. 366–374, 2024, doi: 10.57152/malcom.v4i2.1187.
- [3] M. Syarovy, A. P. Nugroho, and L. Sutiarmo, "Pemanfaatan Model Neural Network Dalam Generasi Baru Pertanian Presisi Di Perkebunan Kelapa Sawit," *War. Pus. Penelit. Kelapa Sawit*, vol. 28, no. 1, pp. 39–54, 2023, doi: 10.22302/iopri.war.warta.v28i1.97.
- [4] R. Novia Yanti and I. L. Hutasuhut, "Potensi Limbah Padat Perkebunan Kelapa Sawit Di Provinsi Riau," *Wahana For. J. Kehutan.*, vol. 15, no. 2, pp. 1–11, 2020, doi: 10.31849/forestra.v15i2.4696.
- [5] E. Safitri, S. Basriati, and K. Sari, "Pengaruh Luas Lahan Dan Produktivitas Terhadap Hasil Produksi Kelapa Sawit Menggunakan Metode Cobb Douglass (Studi Kasus: Ptpn V, Pekanbaru)," *MAP (Mathematics Appl. J.)*, vol. 2, pp. 56–65, Dec. 2020, doi: 10.15548/map.v2i2.2266.
- [6] S. Sarah, Mustakim, R. Novita, and N. E. Rozanda, "Implementation of Fuzzy C-Means and Self-Organizing Map for Data Clustering of Palm Oil," in *2023 International Seminar on Intelligent Technology and Its Applications (ISITIA)*, 2023, pp. 444–449. doi: 10.1109/ISITIA59021.2023.10221173.
- [7] D. Hastari, M. Mustakim, R. Novita, and M. Afdal, "Palm Oil Production Prediction Using Support Vector Regression Algorithm and Long Short-Term Memory," in *2024 International Conference on Circuit, Systems and Communication (ICCS)*, 2024, pp. 1–6. doi: 10.1109/ICCS62074.2024.10616414.
- [8] B. Mohammadi and Z. Aghashariatmadari, "Estimation of solar radiation using neighboring stations through hybrid support vector regression boosted by Krill Herd algorithm," *Arab. J. Geosci.*, vol. 13, no. 10, pp. 1–16, 2020, doi: 10.1007/s12517-020-05355-1.
- [9] S. Wen, B. Wei, J. You, Y. He, J. Xin, and M. A. Varfolomeev, "Forecasting oil production in unconventional reservoirs using long short term memory network coupled support vector regression method: A case study," *Petroleum*, vol. 9, no. 4,



- pp. 647–657, 2023, doi: 10.1016/j.petlm.2023.05.004.
- [10] T. Falatouri, F. Darbanian, P. Brandtner, and C. Udokwu, “Predictive Analytics for Demand Forecasting - A Comparison of SARIMA and LSTM in Retail SCM,” *Procedia Comput. Sci.*, vol. 200, no. 2019, pp. 993–1003, 2022, doi: 10.1016/j.procs.2022.01.298.
- [11] M. Ramadhan, dan fachrie, “Implementasi Algoritma Long Short-Term Memory Pada Sistem Prediksi Hasil Panen Sawit,” *J. Inform. Teknol. dan Sains*, vol. 6, pp. 937–944, 2024, doi: <https://doi.org/10.51401/jinteks.v6i4.4876>.
- [12] R. M. Faris, K. Kurniaji, D. Budiman, and M. W. Kusuma, “Analisis Peramalan Produksi Tanaman Kelapa Sawit Menggunakan Metode Arima pada PTPN Kebun Sukamaju,” *J. Bisnis dan Manaj. West Sci.*, vol. 3, no. 03, pp. 275–290, 2024, doi: <https://doi.org/10.58812/jbmws.v3i03.1537>.
- [13] A. Widiarni, “Penerapan Algoritma Support Vector Regression dalam Memprediksi Produksi dan Produktivitas Kelapa Sawit,” *J. Media Inform. Budidarma*, vol. 7, no. 2, pp. 864–872, 2023, doi: 10.30865/mib.v7i2.6089.
- [14] N. H. Setiawan and Z. Zulkarnain, “Forecasting Palm Oil Production Using Long Short-Term Memory (LSTM) With Time Series Cross Validation (TSCV),” *Int. J. Soc. Serv. Res.*, vol. 4, no. 05, pp. 1237–1251, 2024, doi: 10.46799/ijssr.v4i05.780.
- [15] M. Panahi, N. Sadhasivam, H. R. Pourghasemi, F. Rezaie, and S. Lee, “Spatial prediction of groundwater potential mapping based on convolutional neural network (CNN) and support vector regression (SVR),” *J. Hydrol.*, vol. 588, no. April, p. 125033, 2020, doi: 10.1016/j.jhydrol.2020.125033.
- [16] S. R. Thumu and G. Nellore, “Improving Cryptocurrency Price Prediction Accuracy with MultiKernel Support Vector Regression Approach,” *Int. Res. J. Multidiscip. Technovation*, vol. 6, no. 4, pp. 20–31, 2024, doi: 10.54392/irjmt2443.
- [17] R. A. Fahrezi, M. Y. Wijaya, and N. Fitriyati, “Prediksi Harga Penutupan Saham Bank Central Asia: Implementasi Algoritma Long Short-Term Memory Dan Perbandingannya Dengan Support Vector Machine,” *J. Lebesgue J. Ilm. Pendidik. Mat. Mat. dan Stat.*, vol. 5, no. 1, pp. 452–464, 2024, doi: 10.46306/lb.v5i1.582.
- [18] C. Shoko and C. Sigauke, “Short-term forecasting of COVID-19 using support vector regression: An application using Zimbabwean data,” *Am. J. Infect. Control*, vol. 51, no. 10, pp. 1095–1107, 2023, doi: 10.1016/j.ajic.2023.03.010.
- [19] M. Bansal, A. Goyal, and A. Choudhary, “A comparative analysis of K-Nearest Neighbor, Genetic, Support Vector Machine, Decision Tree, and Long Short Term Memory algorithms in machine learning,” *Decis. Anal. J.*, vol. 3, no. May, p. 100071, 2022, doi: 10.1016/j.dajour.2022.100071.
- [20] H. Xue, C. Guo, G. Dong, C. Zhang, Y. Lian, and Q. Yuan, “Prediction of runoff in the upper reaches of the Hei River based on the LSTM model guided by physical mechanisms,” *J. Hydrol. Reg. Stud.*, vol. 58, no. January, p. 102218, 2025, doi: 10.1016/j.ejrh.2025.102218.
- [21] S. Kumari and P. Muthulakshmi, “SARIMA Model: An Efficient Machine Learning Technique for Weather Forecasting,” *Procedia Comput. Sci.*, vol. 235, no. 2023, pp. 656–670, 2024, doi: 10.1016/j.procs.2024.04.064.
- [22] A. Kumar Dubey, A. Kumar, V. García-Díaz, A. Kumar Sharma, and K. Kanhaiya, “Study and analysis of SARIMA and LSTM in forecasting time series data,” *Sustain. Energy Technol. Assessments*, vol. 47, no. August 2020, p. 101474, 2021, doi: 10.1016/j.seta.2021.101474.
- [23] S. Zhou, “Using SARIMA Method and Random Forest to Predict the Covid-19 Infection Cases,” *Adv. Econ. Manag. Polit. Sci.*, vol. 0, pp. 1–13, 2024, doi: 10.54254/2754-1169/140/2024.GA18764.
- [24] D. Chicco, M. J. Warrens, and G. Jurman, “The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation,” *PeerJ Comput. Sci.*, vol. 7, pp. 1–24, 2021, doi: 10.7717/PEERJ-CS.623.
- [25] S. Jin, “A Comparative Analysis of Traditional and Machine Learning Methods in Forecasting the Stock Markets of China and the US,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 15, no. 4, pp. 1–8, 2024, doi: 10.14569/IJACSA.2024.0150401.