



# Perbandingan Metode *ARIMA* dan *Prophet* dalam Prediksi Harga Cabai Rawit di Provinsi Jawa Timur

Haidar Fadhila Fiqa<sup>1</sup>, Atika Ratna Dewi<sup>2</sup>, Ridwan Pandiya<sup>3</sup>

<sup>1,2</sup>Sains Data, Institut Teknologi Telkom Purwokerto

<sup>1</sup>[20110022@ittelkom-pwt.ac.id](mailto:20110022@ittelkom-pwt.ac.id)

<sup>2</sup>[atika@ittelkom-pwt.ac.id](mailto:atika@ittelkom-pwt.ac.id)

<sup>3</sup>Teknik Informatika, Institut Teknologi Telkom Purwokerto

Corresponding author email: [ridwanpandiya@ittelkom-pwt.ac.id](mailto:ridwanpandiya@ittelkom-pwt.ac.id)

**Abstract:** *The price of cayenne pepper is a seasonal horticultural commodity, so its availability is not consistent throughout the year, causing its price to fluctuate. Fluctuations, instability, and uncertainty in the price of cayenne pepper in the future cause losses for society, especially people who are classified as poor. Therefore, it is necessary to carry out predictive modeling using appropriate methods based on past data. Autoregressive Integrated Moving Average (ARIMA) and Prophet are time series-based prediction methods. The purpose of this study is to compare ARIMA and Prophet to get the optimal model in predicting the price of cayenne pepper in East Java Province for the next three months. ARIMA is effective in modeling trends and seasonal patterns in stable historical data and is good at short-term prediction. Meanwhile, Prophet can predict well for data with or without seasonal effects and is resistant to outliers, missing values, and is good at dealing with trend changes automatically. The results show that ARIMA (7, 1, 7) is better than Prophet with an RMSE value of 18723.92 and MAPE of 19.48%. The prediction results show a periodic decrease in the price of cayenne pepper from May to July 2024. ARIMA results that are better than Prophet can be caused by the data not having outlier values, so the advantages of Prophet in handling outlier data are not relevant. In addition, ARIMA allows for more specific model customization through the selection of parameters (p, d, q) that are optimal for the observed data.*

**Keywords:** *cayenne pepper, East Java, prediction, ARIMA, prophet*

**Abstrak:** Harga cabai rawit merupakan komoditas hortikultura yang bersifat musiman, sehingga ketersediaannya tidak konsisten sepanjang tahun yang menyebabkan harganya mengalami fluktuasi. Adanya fluktuasi, ketidakstabilan, serta tidak ada kepastian harga cabai rawit di masa mendatang menyebabkan kerugian bagi masyarakat, terutama masyarakat yang tergolong miskin. Oleh karena itu, perlu dilakukan pemodelan prediksi menggunakan metode yang tepat berdasarkan data masa lalu. Autoregressive Integrated Moving Average (ARIMA) dan Prophet merupakan salah satu metode prediksi berbasis time series. Tujuan penelitian ini adalah membandingkan ARIMA dan Prophet untuk mendapatkan model yang optimal dalam memprediksi harga cabai rawit di Provinsi Jawa Timur selama tiga bulan mendatang. ARIMA efektif dalam memodelkan tren dan pola musiman dalam data historis yang stabil serta unggul dalam prediksi jangka pendek. Sementara itu, Prophet dapat melakukan prediksi dengan baik untuk data dengan efek musiman maupun tidak serta tahan terhadap outlier, missing value, dan unggul dalam mengatasi perubahan tren secara otomatis. Berdasarkan penelitian, diperoleh hasil bahwa ARIMA (7, 1, 7) lebih unggul daripada Prophet dengan nilai RMSE sebesar 18723.92 dan MAPE sebesar 19.48%. Hasil prediksi menunjukkan adanya penurunan harga cabai rawit secara berkala dimulai dari Mei hingga Juli 2024. Hasil ARIMA yang lebih unggul daripada Prophet dapat disebabkan karena data tidak memiliki nilai outlier, sehingga keunggulan Prophet dalam menangani data outlier tidak relevan. Selain itu, ARIMA memungkinkan penyesuaian model yang lebih spesifik melalui pemilihan parameter (p, d, q) yang optimal untuk data pengamatan.

**Kata kunci:** *cabai rawit, Jawa Timur, prediksi, ARIMA, prophet*

## I. PENDAHULUAN

Salah satu komoditas sayur-sayuran yang paling banyak dibutuhkan di pasaran dalam negeri maupun luar negeri adalah cabai [1]. Badan Pusat Statistik (BPS) mencatat jumlah produksi cabai, baik cabai besar dan cabai rawit di Indonesia tahun 2021 sebesar 2,74 juta ton, dan terjadi peningkatan pada tahun 2022 dengan jumlah produksi sebanyak 3 juta ton. Cabai rawit merupakan salah satu komoditas sayur-sayuran yang banyak dibutuhkan oleh masyarakat, khususnya untuk konsumsi rumah tangga sebagai pelengkap bumbu dapur [2].



Menurut data statistik yang didapat melalui Survei Sosial Ekonomi Nasional (SUSENAS) tercatat bahwa adanya peningkatan konsumsi dari tahun ke tahun pada kebutuhan masyarakat Indonesia. Rata-rata konsumsi cabai rawit dalam seminggu pada tahun 2020 sebanyak 0,034 ons, yang kemudian meningkat menjadi 0,037 ons pada tahun 2021 dan 0,040 ons pada tahun 2022 hingga mencapai 0,042 ons pada tahun 2023 [3]. Hal tersebut dapat dijadikan kesempatan bagi petani untuk membudidayakan terutama pada sektor sayuran, seperti cabai rawit agar dapat menyesuaikan dengan kebutuhan konsumsi masyarakat Indonesia [2]. Permintaan cabai rawit di Indonesia yang meningkat setiap tahunnya juga diakibatkan karena bersamaan dengan jumlah penduduk yang meningkat dan tumbuhnya industri pengolahan cabai rawit [4]. Namun, peningkatan permintaan cabai rawit dapat mengalami fluktuasi yang diakibatkan karena naik turunnya harga cabai rawit. Perlu diketahui bahwa cabai rawit adalah salah satu komoditas hortikultura yang memiliki sifat musiman, sehingga ketersediaannya tidak konsisten sepanjang tahun. Hal tersebut menyebabkan fluktuasi pada harga cabai rawit [5].

Mengacu pada data dari Badan Pusat Statistik tahun 2020 hingga 2023, daerah yang memiliki produksi cabai rawit paling tinggi adalah Provinsi Jawa Timur. Pada tahun 2020 produksi cabai rawit di Jawa Timur mencapai 684.94 ton, tahun 2021 sebanyak 578.88 ton. Produksi cabai rawit di Jawa Timur kembali meningkat pada tahun 2022 sebanyak 646.74 ton dan pada tahun 2023 sebanyak 562.81 ton [6]. Menurut Badan Penelitian dan Pengembangan Pertanian Jawa Timur, cabai merupakan komoditas unggulan Jawa Timur. Cabai rawit merupakan produk pertanian strategis Jawa Timur dengan value chain (rantai pasok) yang panjang. Hal ini dapat menyebabkan harga cabai rawit yang lebih tinggi di tingkat konsumen akhir.

Peningkatan harga cabai rawit yang tinggi dapat merugikan masyarakat, terutama bagi masyarakat yang tergolong miskin. Menurut Peraturan Menteri Perdagangan Nomor 03 Tahun 2023, menjaga kestabilan harga dan ketersediaan pasokan barang kebutuhan pokok merupakan salah satu tugas dari Kementerian Perdagangan. Cabai rawit merupakan salah satu kebutuhan pokok hasil pertanian. Pemerintah harus terus memantau harga kebutuhan bahan pokok karena harga di pasar sering tidak stabil. Fluktuasi harga dan ketidakpastian harga cabai rawit di masa depan menuntut pemerintah untuk merumuskan kebijakan harga bahan pokok. Dalam menyusun kebijakan ini, pemerintah perlu didukung oleh prediksi harga cabai rawit di masa mendatang, sehingga Kementerian Perdagangan dapat mengantisipasi kenaikan dan mengendalikan harga cabai rawit.

Salah satu metode prediksi yang sering digunakan adalah metode prediksi berbasis *time series* (runtun waktu). Prediksi berdasarkan runtun waktu adalah metode prediksi yang dilakukan secara kuantitatif dengan menggunakan data terkait periode waktu tertentu [7]. Diketahui bahwa ARIMA merupakan salah satu teknik prediksi berbasis *time series* yang efektif dalam memodelkan tren dan pola musiman pada data historis yang stabil [8]. Metode ARIMA sering kali digunakan untuk peramalan jangka pendek sebab akurasinya berkurang dalam peramalan jangka panjang dan cenderung memberikan hasil yang datar atau konstan untuk periode yang panjang [9].

Menurut penelitian [10] metode *Prophet* merupakan salah satu metode yang juga baik untuk memprediksi data dalam bentuk *time series*. Model *Prophet* diketahui dapat melakukan prediksi dengan baik untuk data yang memiliki ataupun tidak memiliki efek musiman. Sifat otomatisnya memberikan fleksibilitas pada data runtun waktu yang mengalami perubahan signifikan, sehingga analisis tidak perlu khawatir datanya tidak cocok untuk prediksi dengan *Prophet*. Selain itu, model *Prophet* tahan terhadap *outlier*, *missing value*, dan unggul dalam mengatasi perubahan tren secara otomatis. *Prophet* diketahui efektif untuk memprediksi fluktuasi harga pangan karena menawarkan



tingkat kesalahan prediksi yang lebih rendah dibandingkan metode lainnya serta dapat menambahkan komponen hari libur nasional sebagai salah satu faktor dalam prediksi [11].

Berdasarkan latar belakang masalah tersebut, dilakukan perbandingan dua metode berbeda dalam memprediksi harga cabai rawit di Provinsi Jawa Timur agar mendapatkan metode yang terbaik, yaitu antara metode ARIMA dan *Prophet*. Perbandingan kedua metode tersebut dilakukan untuk mencapai model atau hasil yang optimal. Hasil dari prediksi akan diuji menggunakan *Root Mean Square Error* (RMSE) dan *Mean Absolute Percentage Error* (MAPE) sebagai ukuran ketepatan hasil prediksi.

## II. METODE PENELITIAN

Penelitian dilakukan dengan melakukan tahapan sebagai berikut :

### 1. Identifikasi masalah

Penelitian dimulai dengan mengidentifikasi permasalahan yang akan diselidiki, yaitu dengan menganalisis pola data harga cabai rawit di Provinsi Jawa Timur dari Januari 2020 hingga April 2024. Dengan merinci pola data tersebut, dapat dilakukan identifikasi potensi masalah yang mungkin muncul dalam dataset.

### 2. Pengumpulan data

Data dikumpulkan melalui situs Pusat Informasi Harga Pangan Strategis (PIHPS) Nasional, dengan fokus pada harga cabai rawit di pasar modern Provinsi Jawa Timur. Dalam penelitian ini, data harga cabai rawit di Provinsi Jawa Timur dari Januari 2020 hingga April 2024 yang berjumlah 1582 *record* diperoleh melalui *website* tersebut.

### 3. *Exploratory Data Analysis* (EDA)

*EDA* adalah tahap mempelajari dan memahami data dengan menggunakan teknik aritmatika dan visualisasi grafis yang digunakan untuk mengeksplorasi data dengan tujuan mencari pola dan meringkas informasi yang terkandung dalam data. Dalam proses ini, data dijelajahi melalui berbagai bentuk grafik, plot, dan tabel yang membantu menyajikan ringkasan statistik secara visual. *EDA* membantu untuk mendapatkan pemahaman yang lebih mendalam tentang data yang diamati sebelum masuk ke analisis lebih lanjut atau pembuatan model [12].

### 4. *Autoregressive Integrated Moving Average* (ARIMA)

Model *ARIMA* memiliki tiga parameter, yaitu  $p$ ,  $d$ , dan  $q$ . Parameter  $p$  menggambarkan periode lag yang menentukan berapa banyak periode waktu sebelumnya dari deret waktu yang digunakan dalam persamaan *auto-regression* (*AR*). Parameter  $d$  mewakili jumlah transformasi *differencing* (pembeda) yang diterapkan pada data untuk menghilangkan tren dan atau musiman, sehingga menjadikan deret waktu menjadi stasioner dengan rata-rata dan varian yang konstan dari waktu ke waktu. Parameter  $q$  mencerminkan lag dari komponen kesalahan dalam model *MA* [13]. *ARIMA* dihitung dengan menggunakan persamaan 1.

$$\Phi_p(B)(1 - B)^d Y_t = c + \theta_q(B)e_t \quad (1)$$

Keterangan :

$c$  = konstanta

$e_t$  = kesalahan (galat) pada periode ke- $t$

$(1 - B)^d$  = proses pembedaan orde ke- $d$

$\Phi_p(B)$  =  $(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)$  yaitu operator langkah mundur untuk *AR*



$\theta_q(B) = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q)$  yaitu operator langkah mundur untuk *MA*

Tahap dalam melakukan analisis pada metode *ARIMA* sebagai berikut :

1. Identifikasi stasioner *ARIMA*

Dalam analisis *ARIMA*, tahap pertama adalah mengidentifikasi apakah data yang digunakan stasioner. Uji *Augmented Dickey Fuller (ADF)* dapat digunakan untuk mengetahui stasioner rata-rata. Proses *differencing* digunakan untuk membuat hasil stasioner apabila hasil tidak stasioner dalam rata-rata. Uji *Box-Cox* dapat digunakan untuk mengetahui apakah data stasioner dalam varians. Data dapat dikatakan stasioner dalam varians apabila nilai *rounded value* ( $\lambda$ )  $\geq 1$  [14]. Transformasi digunakan untuk mengatasi ketidakstasioneran dalam varians.

2. Identifikasi model *ARIMA*

Setelah data menjadi stasioner, maka identifikasi plot *ACF* dan plot *PACF* untuk mengetahui model *ARIMA*. Dari hasil plot *ACF* dan *PACF*, dapat ditentukan model awal *ARIMA* (*AR*, *MA*, atau *ARMA*) [15]. Setelah ditentukan model awal *ARIMA*, pilih kombinasi nilai yang memberikan kesalahan atau galat paling minimal [16].

3. Estimasi parameter

Proses estimasi parameter melibatkan pengujian signifikansi untuk memastikan bahwa nilai parameter yang diestimasi memiliki tingkat signifikansi yang cukup untuk mempengaruhi model [15]. Pada uji signifikansi nilai parameter dianggap signifikan jika probabilitasnya (*P*) lebih rendah dari tingkat signifikansi  $\alpha$ . Apabila dengan interval kepercayaan 95%, maka nilai  $\alpha$  adalah 0.05. Jika probabilitas melebihi nilai  $\alpha$ , maka parameter dianggap tidak signifikan dan model tidak dapat digunakan [14].

4. *Diagnostic checking*

Uji asumsi kenormalan dan kebebasan sisaan adalah dua bagian dari pemeriksaan diagnostik. Pada uji kebebasan sisaan model dilakukan dengan uji asumsi *white noise* menggunakan *Ljung-box*. Suatu model dikatakan baik apabila nilai *p-value* pada *Ljung-box* lebih besar dari  $\alpha$  (0.05). Sementara itu, uji asumsi residual berdistribusi normal dilakukan dengan uji *Kolmogorov Smirnov* yang memiliki ketentuan apabila nilai *p-value* pada *Kolmogorov Smirnov* lebih besar dari  $\alpha$  (0.05) maka artinya model sudah memenuhi asumsi kenormalan [17].

5. Evaluasi model

Langkah berikutnya adalah melakukan evaluasi terhadap model yang terpilih tersebut. Evaluasi model dilakukan dengan melihat nilai *RMSE*, *MAPE*, dan *Akaike Information Criterion (AIC)* untuk model yang terpilih. Model yang memiliki nilai *RMSE*, *MAPE*, dan *AIC* terendah akan dipilih sebagai model yang optimal dan akan digunakan dalam prediksi [18].

5. *Prophet*

Algoritma model *prophet* termasuk kedalam *semi-automatic forecast*, yang mana dalam model ini membagi tugas antara manusia dan otomatisasi. Pada model *prophet*, langkah awal dalam proses peramalan adalah melakukan visualisasi data untuk mendapatkan gambaran awal dalam pembentukan model. Langkah kedua adalah membentuk model dengan melakukan penyesuaian parameter tren, musiman, liburan, dan komponen lainnya. Selanjutnya, model peramalan dievaluasi dengan menggunakan metrik evaluasi yang ditetapkan [19].



Model *prophet* menggunakan pendekatan waktu yang dapat dianalisis dalam tiga segmen utama yaitu tren, musiman, dan liburan. Dalam analisisnya, *prophet* memperhatikan dan memodelkan ketiga aspek tersebut untuk menghasilkan peramalan yang akurat dan informatif [19].

$$y(t) = g(t) + s(t) + h(t) + \epsilon_t \quad (2)$$

Keterangan :

$y(t)$  = *additive regression model*

$g(t)$  = fungsi tren yang memodelkan perubahan non-periodik

$s(t)$  = musiman yang mewakili perubahan periodik (mingguan dan tahunan)

$h(t)$  = efek hari libur yang muncul pada waktu tertentu

$\epsilon_t$  = nilai *error* (kesalahan)

#### 6. Root Mean Square Error (RMSE)

*Root Mean Square Error (RMSE)* lebih mudah dipahami dibandingkan dengan *Mean Square Error (MSE)* karena menggunakan skala pengukuran yang sama dengan data yang dianalisis. Semakin rendah nilai *RMSE*, semakin baik kualitas prediksi model tersebut, artinya variasi nilai yang diprediksi mendekati variasi nilai yang diamati dalam data sebenarnya [20]. Perhitungan *RMSE* dapat dilihat pada persamaan 3.

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (X_t - F_t)^2}{n}} \quad (3)$$

Keterangan :

$X_t$  = nilai sebenarnya pada periode  $t$

$F_t$  = nilai prediksi pada periode  $t$

$n$  = jumlah periode prediksi

#### 7. Mean Absolute Percentage Error (MAPE)

*Mean Mean Absolute Percentage Error (MAPE)* adalah persentase kesalahan absolut rata-rata. *MAPE* lebih baik daripada metode sebelumnya karena memberikan informasi dalam bentuk persentase kesalahan atau *error* [20]. Perhitungan *MAPE* dapat dilihat pada persamaan 4.

$$MAPE = \frac{\sum_{i=1}^n \left| \frac{X_t - F_t}{X_t} \right|}{n} \times 100\% \quad (4)$$

Keterangan :

$X_t$  = nilai sebenarnya pada periode  $t$

$F_t$  = nilai prediksi pada periode  $t$

$n$  = jumlah periode prediksi

#### 8. Pemilihan metode terbaik

Setelah mengimplementasikan model *ARIMA* dan *prophet*, langkah selanjutnya adalah memilih metode terbaik yang memiliki nilai kesalahan paling kecil. Model dengan nilai kesalahan paling kecil berdasarkan nilai *RMSE* dan *MAPE* merupakan model yang baik yang dapat digunakan untuk melakukan prediksi.

#### 9. Prediksi

Peramalan merupakan estimasi atau prediksi mengenai nilai variabel untuk periode mendatang. Berdasarkan studi literatur, diketahui bahwa peramalan jangka pendek (3 bulan sampai dengan 1 tahun) cenderung lebih akurat daripada jangka panjang (3 tahun atau lebih) karena dalam jangka



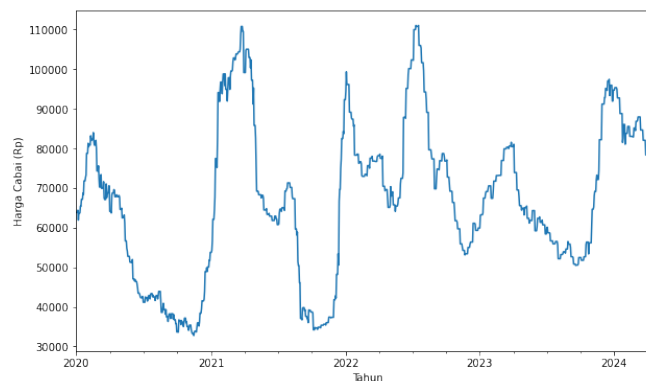
pendek, kondisi umumnya stabil atau berubah dengan lambat. Oleh karena itu, penelitian dilakukan dengan melakukan prediksi jangka pendek selama tiga bulan kedepan dimulai dari 1 Mei 2024 sampai dengan 31 Juli 2024.

### III. HASIL DAN PEMBAHASAN

Pada tahap *EDA*, dilakukan eksplorasi data secara umum, yaitu mengetahui informasi awal terkait data, visualisasi data, mengatasi *missing value*, dan mengatasi data *outlier*. Berdasarkan hasil dari perintah “.info()” di *python* bahwa terdapat *missing value* sebanyak 452 data pada kolom “**HARGA CABAI**”.

Diketahui bahwa data yang hilang tersebut merupakan data pada hari Sabtu dan Minggu. Oleh karena itu, perlu dilakukan *handling missing value* untuk nilai yang hilang tersebut. Penanganan *missing value* pada hari Sabtu dan Minggu dilakukan menggunakan nilai rata-rata pada hari Jumat dan Senin. Hal tersebut dilakukan untuk membantu menjaga konsistensi data. Selain itu, untuk membantu mengurangi bias dan memberikan gambaran yang lebih tepat terkait distribusi harga cabai rawit secara keseluruhan. Dilakukan pengecekan terhadap data *outlier* dan diketahui bahwa data pengamatan tidak mengandung *outlier* maupun nilai ekstrim, sehingga dapat dilanjutkan ke tahap eksplorasi selanjutnya.

Harga cabai rawit di Provinsi Jawa Timur seringkali mengalami perubahan setiap tahunnya. Sejak Januari 2020 hingga April 2024, harga tersebut mengalami fluktuasi yang besar dengan peningkatan dan penurunan harga yang tajam. Pada gambar 1 ditampilkan fluktuasi harga cabai rawit Provinsi Jawa Timur sejak tahun 2020 sampai dengan April 2024.



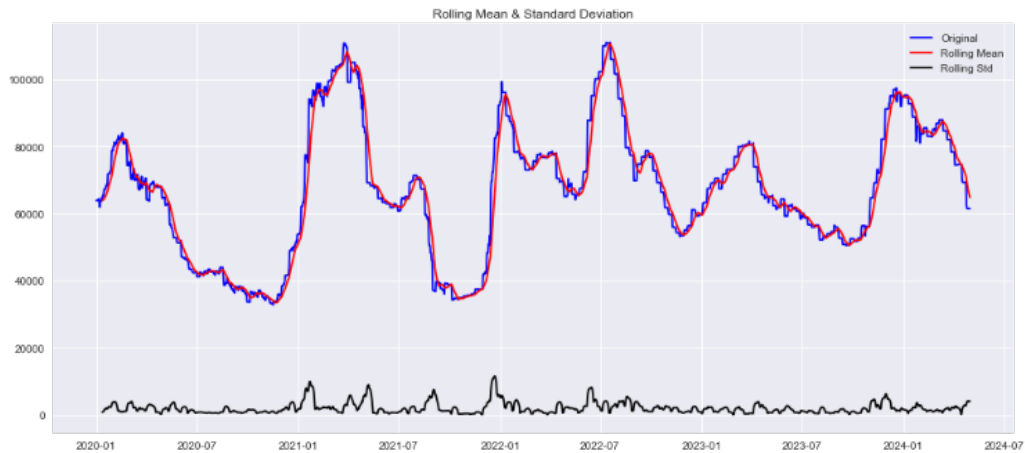
Gambar 1. Harga Cabai Rawit Januari 2020 – April 2024

#### 3.1. Metode *ARIMA*

Tahap dalam implementasi metode *ARIMA* adalah sebagai berikut :

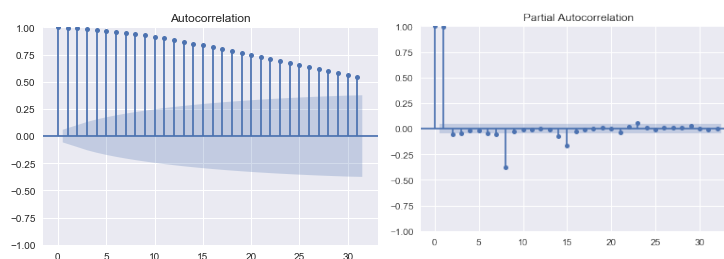
##### 1. Identifikasi stasioner

*Rolling mean* mewakili tren dalam deret waktu sedangkan *rolling standard deviation* mewakili musiman dalam deret waktu. Gambar 2 menunjukkan secara umum bahwa tren mengalami peningkatan dan penurunan secara bertahap seiring waktu dan standar deviasi memiliki nilai yang tinggi pada waktu-waktu tertentu dalam setahun yang mengindikasikan bahwa harga cabai rawit lebih bervariasi pada waktu-waktu tersebut. Secara keseluruhan dapat diidentifikasi bahwa data pengamatan belum stasioner dalam rata-rata dan varians karena *rolling mean* dan *rolling standard deviation* tidak konstan seiring waktu. Oleh karena itu, diperlukan proses transformasi dan *differencing* agar data menjadi stasioner.



Gambar 2. Stasioneritas Data Pengamatan

Uji ADF adalah cara untuk mengetahui stasioner rata-rata. Apabila  $p\text{-value} < \alpha$  (0.05), maka data dinyatakan sudah stasioner dalam rata-rata. Berdasarkan hasil uji ADF, nilai  $p\text{-value}$  adalah 0.0013 lebih kecil dari  $\alpha = 0.05$ , sehingga data harian harga cabai rawit stasioner dalam rata-rata. Hal tersebut bertolak belakang dengan hasil yang ada pada gambar 2. Pengujian ADF diperkuat dengan pengujian terhadap plot ACF dan PACF yang dapat dilihat pada gambar 3.



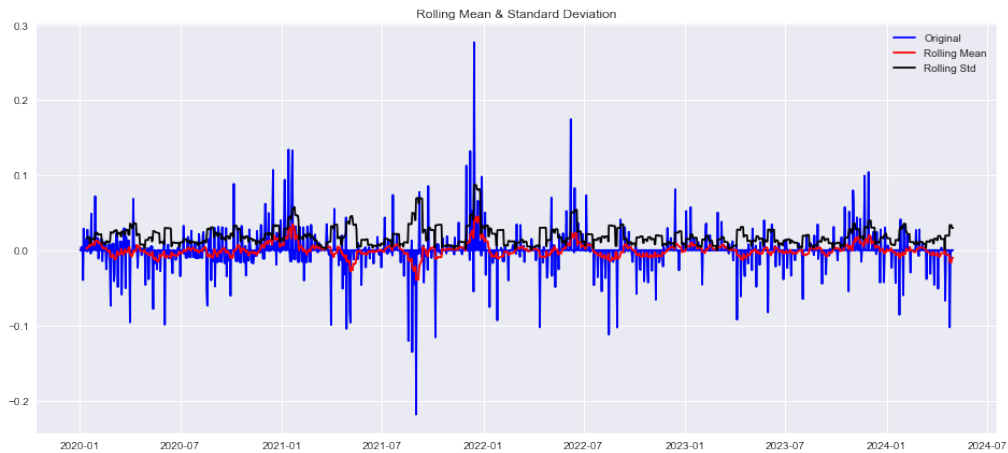
Gambar 3. Plot ACF dan PACF Data Pengamatan

Plot ACF terlihat menunjukkan pola menurun secara perlahan dan membentuk pola tangga. Sementara itu, sebagian besar plot PACF mendekati nilai nol setelah lag 0. Hasil tersebut menunjukkan bahwa data harian harga cabai rawit tidak menunjukkan sifat stasioner dalam rata-rata. Oleh karena itu, secara keseluruhan dapat disimpulkan bahwa data pengamatan masih belum stasioner dalam rata-rata.

Box-Cox adalah cara untuk mengetahui stasioner varians. Apabila nilai rounded value ( $\lambda$ )  $\geq 1$ , maka data dinyatakan sudah stasioner dalam varians. Berdasarkan nilai Box-Cox yang didapatkan, nilai  $\lambda$  adalah 0.70 kurang dari sama dengan 1, sehingga data harian harga cabai rawit tidak stasioner dalam varians.

## 2. Mengatasi data tidak stasioner

Sebelum melakukan analisis lebih lanjut, perlu dilakukan transformasi dan *differencing* untuk mengatasi masalah data yang tidak stasioner dalam rata – rata maupun varians. Salah satu transformasi yang dapat dilakukan adalah dengan log transformasi.

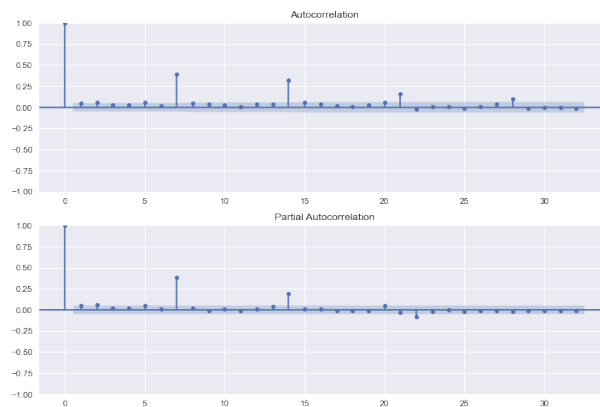


**Gambar 4.** Stasioneritas Setelah Transformasi dan *Differencing*

Gambar 4 adalah hasil plot setelah dilakukan log transformasi dan *differencing*. Berdasarkan plot tersebut dapat terlihat bahwa *rolling mean* dan *rolling std* memiliki nilai yang relatif stabil disekitar nol. Hal tersebut menunjukkan bahwa tidak ada tren yang signifikan dalam data dari waktu ke waktu.

### 3. Identifikasi model

Identifikasi model  $AR(p)$  dan  $MA(q)$  dilakukan berdasarkan plot *ACF* dan *PACF* yang stasioner. Model  $AR(p)$  digambarkan dalam plot *ACF*, sedangkan model  $MA(q)$  digambarkan dalam plot *PACF*. Gambar 5 adalah plot *ACF* dan *PACF* berdasarkan data yang sudah stasioner.



**Gambar 5.** Plot *ACF* dan *PACF* Data Stasioner

Pada plot *ACF* dan *PACF* terlihat bahwa lag ke-1, 2, 5, dan 7 keluar dari batas signifikan. Sebelumnya sudah diketahui bahwa nilai *differencing* adalah sama dengan 1. Oleh karena itu, diperoleh beberapa kombinasi model yang mungkin. Kombinasi tersebut dapat dilihat pada tabel 1. Model dengan nilai *AIC* terendah akan dipilih sebagai model yang optimal untuk digunakan. Dapat disimpulkan bahwa model  $ARIMA(7, 1, 7)$  adalah model yang optimal untuk digunakan.

**Tabel 1.** Nilai *AIC* Dugaan Model  $ARIMA$

Model	Nilai <i>AIC</i>
$ARIMA(1, 1, 1)$	27718.301
$ARIMA(2, 1, 2)$	27715.519
$ARIMA(5, 1, 5)$	27629.276
$ARIMA(7, 1, 7)$	27544.663





#### 4. Estimasi parameter

Estimasi parameter model bertujuan untuk mengevaluasi apakah model yang terpilih memenuhi kriteria yang ditetapkan. Apabila selang kepercayaan 95%, maka nilai  $\alpha$  adalah 0,05. Jika probabilitas melebihi  $\alpha$  ( $P > \alpha$ ), parameter model tersebut dianggap tidak signifikan. Estimasi parameter model *ARIMA* (7, 1, 7) dapat dilihat pada tabel 2. Hasil uji signifikansi menunjukkan bahwa model *ARIMA* (7, 1, 7) signifikan, sehingga model dapat digunakan.

**Tabel 2.** Uji Signifikansi *ARIMA* (5, 1, 5)

Model	Type	P	Tingkat Signifikansi
(7, 1, 7)	AR 7	0.000	Signifikan
	MA 7	0.000	

#### 5. Diagnostic checking

Berdasarkan hasil uji *Ljung-Box*, nilai *p-value* adalah 0.999 yang mana lebih besar dari  $\alpha$  (0.05). Oleh karena itu, dapat disimpulkan bahwa residual memiliki sifat *white noise*. Artinya, model *ARIMA* (7 1, 7) memenuhi untuk digunakan. Sementara itu, hasil yang uji normalitas dengan *Kolmogorov Smirnov* diperoleh, nilai *p-value* adalah  $6.760 \times 10^{-118}$  yang mana lebih kecil dari  $\alpha$  (0.05). Oleh karena itu, uji normalitas tidak memenuhi ketentuan yang sudah ditetapkan.

Menurut yang sudah diuraikan, terdapat salah satu uji yang tidak terpenuhi dalam *ARIMA* yaitu uji normalitas. Namun, berdasarkan nilai *AIC* yang didapatkan dari dugaan model *ARIMA* menunjukkan bahwa model *ARIMA* (7, 1, 7) memiliki nilai *AIC* yang terendah. Selain itu, menurut [21] bahwa residu berguna (tetapi tidak perlu) memiliki dua sifat yang salah satu diantaranya adalah residu berdistribusi normal. Oleh karena itu, *ARIMA* (7, 1, 7) terpilih sebagai model yang paling optimal.

#### 6. Evaluasi model

Langkah pertama yang dilakukan adalah membagi data menjadi data latih dan data uji dengan rasio 80 : 20. Berdasarkan hasil yang diperoleh, nilai *MAPE* adalah 19.48%, maka dapat disimpulkan bahwa kemampuan prediksi adalah baik. Sementara itu, nilai *RMSE* adalah 18723.92 yang berarti nilai rata-rata dari selisih antara nilai prediksi dan nilai aktual adalah 18723.92.

#### 3.2. Metode *Prophet*

Terdapat empat model yang dilakukan dengan menggunakan metode *prophet*. Sebelum menerapkan empat model tersebut, terdapat beberapa hal yang dilakukan, yaitu.

##### a. Mengubah nama kolom

Pengubahan nama kolom dilakukan untuk menyesuaikan dengan format yang ditentukan oleh *prophet*. Kolom yang diubah yaitu kolom “**Tanggal**” diubah menjadi “**ds**” dan kolom “**HARGA CABAI**” diubah menjadi “**y**”.

##### b. Membagi data

Pembagian data dilakukan dengan rasio yang sama seperti pada metode *ARIMA* sebelumnya, yaitu data latih dan data uji sebesar 80 : 20. Setelah itu, data latih dan uji dilakukan pengubahan nama kolom sesuai seperti format pada *prophet*.

##### c. Evaluasi model

Membuat fungsi evaluasi model dengan menggunakan dua metrik kesalahan prediksi, yaitu *RMSE* dan *MAPE*. Hal tersebut dilakukan agar mempermudah pada saat perhitungan evaluasi



model. Sehingga hanya perlu memanggil nama fungsinya saja ketika ingin menampilkan hasil evaluasi model.

Berikut adalah empat model yang dilakukan :

#### 1. Model *prophet* dasar

Model *prophet* dasar adalah model *prophet* yang terbentuk secara *default* tanpa adanya proses *tuning* komponen. Artinya, tidak terdapat komponen hari libur, sehingga hanya ada komponen *trend* dan *seasonal* yang digunakan secara *default*. Langkah dalam menerapkan model *prophet* dasar dimulai dari melakukan inisialisasi model, membuat prediksi, dan melakukan evaluasi model. Berdasarkan hasil yang didapatkan, nilai *MAPE* sebesar 53.05% yang berarti model *prophet* dasar memiliki kemampuan prediksi yang buruk. Nilai *RMSE* didapat sebesar 46295.28 yang berarti nilai rata-rata dari selisih antara nilai prediksi dan nilai aktual adalah 46295.28. Persamaan 5 terbentuk berdasarkan hasil dari model *prophet* dasar.

$$y(t) = ((-1,23427 + \alpha(t)^T \delta)t + (0,572808 + \alpha(t)^T \gamma)) + X(t)\beta + \varepsilon_t \quad (5)$$

#### 2. Model *prophet* tambahan hari libur

Tahapan dalam menerapkan model *prophet* dengan komponen hari libur sama seperti model *prophet* dasar sebelumnya, tetapi terdapat komponen tambahan berupa data hari libur. Penggunaan hari libur Indonesia sudah tersedia dalam *library prophet* dengan menuliskan kode negara “ID”. Pada model ini, diperoleh nilai *MAPE* sebesar 50.80% menunjukkan hasil yang lebih baik jika dibandingkan dengan model sebelumnya. Artinya dengan menggunakan tambahan komponen hari libur bisa membantu meningkatkan akurasi model. Nilai *RMSE* sebesar 44797.89, hasil tersebut menunjukkan nilai yang lebih kecil daripada model *prophet* dasar. Persamaan 6 terbentuk berdasarkan hasil dari model *prophet* tambahan komponen hari libur.

$$y(t) = ((-1,31509 + \alpha(t)^T \delta)t + (0,579533 + \alpha(t)^T \gamma)) + X(t)\beta + Z(t)k + \varepsilon_t \quad (6)$$

#### 3. Model *prophet* tambahan musiman periode bulanan

Model *prophet* dengan tambahan musiman periode bulanan memiliki tahapan yang sama seperti model *prophet* dasar, tetapi terdapat tambahan komponen musiman dengan menggunakan periode bulanan. Musiman ditambahkan ke dalam model dengan menggunakan perintah “.add\_seasonality()”.

Nilai *MAPE* yang diperoleh adalah 52.63%, nilai tersebut menunjukkan hasil yang tidak lebih baik dibandingkan dengan model sebelumnya. Selain itu, nilai *RMSE* sebesar 46025.07 juga menunjukkan angka yang lebih besar jika dibandingkan dengan model sebelumnya. Hasil *RMSE* dan *MAPE* mengindikasikan bahwa dengan menambahkan komponen musiman periode bulanan tidak dapat membantu mengurangi kesalahan prediksi. Persamaan 7 terbentuk berdasarkan hasil dari model *prophet* tambahan komponen musiman periode bulanan.

$$y(t) = ((-1,25766 + \alpha(t)^T \delta)t + (0,575164 + \alpha(t)^T \gamma)) + X(t)\beta + \varepsilon_t \quad (7)$$

#### 4. Model *prophet* tambahan hari libur dan musiman periode bulanan

Model ini merupakan dari dua model sebelumnya. Nilai *MAPE* yang diperoleh adalah 49.81% menunjukkan penurunan nilai dibandingkan dengan model-model sebelumnya. Selain itu, nilai *RMSE* sebesar 44133.11 yang juga lebih rendah daripada ketiga model sebelumnya. Hal tersebut mengindikasikan bahwa model *prophet* dengan kombinasi komponen hari libur dan musiman periode bulanan lebih baik untuk meningkatkan akurasi model. Persamaan 8 terbentuk



berdasarkan hasil dari model *prophet* tambahan komponen hari libur dan musiman periode bulanan.

$$y(t) = ((-1,38284 + \alpha(t)^T \delta)t + (0,585886 + \alpha(t)^T \gamma)) + X(t)\beta + Z(t)k + \varepsilon_t \quad (8)$$

### 3.3. Pemilihan Metode Terbaik

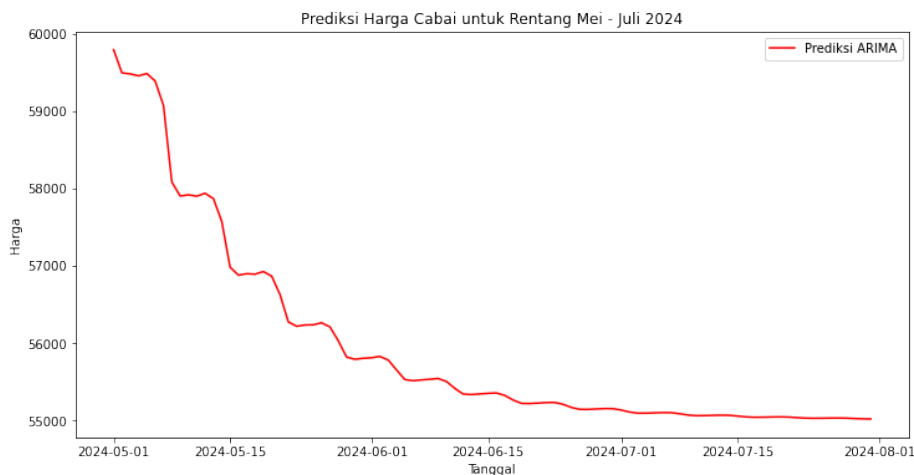
Berdasarkan hasil pengujian seluruh metode peramalan, yaitu *ARIMA* dan *prophet* diperoleh nilai *MAPE* dan *RMSE* seperti pada tabel 3.

**Tabel 3.** Perbandingan Nilai *MAPE* dan *RMSE*

Metode	<i>MAPE</i>	<i>RMSE</i>
<i>ARIMA</i> (7, 1, 7)	19.48%	18723.92
<i>Prophet</i> dasar	53.05%	46295.28
<i>Prophet</i> tambahan hari libur	50.80%	44797.89
<i>Prophet</i> tambahan musiman periode bulanan	52.63%	46025.07
<i>Prophet</i> tambahan hari libur dan musiman periode bulanan	49.81%	44133.11

Model dengan nilai kesalahan paling kecil merupakan model yang baik yang dapat digunakan untuk melakukan prediksi. Oleh karena itu, metode *ARIMA* (7, 1, 7) dipilih sebagai metode yang digunakan untuk melakukan prediksi harga cabai rawit Provinsi Jawa Timur selama tiga bulan mendatang dimulai dari Mei sampai dengan Juli 2024

### 3.4. Prediksi



**Gambar 6.** Prediksi Harga Cabai Rawit Mei – Juli 2024 dengan *ARIMA* (7, 1, 7)

Hasil prediksi menunjukkan bahwa harga cabai rawit mencapai puncak harga pada awal Mei 2024 dan menurun secara berkala setelahnya. Diketahui bahwa puncak harga cabai rawit yang terjadi pada 1 Mei 2024 adalah sebesar Rp59.794. Puncak harga pada awal Mei dapat berhubungan dengan beberapa faktor, antara lain adalah permintaan atau konsumsi cabai rawit yang tinggi untuk perayaan Hari Buruh. Selain itu, disebabkan oleh adanya penurunan produksi cabai rawit karena faktor cuaca atau siklus panen.

Sementara itu, harga cabai rawit paling rendah adalah sebesar Rp55.017 yang terjadi pada 31 Juli 2024. Penurunan harga secara berkala setelah Mei hingga mencapai titik terendah pada akhir Juli dapat disebabkan oleh pasokan cabai rawit yang meningkat pada bulan tersebut karena panen yang lebih baik. Adapun faktor lain, seperti penurunan permintaan terutama setelah bulan puasa dan lebaran. Liburan sekolah yang biasanya ada pada Bulan Juni dan Juli juga dapat mempengaruhi



permintaan cabai rawit. Selama periode tersebut, konsumsi rumah tangga mungkin menurun karena banyak keluarga yang bepergian atau mengurangi aktivitas memasak dirumah, sehingga mempengaruhi penurunan harga cabai rawit.

#### IV. KESIMPULAN

Model prediksi harga cabai rawit Provinsi Jawa Timur terbaik adalah *ARIMA* (7, 1, 7) dengan *RMSE* sebesar 18723.92 dan *MAPE* sebesar 19.48. Berdasarkan hasil prediksi harga cabai rawit di Provinsi Jawa Timur dari Mei hingga Juli 2024 menunjukkan penurunan harga secara berkala. Diperkirakan harga cabai rawit mencapai puncak harga pada 1 Mei 2024 dengan perkilogramnya adalah Rp59.794. Sementara itu, harga terendah terjadi pada 31 Juli 2024 dengan perkilogramnya adalah Rp55.017. Peningkatan dan penurunan harga tersebut dapat disebabkan oleh banyak kombinasi faktor, seperti faktor musim tanam yang belum memasuki masa panen, cuaca buruk, atau peningkatan maupun penurunan permintaan.

#### UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada seluruh pihak yang telah mendukung terwujudnya penelitian ini. Tak lupa, penulis sampaikan ucapan terima kasih kepada Institut Teknologi Telkom Purwokerto.

#### REFERENSI

1. Kementerian Pertanian, *Analisis Kinerja Perdagangan Cabai Merah*. Pusat Data dan Sistem Informasi Pertanian Kementerian Pertanian, 2021. [Online]. Available: [https://satudata.pertanian.go.id/assets/docs/publikasi/Analisis\\_Kinerja\\_Perdagangan\\_Komoditas\\_Cabai\\_Merah\\_Semester\\_I\\_Tahun\\_2021.pdf](https://satudata.pertanian.go.id/assets/docs/publikasi/Analisis_Kinerja_Perdagangan_Komoditas_Cabai_Merah_Semester_I_Tahun_2021.pdf)
2. Z. Arifin, “Kontribusi Usahatani Pembibitan Cabai Rawit Terhadap Pendapatan Rumah Tangga Kelompok Tani Perdi Di Desa Dilem Kecamatan Kepanjen Kabupaten Malang,” *J. Pertan. Cemara*, vol. 18, no. 1, pp. 25–41, 2021, doi: 10.24929/fp.v18i1.1160.
3. Badan Pusat Statistik, “Rata-Rata Konsumsi per Kapita Seminggu Beberapa Macam Bahan Makanan Penting, 2007-2023,” *Badan Pusat Statistik*, 2024. <https://www.bps.go.id/id/statistics-table/1/OTUwIzE=/rata-rata-konsumsi-per-kapita-seminggu-beberapa-macam-bahan-makanan-penting--2007-2023.html> (accessed Jun. 24, 2024).
4. D. Septiadi, N. M. W. Sari, and A. Zainuddin, “Analisis Permintaan Konsumsi Cabai Rawit pada Rumah Tangga di Kota Mataram,” *Agrimor*, vol. 5, no. 2, pp. 36–39, 2020, doi: 10.32938/ag.v5i2.1013.
5. F. Hilmiyah and A. Supriono, “Keterpaduan Pasar dan Transmisi Harga Cabai Rawit di Indonesia: Pendekatan Vector Error Correction Model,” *JSEP (Journal Soc. Agric. Econ.*, vol. 15, no. 2, p. 209, 2022, doi: 10.19184/jsep.v15i2.24690.
6. BPS, “Produksi Tanaman Sayuran,” *Badan Pusat Statistik*. <https://www.bps.go.id/indicator/55/61/1/produksi-tanaman-sayuran.html>
7. I. Ardiansah, I. F. Adiarsa, S. H. Putri, and T. Pujiyanto, “Penerapan Analisis Runtun Waktu pada Peramalan Penjualan Produk Organik menggunakan Metode Moving Average dan Exponential Smoothing,” *J. Tek. Pertan. Lampung (Journal Agric. Eng.*, vol. 10, no. 4, p. 548, 2021, doi: 10.23960/jtep-l.v10i4.548-559.
8. A. D. Milniadi and N. O. Adiwijaya, “ANALISIS PERBANDINGAN MODEL ARIMA DAN LSTM DALAM PERAMALAN HARGA PENUTUPAN SAHAM (STUDI KASUS: 6 KRITERIA KATEGORI SAHAM MENURUT PETER LYNCH),” *SIBATIK J. J. Ilm. Bid. Sos. Ekon. Budaya, Teknol. dan Pendidik.*, vol. 2, no. 6, pp. 1683–1692, 2023, doi: 10.54443/sibatik.v2i6.798.
9. Syaharuddin, Q. S. Akmal, and L. Sucipto, “Metode ARIMA, ARIMAX, dan SARIMA: Sebuah meta-analisis perbedaan tingkat akurasi peramalan data time series,” *J. Inform. Kaputama*, vol. 6, no. 3, pp. 502–509, 2022.
10. M. A. Muzakki, M. A. Sabila, S. Sundari, and B. Wisnuadhi, “Analisis Algoritma Prophet untuk Memprediksi Harga Pangan di Kota Bandung,” *Pros. Ind. Res. Work. Natl. Semin.*, vol. 12, pp. 659–



- 664, 2021.
11. S. J. Taylor and B. Letham, “Forecasting at Scale,” *Am. Stat.*, vol. 72, no. 1, pp. 37–45, 2018, doi: 10.1080/00031305.2017.1380080.
12. M. E.-K. Kesuma and R. Iskandar, “Analisis Toko dan Asal Toko Fashion Pria di Shopee Menggunakan Data Scrapping dan Exploratory Data Analysis,” *Maj. Ilm. Teknol. Elektro*, vol. 21, no. 1, p. 127, 2022, doi: 10.24843/mite.2022.v21i01.p17.
13. H. Alabdulrazzaq, M. N. Alenezi, Y. Rawajfih, B. A. Alghannam, A. A. Al-Hassan, and F. S. Al-Anzi, “On the accuracy of ARIMA based prediction of COVID-19 spread,” *Results Phys.*, vol. 27, p. 104509, 2021, doi: 10.1016/j.rinp.2021.104509.
14. S. Suseno and Suryo Wibowo, “Penerapan Metode ARIMA dan SARIMA Pada Peramalan Penjualan Telur Ayam Pada PT Agromix Lestari Group,” *J. Teknol. dan Manaj. Ind. Terap.*, vol. 2, no. 1, pp. 33–40, 2023, doi: 10.55826/tmit.v2i1.85.
15. T. Yunita, “Peramalan Jumlah Penggunaan Kuota Internet Menggunakan Metode Autoregressive Integrated Moving Average (ARIMA),” *J. Math. Theory Appl.*, vol. 1, no. 2, pp. 16–22, 2020, doi: 10.31605/jomta.v2i1.777.
16. S. R. A. Arifai and Lukman Junaedi, “Prediksi Permintaan Barang Berdasarkan Penjualan Menggunakan Metode Arima Box-Jenkins (Studi Kasus : Pt. Beststamp Indonesia),” *J. E-Bis*, vol. 4, no. 2, pp. 138–146, 2020, doi: 10.37339/e-bis.v4i2.227.
17. Wahyu ngestisari, “Perbandingan Metode ARIMA dan Jaringan Syaraf Tiruan untuk Peramalan Harga Beras,” *Indones. J. Data Sci.*, vol. 1, no. 3, pp. 96–107, 2020, doi: 10.33096/ijodas.v1i3.18.
18. M. PRANATA, D. ANGGRAINI, D. Makbuloh, and A. Rinaldi, “Prediksi Pencurian Sepeda Motor Menggunakan Model Time Series (Studi Kasus: Polres Kotabumi Lampung Utara),” *BAREKENG J. Ilmu Mat. dan Terap.*, vol. 14, no. 3, pp. 425–434, 2020, doi: 10.30598/barekengvol14iss3pp425-434.
19. R. S. Pontoh, S. Zahroh, H. R. Nurahman, R. I. Aprillion, A. Ramdani, and D. I. Akma, “Applied of feed-forward neural network and facebook prophet model for train passengers forecasting,” *J. Phys. Conf. Ser.*, vol. 1776, no. 1, 2021, doi: 10.1088/1742-6596/1776/1/012057.
20. N. N. S. Wahyuni, I. M. Candiasa, and M. Juniantari, “Prediksi Angka Partisipasi Sekolah Menurut Kelompok Umur Di Provinsi Bali Menggunakan Single Moving Average,” *Maju*, vol. 7, no. 2, pp. 100–109, 2020.
21. R. J. Hyndman and G. Athanasopoulos, *Forecasting: Principles and Practice*. Melbourne, Australia: OTexts, 2018. [Online]. Available: <https://otexts.com/fpp2/index.html>