e-ISSN: 2580-3239
p-ISSN: 2085-7993
Volume 22 No. 02 November 2023

Pemrograman Python Untuk Peramalan Data Deret Waktu Menggunakan Metode *Seasonal Autoregressive Integrated Moving Average* (Sarima)

Michelle Selina Buntara, Herlina Napitupulu, Nurul Gusriani
Departemen Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Padjadjaran
Email: michelle19003@mail.unpad.ac.id, herlina@unpad.ac.id, nurul.gusriani@unpad.ac.id

Abstrak

Peramalan deret waktu adalah penggunaan model untuk memprediksi nilai masa depan berdasarkan nilai yang diamati sebelumnya. Model *Seasonal Autoregressive Integrated Moving Average* (SARIMA) merupakan salah satu model yang digunakan untuk peramalan ketika deret waktu univariat menunjukkan variasi musiman. Model SARIMA merupakan bentuk khusus dari model ARIMA yang terdiri dari tiga bagian, yaitu; ‘AR’ yang berarti *Autoregressive*, ‘I’ yang merupakan bagian *differencing*, dan ‘MA’ yang berarti *Moving Average*.Penelitian ini bertujuan untuk memperoleh model SARIMA terbaik melalui beberapa tahap, yaitu; preparasi, identifikasi, penaksiran nilai parameter, dan uji diagnostik. Performa model peramalan diuji menggunakan *mean absolute percentage error* (MAPE).

Kata Kunci: peramalan; deret waktu; SARIMA

Abstract

Time series forecasting is the use of models to predict future values based on previously observed values. The Seasonal Autoregressive Integrated Moving Average (SARIMA) model is one of the models used for forecasting when univariate time series exhibit seasonal variations. The SARIMA model is a special form of the ARIMA model, consisting of three parts: 'AR' stands for Autoregressive, 'I' represents the differencing component, and 'MA' stands for Moving Average. This research aims to obtain the best SARIMA model through several stages, including preparation, identification, parameter value estimation, and diagnostic testing. The forecasting model's performance is assessed using the mean absolute percentage error (MAPE).

Keywords: *forecasting; time series; SARIMA*

e-ISSN: 2580-3239
p-ISSN: 2085-7993
Volume 22 No. 02 November 2023

1 PENDAHULUAN

Peramalan (*forecasting*) adalah seni dan ilmu untuk memperkirakan kejadian di masa depan (Heizer, 2011). Peramalan dapat diimplementasikan dalam berbagai bidang mulai dari pendidikan, kesehatan, pembangunan, ekonomi, sosial, politik, dan lain-lain. Peramalan kerap dilakukan pada data deret waktu untuk memprediksi nilai masa depan. Deret waktu merupakan serangkaian observasi terhadap suatu variabel yang diambil secara beruntun berdasarkan interval waktu yang tetap (Wei, 2006). Analisis deret waktu merupakan suatu metode kuantitatif untuk menentukan statistik dan karakteristik data yang berarti. Peramalan deret waktu adalah penggunaan model untuk memprediksi nilai masa depan berdasarkan nilai yang diamati sebelumnya. Tujuan dari peramalan deret waktu adalah untuk meramalkan peristiwa masa depan berdasarkan pola data peristiwa sebelumnya yang telah diamati dan dikumpulkan pada interval waktu yang teratur (Pham, 2006).

Dalam kumpulan data deret waktu, data tidak teratur dapat muncul. Data tidak teratur ini dihasilkan dari variabel fitur tunggal yang diukur pada laju pengambilan sampel tanpa interval yang konsisten antara pengamatan, baik karena proses manual yang tidak terstruktur, kegagalan perangkat atau sinyal, ataupun penghilangan yang disengaja (Weerakody *et al.*, 2021). Diskontinuitas menimbulkan hambatan yang signifikan untuk skema prediksi deret waktu, yang umumnya memerlukan data kontinu sebagai syarat penggunaan. Hal ini tentu akan mengarah pada penerapan teknik imputasi untuk mengatasi masalah tersebut (Junninen *et al.*, 2004). Banyak metode telah dikembangkan selama bertahun-tahun untuk menangani data deret waktu yang tidak teratur sehingga dapat menghasilkan model prediksi yang mendekati model yang menangani data teratur. Salah satu metode yang populer adalah interpolasi linear. Metode ini memperkirakan nilai yang hilang dengan menghubungkan titik-titik dalam garis lurus. Pada tahun 2015, Saaban, Zainudin, dan Bakar mengevaluasi kemampuan metode interpolasi linear dalam memprediksi data radiasi matahari di Perlis. Berdasarkan studi tersebut, interpolasi linear berhasil untuk memperkirakan nilai yang hilang pada dataset radiasi matahari yang mengandung nilai hilang 5% dan 10%.

Data deret waktu sering menunjukkan variasi musiman, terutama data ekonomi dan bisnis. Pada tahun 1976, Box dan Jenkins mengembangkan model deret waktu yang disebut *Seasonal Autoregressive Integrated Moving Average* atau dikenal dengan SARIMA, khusus untuk deret yang memiliki sifat musiman. Model SARIMA dibentuk dari model ARIMA (*Autoregressive Integrated Moving Average*) dengan memasukkan parameter musiman tambahan ke dalam model. SARIMA merupakan model yang sederhana dan membutuhkan sedikit data (Siami-Namini et al., 2018), sehingga model ini banyak diimplementasikan pada berbagai jenis permasalahan penelitian yang berkaitan dengan *time series*, diantaranya Chen et al (2018) yang meramalkan temperatur rata-rata bulanan kota Nanjing menggunakan SARIMA. Model terbaik yang didapatkan adalah $\mathrm{SARIMA}(1,1,1)(1,0,1)_{12}$ dengan nilai MSE 0.89. Kemudian Harrison et al (2020) yang meramalkan jumlah produksi dan ekspor minyak kelapa sawit Nigeria menggunakan SARIMA dengan model terbaik adalah $\mathrm{SARIMA}(1,0,1)(2,0,0)_{12}$ dan $\mathrm{SARIMA}(2,1,0)(1,0,1)_{12}$. Lalu terdapat juga Pongdatu dan Putra (2018) yang membandingkan metode SARIMA dengan *Holt Winter's exponential smoothing* dalam meramalkan data penjualan toko retail pakaian X. Hasil perbandingan yang didapat adalah metode SARIMA menunjukkan keakuratan yang lebih baik dengan nilai MAD lebih kecil.

e-ISSN: 2580-3239
p-ISSN: 2085-7993
Volume 22 No. 02 November 2023

2 KAJIAN PUSTAKA

2.1 Preparasi Data

Preparasi data merupakan langkah persiapan dan pengolahan data awal sebelum data tersebut dapat dimanfaatkan untuk analisis atau pemodelan. Tujuan utama dari preparasi data adalah memastikan bahwa data yang akan digunakan memiliki kualitas yang optimal, tidak terdapat kesalahan, terorganisir dengan baik, dan siap untuk dilakukan analisis lebih lanjut. Proses preparasi data melibatkan beberapa langkah, antara lain:

1. Pembersihan Data
   Dalam tahap pembersihan data ini, dilakukan penanganan terhadap kesalahan atau ketidaksesuaian dalam data. Contohnya adalah penanganan data hilang atau tidak lengkap dan menangani outlier. Ada beberapa pendekatan statistik yang banyak digunakan untuk menangani nilai-nilai yang hilang dari dataset, seperti menggantikan dengan rata-rata, median, atau mode atribut. Banyak peneliti juga mengusulkan berbagai solusi lain yang menargetkan imputasi data biner, nominal, atau numerik (Khan dan Hoque, 2020). Interpolasi linear merupakan salah satu metode yang cukup banyak digunakan dalam menangani data hilang. Interpolasi linear adalah metode imputasi dasar yang mengasumsikan hubungan linear antara nilai yang hilang dan yang tidak hilang (Zhang dan Chen, 2001). Interpolasi Linear menyesuaikan garis lurus antara titik akhir celah dan memungkinkan nilai yang hilang dihitung secara langsung menggunakan persamaan garis. Persamaan interpolasi linear dapat dinyatakan dalam persamaan 1 (Hamzah *et al*., 2020).

$$f_1(x) = f(x_0) + \frac{f(x_1) - f(x_0)}{x_1 - x_0} \cdot (x - x_0) \qquad (1)$$

2. Pembagian Data
   Dalam tahap pembagian data, dilakukan pemisahan data secara keseluruhan menjadi dua kelompok data yang memiliki fungsi yang berbeda, yaitu data latih dan data validasi. Data latih digunakan untuk melatih model, sementara data validasi digunakan untuk menguji kinerja model dan memvalidasi hasilnya.

2.2 Model SARIMA

Model Seasonal Autoregressive Integrated Moving Average (SARIMA) merupakan bentuk khusus dari model ARIMA yang mendukung data deret waktu univariat dengan unsur musiman. Bagian musiman dari model terdiri dari suku-suku yang sangat mirip dengan komponen model ARIMA, namun melibatkan pergeseran balik dari periode musiman (Hyndman dan Athanasopoulos, 2018). Model SARIMA ditulis dalam bentuk $(p, d, q)(P, D, Q)_S$. (p, d, q) merupakan bagian non-musiman, $(P, D, Q)_S$ merupakan bagian musiman, dan s menunjukkan panjang periode musiman. Persamaan umum model SARIMA$(p, d, q)(P, D, Q)_S$ adalah sebagai berikut (Wei, 2006).

$$\phi_p(B)\Phi_P(B^S)(1-B)^d(1-B^S)^D Z_t = \theta_q(B)\Theta_Q(B^S)e_t \qquad (2)$$

Z_t : Nilai pengamatan pada waktu ke-t $(t = 1, 2, \dots, n)$

$\phi_p(B)$: Parameter *Autoregressive* non-musiman dengan orde p $\left(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p\right)$

$\Phi_{\mathrm{P}}(B^S)$: Parameter *Autoregressive* musiman dengan orde P $(1 - \Phi_1 B^S - \Phi_2 B^{2S} - \cdots - \Phi_P B^{PS})$

$\theta_q(B)$: Parameter *Moving Average* non-musiman dengan orde q $\left(1 - \theta_1 B - \theta B^2 - \cdots - \theta_q B^q\right)$

$\Theta_Q(B^S)$: Parameter *Moving Average* musiman dengan orde Q $\left(1 - \Theta B^S - \Theta_2 B^{2S} - \cdots - \Theta_Q B^{QS}\right)$

$(1 - B)^d$: Operasi matematis dari *differencing* non-musiman

$(1 - B^S)^D$: Operasi matematis dari *differencing* musiman

e_t : Eror pada waktu ke-t

2.3 Stasioneritas

Pengujian stasioneritas data adalah hal yang penting dalam analisis regresi data deret waktu. Stasioneritas merupakan konsep

e-ISSN: 2580-3239
p-ISSN: 2085-7993
Volume 22 No. 02 November 2023

penting dalam analisis deret waktu. Data deret waktu dikatakan stasioner apabila nilai rata-rata dan variansnya tidak mengalami perubahan yang secara sistematik sepanjang waktu atau dengan kata lain, rata-rata dan variansnya konstan (Makridakis et al., 1999).

Augmented Dickey-Fuller Test (ADF) merupakan metode uji yang efisien untuk mengevaluasi stasioneritas deret waktu dalam rata-rata (Wang et al., 2021). Persamaan 3 berikut merupakan uji statistik ADF.

$$\hat{\tau} = \frac{\hat{\delta}}{SE(\hat{\delta})} \tag{3}$$

Metode Box-Cox dapat digunakan untuk menguji kestasioneran data dalam varians. Transformasi Box-Cox merupakan transformasi pangkat berparameter tunggal λ. Data dapat dikatakan stasioner dalam varians apabila nilai $\lambda = 1$. Menurut Box dan Cox (1964), transformasi Box-Cox didefinisikan sebagai persamaan 4 berikut

$$(\lambda) = \begin{cases} \frac{Z^{\lambda}-1}{\lambda}, \lambda \neq 0 \\ \ln Z \ , \lambda = 0 \end{cases} \tag{4}$$

2.4 *Differencing*

Differencing perlu dilakukan jika suatu data deret waktu tidak stasioner dalam rata-rata. Proses differencing dilakukan dengan cara mengurangi nilai data pada suatu periode dengan nilai data periode sebelumnya secara berurutan menggunakan operator backward shift yang dinotasikan sebagai berikut (Hyndman dan Athanasopoulos, 2018).

$$Z_t^d = (1-B)^d Z_t \tag{5}$$

dengan B merupakan operator *backward shift*.

2.5 ACF dan PACF

Dalam analisis deret waktu, *Autocorrelation Function* (ACF) dan Partial *Autocorrelation Function* (PACF) merupakan hal penting untuk menentukan model dari data. ACF berfungsi untuk menaksir orde model MA dan PACF digunakan untuk menaksir orde model AR. Koefisien autokorelasi (ACF) berarti korelasi deret berkala dengan deret berkala itu sendiri yang memiliki selisih waktu k (lag-k). ACF didefinisikan pada persamaan 6 berikut (Wei, 2006).

$$\hat{\rho}_k = \frac{\sum_{t=1}^{n-k}(Z_t-\bar{Z})(Z_{t+k}-\bar{Z})}{\sum_{t=1}^{n}(Z_t-\bar{Z})^2} \tag{6}$$

Rumus PACF didefinisikan pada persamaan 7 berikut.

$$\hat{\beta}_{kk} = \frac{\hat{\rho}_k - \sum_{j=1}^{k-1}\hat{\beta}_{k-1,j}\,\hat{\rho}_{k-j}}{1-\sum_{j=1}^{k-1}\hat{\beta}_{k-1.j}\,\hat{\rho}_j} \tag{7}$$

2.6 Penaksiran Nilai Parameter

Salah satu metode statistik yang dapat digunakan dalam menaksirkan parameter model adalah *Maximum Likelihood Estimation* (MLE) (Wei, 2006). MLE menggunakan prinsip memaksimumkan fungsi likelihood dengan menaksir parameter $\phi, \theta, \Phi\ dan\ \Theta$ pada model SARIMA. Persamaan 8 berikut merupakan fungsi kepadatan peluang dari MLE pada model SARIMA.

$$(e|\phi,\theta,\Phi,\Theta,\sigma_e^2) = (2\pi\sigma_e^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma_e^2}\sum_{t=1}^{n} e_t^2\right) \tag{8}$$

2.7 Uji Signifikansi

Uji signifikansi berfungsi untuk memeriksa apakah suatu parameter dalam sebuah model statistik mempengaruhi variabel dependen secara signifikan atau tidak. Dalam analisis deret waktu menggunakan model SARIMA, uji signifikansi digunakan untuk mengetahui apakah parameter-parameter yang diestimasi pada model SARIMA memiliki pengaruh yang signifikan terhadap data deret waktu yang dianalisis. Uji T merupakan salah satu teknik statistik paling populer yang digunakan untuk menguji signifikansi (Mishra *et al*., 2019). Persamaan 9 sampai 12 berikut merupakan uji statistik T.

$$t_p = \frac{\hat{\phi}_p}{SE(\hat{\phi}_p)} \tag{9}$$

$$t_q = \frac{\hat{\theta}_q}{SE(\hat{\theta}_q)} \tag{10}$$

$$t_P = \frac{\hat{\Phi}_P}{SE(\hat{\Phi}_P)} \tag{11}$$

$$t_Q = \frac{\hat{\Theta}_Q}{SE(\hat{\Theta}_Q)} \tag{12}$$

2.8 Uji Diagnostik

e-ISSN: 2580-3239
p-ISSN: 2085-7993
Volume 22 No. 02 November 2023

Uji diagnostik model dilakukan untuk memeriksa apakah residual model bersifat *white noise* atau tidak. Dalam uji diagnostik dilakukan uji normalitas dan uji autokorelasi.

1. Uji Normalitas

   Uji normalitas bertujuan untuk menguji apakah model yang akan dibangun memiliki distribusi normal atau mendekati normal antara variabel dependen, variabel independen, atau keduanya (Gujarati *et al*., 2006). Uji Jarque-Berra merupakan salah satu uji normalitas berdasarkan pengukuran perbedaan skewness dan kurtosis pada data. Berikut pengujian Jarque-Berra (Abdellatif *et al*., 2018).

   $$JB = \frac{n}{6}\left(S^2 + \frac{(K-3)^2}{4}\right) \qquad (13)$$

   dengan S merupakan *skewness* dan K merupakan Kurtosis.

2. Uji Autokorelasi

   Uji Ljung-Box (Box-Pierce) merupakan uji yang dapat digunakan untuk melihat autokorelasi data deret waktu. Uji Ljung-Box umumnya dilakukan dengan bantuan bahasa pemrograman statistik seperti R atau Python. Secara umum, uji Box-Ljung didefinisikan sebagai (Ljung and Box, 1978).

   $$Q = n(n+2)\sum_{k=1}^{K_{maks}} \frac{\hat{\rho}_k^2}{n-k} \qquad (14)$$

   dengan K_{maks} merupakan lag maksimum dan p, q merupakan orde dari ARMA(p, q).

2.9 Pemilihan Model Terbaik

Akaike Information Criterion (AIC) dapat digunakan untuk memilih model yang paling cocok dengan data. AIC merupakan ukuran dari kecocokan relatif model statistik. Nilai AIC hanya memberikan sarana untuk pemilihan model, namun tidak dapat menjelaskan seberapa cocok model dengan data. AIC didefinisikan pada persamaan 15 berikut.

$$\text{AIC} = 2m - 2\ln(L) \qquad (15)$$

dimana m adalah banyak parameter dalam model statistik dan L adalah nilai maksimum dari fungsi *likelihood* model. AIC diterapkan dalam pemilihan model dimana model dengan nilai AIC paling kecil dipilih sebagai kandidat model terbaik (Akaike, 1998).

2.10 Akurasi Peramalan

Mean Absolute Percentage Error (MAPE) merupakan metrik yang menentukan akurasi sebuah metode peramalan. Metrik ini mewakili persentase dari rata-rata absolut dari setiap periode dalam himpunan data untuk menghitung seberapa akurat jumlah yang diperkirakan dibandingkan dengan jumlah aktual. MAPE dapat dihitung dengan persamaan 16 sebagai berikut (Makridakis et al., 1999).

$$\text{MAPE} = \frac{\sum_{t=1}^{n}\left|\frac{(Z_t - \hat{Z}_t)}{Z_t}\right|}{n} \times 100\% \qquad (16)$$

dengan n menyatakan banyaknya data, $\hat{Z}_t$ menyatakan nilai taksiran peramalan, dan Z_t adalah nilai aktual dengan indeks t.

Semakin kecil nilai MAPE yang dihasilkan, maka semakin akurat peramalannya. Ukuran MAPE merupakan yang paling sering digunakan dan merupakan ukuran ketepatan relatif berupa persentase penyimpangan hasil peramalan. Menurut Lewis (1982), terdapat kategori standar MAPE untuk evaluasi peramalan. Kategori tersebut dapat dilihat pada Tabel 1.

Tabel 1 Kategori standar MAPE

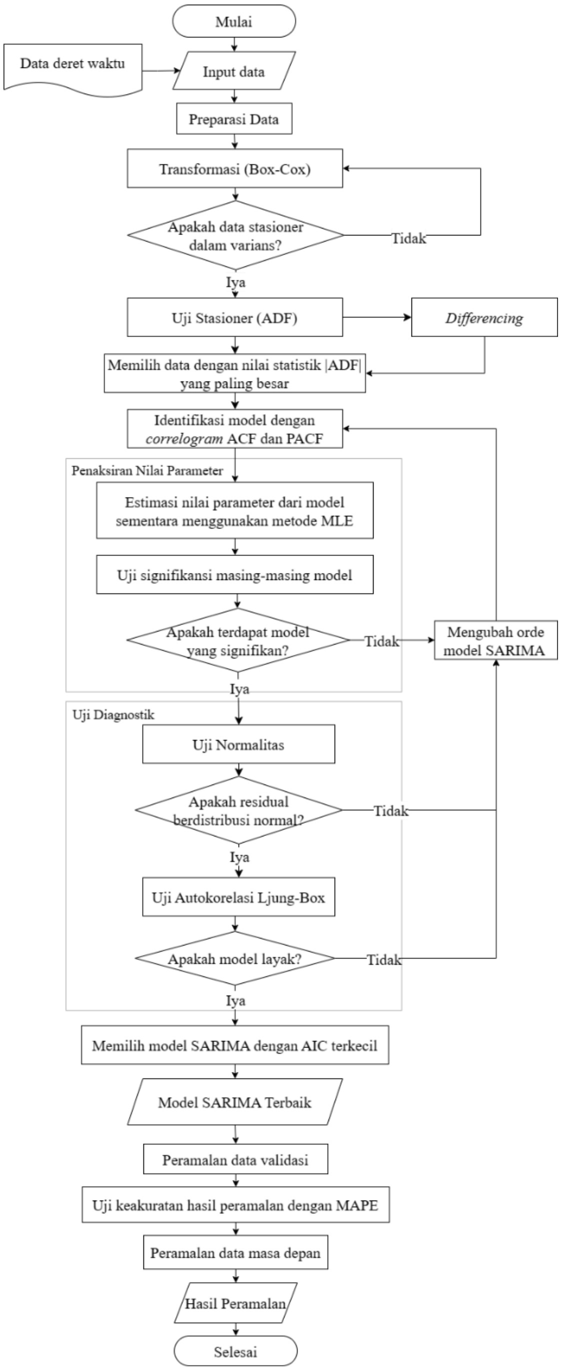
Rentang Nilai MAPE	Kategori Tingkat Akurasi
$\leq 10\%$	Sangat Baik
$10\% < \text{MAPE} \leq 20\%$	Baik
$20\% < \text{MAPE} \leq 50\%$	Layak
50%	Buruk

Sumber: Lewis (1982)

e-ISSN: 2580-3239
p-ISSN: 2085-7993
Volume 22 No. 02 November 2023

3 METODE PENELITIAN

Metode yang digunakan dalam penelitian ini adalah metode SARIMA dengan diagram alir penelitian ditunjukkan oleh Gambar 2.



Gambar 2 Diagram alir penelitian

4 HASIL DAN PEMBAHASAN

4.1 Preparasi Data

Sebelum dilakukan peramalan menggunakan metode SARIMA, perlu dilakukan preparasi data terlebih dahulu. Pada penelitian ini dilakukan pembersihan data dan pembagian data.

1. Memeriksa kelengkapan data keseluruhan menggunakan fungsi Python. *Syntax* untuk memeriksa kelengkapan data adalah sebagai berikut:

```
df.isna().sum()
```

   Jika *output* yang dihasilkan tidak sama dengan nol, berarti ada data yang hilang. Selanjutnya, data yang hilang diubah nilainya menggunakan metode interpolasi linear pada persamaan 1. *Syntax* pengisian data hilang adalah sebagai berikut:

```
df_interpolation = df.copy(deep=True)
df_interpolation.interpolate(limit_dire
ction="both",inplace=True)
```

2. Data yang digunakan dibagi menjadi 2 dataset yang berbeda, yaitu data latih dan data validasi. *Syntax* pembagian data adalah sebagai berikut:

```
train_end = datetime(yyyy,mm,dd)
val_end = datetime(yyyy,mm,dd)
train_data =
df_interpolation[:train_end]
val_data = df_interpolation[train_end +
timedelta(days=1):val_end]
```

4.2 Uji Stasioneritas

Pembuatan model dilakukan ketika data sudah lengkap dan stasioner. Uji Stasioneritas pada Python dilakukan dengan bantuan *library* Scipy dengan modul boxcox untuk transformasi Box-Cox dan *library* Statsmodels dengan modul adfuller untuk uji ADF. *Syntax* untuk uji stasioneritas dalam varians dan rata-rata adalah sebagai berikut:

e-ISSN: 2580-3239
p-ISSN: 2085-7993
Volume 22 No. 02 November 2023

```
df_transformed, lambda_value =
boxcox(train_data)
```

Jika *output* lambda yang dihasilkan tidak sama dengan 1, maka langkah selanjutnya dapat menggunakan data hasil transformasi yaitu df_transformed.

```
result = adfuller(df_transformed)
print('ADF Statistic: %f' % result[0])
print('p-value: %.3f' % result[1])
print('Critical Values:')
for key, value in result[4].items():
      print('\t%s: %.3f' % (key, value))
```

Jika hasil uji ADF menunjukkan bahwa data tidak stasioner dalam rata-rata, maka perlu dilakukan proses *differencing*. *Syntax* untuk *differencing* adalah sebagai berikut:

```
diff_df=diff(df_transformed, k_diff=1,
k_seasonal_diff=True, seasonal_periods=30)
```

4.3 Identifikasi Model SARIMA

Identifikasi model dapat dilakukan dengan melihat plot ACF dan PACF untuk mencari nilai dari orde p, q, P, dan Q pada model SARIMA yang sesuai. Pembentukan plot ACF dan PACF dilakukan dengan menggunakan *library* Statsmodel pada Python. *Syntax* pembentukan kedua plot adalah sebagai berikut:

```
acf_plot=sm.graphics.tsa.plot_acf(diff_df.
squeeze(), lags=10)
pacf_plot=sm.graphics.tsa.plot_pacf(diff_d
f.squeeze(), lags=10)
plt.show()
```

4.4 Pembuatan Model SARIMA

Setelah dilakukan identifikasi model dan didapat nilai orde yang sesuai, data sudah siap untuk dimodelkan dan diramalkan. Berikut *syntax* pembuatan model SARIMA dengan menggunakan *library* Statsmodel:

```
order = (1,1,1)
seasonal_order = (1, 1, 1, 12)
model =
sm.tsa.statespace.SARIMAX(train_data,
order=order,
seasonal_order=seasonal_order)
model_fit = model.fit()

print(model_fit.summary())
```

4.5 Peramalan data validasi

Berikut *syntax* untuk meramalkan data validasi menggunakan model SARIMA:

```
predictions =
model_fit.forecast(len(val_data))
predictions.plot(figsize=(15, 5),
color=color_pal[0], title='Predictions')
```

Syntax untuk membandingkan data aktual dan data prediksi adalah sebagai berikut:

```
ax = val_data.plot(figsize=(15, 5))
predictions.plot(ax=ax, style='-',
xlabel='Tanggal', ylabel='Satuan')
plt.legend(['Data Aktual', 'Data
Taksiran'])
ax.set_title('')
plt.show()
```

4.6 Uji Akurasi MAPE

Setelah data taksiran diperoleh, dilakukan perhitungan MAPE untuk mengetahui seberapa akurat model dalam meramalkan data. Berikut *syntax* untuk perhitungan MAPE:

```
def MAPE(Y_actual,Y_Predicted:
    mape = np.mean(np.abs((Y_actual -
Y_Predicted)/Y_actual))*100
    return mape
LR_MAPE= MAPE(val_data,predictions)
print("MAPE: ",LR_MAPE)
```

e-ISSN: 2580-3239
p-ISSN: 2085-7993
Volume 22 No. 02 November 2023

5 SIMPULAN

Berdasarkan tahapan penelitian yang telah dilakukan, dapat disimpulkan bahwa peramalan data univariat dengan unsur musiman dapat dilakukan menggunakan metode SARIMA dengan bahasa pemrograman Python. Untuk penelitian selanjutnya diharapkan dapat melakukan analisis terhadap faktor-faktor yang memengaruhi variabel dependen. Faktor-faktor interverensi tersebut dapat diterapkan menggunakan metode interverensi SARIMA atau SARIMAX (SARIMA dengan variabel eksternal *exogenous*) yang memungkinkan penggunaan lebih dari satu variabel eksternal sehingga peramalan dapat menjadi lebih akurat.

DAFTAR PUSTAKA

Abdellatif, D., El Moutaouakil, K. dan Satori, K. (2018) ‘Clustering and Jarque-Bera Normality Test to Face Recognition’, Procedia Computer Science, 127, pp. 246–255. Tersedia di: https://doi.org/https://doi.org/10.1016/j.procs.2018.01.120.

Akaike, H. (1998) ‘Information theory and an extension of the maximum likelihood principle’, Selected papers of hirotugu akaike, pp. 199–213.

Box, G.E.P. dan Cox, D.R. (1964) ‘An analysis of transformations’, Journal of the Royal Statistical Society: Series B (Methodological), 26(2), pp. 211–243.

Box, G.E.P., Jenkins, G.M., Reinsel, G.C. dan Ljung, G.M. (2015) Time series analysis: forecasting and control. John Wiley & Sons.

Chen, P., Niu, A., Liu, D., Jiang, W. dan Ma, B. (2018) ‘Time series forecasting of temperatures using SARIMA: An example from Nanjing’, in IOP conference series: materials science and engineering, p. 52024.

Farhan, J. dan Ong, G.P. (2018) ‘Forecasting seasonal container throughput at international ports using SARIMA models’, Maritime Economics & Logistics, 20, pp. 131–148.

Franses, P.H., Dijk, D. van dan Opschoor, A. (2014) Time Series Models for Business and Economic Forecasting. 2nd edn. Cambridge University Press. Tersedia di: https://doi.org/10.1017/CBO9781139049894. [7] Weerakody et al., 2021

Harrison, E., Emeka, I. dan Cecilia, E. (2020) ‘African Journal of Mathematics and Statistics Studies: Time Series Models of Crude Oil Production and Export in Nigeria’, 3(1), pp. 1–24. Tersedia di: www.abjournals.org.

Heizer, J. (2011) ‘Operations Management 10th ed-Jay Heizer, Barry Render’. Pearson.

Hyndman, R.J. dan Athanasopoulos, G. (2018) Forecasting: principles and practice. Melbourne: OTexts.

Lewis, C.D. (1982) Industrial and business forecasting methods: A practical guide to exponential smoothing and curve fitting. London: Butterworth-Heinemann.

Ljung, G.M. dan Box, G.E.P. (1978) ‘On a measure of lack of fit in time series models’, Biometrika, 65(2), pp. 297–303.

Makridakis, S., Wheelwright, S.C. dan McGee, V.E. (1999) Metode dan aplikasi peramalan. Jakarta: Erlangga.

Mishra, P., Singh, U., Pandey, C.M., Mishra, P. dan Pandey, G. (2019) ‘Application of student’s t-test, analysis of variance,

e-ISSN: 2580-3239
p-ISSN: 2085-7993
Volume 22 No. 02 November 2023

and covariance', Annals of cardiac anaesthesia, 22(4), p. 407.

Pongdatu, G.A.N. dan Putra, Y.H. (2018) 'Seasonal time series forecasting using SARIMA and Holt Winter's exponential smoothing', pada IOP Conference Series: Materials Science and Engineering, p. 12153.

Rao, S. (1985) 'An empirical comparison of sales forecasting models', Journal of Product Innovation Management: An International Publication of The Product Development & Management Association, 2(4), pp. 232–242.

Saaban, A., Zainudin, L. dan Bakar, M.N.A. (2015) 'Evaluation of linear interpolation method for missing value on solar radiation dataset in Perlis', pada AIP Conference Proceedings, p. 50024.

Weerakody, P.B., Wong, K.W., Wang, G. dan Ela, W. (2021) 'A review of irregular time series data handling with gated recurrent neural networks', Neurocomputing, 441, pp. 161–178.

Wei, W. (2006) Time Series Analysis: Univariate and Multivariate Methods, 2nd edition, 2006. Boston: Pearson Addison Wesley.