



Model *Hybrid* SARIMA-ELM Untuk Peramalan Jumlah Penumpang Kereta Api di Stasiun Bandung

Elisa Febriyani¹, Fitriani Agustina², dan Nar Herrhyanto^{3*}

^{1, 2, 3}Program Studi Matematika, Fakultas Pendidikan Matematika dan Ilmu Pengetahuan Alam, Universitas Pendidikan, Indonesia

*Correspondence: narherrhyanto@upi.edu

ABSTRAK

Penelitian ini bertujuan untuk mengetahui penerapan model *Hybrid* SARIMA-ELM pada data jumlah penumpang kereta api serta mengetahui tingkat akurasi peramalannya. Peramalan didasarkan pada data jumlah penumpang yang turun di Stasiun Bandung pada periode sebelum Covid-19 (Agustus 2014-Desember 2018) dan periode saat Covid-19 (Januari 2020-Desember 2021). Peramalan diawali dengan mencari model SARIMA terbaik, kemudian residual dari SARIMA dimodelkan dengan ELM. Hasil peramalan model *Hybrid* SARIMA-ELM diukur keakuratannya menggunakan MAPE. Model SARIMA terbaik, yaitu SARIMA $(0,1,1)(0,1,0)^{12}$ untuk fase sebelum Covid-19 dan model ARIMA $(2,0,0)$ untuk fase saat Covid-19. Dengan perbandingan data 80:20, 3 fitur *input*, 1 *neuron hidden layer*, dan 1 *output* pada model ELM, diperoleh nilai MAPE sebesar 5,27% yang berarti peramalan berakurasi tinggi untuk model sebelum Covid-19 dan MAPE sebesar 13,43% yang berarti peramalan baik untuk model saat Covid-19.

© 2024 Kantor Jurnal dan Publikasi UPI

ABSTRACT

This study aims to determine the application of the Hybrid SARIMA-ELM model to train passenger data and determine the level of forecasting accuracy. Forecasting is based on data on the number of passengers disembarking at Bandung Station in the period before Covid-19 (August 2014-December 2018) and the period during Covid-19 (January 2020-December 2021). Forecasting begins with finding the best SARIMA model, then the residuals from SARIMA are modeled with ELM. Hybrid SARIMA-ELM model forecasting results are measured for accuracy using MAPE. The best SARIMA model is SARIMA $(0,1,1)(0,1,0)^{12}$ for the best phase before Covid-19 and ARIMA $(2,0,0)$ model for the phase during Covid-19. With a data ratio of 80:20, 3 input features, 1 hidden layer neuron, and 1 output in the ELM model, a MAPE value of 5.27% is obtained, which means that forecasting is highly accurate for the model before Covid-19 and a MAPE of 13.43%, which means good forecasting for the model during Covid-19.

© 2024 Kantor Jurnal dan Publikasi UPI

INFORMASI ARTIKEL

Sejarah Artikel:

Diterima 14 Oktober 2024
Direvisi 29 Oktober 2024
Disetujui 1 November 2024
Tersedia online 9 November 2024
Dipublikasikan 9 November 2024

Kata Kunci:

ELM,
Jumlah Penumpang,,
Peramalan,
SARIMA.

Keywords:

ELM,
Forecasting,
Number of Passengers,
SARIMA.

1. PENDAHULUAN

Salah satu elemen penting dalam suatu negara adalah transportasi, karena transportasi merupakan penunjang dalam mendukung pertumbuhan ekonomi, pengembangan wilayah, dan pemersatu wilayah Negara Kesatuan Republik Indonesia untuk mewujudkan wawasan nusantara, serta memperkuat ketahanan nasional (Undang-Undang Republik Indonesia Nomor 23 Tahun 2007 Tentang Perkeretaapian, 2007). Kereta api menjadi pilihan masyarakat Indonesia, karena selain cepat, juga dapat terhindar dari kemacetan yang merupakan permasalahan yang sampai saat ini terjadi terutama di kota-kota wisata seperti Bandung. Jumlah penumpang kereta api selalu bertambah setiap tahunnya. Namun pada tahun 2020, berdasarkan data BPS terdapat pola menurun yang sangat drastis hingga 83,55%. Adapun peningkatan jumlah penumpang kereta api terjadi pada bulan tertentu, seperti bulan-bulan yang termasuk hari libur sekolah, seperti libur Hari Raya Idul Fitri, maupun Natal dan tahun baru. Karena adanya pola yang sama setiap tahunnya, maka data tersebut dikatakan memiliki pola musiman. Untuk menaksir lonjakan jumlah penumpang kereta api, diperlukan adanya peramalan untuk memprediksi jumlah penumpang kereta api pada periode selanjutnya agar PT. Kereta Api Indonesia (PERSERO) tetap mampu memberikan pelayanan terbaik demi kenyamanan penumpang dengan cara meningkatkan sarana dan prasarannya. Dengan adanya peramalan diharapkan ramalan yang dibuat bisa meminimumkan kesalahan peramalan, artinya perbedaan antara kenyataan dengan ramalan tidak jauh berbeda (Mulyani, et al., 2021).

Metode peramalan *time series* yang dapat digunakan untuk meramalkan data dengan pola musiman adalah *Seasonal Autoregressive Integrated Moving Average* (SARIMA) (Chen, et al., 2018). Model SARIMA mengasumsikan model yang linear, sehingga menyebabkan model SARIMA tidak menangkap pola-pola non-linear yang umumnya terdapat pada *time series*. Oleh karena itu, diperlukan sebuah model yang dapat menangkap pola-pola non-linear. Model *Extreme Learning Machine* (ELM) baik digunakan dalam penelitian untuk data yang berpola non-linear (Simbolon, et al., 2021).

Penelitian mengenai peramalan menggunakan metode SARIMA telah dilakukan oleh Utomo dan Fanani (2020) diperoleh nilai MAPE sebesar 6,26%. Lalu, Izati, et al. (2019) menggunakan FFNN dengan metode ELM untuk prediksi harga emas dan menghasilkan peramalan yang mendekati nilai aktualnya. Kemudian perbandingan metode SARIMA dan ELM dilakukan oleh Bayu dan Edy (2021), hasil menunjukkan model SARIMA cocok digunakan pada peramalan jumlah wisatawan mancanegara ke Bali dibandingkan dengan model ELM. Model SARIMA dan ELM memiliki kelebihan dan kekurangannya masing-masing. Sebuah model gabungan (*hybrid*) dirancang untuk saling menutupi kekurangan dari tiap model dan memanfaatkan kelebihan dari masing-masing model untuk meningkatkan akurasi dari hasil peramalannya (Rowan, et al., 2022).

Berdasarkan uraian di atas, maka penelitian ini akan mengkaji lebih dalam mengenai model *hybrid* SARIMA-ELM untuk peramalan jumlah penumpang kereta api di Stasiun Bandung.

2. METODE

2.1 Data Penelitian

Data yang digunakan Data yang digunakan adalah data sekunder berupa data bulanan jumlah penumpang kereta api yang turun di Stasiun Bandung yang diperoleh dari PT. KAI

melalui aplikasi PPID PT KAI. Data dibagi menjadi dua periode, yaitu periode sebelum Covid-19 (Agustus 2014-Desember 2018) dan periode saat Covid-19 (Januari 2020-Desember 2022).

2.2 Langkah Penelitian

1. Analisis dengan Metode SARIMA

Analisis pada penelitian ini dilakukan dengan langkah-langkah sebagai berikut.

- a. Membuat plot *time series*
- b. Mengidentifikasi kestasioneran pada data.
 - Data dikatakan stasioner dalam varians, apabila nilai *rounded value* dari plot Box-Cox bernilai satu. Jika data belum stasioner dalam variansnya, maka dilakukan transformasi Box-Cox pada data.
 - Data dikatakan stasioner dalam rata-rata, apabila nilai *p-value* pada uji *Augmented Dickey-Fuller* (ADF) lebih kecil dari taraf signifikansi ($\alpha=0,05$)
- c. Menentukan model sementara dengan melihat grafik ACF dan PACF dari data yang telah stasioner dalam varian dan rata-rata.
- d. Melakukan estimasi parameter dan uji signifikansi parameter pada model yang diperoleh. Hipotesis dari uji signifikansi parameter adalah sebagai berikut

Hipotesis :

H_0 : Parameter model tidak signifikan

H_1 : Parameter model signifikan

Kriteria Pengujian :

Tolak H_0 jika nilai *p-value* < α dengan $\alpha = 0,05$.

- e. Menguji kelayakan model, bertujuan untuk membuktikan bahwa model layak dan merupakan model terbaik yang digunakan untuk peramalan. Model dikatakan layak apabila residual telah memenuhi asumsi white noise dan berdistribusi normal.

- Uji *white noise*

Hipotesis :

H_0 : Residual memenuhi asumsi white noise

H_1 : Residual tidak memenuhi asumsi white noise

Kriteria Pengujian :

Tolak H_0 jika *p-value* < α dengan $\alpha = 0,05$.

- Uji normalitas

Hipotesis :

H_0 : Residual berdistribusi normal

H_1 : Residual berdistribusi tidak normal

Kriteria Pengujian :

Tolak H_0 jika *p-value* < α dengan $\alpha = 0,05$.

- f. Memilih model terbaik dengan menggunakan nilai AIC terkecil.
- g. Melakukan peramalan

2. Analisis dengan metode ELM

Variabel *Input* yang digunakan pada metode ELM ini ialah residual dari hasil peramalan pada metode SARIMA pada tahap sebelumnya, sehingga model akhir yang diperoleh adalah model *Hybrid* SARIMA-ELM. Langkah-langkahnya adalah sebagai berikut.

- a. Melakukan normalisasi pada data sesuai dengan fungsi aktivasi yang digunakan. Pada penelitian ini fungsi aktivasi yang digunakan adalah *sigmoid biner*, dimana outputnya berupa bilangan riil antara 0 dan 1, sehingga persamaannya:

$$x'_t = \frac{0,8 (x_t - (x))}{((x) - (x))} + 0,1 \tag{1}$$

- b. Menentukan data *training* dan data *testing* dengan perbandingan 80:20. Data *training* diolah menjadi data latih dengan 3 fitur disertai target latih yang nantinya akan digunakan pada proses *training*, sedangkan data *testing* diolah menjadi data uji dengan 3 fitur disertai target uji yang nantinya digunakan pada proses *testing*.

Tabel 1. Data Latih dan Target Latih

| No | Data Latih | Target Latih |
|-----|---------------------------------------|-----------------------|
| 1. | Data pada bulan ke-1 s.d bulan ke-3 | Data pada bulan ke-4 |
| 2. | Data pada bulan ke-2 s.d bulan ke-4 | Data pada bulan ke-5 |
| ... | ... | ... |
| 39. | Data pada bulan ke-39 s.d bulan ke-41 | Data pada bulan ke-42 |

Tabel 2. Data Uji dan Target Uji

| No | Data Uji | Target Uji |
|-----|---------------------------------------|-----------------------|
| 1. | Data pada bulan ke-43 s.d bulan ke-45 | Data pada bulan ke-46 |
| 2. | Data pada bulan ke-44 s.d bulan ke-46 | Data pada bulan ke-47 |
| ... | ... | ... |
| 11. | Data pada bulan ke-50 s.d bulan ke-52 | Data pada bulan ke-53 |

c. Proses *Training*

Proses *training* bertujuan untuk mendapatkan nilai *output weigh* (β) menggunakan data latih dan target latih, dengan langkah sebagai berikut (Huang et al., 2006).

- 1) Langkah pertama yaitu melakukan inisialisasi *input weigh* (bobot awal) dan bias secara acak dengan nilai antara 0 dan 1. Bobot berupa matriks W_{mn} , dalam bentuk *array* ukuran $m \times n$ dengan m merupakan jumlah *neuron* pada *hidden layer* dan n merupakan jumlah *input neuron*. Sedangkan bias berukuran $1 \times$ jumlah *neuron* pada *hidden layer* (Sugianto, et al., 2018).
- 2) Menghitung matriks H_{init} menggunakan persamaan:

$$H_{init\ train} = X_{train}W^T + b \tag{2}$$

dimana X merupakan data latih, W merupakan matriks bobot, dan b merupakan bias.

Selanjutnya menghitung fungsi aktivasi dari nilai H_{init} . Nilai *output hidden layer* yang dihitung menggunakan fungsi aktivasi dilambangkan dengan H . Persamaan fungsi aktivasi yang digunakan adalah *sigmoid biner* sebagai berikut.

$$H = \frac{1}{1 + \exp(-H_{init\ train})} \tag{3}$$

- 3) Menghitung matriks *Moore-Penrose Generalized Invers* dengan persamaan:

$$H^+ = (H^T H)^{-1} H^T \tag{4}$$

- 4) Menghitung nilai *output weigh* (bobot keluaran) dengan persamaan:

$$\beta = H^+ Y \tag{5}$$

dimana Y merupakan data target latih.

d. Proses *Testing*

Pada proses *testing* nilai *input weigh*, bias, serta *output weigh* (bobot keluaran) yang digunakan yaitu nilai yang didapat dari proses *training*, hanya saja data yang digunakan, yaitu data uji dan target uji. Adapun langkah-langkah yang perlu dilakukan adalah sebagai berikut.

- 1) Menghitung matriks H_{init} dan matriks *output hidden layer H* menggunakan persamaan yang sama pada proses *training* namun menggunakan data *testing*.
- 2) Menghitung hasil peramalan melalui perkalian matriks *output hidden layer* dengan *output weigh* (bobot keluaran) dari proses *training* menggunakan persamaan:

$$\hat{Y} = H\beta \tag{6}$$

- 3) Melakukan denormalisasi data dari hasil peramalan untuk mendapatkan nilai aktual dengan persamaan:

$$\hat{x}_t = \frac{(\hat{x}'_t - 0,1)((x) - (x))}{0,8} + \min(x) \tag{7}$$

3. Memperoleh hasil peramalan metode *Hybrid SARIMA-ELM*.
4. Menghitung akurasi peramalan menggunakan nilai MAPE dengan rumus sebagai berikut.

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{Z_t - \hat{Z}_t}{Z_t} \right| \times 100\% \tag{8}$$

dengan Z_t adalah nilai aktual, \hat{Z}_t adalah nilai peramalan, dan n adalah banyak data.

Adapun kriteria nilai MAPE, yaitu:

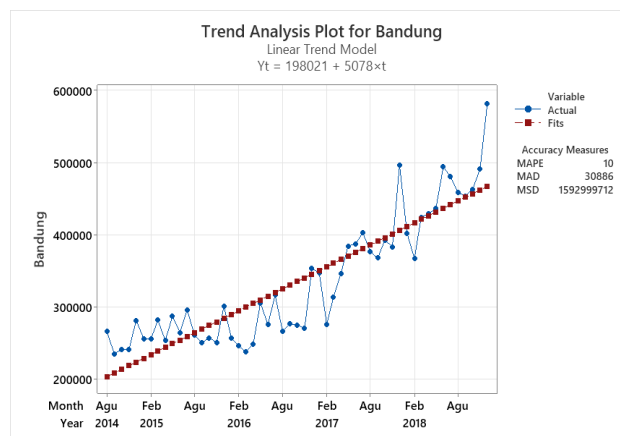
- a. < 10% = Berakurasi tinggi (*Highly Accurate*)
- b. 11% - 20% = Peramalan baik (*Good Forecast*)
- c. 21% - 50% = Peramalan dengan alasan (*Reasonable Forecast*)
- d. > 50% = Peramalan tidak akurat (*Inaccurate Forecast*)

3. HASIL DAN PEMBAHASAN

Data pada penelitian ini dibagi menjadi dua kasus. Kasus yang pertama, yaitu jumlah penumpang kereta api sebelum pandemi Covid-19 dan kasus yang kedua, yaitu jumlah penumpang kereta api saat pandemi Covid-19.

3.1 Pemodelan Jumlah Penumpang Kereta Api Fase sebelum Covid-19

Data yang digunakan adalah data jumlah penumpang kereta api yang turun di Stasiun Bandung pada periode Agustus 2014 sampai dengan Desember 2018 yang diperoleh dari PT. Kereta Api (PERSERO). Berikut merupakan plot *time series* data jumlah penumpang yang turun di Stasiun Bandung.

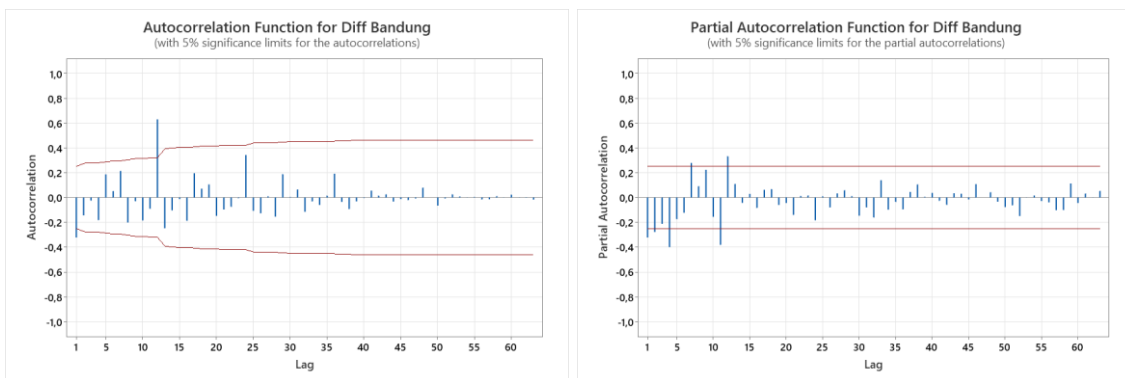


Gambar 1. *Trend Analysis Plot* Jumlah Penumpang yang Turun di Stasiun Bandung Sebelum Covid-19

Berdasarkan Gambar 1, terlihat bahwa data jumlah penumpang kereta api yang turun di Stasiun Bandung mengalami trend yang cenderung naik. Hal tersebut menandakan data tidak

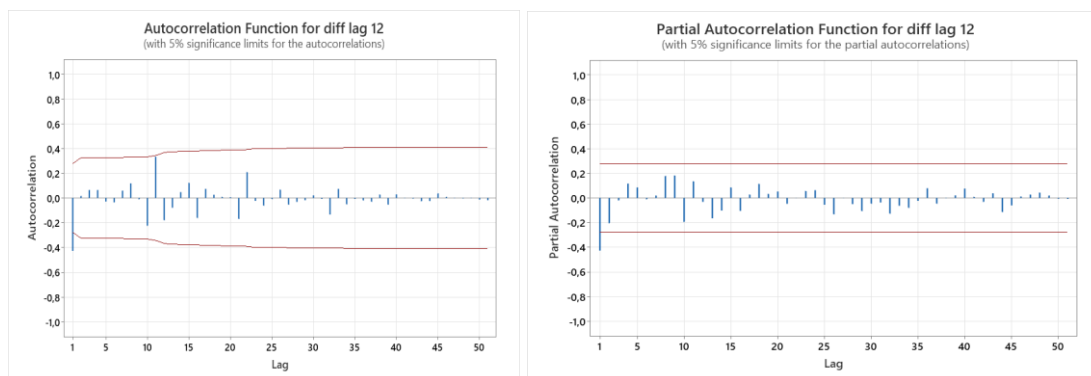
stasioner dalam variansi maupun rata-rata sehingga perlu dilakukan transformasi Box-Cox dan differencing pada data. Selain itu, terdapat pola musiman dilihat dari plot yang naik turun dalam jangka waktu yang tetap, dimana dari bulan November ke Desember mengalami kenaikan dan dari Desember ke Januari mengalami penurunan.

Selanjutnya akan dibentuk model sementara berdasarkan lag yang *cuts off* maupun *dies down* pada diagram ACF dan PACF. Berikut merupakan plot ACF dan PACF differencing pertama data jumlah penumpang kereta api di Stasiun Bandung yang telah di transformasi.



Gambar 2. Plot ACF dan PACF Data Differencing 1 Non-musiman

Pada Gambar 2, ACF *cuts off* pada lag pertama sehingga diperoleh model MA(1), sedangkan untuk PACF terlihat *dies down* untuk bagian non-musiman, sehingga dapat dimodelkan untuk model non-musiman, yaitu (0,1,1). Selanjutnya untuk bagian musimannya terlihat pada lag ke-12 signifikan, sehingga perlu adanya differencing pada lag 12. Diagram ACF dan PACF setelah dilakukan differencing pada lag 12 ditampilkan pada Gambar 3.



Gambar 3. Plot ACF dan PACF Data Differencing 1 Musiman

Berdasarkan Gambar 3, ACF dan PACF *cuts off* setelah lag 1, sehingga dapat diprediksi model MA(1)¹² dan AR(1)¹² untuk bagian musimannya. Berdasarkan ACF dan PACF tersebut dapat dimodelkan untuk model bagian musimannya, yaitu (1,1,1)¹². Sehingga diperoleh model SARIMA sementara, yaitu SARIMA (0,1,1)(1,1,1)¹² dan model lain yang mungkin terbentuk, yaitu SARIMA (0,1,1)(0,1,1)¹², SARIMA (0,1,1)(1,1,0)¹² dan SARIMA (0,1,1)(0,1,0)¹²

Selanjutnya akan dilakukan uji signifikansi parameter dari model sementara yang telah diperoleh. Adapun hasil pengujian parameter untuk model-model tersebut terdapat pada Tabel 3.

Tabel 3. Penaksiran Parameter Model Sementara

| Model Sementara | Parameter | Estimasi Parameter | P-Value | Ket. |
|--|-----------|--------------------|-----------|------------------|
| SARIMA (0,1,1)(1,1,1) ¹² | MA 1 | -0,5463 | 4,266e-07 | Signifikan |
| | SAR 1 | 0,7468 | 0,00671 | Signifikan |
| | SMA 1 | -0,9988 | 0,37702 | Tidak signifikan |
| SARIMA (0,1,1)(0,1,1) ¹² | MA 1 | -0,5532 | 1,459e-07 | Signifikan |
| | SMA 12 | -0,1246 | 0,4909 | Tidak signifikan |
| SARIMA (0,1,1)(1,1,0) ¹² | MA 1 | -0,5519 | 1,771e-07 | Signifikan |
| | SAR 12 | -0,1420 | 0,4634 | Tidak signifikan |
| SARIMA (0,1,1)(0,1,0) ¹² | MA1 | -0,5537 | 1,476e-07 | Signifikan |

Berdasarkan Tabel 3, diketahui parameter model SARIMA (0,1,1)(0,1,0)¹² memiliki nilai *p-value* yang lebih kecil dari taraf signifikan ($\alpha=0,05$) sehingga parameter dapat digunakan pada model. Oleh karena itu, model SARIMA (0,1,1)(0,1,0)¹² lolos ke tahap selanjutnya.

Tahap selanjutnya yaitu uji kelayakan model. Model dikatakan layak, apabila residual telah memenuhi asumsi *white noise* dan berdistribusi normal. Adapun hasil uji tersebut ditampilkan pada tabel di bawah ini.

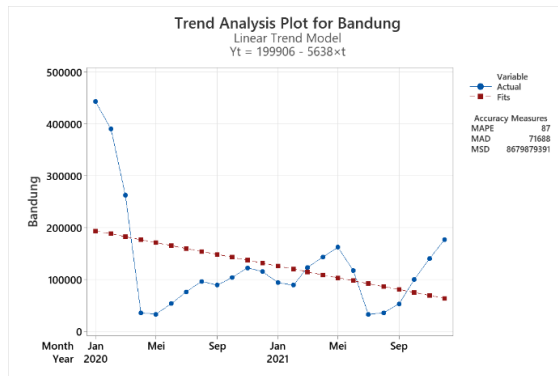
Tabel 4. Hasil Uji Kelayakan Model

| Uji White Noise | | | | Uji Normalitas Residual | |
|--|-----|------------|---------|-------------------------|---------|
| Model | Lag | Chi-Square | P-Value | Kolmogorov-Smirnov | P-Value |
| SARIMA (0,1,1)(0,1,0) ¹² | 6 | 3,8015 | 0,7035 | 0,078 | > 0,150 |
| | 12 | 16,772 | 0,1584 | | |
| | 18 | 26,062 | 0,09836 | | |
| | 24 | 31,589 | 0,1375 | | |
| | 36 | 34,807 | 0,5253 | | |
| | 48 | 35,194 | 0,9156 | | |

Berdasarkan Tabel 4, nilai-nilai autokorelasi residual memiliki *p-value* yang lebih besar dari taraf signifikan ($\alpha=0,05$), sehingga model dikatakan memenuhi asumsi *white noise*. Kemudian nilai *p-value* pada uji Kolmogorov-Smirnov bernilai > 0,150 yang berarti lebih besar dari taraf signifikan ($\alpha=0,05$), sehingga dapat dikatakan model SARIMA (0,1,1)(0,1,0)¹² telah berdistribusi normal dan layak dijadikan model terbaik untuk selanjutnya dilakukan peramalan.

3.2 Pemodelan Jumlah Penumpang Kereta Api Fase Covid-19

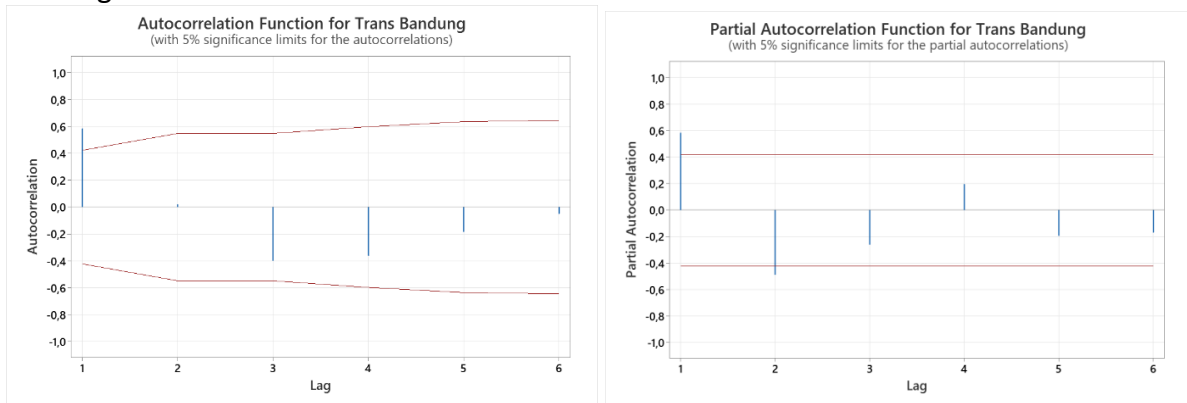
Data yang digunakan adalah data jumlah penumpang kereta api yang turun di Stasiun Bandung pada periode Januari 2020 sampai dengan Desember 2021. Berikut merupakan plot *time series* data jumlah penumpang yang turun di Stasiun Bandung Saat Covid-19.



Gambar 4. Trend Analysis Plot Jumlah Penumpang yang Turun di Stasiun Bandung

Pada Gambar 4, terlihat bahwa data jumlah penumpang kereta api yang turun di Stasiun Bandung cenderung mengalami trend turun. Selain itu, tidak terdapat unsur musiman. Hal tersebut dikarenakan adanya kebijakan pemerintah terkait larangan bepergian ke luar kota dan juga pembatasan kuota penumpang transportasi umum dalam rangka upaya pencegahan penyebaran Covid-19.

Gambar 5 merupakan plot ACF dan PACF data jumlah penumpang kereta api di Stasiun Bandung.



Gambar 5. Plot ACF dan PACF Data Jumlah Penumpang Kereta Api di Stasiun Bandung Saat Covid-19

Pada Gambar 5, ACF *cuts off* pada lag pertama sehingga diperoleh model MA(1) dan untuk PACF terlihat *cuts off* pada lag kedua, sehingga diperoleh model AR(2) untuk bagian non musiman. Berdasarkan ACF dan PACF tersebut dapat dimodelkan untuk model non musiman, yaitu ARIMA (2,0,1). Adapun model sementara lain yang mungkin, yaitu ARIMA (2,0,0) dan ARIMA (0,0,1).

Selanjutnya akan dilakukan uji signifikansi parameter dari model sementara yang telah diperoleh. Adapun hasil pengujian parameter untuk model-model tersebut adalah sebagai berikut.

Tabel 5. Penaksiran Parameter Model Sementara

| Model | Parameter | Estimasi Parameter | P-Value | Ket. |
|---------------|-----------|--------------------|-----------|------------------|
| ARIMA (2,0,1) | AR 1 | 1,4379 | 2,958e-16 | Signifikan |
| | AR 2 | -0,8273 | 3,834e-06 | Signifikan |
| | MA1 | -0,109 | 0,6898 | Tidak signifikan |
| ARIMA (2,0,0) | AR 1 | 1,3868 | 2,2e-16 | Signifikan |
| | AR 2 | -0,7741 | 3,243e-06 | Signifikan |
| ARIMA (0,0,1) | MA 1 | 0,9458 | 0,0003254 | Signifikan |

Berdasarkan Tabel 5, parameter pada model ARIMA (2,0,0) dan ARIMA (0,0,1) memiliki nilai p -value yang lebih kecil dari taraf signifikan ($\alpha=0,05$) sehingga parameter dapat digunakan pada kedua model tersebut. Oleh karena itu, model ARIMA (2,0,0) dan ARIMA (0,0,1) lolos ke tahap selanjutnya.

Tahap selanjutnya, yaitu uji kelayakan model. Adapun hasil uji tersebut ditampilkan pada Tabel 6.

Tabel 6. Hasil Uji Kelayakan Model

| Model | Uji White Noise | | | Uji Normalitas Residual | |
|---------------|-----------------|------------|---------|-------------------------|---------|
| | Lag | Chi-Square | P-Value | Kolmogorov-Smirnov | P-Value |
| ARIMA (2,0,0) | 6 | 4,3634 | 0,6276 | 0,21833 | 0,1745 |
| | 12 | 6,5509 | 0,8858 | | |
| | 18 | 13,006 | 0,7912 | | |
| ARIMA (0,0,1) | 6 | 10,058 | 0,1222 | 0,14436 | 0,6471 |
| | 12 | 12,996 | 0,3694 | | |
| | 18 | 20,602 | 0,2999 | | |

Berdasarkan Tabel 6, diketahui nilai-nilai autokorelasi residual pada kedua model memiliki p -value yang lebih besar dari taraf signifikan ($\alpha=0,05$), sehingga kedua model dikatakan memenuhi asumsi *white noise*. Kemudian nilai p -value uji Kolmogorov-Smirnov pada kedua model lebih besar dari taraf signifikan ($\alpha=0,05$), sehingga dapat dikatakan model telah berdistribusi normal.

Karena terdapat 2 model ARIMA yang layak digunakan, maka akan dipilih model terbaik berdasarkan nilai AIC terkecil. Adapun nilai AIC dari kedua model disajikan pada Tabel 7.

Tabel 7. Nilai AIC

| Model | AIC |
|---------------|--------|
| ARIMA (2,0,0) | 594,18 |
| ARIMA (0,0,1) | 609,06 |

Berdasarkan Tabel 7, diketahui nilai AIC model ARIMA (2,0,0) lebih kecil daripada model ARIMA (0,0,1) sehingga model ARIMA (2,0,0) merupakan model terbaik untuk meramalkan data jumlah penumpang kereta api di Stasiun Bandung.

3.3 Pemodelan *Hybrid* Jumlah Penumpang Kereta Api Fase Sebelum Covid-19

1. Proses *Training*

Dengan bantuan program R diperoleh nilai bobot dan bias yang disajikan pada Tabel 8 dan Tabel 9.

Tabel 8. Matriks Bobot

| Matriks W | 1 | 2 | 3 |
|-----------|-----------|-----------|-----------|
| 1 | 0,2709123 | 0,7025975 | 0,8831847 |

Tabel 9. Matriks Bias

| Matriks Bias | 1 |
|--------------|-----------|
| 1 | 0,0514347 |

Selanjutnya dilakukan perhitungan nilai keluaran *hidden layer* dengan menggunakan fungsi H_{init} sesuai Persamaan 2, sehingga diperoleh matriks H_{init} yang disajikan pada Tabel 10.

Tabel 10. Matriks Fungsi *Hidden Layer*

| | |
|---------------------------|-----------|
| Matriks $H_{init\ train}$ | 1 |
| 1 | 1,0534705 |
| 2 | 1,0531886 |
| 3 | 1,0533076 |
| ... | ... |
| 39 | 1,3196810 |

Perhitungan menggunakan fungsi aktivasi sigmoid biner sesuai Persamaan 3, sehingga diperoleh matriks H (Tabel 11).

Tabel 11. Matriks Keluaran *Hidden Layer*

| | |
|-------------|-----------|
| Matriks H | 1 |
| 1 | 0,7414408 |
| 2 | 0,7413867 |
| 3 | 0,7414096 |
| ... | ... |
| 39 | 0,7891286 |

Lalu dihitung matriks *moore-penrose generalized invers* (H^+) menggunakan Persamaan 4, sehingga diperoleh hasil yang disajikan pada Tabel 12.

Tabel 12. Matriks *Moore-penrose Generalized Invers*

| | | | | | |
|---------------|------------|------------|------------|-----|------------|
| Matriks H^+ | 1 | 2 | 3 | ... | 39 |
| 1 | 0,03328297 | 0,03328054 | 0,03328157 | ... | 0,03542366 |

Selanjutnya dilakukan perkalian matriks *moore-penrose generalized invers* (H^+) dengan matriks target latih, sehingga diperoleh nilai *output weight* berikut.

Tabel 13. Matriks *Output Weight*

| | |
|-----------------|-----------|
| Matriks β | 1 |
| 1 | 0,7586177 |

2. Proses *Testing*

Dengan menggunakan data uji, diperoleh matriks H_{init} sebagai berikut.

Tabel 14. Matriks Fungsi *Hidden Layer*

| | |
|-----------------------------|-----------|
| Matriks $H_{init\ testing}$ | 1 |
| 1 | 1,0298946 |
| 2 | 0,7731349 |
| 3 | 1,0219145 |
| ... | ... |
| 8 | 1,1622545 |

Perhitungan menggunakan fungsi aktivasi *sigmoid biner* sesuai Persamaan 3, sehingga diperoleh matriks H berikut.

Tabel 15. Matriks Keluaran *Hidden Layer*

| | |
|-------------|-----------|
| Matriks H | 1 |
| 1 | 0,7368955 |
| 2 | 0,6841987 |
| 3 | 0,7353454 |
| ... | ... |
| 8 | 0,7617421 |

Setelah didapatkan hasil dari fungsi aktivasi, langkah selanjutnya dihitung nilai *output layer* dengan cara mengalikan matriks H dengan *output weight* diperoleh dari proses *training*. Tabel 16 merupakan hasil *output layer* berupa nilai peramalan.

Tabel 16. Hasil Peramalan

| Peramalan | 1 |
|-----------|-----------|
| 1 | 0,5590219 |
| 2 | 0,5190452 |
| 3 | 0,5578460 |
| ... | ... |
| 8 | 0,5778711 |

Hasil dari *output layer* akan diubah kedalam bentuk data asli dengan perhitungan denormalisasi berdasarkan Persamaan 7. Hasil dari proses denormalisasi data ditunjukkan pada Tabel 17.

Tabel 17. Hasil Denormalisasi Data Peramalan

| Denormalisasi | 1 |
|---------------|-------|
| 1 | 3078 |
| 2 | -3172 |
| 3 | 2894 |
| ... | ... |
| 8 | 6026 |

3.4 Pemodelan *Hybrid* Jumlah Penumpang Kereta Api Fase Covid-19

1. Proses *Training*

Dengan menggunakan program R diperoleh nilai bobot dan bias (Tabel 18).

Tabel 18. Matriks Bobot

| Matriks W | 1 | 2 | 3 |
|-----------|-----------|-----------|-----------|
| 1 | 0,8811066 | 0,0670966 | 0,6445244 |

Tabel 19. Matriks Bias

| Matriks Bias | 1 |
|--------------|-----------|
| 1 | 0,2860588 |

Selanjutnya dilakukan perhitungan H_{init} sehingga diperoleh matriks H_{init} (Tabel 20).

Tabel 20. Matriks Fungsi *Hidden Layer*

| Matriks H_{init} train | 1 |
|--------------------------|-----------|
| 1 | 1,3492638 |
| 2 | 0,7538770 |
| 3 | 1,2602938 |
| ... | ... |
| 16 | 0,7629591 |

Perhitungan menggunakan fungsi aktivasi sigmoid biner, sehingga diperoleh matriks H (Tabel 21).

Tabel 21. Matriks Keluaran *Hidden Layer*

| Matriks H | 1 |
|-----------|-----------|
| 1 | 0,7940092 |
| 2 | 0,6800229 |

| | |
|-----|-----------|
| 3 | 0,7790767 |
| ... | ... |
| 16 | 0,6819958 |

Lalu dihitung matriks *moore-penrose generalized invers* (H^+), sehingga diperoleh hasil sebagai berikut.

Tabel 22. Matriks Moore-penrose Generalized Invers

| Matriks H^+ | 1 | 2 | 3 | ... | 16 |
|---------------|------------|------------|------------|-----|------------|
| 1 | 0,09805162 | 0,08397553 | 0,09620761 | ... | 0,08421917 |

Sehingga diperoleh matriks *output weight* berukuran 1 x 1.

Tabel 23. Hasil Output Weight

| Matriks β | 1 |
|-----------------|-----------|
| 1 | 0,4866174 |

2. Proses Testing

Dengan menggunakan data uji, diperoleh matriks H_{init} sebagai berikut.

Tabel 24. Matriks Fungsi Hidden Layer

| Matriks $H_{init\ testing}$ | 1 |
|-----------------------------|-----------|
| 1 | 1,0094374 |
| 2 | 0,8158175 |

Perhitungan menggunakan fungsi aktivasi *sigmoid biner*, sehingga diperoleh matriks H berikut.

Tabel 25. Matriks Keluaran Hidden Layer

| Matriks H | 1 |
|-------------|-----------|
| 1 | 0,7329100 |
| 2 | 0,6933478 |

Tabel berikut merupakan hasil *output layer* berupa nilai peramalan.

Tabel 26. Hasil Peramalan

| Peramalan | 1 |
|-----------|-----------|
| 1 | 0,3566467 |
| 2 | 0,3373951 |

Hasil dari proses denormalisasi data ditunjukkan pada Tabel 27.

Tabel 27. Hasil Denormalisasi Data Peramalan

| Denormalisasi | 1 |
|---------------|--------|
| 1 | -8148 |
| 2 | -13254 |

3.4 Hasil Peramalan Model Hybrid

Tabel 28 merupakan hasil peramalan jumlah penumpang kereta api yang turun di Stasiun Bandung yang berasal dari *output* program R.

Tabel 28. Hasil Peramalan Model Hybrid Sebelum Covid-19

| No | Periode | Hasil Peramalan | Data Aktual |
|----|--------------|-----------------|-------------|
| 1 | Januari 2019 | 493.778 | 486.711 |

| | | | |
|---|---------------|---------|---------|
| 2 | Februari 2019 | 452.367 | 435.684 |
| 3 | Maret 2019 | 515.736 | 488.888 |
| 4 | April 2019 | 513.895 | 509.367 |
| 5 | Mei 2019 | 521.207 | 492.410 |
| 6 | Juni 2019 | 579.277 | 555.382 |
| 7 | Juli 2019 | 571.236 | 523.272 |
| 8 | Agustus 2019 | 541.038 | 486.596 |

Berdasarkan Tabel 28, dapat dilihat jumlah penumpang kereta api yang turun di Stasiun Bandung pada tahun 2019 paling banyak terjadi pada bulan Juni mencapai 579.277 jiwa. Hal tersebut dikarenakan pada bulan Juni 2019 terdapat libur Hari Raya Idul Fitri 1440 Hijriah.

Tabel 29. Hasil Peramalan Model *Hybrid* Saat Covid-19

| No | Periode | Hasil Peramalan | Data Aktual |
|----|---------------|-----------------|-------------|
| 1 | Januari 2022 | 179.471 | 184.752 |
| 2 | Februari 2022 | 160.048 | 129.059 |

Berdasarkan Tabel 29, diketahui peramalan untuk periode 2 kedepan mencapai 179.471 jiwa pada bulan Januari 2022 dan mengalami penurunan pada bulan Februari menjadi 160.048 jiwa.

3.4 Akurasi Hasil Peramalan

Selanjutnya akan dilihat keakuratan hasil peramalan menggunakan nilai *Mean Absolute Prediction Error* (MAPE). Tabel 30 merupakan hasil dari uji MAPE.

Tabel 30. Nilai MAPE Peramalan Sebelum Covid-19

| Stasiun | Model | MAPE | Kriteria |
|---------|------------------------------------|-------|-------------------|
| Bandung | SARIMA(0,1,1)(0,1,0) ¹² | 5,57% | Berakurasi Tinggi |
| | <i>Hybrid</i> SARIMA-ELM | 5,27% | Berakurasi Tinggi |

Berdasarkan Tabel 30, diketahui nilai MAPE model *hybrid* lebih kecil dibandingkan dengan nilai MAPE model SARIMA. Sehingga dapat disimpulkan bahwa penerapan model *hybrid* pada peramalan jumlah penumpang kereta api yang turun di Stasiun Bandung pada fase sebelum Covid-19 dapat meningkatkan akurasi peramalan dibandingkan dengan hanya menggunakan model SARIMA saja.

Tabel 31. Nilai MAPE Peramalan Saat Covid-19

| Stasiun | Model | MAPE | Kriteria |
|---------|-------------------------|--------|----------|
| Bandung | ARIMA (2,0,0) | 17,91% | Baik |
| | <i>Hybrid</i> ARIMA-ELM | 13,43% | Baik |

Berdasarkan Tabel 31, diketahui nilai MAPE model *hybrid* lebih kecil dibandingkan dengan nilai MAPE model ARIMA. Sehingga dapat disimpulkan bahwa penerapan model *hybrid* pada peramalan jumlah penumpang kereta api yang turun di Stasiun Bandung pada fase saat Covid-19 dapat meningkatkan akurasi peramalan dibandingkan dengan hanya menggunakan model ARIMA saja.

4. KESIMPULAN

Penerapan model *hybrid* diawali dengan mencari model SARIMA terbaik, dimana diperoleh model SARIMA (0,1,1)(0,1,0)¹² untuk data jumlah penumpang kereta api yang turun di Stasiun Bandung sebelum Covid-19 dan model ARIMA (2,0,0) untuk data saat Covid-19. Residual dari

masing-masing model diolah menggunakan model *Extreme Learning Machine* (ELM). Dengan perbandingan data 80:20, 3 fitur input, 1 neuron hidden layer, dan 1 output, diperoleh peramalan dari model hybrid sebelum Covid-19 untuk periode Januari 2019- Agustus 2019 dan peramalan dari model hybrid saat Covid-19 untuk periode Januari 2022 dan Februari 2022. Berdasarkan kriteria MAPE, diperoleh peramalan berakurasi tinggi untuk fase sebelum Covid-19 dengan nilai MAPE sebesar 5,27% dan peramalan baik untuk fase saat Covid-19 dengan nilai MAPE 13,43%. Penelitian selanjutnya dapat menambahkan pengujian untuk menentukan jumlah neuron pada hidden layer untuk mengetahui neuron yang diperlukan agar mendapatkan hasil yang lebih optimal.

5. DAFTAR PUSTAKA

- Chen, P., Niu, A., Liu, D., Jiang, W., & Ma, B. (2018). Time series forecasting of temperatures using SARIMA: An Example from Nanjing. *IOP Conference Series: Materials Science and Engineering*, 394, 052024.
- Huang, G.-B., Zhu, Q.-Y., & Siew, C.-K. (2006). Extreme learning machine: theory and applications. *Neurocomputing*, 70(1), 489–501.
- Izati, N. A., Warsito, B., & Widiari, T. (2019). Prediksi harga emas menggunakan feed forward neural network dengan metode extreme learning machine. *Jurnal Gaussian*, 8(2), 171-183.
- Moreno, J. J. M., Pol, A. P., Abad, A. S., & Blasco, B. C. (2013). Using the R-MAPE index as a resistant measure of forecast accuracy. *Psicothema*, 25(4), 500-506.
- Mulyani, S., Hayati, D., & Sari, A. N. (2021). Analisis metode peramalan (Forecasting) penjualan sepeda motor honda dalam menyusun anggaran penjualan pada PT Trio Motor Martadinata Banjarmasin. *Dinamika Ekonomi: Jurnal Ekonomi dan Bisnis*, 14(1), 178-188.
- Prianda, B. G., & Widodo, E. (2021). Perbandingan metode seasonal arima dan extreme learning machine pada peramalan jumlah wisatawan mancanegara ke Bali. *Barekeng: Jurnal Ilmu Matematika dan Terapan*, 15(4), 639-650.
- Rowan, R., Muflikah, L., & Cholissodin, I. (2022). Peramalan kasus positif COVID-19 di Jawa Timur menggunakan metode Hybrid ARIMA-LSTM. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 6(9), 4146-4153.
- Simbolon, S. N., Suprijadi, J., & Arisanti, R. (2021). Penerapan metode extreme learning machine dalam meramalkan temperatur udara harian kota Jakarta Pusat. *Prosiding Seminar Nasional Statistika*, 10(22).
- Sugianto, N.A., cholissodin, I. & widodo, A.W., 2018. Klasifikasi keminatan menggunakan algoritme extreme learning machine dan particle swarm optimization untuk seleksi fitur (Studi kasus: Program Studi Teknik Informatika FILKOM UB). *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 2(5), 1856–1865.
- Utomo, P., & Fanani, A. (2020). Peramalan jumlah penumpang kereta api di Indonesia menggunakan metode Seasonal Autoregressive Integrated Moving Average (SARIMA). *Jurnal Mahasiswa Matematika ALGEBRA*, 1(1), 169-178.