

BAB III

METODE PENELITIAN

3.1 Jenis penelitian

Penelitian ini merupakan jenis penelitian kuantitatif karena penelitian ini bertujuan untuk mengolah data kuantitatif (angka) menggunakan model-model matematis.

3.2 Jenis dan Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder (*time series*) berupa data jumlah penumpang kereta api yang turun di tiga stasiun, yaitu Stasiun Bandung, Stasiun Yogyakarta dan Stasiun Malang yang diperkirakan menjadi tujuan wisata bagi para wisatawan. Data tersebut diperoleh dari PT. Kereta Api Indonesia (Persero) melalui aplikasi PPID PT KAI. Data yang diambil adalah data bulanan dari Agustus 2014 sampai dengan Juni 2022. Pada penelitian ini data akan dibagi menjadi dua, yaitu data jumlah penumpang sebelum pandemi Covid-19 dan saat pandemi Covid-19.

3.3 Variabel Penelitian

Variabel penelitian yang digunakan dalam penelitian ini adalah jumlah penumpang kereta api yang turun di Stasiun Bandung, Yogyakarta dan Malang periode sebelum Covid-19 (Agustus 2014-Desember 2018) dan periode saat Covid-19 (Januari 2020-Desember 2021).

3.4 Metode Analisis Data

Metode analisis data yang digunakan dalam penelitian ini adalah metode *Hybrid SARIMA- Extreme Learning Machine* (ELM), dimana model SARIMA digunakan untuk model linear dan model ELM digunakan untuk model non-linear. Langkah-langkahnya adalah sebagai berikut.

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 variansinya, 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 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* $< \alpha$ dengan $\alpha = 0,05$.

- e. Menguji kelayakan model (uji white noise dan uji normalitas)

- Uji *white noise*

Hipotesis :

$H_0: \rho_1 = \rho_2 = \dots = \rho_k = 0$ (residu memenuhi asumsi *white noise*)

$H_1: \exists \rho_k \neq 0, k = 1, 2, \dots, K$ (residu tidak memenuhi asumsi *white noise*)

Kriteria Pengujian :

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

- f. Memilih model terbaik dengan menggunakan nilai AIC terkecil
 - g. Melakukan peramalan untuk beberapa periode kedepan
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:

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

- b. Menentukan data latih dan target latih untuk proses *training* serta data uji untuk dan target uji untuk proses *testing*.

Tabel 3. 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
...
15	Data pada bulan ke-41 s.d bulan ke-43	Data pada bulan ke-42

Tabel 3. 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
...
14	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 dkk., 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*.
- 2) Menghitung matriks H_{init} menggunakan persamaan:

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

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:

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

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

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

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

$$\beta = H^+ Y$$

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:

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

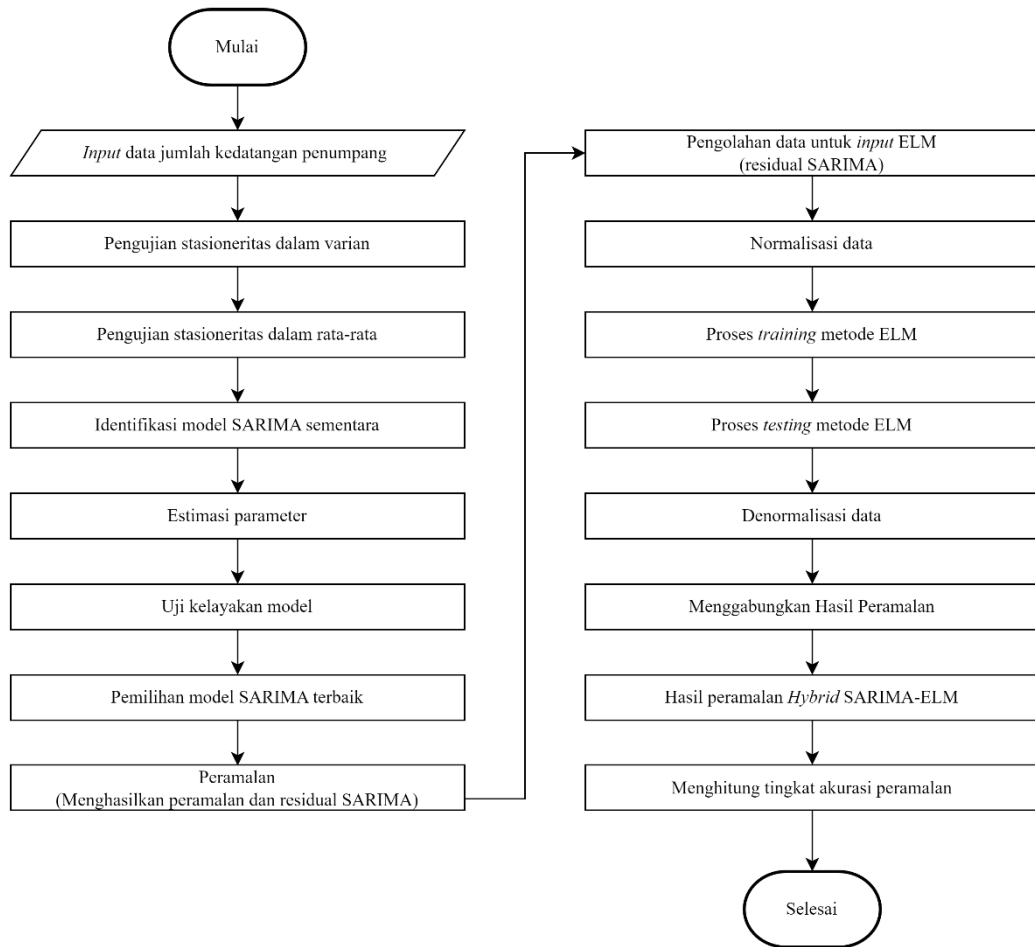
- 3) Melakukan dernationalisasi data dari hasil peramalan untuk mendapatkan nilai aktual dengan persamaan:

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

3. Memperoleh hasil peramalan metode *Hybrid SARIMA-ELM*
4. Menghitung nilai MAPE untuk menghitung akurasi peramalan menggunakan persamaan:

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

3.5 Alur Penelitian



Gambar 3.1 Flow Chart Hybrid SARIMA-ELM