

BAB III

METODE PENELITIAN

3.1 Data dan Variabel Penelitian

Data dalam penelitian ini adalah data sekunder berupa data rata-rata bulanan curah hujan dan kelembaban udara di Kota Bandung. Rentang waktu yang digunakan sebagai data *input* dimulai dari bulan Januari 2008 sampai Desember 2022 yang diambil dari Badan Pusat Statistik (BPS) Kota Bandung. Penelitian ini menggunakan model runtun waktu dengan dua variabel yang digunakan, yaitu variabel endogen dan variabel eksogen. Variabel endogen berupa curah hujan dan variabel eksogennya berupa kelembaban udara dalam jangka waktu yang sama.

Pada model hybrid, setelah model pertama yaitu SARIMAX berhasil dijalankan, residu dari hasil model tersebut dibagi menjadi dua bagian yaitu data latih dan data uji. Data latih digunakan untuk pelatihan model LSTM, sedangkan data uji digunakan sebagai bahan penentuan keakuratan peramalan.

3.2 Langkah Analisis

Langkah-langkah analisis data diolah dengan menggunakan bantuan bahasa pemrograman *Python* pada *Google Colab* dan *Microsoft Excel*. Adapun langkah-langkah analisis yang dilakukan pada penelitian ini adalah sebagai berikut.

3.3.1 Pembentukan Model SARIMAX

Berikut adalah langkah-langkah identifikasi model SARIMAX $(p, d, q)(P, D, Q)s(B)$.

- 1) Membuat plot data deret waktu dari data awal.
- 2) Menguji linearitas data dengan menggunakan Ramsey's RESET Test.
- 3) Melakukan analisis time series FAK dan FAKP.
- 4) Menguji kestasioneritasan data.
- 5) Mengidentifikasi komponen ARIMA (p, d, q) .
- 6) Mengidentifikasi model SARIMAX $(p, d, q)(P, D, Q)s(B)$.
- 7) Pemilihan model terbaik dengan uji signifikansi parameter dan uji diagnostik *white noise* L'Jung Box dan perbandingan nilai AIC.

- 8) Membentuk model SARIMAX terbaik dengan persamaan 2.16.

$$\varphi_p(B)\phi_p(B^s)(1-B)^d(1-B^s)^D Z_t = \theta_q(B)\vartheta_Q(B^s)a_t + \beta X_t$$

dimana:

Z_t : nilai deret waktu pada waktu t

μ : konstanta

$\varphi_p(B)$: parameter *autoregressive*

$\phi_p(B^s)$: parameter *autoregressive* musiman

a_t : *error* (kesalahan) pada waktu t

$\theta_q(B)$: parameter *moving average*

$\vartheta_Q(B^s)$: parameter *moving average* musiman

X_t : variabel eksogen pada waktu t

β : parameter yang menghubungkan variabel eksogen dengan deret waktu

- 9) Melakukan peramalan.

- 10) Mengevaluasi hasil peramalan berdasarkan *Root Mean Squared Error* (RMSE).

3.3.2 Pembentukan Model LSTM

Berikut adalah langkah-langkah pembentukan model LSTM

- 1) Melakukan *preprocessing* data dengan normalisasi data.

Metode ini menghasilkan nilai pada selang 0 - 1 dengan rumus sebagai berikut.

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (3.1)$$

dengan

x' : data normalisasi

x : data asli

x_{min} : nilai minimum dari data per kolom

x_{max} : nilai maximum dari data per kolom

- 2) Membagi data menjadi data *training* 80% dan data *testing* 20%.

- 3) Inisialisasi model LSTM .

- 4) Menentukan nilai *window size*.

- 5) Menentukan nilai *batch size*, *neurons*, dan *epochs* yang optimal.

- 6) Membangun model LSTM.
- 7) Menentukan nilai parameter yang optimal pada masing-masing gerbang LSTM.
- 8) Peramalan menggunakan model LSTM untuk data *testing*.
- 9) Transformasi kembali hasil prediksi ke skala awal.
- 10) Mengevaluasi hasil peramalan data *testing* berdasarkan *Root Mean Squared Error* (RMSE).

3.3.3 Pembentukan Model SARIMAX-LSTM

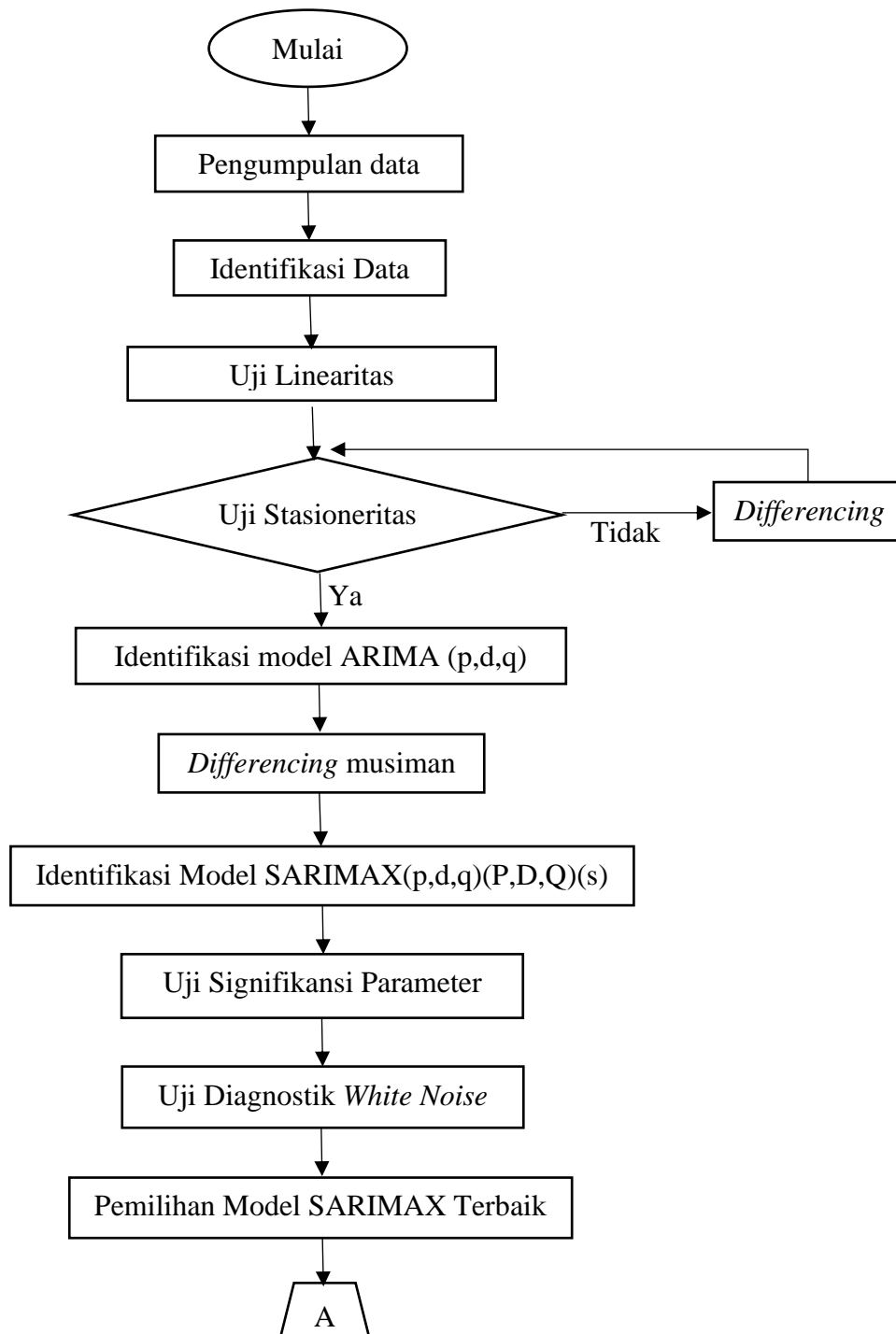
Langkah-langkah untuk melakukan peramalan dengan metode *Seasonal Autoregressive Integrated Moving Average with Exogenous Variables* (SARIMAX) dan *Long Short-Term Memory* (LSTM) adalah sebagai berikut.

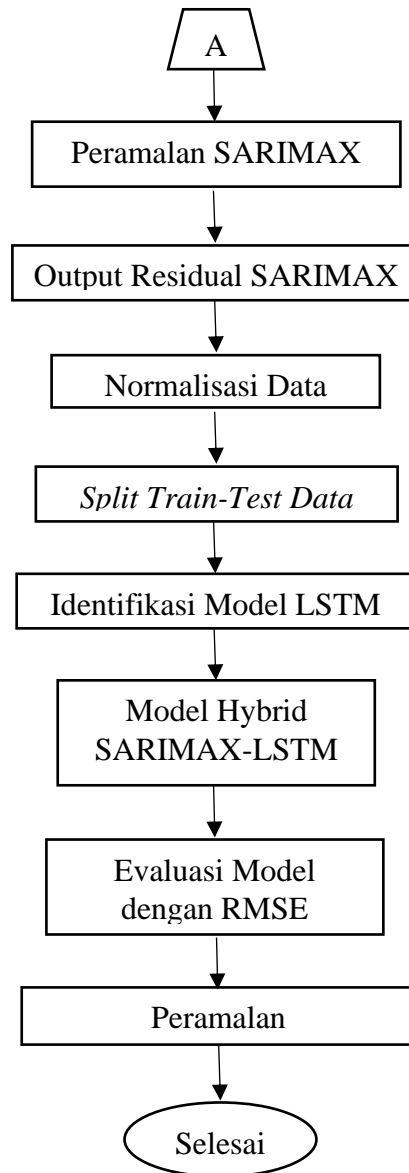
- 1) Membuat plot data deret waktu dari data awal.
- 2) Menguji linearitas data dengan menggunakan Ramsey's RESET Test.
- 3) Melakukan analisis time series FAK dan FAKP.
- 4) Menguji kestasioneritasan data.
- 5) Mengidentifikasi komponen ARIMA (p, d, q) .
- 6) Mengidentifikasi model SARIMAX $(p, d, q)(P, D, Q)s(B)$.
- 7) Pemilihan model terbaik dengan uji signifikansi parameter dan uji diagnostik *white noise* L'Jung Box dan perbandingan nilai AIC.
- 8) Membentuk model SARIMAX terbaik yang telah diperoleh.
- 9) Ambil residu dari *output* model SARIMAX.
- 10) Normalisasi data residual sehingga menghasilkan rentang nilai 0-1.
- 11) Pisahkan residu yang telah dinormalisasi menjadi 80% data *training* dan 80% data *testing*.
- 12) Inisialisasi model LSTM.
- 13) Menentukan nilai *window size*, *batch size*, *neurons*, dan *epochs* yang optimal untuk memperoleh model LSTM terbaik dengan melakukan eksperimen dan membandingkan nilai RMSEnya. Rentang nilai eksperimennya adalah sebagai berikut.

- *window size* : 12
- *batch size* : 16, 32, 64, 128
- *neurons* : 32, 64, 128, 256
- *epochs* : 25, 50, 75, 100

- 14) Membangun model LSTM terbaik.
- 15) Melakukan peramalan LSTM dengan data *training* untuk data *testing*.
- 16) Transformasikan kembali data hasil peramalan ke skala awal.
- 17) Gabungkan hasil peramalan dari model SARIMAX dan LSTM.
- 18) Mengevaluasi hasil peramalan data *testing* berdasarkan *Root Mean Squared Error* (RMSE).
- 19) Melakukan peramalan.

3.4 Alur Penelitian





Gambar 3.1 Alur Penelitian SARIMAX-LSTM