

BAB III

METODOLOGI PENELITIAN

3.1 Pendekatan Penelitian

Penelitian ini menggunakan pendekatan dengan jenis penelitian kuantitatif deskriptif. Metode penelitian Kuantitatif merupakan metode penelitian yang digunakan untuk meneliti populasi atau sampel tertentu dengan menggunakan instrumen dan analisis data yang bersifat kuantitatif/statistik untuk menguji hipotesis. Pendekatan penelitian ini menggunakan pendekatan kuantitatif karena menggunakan angka, mulai dari pengumpulan data, analisis terhadap data tersebut, serta dilihat dari hasilnya. Kelebihan penelitian kuantitatif adalah penelitian ini dapat digunakan untuk menduga, akurat karena berupa angka, hubungan antara variabel lebih jelas, dan menyederhanakan permasalahan yang kompleks.

3.2 Tempat dan Waktu Penelitian

Penelitian ini dilakukan di Laboratorium Komputer Jurusan Matematika Universitas Bangka Belitung. Adapun waktu yang digunakan pada penelitian ini dimulai dari Januari 2025 sampai dengan Mei 2025.

3.3 Jenis dan Sumber Data

3.3.1 Jenis Data

Jenis data yang digunakan pada penelitian ini yaitu data sekunder yang berupa data *time series* (deret waktu). *Time series* (deret waktu) merupakan data yang disusun berdasarkan urutan waktu atau data yang dikumpulkan dari waktu ke waktu, waktu yang digunakan dapat berupa hari, minggu, bulan, tahun. Data sekunder adalah data yang peneliti kumpulkan secara tidak langsung karena data diperoleh melalui perantara seperti dari *website*, dokumen – dokumen penting, buku, dan sebagainya.

3.3.2 Sumber Data

Data penelitian yang digunakan merupakan data *time series* yang diperoleh dari *website* Badan Pusat Statistika (BPS) Provinsi Kepulauan Bangka Belitung. data Produk Domestik Regional Bruto (PDRB) menggunakan indikator PDRB Atas Dasar Harga Konstan (ADHK) dari tahun 2010 samapi dengan tahun 2025 triwulan

1 yang tersedia pada link Badan Pusat Statistik (BPS) Provinsi Kepulauan Bangka Belitung <https://babel.bps.go.id/>.

3.4 Teknik Pengumpulan Data

Teknik pengumpulan data dalam penelitian ini menggunakan teknik dokumentasi. Teknik dokumentasi merupakan salah satu sumber data sekunder yang diperlukan dalam sebuah penelitian. Dokumentasi bisa berbentuk tulisan angka, gambar, buku, asip, dokumen serta keterangan yang dapat mendukung penelitian (Sugiyono, 2015)

3.5 Teknik Analisis Data

Teknik analisis data yang digunakan yaitu metode SARIMA dan SVR yang akan memprediksi data penelitian dengan beberapa tahapan sehingga memperoleh model dan estimasi yang tepat. Adapun tahapan – tahapan penelitian ini sebagai berikut:

1. Pengumpulan Data

Data yang digunakan dalam penelitian ini diperoleh melalui situs resmi Badan Pusat Statistik (BPS) Provinsi Kepulauan Bangka Belitung. Data yang digunakan adalah data PDRB Atas Dasar Harga Konstan (ADHK) menurut lapangan usaha dari tahun 2010 triwulan I sampai dengan tahun 2025 triwulan I.

2. Identifikasi Data

Pada tahapan identifikasi data ini bertujuan untuk mengetahui apakah data tersebut mengandung pola musiman. Pada tahapan ini dilakukan dengan cara membuat *plot* dalam bentuk grafik.

3. Analisis Statistika Deskriptif

Pada tahapan analisis statistika deskriptif ini bertujuan untuk mendapatkan gambaran tentang data yang digunakan pada penelitian ini yang dapat dilihat dari nilai rata – rata (mean), standar deviasi, maksimum, dan minimum dari data yang digunakan yaitu data PDRB ADHK menurut lapangan usaha.

4. Analisis Metode SARIMA.

Adapun tahapan penelitian data pada metode SARIMA sebagai berikut:

a. Identifikasi Stasioneritas Data

Data diuji kestasioneritasnya dengan menggunakan Uji ADF untuk menguji Data stasioner dalam rata-rata. ADF akan memberikan nilai statistik dan p-value:

Jika $p\text{-value} < 0.05$, maka H_0 ditolak, data stasioner.

Jika $p\text{-value} \geq 0.05$, maka H_0 gagal ditolak, data tidak stasioner.

Uji ADF bisa dilakukan dengan atau tanpa trend, tergantung sifat datanya. Jika data tidak stasioner, maka bisa dilakukan *differencing* (pembedaan) untuk membuatnya stasioner.

Uji Box-Cox menguji stasioner dalam varians. Box-Cox adalah transformasi matematis untuk menstabilkan varians dari data. Ini membantu ketika data menunjukkan heteroskedastisitas (varian berubah-ubah), misalnya tren naik disertai fluktuasi yang makin besar.

b. Uji ACF dan PACF

ACF adalah perhitungan autokorelasi berdasarkan varians dan kovariansnya. Autokorelasi merupakan cara untuk melihat adanya hubungan atau korelasi pada data yang sama antar waktu. *Partial Autocorrelative Function* (PACF) adalah suatu persamaan yang berfungsi sebagai penghitung ukuran kekuatan dari variabel Y_t dan Y_{t+k} .

c. Pemilihan Model Terbaik

SARIMA (*Seasonal Autoregressive Integrated Moving Average*) adalah Perluasan dari model ARIMA. Persamaan model SARIMA dituliskan sebagaimana pada persamaan [2.11].

d. Menghitung nilai MAPE

Menghitung nilai MAPE bertujuan untuk melihat Tingkat akurasi model Sebelum diprediksi. Nilai MAPE dapat dilihat berdasarkan tingkat signifikansinya jika semakin kecil (mendekati 0) maka nilai MAPE semakin baik untuk melakukan prediksi. Adapun jika nilai MAPE lebih dari 50 persen maka nilai MAPE termasuk kriteria sangat buruk sehingga tidak bisa untuk melakukan prediksi.

e. Menghitung nilai RMSE

RMSE adalah salah satu metrik yang digunakan untuk mengukur seberapa baik suatu model memprediksi data dibandingkan dengan nilai aktual. RMSE menghitung akar dari rata-rata kuadrat perbedaan antara nilai yang diprediksi oleh model dan nilai aktual. Jika nilai RMSE kurang dari 1 maka nilai RMSE termasuk kriteria baik.

f. Prediksi

Hasil prediksi dilakukan setelah menemukan model SARIMA terbaik untuk melakukan prediksi pada data PDRB ADHK Provinsi Kepulauan Bangka Belitung periode Januari 2024 hingga Desember 2027.

5. Analisis Metode *Support Vector Regression* (SVR)

Adapun tahapan penelitian pada metode *Support Vector Regression* (SVR) sebagai berikut:

a. Normalisasi Data Proses

normalisasi dilakukan untuk mengurangi perbedaan jangkauan data yang jauh antara variabel. Nilai rentang dari data memiliki ciri sangat bervariasi, 26 sehingga nilai yang diperoleh harus diskalakan kedalam batas nilai tertentu supaya tidak terdapat data yang terlalu besar maupun data yang terlalu kecil yang mempengaruhi hasil. Pada normalisasi dimana fungsi ini bernilai antara 0 sampai dengan 1. Proses normalisasi hanya digunakan pada kernel *sigmoid* saja sebagai fungsi aktivasi.

b. Pembagian Data

Data akan dibagi menjadi data latih dan data uji dengan persentase 70 persen : 30 persen. Data latih digunakan untuk menentukan model terbaik sedangkan data uji untuk menguji keakuratan model serta mendapatkan hasil prediksi.

c. Penentuan Kernel dan Parameter

Pemodelan SVR untuk membentuk model SVR dengan menggunakan data latih kemudian data tersebut digunakan untuk menentukan kernel dan parameter terbaik. Parameter SVR yang dicari dengan optimasi ini adalah nilai C , γ (*gamma*) dan nilai ϵ (*epsilon*) yang paling optimal, yang mana nilai dari parameter C , γ dan ϵ akan mempengaruhi presisi SVR secara signifikan. Model terbaik dapat dilihat dari nilai MAPE yang rendah pada data latih, sehingga model terbaik yang dihasilkan dapat diimplementasikan pada data uji. Setelah masing – masing kernel mendapatkan model terbaik, maka dari ke – empat model tersebut akan dipilih satu untuk dilakukan prediksi data PDRB ADHK Provinsi Kepulauan Bangka Belitung.

d. Pemilihan Model Terbaik

Pada tahapan pemilihan model SVR terbaik, proses dilakukan dengan mengevaluasi performa model yang dibentuk berdasarkan data latih untuk setiap

kombinasi kernel dan parameter yang digunakan. Apabila model dari data latih belum menunjukkan performa yang optimal atau belum sesuai dengan kriteria evaluasi seperti nilai MAPE dan RMSE yang rendah, maka dilakukan pengulangan proses pelatihan dengan penyesuaian nilai parameter dan pemilihan jenis kernel yang berbeda yang dilakukan melalui metode grid search atau trial and error.

e. Menghitung Nilai MAPE

Menghitung nilai MAPE bertujuan untuk melihat tingkat akurasi model sebelum diprediksi. Nilai MAPE dapat dilihat berdasarkan tingkat signifikansinya jika semakin kecil (mendekati 0) maka nilai MAPE semakin baik untuk melakukan prediksi. Adapun jika nilai MAPE lebih dari 50 persen maka nilai MAPE termasuk kriteria sangat buruk sehingga tidak bisa untuk melakukan prediksi.

f. Menghitung Nilai RMSE

RMSE adalah salah satu metrik yang digunakan untuk mengukur seberapa baik suatu model memprediksi data dibandingkan dengan nilai aktual. RMSE menghitung akar dari rata-rata kuadrat perbedaan antara nilai yang diprediksi oleh model dan nilai aktual. Jika nilai RMSE kurang dari 1 maka nilai RMSE termasuk kriteria baik

g. Prediksi

Prediksi dilakukan setelah dihasilkan model SVR terbaik melalui tahapan pada proses pelatihan data. Model SVR yang terbaik akan digunakan untuk memprediksi data PDRB tahun 2025 triwulan II hingga tahun 2028 triwulan IV.

h. Denormalisasi

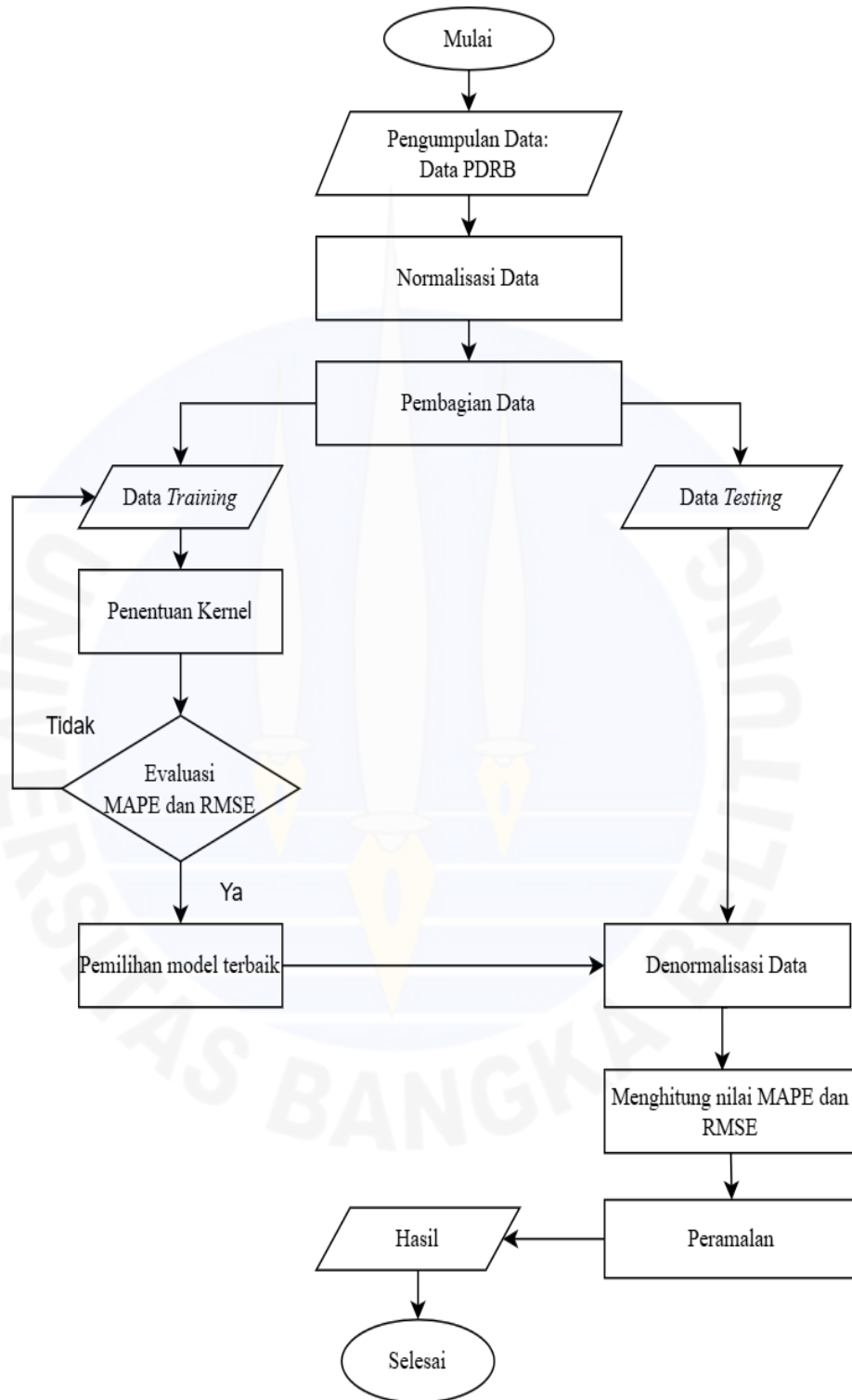
Denormalisasi dilakukan untuk mengembalikan data yang telah dinormalisasi ke bentuk data asli

6. Perbandingan Hasil Prediksi

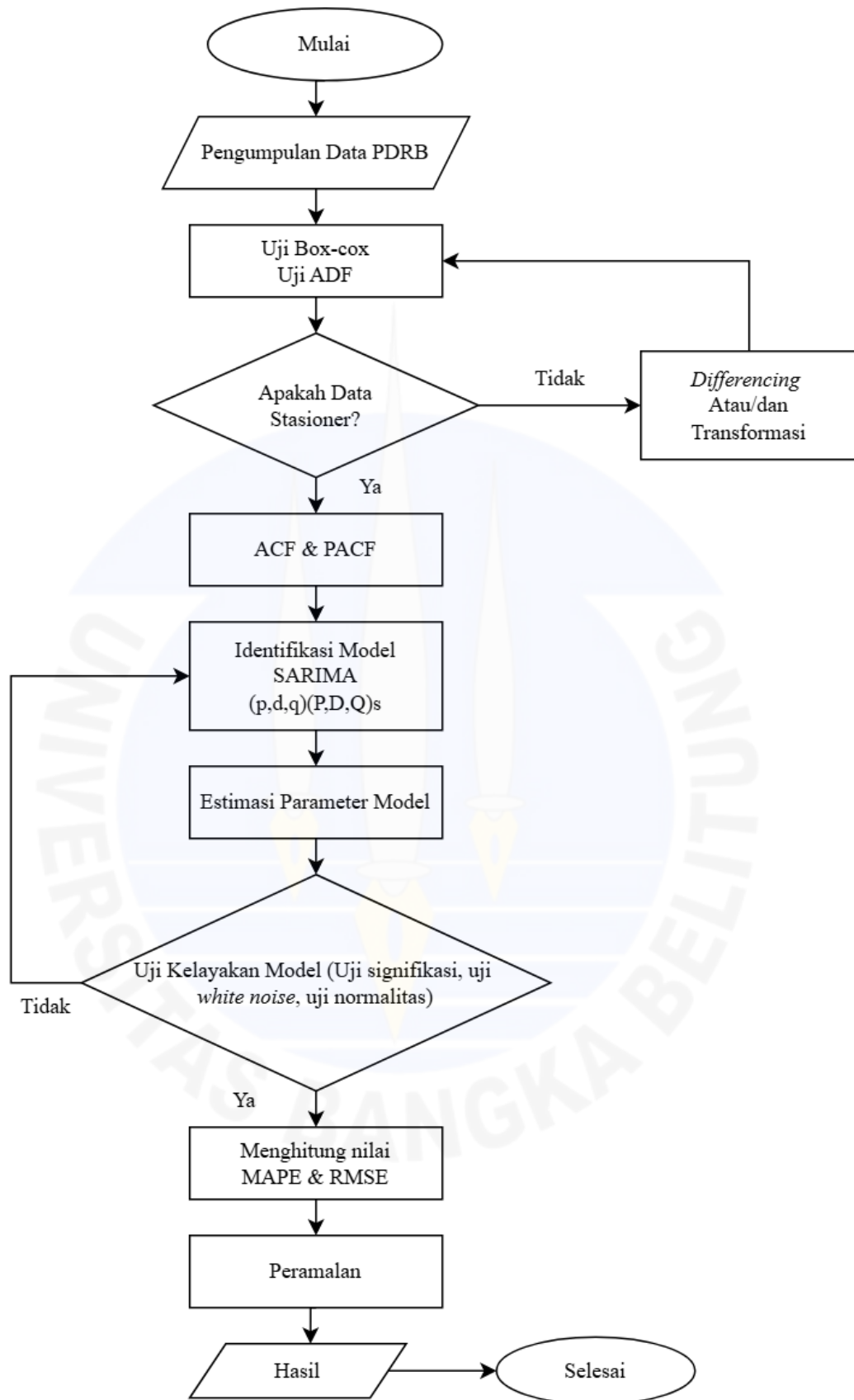
Perbandingan hasil prediksi antara metode SARIMA dan SVR dilakukan dengan tujuan untuk mengevaluasi kinerja masing-masing model dalam menghasilkan nilai PDRB ADHK. Evaluasi ini mencakup akurasi hasil prediksi, kemampuan model dalam menangkap pola data historis. Dapat diperoleh metode yang paling sesuai untuk digunakan dalam memproyeksikan nilai PDRB ADHK pada periode Januari 2024 hingga Desember 202

3.6 Diagram Alir Penelitian

Diagram alir penelitian ini dapat dilihat pada gambar 3.1 dan 3.2 dibawah ini.



Gambar 3.1 Diagram Alir Penelitian Metode SVR



Gambar 3.2. Diagram Alir Penelitian Metode SARIMA