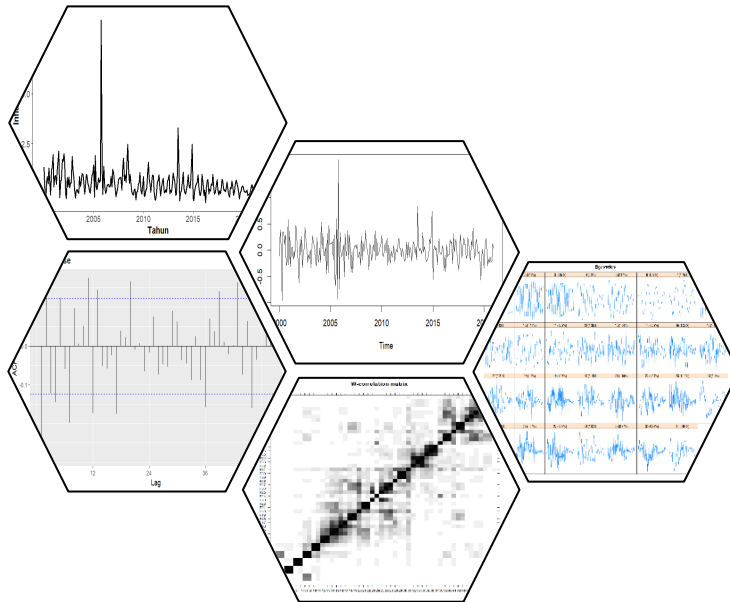


**PERBANDINGAN METODE *HYBRID* SSA-SVR DAN *HYBRID* SSA-ARIMA DALAM
MEMPREDIKSI INFLASI DI INDONESIA**

**COMPARISON OF *HYBRID* SSA-SVR AND *HYBRID* SSA-ARIMA METHODS FOR
PREDICTING INFLATION IN INDONESIA**



NUR IKHWANA

H062221012



**PROGRAM STUDI MAGISTER STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS HASANUDDIN
MAKASSAR
2024**

**PERBANDINGAN METODE *HYBRID* SSA-SVR DAN *HYBRID* SSA-ARIMA DALAM
MEMPREDIKSI INFLASI DI INDONESIA**

**NUR IKHWANA
H062221012**



**PROGRAM STUDI MAGISTER STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS HASANUDDIN
MAKASSAR
2024**

**PERBANDINGAN METODE *HYBRID* SSA-SVR DAN *HYBRID* SSA-ARIMA DALAM
MEMPREDIKSI INFLASI DI INDONESIA**

Tesis

Sebagai salah satu syarat untuk mencapai gelar magister

Program Studi Magister Statistika

Disusun dan diajukan oleh

NUR IKHWANA
H062221012

kepada

**PROGRAM STUDI MAGISTER STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS HASANUDDIN
MAKASSAR
2024**

TESIS

**PERBANDINGAN METODE HYBRID SSA-SVR DAN HYBRID SSA-ARIMA
DALAM MEMPREDIKSI INFLASI DI INDONESIA**

NUR IKHWANA

H062221012

Telah dipertahankan di hadapan Panitia Ujian Magister pada 08 Agustus 2024 dan dinyatakan telah memenuhi syarat kelulusan


pada

Program Studi Magister Statistika
Departemen Statistika
Fakultas Matematika Dan Ilmu Pengetahuan Alam
Universitas Hasanuddin
Makassar

Mengesahkan:

Pembimbing Utama

Pembimbing Pendamping

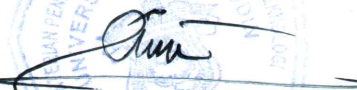


Dr. Erna Tri Herdiani, S.Si., M.Si.
NIP. 19750429 200003 2 001

Prof. Dr. Dr. Georgina M. Tinungki, M.Si.
NIP. 19620926 198702 2 001

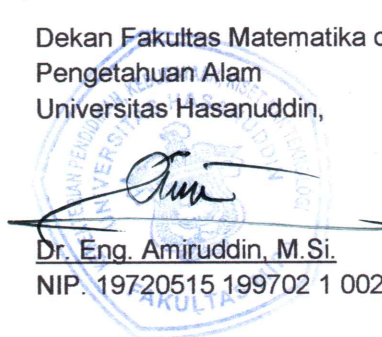
Ketua Program Studi
Magister Statistika,

Dekan Fakultas Matematika dan Ilmu
Pengetahuan Alam
Universitas Hasanuddin,



Dr. Erna Tri Herdiani, S.Si., M.Si.
NIP. 19750429 200003 2 001

Dr. Eng. Amiruddin, M.Si.
NIP. 19720515 199702 1 002



**PERNYATAAN KEASLIAN TESIS
DAN PELIMPAHAN HAK CIPTA**

Dengan ini saya menyatakan bahwa, tesis berjudul "Perbandingan Metode *Hybrid* SSA-SVR dan *Hybrid* SSA-ARIMA dalam Memprediksi Inflasi di Indonesia" adalah benar karya saya dengan arahan dari tim pembimbing (Dr. Erna Tri Herdiani, S.Si., M.Si. sebagai Pembimbing Utama dan Prof. Dr. Dr. Georgina Maria Tinungki, M.Si. sebagai Pembimbing Pendamping). Karya ilmiah ini belum diajukan dan tidak sedang diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka tesis ini. Sebagian dari isi tesis ini akan di publikasikan di *Journal of Modern Applied Statistical Methods* sebagai artikel dengan judul "*A hybrid Model of Singular Spectrum Analysis and Support Vector Regression to Predict Inflation in Indonesia*". Apabila di kemudian hari terbukti atau dapat dibuktikan bahwa sebagian atau keseluruhan tesis ini adalah karya orang lain, maka saya bersedia menerima sanksi atas perbuatan tersebut berdasarkan aturan yang berlaku.

Dengan ini saya melimpahkan hak cipta (hak ekonomis) dari karya tulis saya berupa tesis ini kepada Universitas Hasanuddin.

Makassar, 08 Agustus 2024



Nur Ikhwana
NIM H062221012

UCAPAN TERIMA KASIH

Bismillahirrohmanirrohim,

Tiada kata yang terurai indah yang terlontarkan dari lisan ini, selain puji syukur kehadiran Allah SWT, Tuhan semesta alam yang selalu menganugerahkan kesehatan dan kesempatan untuk melalui setiap perjalanan hidup ini. Berkat kasih sayang dan pertolonganNya penulis dapat menyelesaikan tesis yang berjudul “**Perbandingan Metode Hybrid SSA-SVR dan Hybrid SSA-ARIMA dalam Memprediksi Inflasi di Indonesia**”. Sholawat terbingkai salam senantiasa terlimpah curahkan kepada Nabiullah Muhammad SAW yang membawa kita dari gelapnya zaman kebodohan menuju terang benderangnya zaman yang kaya akan ilmu pengetahuan seperti saat ini.

Kelancaran dalam penyusunan tesis ini tidak terlepas dari doa, dukungan, serta pengorbanan dari orang yang paling berjasa dalam hidup penulis yaitu Orang tua penulis, Ayahanda H. Parawansa dan Ibunda Hj. Hasni H Mansur, Kakek dan Nenek (H. Lido, Hj. Nanang, Hj. Marhana, dan H. Mansur), Adik (Muh. Ikhwan), serta seluruh keluarga besar yang tidak henti-hentinya melangitkan doa doa terbaik untuk penulis. Saat ini, bertambah satu doanya yang saya jemput ke bumi.

Penulis menyadari bahwa tesis ini tidak dapat diselesaikan dengan baik tanpa bantuan dari berbagai pihak. Ucapan terima kasih yang sebesar-besarnya kepada Ibu Dr. Erna Tri Herdiani, S.Si., M.Si. dan Ibu Prof. Dr. Dr. Georgina Maria Tinungki, M.Si. selaku pembimbing yang senantiasa meluangkan waktu memberi bimbingan, arahan, saran, dan memotivasi penulis dalam penyusunan dan penyelesaian tesis ini. Juga kepada Bapak Dr. Nirwan Ilyas, M.Si., Ibu Prof. Dr. Nurtiti Sunusi, S.Si., M.Si., serta Ibu Dr. Anna Islamiyati, S.Si., M.Si, selaku penguji yang senantiasa memberikan saran dan masukan demi perbaikan tesis ini.

Penulis juga menyampaikan rasa syukur dan terima kasih yang sebesar besarnya kepada:

1. Bapak Prof. Dr. Ir. Jamaluddin Jompa, M.Sc. selaku Rektor Universitas Hasanuddin.
2. Bapak Dr. Eng. Amiruddin, M.Si. selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam beserta seluruh jajarannya.
3. Ibu Dr. Anna Islamiyati, S.Si., M.Si. selaku Ketua Departemen Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam
4. Ibu Dr. Erna Tri Herdiani, S.Si., M.Si. selaku Ketua Program Studi Magister Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam.
5. Bapak Ibu Dosen, staf, serta seluruh civitas akademika Departemen Statistika FMIPA UNHAS, yang telah mendidik, memberi banyak ilmu dan berbagi pengalaman kepada penulis selama masa perkuliahan.
6. Kawan-kawan seperjuangan, Nalto Batty Mangiri, Rizka Febrianti, Mahrani, serta seluruh teman teman di Program Studi Magister Statistika. Terima kasih atas segala bantuan dan kebersamaan dalam menghadapi masa-masa terindah maupun masa masa sulit dalam menuntut ilmu di kampus FMIPA UNHAS.
7. Sahabat sahabat tercinta, Mita, Nini, Jija, Ana, Pute, Fitri, Mubi, Tetti, Aja', teman-teman Signifikan, serta seluruh pihak yang tidak bisa penulis sebutkan satu persatu, semoga segala kebaikan dan bantuan bernilai ibadah disisi Allah SWT.

Tak lupa penulis juga mengucapkan terima kasih yang sebesar-besarnya kepada pimpinan Program Studi Statistika FMIPA UNM, Ibu Dr. Ruliana, S.Pd., M.Si. yang telah

memberikan izin kepada penulis untuk melanjutkan studi dan memberikan banyak bantuan serta memotivasi penulis hingga penyelesaian studi. Kepada Bapak Rahmat H.S., S.Pd., M.Si., Bapak Zulkifli Rais, S.Pd., M.Si., Bapak Ansari Saleh Ahmar, S.Si., M.Sc., Ph.D. yang senantiasa meluangkan waktu memberikan penulis bantuan, bimbingan, arahan, serta memotivasi penulis selama masa perkuliahan hingga penyusunan dan penyelesaian tesis ini. Kepada partner kerja penulis, Ibnu Abdul Hadi yang telah banyak membantu meringankan pekerjaan penulis selama penulis berkuliah. Kakak-kakak "*Supporting Team*" yang selalu menjadi pendengar dan memberi bantuan kepada penulis dalam menyelesaikan tesis ini. Serta, seluruh civitas akademika Program Studi Statistika FMIPA UNM yang senantiasa menjadi tempat bagi penulis untuk belajar dan bertanya, terima kasih atas segala dukungan dan motivasi yang diberikan kepada penulis agar bisa terus belajar dan mengembangkan diri.

Akhir kata, Semoga tulisan ini dapat membawa manfaat bagi pembaca dan bagi kemajuan ilmu pengetahuan.

Penulis,

Nur Ikhwana

ABSTRAK

NUR IKHWANA. **Perbandingan Metode *Hybrid* SSA-SVR dan *Hybrid* SSA-ARIMA dalam Memprediksi Inflasi di Indonesia** (dibimbing oleh Erna Tri Herdiani dan Georgina Maria Tinungki).

Latar Belakang. Penggunaan dua metode secara bertahap dikenal dengan metode *hybrid*. Dimana, hasil yang diperoleh dari metode pertama selanjutnya digunakan pada metode kedua. Metode *hybrid* bertujuan memanfaatkan kelebihan dari masing masing metode untuk mendapatkan kemampuan yang baik dalam hal ini ramalan yang akurat. Berbagai macam metode *hybrid* dengan melibatkan SSA telah banyak dilakukan, diantaranya *hybrid* SSA-SVR dan SSA-ARIMA. **Tujuan.** Penelitian ini bertujuan untuk membandingkan hasil prediksi Inflasi di Indonesia dengan menggunakan metode *hybrid* SSA-SVR dan *Hybrid* SSA-ARIMA sehingga dapat memperoleh model terbaik untuk memprediksi Inflasi di Indonesia. **Metode.** Penelitian dibagi menjadi dua tahap yaitu: 1) Pemodelan *Hybrid* SSA-SVR dan 2) Pemodelan *hybrid* SSA-ARIMA. Analisis dilakukan menggunakan R Studio 2023.03.0+386. **Hasil.** Metode *Hybrid* SSA-SVR menghasilkan akurasi RMSE sebesar 0,1736 yang diperoleh dengan fungsi kernel linear dengan parameter $C = 128$, $\gamma = 0.00390625$, $\varepsilon = 0,01$. Metode *hybrid* SSA-ARIMA menghasilkan akurasi RMSE sebesar 0,3609 yang diperoleh dari model ARIMA (3,0,3). **Kesimpulan.** Model terbaik dalam memprediksi inflasi di Indonesia adalah model *hybrid* SSA-SVR karena menghasilkan akurasi RMSE yang lebih kecil dibandingkan dengan model SSA-ARIMA.

Kata Kunci: Inflasi; *Singular Spectrum Analysis* (SSA); *Support Vector Regression* (SVR); *Autoregressive Moving Average* (ARIMA).

ABSTRACT

NUR IKHWANA. **Comparison of *Hybrid SSA-SVR* and *Hybrid SSA-ARIMA* Methods for Predicting Inflation in Indonesia** (supervised by Erna Tri Herdiani and Georgina Maria Tinungki).

Background. The use of two methods in stages is known as the *hybrid* method. Where, the results obtained from the first method are then used in the second method. The *hybrid* method aims to utilize the advantages of each method to get good capabilities in this case an accurate forecast. Various kinds of *hybrid* methods involving SSA have been carried out, including *hybrid* SSA-SVR and SSA-ARIMA. **Aim.** This study aims to compare the results of Inflation prediction in Indonesia using *hybrid* SSA-SVR and *Hybrid* SSA-ARIMA methods to obtain the best model for predict Inflation in Indonesia. **Methods.** The research is divided into two stages, namely: 1) *Hybrid* SSA-SVR modeling and 2) SSA-ARIMA *hybrid* modeling. The analysis was conducted using R Studio 2023.03.0+386. **Results.** *Hybrid* SSA-SVR method produces an RMSE accuracy of 0.1736 obtained with a linear kernel function with parameters $C = 128$, $\gamma=0.00390625$, $\epsilon=0.01$. The *hybrid* SSA-ARIMA method produces an RMSE accuracy of 0.3609 obtained from the ARIMA (3,0,3) model. **Conclusion.** The best model for predicting inflation in Indonesia is the *hybrid* SSA-SVR model because it produces smaller RMSE accuracy compared to the SSA-ARIMA model.

Keywords: Inflation; Singular Spectrum Analysis (SSA); Support Vector Regression (SVR); Autoregressive Moving Average (ARIMA).

DAFTAR ISI

	Halaman
HALAMAN JUDUL.....	i
HALAMAN PENGANTAR.....	ii
LEMBAR PENGESAHAN.....	iii
PERNYATAAN KEASLIAN TESIS.....	iv
UCAPAN TERIMA KASIH.....	v
ABSTRAK.....	vii
ABSTRACT.....	viii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	x
DAFTAR GAMBAR.....	xi
DAFTAR LAMPIRAN.....	xii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	2
1.3 Batasan Masalah.....	2
1.4 Tujuan Penelitian.....	3
1.5 Manfaat Penelitian.....	3
1.6 Landasan Teori.....	3
BAB II METODE PENELITIAN.....	19
2.1 Sumber Data.....	19
2.2 Identifikasi Variabel.....	19
2.3 Tahapan Analisis Data.....	19
2.4 Diagram Alir.....	21
BAB III HASIL DAN PEMBAHASAN.....	24
3.1 Deskripsi Data.....	24
3.2 Pemodelan <i>Hybrid</i> SSA-SVR.....	25
3.3 Pemodelan <i>Hybrid</i> SSA-ARIMA.....	47
3.4 Penentuan Model Terbaik.....	51
BAB IV KESIMPULAN DAN SARAN.....	54
4.1 Kesimpulan.....	54
4.2 Saran.....	54
DAFTAR PUSTAKA.....	55
LAMPIRAN.....	58

DAFTAR TABEL

Nomor Urut	Halaman
1. Model runtun waktu AR, MA, dan ARMA	3
2. Nilai lambda (λ) beserta jenis transformasi	4
3. Plot ACF dan PACF	5
4. Fungsi Kernel.....	15
5. Deskripsi Data Inflasi Indonesia 2000-2022.....	24
6 Diagonal averaging Data Simulasi	29
7. Hasil Evaluasi Model Hybrid SSA-SVR Data Simulasi.....	31
8. Hasil pelacakan beberapa nilai L	31
9. Pelacakan nilai pada sekitar nilai 100	32
10. Eigenvalues (λ_i) dan Singular value (λ_i).....	33
11. Eigen vektor	33
12. Komponen Utama	34
13. Pengelompokan Eigentriple setiap komponen	41
14. Hasil Diagonal averaging	41
15. Hasil evaluasi model hybrid SSA-SVR pada setiap kernel	46
16. Hasil Peramalan Metode Hybrid SSA-SVR.....	46
17. Estimasi Parameter Model.....	49
18. Hasil Pengujian Asumsi White Noise	49
19. Hasil Uji Shapiro Wilk.....	50
20. Nilai AIC masing-masing model	50
21. Hasil Peramalan Hybrid SSA-ARIMA.....	51
22. Nilai RMSE model hybrid SSA-SVR dan SSA-ARIMA.....	51
23. Hasil Peramalan SSA-SVR dan SSA-ARIMA	52

DAFTAR GAMBAR

Nomor Urut	Halaman
1. Ilustrasi Support Vector Regression.....	9
2. <i>Intensive zone</i> (a) <i>original input space</i> , dan (b) <i>feature space</i>	10
3. Batas Error dalam SVR.....	12
4. Diagram Alir Penelitian	21
5. Diagram alir metode hybrid SSA-SVR	22
6. Diagram alir metode hybrid SSA-ARIMA	23
7. Plot data Inflasi Indonesia.....	24
8. PACF Komponen Trend Data Simulasi.....	30
9. PACF Komponen Musiman Data Simulasi	30
10. Plot singular value.....	35
11. Plot eigen triple	39
12. Matriks korelasi	40
13. Tren yang di rekonstruksi.....	42
14. Seasonal yang direkonstruksi	42
15. Noise yang direkonstruksi.....	43
16. Matriks Korelasi hasil rekonstruksi.....	43
17. PACF Komponen Tren.....	44
18. PACF Komponen Seasonal	45
19. PACF Komponen Noise	45
20. Plot Data Komponen Noise.....	47
21. Plot ACF Data Komponen Noise.....	48
22. Plot PACF Data Komponen Noise	48
23. Plot perbandingan Data Aktual, Model <i>Hybrid SSA-SVR</i> , dan <i>Hybrid SSA-ARIMA</i>	52

DAFTAR LAMPIRAN

Nomor Urut	Halaman
1. Data Inflasi Indonesia	58
2. Syntax SSA-SVR	59
3. Syntax SSA-ARIMA	67
4. Riwayat Hidup.....	74

BAB I

PENDAHULUAN

1.1 Latar Belakang

Ekonomi makro adalah ilmu yang mempelajari fenomena-fenomena dalam perekonomian secara luas, seperti inflasi, pengangguran, dan pertumbuhan ekonomi (Mankiw, 2018). Permasalahan ekonomi yang banyak terjadi di beberapa negara baik negara maju maupun negara berkembang seperti Indonesia adalah inflasi. Menurut Bank Indonesia (2020) inflasi dapat diartikan sebagai kenaikan harga barang dan jasa secara umum dan terus menerus dalam jangka waktu tertentu. Inflasi yang rendah dan stabil merupakan prasyarat bagi pertumbuhan ekonomi nasional yang berkesinambungan dimana pada akhirnya memberikan manfaat bagi peningkatan kesejahteraan masyarakat. Untuk mengembangkan sistem ekonomi tersebut diperlukan upaya diantaranya adalah pemanfaatan data sebagai salah satu dasar untuk perencanaan ekonomi nasional. Tujuannya adalah memanfaatkan data yang tersedia untuk membantu pengambilan kebijakan dengan harapan pembangunan ekonomi dapat mencapai sasaran secara efisien.

Terjadinya suatu peristiwa terkadang tidak sesuai dengan waktu yang diharapkan untuk terjadi. Kondisi ini menjadi alasan utama mengapa peramalan diperlukan. Hal ini sejalan dengan Fajar & Rachmad (2020) bahwa agar dapat mengambil keputusan atau kebijakan yang tepat, diperlukan peramalan untuk menentukan kapan suatu peristiwa akan terjadi atau diharapkan terjadi.

Peramalan merupakan prediksi terhadap kejadian di masa yang akan datang dengan memperhatikan kejadian masa lalu maupun saat ini. Salah satu model peramalan kuantitatif adalah model runtun waktu. Model runtun waktu berupaya untuk meramalkan kondisi masa yang akan datang dengan menggunakan informasi historis dan mengekstrapolasikan pola tersebut ke masa depan. Metode yang dapat digunakan untuk peramalan antara lain adalah metode *Autoregressive Integrated Moving Average* (ARIMA) (Montgomery et al., 2008).

ARIMA merupakan salah satu metode yang penting dalam melakukan proses peramalan suatu data deret waktu tertentu (Chatfield, 2001). Model ARIMA memiliki asumsi yang harus dipenuhi sebelum penyusunan model dilakukan. Asumsi tersebut diantaranya adalah kestasioneran data terhadap mean dan varians. Namun dalam aplikasinya, tidak semua data dapat memenuhi asumsi tersebut sehingga membutuhkan alternatif lain dalam pemecahannya. Pendekatan alternatif yang berkembang adalah pendekatan non parametrik yang tidak membutuhkan asumsi – asumsi pada ARIMA.

Salah satu metode analisis runtun waktu non-parametrik yang dapat digunakan adalah *Singular Spectrum Analysis* (SSA). Metode SSA bekerja dengan menguraikan data runtun waktu menjadi suatu penjumlahan dari komponen – komponen kecil independen yaitu komponen *trend*, musiman, dan *noise* (Golyandina et al., 2001).

Hasil yang diperoleh dari metode SSA dapat digunakan untuk analisis lebih lanjut. Tujuannya adalah untuk mendapatkan prediksi yang lebih akurat (Vahabie et al., 2007) dan (Zhang et al., 2011). Penggunaan dua metode secara bertahap dikenal dengan metode *hybrid*. Dimana, hasil yang diperoleh dari metode pertama selanjutnya

digunakan pada metode kedua. Metode *hybrid* bertujuan memanfaatkan kelebihan dari masing masing metode untuk mendapatkan kemampuan yang baik dalam hal ini ramalan yang akurat (Muslim, 2017).

Berbagai macam metode *hybrid* dengan melibatkan SSA telah banyak dilakukan, diantaranya Lahmiri (2018) yang meramalkan harga saham menggunakan metode *hybrid* SSA-SVR dengan kesimpulan bahwa metode tersebut memiliki kemampuan peramalan yang lebih baik dibandingkan dengan model tunggal. Penelitian lain dilakukan oleh Xiao et al., (2019) yang juga melakukan pemodelan harga saham menggunakan metode *hybrid* SSA-SVR dan menghasilkan kesimpulan bahwa metode *hybrid* SSA-SVR lebih unggul dibandingkan metode *Ensemble Empirical Mode Decomposition-SVR* karena menghasilkan prediksi lebih tepat. Penelitian lain juga dilakukan oleh (Athoillah et al., 2021) yang melakukan peramalan curah hujan dengan model *hybrid* SSA-SVR yang menghasilkan kesimpulan bahwa prediksi metode SSA-SVR lebih mendekati data aktual dibandingkan dengan beberapa model *nonhybrid/tunggal*.

Metode *hybrid* lain yang melibatkan SSA seperti yang telah dilakukan oleh oleh Arumsari et al., (2021) yang menggunakan metode SSA-ARIMA menyimpulkan bahwa metode tersebut adalah kombinasi metode peramalan yang baik untuk meningkatkan akurasi peramalan dan cocok untuk data ekonomi yang cenderung memiliki pola tren dan musiman seperti inflasi. Penelitian lain juga dilakukan oleh Darmawan et al. (2022) yang melakukan penelitian *hybrid* SSA-ARIMA untuk data musiman yang menghasilkan kesimpulan bahwa metode *hybrid* adalah metode yang akurat karena memiliki nilai MAPE yang kecil. Fajar & Rachmad (2020) juga melakukan penelitian menggunakan model *hybrid* SSA-ARIMA untuk meramalkan pertumbuhan ekonomi dan menghasilkan kesimpulan bahwa metode SSA-ARIMA lebih baik dibandingkan metode SSA dan ARIMA karena memiliki nilai RMSE yang lebih kecil.

Berdasarkan latar belakang tersebut, penulis tertarik untuk membandingkan peramalan metode *hybrid* SSA-SVR dan SSA-ARIMA dalam memprediksi inflasi di Indonesia dengan judul "**Perbandingan Metode *Hybrid* SSA-SVR dan *Hybrid* SSA-ARIMA dalam memprediksi inflasi di Indonesia**".

1.2 Rumusan Masalah

Berdasarkan latar belakang penelitian di atas, rumusan masalah pada penelitian ini adalah:

1. Bagaimana perbandingan Metode *Hybrid* SSA-SVR dan *Hybrid* SSA-ARIMA dalam memprediksi inflasi di Indonesia?
2. Model manakah yang lebih baik dalam memprediksi inflasi Indonesia?

1.3 Batasan Masalah

Penelitian ini dibatasi pada:

1. Data yang digunakan adalah data inflasi Indonesia dalam satuan bulanan periode Januari 2000 sampai dengan Desember 2022.
2. Membandingkan kemampuan prediksi metode *hybrid* SSA-SVR dan *hybrid* SSA-ARIMA.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah sebagai berikut:

1. Membandingkan hasil prediksi Inflasi di Indonesia dengan menggunakan metode *hybrid* SSA-SVR dan *Hybrid* SSA-ARIMA.
2. Menentukan model terbaik hasil prediksi Inflasi di Indonesia.

1.5 Manfaat Penelitian

Manfaat dari penelitian ini adalah sebagai berikut.

1. Memberikan pengetahuan tentang metode *hybrid* SSA-SVR dan *Hybrid* SSA-ARIMA dalam memprediksi inflasi di Indonesia.
2. Memberikan informasi kepada para pelaku ekonomi terkait inflasi di Indonesia.

1.6 Landasan Teori

1.6.1 Analisis Runtun Waktu

Analisis runtun waktu pertama kali diperkenalkan oleh Box & Jenkins, (1976) melalui bukunya yang berjudul *Time Series Analysis: Forecasting and Control*. Analisis runtun waktu merupakan suatu metode peramalan untuk di masa yang akan datang yang dilakukan berdasarkan nilai atau data di masa lalu dari suatu variabel dan kesalahan (*error*) di masa lalu (Makridakis et al., 1983). Runtun waktu (*time series*) merupakan serangkaian data pengamatan yang terjadi berdasarkan indeks waktu secara berurutan dengan interval waktu tetap (Aswi dan Sukarna, 2006).

1.6.2 Autoregressive Moving Average (ARIMA)

Model runtun waktu ARIMA merupakan model deret waktu yang pertama kali diperkenalkan pada tahun 1970-an oleh Box dan Jenkins sehingga ARIMA juga sering disebut metode runtun waktu Box-Jenkins. Model Box-Jenkins terdiri dari model stasioner: *Autoregressive* (AR), *Moving Average* (MA), *Autoregressive - Moving Average* (ARMA), dan model nonstasioner: *Autoregressive Integrated Moving Average* (ARIMA).

Tabel 1 Model runtun waktu AR, MA, dan ARMA

Model	Persamaan
AR (p)	$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p}$
MA (q)	$Y_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots + \theta_q a_{t-q}$
ARMA (p, q)	$\phi_p(B)Y_t = \theta_q(B)a_t$

dengan,

ϕ_p : Parameter model AR orde p

θ_q : Parameter model MA orde q

a_t : Nilai galat pada saat t

B : Operator *backshift*

Dalam analisis runtun waktu ARIMA, terdapat tiga langkah utama yang harus dilakukan dalam memilih model yang cocok untuk melakukan estimasi dan peramalan data yaitu identifikasi model, estimasi parameter, dan uji diagnosis residual (Box et al., 2008).

Pada umumnya, data runtun waktu yang ditemukan cenderung tidak stasioner dalam rata-rata maupun variansi. Sehingga sebelum mengkonstruksi model ARIMA, data deret waktu harus distasionerkan terlebih dahulu menggunakan dua cara, yaitu:

1. Differencing

Proses *differencing* merupakan suatu metode yang bertujuan untuk menstasionerkan data deret waktu yang tidak stasioner dalam rata-rata. Proses *differencing* dapat dilakukan sebanyak d kali dan proses ini memiliki operasi sebagai berikut:

$$(1 - B)^d Y_t = (1 - B)^{d-1} Y_t - (1 - B)^{d-1} Y_{t-1} \quad (1)$$

2. Transformasi Box-Cox

Transformasi Box-Cox adalah suatu metode transformasi data yang dibutuhkan jika data cenderung tidak stasioner dalam variansi. Kestasioneran data deret waktu dalam variansi dapat dilihat berdasarkan plot Box-Cox. Parameter dapat diestimasi menggunakan metode maximum likelihood dan akan dipilih nilai λ yang dapat meminimumkan jumlah kuadrat galat yang mengakibatkan hasil transformasi dari $\{Y_t\}$ akan semakin baik. Formulasi dari transformasi box-cox sebagai berikut:

$$T(Y_t) = \frac{Y_t^\lambda - 1}{\lambda} \quad (2)$$

Tabel 2 Menunjukkan nilai-nilai lambda (λ) beserta jenis transformasi yang umumnya digunakan pada data.

Tabel 2 Nilai lambda (λ) beserta jenis transformasi

Nilai λ	Jenis Transformasi
-1	$\frac{1}{Y_t}$
-0,5	$\frac{1}{\sqrt{Y_t}}$
0	$\ln Y_t$
0,5	$\sqrt{Y_t}$
1	Y_t

Model ARIMA merupakan model ARMA yang telah mengalami proses *differencing* sebanyak d dan dilambangkan sebagai ARIMA (p, d, q) yang dinyatakan sebagai berikut:

$$\phi_p(B)(1 - B)^d Z_t = \theta_q(B) a_t \quad (3)$$

Dengan,

ϕ_p : Parameter model AR orde p

θ_q : Parameter model MA orde q

a_t : Nilai galat pada saat t

B : Operator backshift t

$(1 - B)^d$: operator differencing

Dalam tahap identifikasi model, digunakan 2 fungsi autokorelasi untuk menentukan orde dari model ARIMA (p, d, q), yaitu: *Autocorrelation Function* (ACF) dan *Partial Autocorrelation Function* (PACF). Penentuan orde AR dan MA dilakukan dengan melihat pola dari plot ACF dan PACF tersebut dengan karakteristik sebagaimana Tabel 3 dibawah ini.

Tabel 3 Plot ACF dan PACF

Model	ACF	PACF
AR (p)	Turun cepat membentuk eksponensial	Terpotong setelah lag ke-p
MA (q)	Terpotong setelah lag ke-q	Turun cepat membentuk eksponensial
ARMA (p, q)	Turun cepat membentuk eksponensial	Turun cepat membentuk eksponensial
AR (p) atau MA (q)	Terpotong setelah lag ke-q	Terpotong setelah lag ke-p

1.6.3 Singular Spectrum Analysis (SSA)

Singular Spectrum Analysis (SSA) adalah teknik analisis deret waktu dan peramalan yang menggabungkan unsur analisis klasik *time series*, *multivariate statistics*, *multivariate geometric*, *dynamical system*, dan *signal processing* (Golyandina et al., 2001). SSA merupakan salah satu metode analisis deret waktu nonparametrik yang menguraikan data deret waktu menjadi beberapa komponen yang mengandung pola tertentu seperti tren, musiman, siklus, dan noise.

SSA terdiri dari dua tahap yaitu: Dekomposisi dan Rekonstruksi. Tahap dekomposisi terdiri dari tahapan *Embedding* dan *Singular Value Decomposition* (SVD) sedangkan tahap rekonstruksi terdiri dari tahapan *Grouping* dan *Diagonal Averaging*.

1. Dekomposisi

Dalam tahap dekomposisi terdapat dua langkah, yaitu *embedding* dan *Singular Value Decomposition* (SVD).

a. *Embedding*

Pada tahap *embedding*, data deret waktu diubah ke dalam bentuk matriks lintasan (*trajectory matrix*) yakni mentransformasi data yang berdimensi satu (vektor) menjadi data yang multidimensional (matriks). Misalkan data deret waktu dengan panjang N , tanpa data hilang dinyatakan dengan $\mathbf{X} = \{x_1, x_2, \dots, x_N\}$, data di transformasi ke dalam bentuk matriks berukuran $L \times K$. Parameter L merupakan panjang jendela (*windows length*) bernilai $2 < L < N/2$. Penentuan nilai L optimal mengikuti (Golyandina et al., 2018), salah satunya jika suatu deret waktu diketahui periode musimannya maka L dengan nilai musiman dan kelipatannya dapat memberikan hasil penguraian lebih baik dan K diperoleh $K = N - L + 1$. Dari proses *embedding* ini diperoleh sebanyak K vektor lag waktu

$$X_i = (X_i, X_{i+1}, \dots, X_{i+L-1})^T, 1 \leq i \leq K. \quad (4)$$

Dalam bentuk matriks dapat ditulis sebagai berikut:

$$\mathbf{X} = [X_1, X_2, \dots, X_K] = \begin{bmatrix} x_1 & x_2 & \dots & x_K \\ x_2 & x_3 & \dots & x_{K+1} \\ \vdots & \dots & \ddots & \vdots \\ x_L & x_{L+1} & \dots & x_{K+L-1} \end{bmatrix} \quad (5)$$

Matriks \mathbf{X} disebut juga matriks lintasan dimana semua elemen anti diagonalnya bernilai sama. Jadi pada tahap ini output yang diperoleh adalah matriks lintasan yang berukuran $L \times K$.

b. *Singular Value Decomposition (SVD)*

Langkah kedua dalam dekomposisi adalah membuat *Singular Value Decomposition (SVD)* dari matriks lintasan. SVD bertujuan untuk memperoleh pemisahan komponen dalam dekomposisi dari data deret waktu. SVD pada penerapannya memiliki persamaan dengan Analisis Komponen Utama (*Principal Component Analysis*) yaitu berfungsi untuk mereduksi komponen dari data awal serta mengurangi dimensi. Hasil dari langkah SVD adalah *singular value*, *eigenvector*, dan *principal component* dari matriks lintasan.

Misalkan $\mathbf{S} = \mathbf{X}\mathbf{X}^T$ dan $\lambda_1, \dots, \lambda_L$ adalah nilai eigen dengan $\lambda_1 \geq \dots \geq \lambda_L \geq 0$ dan $U_1 \dots U_L$ adalah vektor eigen dari masing – masing nilai eigen. Misalkan $d = \max \{i, \lambda_i > 0\}$ merupakan pangkat dari matriks \mathbf{X} . Jika $V_i = \mathbf{X}^T U_i / \sqrt{\lambda_i}$ untuk $i = 1 \dots d$ maka SVD dari matriks lintasan \mathbf{X} adalah

$$\begin{aligned} \mathbf{X} &= X_1 + X_2 + \dots + X_d \\ \mathbf{X} &= U_1 \sqrt{\lambda_1} V_1^T + U_2 \sqrt{\lambda_2} V_2^T + \dots + U_d \sqrt{\lambda_d} V_d^T \\ \mathbf{X} &= \sum_{i=1}^d U_i \sqrt{\lambda_i} V_i^T. \end{aligned} \quad (6)$$

Sekumpulan nilai $\sqrt{\lambda_i} U_i V_i^T$ disebut *eigen triple* (ET)

Konsep dasar pada tahap ini adalah mendapatkan barisan matriks dari matriks \mathbf{S} , dimana pada masing-masing matriks dalam barisan tersebut mengandung *eigen triple* yang terdiri atas *eigenvector* U_i , *singular value* $\sqrt{\lambda_i}$ dan komponen utama V_i^T . Komponen utama menggambarkan karakteristik pada masing-masing matriks barisan tersebut. *Eigenvector* adalah vektor kolom bukan nol yang bila dikalikan dengan suatu matriks berukuran $n \times n$ akan menghasilkan vektor lain yang memiliki nilai kelipatan dari *eigenvector* itu sendiri. *Singular value* merupakan akar kuadrat dari *eigenvalue* yang merupakan nilai karakteristik dari suatu matriks berukuran $n \times n$ (Aryani & Maisyitah, 2015).

2. Rekonstruksi

Pada tahap rekonstruksi menggunakan parameter *grouping effect* (r). *Grouping effect* (r) merupakan parameter kedua dalam SSA yang berperan penting dalam menentukan pola pada plot data. Setelah penggunaan parameter L dalam dekomposisi, pada hasil SVD

akan menyajikan serangkaian seri awal yang telah dipisahkan dengan baik (Sakinah, 2012). Dalam tahap rekonstruksi terdapat dua langkah yang harus dilakukan, yaitu langkah *Grouping* kemudian dilanjutkan dengan pembentukan deret rekonstruksi berdasarkan hasil yang diperoleh pada langkah *diagonal averaging*.

a. *Grouping*

Pada langkah ini akan dilakukan pengelompokan hasil dekomposisi matriks lintasan yang berukuran $L \times K$ dengan tujuan untuk memisahkan komponen aditif SVD ke dalam beberapa sub kelompok, yaitu tren, musiman, dan noise (Darmawan, 2016). Pengelompokan dilakukan dengan melihat plot nilai singular (akar ciri), plot vector ciri dan plot korelasi terboboti. Pengelompokan akan memisahkan himpunan indeks $\{1, \dots, d\}$ menjadi m himpunan bagian saling lepas I_1, \dots, I_m . Misalkan $I_m = \{i_1, \dots, i_p\}$ dihasilkan dari kelompok I yang bersesuaian maka $\mathbf{X}_{I_m} = \mathbf{X}_{I_1} + \dots + \mathbf{X}_{I_p}$. dengan melakukan prosedur yang sama untuk setiap kelompok I_1, \dots, I_m diperoleh

$$\mathbf{X} = \mathbf{X}_{I_1} + \dots + \mathbf{X}_{I_m} \quad (7)$$

Prosedur pemilihan himpunan I_1, \dots, I_m ini disebut pengelompokan *eigentriple*.

b. *Diagonal Averaging*

Setelah melakukan grouping, tahap selanjutnya akan dilakukan transformasi dari hasil pengelompokan ke dalam deret baru dengan panjang N . Tujuan dari tahap ini adalah mendapatkan singular value dari komponen-komponen yang telah dipisahkan, kemudian akan digunakan dalam peramalan. Misalkan terdapat matriks \mathbf{Y} :

$$\mathbf{Y} = \begin{bmatrix} y_{11} & y_{12} & \dots & y_{1,K} \\ y_{21} & y_{22} & \dots & y_{2,k} \\ \vdots & \vdots & \ddots & \vdots \\ y_{L,1} & y_{L,2} & \dots & y_{L,K} \end{bmatrix} \quad (8)$$

Jika \mathbf{Y} adalah matriks $L \times K$ dengan elemen y_{ij} , $1 \leq i \leq L$, $1 \leq j \leq K$ maka $L^* = \min(L, K)$ dan $K^* = \max(L, K)$, dan $N = L + K - 1$. jika $L < K$ maka $y_{ij}^* = y_{ij}$ dan sebaliknya jika $L > K$, $y_{ij}^* = y_{ji}$.

Diagonal Averaging diperoleh dengan persamaan berikut:

$$y_k = \begin{cases} \frac{1}{k} \sum_{m=1}^k y_{m,k-m+1}^* & ; 1 \leq k \leq L^* \\ \frac{1}{L^*} \sum_{m=1}^{L^*} y_{m,k-m+1}^* & ; L^* \leq k \leq K^* \\ \frac{1}{N-k+1} \sum_{m=1}^{N-K^*+1} y_{m,k-m+1}^* & ; K^* \leq k \leq N \end{cases} \quad (9)$$

Ini bersesuaian dengan merata-ratakan elemen matriks di sepanjang antidiagonal $i + j = k + 1$ sebagai contoh misalkan $k = 1$ maka $y_1 = y_{1,1}$, untuk $k = 2$ maka $y_2 = (y_{1,2}, y_{2,1})/2$.

Diagonal averaging pada persamaan (9) diterapkan untuk semua komponen matriks X_{I_j} pada persamaan (7) menghasilkan sebuah series $\tilde{X}^{(k)} = (\tilde{x}_1^{(k)}, \tilde{x}_2^{(k)}, \dots, \tilde{x}_N^{(k)})$. Sehingga series x_1, x_2, \dots, x_N didekomposisi menjadi sebuah penjumlahan dari m series rekonstruksi:

$$y_n = \sum_{k=1}^m \tilde{x}_n^{(k)}, n = 1, 2, \dots, N \quad (10)$$

3. Peramalan *Singular Spectrum Analysis*

Terdapat dua metode peramalan dalam *Singular Spectrum Analysis*, yaitu metode rekuren (*R-Forecasting*) dan metode vektor (*V-Forecasting*). Metode rekuren adalah metode dasar yang sering digunakan karena relatif lebih mudah (Golyandina et al., 2001). Metode vektor merupakan hasil modifikasi dari metode rekuren. Perbedaan antara *R-Forecasting* dan *V-Forecasting* adalah peramalan dengan *Rforecasting* melakukan kontinuitas secara langsung (dengan bantuan LRF), sedangkan peramalan dengan *V-forecasting* berhubungan dengan *L-continuation*. Hal ini menyebabkan dalam *approximate continuation*-nya biasanya memberikan hasil yang berbeda (Golyandina et al., 2001). Dalam peramalan *Singular Spectrum Analysis*, model dibangun dengan bantuan *Linear Recurrent Formula* (LRF) dari deret waktu hasil rekonstruksi. Misalkan ditentukan M titik waktu ke depan untuk peramalan, maka model ramalannya:

$$g_i = \begin{cases} y_i & ; i = 1, \dots, N \\ \sum_{j=1}^{L-1} a_j g_{i-j} & ; i = N + 1, \dots, N + M \end{cases} \quad (11)$$

Penduga koefisien LRF (\hat{a}_j) diperoleh menggunakan vektor ciri pada tahapan SVD. Misalkan $U = (u_1, u_2, \dots, u_{L-1}, u_L)^T$, $U^* = (u_1, u_2, \dots, u_{L-1})^T$, dan $\pi_i = u_L$ adalah komponen terakhir vektor U . Vektor penduga koefisien LRF $R = (a_{L-1}, \dots, a_1)^T$ diperoleh persamaan

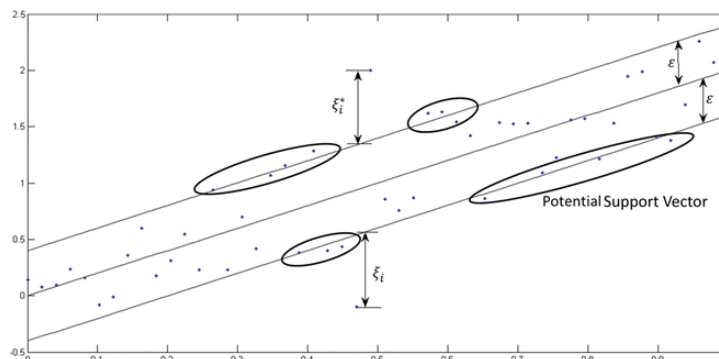
$$R = \frac{1}{1 - v^2} \sum_{i=1}^r \pi_i U_i^* \quad (12)$$

Dengan $v^2 = \sum_{i=1}^r \pi_i$, dan r adalah parameter effect grouping (banyaknya ET yang masuk ke dalam tren dan musiman)

1.6.4 *Support Vector Regression (SVR)*

Support Vector Regression (SVR) pertama kali diperkenalkan oleh Vladimir N Vapnik pada tahun 1995. *SVR* merupakan pengembangan dari *Support Vector Machine (SVM)* yang digunakan untuk memecahkan masalah regresi. *SVR* dapat menghasilkan nilai prediksi yang baik dan bisa mendapatkan suatu fungsi dengan tingkat kesalahan yang kecil karena *SVR* mempunyai kemampuan dalam mengatasi masalah *overfitting*. *Overfitting* merupakan kondisi dimana suatu model mencoba untuk meregresikan setiap

titik data training secara akurat sehingga biasanya model tersebut akan menangkap data *noise* (data yang memuat *error*) yang seharusnya diabaikan. Kondisi ini akan mengakibatkan hasil prediksi yang baik pada data *training*, namun apabila menggunakan data *testing* akan diperoleh hasil prediksi yang buruk (Smola & Scholkopf, 2004). Konsep dari SVR adalah agar sebanyak mungkin data masuk ke dalam suatu zona, dengan tetap meminimalkan nilai epsilon (ϵ). Ilustrasi SVR dapat dilihat pada Gambar 1 (Awad & Khanna, 2015).



Gambar 1 Ilustrasi Support Vector Regression

Gambar 1 menunjukkan sebuah *hyperplane* atau garis yang terletak tepat pada tengah-tengah antara garis batas (-) dan garis batas (+). Terlihat ada beberapa titik-titik (*data points*) yang dilingkari yang menjadi *potential support vector* atau titik-titik yang dapat menjadi pembatas, sehingga titik data dapat masuk ke dalam satu zona dengan sebisa mungkin tetap meminimumkan nilai epsilon (ϵ). Pada Gambar 1, epsilon (ϵ) merupakan jarak antara *hyperplane* dengan garis batas (-) dan garis batas (+).

Kelebihan dari penggunaan metode SVR dibandingkan dengan metode regresi linear (*linear regression*) yaitu pada regresi linear menghasilkan fungsi linear berupa garis lurus, sedangkan pada algoritma SVR menghasilkan tren data yang bergelombang mengikuti jalur data yang terbentuk, sehingga mampu menghasilkan prediksi yang lebih akurat (Cahyono et al., 2019). Misalkan diberikan m data set *training* $(x_i, y_i), i = 1, 2, \dots, m$, dimana x_i adalah vector input $x = \{x_1, x_2, \dots, x_m\} \subset \mathbb{R}^n$, y_i adalah vektor output $y = \{y_1, y_2, \dots, y_m\} \subset \mathbb{R}$, dan m adalah banyaknya data *training*. Tujuan dari SVR adalah menemukan suatu fungsi $f(x)$ sebagai *hyperplane* berupa fungsi regresi yang mempunyai deviasi maksimum sebesar ϵ atau yang mempunyai *error* kurang dari sama dengan ϵ untuk mendapatkan nilai output y_i dari semua data *training* dan pada saat yang sama juga mencari fungsi yang menghasilkan batas yang minimum. Semua *error* (selisih antara output fungsi dengan nilai output aktual) akan diterima asalkan nilainya kurang dari ϵ , sedangkan nilai kesalahan tidak diterima apabila nilainya melebihi ϵ (Smola & Scholkopf, 2004). Misal $f(x)$ adalah fungsi linier dalam bentuk berikut:

$$f(x) = \mathbf{w}^T \mathbf{x} + b, \quad \mathbf{w} \in \mathbb{R}^n, \mathbf{x} \in \mathbb{R}^n, b \in \mathbb{R} \quad (13)$$

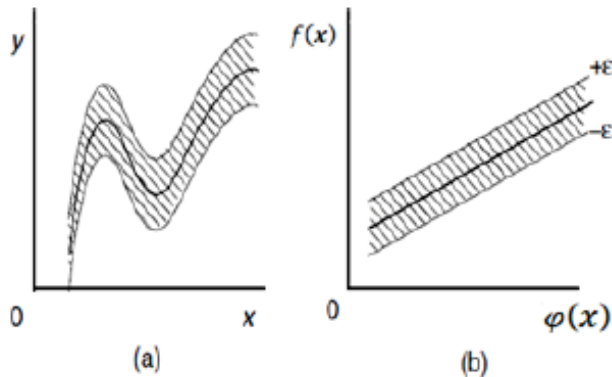
Dengan,

$f(x)$: Fungsi regresi SVR sebagai nilai prediksi output

w^T : Vektor pembobot
 x : Vektor input
 b : Bias berupa konstanta/intersep model

Menurut (Amanda et al., 2014) pada regresi terdapat residual, misalkan residual (r) didefinisikan dengan mengurangkan output scalar y terhadap estimasi $f(x)$ yaitu $r = y - f(x)$ dengan:

$$E(r) = \begin{cases} 0, & \text{untuk } |r| \leq \varepsilon \\ |r| - \varepsilon, & \text{untuk yang lain.} \end{cases}$$



Gambar 2 Intensive zone (a) original input space, dan (b) feature space

$f(x) = w^T x + b = \pm \varepsilon$ merupakan jarak terjauh *support vector* dari *hyperplane* yang disebut sebagai margin. Memaksimumkan margin akan meningkatkan peluang data ke dalam radius $\pm \varepsilon$. Jarak dari *hyperplane* $f(x) = w^T x + b = 0$ ke data (x, y) yaitu $\frac{|D(x, y)|}{\|w^*\|}$, dimana:

$$w^* = \begin{pmatrix} 1 \\ -w^T \end{pmatrix} = (1 - w^T)^T$$

dengan,

w : Vektor pembobot

w^* : Vektor pembobot baru

Diasumsikan bahwa jarak maksimum data terhadap *hyperplane* adalah δ . Estimasi yang ideal akan terpenuhi dengan:

$$\frac{|D(x, y)|}{\|w^*\|} \leq \delta$$

$$|D(x, y)| \leq \delta \|w^*\|$$

$$\delta \|w^*\| \leq \varepsilon$$

Oleh karena itu, supaya δ maksimum, diperlukan $\|w^*\|$ yang minimum. Optimasi penyelesaian masalah dilakukan dengan bentuk *quadratic programming*:

$$\min \frac{1}{2} \|w\|^2 \tag{14}$$

dengan batasan:
$$\begin{cases} y_i - w^T x_i - b \leq \varepsilon, i = 1, 2, \dots, m \\ w^T x_i + b - y_i \leq \varepsilon, i = 1, 2, \dots, m \end{cases}$$

dengan:

y_i : nilai aktual periode ke- i
 w^T : vektor pembobot
 x_i : data input periode ke- i
 b : bias berupa konstanta/intersep model
 ε : deviasi atau derajat toleransi terhadap *error*

Faktor $\|w\|^2$ dinamakan regularisasi. Meminimumkan $\|w\|^2$ akan membuat fungsi setipis (*flat*) mungkin, sehingga bisa mengontrol kapasitas fungsi (*function capacity*). Asumsikan bahwa terdapat fungsi f yang dapat mendekati semua titik (x_i, y_i) dengan presisi ε . Dalam kasus ini diasumsikan bahwa semua titik berada dalam margin atau dalam rentang $f(x) \pm \varepsilon$ (*feasible*). Pada kondisi *infeasible constraints*, masalah optimasi ditunjukkan dengan adanya beberapa titik yang keluar dari rentang $f(x) \pm \varepsilon$. Penyelesaian masalah optimasi pada kondisi *infeasible* ini yaitu dengan menambahkan variable *slack* ξ dan ξ^* (Santosa, 2007). Formulasi dari permasalahan optimasi setelah ditambahkan variable *slack* yaitu sebagai berikut:

$$\min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^m (\xi + \xi^*) \quad (15)$$

dengan batasan:
$$\begin{cases} y_i - w^T x_i - b \leq \varepsilon + \xi_i & ; i = 1, 2, \dots, m \\ w^T x_i + b - y_i \leq \varepsilon + \xi_i^* & ; i = 1, 2, \dots, m \\ \xi & ; \xi^* \geq 0 \end{cases}$$

dengan,

y_i : nilai aktual periode ke- i
 w^T : vektor pembobot
 x_i : data input periode ke- i
 b : bias berupa konstanta/intersep model
 ε : deviasi atau derajat toleransi terhadap *error*
 ξ, ξ^* : *slack variables*

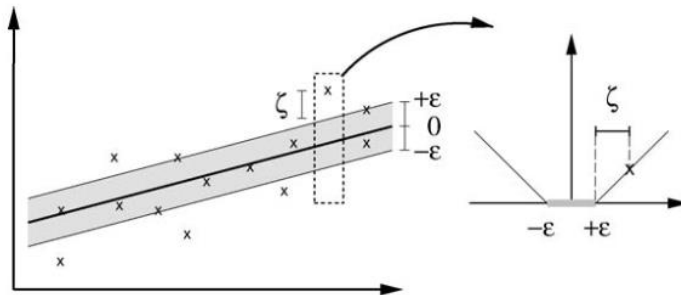
Fungsi yang menunjukkan hubungan antara *error*, apakah *error* tersebut dikenai penalti atau tidak disebut *loss function*. Perbedaan *loss function* akan menghasilkan formula SVR yang berbeda (Santosa, 2007). *Loss function* yang paling sederhana yaitu ε -*insensitive loss function* dengan formulasi matematikanya sebagai berikut:

$$L_\varepsilon(y) = \begin{cases} 0, & \text{untuk } |f(x) - y| < \varepsilon \\ |f(x) - y| - \varepsilon, & \text{untuk yang lain} \end{cases} \quad (16)$$

L_ε : ε -*insensitive loss function*,
 E : deviasi atau derajat toleransi terhadap *error*.

Konstanta C (*cost*) > 0 menentukan tawar menawar (*trade-off*) antara ketipisan fungsi $f(x)$ dengan batas atas deviasi yang lebih besar dari ε yang masih ditoleransi atau dengan kata lain konstanta C merupakan nilai penalti jika deviasi lebih besar dari batas *error*. Semua titik yang berada di luar batas ε akan dikalikan dengan C . Parameter C

berfungsi untuk menentukan berapa besar deviasi dari parameter ε (epsilon) yang masih bisa ditoleransi, yaitu semakin besar nilai C maka nilai toleransi yang masih diijinkan juga semakin besar. Sementara itu, parameter ε fungsi untuk mengontrol lebar zona regresi dalam pemrosesan data. Semakin besar nilai ε maka estimasi regresi semakin datar mendekati regresi linear. Smola dan Scholkopf (2004) mengilustrasikan *Support Vector Regression* dengan ε -insensitive loss function seperti pada Gambar 3.



Gambar 3 Batas Error dalam SVR

Gambar 3 menggambarkan situasi secara grafis penambahan *variable slack* ξ dan ξ^* . Variabel *slack* yang terletak dalam selang ε akan bernilai nol, kemudian semakin jauh jarak *variable slack* terhadap batas ε maka nilainya semakin besar. *Support vector* adalah titik-titik dalam data training yang terletak tepat garis batas dan di luar batas ε . Semakin kecil nilai ε , maka semakin tinggi nilai *variable slack*, akurasi semakin tinggi, dan jumlah *support vector* semakin banyak sehingga memberi pengaruh yang lebih besar pada fungsi optimasi. Sebaliknya, semakin besar nilai ε , maka semakin kecil nilai variabel *slack*, akurasi semakin rendah, dan jumlah *support vector* semakin sedikit.

Selain itu, SVR juga melakukan perhitungan terhadap nilai error. Jika nilai error berada dalam batas garis $\pm\varepsilon$ maka nilai error tersebut akan bernilai 0, sedangkan jika nilai error tersebut berada di luar batas $\pm\varepsilon$ maka nilai error dihitung sebagai $error - \varepsilon$. Apabila terdapat titik di luar margin $\pm\varepsilon$ atau di luar tabung(tube) harus ditambahkan variabel *slack*, sehingga nilai error tersebut merupakan $error - \varepsilon$ dan dikenai penalti sebesar C .

Masalah optimasi pada persamaan (15) dapat diselesaikan lebih mudah dengan bentuk formulasi dualnya menggunakan perkalian Lagrange (*Lagrange multipliers*) dengan satu pengali Lagrange untuk setiap kendala, yaitu sebagai berikut (Smola & Scholkopf, 2004):

$$L(w, b, \xi, \xi^*, \alpha_i, \alpha_i^*, \eta_i, \eta_i^*) = \frac{1}{2} \|w\|^2 + C \left(\sum_{i=1}^m (\xi + \xi^*) \right) - \sum_{i=1}^m \eta_i \xi_i + \eta_i^* \xi_i^* - \sum_{i=1}^m \alpha_i (\varepsilon + \xi_i - y_i + (w^T x_i + b)) - \quad (17)$$

$$\sum_{i=1}^m \alpha_i^* (\varepsilon + \xi_i^* - y_i + (w^T x_i + b))$$

Dengan L_p dinamakan fungsi lagrange untuk primal problem, $\eta_i, \eta_i^*, \alpha_i, \alpha_i^* \geq 0$ adalah *Lagrange Multiplier* yang digunakan untuk menyelesaikan permasalahan optimasi dengan kendala (*constrained optimization*), di mana kendala tersebut akan dikonversi menjadi masalah optimasi tanpa kendala (*unconstrained optimization*). Pencarian nilai parameter *hyperplane* yang optimal ditentukan dengan menggunakan Karush Kuhn Tucker (KKT) yang menjadikan turunan dari variabel L_p terhadap w, b, ξ, ξ^* sama dengan nol sebagai berikut: (Hong, 2009)

$$\frac{\partial L_p}{\partial b} = \sum_{i=1}^m (\alpha_i - \alpha_i^*) = 0, \quad (18)$$

$$\frac{\partial L_p}{\partial b} = w - \sum_{i=1}^m (\alpha_i - \alpha_i^*) x_i = 0 \Rightarrow w = \sum_{i=1}^m (\alpha_i - \alpha_i^*) x_i, \quad (19)$$

$$\frac{\partial L_p}{\partial \xi_i} = C - \alpha_i - \eta_i = 0 \Rightarrow \eta_i = C - \alpha_i, \quad (20)$$

$$\frac{\partial L_p}{\partial \xi_i^*} = C - \alpha_i^* - \eta_i^* = 0 \Rightarrow \eta_i^* = C - \alpha_i^*, \quad (21)$$

Formulasi dual untuk persamaan (17) diperoleh dengan mensubsitusikan persamaan (18), (19), (20), (21) ke dalam persamaan primal problem pada persamaan (17), sehingga diperoleh formulasi dual untuk masalah optimasi dari SVR sebagai berikut:

$$\begin{aligned} \max L_d(\alpha_i, \alpha_i^*) = & -\frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) x_i^T x_j \\ & -\varepsilon \sum_{i=1}^m (\alpha_i + \alpha_i^*) + \sum_{i=1}^m y_i (\alpha_i - \alpha_i^*) \end{aligned} \quad (22)$$

Dengan batasan:

$$\begin{aligned} \sum_{i=1}^m y_i (\alpha_i - \alpha_i^*) &= 0, \\ 0 \leq \alpha_i \leq C, i &= 1, 2, \dots, m, \\ 0 \leq \alpha_i^* \leq C, i &= 1, 2, \dots, m. \end{aligned}$$

Representasi formulasi dual yang dihasilkan merupakan bentuk *Quadratic programming* juga, namun memiliki kendala yang lebih sederhana (*boundconstrained optimization*). Penyelesaian masalah *primal* dan *dual* pada persamaan dan (17) dan (22) memberikan solusi yang sama karena merupakan masalah optimasi *convex*. Setelah

menyelesaikan *dual problem* diperoleh solusi optimal dari kendala estimasi parameter w dalam bentuk koefisien *Lagrange* α_i dan α_i^* yang dapat ditulis kembali sebagai:

$$w = \sum_{i=1}^m (\alpha_i + \alpha_i^*) x_i \quad (23)$$

sehingga diperoleh *optimal hyperplane* diformulasikan sebagai berikut:

$$f(x_i) = \sum_{i=1}^m (\alpha_i - \alpha_i^*) x_i^T x_j + b \quad (24)$$

Kompleksitas fungsi yang direpresentasikan dengan *support vector* hanya bergantung pada banyaknya *support vector*. Algoritma lengkap *support vector* dapat diuraikan dalam perkalian skalar (*dot product*) antar data. Pada saat mengevaluasi fungsi $f(x)$, maka tidak perlu menghitung w secara eksplisit.

Menurut Santosa (2007) pada umumnya data dalam dunia nyata kebanyakan bersifat *non-linear*, sehingga jarang yang bersifat *linear separable*. SVR untuk kasus *non-linear* memberikan pendekatan alternatif dengan cara melakukan pemetaan data x dari *input space* ke *feature space* dengan dimensi yang lebih tinggi melalui suatu fungsi φ sehingga $\varphi: x \mapsto \varphi(x)$.

Akan tetapi, transformasi φ pada umumnya tidak diketahui dan sulit untuk dipahami, sehingga akan kesulitan dalam pemetaannya. Menurut Vapnik (1995) masalah ini dapat diatasi dengan *kernel trick* yaitu perkalian skalar (*dot product*) $\varphi^T(x_i)\varphi(x_j)$ dalam *feature space* dapat digantikan dengan fungsi kernel, yaitu:

$$K(x_i, x_j) = \varphi^T(x_i)\varphi(x_j) \quad (25)$$

dengan fungsi kernel tersebut mampu mendefinisikan secara implisit transformasi φ . *Kernel trick* memberikan berbagai kemudahan, untuk menemukan titik – titik *support vector* dalam proses pembelajaran SVM, cukup dengan mengetahui fungsi kernel yang digunakan, sehingga tidak perlu mengetahui wujud dari fungsi non linear φ (Nugroho, 2003).

Fungsi regresi SVR untuk kasus *non-linear* pada persamaan (24) dapat dituliskan kembali sebagai berikut:

$$f(x_i) = \sum_{i=1}^m (\alpha_i - \alpha_i^*) K(x_i, x_j) + b. \quad (26)$$

dengan:

- α_i, α_i^* : Pengali lagrange,
- $K(x_i, x_j)$: fungsi kernel
- b : konstanta

1. Fungsi Kernel

Dalam aplikasinya, hubungan antara ruang input dan output biasanya nonlinier. Oleh karena itu, korelasi data ditemukan melalui metode pemetaan nonlinier yang dipetakan oleh SVR ruang input ke ruang dimensi yang lebih tinggi (ruang fitur) menggunakan fungsi kernel. Fungsi kernel yang sering digunakan pada model SVR antara lain kernel linier, kernel polinomial, dan kernel RBF (Danandeh Mehr et al., 2019). Rumus fungsi kernel tersebut sebagai berikut:

Tabel 4 Fungsi Kernel

Jenis Kernel	Fungsi
Linear	$K(x_i, x_j) = (x_i^T - x_j)$
Polynomial	$K(x_i, x_j) = (x_i \cdot x_j + 1)^p$
Radial Basis Function (RBF)	$K(x_i, x_j) = \exp(-\gamma(x_i - x_j)^2)$

2. Grid Search Optimization

Salah satu langkah menentukan parameter optimal dari model SVR adalah menggunakan *Grid Search Optimization*. *Grid Search Optimization* ini membagi jangkauan parameter yang akan di optimalkan kedalam grid dan melintasi semua titik untuk mendapatkan parameter yang optimal (Yasin et al., 2014). *Grid search* akan melatih beberapa pasang model dan akhirnya akan memilih pasangan dengan rata-rata error terkecil.

Dalam aplikasinya, *Grid Search Optimization* biasanya diukur dengan cross validation pada data training. *Cross validation* adalah pengujian standar yang dilakukan untuk memprediksi error rate. Data *training* dibagi secara random kedalam beberapa bagian dengan perbandingan yang sama kemudian error rate dihitung bagian demi bagian, selanjutnya hitung rata-rata seluruh error rate untuk mendapatkan error rate secara keseluruhan (Santosa, 2007).

1.6.5 Penentuan Model Terbaik

Penentuan model terbaik menggunakan *Root Mean Square Error* (RMSE), Model terbaik memiliki nilai RMSE yang mendekati 0. Nilai sisaan digunakan untuk mengetahui besarnya simpangan nilai dugaan terhadap nilai aktual. Perhitungan RMSE mengikuti persamaan (Wei, 2006):

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}}$$

Dengan,

n : Jumlah data

y_i : Nilai sebenarnya pada data ke-i

\hat{y}_i : Nilai prediksi pada data ke-i

1.6.6 Inflasi

Inflasi diartikan sebagai kenaikan jumlah uang beredar atau kenaikan likuiditas dalam suatu perekonomian. Pengertian tersebut mengacu pada gejala umum yang ditimbulkan oleh adanya kenaikan jumlah uang beredar yang diduga telah menyebabkan adanya kenaikan harga-harga. Dalam perkembangan lebih lanjut, inflasi secara singkat dapat diartikan sebagai suatu kecenderungan meningkatnya harga-harga barang dan jasa secara umum dan terus-menerus.

Pengukuran perubahan inflasi dari waktu ke waktu umumnya menggunakan suatu angka indeks. Angka indeks tersebut disusun dengan memperhitungkan sejumlah barang dan jasa yang akan digunakan untuk menghitung besarnya angka inflasi. Kelompok barang dan jasa yang dipilih tersebut diberi bobot sesuai dengan intensitas penggunaannya oleh masyarakat. Dengan demikian perubahan harga barang dan jasa yang memiliki bobot besar pada penghitungan inflasi akan memberikan dampak yang lebih besar.

Perkembangan kenaikan harga sejumlah barang dan jasa secara umum dalam suatu periode waktu ke waktu tersebut disebut sebagai laju inflasi (inflation rate). Laju inflasi pada umumnya dinyatakan dalam angka persentase (%). Laju inflasi dapat terjadi pada tingkat yang ringan, sedang, berat, dan hiperinflasi. Menurut Wikipedia, inflasi ringan terjadi apabila kenaikan harga berada di bawah 10%; inflasi sedang antara 10 - 30%; dan inflasi berat antara 30 -100% per tahun; dan hiperinflasi atau inflasi tidak terkendali terjadi apabila kenaikan harga berada di atas 100% setahun. Namun demikian, angka-angka inflasi tersebut pada umumnya bersifat relatif dan tidak ada suatu standar yang umum. (Suseno & Astiyah, 2009)

Laju inflasi dihitung berdasarkan angka indeks yang disusun dari harga barang dan jasa yang dikonsumsi oleh masyarakat dan disebut sebagai Indeks Harga Konsumen (IHK). Cara perhitungannya adalah dengan membandingkan IHK pada tahun dasar dengan IHK periode pengamatan. IHK di Indonesia dihitung menggunakan formula Laspeyres yang dimodifikasi sebagai berikut: (Badan Pusat Statistik, 2022).

$$IHK_n = \frac{\sum_{i=1}^k \frac{P_{ni}}{P_{(n-1)i}} P_{(n-1)i} \cdot Q_{oi}}{\sum_{i=1}^k P_{oi} \cdot Q_{oi}} \times 100$$

Dengan,

IHK_n : Indeks Harga Konsumen bulan ke-n

P_{ni} : Harga jenis barang/jasa i pada bulan ke (n)

$P_{(n-1)i}$: Harga jenis barang/jasa i pada bulan ke (n-1)

$P_{(n-1)i} \cdot Q_{oi}$: Nilai Konsumsi (NK) jenis barang/jasa i pada bulan ke (n-1)

$\frac{P_{ni}}{P_{(n-1)i}}$: Relatif harga (RH) jenis barang/jasa i pada bulan ke (n)

$P_{oi} \cdot Q_{oi}$: Nilai konsumsi (NK) jenis barang/jasa i pada tahun dasar

k : Jumlah jenis barang/jasa yang tercakup dalam paket komoditas IHK

Presentase perubahan indeks atau laju inflasi atau deflasi setiap bulan diperoleh dengan mengurangkan indeks suatu bulan dengan indeks bulan sebelumnya, kemudian hasilnya dibagi dengan indeks bulan sebelumnya dan dikalikan 100. Laju inflasi atau deflasi tersebut juga dapat dihitung dari indeks suatu bulan dibagi dengan indeks bulan

sebelumnya, hasilnya dikurangi dengan 1, dan dikalikan dengan 100. Perhitungan di atas dijabarkan dengan rumus sebagai berikut: (Badan Pusat Statistik, 2022).

$$Inf_n = \frac{IHK_n - IHK_{(n-1)}}{IHK_{(n-1)}} \times 100$$

Atau

$$Inf_n = \left(\frac{IHK_n}{IHK_{(n-1)}} - 1 \right) \times 100$$

Dengan ,

Inf_n : Laju inflasi/deflasi bulan ke (n)

IHK_n : Indeks Harga Konsumen bulan ke (n)

$IHK_{(n-1)}$: Indeks Harga Konsumen bulan ke (n-1)

Selain dihitung berdasarkan IHK, inflasi juga dapat dihitung berdasarkan Indeks Biaya Hidup (IBH), Indeks Harga Produsen (IBP), Indeks Harga Perdagangan Besar (IHPB), atau juga dapat dihitung dengan deflator Produk Domestik Bruto (PDB deflator). Indeks Harga Produsen (IHP) merupakan indeks yang mengukur perubahan harga yang diterima produsen domestik untuk barang yang mereka hasilkan. IHP mengukur tingkat harga yang terjadi pada level produsen. Indeks Harga Perdagangan Besar (IHPB) mengukur perubahan harga yang terjadi antara penjual/pembeli besar pertama dan pembeli/penjual besar berikutnya dalam jumlah besar pada pasar pertama. Dibeberapa negara termasuk Indonesia, IHPB merupakan indikator yang menggambarkan pergerakan harga dari komoditas-komoditas yang diperdagangkan di suatu daerah. PDB deflator mengukur perubahan harga dalam perekonomian secara keseluruhan. Cakupan perubahan harga yang diukur dalam PDB deflator lebih luas dibandingkan IHK maupun IHPB. Angka deflator tersebut dihitung dengan membandingkan PDB nominal pada suatu tahun tertentu dengan PDB pada tahun tertentu yang ditetapkan (tahun dasar). Deflator PDB menunjukkan besarnya perubahan harga dari semua barang baru, barang produksi lokal, barang jadi dan jasa. (Suseno & Astiyah, 2009)

Secara umum, Inflasi di Indonesia dapat disebabkan oleh berbagai faktor baik dari sisi permintaan (*Demand Pull Inflation*), penawaran (*Cost Push Inflation*), maupun dari ekspektasi.

1. *Demand Pull Inflation* timbul karena adanya permintaan yang tidak diimbangi dengan kondisi peningkatan tingkat produksi, akibatnya sesuai hukum permintaan, bila permintaan banyak sementara penawaran tetap maka harga akan naik. Dan bila hal ini berlangsung terus menerus, akan menyebabkan inflasi yang berkepanjangan.
2. *Cost Push Inflation* adalah kenaikan inflasi karena meningkatnya biaya input atau biaya faktor produksi. Akibat naiknya biaya produksi tersebut, ada dua hal yang dapat dilakukan oleh produsen, yaitu dengan menaikkan harga produknya dengan jumlah penawaran yang sama atau harga produk naik karena penurunan jumlah produksi.
3. Ekspektasi inflasi adalah faktor yang dipengaruhi oleh persepsi dan harapan masyarakat serta pelaku ekonomi terhadap tingkat inflasi di masa depan. Faktor ini dapat mempengaruhi keputusan konsumen, investor, dan pelaku ekonomi lainnya.

Kontribusi dari masing-masing faktor yang mempengaruhi inflasi tersebut tidak selalu sama dari waktu ke waktu. Dengan kata lain, faktor utama yang mempengaruhi inflasi bisa berbeda dari waktu ke waktu. Oleh karena itu, Pemerintah dan Bank Indonesia menempuh tujuh langkah strategis untuk mengendalikan inflasi, yaitu: (Bank Indonesia, 2021)

1. Melaksanakan kebijakan moneter dan fiskal yang konsisten dengan upaya mendukung pengendalian inflasi dan mendorong pertumbuhan ekonomi;
2. Mengendalikan inflasi kelompok *Volatile Food* agar dapat terkendali di bawah 5%, dengan fokus pada komoditas beras, aneka cabai, dan aneka bawang;
3. Menjaga ketersediaan pasokan dan kelancaran distribusi pangan untuk memitigasi risiko jangka pendek, termasuk mengantisipasi pergeseran musim panen dan peningkatan permintaan menjelang Hari Besar Keagamaan Nasional (HBKN);
4. Memperkuat ketahanan pangan melalui upaya peningkatan produktivitas dan hilirisasi pangan;
5. Memperkuat ketersediaan data pasokan pangan untuk mendukung perumusan kebijakan pengendalian inflasi;
6. Memperkuat sinergi Tim Pengendalian Inflasi Pusat dan Daerah (TPIP-TPID) antara lain melalui Gerakan Nasional Pengendalian Inflasi Pangan (GNPIP); serta
7. Memperkuat komunikasi untuk menjaga ekspektasi inflasi.

Bagi masyarakat, dalam menghadapi dampak inflasi dan menjaga daya beli, upaya yang dapat dilakukan diantaranya adalah:

1. Mengelola anggaran dengan baik;
2. Menghemat energi dan sumberdaya;
3. Investasi dan tabungan;
4. Meningkatkan keterampilan dan pengetahuan;
5. Diversifikasi pendapatan/pendapatan tambahan;
6. Hidup sederhana dan mandiri.

BAB II

METODE PENELITIAN

2.1 Sumber Data

Data yang digunakan dalam penelitian ini diperoleh dari Badan Pusat Statistik Indonesia. Periode data yang diteliti adalah data bulanan dari Januari 2000 hingga Desember 2022.

2.2 Identifikasi Variabel

Variabel yang digunakan dalam penelitian ini adalah inflasi. Inflasi didefinisikan sebagai tingkat kecenderungan naiknya harga barang dan jasa secara umum secara terus menerus dalam jangka waktu tertentu.

2.3 Tahapan Analisis Data

A. Eksplorasi dan Pre-processing data

1. Melakukan eksplorasi data inflasi untuk mengetahui karakteristik data inflasi.
2. Membagi data inflasi menjadi dua bagian, yaitu data *training* dan data *testing* dengan proporsi 95% dan 5%. Data *training* dimulai dari tahun 2000-2021 berjumlah 264 data sedangkan data *testingnya* dari tahun 2022 yang berjumlah 12 data. Data *training* digunakan untuk menyusun model SSA-SVR dan SSA-ARIMA, sedangkan data *testing* digunakan untuk menguji kemampuan dari model.

B. Tahapan Analisis Pemodelan *hybrid* SSA-SVR

1. *Embedding*, menentukan parameter *window length* (L) dengan rentang $2 < L < \frac{N}{2}$ untuk membentuk matriks lintasan X bagi data *training*.

$$X = [X_1, X_2, \dots, X_K] = \begin{bmatrix} x_1 & x_2 & \dots & x_K \\ x_2 & x_3 & \dots & x_{K+1} \\ \vdots & \vdots & \vdots & \vdots \\ x_L & x_{L+1} & \dots & x_{K+L-1} \end{bmatrix}$$

2. Menentukan nilai L optimal mengikuti Golyandina *et al.* (2018) dengan proses *trial and error*. Setelah diperoleh nilai L optimal maka terbentuk matriks lintasan X berukuran $L \times K$, dengan $K = N - L + 1$.
3. Penguraian dengan metode singular value decomposition (SVD). Menguraikan matriks lintasan X dari tahapan *embedding* dengan metode SVD ke dalam bentuk,

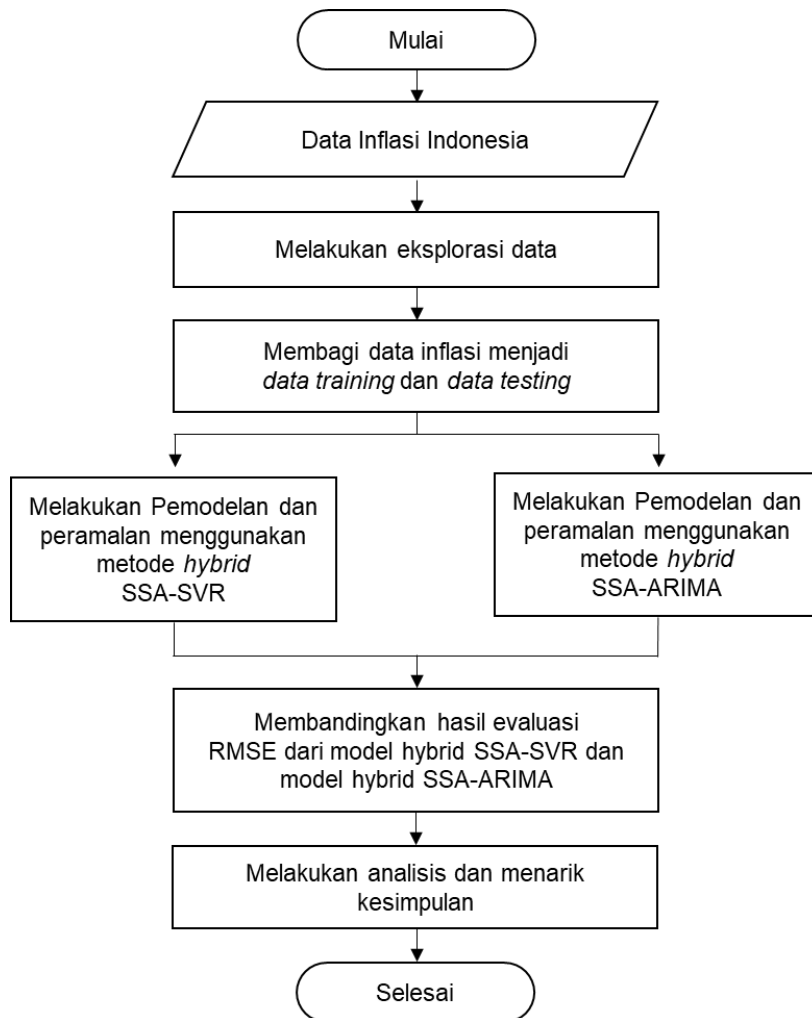
$$X = U\Sigma V^T = \sum_{i=1}^L \sqrt{\lambda_i} U_i V_i^T$$

Dengan λ_i ($i = 1, \dots, L$) adalah nilai eigen dari matriks $S = XX^T$. Dalam metode SVD $\sqrt{\lambda_i}$ disebut nilai singular, sedangkan vector U_i dan V_i adalah vector singular kiri dan kanan. Kumpulan dari $\sqrt{\lambda_i}$, U_i , V_i disebut *eigen triple*.

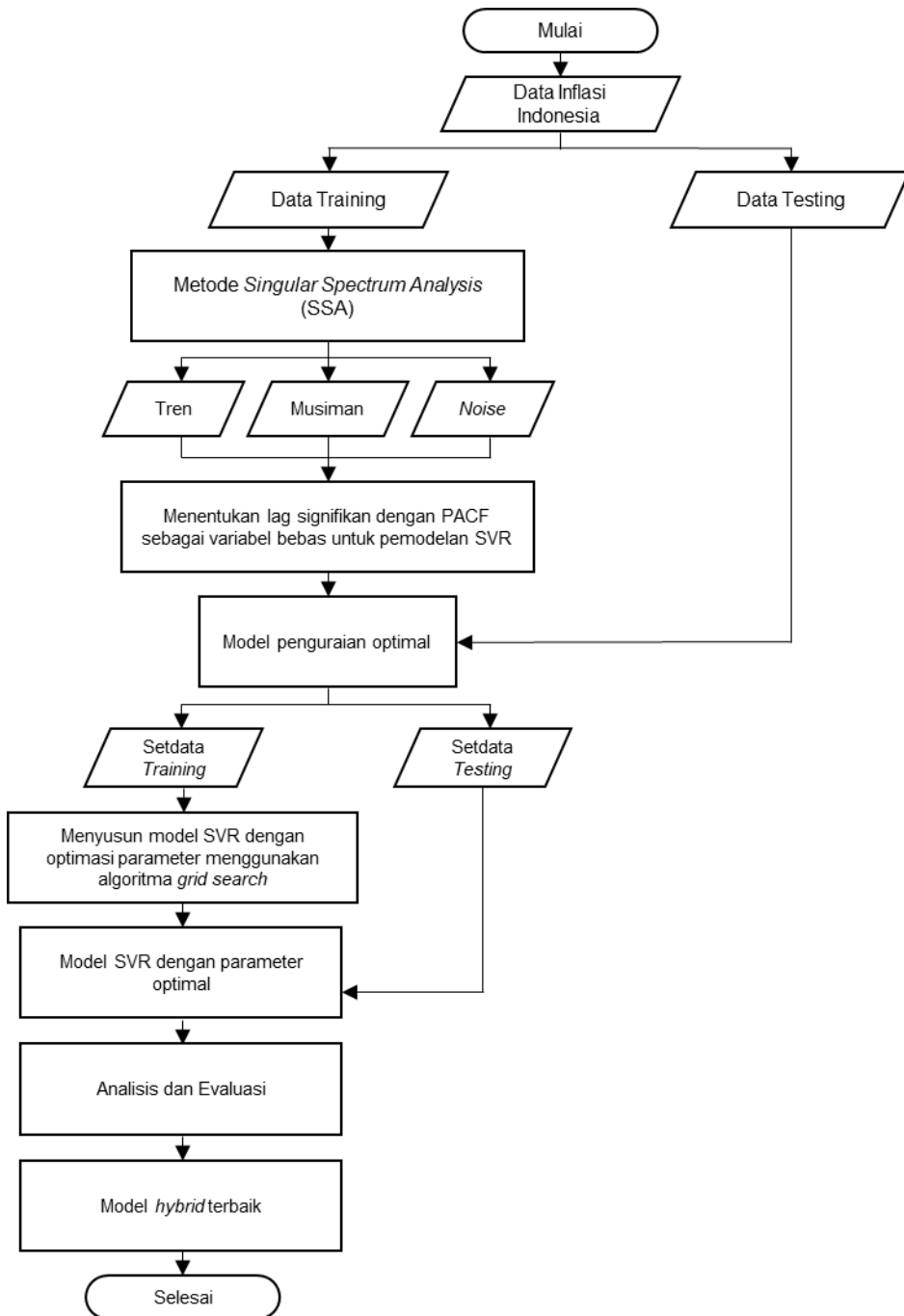
4. *Grouping*, mengelompokkan *eigen triple* ke dalam komponen tren, musiman, dan *noise* dengan cara identifikasi plot nilai singular dan vektor eigen.

5. *Diagonal averaging* akan merekonstruksi matriks hasil pengelompokan menjadi data deret waktu tren, musiman dan *noise*.
 6. Menguji komponen *noise* yang dihasilkan metode SSA pada data training *hybrid* apakah termasuk *white noise* menggunakan plot PACF dan uji *Ljung-Box*.
 7. Menentukan *lag* waktu yang signifikan dari setiap komponen hasil penguraian dengan plot PACF yang akan digunakan sebagai variabel bebas untuk model SVR.
 8. Melakukan *pre-processing* pada data testing sebagaimana pada langkah 1 sampai 7.
 9. Membuat setdata untuk pemodelan SVR dengan variable bebas diperoleh dari gabungan *lag* waktu yang signifikan tiap komponen sebagaimana pada langkah 7. Sedangkan variable terikat diperoleh dari data lag waktu yang menyesuaikan dengan variabel bebas. Hasil penggabungan data ini selanjutnya disebut setdata training.
 10. Melakukan hal yang sama pada langkah 9 dengan menggunakan data testing (langkah 8) yang selanjutnya disebut sebagai setdata testing.
 11. Menyusun model SVR menggunakan setdata training dengan fungsi kernel Linear, Polynomial, dan RBF dengan validasi silang 10 lipat. Optimasi parameter yang digunakan yaitu algoritma *grid search* dengan rentang parameter yang dioptimasi adalah *gamma* (2^{-8} sampai 2^3), *cost* (2^{-5} sampai 2^8), dan *epsilon* (0,1; 0,01; 0,001).
 12. Melakukan pengujian model SVR pada langkah 11 menggunakan setdata *testing*.
 13. Menghitung nilai RMSE.
- C. Pemodelan *Hybrid* SSA-ARIMA
1. Melakukan pemodelan SSA seperti pada tahapan pemodelan SSA-SVR (B1 – B5)
 2. Melakukan analisis *time series* dengan ARIMA menggunakan data komponen *noise* dari hasil rekonstruksi metode SSA, dengan tahapan sebagai berikut :
 - a) Mengecek stasioneritas data komponen *noise*,
 - b) Membuat plot ACF dan PACF untuk identifikasi bentuk model ARIMA.
 - c) Mengestimasi parameter untuk melihat hasil uji signifikansi parameter berdasarkan model yang terpilih.
 - d) Melakukan pemeriksaan diagnostik untuk *white noise* dan normalitas
 - e) Jika diperoleh lebih dari satu model ARIMA, maka dilakukan pemilihan model terbaik dengan menggunakan kriteria AIC.
 - f) Melakukan peramalan dengan model ARIMA.
 3. Melakukan peramalan model *Hybrid* SSA-ARIMA dengan cara menggabungkan hasil peramalan SSA dan ARIMA
 4. Menghitung akurasi peramalan dengan menggunakan RMSE dari model *hybrid* SSA-ARIMA.

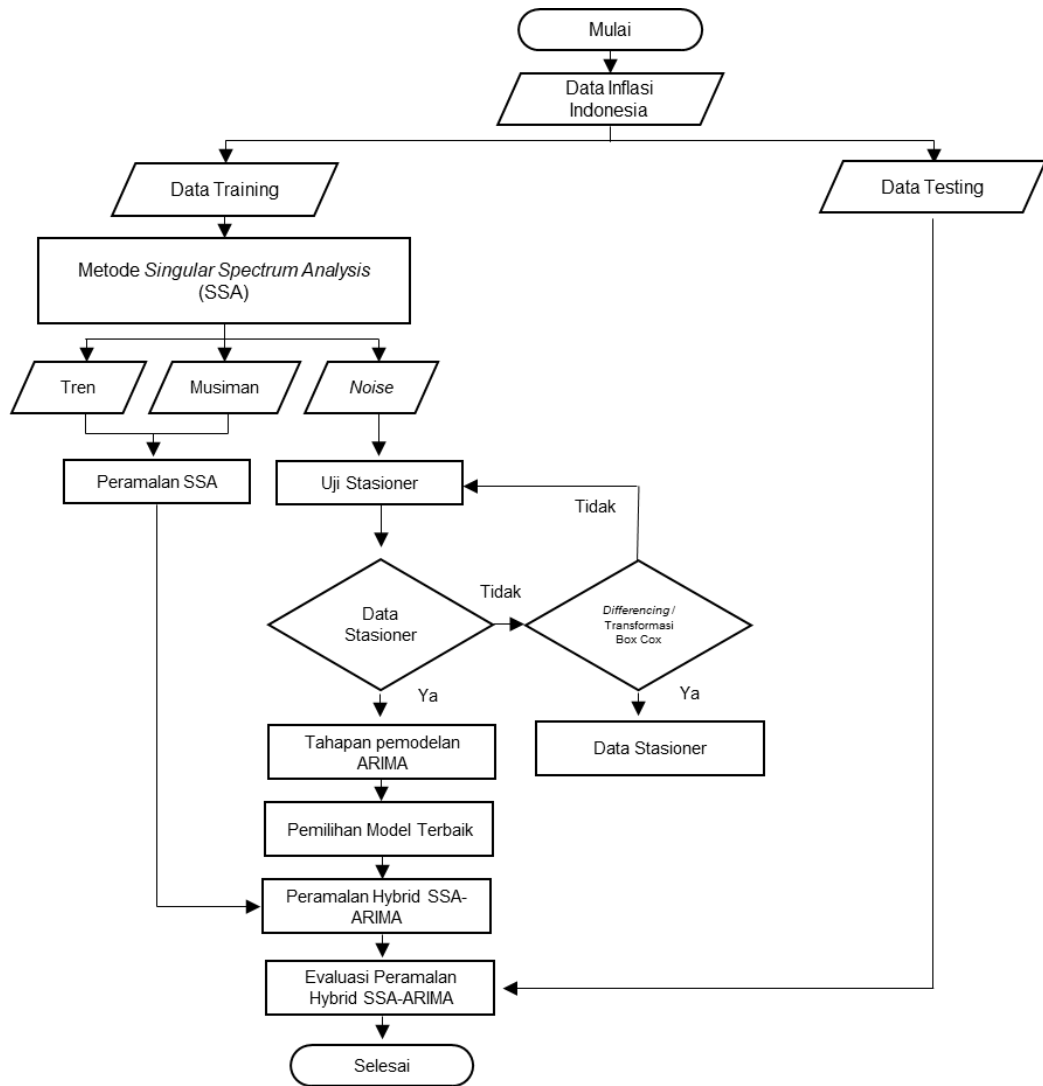
2.4 Diagram Alir



Gambar 4 Diagram Alir Penelitian



Gambar 5 Diagram alir metode *hybrid* SSA-SVR



Gambar 6 Diagram alir metode *hybrid* SSA-ARIMA