PENERAPAN METODE *HYBRID AUTOREGRESSIVE INTEGRATED MOVING AVERAGE* DENGAN *SUPPORT VECTOR REGRESSION* (ARIMA-SVR) PADA HARGA *ETHEREUM*

(Skripsi)

Oleh

INSYAFIATUL M.D ASTUTI

2217031102

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2026

ABSTRACT

APPLICATION OF THE HYBRID AUTOREGRESSIVE INTEGRATED MOVING AVERAGE WITH SUPPORT VECTOR REGRESSION (ARIMA-SVR) METHOD TO ETHEREUM PRICES

By

INSYAFIATUL M.D ASTUTI

The hybrid Autoregressive Integrated Moving Average with Support Vector Regression (ARIMA-SVR) method is a time series forecasting method capable of capturing both linear and nonlinear patterns simultaneously.This study aims to apply the ARIMA-SVR hybrid model to forecast Ethereum prices and determine the accuracy of the ARIMA-SVR hybrid model obtained for Ethereum prices. The data used consists of Ethereum closing prices from December 11, 2020, to December 10, 2025, the study began by dividing the data into training and testing sets using three data partitioning schemes is 70%:10%, 80%:20%, dan 90%:10%. The results indicate that the best model is the ARIMA(2,1,2)-SVR with optimal parameters C = 0.8125, γ = 5, and ε = 0.125, under the 90% training dan 10% testing data split. The accuracy of the ARIMA(2,1,2)-SVR hybrid model is demonstrated by the Mean Absolute Percentage Error (MAPE) values obtained, which are 2.83% for the training data and 2.72% for the testing.

Keywords : ARIMA, SVR, Hybrid ARIMA-SVR, Ethereum

ABSTRAK

PENERAPAN METODE *HYBRID AUTOREGRESSIVE INTEGRATED MOVING AVERAGE* DENGAN *SUPPORT VECTOR REGRESSION* (ARIMA-SVR) PADA HARGA ETHEREUM

Oleh

INSYAFIATUL M.D ASTUTI

Metode *hybrid Autoregressive Integrated Moving Average* dengan *Support Vector Regression* (ARIMA-SVR) merupakan salah satu metode untuk peramalan deret waktu yang mampu menangkap pola linear dan nonlinear secara bersamaan. Penelitian ini bertujuan menerapkan model *hybrid* ARIMA-SVR untuk meramalkan harga *Ethereum* dan mengetahui akurasi model *hybrid* ARIMA-SVR yang diperoleh pada harga *Ethereum*. Data yang digunakan yaitu data harga penutupan *Ethereum* pada rentang waktu 11 Desember 2020 sampai 10 Desember 2025, penelitian dimulai dengan membagi data menjadi data *training* dan data *testing* dengan tiga skema pembagian data yaitu 70%:10%, 80%:20%, dan 90%:10%. Hasil penelitian menunjukkan model terbaik yaitu ARIMA(2,1,2)-SVR dengan parameter terbaik C sebesar 0.8125, parameter γ sebesar 5, dan parameter ε sebesar 0.125 pada skema pembagian data 90% data *training* dan 10% data *testing*. Akurasi model *hybrid* ARIMA(2,1,2)-SVR ditunjukkan oleh nilai *Mean Absolute Percentage Error* (MAPE) yang diperoleh yaitu 2.83% untuk data *training* dan 2.72% untuk data *testing*.

Kata Kunci : ARIMA, SVR, *Hybrid* ARIMA-SVR, *Ethereum*

PENERAPAN METODE *HYBRID AUTOREGRESSIVE INTEGRATED MOVING AVERAGE* DENGAN *SUPPORT VECTOR REGRESSION* (ARIMA-SVR) PADA HARGA *ETHEREUM*

Oleh

INSYAFIATUL M.D ASTUTI

Skripsi

Sebagai Salah Satu Syarat Untuk Mencapai Gelar
SARJANA MATEMATIKA

Pada

Jurusan Matematika
Fakultas Matematika dan Ilmu Pengetahuan Alam

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2026

Judul Skripsi	:	**PENERAPAN METODE *HYBRID AUTOREGRESSIVE INTEGRATED MOVING AVERAGE* DENGAN *SUPPORT VECTOR REGRESSION* (ARIMA-SVR) PADA HARGA *ETHEREUM***
Nama Mahasiswa	:	Insyafiatul M.D Astuti
NPM	:	**2217031102**
Jurusan / Program Studi	:	**Matematika**
Fakultas	:	**Matematika dan Ilmu Pengetahuan Alam**

MENYETUJUI,

1. Komisi Pembimbing

Drs. Nusyirwan, M.Si.	**Dina Eka Nurvazly, S.Pd., M.Si.**
NIP. 196610101992031028	NIP. 199311062019032018

2. Ketua Jurusan Matematika

Dr. Aang Nuryaman, S.Si., M.Si.
NIP. 197403162005011001

MENGESAHKAN

1. Tim Penguji

Ketua	: **Drs. Nusyirwan, M.Si.**	……………..
Sekretaris	: **Dina Eka Nurvazly, S.Pd., M.Si.**	……………..
Penguji Bukan Pembimbing	: **Widiarti, S.Si., M.Si.**	……………..

2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung

Dr. Eng Heri Satria, S.Si., M.Si.
NIP. 197110012005011002

Tanggal Lulus Ujian Skripsi: 15 Juni 2026

PERNYATAN SKRIPSI MAHASISWA

Saya yang bertanda tangan di bawah ini:

Nama Mahasiswa	: **Insyafiatul M.D Astuti**
Nomor Pokok Mahasiswa	: **2217031102**
Jurusan	: **Matematika**
Fakultas	: **Matematika dan Ilmu Pengetahuan Alam**
Judul Skripsi	: **Penerapan Metode *Hybrid Autoregressive Integrated Moving Averange* dengan *Support Vector Regression* (ARIMA-SVR) pada Harga *Ethereum***

Dengan ini menyatakan bahwa skripsi ini adalah hasil pekerjaan saya sendiri. Apabila di kemudian hari terbukti bahwa skripsi ini merupakan hasil salinan atau dibuat oleh orang lain, maka saya bersedia menerima sanksi sesuai dengan ketentuan akademik yang berlaku.

Bandar Lampung, 15 Juni 2026

Penulis

Insyafiatul M.D Astuti

NPM 2217031102

RIWAYAT HIDUP

Penulis memiliki nama lengkap Insyafiatul M.D Astuti yang lahir di Margoyoso pada tanggal 14 Mei 2004. Penulis merupakan anak sulung dari tiga bersaudara.

Penulis menempuh pendidikan di TK Aisyah Margoyoso pada tahun 2008-2010, kemudian melanjutkan pendidikan Sekolah di SDN 1 Margoyoso pada tahun 2010-2016. Penulis melanjutkan pendidikan Sekolah Menengah Pertama di SMP Syubbanul Wathon Tegalrejo Magelang pada tahun 2016-2019, lalu melanjutkan Sekolah Menengah Atas di SMA Syubbanul Wathon Tegalrejo Magelang pada tahun 2019-2022. Penulis diterima sebagai mahasiswa program studi S1 Matematika dan Ilmu Pengetahuan Alam (FMIPA), Universitas Lampung pada tahun 2022 melalui jalur Seleksi Bersama Masuk Perguruan Tinggi Negeri (SBMPTN).

Pada awal tahun 2025, penulis melaksanakan Kuliah Kerja Nyata (KKN) sebagai pengabdian masyarakat selama 30 hari di Desa Padan, Kecamatan Penengahan, Kabupaten Lampung Selatan. Penulis melaksanakan Kerja Praktik (KP) di PT. Kereta Api Indonesia (KAI) Persero Divre IV Tanjungkarang selama 40 hari yang dilaksanakan pada bulan Mei sampai Juli 2025.

KATA INSPIRASI

"Sungguh atas kehendak Allah semua ini terwujud, tiada kekuatan kecuali dengan pertolongan Allah."

(Q.S Al-Kahfi: 39)

"Maka sesungguhnya beserta kesulitan ada kemudahan, sesungguhnya beserta kesulitan itu ada kemudahan."

(Q.S Al-Insyirah: 5-6)

"I will no longer go full throttle. Instead I will walk slowly enjoying every steps of the journey. At my own pace, following my own rhythms."

(Mark Lee)

"if what you want to do doesn't work, keep going until you can."

(Na Jaemin)

"Don't change, let go. Be brave, be strong. But take it easy, no rush I know you got it. Go, fail forward. Go get your voice heard, don't fold. Find yourself again into the UNKNOWN."

(NCT Dream)

PERSEMBAHAN

Segala puji dan syukur atas kehadirat Allah SWT, atas limpahan nikmat, rahmat, serta karunia-Nya sehingga skripsi ini dapat selesai dengan baik di waktu yang tepat. Shalawat beserta salam selalu tercurah kepada junjungan kita Nabi Muhammad SAW. Saya persembahkan rasa terima kasih dengan penuh rasa syukur dan bahagia kepada:

Bapak dan Ibu Tersayang

Terima kasih kepada kedua orang tua saya atas segala kasih sayang, pengorbanan, do'a, ridho, dan dukungan selama proses saya menjadi pribadi yang menjadi lebih baik setiap harinya. Terima kasih untuk segala ungkapan syukur yang ditanamamkan, sehingga selalu merasa cukup.

Dosen Pembimbing dan Pembahas

Terima kasih kepada dosen pembimbing dan pembahas atas waktu, ilmu, dan dukungan dalam proses penyelesaian skripsi.

Orang-orang Terdekat

Terima kasih atas dukungan, semangat, waktu, dan tenaga kalian yang senantiasa menemani.

Almameter Tercinta

Universitas Lampung

SANWACANA

Alhamdullillah, puji dan syukur penulis panjatkan kepada Allat SWT atas limpahan rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi ini yang berjudul "Penerapan Metode *Hybrid Autoregressive Integrated Moving Average* dengan *Support Vector Regression* (ARIMA-SVR) pada Harga Penutupan *Ethereum*" dengan baik dan lancar serta tepat pada waktu yang telah ditentukan. Shalawat serta salam semoga senantiasa tercurahkan kepada Nabi Muhammad SAW.

Dalam proses penyusunan skripsi ini, banyak pihak yang telah membantu memberikan bimbingan, dukungan, arahan, motivasi serta saran sehingga skripsi ini dapat terselesaikan. Oleh karena itu, dalam kesempatan ini penulis mengucapkan terima kasih kepada:

1. Bapak Drs. Nusyirwan, M.Si. selaku Pembimbing I yang banyak meluangkan waktu untuk memberi bimbingan, dan saran kepada penulis selama proses skripsi.
2. Ibu Dina Eka Nurvazly, S.Pd., M.Si. selaku Pembimbing II yang telah memberikan arahan, saran, serta dukungan sampai penulis dapat menyelesaikan semua proses skripsi sampai akhir.
3. Ibu Widiarti, S.Si., M.Si. selaku Penguji yang telah bersedia memberikan saran, serta masukan yang membangun sehingga skripsi ini selesai.
4. Bapak Dr. Aang Nuryaman, S.Si., M.Si. selaku dosen Pembimbing Akademik dan Ketua Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung yang banyak memberi nasehat selama masa perkuliahan.

5. Bapak Dr. Eng. Heri Satria, S.Si., M.Si. selaku dekan Fakultas Matematika dan Ilmu Pengetahua Alam Universitas Lampung.
6. Seluruh dosen, staff, dan karyawan Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahua Alam Universitas Lampung.
7. Bapak dan Ibu serta kedua adik yang sangat saya sayangi dan cintai. Terima kasih atas segala bentuk dukungan, kasih sayang, nasehat dan do'a yang tidak pernah putus serta selalu mengiringi langkah penulis setiap harinya.
8. Teman-teman "Subleg" yaitu Putri, Elsa, dan Elisabeth yang senantiasa menerima ajakan main, selalu menghibur dengan lelucon-lelucon sederhana, dan menemani selama perkuliahan sampai skripsi ini selesai.
9. Teman-teman SMP dan SMA terutama Natiq, dan Naila yang dengan suka hati memberi dukungan, mendengarkan segala *yapping*, serta memberi banyak nasehat kepada penulis.
10. 7*Dream* terutama Mark Lee dan Na Jaemin yang banyak memberi motivasi melalui karya-karya yang indah.
11. Semua orang terdekat yang tidak bisa disebutkan satu persatu, terima kasih atas segala dukungan, kebersamaan, dan pelajaran berharga yang diberikan melalui pertemuan singkat serta obrolan menyenangkan.

Semoga skripsi ini memberikan banyak manfaat untuk kita semua. Penulis menyadari bahwa skripsi ini masih sangat jauh dari kata sempurna, sehingga kritik dan saran yang membangun sangat diharapkan guna menyempurnakan skripsi ini.

Bandar Lampung, 15 Juni 2025
Penulis

Insyafiatul M.D Astuti
2217031102

DAFTAR ISI

DAFTAR TABEL

Tabel Halaman

DAFTAR GAMBAR

I. PENDAHULUAN

1.1 Latar Belakang

Time series merupakan rangkaian data yang dikumpulkan berdasarkan urutan waktu (Montgomery et al., 2008). Salah satu metode yang banyak digunakan dalam analisis deret waktu adalah *Autoregressive Integrated Moving Average* (ARIMA), yang menggabungkan proses *Autoregressive* (AR), *Moving Average* (MA), dan *differencing (Integrated)* untuk mengatasi data tidak stasioner. Model ARIMA efektif untuk peramalan jangka pendek, namun akurasinya menurun ketika diterapkan pada data dengan pola nonlinear (Purnama, 2021). Penelitian oleh Karnadi & Prastyo (2023) untuk model ARIMA menunjukkan bahwa nilai MAPE dan RMSE pada data *testing Bitcoin, Ethereum,* dan *Binance* meningkat signifikan dibanding data *training*, sedangkan ARIMAX hanya memberikan sedikit perbaikan dibanding dengan metode ARIMA, oleh karena itu, diperlukan pendekatan yang mampu menangani pola data kompleks, seperti model *hybrid* yang menggabungkan ARIMA dengan metode nonlinear seperti *Support Vector Regression* (SVR).

Support Vector Regression (SVR) merupakan pengembangan dari *Support Vector Machine* (SVM) yang membangun fungsi regresi berupa *hyperplane* dengan meminimalkan kesalahan (Vapnik, 2000). Menurut Vapnik (2000) SVR bertujuan untuk menemukan fungsi yang mampu memprediksi data dengan galat sekecil mungkin melalui pemetaan data ke ruang berdimensi tinggi menggunakan fungsi nonlinear serta menggunakan prinsip *Structural Risk Minimization* untuk menentukan *hyperplane* terbaik sehingga memperoleh model yang akurat. SVR membutuhkan parameter yang relatif sedikit namun tetap mampu menghasilkan solusi yang optimal dan akurasi tinggi pada data nonlinear (Levis & Papageorgiou, 2005), sehingga keterbatasan yang dimiliki ARIMA dapat diselesaikan menggunakan metode SVR. Atmaja & Hakim (2022) melakukan penerapan model SVR untuk meramalkan harga penutupan Solana.

Zhang (2001) pertama kali mengenalkan model *hybrid* untuk peramalan deret waktu dengan menggabungkan *Autoregressive Integrated Moving Average* (ARIMA) dan *Artificial Neural Network* (ANN). Model *hybrid* memiliki keunggulan karena mampu menangkap pola linear dan nonlinear secara bersamaan, sehingga menghasilkan akurasi yang lebih baik. Model *hybrid* ARIMA-SVR merupakan salah satu model *hybrid* dalam analisis deret waktu yang dapat menangkap pola linear dan nonlinear secara bersamaan.

Berdasarkan penelitian yang dilakukan oleh Purnama (2021) pada harga emas selama pandemi COVID-19, model *hybrid* ARIMA-SVR memperoleh akurasi yang lebih baik dibandingkan penggunaan model ARIMA. Gusthvi dkk., (2022) menerapkan model *hybrid* ARIMA-SVR untuk melakukan analisis pada harga *Bitcoin* juga menghasilkan akurasi yang lebih baik dibanding menggunakan model ARIMA tunggal. Sihombing dkk., (2022) juga melakukan penelitian menggunakan metode ARIMA-SVR yang menunjukkan nilai MAPE model *hybrid* lebih baik dibanding ARIMA dalam peramalan Indeks Harga Saham Gabungan (IHSG). Penelitian yang dilakukan Albarr & Kusumawati (2023) menunjukkan keunggulan model *hybrid* ARIMA-SVR dalam peramalan saham PT Unilever Indonesia Tbk (UNVR.JK), PT Perusahaan Gas Negara Tbk (PGAS.JK), dan PT Telekomunikasi Indonesia Tbk (TLKM.JK) dibandingkan model *hybrid* ARIMA-GARCH maupun SVR tunggal.

Perkembangan teknologi digital ikut mendorong meningkatnya pengguna mata uang kripto sebagai alat investasi. *Ethereum* merupakan salah satu *cryptocurrency* terbesar yang digunakan dalam *blockchain* dan memiliki volatilitas harga yang tinggi (Yunizar dkk., 2023). Fluktuasi harga yang tajam menyebabkan aktivitas jual beli aset kripto seperti *Ethereum* memiliki resiko yang sangat besar dan penuh ketidakpastian (Dani & Syauqi, 2023), hal ini menunjukkan pentingnya peramalan harga *Ethereum* untuk pertimbangan dalam pengambilan keputusan untuk melakukan transaksi. Peramalan dapat dilakukan menggunakan model yang mampu menangani karakteristik data kompleks, oleh karena itu, penelitian ini menerapkan model *hybrid* ARIMA-SVR pada harga penutupan *Ethereum* dan melakukan peramalan menggunakan data harian *Ethereum*.

1.2 Tujuan Penelitian

Adapun tujuan yang diharapkan dari penelitian ini yaitu:

1. Menerapkan model *hybrid* ARIMA-SVR untuk meramalkan harga *Ethereum*.
2. Mengetahui akurasi model *hybrid* ARIMA-SVR yang diperoleh pada harga penutupan *Ethereum*.

1.3 Manfaat penelitian

Manfaat yang dapat diharapkan dari penelitian ini yaitu:

1. Memberikan informasi kepada investor, dan pelaku pasar terkait hasil peramalan harga *Ethereum* menggunakan model ARIMA-SVR.
2. Sumber referensi bagi peneliti selanjutnya dalam menggunakan metode *hybrid* ARIMA-SVR.
3. Menambah sumber ilmu pengetahuan bagi penulis dan pembaca.

II. TINJAUAN PUSTAKA

2.1 Analisis Deret Waktu

Deret waktu merupakan data hasil pengamatan yang dikumpulkan secara berurutan berdasarkan rentang waktu tertentu (Box et al., 2016), hal ini menunjukkan bahwa data deret waktu memiliki keterkaitan antara peristiwa masa kini dengan masa lalu yang diukur selama periode tertentu. Analisis deret waktu bertujuan untuk memahami pola perubahan data dari waktu ke waktu serta meramalkan nilai di masa mendatang dengan menggunakan data historis dan faktor lain yang berhubungan (Cryer & Chan, 2008). Menurut Makridakis et al., (1983) terdapat empat jenis pola deret waktu yang dapat dipertimbangkan dalam pemilihan metode analisis yang tepat yaitu:

1. Pola horizontal adalah pola data yang berfluktuasi sekitar nilai rata-rata yang konstan atau tetap.
2. Pola musiman adalah pola yang muncul akibat pengaruh faktor musiman yang berulang secara teratur.
3. Pola siklis yaitu pola data yang menunjukkan perubahan jangka panjang sebagai dampak dari siklus ekonomi.
4. Pola tren yaitu pola yang menggambarkan arah perubahan jangka panjang berupa peningkatan atau penurunan secara konsisten seiring waktu.

2.2 Stasioneritas Data

Stasioneritas data merupakan asumsi terpenting yang harus dipenuhi dalam analisis deret waktu (Cryer & Chan, 2008). Menurut Montgomery et al., (2008) deret waktu dapat dikatakan stasioner apabila karakteristik statistiknya, seperti rata-rata, varians, tetap konstan terhadap waktu atau dengan kata lain, perubahan waktu pengamatan tidak memengaruhi sifat-sifat statistik data. Ada dua macam stasioneritas dalam analisis deret waktu yaitu stasioner terhadap ragam dan stasioner terhadap rata-rata.

2.2.1 Stasioneritas terhadap Ragam

Menurut Wei (2006) data yang belum memenuhi asumsi stasioner terhadap ragam maka perlu ditransformasi menggunakan *Box-Cox*. Model transformasi *Box-Cox* dikenalkan pada tahun 1964 oleh George Box dan David Cox. Bentuk transformasi ini yaitu:

$$T(Z_t) = \frac{Y_t^{\lambda}-1}{\lambda} \tag{2.1}$$

dengan:

Y_t = pengamatan periode ke-t

λ = parameter transformasi

Tabel 1. memberikan contoh beberapa nilai λ yang umum dipakai beserta bentuk transformasi *Box-Cox* yang sesuai (Wei, 2006):

Tabel 1. Nilai Transformasi *Box-Cox*

Nilai λ	**-1.0**	**-0.5**	**0**	**0.5**	**1.0**
Transformasi	$\frac{1}{Y_t}$	$\frac{1}{\sqrt{Y_t}}$	$\ln Y_t$	$\sqrt{Y_t}$	Y_t

2.2.2 Stasioneritas terhadap Rata-Rata

Uji akar unit *Augmented Dickey Fuller* (ADF) digunakan untuk menentukan data stasioner terhadap rata-rata (Wei, 2006). Persamaan uji ADF yaitu:

$$\Delta Y_t = \phi Y_{t-1} + \sum_{j=1}^{p=1} \alpha_j \, \Delta Y_{t-j} + e_t \tag{2.2}$$

dengan:

ϕ = parameter

p = panjang *lag*

α_j = koefisien *lag* ke-*j*

e_t = galat pada periode ke-*t*

Uji stasioner menggunakan uji *Augmented Dickey Fuller* (ADF) melalui persamaan dan hipotesis sebagai berikut (Wei, 2006).

1. Uji Hipotesis

   $H_0 : \phi \geq 0$ (data tidak stasioner)

   $H_1 : \phi < 0$ (data stasioner)

2. Taraf Signifikan

   $\alpha = 5$ % atau 0,05

3. Daerah Kritis

   Jika nilai *p-value* $<$ α, maka tolak H_0, artinya data stasioner.

   Jika nilai *p-value* $>$ α, maka tidak tolak H_0, artinya data tidak stasioner.

4. Statistik Uji

   Statistik uji stasioner dapat diformulasikan pada persamaan berikut:

   $$T = \frac{\hat{\phi}}{SE(\hat{\phi})} \tag{2.3}$$

   dengan:

   $\hat{\phi}$ = penduga parameter ϕ

   $SE(\hat{\phi})$ = standar *error* $\hat{\phi}$

5. Keputusan
6. Kesimpulan

Jika data tidak stasioner terhadap rata-rata maka perlu dilakukan *differencing.* Menurut Cryer & Chan (2008), *differencing* dilakukan dengan cara mengurangkan nilai pengamatan pada waktu Y_t dari nilai pengamatan pada waktu sebelumnya, yaitu Y_{t-1}, menggunakan proses *backward shift* dengan operator B yaitu operator yang menggeser data satu langkah ke belakang pada deret waktu untuk membuat deret baru. Secara umum persamaan *backward shift* (B) sebagai berikut:

$$B^d Y_t = Y_{t-d} \tag{2.4}$$

untuk d dan $t = 1, 2, \ldots, n$ dengan $t > d$. *Differencing* orde pertama $(d = 1)$ dan orde kedua $(d = 2)$ dapat dinyatakan dalam persamaan sebagai berikut :

$$\Delta Y_t = Y_t - Y_{t-1} \tag{2.5}$$

$$\Delta^2 Y_t = Y_t - 2Y_{t-1} + Y_{t-2} \tag{2.6}$$

sehingga bentuk umum *differencing* orde ke-*d* untuk mencapai kestasioneran terhadap rata-rata dapat dinyatakan dalam persamaan (2.7).

$$\Delta^d Y_t = (1 - B)^d Y_t \tag{2.7}$$

2.3 *Autocorrelation Function* (ACF)

Autocorrelation function (ACF) adalah ukuran korelasi antara dua nilai pada deret waktu Y_t dan Y_{t+k} dengan kovarians antar keduanya hanya bergantung pada jarak waktu (*lag*) ke-*k* (Wei, 2006). Fungsi autokorelasi dapat dituliskan dalam persamaan sebagai berikut:

$$\hat{\rho}_k = \frac{\sum_{t=k+1}^{n}(Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{\sum_{t=1}^{n}(Y_t - \bar{Y})^2}, k = 1,2, \ldots, n \tag{2.8}$$

dengan:

Y_t = pengamatan periode ke-*t*

Y_{t+k} = pengamatan periode ke-(*t*+*k*)

$\bar{Y}$ = rata-rata pengamatan Y_t

$\hat{\rho}_k$ = koefisien autokorelasi sampel pada *lag* ke-k

2.4 *Partial Autocorrelation Function* (PACF)

Menurut Wei (2006) *Partial Autocorrelation Function* (PACF) digunakan untuk mengukur korelasi antara dua nilai deret waktu Y_t dan Y_{t+k} dengan mengabaikan pengaruh nilai-nilai antara keduanya (Y_{t+1} hingga Y_{t+k-1}), sehingga fungsi autokorelasi parsial atau *partial autocorrelation function* dapat dituliskan sebagai berikut:

$$\hat{\phi}_{kk} = \frac{\hat{\rho}_k - \sum_{j=1}^{k-1} \hat{\phi}_{k-1,j} \hat{\rho}_{k-j}}{1 - \sum_{j=1}^{k} \hat{\phi}_{k-1,j} \hat{\rho}_j}, k = 1,2,\dots,n \quad (2.9)$$

dengan:

$\hat{\phi}_{kk}$ = koefisien autokorelasi parsial *lag* ke-k

$\hat{\rho}_k$ = koefisien autokorelasi sampel pada *lag* ke-k

$\hat{\rho}_j$ = koefisien autokorelasi sampel pada *lag* ke-j

$\hat{\rho}_{k-j}$ = koefisien autokorelasi sampel pada *lag* ke-(k-j)

2.5 *Autoregressive Integrated Moving Average* (ARIMA)

Model *Autoregressive Integrated Moving Average* (ARIMA) atau biasa disebut metode Box-Jenkins pertama kali dikembangkan oleh George Box dan Gwilym Jenkins pada tahun 1970 (Box et al., 2016). Menurut Box et al., (2016), model ARIMA (p,d,q) dibentuk dari operator *Autoregressif* (AR) berorde p, operator *Differencing* sebanyak d, dan operator *Moving Average* (MA) berorde q.

- *Autoregressive* (AR)

Model *Autoregressive* (AR) yaitu model deret waktu dengan nilai saat ini dipengaruhi oleh nilai-nilai sebelumnya pada variabel yang sama (Cryer & Chan, 2008). Persamaan model *Autoregressive* (AR) berorde p yaitu:

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + e_t \qquad (2.10)$$

dengan:

ϕ_0 = konstanta rata-rata
$\phi_1, \phi_2, \dots, \phi_p$ = parameter AR
e_t = galat pada periode ke-*t*

- *Moving Average* (MA)

Menurut Cryer & Chan (2008) model *Moving Average* (MA) merupakan model deret waktu yang menjelaskan bahwa nilai saat ini dipengaruhi oleh kesalahan acak sebelumnya dan kesalahan acak saat ini. *Moving Average* (MA) berorde q dapat dinyatakan dalam persamaan:

$$Y_t = \theta_0 + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q} \qquad (2.11)$$

dengan:

θ_0 = konstanta
θ_q = parameter MA
e_t = galat pada periode ke-*t*

- Model *Autoregressive Moving Average* (ARMA)

Model ARMA (p, q) atau *Autoregressive Moving Average* merupakan model gabungan dari *Autoregressive* (AR) dan *Moving Average* (MA) dengan orde masing-masing p dan q (Box et al., 2016), dapat dituliskan dalam persamaan:

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q} \qquad (2.12)$$

dengan:

ϕ_0 = konstanta

$\phi_1, \phi_2, \dots, \phi_p$ = parameter AR

$\theta_1, \theta_2, \dots, \theta_q$ = parameter MA

e_t = galat pada periode ke-*t*

- Model *Autoregressive Integrated Moving Average* (ARIMA)

Model ARIMA dapat ditulis dengan orde (p, d, q) yang terdiri dari model *Autoregressive* orde (p), hasil *differencing* pada data berorde (d) dan (q) menunjukkan model *Moving Average* yang digunakan untuk memodelkan data deret waktu (Cryer & Chan, 2008), sehingga didapat persamaan:

$$\phi_p(B)(1-B)^d Y_t = \theta_q(B) e_t \quad (2.13)$$

atau jika ARIMA (*p*,1,*q*) dapat dituliskan sebagai berikut:

$$\Delta Y_t = \phi_1 \Delta Y_{t-1} + \cdots + \phi_p \Delta Y_{t-p} + e_t - \theta_1 e_{t-1} - \cdots - \theta_q e_{t-q} \quad (2.14)$$

dengan:

ϕ_p = parameter AR

θ_q = parameter MA

B = operator *backward shift*

$(1-B)^d$ = diferensiasi orde ke-d

e_t = galat pada periode ke-*t*

ΔY_t = $Y_t - Y_{t-1}$

2.6 Estimasi Parameter Model ARIMA

Estimasi parameter model ARIMA (p, d, q) dilakukan pada data yang telah stasioner menggunakan metode momen, kuadrat terkecil, dan *maximum likelihood* (Cryer & Chan, 2008). Estimasi parameter model pada penelitian ini menggunakan metode *maximum likelihood*. Pendekatan menggunakan *Maximum Likelihood Estimation*

(MLE) berdasarkan asumsi bahwa *error* e_t berdistribusi normal identik dan independen dengan rataan nol dan varians σ_e, yang dinyatakan sebagai:

$$e_t \sim N(0, \sigma_e^2)$$

maka fungsi kepadatan peluang dari e_t dapat dituliskan sebagai:

$$f(e_t|\sigma_e^2) = \frac{1}{\sqrt{2\pi\sigma_e^2}} exp\left(-\frac{e_t^2}{2\sigma_e^2}\right)$$

galat diasumsikan saling independen, maka fungsi *likelihood* diperoleh dari hasil perkalian fungsi kepadatan peluang masing-masing galat, sehingga diperoleh:

$$\begin{aligned} L(\phi, \theta, \sigma_e^2) &= \prod_{t=1}^{n} f(e_t|\sigma_e^2) \\ &= (2\pi\sigma_e^2)^{-\frac{n}{2}} exp\left(-\frac{1}{2\sigma_e^2}\sum_{t=1}^{n} e_t^2\right) \end{aligned} \quad (2.15)$$

dengan:

e_t = galat pada periode ke-*t*

σ_e^2 = varians galat

Misalkan $Y = (Y_1, Y_2, \ldots, Y_n)$ dan nilai awal deret waktu $Y_* = \left(Y_{1-p}, \ldots, Y_0\right)$ serta galat awal $e_* = \left(e_{1-q}, \ldots, e_0\right)$ diketahui, maka fungsi *log-likelihood* bersyarat dapat dinyatakan dalam persamaan sebagai berikut:

$$\ln L_* (\phi, \theta, \sigma_e^2) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma_e^2) - \frac{1}{2\sigma_e^2} S_*(\phi, \theta) \quad (2.16)$$

dengan

$$S_*(\phi, \theta) = \sum_{t=1}^{n} e_t^2 \,(\phi, \theta | Y_*, e_*, Y)$$

merupakan fungsi kuadrat bersyarat (*conditional sum of square)*, selanjutnya persamaan (2.16) diturunkan terhadap σ_e^2 sehingga diperoleh:

$$\frac{\partial \ln L_*}{\partial \sigma_e^2} = -\frac{n}{2\sigma_e^2} + \frac{S_*(\phi,\theta)}{2(\sigma_e^2)^2} \quad (2.17)$$

$$-\frac{n}{2\sigma_e^2} = -\frac{S_*(\phi,\theta)}{2(\sigma_e^2)^2}$$

$$\sigma_e^2 = \frac{1}{n} S_*(\phi, \theta) \tag{2.18}$$

persamaan (2.18) disubtitusi ke dalam persamaan (2.16) menghasilkan fungsi *log-likelihood* terkonsentrasi diperoleh:

$$\ln l_*(\phi, \theta) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2} - \frac{n}{2}\ln\left(\frac{S_*(\phi,\theta)}{n}\right) \tag{2.19}$$

dua suku pertama pada persamaan (2.19) merupakan konstanta yang tidak bergantung pada parameter ϕ dan θ. Proses optimasi fungsi *log-likelihood* monoton naik dan koefisien $-\frac{n}{2}$ bernilai negatif, maka memaksimumkan fungsi *log-likelihood* ekuivalen dengan meminimumkan fungsi kuadrat bersyarat $S_*(\phi, \theta)$, sehingga estimator maksimum likelihood untuk parameter AR dan MA diperoleh:

$$\left(\hat{\phi}, \hat{\theta}\right) = \arg\min_{\phi,\theta} S_*(\phi, \theta) \tag{2.20}$$

Nilai $\hat{\phi}$ dan $\hat{\theta}$ yang diperoleh merupakan estimasi parameter ARIMA, sehingga nilai hasil estimasi ini yang akan digunakan untuk membangun model ARIMA terbaik (Box et al., 2016).

Kelayakan suatu parameter dalam model ARIMA (*p*,*d*,*q*) ditentukan oleh uji signifikansi parameter model (Cryer & Chan, 2008). Model *Autoregressive* (*p*) untuk pengujian signifikan parameter menggunakan hipotesis sebagai berikut:

1. Uji Hipotesis

   $H_0 : \phi_p = 0$ (parameter model AR tidak signifikan)

   $H_1 : \phi_p \neq 0$ (parameter model AR signifikan)

2. Taraf Signifikan

   $\alpha = 5$ % atau 0,05

3. Daerah Kritis

   Jika nilai $|t_{hitung}| > t_{\left(\frac{\alpha}{2}, n-p\right)}$ atau jika *p-value* < α, maka tolak H_0, artinya parameter model AR signifikan.

Jika nilai $|t_{hitung}| < t_{\left(\frac{\alpha}{2},n-p\right)}$ atau jika *p-value* > α, maka tidak tolak H_0, artinya parameter model AR tidak signifikan.

4. Statistik Uji

   Statistik uji signifikan parameter model AR dalam persamaan berikut:

$$T = \frac{\hat{\phi}_p}{SE(\hat{\phi}_p)} \tag{2.21}$$

   dengan:

   $\hat{\phi}_p$ = parameter model AR

   $SE(\hat{\phi}_p)$ = standar *error* parameter model AR

5. Keputusan
6. Kesimpulan

Model *Moving Average* (q) menggunakan hipotesis sebagai berikut:

1. Uji Hipotesis

   $H_0 : \theta_q = 0$ (parameter model MA tidak signifikan)

   $H_1 : \theta_q \neq 0$ (parameter model MA signifikan)

2. Taraf Signifikan

   $\alpha = 5$ % atau 0,05

3. Daerah Kritis

   Jika nilai $|t_{hitung}| > t_{\left(\frac{\alpha}{2},n-p\right)}$ atau jika *p-value* < α, maka tolak H_0, artinya parameter model MA signifikan.

   Jika nilai $|t_{hitung}| < t_{\left(\frac{\alpha}{2},n-p\right)}$ atau jika *p-value* > α, maka tidak tolak H_0, artinya parameter model MA tidak signifikan.

4. Statistik Uji

   Statistik uji signifikan parameter model MA dalam persamaan berikut:

$$T = \frac{\hat{\theta}_q}{SE(\hat{\theta}_q)} \tag{2.22}$$

   dengan:

   $\hat{\theta}_q$ = parameter model MA

$SE(\hat{\theta}_q)$ = standar *error* parameter model MA

5. Keputusan
6. Kesimpulan

2.7 Uji Asumsi *White Noise*

Uji diagnostik dilakukan menggunakan uji *white-noise* untuk mengetahui apakah residual model memiliki varians konstan dan tidak saling berkorelasi (Karnadi & Prastyo, 2023). Menurut Cryer & Chan (2008) suatu proses dikatakan memenuhi asumsi *white-noise* jika residual bersifat acak yang tidak menunjukkan pola autokorelasi. Uji yang digunakan yaitu uji *Ljung-Box* yang berfungsi untuk mengevaluasi apakah residual model ARIMA telah sesuai dengan data. Hipotesis yang digunakan yaitu:

1. Uji Hipotesis

   H_0: $\rho_1 = \rho_2 = \cdots = \rho_k = 0$ (*white-noise).*

   H_1: minimal ada satu nilai $\rho_k \neq 0$ dengan k = 1, 2, ..., k (tidak *white-noise).*
2. Taraf Signifikan

   $\alpha = 5$ % atau 0,05
3. Daerah Kritis

   Tolak H_0 jika nilai $Q > \chi^2_{(\alpha;K-p-q)}$ atau jika *p-value* < α.

   Tidak tolak H_0 jika nilai $Q < \chi^2_{(\alpha;K-p-q)}$ atau jika *p-value* > α.
4. Statistik Uji

$$Q = n_r(n_r + 2)\sum_{k=1}^{K} \frac{\hat{\rho}_k^2}{n_r - k} \tag{2.23}$$

dengan:

$\hat{\rho}_k$ = penduga autokorelasi residual *lag* ke-k

k = banyaknya autokorelasi yang diuji

p = banyaknya parameter AR pada model

q = banyaknya parameter MA pada model

n_r = jumlah residual

2.8 Uji Asumsi Normal

Uji asumsi yang perlu dilakukan selain uji *white noise* yaitu uji normalitas residual. Menurut Cryer & Chan (2008) uji normalitas dapat dilakukan menggunakan uji *Jarque-Bera* yang didasarkan pada karakteristik distribusi normal yaitu *skewness* dan *kurtosis* sama dengan nol. Uji *Jarque-Bera* menggunakan hipotesis sebagai berikut:

1. Uji Hipotesis
   H_0: $e_t \sim N(0, \sigma^2)$ residual berdistribusi normal.
   H_1: $e_t \nsim N(0, \sigma^2)$ residual tidak berdistribusi normal.
2. Taraf Signifikan
   $\alpha = 5$ % atau 0,05
3. Daerah Kritis
   Tidak tolak H_0 jika nilai $JB < \chi^2_{(2)}$ atau jika *p-value* > α.
4. Statistik Uji

$$JB = \frac{\hat{S}^2}{6/n_r} + \frac{\left(\hat{K}^2 - 3\right)^2}{24/n_r} \tag{2.24}$$

dengan:

S = *skewness*

K = kurtosis

n_r = jumlah residual

2.9 Pemilihan Model Terbaik

Pemilihan model ARIMA dapat dilakukan menggunakan kriteria informasi seperti *Akaike's Information Criterion* (AIC) dengan membandingkan berbagai model hasil estimasi *maximum likelihood* dan model terbaik dipilih berdasarkan nilai AIC terkecil (Box et al., 2016). Kriteria AIC dirumuskan sebagai berikut:

$$AIC = n_r \log \sigma^2 + 2M \tag{2.25}$$

keterangan:

σ^2 = varian dari residual model

M = banyaknya parameter pada model

n_r = jumlah residual

2.10 Uji Nonlinearitas

Uji nonlinearitas dapat dilakukan menggunakan uji *Lagrange Multiplier* (LM), salah satu uji *Lagrange Multiplier* (LM) yang dikembangkan melalui model *Neural Network* yaitu uji Terasvirta (Terasvirta et al., 1993). Statistik uji yang digunakan dalam uji Terasvirta yaitu:

$$F_{hit} = \frac{\frac{SSR_0 - SSR_1}{m}}{\frac{SSR_1}{(n-p-1-m)}} \tag{2.26}$$

dengan :

SSR_0 = jumlah nilai kuadrat residual

SSR_1 = jumlah kuadrat residual yang dihitung menggunakan ekspansi deret Taylor

m = suku kubik atau kuadratik yang merupakan ekspansi deret Taylor

n = jumlah pengamatan

Hipotesis :

H_0 : *f (x)* merupakan fungsi linear dalam *x*

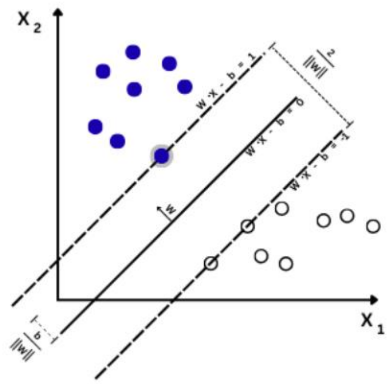
H_1 : *f (x)* merupakan fungsi nonllinear dalam *x*

Keputusan:

Jika $F_{hit} > F_{(n-p-1-m)}$ atau *p-value* < α makan tolak H_0.

2.11 *Support Vector Machine* (SVM)

Support Vector Machine (SVM) pertama kali dikenalkan oleh Vapnik pada tahun 1992 untuk menyelesaikan masalah klasifikasi dan melakukan estimasi terhadap fungsi nonlinear yang bertujuan mencari bidang pemisah *(hyperplane)* terbaik dengan *margin* terbesar sehingga dua kelompok data dapat dipisahkan secara optimal (Vapnik, 2000). Konsep SVM dapat dijelaskan melalui Gambar 1.



Gambar 1. Visualisai Konsep SVM

Berdasarkan Gambar 1, data terbagi menjadi dua kelas yaitu positif (+1) dan negatif (-1). Lingkaran biru menandakan titik data pada kelas positif, dan lingkaran putih menunjukkan kelas negatif. *Hyperplane* terbaik di dapat dengan mengukur *margin*, yaitu jarak antara *hyperplane* dengan titik data terdekat pada masing-masing kelas. Titik-titik data terdekat disebut *support vector* yang merupakan penentu dari *hyperplane* optimal. Posisi *hyperplane* yang berada tepat di tengah antara kedua kelas merupakan *hyperplane* optimal, sedangkan titik yang berada pada garis putus-putus merupakan *support vector*.

Misalkan data *training* $(\boldsymbol{x_i}, y_i)$ untuk $i = 1, 2, \ldots, n$ dengan $\boldsymbol{x_i} \in \mathbb{R}^n$ merupakan vektor fitur dan $y_i \in \{+1, -1\}$ merupakan label kelas. Pemisahan dapat dilakukan oleh sebuah *hyperplane* sebagai berikut (Vapnik, 2000):

$$\boldsymbol{w}.\boldsymbol{x}_i + b = 0 \tag{2.27}$$

dengan:

$\boldsymbol{w}$ = vektor bobot

b = koefisien bias

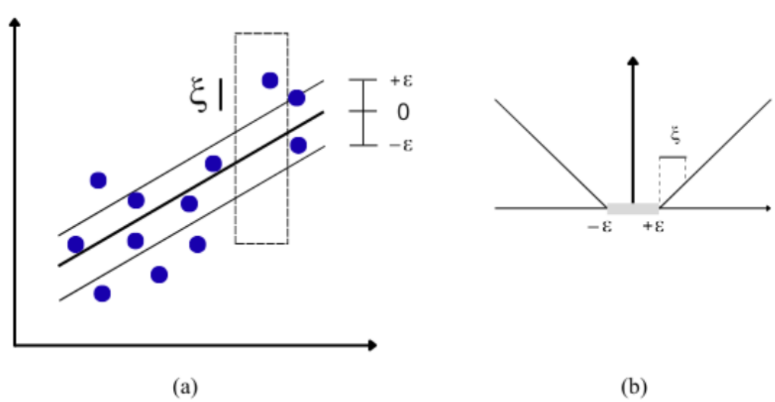
Menurut Vapnik (2000) *hyperplane* dapat dikatakan optimal apabila memiliki *margin* yang besar diperoleh dengan memaksimalkan jarak antar *hyperplane* dan titik data terdekat pada masing-masing kelas. Kedua kelas dapat dibedakan melalui kondisi pemisah yang dapat dituliskan sebagai berikut:

$$\boldsymbol{w}.\boldsymbol{x}_i + b \geq +1 \text{ untuk } y_i = +1 \quad (2.28)$$

$$\boldsymbol{w}.\boldsymbol{x}_i + b \leq -1 \text{ untuk } y_i = -1 \quad (2.29)$$

2.12 *Support Vector Regression* (SVR)

Support Vector Regression (SVR) merupakan pengembangam dari metode SVM yang digunakan untuk menyelesaikan permasalahan regresi. Tujuan utama SVR yaitu menemukan fungsi yang mampu memprediksi data dengan galat sekecil mungkin melalui pemetaan data ke ruang berdimensi tinggi menggunakan fungsi nonlinear, kemudian mencari *hyperplane* terbaik berdasarkan prinsip *Structural Risk Minimization* untuk memperoleh model yang akurat (Vapnik, 2000). Konsep SVR dapat divisualisasikan melalui Gambar 2 berikut:



Gambar 2. Visualisasi Konsep SVR

Gambar 2 dapat menunjukkan *hyperplane* yang ditandai dengan garis tebal, sedangkan dua garis yang mengapitnya merupakan *soft margin*. Jarak antara *hyperplane* dan *soft margin* bernilai ε, dan titik-titik yang berada di antara ε sampai $-\varepsilon$ disebut *support vector*. Titik yang melewati batas *soft margin* memerlukan penambahan variabel *slack* ξ sebagai bentuk toleransi.

Menurut Smola & Scholkopf (2004) ide awal metode SVR adalah menggunakan sejumlah data *training* sebanyak n, $(\boldsymbol{x_i}, y_i)$ dengan $\boldsymbol{x_i} \in \mathbb{R}^d$ yang merupakan vektor *input* dari data ke-i dengan $i = 1, 2, \dots, n$ dan dimensi d, sedangkan y_i merupakan nilai target yang ingin diprediksi. Fungsi regresi pada SVR dapat dituliskan dalam persamaan:

$$f(x) = \langle \boldsymbol{w}.\boldsymbol{x} \rangle + b \tag{2.30}$$

dengan:

$\boldsymbol{w}$ = vektor bobot

b = koefisien bias

$\boldsymbol{x}$ = vektor *input*

$f(x)$ = fungsi SVR

Koefisien $\boldsymbol{w}$ dan b dapat diestimasikan dengan meminimalkan fungsi resiko dalam persamaan berikut:

$$R\big(f(\boldsymbol{x})\big) = C \sum_{i=1}^{n} E_\varepsilon(y_i - f(\boldsymbol{x}_i)) + \frac{1}{2} \|\boldsymbol{w}\|^2 \tag{2.31}$$

dengan $\|\boldsymbol{w}\|$ merupakan fungsi yang diminimumkan untuk menjaga agar model tetap setipis *(falt)* mungkin (Vapnik, 2000). Menurut Smola & Scholkopf (2004), konstanta *C (Cost)* > 0 adalah nilai tawar antara ketipisan fungsi *f* dengan batas deviasi lebih dari ε yang dapat ditoleransi. E_ε adalah *ε-insensitive loss function* yang dapat dituliskan seperti:

$$E_\varepsilon\big(y_i - f(\boldsymbol{x_i})\big) = \begin{cases} |y_i - f(\boldsymbol{x_i})| - \varepsilon, \text{untuk } |y_i - f(\boldsymbol{x_i})| \geq \varepsilon \\ 0 \qquad\qquad\qquad\qquad , \text{lainnya} \end{cases} \tag{2.32}$$

dengan kendala:

$$y_i \leq f(\boldsymbol{x_i}) + \varepsilon$$
$$y_i \geq f(\boldsymbol{x_i}) - \varepsilon$$

Fungsi *f* dalam model SVR diasumsikan bisa mendekati semua titik data $(\boldsymbol{x_i}, y_i)$ dengan tingkat akurasi ε (Smola & Scholkopf, 2004). Jika titik data berada pada rentang $f \pm \varepsilon$ maka titik itu disebut *feasible* (layak) dan jika titik berada di luar rentang $f \pm \varepsilon$ maka disebut *infeasible* (tidak layak), untuk menangani masalah titik yang *infeasible* perlu menambahkan variabel *slack* positif ξ_i, ξ_i^*. Variabel ini membantu mengatasi kendala yang tidak bisa dipengaruhi dalam proses optimasi, sehingga persamaan (2.30) dapat ditransformasi menjadi bentuk persamaan:

$$min\frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i^*) \tag{2.33}$$

dengan kendala:

$$y_i \leq f(\boldsymbol{x_i}) + \varepsilon + \xi_i$$
$$y_i \geq f(\boldsymbol{x_i}) - \varepsilon - \xi_i^*$$
$$\xi_i, \xi_i^* \geq 0$$

Persamaan 2.32 dapat diselesaikan menggunakan fungsi *Lagrange Multiplier* (LM) yang diformulasikan dalam bentuk pemrograman kuadratik *(quadratic programming)* untuk memperoleh *hyperplane* terbaik (Vapnik, 2000), sehingga fungsi *Lagrange* dapat dinyatakan sebagai berikut:

$$L = \left(\frac{1}{2}\|\boldsymbol{w}\|^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i^*)\right) - \left(\sum_{i=1}^{n}(\eta_i\xi_i + \eta_i^*\xi_i^*)\right) - \left(\sum_{i=1}^{n}\alpha_i(\varepsilon + \xi_i + \langle\boldsymbol{w}.x\rangle + b - y_i)\right) - \left(\sum_{i=1}^{n}\alpha_i^*(\varepsilon + \xi_i^* - \langle\boldsymbol{w}.x\rangle + b + y_i)\right) \tag{2.34}$$

dengan :

L = fungsi *Langrange*

$\alpha_i, \alpha_i^*, \eta_i, \eta_i^*$ = *Langrange Multiplier* (koefisien *lagrange*)

Solusi optimal diperoleh dengan menurunkan fungsi L terhadap variabel $\boldsymbol{w}, b, \xi_i$, dan ξ_i^*.

$$\frac{\partial L}{\partial w} = w - \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)x_i = 0 \tag{2.35}$$

$$\frac{\partial L}{\partial b} = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*) = 0 \tag{2.36}$$

$$\frac{\partial L}{\partial \xi_i} = C - \alpha_i - \eta_i = 0 \tag{2.37}$$

$$\frac{\partial L}{\partial \xi_i^*} = C - \alpha_i^* - \eta_i^* = 0 \tag{2.38}$$

Hasil turunan parsial pada persamaan (2.35), (2.36), (2.37), dan (2.38) disubsitusikan kembali ke dalam persamaan (2.34). Setelah dilakukan penyederhanaan, diperoleh bentuk masalah dual model SVR sebagai berikut:

$$\max \begin{cases} -\frac{1}{2}\sum_{i,j=1}^{n}(\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)\langle \boldsymbol{x_i}, \boldsymbol{x_j}\rangle \\ -\varepsilon \sum_{i=1}^{n}(\alpha_i - \alpha_i^*) + \sum_{i=1}^{n} y_i(\alpha_i - \alpha_i^*) \end{cases} \tag{2.39}$$

dengan kendala:

$$\sum_{i=1}^{n}(\alpha_i - \alpha_i^*) = 0$$

$$0 \le \alpha_i \le C \qquad \forall_i = 1, 2, \dots, n$$

$$0 \le \alpha_i^* \le C \qquad \forall_i = 1, 2, \dots, n$$

Solusi optimal untuk parameter $\boldsymbol{w}$ kemudian diperoleh dalam bentuk koefisien *Lagrange Multiplier* α_i, dan α_i^* yang dirumuskan sebagai:

$$\boldsymbol{w} = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)\boldsymbol{x_i} \tag{2.40}$$

Menurut Smola & Scholkopf (2004) nilai b (bias) dapat dihitung menggunakan kondisi *Karush-Kuhn-Tucker* (KKT) yang menyatakan bahwa pada titik solusi optimal berlaku:

$$\alpha_i\big(\varepsilon + \xi_i - y_i + f(\boldsymbol{x_i})\big) = 0 \tag{2.41}$$

$$\alpha_i^*\big(\varepsilon + \xi_i^* + y_i - f(\boldsymbol{x_i})\big) = 0 \tag{2.42}$$

$$(C - \alpha_i)\xi_i = 0 \tag{2.43}$$

$$(C - \alpha_i^*)\xi_i^* = 0 \tag{2.44}$$

Berdasarkan kondisi tersebut, nilai b dapat dihitung menggunakan $b = y_i - \boldsymbol{w}\boldsymbol{x_i} - \varepsilon$ untuk $0 \leq \alpha_i \leq C$ atau $b = -y_i + \boldsymbol{w}\boldsymbol{x_i} - \varepsilon$ untuk $0 \leq \alpha_i^* \leq C$, selanjutnya persamaan (2.30) dan (2.40) akan disubsitusikan sehingga diperoleh persamaan baru sebagai berikut:

$$f(x) = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)\left(\boldsymbol{x_i}.\boldsymbol{x_j}\right) + b \tag{2.45}$$

Pada permasalahan nonlinear, data *input* $\boldsymbol{x_i}$ dan $\boldsymbol{x_j}$ akan ditransformasikan ke *future space* menggunakan fungsi φ dengan $\varphi\colon \mathrm{R} \to \mathrm{R}^z$ artinya nilai vektor $\boldsymbol{x_i}$ dan $\boldsymbol{x_j}$ akan dipetakan ke dalam φ (Smola & Scholkopf, 2004), sehingga diperoleh persamaan nonlinear sebagai berikut:

$$\boldsymbol{w} = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)\varphi(\boldsymbol{x_i}) \tag{2.46}$$

sehingga diperoleh persamaan sebagai berikut:

$$f(x) = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)\varphi(\boldsymbol{x_i})\varphi\left(\boldsymbol{x_j}\right) + b \tag{2.47}$$

Maka φ dapat dituliskan dalam fungsi kernel *K* dalam persamaan (2.48).

$$f(x) = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)K\left(\boldsymbol{x_i}.\boldsymbol{x_j}\right) + b \tag{2.48}$$

dengan $\mathrm{K}\left(\boldsymbol{x_i},\boldsymbol{x_j}\right)$ adalah fungsi kernel yang didefinisikan sebagai $\mathrm{K}\left(\boldsymbol{x_i},\boldsymbol{x_j}\right) = \varphi(\boldsymbol{x_i})\varphi(\boldsymbol{x_j})$ digunakan dalam metode SVM maupun SVR (Vapnik, 2000).

2.13 Fungsi Kernel

Fungsi kernel digunakan untuk menerapkan model ke dalam ruang fitur dengan dimensi yang lebih tinggi secara tidak langsung untuk menentukan pemetaan dari ruang *input* ke ruang fitur tersebut (Annas et al., 2023). Permasalahan nonlinear dengan dimensi yang lebih tinggi dapat diatasi dengan mengganti *inner product* antara vektor $\boldsymbol{x_i}$ dan $\boldsymbol{x_j}$ dengan fungsi kernel. Menurut Purnama (2021), akurasi regresi dapat dipengaruhi oleh pemilihan fungsi kernel yang tepat dan parameter yang digunakan. Fungsi kernel yang digunakan pada metode SVR yaitu:

Tabel 2. Fungsi Kernel

No.	**Jenis Kernel**	**Formula**
1.	Linear	$K(\boldsymbol{x_i}, \boldsymbol{x_j}) = \boldsymbol{x_i^T} \boldsymbol{x_j}$
2.	Polinomial	$K(\boldsymbol{x_i}, \boldsymbol{x_j}) = (\boldsymbol{x_i^T} \boldsymbol{x_j} + 1)^P$
3.	*Radial Basis Function* (RBF)	$K(\boldsymbol{x_i}, \boldsymbol{x_j}) = \exp(-\gamma(\boldsymbol{x_i} - \boldsymbol{x_j})^2)$

dengan:

$\boldsymbol{x_i}, \boldsymbol{x_j}$ = vektor dari dua data set

P = derajat polinomial

γ = parameter gamma

Menurut Pai & Lin (2005) fungsi kermel *Radial Basis Function* (RBF) cenderung menunjukkan performa yang lebih unggul dalam mengatasi data nonlinear, sehingga penelitian ini menggunakan fungsi kernel *Radial Basis Function* (RBF). Dalam SVR fungsi kernel *Radial Basis Function* (RBF) memiliki tiga parameter yang harus ditetapkan yaitu *cost (C), gamma* (γ), dan *epsilon* (ε).

2.14 Metode *Grid-Search*

Metode *grid search* perlu dilakukan untuk mencari parameter optimal pada model SVR (Yasin dkk., 2014). *Grid search* bertujuan untuk mengidentifikasi parameter optimal pada data *training*, sehingga model yang didapat mampu memprediksi secara akurat pada data *testing*. Menurut Hsu et al., (2004) metode *grid search* memerlukan indikator evaluasi performa yang diukur oleh *cross validation* pada data *training* untuk memastikan generalisasi model terhadap data. Langkah *cross validation* dilakukan dengan membagi secara acak data *training* menjadi *n subset* yang berukuran sama, selanjutnya bagian *n-1* akan menjadi data *training* dan satu bagian lainnya menjadi data *testing*. Langkah ini akan dilakukan sebanyak *n* pengulangan pada setiap kombinasi data *training* dan data *testing*. Melalui uji *cross validation* nilai akurasi dari setiap iterasi akan dirata-rata untuk mendapat estimasi nilai akurasi akhir, dan pasangan parameter yang memiliki akurasi terbaik merupakan parameter yang optimal.

2.15 Model *Hybrid* ARIMA-SVR

Model *hybrid* ARIMA-SVR merupakan penggabungan antara model ARIMA dan SVR dengan memanfaatkan keunggulan masing-masing dari metode dalam peramalan. Menurut Albarr & Kusumawati (2023) model *hybrid* ARIMA-SVR terdiri dari dua komponen utama, yaitu komponen linear yang akan diselesaikan oleh ARIMA dan komponen nonlinear akan diselesaikan menggunakan SVR. Pendekatan model *hybrid* mampu memberikan hasil peramalan yang lebih akurat dibandingkan hanya menggunakan salah satu model secara terpisah (Pai & Lin, 2005). Model *hybrid* ARIMA-SVR dapat dinyatakan dalam persamaan:

$$Y_t = L_t + N_t \tag{2.49}$$

dengan Y_t adalah data runtun waktu ke-t, L_t adalah komponen linear dan N_t merupakan komponen nonlinear. Langkah selanjutnya yaitu data akan dimodelkan dengan model linear ARIMA sehingga diperoleh residual pada waktu ke-*t* dari model ARIMA (Zhang, 2001), sebagai berikut:

$$\varepsilon_t = Y_t - \hat{L}_t \tag{2.50}$$

dengan ε_t adalah residual ARIMA dan $\hat{L}_t$ menunjukkan nilai estimasi model ARIMA pada waktu ke-t. Residual ε_t yang mengandung komponen nonlinear, kemudian dimodelkan menggunakan SVR (Pai & Lin, 2005). Model SVR secara umum dapat dinyatakan sebagai fungsi nonlinear sebagai berikut:

$$\varepsilon_t = f(\varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-n}) + \Delta_t \tag{2.51}$$

dengan f merupakan fungsi nonlinear yang dimodelkan oleh SVR, dan Δ_t adalah *random error*. Apabila model ARIMA dan SVR berhasil didapat, maka hasil peramalan model *hybrid* ARIMA-SVR dapat dinyatakan dalam persamaan berikut:

$$\hat{Y}_t = \hat{L}_t + \hat{N}_t \tag{2.52}$$

dengan $\hat{L}_t$ dan $\hat{N}_t$ masing-masing merupakan nilai peramalan komponen linear dan nonlinear yang diperoleh dari model ARIMA dan SVR pada waktu ke-t, dan $\hat{Y}_t$ merupakan nilai ramalan dari model *hybrid* ARIMA-SVR.

2.16 Akurasi Metode Peramalan

Metode peramalan membutuhkan akurasi untuk mengevaluasi hasil dari prediksi yang sudah dilakukan. Penelitian ini menggunakan *Mean Absolute Percentage Error* (MAPE) sebagai metode yang dipilih untuk mengukur tingkat kesalahan model. Berikut adalah persamaan untuk menentukan nilai MAPE (Makridakis et al., 1983):

$$\text{MAPE} = \frac{\sum_{t=1}^{N}\left|\frac{Y_t - \hat{Y}_t}{Y_t}\right|}{n} \times 100\% \tag{2.53}$$

dengan $\hat{Y}_t$ adalah hasil peramalan dan n adalah banyaknya data. Nilai MAPE memiliki kriteria sebagai berikut:

Tabel 3. Kriteria Nilai MAPE

Nilai MAPE	**Tingkat Akurasi**
$\leq$10%	Sangat baik
10% < MAPE $\leq$ 20%	Baik
20% < MAPE $\leq$ 50%	Cukup
>50%	Buruk

2.17 *Ethereum*

Cryptocurrency adalah mata uang digital yang diciptakan dari rangkaian kode atau disebut *blockchain* (Rizkilloh & Widiyanesti, 2022). Mata uang kripto bisa digunakan sebagai alat pembayaran yang transaksinya dapat dilakukan secara virtual atau melalui internet (Chuen et al., 2018), salah satunya yaitu *Ethereum*. *Ethereum* pertama kali dikenalkan oleh Vitalik Buterin pada tahun 2014 dengan menerbitkan *white paper* dan resmi diluncurkan pada tahun 2015 oleh Buterin dan Joe Lubin. *Ethereum* memiliki mata uang kripto *ether* atau (ETH). *Ethereum* sendiri merupakan platform perangkat lunak berbasis *blockchain* yang bersifat terdesentralisasi, sehingga memungkinkan pembuatan dan pengoperasian *Smart Contract and Distributed Applications* (Dapps) tanpa hambatan penipuan kontrol, atau campur tangan pihak ketiga (Yunizar dkk., 2023).

III. METODOLOGI PENELITIAN

3.1 Waktu dan Tempat Penelitian

Penelitian ini dilakukan pada semester ganjil di Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung tahun ajaran 2025/2026.

3.2 Data dan Sumber Data

Data pada penelitian ini menggunakan data sekunder, yaitu harga penutupan *Ethereum* (ETH) dalam US Dollar yang tersedia di *website Investing.com*. Rentang waktu yang diambil yaitu dari 11 Desember 2020 sampai 10 Desember 2025 dengan variabel tanggal, dan penutupan. Periode tersebut diambil karena menyediakan data deret waktu yang cukup panjang dan kontinu untuk menangkap pola tren dan ketergantungan waktu pada data. Penggunaan variabel penutupan dipilih karena mencerminkan kondisi akhir pasar pada setiap pengamatan. Data berjumlah 1827 yang diperoleh dari link berikut https://id.investing.com/crypto/ethereum/historical-data.

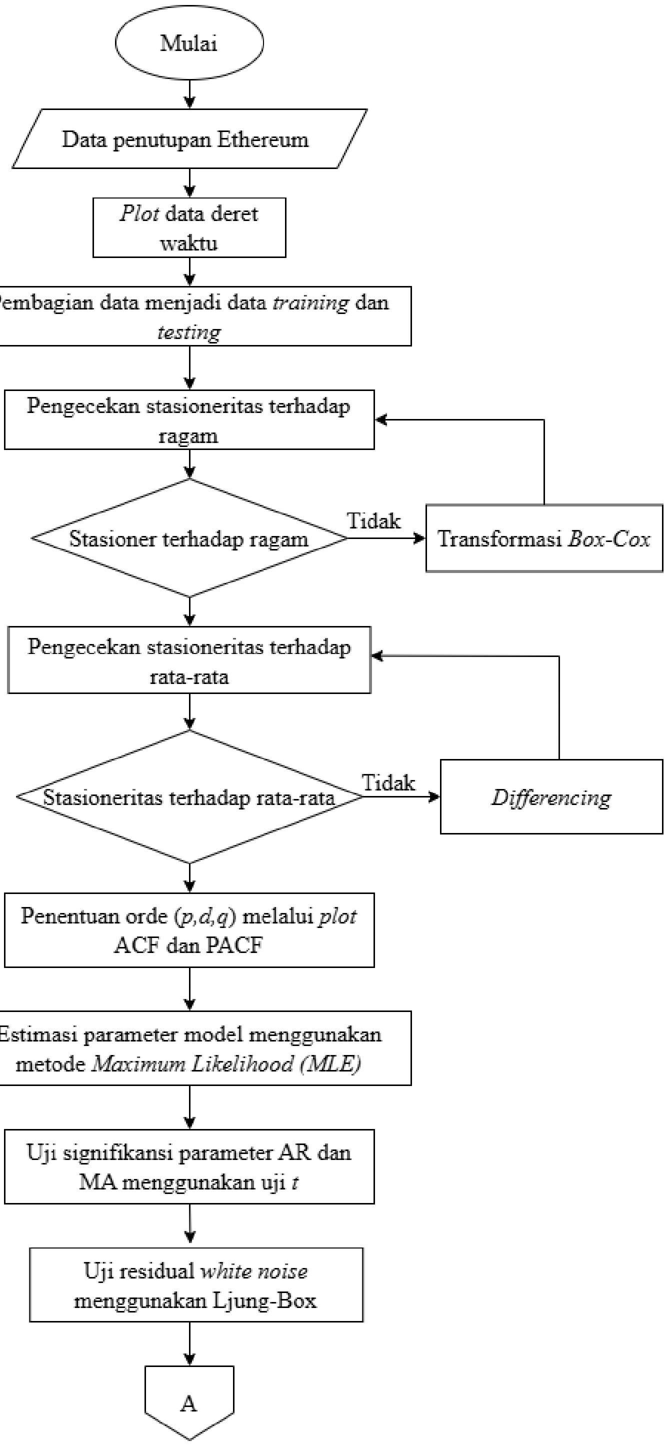
3.3 Metode Penelitian

Analisis data pada penelitian ini meliputi beberapa tahap. Tahap pertama yang dilakukan yaitu melakukan pemodelan data menggunakan model linear ARIMA. Selanjutnya pada tahap kedua dilakukan pemodelan SVR dari residual hasil ramalan model ARIMA. Setelah mendapat model dengan pola linear dan pola nonlinear dari data maka dapat dilakukan peramalan menggunakan metode *hybrid* ARIMA-SVR. Langkah-langkah untuk peramalan menggunakan metode *hybrid* ARIMA-SVR dapat dilihat pada bagan alir sebagai berikut:

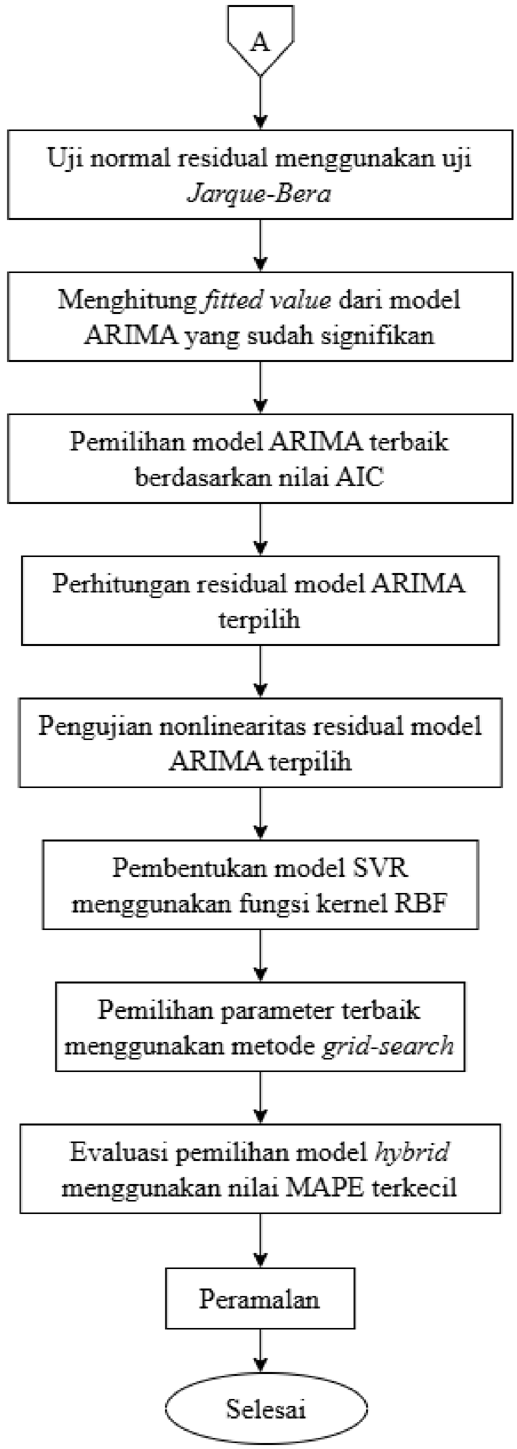
1. Melakukan visualisai data harga penutupan *Ethereum* (ETH) untuk melihat pola data secara visual.
2. Membagi data menjadi dua skema yaitu data *training* dan data *testing* dengan skema 70%:30%, 80%:20% *,* dan 90%:10%.
3. Menguji kestaioneran terhadap ragam menggunakan transformasi *Box-Cox*.
4. Menguji kestasioneran terhadap rata-rata menggunakan uji ADF serta melakukan *differencing* jika data belum stasioner.
5. Menentukan model ARIMA (p,d,q) dengan melihat *plot* ACF dan PACF.
6. Melakukan estimasi parameter model menggunakan metode *Maximum Likelihood* (MLE).
7. Menguji signifikasi parameter AR dan MA menggunakan uji t.
8. Menguji apakah residual bersifat *white noise* menggunakan uji *Ljung-Box*.
9. Menguji residual bersifat normal menggunakan uji *Jarque-Bera*.
10. Memilih model terbaik berdasarkan nilai AIC.
11. Menghitung *fitted value* dari model ARIMA yang sudah signifikan.
12. Menghitung residual dari model ARIMA yang digunakan untuk *input* model SVR.
13. Melakukan uji nonlinearitas terhadap residual model ARIMA menggunakan uji Terasvirta.
14. Membangun model SVR menggunakan input residual model ARIMA.

15. Menggunakan fungsi kernel RBF untuk mengoptimalkan parameter *gamma, cost,* dan *epsilon*.
16. Melakukan *grid search* untuk mendapatkan model parameter SVR yang optimal.
17. Memilih model terbaik berdasarkan nilai MAPE terkecil pada data *training*.
18. Membentuk model *hybrid* ARIMA-SVR dengan menggabungkan *output* ARIMA dan SVR.
19. Melakukan peramalan harga penutupan *Ethereum* (ETH) menggunakan model *hybrid* ARIMA-SVR.

Berikut diberikan diagram alir dari penelitian ini:



Gambar 3. Diagram Alir Model ARIMA (p,d,q)



Gambar 4. Diagram Alir Model *Hybrid* ARIMA-SVR

V. KESIMPULAN

Berdasarkan hasil penelitian menggunakan model *hybrid* ARIMA-SVR pada harga penutupan *Ethereum* dapat disimpulkan bahwa:

1. Model *hybrid* ARIMA-SVR berhasil diterapkan pada harga penutupan *Ehereum* pada rentang waktu 11 Desember 2020 hingga 10 Desember 2025 dengan menggabungkan metode ARIMA dan SVR untuk mengidentifikasi pola linear dan nonlinear menunjukkan bahwa model memiliki performa yang baik. Model *hybrid* paling efektif yaitu ARIMA (2,1,2)-SVR dengan parameter terbaik C sebesar 0.8125, parameter γ sebesar 5, dan parameter ε sebesar 0.125, dengan skema pembagian data terbaik yaitu 90% untuk data *training* dan 10% untuk data *testing*.

2. Penelitian ini memperoleh tingkat kesalahan sebesar 2.83% untuk data *training* dan 2.72% untuk data *testing*. Nilai MAPE berada di bawah 10% menunjukkan bahwa model memiliki tingkat kesalahan yang rendah. Oleh karena itu, model yang dihasilkan memiliki performa yang baik dan layak digunakan untuk peramalan harga penutupan *Ethereum*.

DAFTAR PUSTAKA

Albarr, H., & Kusumawati, R. (2023). Hybrid Autoregressive Integrated Moving Average-Support Vector Regression For Stock Price Forecasting. *Jurnal Matematika, Sains, dan Teknologi*, *24*(2), 1–17.

Annas, S., Rais, Z., Aswi, A., Indrayasaro, & Nurfajriani. (2023). Implementation of Support Vector Regression (SVR) Analysis in Predicting Gold Prices in Indonesia. *Proceddings of the 5th International Conference on Statistics, Mathematics, Teaching, and Research 2023*, 97–107. https://doi.org/10.2991/978-94-6463-332-0_12

Atmaja, D. M. U., & Hakim, A. R. (2022). Peramalan Harga Mata Uang Kripto Solana Menggunakan Metode Support Vector Regression (SVR). *Jurnal Media Elektro*, *9*(2), 97–104. https://doi.org/10.35508/JME.V0I0.8117-

Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2016). *Time Series Analysis Forexasting and Control* (Fifth Edit). John Wiley & Sons, Inc., Hoboken, New Jersey.

Carista, N., Mahadesyawardani, A., & Dwiyanto, A. S. (2025). Forecasting Nickel Prices in Indonesia Using ARIMA , SVR , and Hybrid ARIMA-SVR Approach. *Jurnal Matematika, Statistika dan Komputasi*, *21*(3), 627–645. https://doi.org/10.20956/j.v21i3.42558

Chuen, D. L. K., Guo, L., & Wang, Y. (2018). Cryptocurrency: A New Investment Opportunity? *The Journal of Alternative Investments*, *20*(3), 16–40.

Cryer, J. D., & Chan, K.-S. (2008). *Time Series Analysis With Applications in R* (Second Edi). Springer Science+Business Media, LLC.

Dani, M. F. R., & Syauqi, A. (2023). Pemrograman Finansial Untuk Memprediksi Volatilitas Nilai Mata Uang Kripto Berbasis Deep Learning Melalui Implementasi Metode LSTM (Studi Kasus : Bitcoin , Ethereum , Tether Dan Binance Coin). *Jurnal Sistem Informasi dan Teknologi Peradaban (JSITP)*, *4*(1), 16–23.

Gusthvi, W., Roza, A. A., Mallisa, A. W., & Aryanto. (2022). Analisis Time Series Menggunakan Model Hybrid ARIMA-SVR pada Harga Bitcoin. *Cenderawasih Journal of Statistics and Data Science*, *1*(1), 16–25.

Hsu, C., Chang, C., & Lin, C. (2004). A Practical Guide to Support Vector Classification. *Department of Computer Science National Taiwan University*, *2*(3), 1–20.

Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and Practice* (Third Edit). Monash University, Australia. https://otexts.org/fpp2/

Karnadi, E. V., & Prastyo, D. D. (2023). Peramalan Harga Bitcoin Ethereum, dan Binance Menggunakan Metode Arima (Autoregressive Integrated Moving Average) dengan Variabel Eksogen. *Jurnal Sains dan Seni ITS*, *12*(5), 326–332. https://doi.org/10.12962/j23373520.v12i5.132855

Levis, A. A., & Papageorgiou, L. G. (2005). Customer Demand Forecasting Via Support Vector Regression Analysis. *Chemical Engineering Research and Design*, *83*(A8), 1009–1018. https://doi.org/10.1205/cherd.04246

Makridakis, S., Wheelwright, S. C., & McGee, V. E. (1983). *Forecasting Methods and Applications* (Second Edi). John Wiley & Sons, Inc., Hoboken, New Jersey.

Montgomery, D. C., Jennings, C. L., & Murat, K. (2008). *Introduction to Time Series Analysis and Forecasting*. John Wiley & Sons, Inc., Hoboken, New Jersey.

Pai, P. F., & Lin, C. S. (2005). A Hybrid ARIMA and Support Vector Machines Model in Stock Price Forecasting. *Omega the International Journal of Management Science*, *33*(2005), 497–505. https://doi.org/10.1016/j.omega.2004.07.024

Purnama, D. I. (2021). Peramalan Harga Emas Saat Pandemi Covid-19 Menggunakan Model Hybrid Autoregressive Integrated Moving Average - Support Vector Regression. *Jambura Journal Of Mathematics*, *3*(1), 52–65.

Rizkilloh, M. F., & Widiyanesti, S. (2022). Prediksi Harga Crytocurrency Menggunakan Algoritma Long Short Term Memory (LSTM). *Jurnal Rekayasa Sistem dan Teknologi Informasi*, *6*(1), 25–31. https://doi.org/https://doi.org/10.29207/resti.v6i1.3630

Rosyada, N. S., Setiawan, H., & Irvani, M. H. (2025). *Klasifikasi Kelayakan Penerima Bantuan Sosial dengan Metode K-Nearest Neighbors*. *11*(02), 190–199.

Sihombing, V. C., Martha, S., & Huda, N. M. (2022). Analisis Metode Hybrid ARIMA-SVR pada Indeks Harga Saham Gabungan. *Buletin Ilmiah Math, Stat, dan Terapannya (Bimaster), 11*(3), 413–422.

Smola, A. J., & Scholkopf, B. (2004). A Tutorial on Support Vector Regression. *Statistic and Computing*, *14*(1), 199–222. https://doi.org/10.1210/me.10.7.813

Terasvirta, T., Lin, C.-F., & Granger, C. W. J. (1993). Power of the Neutral Network Linearity Test. *Journal of Time Series Analysis*, *14*(2), 209–220.

Vapnik, V. N. (2000). *The Nature of Statistical Learning Theory* (Second Edi). Springer-Verlag, New York.

Wei, W. W. S. (2006). *Time Series Analysis Univariate and Multivariate Methods* (Second Edi). Pearson Education, Inc., USA.

Yasin, H., Prahutama, A., & Utami, T. W. (2014). Prediksi Harga Saham Menggunakan Support Vector Regression dengan Algoritma Grid Search. *Media Statistika*, *7*(1), 29–35.

Yunizar, A., Rismawan, T., & Midyanti, D. M. (2023). Penerapan Metode Recurrent Neutral Network Model Gated Recurrent Unit untuk Prediksi Harga Cryptocurrency. *Jurnal Komputer dan Aplikasi*, *11*(01), 32–41.

Zhang, G. P. (2001). Time Series Forecasting Using A Hybrid ARIMA and Neural Network Model. *Neurocomputing*, *50*(17), 159–175.