

**PERAMALAN HARGA SAHAM *APPLE* DENGAN METODE *HYBRID*  
*VECTOR AUTOREGRESSIVE INTEGRATED MOVING AVERAGE*  
(*VARIMA*) - *LONG SHORT TERM MEMORY* (*LSTM*)**

(Skripsi)

Oleh

**DEA SAFITRI**



**JURUSAN MATEMATIKA  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS LAMPUNG  
BANDAR LAMPUNG  
2023**

## **ABSTRACT**

### **FORECASTING APPLE STOCK PRICE WITH THE METHOD *HYBRID VECTOR AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (VARIMA) - LONG SHORT TERM MEMORY (LSTM)***

**By**

**DEA SAFITRI**

Stock is one of the long-term financial instruments that are sold buy in the capital market and is a popular alternative for investors to make an investment. Stock prices can increase and decrease in a certain period of time so that it can give rise to certain patterns. Investors need forecasting to estimate stock prices in the future to maximize profits. One method forecasting time series data that can be used is VARIMA, however VARIMA is limited to linear patterns of time series data. Forecasting can also be carried out using deep learning methods, one of which is LSTM. The LSTM method is capable of capturing linear patterns or non-linear data patterns. By Therefore, this study uses the VARIMA-LSTM hybrid method for stock price forecasting. VARIMA-LSTM hybrid method consists of 2 models. The main thing is that the first model is a model built with predictive data from VARIMA while the second model was built with residual data from VARIMA. The best VARIMA-LSTM hybrid model in this study produce an RMSE of 34.0735 and a MAPE of 0.0297 with using splitting data 80% training data and 20% testing data.

**Keywords** : Stock, Prediction, Forecasting, VARIMA, LSTM, *Hybrid VARIMA LSTM*

## **ABSTRAK**

### **PERAMALAN HARGA SAHAM *APPLE* DENGAN METODE *HYBRID VECTOR AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (VARIMA) - LONG SHORT TERM MEMORY (LSTM)***

**Oleh**

**DEA SAFITRI**

Saham adalah salah satu instrument keuangan jangka panjang yang diperjual belikan di pasar modal dan merupakan alternatif yang populer bagi investor untuk melakukan investasi. Harga saham dapat mengalami kenaikan dan penurunan dalam jangka waktu tertentu sehingga dapat menimbulkan beberapa pola tertentu. Para investor membutuhkan peramalan untuk memperkirakan harga saham di masa yang akan datang untuk memaksimalkan keuntungan. Salah satu metode peramalan data deret waktu yang dapat digunakan adalah VARIMA, namun VARIMA terbatas pada pola linier data deret waktu. Peramalan dapat juga dilakukan menggunakan metode *deep learning*, salah satunya yaitu LSTM. Metode LSTM mampu menangkap pola linier atau pola non linier data. Oleh karena itu, penelitian ini menggunakan metode *hybrid* VARIMA-LSTM untuk peramalan harga saham. Metode *hybrid* VARIMA-LSTM terdiri dari 2 model utama yaitu model pertama merupakan model yang dibangun dengan data prediksi dari VARIMA sedangkan model kedua dibangun dengan data residual dari VARIMA. Model *hybrid* VARIMA-LSTM terbaik pada penelitian ini menghasilkan RMSE sebesar 34.0735 dan MAPE sebesar 0.0297 dengan menggunakan *splitting* data 80% data *training* dan 20% data *testing*.

**Kata kunci** : Saham, Prediksi, Peramalan, VARIMA, LSTM, *Hybrid* VARIMA  
LSTM

**PERAMALAN HARGA SAHAM *APPLE* DENGAN METODE *HYBRID VECTOR AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (VARIMA) - LONG SHORT TERM MEMORY (LSTM)***

Oleh

**DEA SAFITRI**

Skripsi

Sebagai Salah Satu Syarat untuk Memperoleh Gelar  
**SARJANA MATEMATIKA**

Pada

**Jurusan Matematika**  
**Fakultas Matematika dan Ilmu Pengetahuan Alam**



**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM**  
**UNIVERSITAS LAMPUNG**  
**BANDAR LAMPUNG**  
**2023**

Judul Skripsi : **PERAMALAN HARGA SAHAM APPLE  
DENGAN METODE *HYBRID VECTOR  
AUTOREGRESSIVE INTEGRATED MOVING  
AVERAGE (VARIMA) – LONG SHORT  
TERM MEMORY (LSTM)***

Nama Mahasiswa : **Dea Safitri**

Nomor Pokok Mahasiswa : **1917031006**

Jurusan : **Matematika**

Fakultas : **Matematika dan Ilmu Pengetahuan Alam**



1. Komisi Pembimbing

**Dian Kurniasari, S.Si., M.Sc.**  
NIP. 196903051996032001

**Prof. Drs. Mustofa Usman, M.A., Ph.D.**  
NIP. 195701011984031020

2. Ketua Jurusan Matematika

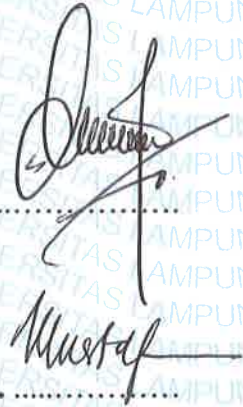
**Dr. Aang Nuryaman, S.Si., M.Si.**  
NIP. 197403162005011001



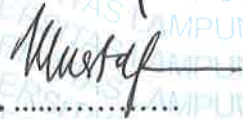
**MENGESAHKAN**

**1. Tim Penguji**

**Ketua : Dian Kurniasari, S.Si., M.Sc.**



**Sekretaris : Prof. Drs. Mustofa Usman, M.A., Ph.D.**



**Penguji  
Bukan Pembimbing : Ir. Warsono, M.S., Ph.D.**



**2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam  
Universitas Lampung**



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

**Tanggal Lulus Ujian Skripsi : 28 Juli 2023**

## PERNYATAAN

Yang bertandatangan di bawah ini :

Nama : **Dea Safitri**

Nomor Pokok Mahasiswa : **1917031006**

Jurusan : **Matematika**

Judul Skripsi : **Peramalan Harga Saham *Apple* dengan Metode *Hybrid Vector Autoregressive Integrated Moving Average (VARIMA) – Long Short Term Memory (LSTM)***

Dengan ini menyatakan bahwa penelitian 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, 28 Juli 2023  
Yang Menyatakan,



**Dea Safitri**  
NPM. 1917031006

## **RIWAYAT HIDUP**

Penulis bernama lengkap Dea Safitri lahir di Giri Rejo, 17 Desember 2002. Penulis merupakan putri dari Bapak Legito dan Ibu Titik Wahyuti. Penulis merupakan anak pertama dari dua bersaudara, memiliki adik yang bernama Muhammad Khadzik Halwani.

Penulis menempuh pendidikan di Kelompok Bermain Permata Hati pada tahun 2007 sampai 2008. Pendidikan sekolah dasar di SD Negeri 4 Wiyono pada tahun 2008 sampai 2014. Pendidikan menengah pertama di SMP Negeri 1 Pesawaran pada tahun 2014 sampai 2017. Pendidikan menengah atas di SMA Negeri 1 Gedong Tataan pada tahun 2017 sampai 2019.

Pada tahun 2019 penulis terdaftar sebagai Mahasiswi Program Studi S1 Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung melalui jalur Seleksi Nasional Masuk Perguruan Tinggi Negeri (SNMPTN). Selama menjadi mahasiswi penulis aktif di beberapa organisasi yaitu sebagai Anggota Bidang Minat dan Bakat Himpunan Mahasiswa Jurusan Matematika (HIMATIKA) tahun 2020, Anggota Koperasi Mahasiswa Unila tahun 2020, Bidang 1 Gugus Fakultas Koperasi Mahasiswa Unila 2021. Selama menjadi mahasiswi penulis juga aktif dalam kegiatan kepanitiaan Dies Natalis Jurusan Matematika (DINAMIKA) yaitu sebagai Anggota Divisi Acara DINAMIKA XXI, Anggota *Talent Squad* pada acara Anniversary Koperasi Mahasiswa Unila ke 39 dan 40, Sekretaris di acara WEBINAR KTI dan ESSAY, serta menjadi Staff PSDA Koperasi Mahasiswa Unila pada Tahun



2020. Sekarang penulis masih aktif menjadi anggota Majelis Sabuk Hitam (MSH) KKI Lampung dan menjadi pelatih karate di Dojo KKI Gedong Tataan. Pada awal tahun 2022, penulis melakukan Kerja Praktik (KP) di Kantor BPJS Kesehatan Kota Metro. Pada pertengahan tahun 2022, sebagai bentuk pengabdian kepada masyarakat, penulis melaksanakan Kuliah Kerja Nyata (KKN) selama 40 hari di Desa Sumber Mulyo, Kecamatan Sumber Rejo, Kabupaten Tanggamus.

## **KATA INSPIRASI**

“Allah tidak akan pernah membebani seseorang melainkan sesuai dengan kesanggupannya”

**(QS. Al-Baqarah : 286)**

“Karena sesungguhnya sesudah kesulitan ada kemudahan”

**(QS. Al-Insyirah : 5)**

“Terkadang Allah akan membiarkanmu merasakan pahitnya dunia, sehingga kamu dapat menghargai dan mencicipi nikmatnya iman”

**(Omar S.)**

“Kebahagiaanmu tidak ditentukan oleh orang lain, melainkan berasal dari dirimu sendiri”

**(Anonymous)**

“Apapun yang menjadi takdirmu, akan mencari jalannya menemukanmu”

**(Ali bin Abi Thalib)**

“Jangan pernah takut akan gagal karena kegagalan awal dari kesuksesan”

**(Dea Safitri)**

## **PERSEMBAHAN**

Dengan mengucap puji dan syukur kehadiran Allah SWT yang telah memberikan taufik dan hidayah-Nya untuk menyelesaikan skripsi ini, kupersembahkan karya kecil dan sederhana ini kepada :

### **Bapak dan Mamak Tercinta**

Yang selalu bekerja dengan keras agar penulis dapat menempuh pendidikan dan mendapat gelar sarjana, yang tidak pernah lelah untuk selalu mendoakan, memberikan dukungan, nasehat dan kasih sayang yang tidak mungkin terbalas oleh apapun.

### **Adik Tersayang**

Yang telah memberikan semangat, motivasi, doa dan dukungan.

### **Dosen Pembimbing dan Penguji**

Yang senantiasa meluangkan waktu untuk mengarahkan dan memotivasi penulis

### **Sahabat-sahabatku**

Yang selalu memberikan doa, dukungan, motivasi, canda dan tawa yang telah menemani penulis dalam setiap langkahnya

**Almamater Tercinta, Universitas Lampung**

## SANWACANA

Puji dan syukur penulis panjatkan atas kehadiran Allah SWT karena berkat taufik dan hidayah-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul “Peramalan Harga Saham *Apple* dengan Metode *Hybrid Vector Autoregressive Integrated Moving Average (VARIMA) – Long Short Term Memory (LSTM)*”.

Skripsi ini dibuat sebagai salah satu syarat memperoleh gelar Sarjana Matematika di Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung. Dalam proses penyusunan skripsi ini, banyak pihak yang telah membantu memberikan bimbingan, dukungan, motivasi, serta saran sehingga skripsi ini dapat terselesaikan. Oleh karena itu penulis ingin mengucapkan terima kasih kepada :

1. Ibu Dian Kurniasari, S.Si., M.Sc., selaku Pembimbing I yang selalu bersedia memberikan waktu, arahan, bimbingan, saran serta dukungan kepada penulis sehingga dapat menyelesaikan skripsi ini.
2. Bapak Prof. Drs. Mustofa Usman, M.A., Ph.D., selaku Pembimbing II yang telah bersedia memberikan waktu, arahan, bimbingan, dan dukungan kepada penulis.
3. Bapak Ir. Warsono, M.S., Ph.D., selaku Penguji yang telah bersedia memberikan kritik dan saran serta evaluasi kepada penulis.
4. Ibu Dr. Fitriani, S.Si., M.Sc., selaku Pembimbing Akademik yang selalu bersedia memberikan bimbingan, saran serta dukungan kepada penulis pada hal yang berkaitan dengan akademik.
5. Bapak Aang Nuryaman, S.Si., M.Si., selaku Ketua Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.

6. Bapak Dr. Eng. Heri Satria, S.Si., M.Si., selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
7. Seluruh dosen, staff dan karyawan Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
8. Bapak, Mamak, Adek dan keluarga besar terima kasih atas cinta, kasih sayang, doa serta dukungan yang tiada hentinya kepada penulis.
9. Sania, Madha, dan Dewi sebagai teman dan saudara yang sering direpotkan oleh penulis.
10. Alfa Khoirul Umami, Fitriana, Risa, Putri, Candra, dan Vista yang selalu memberikan semangat, saran dan menjadi tempat berkeluh kesah penulis.
11. Teman-teman seperjuangan yaitu Nurul, Irma, Mia, Tiara, Azza, Zida, Anin, Silvi, Adel, Clara dan Fiqih yang selalu berbagi suka duka, saling membantu dan saling memotivasi.
12. Anisa Fitriyani, Silvia Rizki Aulia, Rara Gusti Rahmawati, Irma Destiana, dan Luailiyatuzzahrok yang telah membuat perkuliahan penulis lebih berwarna.
13. Fatihul Qulub, Karate Dojo KKI Gedong Tataan, dan IPNU IPPNU Pesawaran sebagai organisasi yang telah menghibur dan memberi semangat penulis.
14. Teman-teman Jurusan Matematika Angkatan 2019.
15. Semua pihak yang telah membantu penulis dalam menyelesaikan skripsi ini.

Penulis menyadari bahwa skripsi ini masih jauh dari kesempurnaan. Oleh karena itu, penulis mengharapkan kritik dan saran yang membangun. Semoga skripsi ini dapat memberikan manfaat bagi pembaca.

Bandar Lampung, 28 Juli 2023  
Penulis

**Dea Safitri**



## DAFTAR ISI

Halaman

<b>DAFTAR TABEL.....</b>	<b>vi</b>
--------------------------	-----------

<b>DAFTAR GAMBAR.....</b>	<b>viii</b>
---------------------------	-------------

<b>I. PENDAHULUAN .....</b>	<b>1</b>
1.1 Latar Belakang dan Masalah.....	1
1.2 Tujuan .....	5
1.3 Manfaat .....	5
<b>II. TINJAUAN PUSTAKA.....</b>	<b>6</b>
2.1 Peramalan.....	6
2.2 Analisis <i>Time Series</i> .....	6
2.3 Model <i>Time Series</i> Univariat .....	7
2.3.1 <i>Autoregressive (AR)</i> .....	7
2.3.2 <i>Moving Average (MA)</i> .....	8
2.3.3 <i>Autoregressive Moving Average (ARMA)</i> .....	8
2.3.4 <i>Autoregressive Intergated Moving Average (ARIMA)</i> .....	9
2.4 Model <i>Time Series</i> Multivariat .....	10
2.4.1 <i>Vector Autoregressive (VAR)</i> .....	10
2.4.2 <i>Vector Moving Average (VMA)</i> .....	10
2.4.3 <i>Vector Autoregressive Moving Average (VARMA)</i> .....	11
2.4.4 <i>Vector Autoregressive Intergated Moving Average (VARIMA)</i> ...	11
2.5 Stasioneritas Data.....	12
2.6 <i>Scaling Data</i> .....	13
2.7 Uji Asumsi Residual .....	15
2.7.1 Uji Asumsi Residual <i>White Noise</i> .....	15
2.7.2 Uji Asumsi Distribusi Normal Multivariat .....	16
2.8 <i>Recurrent Neural Network (RNN)</i> .....	17
2.9 <i>Long Short Term Memory (LSTM)</i> .....	18
2.10 <i>Hyperparameter</i> .....	21
2.11 <i>Hybrid VARIMA-LSTM</i> .....	22

2.12 Fungsi Aktivasi .....	23
2.13 Evaluasi Model .....	25
2.13.1 <i>Akaike Information Criteria (AIC)</i> .....	25
2.13.2 <i>Mean Absolute Percentage Error (MAPE)</i> .....	26
2.13.3 <i>Root Mean Square Error (RMSE)</i> .....	26
2.14 Saham.....	27
<b>III. METODOLOGI PENELITIAN .....</b>	<b>28</b>
3.1 Waktu dan Tempat Penelitian .....	28
3.2 Data Penelitian .....	28
3.3 Metode Penelitian .....	29
<b>IV. PEMBAHASAN .....</b>	<b>32</b>
4.1. Proses <i>Input Data</i> .....	32
4.2. Prediksi Data dengan VARIMA .....	34
4.2.1 <i>Preprocessing Data</i> .....	35
4.2.2 Uji Stasioneritas Data .....	35
4.2.3 Pembentukan Model VARIMA .....	38
4.2.4 Estimasi Parameter.....	40
4.2.5 Prediksi VARIMA .....	41
4.2.6 Residual VARIMA .....	43
4.2.7 Uji Asumsi Residual VARIMA .....	44
4.2.8 Peramalan VARIMA .....	45
4.3. Prediksi <i>Hybrid VARIMA-LSTM</i> .....	47
4.3.1 <i>Splitting data</i> .....	47
4.3.2 <i>Scaling data</i> .....	48
4.3.3 Membangun Model <i>Hybrid</i> Prediksi VARIMA - LSTM .....	48
4.3.4 <i>Hyperparameter Tuning</i> Model <i>Hybrid</i> Prediksi VARIMA-LSTM..	49
4.3.5 Membangun Model <i>Hybrid</i> Residual VARIMA-LSTM .....	50
4.3.6 <i>Hyperparameter Tuning</i> Model <i>Hybrid</i> Residual VARIMA-LSTM .	51
4.3.7 Prediksi Model <i>Hybrid</i> Prediksi VARIMA-LSTM .....	52
4.3.8 Prediksi Model <i>Hybrid</i> Residual VARIMA-LSTM .....	56
4.4. Prediksi <i>Hybrid</i> .....	59
4.5. Hasil Peramalan <i>Hybrid VARIMA-LSTM</i> .....	62
4.5.1 Hasil Peramalan Model <i>Hybrid</i> Prediksi VARIMA-LSTM .....	62
4.5.2 Hasil Peramalan Model <i>Hybrid</i> Residual VARIMA-LSTM .....	64
4.5.3 Hasil Peramalan Model <i>Hybrid</i> .....	65
4.6 Perbandingan Peramalan VARIMA dan <i>Hybrid VARIMA-LSTM</i> .....	68
<b>V. KESIMPULAN .....</b>	<b>69</b>
<b>DAFTAR PUSTAKA .....</b>	<b>70</b>

## DAFTAR TABEL

Tabel	Halaman
1. Data Harga Saham <i>Apple</i> .....	29
2. <i>Missing Value</i> .....	35
3. Uji Stasioner Awal.....	37
4. Uji Stasioner Setelah <i>Differencing</i> .....	38
5. Pemilihan Model Terbaik VARIMA.....	41
6. Plot Data Prediksi VARIMA.....	42
7. Uji Ljung-Box.....	44
8. Uji <i>Jarque-Bera Test of Normality</i> .....	45
9. Plot Data Peramalan VARIMA.....	46
10. <i>Splitting</i> Data.....	47
11. Plot Data Prediksi Model <i>Hybrid</i> Prediksi VARIMA- LSTM untuk <i>splitting</i> data 70% data <i>training</i> dan 30% data <i>testing</i> .....	53
12. Plot Data Prediksi Model <i>Hybrid</i> Prediksi VARIMA- LSTM untuk <i>splitting</i> data 80% data <i>training</i> dan 20% data <i>testing</i> .....	54
13. Plot Data Prediksi Model <i>Hybrid</i> Residual VARIMA- LSTM untuk <i>splitting</i> data 70% data <i>training</i> dan 30% data <i>testing</i> .....	56
14. Plot Data Prediksi Model <i>Hybrid</i> residual VARIMA- LSTM untuk <i>splitting</i> data 80% data <i>training</i> dan 20% data <i>testing</i> .....	57
15. Plot Data Prediksi Model <i>Hybrid</i> untuk <i>splitting</i> data 70% data <i>training</i> dan 30% data <i>testing</i> .....	59
16. Plot Data Prediksi Model <i>Hybrid</i> untuk <i>splitting</i> data 80% data <i>training</i> dan 20% data <i>testing</i> .....	60
17. Plot Data Peramalan Model <i>Hybrid</i> untuk <i>splitting</i> data 70% data <i>training</i> dan 30% data <i>testing</i> .....	66

18. Plot Data Peramalan Model <i>Hybrid</i> untuk <i>splitting</i> data 80% data <i>training</i> dan 20% data <i>testing</i> .....	67
19. Uji <i>T-test</i> .....	68

## DAFTAR GAMBAR

Gambar	Halaman
1. Struktur Sel LSTM.....	17
2. Diagram Alir Metode Penelitian.....	31
3. Plot Data Asli Harga <i>Open</i> Saham <i>Apple</i> .....	32
4. Plot Data Asli Harga <i>High</i> Saham <i>Apple</i> .....	33
5. Plot Data Asli Harga <i>Low</i> Saham <i>Apple</i> .....	33
6. Plot Data Asli Harga <i>Close</i> Saham <i>Apple</i> .....	34
7. Plot Data Stasioner.....	38
8. Plot ACF dan PACF Variabel <i>Open</i> .....	39
9. Plot ACF dan PACF Variabel <i>High</i> .....	39
10. Plot ACF dan PACF Variabel <i>Low</i> .....	40
11. Plot ACF dan PACF Variabel <i>Close</i> .....	40
12. Plot Data Residual VARIMA.....	44
13. <i>Hyperparameter tuning</i> model <i>Hybrid</i> prediksi VARIMA- LSTM untuk <i>splitting</i> data 70% data <i>training</i> dan 30% data <i>testing</i> .....	49
14. <i>Hyperparameter tuning</i> model <i>Hybrid</i> prediksi VARIMA- LSTM untuk <i>splitting</i> data 80% data <i>training</i> dan 20% data <i>testing</i> .....	50
15. <i>Hyperparameter tuning</i> model <i>Hybrid</i> residual VARIMA- LSTM untuk <i>splitting</i> data 70% data <i>training</i> dan 30% data <i>testing</i> .....	51
16. <i>Hyperparameter tuning</i> model <i>Hybrid</i> residual VARIMA- LSTM untuk <i>splitting</i> data 80% data <i>training</i> dan 20% data <i>testing</i> .....	52
17. Plot peramalan model <i>Hybrid</i> prediksi VARIMA- LSTM untuk <i>splitting</i> data 70% data <i>training</i> dan 30% data <i>testing</i> .....	62
18. Plot peramalan model <i>Hybrid</i> prediksi VARIMA- LSTM untuk <i>splitting</i> data 80% data <i>training</i> dan 20% data <i>testing</i> .....	63



19. Plot peramalan model <i>Hybrid</i> residual VARIMA- LSTM untuk <i>splitting</i> data 70% data <i>training</i> dan 30% data <i>testing</i> .....	64
20. Plot peramalan model <i>Hybrid</i> residual VARIMA- LSTM untuk <i>splitting</i> data 80% data <i>training</i> dan 20% data <i>testing</i> .....	65

## I. PENDAHULUAN

### 1.1 Latar Belakang dan Masalah

Perencanaan yang efektif dan efisien memerlukan alat bantu peramalan yang baik (Ningtyas, 2018). Peramalan merupakan suatu pendekatan dalam memperkirakan secara kuantitatif mengenai suatu kejadian yang akan terjadi pada beberapa periode ke depan, didasarkan dengan data histori yang saling terkait dan relevan serta telah terjadi di masa sebelumnya (Ahmad, 2020). Peramalan digunakan untuk meminimumkan ketidakpastian sebuah masalah. Salah satu metode peramalan yang banyak dikembangkan saat ini adalah metode peramalan data runtun waktu (*time series*).

Pada dasarnya terdapat dua metode peramalan yaitu metode peramalan kuantitatif dan metode peramalan kualitatif (Susilawati & Sunendiari, 2022). Metode peramalan data *time series* dapat juga digolongkan berdasarkan banyaknya peubah yang menjadi pengamatan yaitu univariat dan multivariat. Metode peramalan univariat merupakan metode yang hanya memiliki satu variabel pengamatan, sedangkan multivariat memiliki lebih dari satu variabel pengamatan dan terdapat korelasi antarvariabel.

Salah satu metode deret waktu univariat yang sedang berkembang dan umum digunakan saat ini adalah metode ARIMA. Metode ARIMA merupakan metode yang tidak mengabaikan kaidah-kaidah deret waktu seperti pengujian

stasioneritas, penaksiran parameter, dan pemeriksaan diagnostik (Susilawati & Sunendiari, 2022). Jika dibandingkan dengan metode-metode peramalan lainnya proses perhitungan menggunakan metode ARIMA cukup kompleks. ARIMA terbentuk dari tiga model yaitu AR (*Autoregressive*), MA (*Moving Average*), dan ARMA (*Autoregressive and Moving Average*). Model AR (*Autoregressive*) yaitu suatu model yang menjelaskan pergerakan suatu variabel melalui variabel itu sendiri dan model MA (*Moving Average*) yaitu model yang melihat pergerakan variabelnya melalui residual di masa lalu (Namini, dkk., 2018).

Metode data deret waktu yang dapat digunakan untuk peramalan data multivariat adalah VARIMA. VARIMA merupakan model deret waktu peubah multivariat pengembangan dari metode ARIMA (Rusyana, dkk., 2020). VARIMA digunakan untuk menjelaskan hubungan antara pengamatan dan galat pada suatu peubah pada waktu tertentu dengan pengamatan dan galat pada peubah itu sendiri dan peubah lain pada waktu sebelumnya. Mirip dengan ARIMA, VARIMA terbatas pada pola linier data deret waktu. VARIMA terbentuk dari tiga model yaitu VAR (*Vector Autoregressive*), VMA (*Vector Moving Average*), dan VARMA (*Vector Autoregressive and Moving Average*).

VAR (*Vector Autoregressive*) merupakan model deret waktu multivariat yang dapat digunakan untuk peramalan dengan beberapa variabel yang saling mempengaruhi (Dissanayeka, dkk., 2021). VAR adalah pengembangan model AR dengan melibatkan lebih dari satu variabel (Pradnyandita, dkk., 2022). VAR memiliki asumsi bahwa data atau variabel yang digunakan harus stasioner (Kaur, dkk., 2021). Jika data tidak stasioner maka data harus di ubah menjadi stationer. Selain VAR, VMA dan VARMA juga memiliki poses analisis dengan data yang harus stasioner. VARMA merupakan penerapan model ARMA ke deret waktu multivariat (Caliwag & Lim, 2019). VARMA juga terbatas pada pola linier data deret waktu. Tidak seperti VAR, VMA dan VARMA, VARIMA dapat

menggunakan data deret waktu yang tidak stasioner, yang kemudian dilakukan *differencing* agar data tersebut stasioner.

Peramalan dapat juga dilakukan menggunakan metode *deep learning*. Salah satu jenis dari metode *deep learning* adalah RNN. RNN adalah salah satu algoritma yang memiliki nilai akurasi yang cukup baik dalam melakukan prediksi data berbentuk time series. Metode RNN memiliki beberapa jenis salah satunya yaitu metode LSTM. LSTM merupakan salah satu perubahan dilakukan pada RNN dengan menambahkan memori sel yang dapat menyimpan informasi dalam jangka Panjang. Metode ini mampu mempelajari depensi jangka panjang. LSTM juga merupakan kerangka kerja komputasi yang fleksibel, digunakan untuk memodelkan masalah nonlinier. Tidak seperti VARIMA, LSTM mampu menangkap pola nonlinier dalam data (Caliwag & Lim, 2019). Namun, diperlukan sejumlah data besar untuk mempelajari nonlinier. Dengan demikian LSTM saja tidak mampu menangani pola linier dan nonlinier secara bersamaan (Caliwag & Lim, 2019).

Pengembangan prediksi model multivariat yang berdasarkan model VARIMA dan menggunakan metode LSTM yaitu metode *hybrid* VARIMA-LSTM. Model VARIMA digunakan untuk menganalisis deret waktu multivariat dan memprediksi nilai di masa depan. Kemudian hasil residu dan prediksi yang diperoleh dari hasil model VARIMA digunakan untuk model masukan metode LSTM yang secara bersamaan memperkirakan nilai prediksi masa depan. Kombinasi antara model VARIMA dan LSTM digunakan untuk mengatasi kelemahan penggunaan VARIMA dan LSTM secara independen.

Salah satu implementasi dari metode *hybrid* VARIMA-LSTM yaitu untuk meramalkan harga saham. Saham adalah salah satu instrument keuangan jangka panjang yang diperjual belikan di pasar modal dan merupakan alternatif yang populer bagi investor untuk melakukan investasi (Prasetya, dkk., 2020). Tren

investasi atau trading saham sangat meningkat dan banyak peminat, namun banyak masyarakat yang belum paham dalam memperoleh keuntungan dan lainnya (Pratama, dkk., 2022). Salah satu resiko dalam investasi saham adalah turunnya harga saham. Para investor harus melakukan analisis pembelian saham untuk mengurangi resiko tersebut. Tidak ada saham yang terus menerus mengalami kenaikan maupun penurunan. Harga saham dapat mengalami kenaikan dan penurunan dalam jangka waktu tertentu sehingga dapat menimbulkan beberapa pola tertentu. Oleh karena itu peramalan dibutuhkan para investor untuk memperkirakan harga saham di masa yang akan datang untuk memaksimalkan keuntungan.

Penelitian terdahulu menggunakan metode VARIMA dilakukan oleh Rusyana, dkk. (2020) untuk prediksi curah hujan menghasilkan nilai perkiraan curah hujan, suhu, kelembaban dan kecepatan angin tidak jauh berbeda dari nilai sebenarnya dan nilai MAPE tiap variabel rendah. Lusia dan Ambarwati (2018) melakukan penelitian tentang perbandingan peramalan univariat dan multivariat ARIMA pada indeks harga saham gabungan. Metode yang digunakan adalah ARIMA dan VARIMA menghasilkan bahwa model VARIMA merupakan model terbaik berdasarkan ketepatan RMSE testing dan ketepatan *high* maupun *low*.

Pratama, dkk. (2022) melakukan penelitian menggunakan metode LSTM untuk memprediksi harga saham menghasilkan nilai MAPE tiap variabel kurang dari 10% maka akurasi sudah sangat baik. Penelitian selanjutnya yaitu penelitian menggunakan kombinasi ARIMA-LSTM dilakukan oleh Dave, dkk. (2021) untuk peramalan ekspor Indonesia. Dari metode hybrid ARIMA-LSTM mendapatkan hasil bahwa nilai MAPE sebesar 7,38% yaitu kurang dari 10%, sehingga metode tersebut sangat baik digunakan. Kemudian penelitian yang dilakukan oleh Caliwag dan Lim (2019) dengan kombinasi metode VARMA dan LSTM untuk keamanan baterai lithium-ion dan peramalan tegangan keluaran dalam aplikasi sepeda motor listrik. Hasil kombinasi VARMA-LSTM adalah peramalan



menggunakan metode *hybrid* VARMA-LSTM memiliki nilai RMSE yang lebih kecil dibandingkan dengan VARMA atau LSTM sendiri.

Berdasarkan pemaparan di atas, maka penulis ingin meramalkan harga saham dalam penelitian yang berjudul “Peramalan Harga Saham *Apple* dengan Metode *Hybrid Vector Autoregressive Integrated Moving Average* (VARIMA) - *Long Short Term Memory* (LSTM)”.

## **1.2 Tujuan**

1. Mengetahui definisi mendasar untuk membangun model VARIMA dan LSTM terhadap harga saham *apple*.
2. Melakukan pengembangan dan mengevaluasi metode *hybrid* VARIMA-LSTM untuk mendapatkan hasil yang lebih akurat.
3. Mengetahui peramalan harga saham *apple* menggunakan metode *hybrid* VARIMA-LSTM.

## **1.3 Manfaat**

1. Memperoleh pengetahuan tentang membangun model dan prediksi harga saham *apple*.
2. Dapat mengembangkan dan mengevaluasi metode *hybrid* VARIMA-LSTM untuk mendapatkan hasil yang akurat.
3. Dapat menjadi sumber ilmu dan referensi bagi pembaca.

## **II. TINJAUAN PUSTAKA**

### **2.1 Peramalan**

Dalam kehidupan sosial segala sesuatu yang ada itu tidak pasti dan sulit untuk diperkirakan dengan tepat. Banyak keputusan penting yang dilakukan secara pribadi, instansi, maupun perusahaan untuk kejadian dimasa yang akan datang, sehingga peramalan diperlukan untuk keadaan lingkungan dimasa depan.

Peramalan berasal dari kata ramalan yang artinya suatu situasi atau kondisi yang diperkirakan akan terjadi pada masa yang akan datang. Sedangkan peramalan merupakan bentuk kegiatannya. Ramalan dapat didasarkan dengan berbagai macam metode, metode tersebut digunakan sesuai dengan informasi yang akan diramal serta tujuan yang akan dicapai. Metode peramalan adalah suatu pendekatan dalam memperkirakan secara kuantitatif mengenai suatu kejadian yang akan terjadi pada beberapa periode kedepan, didasarkan dengan data histori yang saling terkait dan relevan serta telah terjadi dimasa sebelumnya (Ahmad, 2020).

### **2.2 Analisis *Time Series***

Analisis *time series* merupakan bidang penelitian yang dapat digunakan untuk bidang bisnis, ekonomi, keuangan, dan ilmu komputer sebagai aplikasinya.

Analisis *time series* sendiri digunakan ketika data yang diperoleh berkaitan dengan waktu. Tujuan dari analisis *time series* ini adalah mempelajari pengamatan deret waktu dan membangun suatu model untuk menggambarkan struktur data dan memprediksi nilai deret waktu di masa yang akan datang (Namini, dkk., 2018). Memahami suatu pola *trend* suatu data yang terjadi dari waktu ke waktu dapat menggunakan analisis *time series*. Dengan metode analisis *time series* yang tepat dapat memprediksi suatu kejadian di masa yang akan datang dengan lebih baik. Jika suatu analisis *time series* hanya terdapat satu variabel data deret waktu maka dapat disebut *time series* univariat. Sedangkan jika memiliki lebih dari satu variabel dalam data deret waktu yang terlibat maka disebut *time series* multivariat.

## 2.3 Model Time Series Univariat

### 2.3.1 Autoregressive (AR)

*Autoregressive* (AR) merupakan model regresi yang menggunakan ketergantungan antara pengamatan dan pengamatan pada periode sebelumnya. Model AR dapat digunakan untuk memprediksi nilai masa lalu pada periode tertentu. model AR dinotasikan dengan (p). Data yang digunakan untuk model AR ini adalah data yang stasioner.

Secara umum model AR mempunyai bentuk sebagai berikut (Wei, 2006) :

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + a_t \quad (2.1)$$

dengan:

$Z_t$  = deret waktu stasioner

$Z_{t-1}, \dots, Z_{t-p}$  = nilai masa lalu yang berhubungan

$\phi_1, \dots, \phi_t$  = koefisien atau parameter dari model *autoregressive*

$a_t$  = residual pada waktu t

### 2.3.2 Moving Average (MA)

*Moving average* (MA) merupakan pendekatan yang memperhitungkan ketergantungan antara pengamatan dan kesalahan ketika model rata-rata bergerak. Model MA dinotasikan dengan (q). Data yang digunakan untuk model AR ini adalah data yang stasioner.

Secara umum model MA mempunyai bentuk sebagai berikut (Wei, 2006) :

$$Z_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_t a_{t-q} \quad (2.2)$$

dengan:

$Z_t$  = deret waktu stasioner

$a_{t-1}, \dots, a_{t-q}$  = nilai residual pada masa lampau

$\theta_1, \dots, \theta_t$  = koefisien atau parameter dari model *moving average*

$a_t$  = residual pada waktu t

### 2.3.3 Autoregressive Moving Average (ARMA)

*Autoregressive moving average* (ARMA) merupakan penggabungan dari model *autoregressive* (AR) dan *moving average* (MA). Model ARMA dinotasikan dengan (p,q). Data yang digunakan untuk model ARMA ini adalah data yang stasioner.

Secara umum model ARMA mempunyai bentuk sebagai berikut (Wei, 2006) :

$$Z_t = \phi_1 Z_{t-1} + \dots + \phi_p Z_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_t a_{t-q} \quad (2.3)$$

dengan:

$Z_t$  = deret waktu stasioner

$a_{t-1}, \dots, a_{t-q}$  = nilai residual pada masa lampau

$\theta_1, \dots, \theta_t$  = koefisien atau parameter dari model *moving average*

$Z_{t-1}, \dots, Z_{t-p}$  = nilai masa lalu yang berhubungan

$\phi_1, \dots, \phi_t$  = koefisien atau parameter dari model *autoregressive*

$a_t$  = residual pada waktu t

### 2.3.4 *Autoregressive Intergated Moving Average (ARIMA)*

*Autoregressive integrated moving average (ARIMA)* merupakan model umum dari ARMA yang mengalami proses *diifrencing*. Terdapat tiga parameter yang mewakili ARIMA yaitu model (AR, I, MA) sehingga dapat dinotasikan (p, d, q). ARIMA ini terjadi ketika data deret waktu yang tidak stasioner.

Secara umum model ARIMA mempunyai bentuk sebagai berikut (Wei, 2006) :

$$Z_t = \phi_1 Z_{t-1} + \dots + \phi_p Z_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (2.4)$$

dengan:

$Z_t$  = deret waktu stasioner

$a_{t-1}, \dots, a_{t-q}$  = nilai residual pada masa lampau

$\theta_1, \dots, \theta_t$  = koefisien atau parameter dari model *moving average*

$Z_{t-1}, \dots, Z_{t-p}$  = nilai masa lalu yang berhubungan

$\phi_1, \dots, \phi_t$  = koefisien atau parameter dari model *autoregressive*

$a_t$  = residual pada waktu t



## 2.4 Model *Time Series* Multivariat

### 2.4.1 *Vector Autoregressive* (VAR)

*Vector autoregressive* (VAR) merupakan generalisasi dari *autoregressive* (AR) dengan jumlah variabel lebih dari satu. VAR merupakan sistem persamaan dinamis, dengan estimasi periode waktu yang dimodifikasi tergantung pada pergerakan variabel dan variabel lain yang terlibat pada periode sebelumnya (Rusyana, dkk., 2020). VAR ini menggunakan data yang harus stasioner.

Secara umum model VAR mempunyai bentuk sebagai berikut (Wei, 2006) :

$$\mathbf{Z}_t = \boldsymbol{\phi}_1 \mathbf{Z}_{t-1} + \boldsymbol{\phi}_2 \mathbf{Z}_{t-2} + \dots + \boldsymbol{\phi}_p \mathbf{Z}_{t-p} + \mathbf{a}_t \quad (2.5)$$

dengan:

$$\begin{aligned} \mathbf{Z}_t &= [Z_{1,t}, Z_{2,t}, \dots, Z_{N,t}]^T \text{ vektor deret waktu stasioner ukuran } N \times 1 \\ \boldsymbol{\phi}_1, \dots, \boldsymbol{\phi}_t &= \text{matriks parameter } \textit{autoregressive} \text{ berukuran } N \times N \\ \mathbf{a}_t &= [a_{1,t}, a_{2,t}, \dots, a_{N,t}]^T \text{ vektor error ukuran } N \times 1 \end{aligned}$$

### 2.4.2 *Vector Moving Average* (VMA)

*Vector moving average* (VMA) merupakan generalisasi dari *moving average* (MA) dengan jumlah variabel lebih dari satu. VMA merupakan sistem persamaan dinamis, dengan estimasi periode waktu yang dimodifikasi tergantung pada kesalahan variabel dan kesalahan lain dari variabel yang terlibat pada periode sebelumnya (Rusyana, dkk., 2020). VMA ini menggunakan data yang harus stasioner.

Secara umum model VMA mempunyai bentuk sebagai berikut (Wei, 2006) :

$$\mathbf{Z}_t = \mathbf{a}_t - \boldsymbol{\theta}_1 \mathbf{a}_{t-1} - \boldsymbol{\theta}_2 \mathbf{a}_{t-2} - \dots - \boldsymbol{\theta}_t \mathbf{a}_{t-q} \quad (2.6)$$

dengan:

$\mathbf{Z}_t = [Z_{1,t}, Z_{2,t}, \dots, Z_{N,t}]^T$  vektor deret waktu stasioner ukuran  $N \times 1$

$\mathbf{a}_t = [a_{1,t}, a_{2,t}, \dots, a_{N,t}]^T$  vektor error ukuran  $N \times 1$

$\theta_1, \dots, \theta_t =$  matriks parameter *moving average* berukuran  $N \times N$

### 2.4.3 Vector Autoregressive Moving Average (VARMA)

*Vector autoregressive moving average* (VARMA) merupakan model deret waktu yang dapat digunakan untuk lebih dari satu variabel. VARMA adalah kombinasi dari orde p pada model VAR dan orde q pada model VMA sehingga menjadi sebuah model VARMA (p,q) (Aulia, dkk., 2021). Data yang digunakan dalam model VARMA adalah data yang stasioner.

Secara umum model VARMA mempunyai bentuk sebagai berikut (Wei, 2006) :

$$\mathbf{Z}_t = \phi_1 \mathbf{Z}_{t-1} + \dots + \phi_p \mathbf{Z}_{t-p} + \mathbf{a}_t - \theta_1 \mathbf{a}_{t-1} - \dots - \theta_t \mathbf{a}_{t-q} \quad (2.7)$$

dengan:

$\mathbf{Z}_t = [Z_{1,t}, Z_{2,t}, \dots, Z_{N,t}]^T$  vektor deret waktu stasioner ukuran  $N \times 1$

$\mathbf{a}_t = [a_{1,t}, a_{2,t}, \dots, a_{N,t}]^T$  vektor error ukuran  $N \times 1$

$\phi_1, \dots, \phi_t =$  matriks parameter *autoregressive* berukuran  $N \times N$

$\theta_1, \dots, \theta_t =$  matriks parameter *moving average* berukuran  $N \times N$

### 2.4.4 Vector Autoregressive Intergated Moving Average (VARIMA)

*Vector autoregressive integrated moving average* (VARIMA) merupakan model VARMA yang mengalami proses *diiferencing* (Rusyana, dkk., 2020). VARIMA terjadi ketika data yang diperoleh tidak stasioner, sehingga diperlukan proses *differencing* agar data yang digunakan menjadi data yang stasioner.

Secara umum model VARIMA mempunyai bentuk sebagai berikut (Wei, 2006) :

$$\mathbf{Z}_t = \boldsymbol{\phi}_1 \mathbf{Z}_{t-1} + \dots + \boldsymbol{\phi}_p \mathbf{Z}_{t-p} + \mathbf{a}_t - \boldsymbol{\theta}_1 \mathbf{a}_{t-1} - \dots - \boldsymbol{\theta}_q \mathbf{a}_{t-q} \quad (2.8)$$

dengan:

$\mathbf{Z}_t = [Z_{1,t}, Z_{2,t}, \dots, Z_{N,t}]^T$  vektor deret waktu stasioner ukuran  $N \times 1$

$\mathbf{a}_t = [a_{1,t}, a_{2,t}, \dots, a_{N,t}]^T$  vektor error ukuran  $N \times 1$

$\boldsymbol{\phi}_1, \dots, \boldsymbol{\phi}_t$  = matriks parameter *autoregressive* berukuran  $N \times N$

$\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_t$  = matriks parameter *moving average* berukuran  $N \times N$

## 2.5 Stasioneritas Data

Stasioneritas data dibutuhkan dalam suatu analisis data *time series* untuk melakukan analisis selanjutnya (Aulia, dkk., 2021). Dikatakan deret waktu yang stasioner jika tidak terjadi kenaikan atau penurunan nilai secara tajam dan fluktuasi data di sekitar rata-rata dan varian. Dengan kata lain data tersebut bergerak stabil dan konvergen di sekitar rata-rata dengan simpangan baku yang kecil. Jika data tersebut tidak memenuhi asumsi stasioner maka harus dimodifikasi agar data tersebut menjadi data stasioner. Stasioner data dapat dilihat secara visual melalui plot data. Selain menggunakan plot, stasioner data dapat dilihat dengan melakukan uji stasioner.

Uji stasioneritas data dapat dilakukan dengan menggunakan uji akar unit. Salah satu metode akar unit yang dapat digunakan yaitu *Augmented Dickey-Fuller test*. Berdasarkan uji akar unit yaitu *Augmented Dickey-Fuller test* pada data untuk menguji kestasioneran data yakni sebagai berikut:

1. Hipotesis:

- $H_0: \rho = 1$  deret waktu tidak stasioner
- $H_1: \rho < 1$  deret waktu stasioner

2. Taraf signifikansi

- $\alpha = 5\% = 0,05$

3. Kriteria uji:

- Tolak  $H_0$  jika  $ADF_{hitung} > ADF_{tabel}$  atau  $p\text{-value} < \alpha$
- Tidak tolak  $H_0$  jika  $ADF_{hitung} < ADF_{tabel}$  atau  $p\text{-value} > \alpha$

4. Statistik uji:

$$ADF_{hitung} = \frac{\delta}{Se(\delta)} \quad (2.9)$$

dengan :

$\delta$  = nilai dugaan parameter *autoregressive*

$Se(\delta)$  = *standar error* dari  $\delta$

5. Keputusan

6. Kesimpulan

Data yang tidak stasioner dapat diubah menjadi data yang stasioner dengan melakukan proses *differencing*. Proses *differencing* merupakan perubahan data saat ini dikurang dengan data sebelumnya sehingga diperoleh  $d=1$ , dapat dilihat pada persamaan berikut:

$$\Delta Z_t = Z_t - Z_{t-1} \quad (2.10)$$

dengan:

$\Delta Z_t$  = *differencing* data

$Z_t$  = data periode saat ini

$Z_{t-1}$  = data periode sebelumnya

## 2.6 *Scaling* Data

*Scaling* data merupakan teknik mengubah nilai numerik dataset menjadi skala umum. Teknik digunakan untuk meminimalkan suatu kesalahan atau *error*. *Scaling* data ini juga berfungsi untuk memperkecil ukuran data tanpa mengubah

data aktual. Selain itu, *scaling* data dilakukan untuk menghindari penyimpangan dan ketidak konsistena data (Aldi, dkk., 2018). *Scaling* data ini digunakan pada bagian *preprocessing*. Teknik yang dapat digunakan pada *scaling* data adalah sebagai berikut:

### 1. *Min-max scaler*

*Min-max scaler* merupakan teknik yang dilakukan pada dataset dengan mengubah data aktual menjadi nilai dengan *range* interval [0,1]. Teknik ini digunakan jika standar deviasinya kecil dan ketika bukan distribusi normal. Teknik *scaling* data pada *min-max scaler* menggunakan persamaan sebagai berikut (Aldi, dkk., 2018):

$$X' = \frac{(x - \min_x)}{(\max_x - \min_x)} \quad (2.11)$$

dengan:

$X'$  = data hasil normalisasi

$x$  = data asli

$\min_x$  = nilai minimum dari data x

$\max_x$  = nilai maximum dari data x

### 2. *Standardscaler*

Teknik *Standardscaler* didasarkan pada nilai *mean* dan standar deviasinya (Ambarwari, dkk., 2020). *Standardscaler* pada data set merupakan teknik yang mengubah skala distribusi nilai, sehingga nilai *mean* adalah 0 dan standar deviasinya adalah 1. Teknik ini digunakan jika terdapat asumsi data berdistribusi normal. Teknik *scaling* data pada *standardscaler* menggunakan persamaan sebagai berikut (Ambarwari, dkk., 2020):

$$X' = \frac{x - \text{mean}_x}{\text{std}_x} \quad (2.12)$$

dengan:

- $X'$  = data hasil normalisasi  
 $x$  = data asli  
 $mean_x$  = nilai rata-rata dari data x  
 $std_x$  = nilai standar deviasi dari data x

## 2.7 Uji Asumsi Residual

Uji asumsi residual digunakan untuk mendapatkan kelayakan pada model. Model univariat dan multivariat memiliki asumsi yang sebaiknya dipenuhi yaitu residual *white noise* dan berdistribusi normal. Namun pengujian asumsi ini dapat diabaikan dalam melakukan peramalan, karena pada peramalan hal yang perlu diperhatikan yaitu kemampuan model dalam melakukan peramalan.

### 2.7.1 Uji Asumsi Residual *White Noise*

Uji asumsi residual *white noise* memiliki arti bahwa tidak ada korelasi dari residual dalam model atau dapat dikatakan residual dari model tersebut saling bebas (Rosyidah, dkk., 2017). Pengujian *white noise* dapat dilakukan menggunakan uji Ljung-Box. Berdasarkan uji Ljung-Box pada data untuk menguji korelasi data residual yakni sebagai berikut:

1. Hipotesis:
  - $H_0$  = tidak ada korelasi dari residual
  - $H_1$  = terdapat korelasi dari residual
2. Taraf signifikansi
  - $\alpha = 5\% = 0,05$
3. Kriteria uji:
  - Tolak  $H_0$  jika  $Q_{hitung} > X_{tabel}^2$  atau  $p\text{-value} < \alpha$

- Tidak tolak  $H_0$  jika  $Q_{hitung} < X_{tabel}^2$  atau  $p\text{-value} > \alpha$

4. Statistik uji:

$$Q_{hitung} = n(n + 2) \sum_{i=1}^m \frac{\rho_i^2}{(n - i)} \quad (2.13)$$

dengan :

$n$  = banyaknya data pengamatan

$\rho_i$  = autokorelasi lag ke  $i$

$m$  = jumlah lag

5. Keputusan

6. Kesimpulan

### 2.7.2 Uji Asumsi Distribusi Normal Multivariat

Uji asumsi distribusi normal multivariat memiliki arti bahwa residual dari peramalan memiliki distribusi normal. Pengujian distribusi normal dapat dilakukan menggunakan uji *Jarque-Bera Test of Normality*. Pengujian data residual normal multivariat yakni sebagai berikut ( Jarque & Bera, 1987):

1. Hipotesis:

- $H_0$  = residual berdistribusi normal multivariat
- $H_1$  = residual tidak berdistribusi normal multivariat

2. Taraf signifikansi

- $\alpha = 5\% = 0,05$

3. Kriteria uji:

- Tolak  $H_0$  jika  $JB_{hitung} > X_{tabel}^2$  atau  $p\text{-value} < \alpha$
- Tidak tolak  $H_0$  jika  $JB_{hitung} < X_{tabel}^2$  atau  $p\text{-value} > \alpha$

4. Statistik uji:

$$JB_{hitung} = \left[ \frac{N}{6} b_1^2 + \frac{N}{24} (b_2 - 3)^2 \right] \quad (2.14)$$

dengan :

$N$  = banyaknya data pengamatan

$b$  = *expected skewness*

5. Keputusan
6. Kesimpulan

## 2.8 *Recurrent Neural Network (RNN)*

*Recurrent Neural Network (RNN)* merupakan salah satu jenis dari *Artificial Neural Network* untuk memproses data sekunsial, seperti data teks, audio, dan video. Fitur khas dari RNN merupakan koneksi siklik yang memungkinkan RNN untuk memiliki kapasitas untuk memperbaharui keadaan saat ini berdasarkan status masa lalu dan data *input* saat ini (Yu, dkk., 2019) . RNN juga merupakan jaringan yang terdiri dari sel berulang standar seperti sel sigma dan sel tanh. RNN disebut berulang karena RNN melakukan tugas yang sama untuk setiap elemen urutan, dengan karakteristik memanfaatkan informasi yang ditangkap sebelumnya untuk memprediksi data berurutan yang tidak terlihat di masa depan (Namini, dkk., 2018).

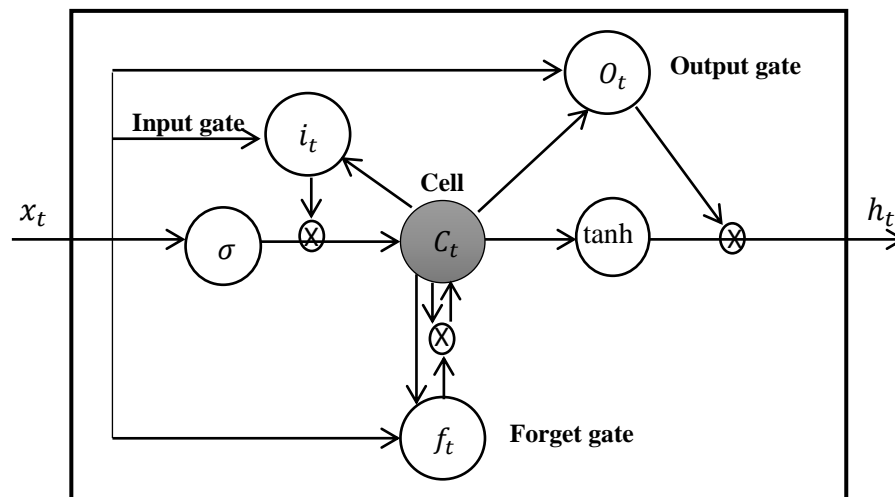
RNN memiliki struktur yang terdiri dari *input layer*, *hidden layer*, dan *output layer*. Pada dasarnya model RNN memiliki satu arah aliran informasi dari unit *input* ke unit tersembunyi, dan sintesis aliran informasi satu arah dari unit penyembunyian sementara sebelum ke unit penyembunyian saat ini (Yin, dkk., 2017). RNN dapat digunakan untuk *supervised classification learning*. Ketika memiliki interval waktu jangka panjang RNN tidak mampu menyimpan informasi sebelumnya dengan baik, hal itu disebabkan karena masalah gradien menghilang (Tian, dkk., 2018). Hochreiter & Schmidhuber (1997) mengusulkan *Long Short Term Memory (LSTM)* untuk mengatasi kelemahan ini dan meningkatkan kinerja RNN menjadi lebih baik.



## 2.9 Long Short Term Memory (LSTM)

Hochreiter & Schmidhuber mengusulkan *Long Short Term Memory* (LSTM) pada tahun 1997 yang merupakan pengembangan dari RNN. LSTM muncul karena terdapat ketidakpuasan dalam RNN untuk memproses data jangka Panjang. LSTM juga dikenal sebagai pengembangan dari RNN karena memiliki struktur yang sama. Struktur LSTM terdiri dari *input layer*, *hidden layer*, dan *output layer*. Perbedaan struktur LSTM dengan RNN terdapat pada *hidden layer*. Pada RNN hanya terdiri dari satu *layer* sederhana dengan fungsi aktivasi tanh, sedangkan LSTM terdiri dari dua fungsi aktivasi yakni fungsi aktivasi tanh dan fungsi aktivasi sigmoid.

LSTM menggabungkan memori jangka pendek dengan memori jangka panjang melalui kontrol gerbang (Tian, dkk., 2018). Struktur gerbang merupakan sebuah cara model LSTM dalam menyaring informasi mempertahankan dan memperbaharui keadaan sel memori. Unit umum LSTM terdiri dari *input gate*, *output gate*, dan *forget gate* Setiap sel memori memiliki satu lapisan tanh dan tiga lapisan sigmoid (Qiu, dkk., 2020).



Gambar 1. Struktur sel pada LSTM (Sumber: Chung & Shin, 2018)

*Forget gate* merupakan gerbang pertama dalam LSTM untuk menyaring informasi yang akan dihilangkan atau dibuang dari status sel. Keputusan ini dibuat oleh lapisan sigmoid. Dapat dilihat pada sel memori yang menerima *output*  $h_{t-1}$  dari waktu sebelumnya dan informasi eksternal  $x_t$  dari waktu saat ini serta digabungkan menjadi vektor panjang  $[h_{t-1}, x_t]$ . Hasil *output* gerbang ini menampilkan angka antara 0 dan 1 untuk setiap angka dalam *cell state*  $C_{t-1}$ . Untuk angka 1 mewakili “sepenuhnya simpan” sedangkan angka 0 mewakili “sepenuhnya singkirkan”. Rumus yang digunakan pada gerbang tersebut adalah :

$$f_t = \sigma (W_f \cdot [h_{t-1}, x_t]b_f) \quad (2.15)$$

dengan:

- $f_t$  = *forget gate*
- $\sigma$  = fungsi sigmoid
- $W_f$  = bobot pada *forget gate*
- $h_{t-1}$  = hasil *output* pada *time step* t-1
- $x_t$  = *input* pada *time step* t
- $b_f$  = bias pada *forget gate*

*Input gate* digunakan untuk menentukan banyaknya *input* jaringan  $x_t$  yang disimpan ke dalam *cell state*  $C_t$ . *Input gate* mempunyai dua fungsi, pertama menggunakan fungsi sigmoid untuk menentukan nilai yang akan diperbaharui. Sedangkan yang kedua yaitu menggunakan fungsi tanh untuk membuat vektor nilai baru yang kemudian ditambahkan pada nilai *cell state*  $C'_t$ . Kemudian kedua bagian tersebut digabungkan sehingga dapat membuat pembaruan informasi pada *cell state*. Rumus yang digunakan pada *input gate* adalah :

$$i_t = \sigma (W_i \cdot [h_{t-1}, x_t]b_i) \quad (2.16)$$

$$C'_t = \tanh (W_c \cdot [h_{t-1}, x_t]b_c) \quad (2.17)$$

dengan:

- $i_t$  = *input gate*
- $\sigma$  = fungsi sigmoid
- $W_i$  = bobot pada *input gate*

- $h_{t-1}$  = hasil *output* pada *time step* t-1  
 $x_t$  = *input* pada *time step* t  
 $b_i$  = bias pada *input gate*  
 $C'_t$  = kandidat nilai *cell state* baru yang akan ditambahkan ke  $C_{t-1}$   
 $\tanh$  = fungsi tanh  
 $W_c$  = bobot pada operasi *cell state* baru  
 $b_c$  = bias pada operasi *cell state* baru

Pembaharuan pada *cell state*  $C_{t-1}$  menjadi *cell state* baru  $C_t$  menggunakan hasil yang di dapat melalui operasi *forget gate* dan *input gate*. Melakukan operasi perkalian antara *forget gate*  $f_t$  dengan *cell state* waktu sebelumnya. Kemudian hasil operasi perkalian tersebut ditambahkan dengan hasil dari perkalian antara *input gate*  $i_t$  dengan nilai *cell state* baru pada hasil perhitungan *input gate*  $C'_t$ . Rumus yang digunakan pada operasi ini adalah:

$$C_t = f_t * C_{t-1} + i_t * C'_t \quad (2.18)$$

dengan:

- $C_t$  = *Cell state* pada *time step* t  
 $f_t$  = *forget gate*  
 $C_{t-1}$  = *Cell state* pada *time step* t-1  
 $C'_t$  = Nilai *cell state* baru pada hasil perhitungan *input gate*  
 $i_t$  = *input gate*

*Output gate* didasarkan pada *cell state* tetapi menjadi versi yang sudah disaring. Pada *output gate* yang pertama dilakukan yaitu menjalankan fungsi sigmoid untuk menentukan bagian dari *cell state* yang akan dihasilkan. Kemudian menjalankan fungsi tanh pada *cell state* dan melakukan perkalian dengan keluaran dari sigmoid *gate*, sehingga hanya menampilkan bagian keluaran akhir. Rumus yang digunakan pada *output gate* tersebut adalah:

$$o_t = \sigma ( W_o \cdot [h_{t-1}, x_t] b_o ) \quad (2.19)$$

$$h_t = o_t * \tanh(C_t) \quad (2.20)$$

dengan:

- $o_t$  = *output gate*
- $\sigma$  = fungsi sigmoid
- $W_o$  = bobot pada *output gate*
- $h_{t-1}$  = hasil *output* pada *time step* t-1
- $x_t$  = *input* pada *time step* t
- $b_o$  = bias pada *output gate*
- $h_t$  = hasil *output* pada *time step* t
- $\tanh$  = fungsi tanh

## 2.10 Hyperparameter

*Hyperparameter* harus dipertimbangkan saat membangun model karena dapat mempengaruhi kinerja model dan nilai akurasinya. Beberapa *hyperparameter* pada *deep learning* adalah sebagai berikut (Hikmaturokhman, dkk., 2022):

- a. Jumlah *hidden layer*.
- b. Jumlah *hidden neuron* pada setiap *hidden layer*.
- c. Ukuran *batch* (mewakili jumlah data *train* selama periode)
- d. Iterasi, misalnya terdapat 10.000 dataset dengan ukuran *batch* 200 maka satu periode terdiri dari 50 iterasi (10.000 dibagi 200).
- e. *Epoch* (mewakili satu set iterasi).
- f. Tingkat pembelajaran.
- g. Parameter regularisasi (keteraturan parameter)

*Hyperparameter* merupakan variabel konfigurasi di luar model yang nilainya sulit diperkirakan dari data. Artinya, *hyperparameter* tidak dapat dipelajari langsung dari data dalam standar model *training*. Sebagai gantinya, *machine learning engineer* harus dapat menentukan *hyperparameter* sebelum *training* dan

dilakukan *trial and error* untuk memperoleh nilai prediksi yang terbaik (Hikmaturokhman, dkk., 2022).

### 2.11 Hybrid VARIMA-LSTM

Metode *hybrid* merupakan pengembangan metode peramalan dengan menggabungkan dua metode atau lebih. Metode *hybrid* ini diharapkan dapat bekerja sesuai dengan keahlian dari masing-masing metode sehingga dapat mengatasi kelemahan. Kemudian model yang dihasilkan dapat memberikan prediksi yang lebih akurat dibanding model individual (Dave, dkk., 2021). Diasumsikan bahwa model kombinasi deret waktu terdapat dua komponen dasar, yaitu komponen linier dan nonlinier yang dapat dituliskan sebagai berikut (Zhang, 2003):

$$Y_t = L_t + NL_t + a_t \quad (2.21)$$

dengan :

- $Y_t$  = data pengamatan deret waktu ke  $t$
- $L_t$  = komponen linear ke  $t$
- $NL_t$  = komponen nonlinier ke  $t$
- $t$  = indeks waktu
- $a_t$  = *error*

Metode VARIMA dan LSTM memiliki kelemahan dan kelebihan masing-masing. VARIMA mengalami kesulitan saat menemui waktu nonlinier. Sedangkan LSTM mampu menangkap pola nonlinier dalam data (Caliwag & Lim, 2019). Namun, diperlukan sejumlah data besar untuk mempelajari nonlinier. Sehingga LSTM saja tidak mampu menangani pola linier dan nonlinier secara bersamaan. Berdasarkan faktor tersebut model *hybrid* dapat diterapkan.

Metode *hybrid* VARIMA-LSTM memiliki dua tahap, tahap pertama yaitu penggunaan metode VARIMA untuk memodelkan bagian linier. Sedangkan tahap kedua yaitu menggunakan metode LSTM untuk memodelkan bagian nonlinier. Kemudian menggabungkan kedua model, baik dari model linier maupun nonlinier. Proses pembangunan model terdapat tiga langkah yaitu sebagai berikut

1. Model VARIMA dibangun menggunakan data aktual untuk memperkirakan komponen linier
2. Data Residual yang diperoleh dari VARIMA, kemudian digunakan pada LSTM untuk memperkirakan komponen nonlinier
3. Menggabungkan keduanya untuk melakukan prediksi dan peramalan

Persamaan peramalan *hybrid* VARIMA-LSTM dapat dituliskan yaitu sebagai berikut :

$$Y'_t = L'_t + N'_t \quad (2.22)$$

dengan :

$Y'_t$  = nilai peramalan dari model *hybrid* ke  $t$

$L'_t$  = nilai peramalan dari komponen linear ke  $t$

$N'_t$  = nilai peramalan dari komponen nonlinear ke  $t$

$t$  = indeks waktu

## 2.12 Fungsi Aktivasi

Fungsi yang menggambarkan hubungan antara tingkat aktivitas internal yang dapat berupa bentuk linier maupun non linier disebut fungsi aktivasi (Saputra & Prabowo, 2018). Fungsi aktivasi berfungsi sebagai penentu sebuah neuron harus diaktifkan atau tidak. Fungsi aktivasi bekerja dengan menunggu rangsangan dan meresponnya dengan efek tertentu, sehingga fungsi aktivasi juga sebagai penyaring nilai yang masuk menjadi keluaran dengan rentang tertentu. Rentang

tersebut mengubah data yang sebelumnya non linier menjadi data yang linier sehingga proses selanjutnya dapat dilakukan menjadi lebih mudah (Ivan & Purnomo, 2022). Fungsi aktivasi nonlinier yang sering digunakan adalah sebagai berikut :

### 1. Fungsi sigmoid

Fungsi sigmoid merupakan fungsi yang digunakan untuk melakukan pembaruan informasi pada suatu proses (Ivan & Purnomo, 2022). Pada fungsi sigmoid jika semakin kecil *input* maka output yang didapat akan mendekati nol sedangkan semakin besar *input* maka output akan mendekati nilai satu. Fungsi sigmoid menghasilkan klasifikasi biner yaitu probabilitas positif dan negative antara nilai 0 dan 1 (Pasaribu, dkk., 2020).

Fungsi sigmoid menggunakan persamaan sebagai berikut (Akbar, dkk., 2022):

$$f(x) = \frac{1}{1+e^{-x}} \quad (2.23)$$

dengan:

$$\begin{aligned} x &= \text{data} \\ e &= \text{bilangan euler} \end{aligned}$$

### 2. Fungsi tanh

Fungsi tanh adalah fungsi yang digunakan untuk menyimpan nilai baru dari proses fungsi sigmoid. Fungsi tanh merupakan fungsi nol-pusat yang lebih halus dengan rentang nilainya adalah antara -1 hingga 1. Keuntungan fungsi ini adalah memberikan output terpusat nol yang membantu proses propagasi balik.

Fungsi tanh menggunakan persamaan sebagai berikut (Akbar, dkk., 2022):

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (2.24)$$

dengan:

$$\begin{aligned} x &= \text{data} \\ e &= \text{bilangan euler} \end{aligned}$$

## 2.13 Evaluasi Model

Evaluasi model digunakan untuk menentukan apakah model yang digunakan baik atau tidak dalam menangani data tersebut. Diperlukannya evaluasi model yaitu untuk mengetahui seberapa jauh ketepatan model tersebut dalam memprediksi data yang telah diketahui. Untuk mengevaluasi ketepatan model dapat menggunakan metode yaitu *Akaike Information Criteria (AIC)*, *Mean Absolute Percentage Error (MAPE)*, dan *Root Mean Square Error (RMSE)*.

### 2.13.1 Akaike Information Criteria (AIC)

*Akaike Information Criteria (AIC)* merupakan salah satu kriteria pemilihan model terbaik dengan mempertimbangkan banyaknya parameter dalam model. Model dengan kinerja yang terbaik ditentukan berdasarkan nilai AIC minimum (Pertiwi, dkk., 2021). Rumus mendapatkan nilai AIC adalah sebagai berikut :

$$AIC(p, q) = \ln |\Sigma| + \frac{2k^2(p+q)}{T} \quad (2.25)$$

dengan:

- $ln$  = natural log
- $p$  = lag untuk AR
- $q$  = lag untuk MA
- $k$  = banyak variabel
- $T$  = banyak pengamatan
- $\Sigma$  = matriks kovarian residual



### 2.13.2 Mean Absolute Percentage Error (MAPE)

*Mean Absolute Percentage Error* (MAPE) digunakan untuk mengetahui persentase hasil pendugaan terhadap hasil aktual. Dengan kata lain MAPE merupakan rata-rata kesalahan mutlak selama periode tertentu yang kemudian dikalikan dengan 100% agar mendapatkan hasil secara persentase (Pertiwi, dkk., 2021). Rumus yang dipakai untuk menghitung *Mean Absolute Percentage Error* (MAPE) adalah :

$$MAPE = \left( \sum_{t=1}^n \left| \frac{Y_t - \hat{Y}_t}{\hat{Y}_t} \right| \times 100\% \right) \quad (2.26)$$

dengan:

- $n$  = banyaknya data yang diamati
- $Y_t$  = nilai data aktual pada periode ke- $t$
- $\hat{Y}_t$  = nilai data hasil peramalan pada periode ke- $t$

### 2.13.3 Root Mean Square Error (RMSE)

*Root Mean Square Error* (RMSE) merupakan nilai akar kuadrat dari selisih nilai prediksi dengan nilai data actual (Temur, dkk., 2019). Hasil RMSE menunjukkan seberapa jauh nilai akar kuadrat dari nilai asli dan nilai prediksi model. Rumus yang dipakai untuk menghitung *Root Mean Square Error* (RMSE) adalah :

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (Y_t - \hat{Y}_t)^2}{n}} \quad (2.27)$$

dengan:

- $n$  = banyaknya data yang diamati
- $Y_t$  = nilai data aktual pada periode ke- $t$
- $\hat{Y}_t$  = nilai data hasil peramalan pada periode ke- $t$

## 2.14 Saham

Saham adalah salah satu instrument keuangan jangka panjang yang diperjual belikan di pasar modal dan merupakan alternatif yang populer bagi investor untuk melakukan investasi (Prasetya, dkk., 2020). Tren investasi atau trading saham sangat meningkat dan banyak peminat, namun banyak masyarakat yang belum paham dalam memperoleh keuntungan dan lainnya (Pratama, dkk., 2022). Salah satu resiko dalam investasi saham adalah turunnya harga saham. Masyarakat harus melakukan analisis pembelian saham untuk mengurangi resiko tersebut. tidak ada saham yang terus menerus mengalami kenaikan maupun penurunan. Harga saham dapat mengalami kenaikan dan penurunan dalam jangka waktu tertentu sehingga dapat menimbulkan beberapa pola tertentu. Salah satu saham yang menarik para investor adalah saham *apple*.

Perusahaan *apple* pernah dinobatkan oleh majalah fortune edisi 17 Maret 2008 sebagai perusahaan yang paling dikagumi di dunia (World Most Admired Companies). *Apple* merupakan perusahaan multinasional yang bergerak pada bidang perancangan, pengembangan, dan penjualan barang – barang yang meliputi elektronik, perangkat lunak komputer serta komputer pribadi. Pada saat ini *apple* sangat terkenal dan menjadi brand yang paling diminati khususnya pada produk *iphone* dan *ipod*. Namun harga saham *apple* yang tidak menentu dari waktu ke waktu menyebabkan para investor harus menyiapkan strategi agar tidak mengalami kerugian yang banyak. Harga saham yang akan datang dapat diperkirakan dengan menggunakan metode peramalan untuk membantu investor dalam melakukan investasi.

### III. METODOLOGI PENELITIAN

#### 3.1 Waktu dan Tempat Penelitian

Penelitian ini dilaksanakan pada semester ganjil tahun ajaran 2022/2023 bertempat di Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.

#### 3.2 Data Penelitian

Data yang digunakan dalam penelitian ini merupakan data harian harga saham *apple* yang tersedia di situs online pada Yahoo *finance* dengan link <https://finance.yahoo.com/quote/AAPL?p=AAPL&.tsrc=fin-srch>. Data yang tersedia memiliki 5 variabel yaitu *open*, *high*, *low*, *close*, dan *volume*. *Open* merupakan harga pembukaan dari *apple*, *low* merupakan harga terendah dari *apple*, *high* merupakan harga tertinggi dari *apple*, *close* merupakan harga penutupan dari *apple*, dan *volume* merupakan banyaknya volume saham *apple* yang diperdagangkan. Data yang digunakan pada penelitian ini merupakan data harian yang diambil selama 5 tahun yaitu periode Januari 2017 hingga Mei 2022. Dalam penelitian ini hanya menggunakan 4 variabel yaitu variabel *open*, *high*, *low*, dan *close* karena sesuai dengan metode VARIMA untuk meramalkan harga saham. Data dapat diunduh secara langsung dalam bentuk CSV.

Tabel 1. Data Harga Saham *Apple*

<b>Periode</b>	<b>Open</b>	<b>High</b>	<b>Low</b>	<b>Close</b>	<b>Volume</b>
3/1/2017	28.950001	29.085	28.690001	29.0375	115127600
4/1/2017	28.9625	29.127501	28.9375	29.004999	84472400
5/1/2017	28.98	29.125	28.952499	29.1525	88774400
6/1/2017	29.195	29.540001	29.1175	29.477501	127007600
⋮	⋮	⋮	⋮	⋮	⋮
20/5/2022	139.089996	140.699997	132.610001	137.589996	137194600

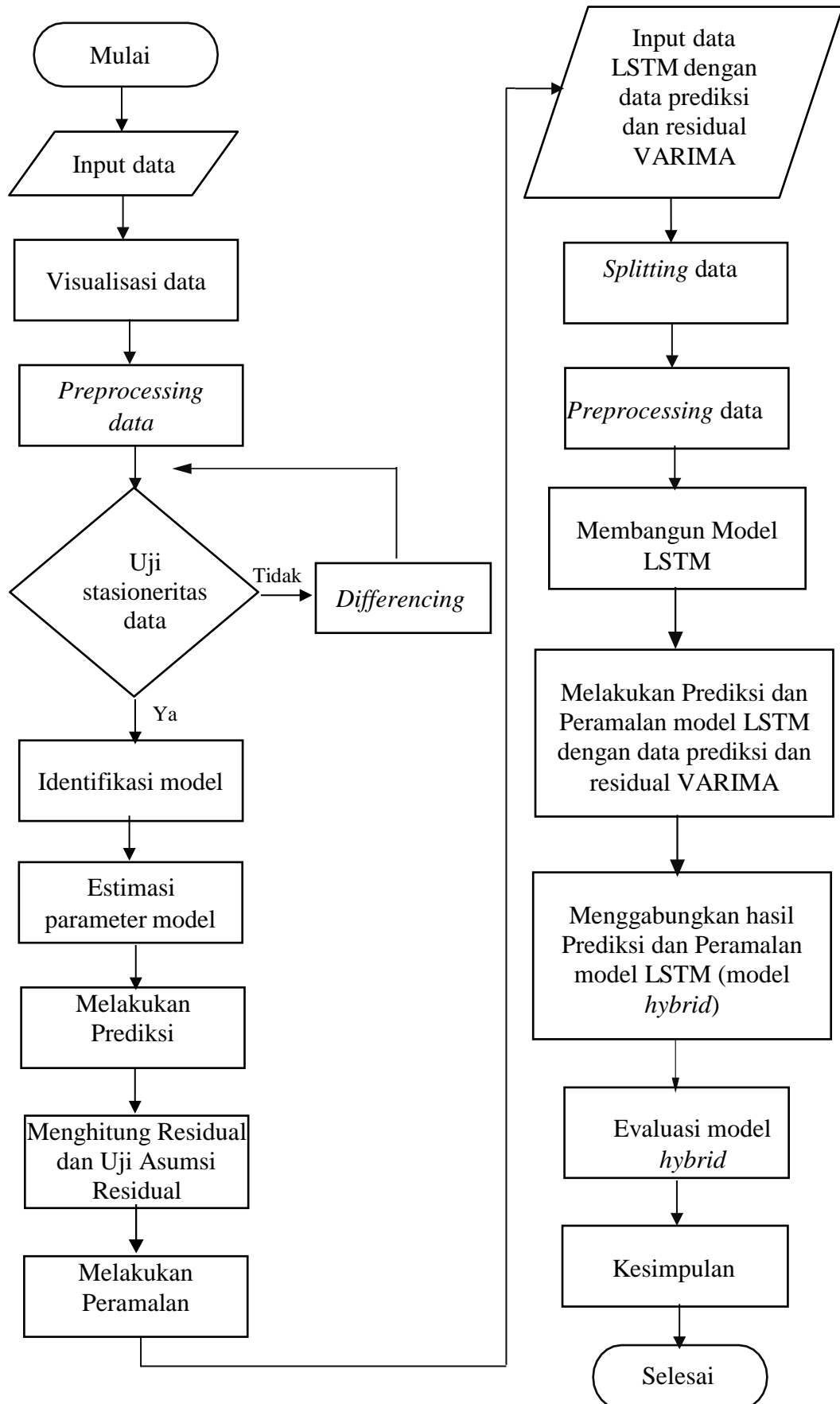
### 3.3 Metode Penelitian

Penelitian ini menggunakan studi literatur secara sistematis yang diperoleh dari buku, jurnal, dan juga media internet untuk mendapatkan informasi yang sebanyak mungkin. Untuk mempermudah perhitungan penelitian ini menggunakan program *google colab*.

Adapun tahapan yang dilakukan penelitian ini antara lain:

1. *Input* data harga saham *apple*
2. Melakukan visualisasi data  
Visualisasi data digunakan untuk melihat variabel pada data harga saham *apple*. Kemudian melihat plot data tersebut.
3. Melakukan *preprocessing* data  
*Preprocessing* data digunakan melihat ada atau tidak *missing value* pada data tersebut.
4. Melakukan uji stasioneritas data  
Uji stasioneritas data pada rata-rata menggunakan uji akar unit yaitu *Dickey-Fuller test*. Jika data tersebut *non-stasioner* maka dilakukan *differencing* data.

5. Pembentukan model VARIMA  
Model VARIMA memiliki orde  $p$ ,  $d$ ,  $q$  yaitu dapat dilihat dari plot PACF, *differencing*, dan plot ACF.
6. Estimasi Parameter  
Estimasi parameter digunakan untuk menentukan model terbaik berdasarkan nilai AIC yang terkecil.
7. Melakukan prediksi pada data harga saham *apple*
8. Menghitung nilai residual dari model terbaik
9. Melakukan Uji Asumsi Residual  
Uji asumsi residual yang digunakan adalah uji Ljung-Box untuk asumsi *white noise* dan uji *Jarque-Beta Test of Normality* untuk asumsi normal multivariat.
10. Melakukan *input* data dengan metode LSTM  
Data yang di *input* pada metode LSTM adalah data hasil prediksi dan residual dari VARIMA.
11. Melakukan *splitting* data  
*Splitting* untuk kedua data dengan skema 70% *training*, 30% *testing* dan 80% *training*, 20% *testing*.
12. Melakukan *Scaling* data  
*Scaling* data pada data prediksi dan residual dari VARIMA menggunakan *standardscaler*.
13. Membangun model *hybrid* LSTM  
Membangun dua model utama LSTM untuk melakukan proses *hybrid*. Penentuan parameter terbaik kedua model LSTM yaitu untuk data prediksi dan residual dari VARIMA, dilakukan menggunakan *hyperparameter tuning*.
14. Melakukan prediksi dan peramalan dengan model pertama LSTM menggunakan data prediksi VARIMA.
15. Melakukan prediksi dan peramalan dengan model kedua LSTM menggunakan data residual VARIMA.
16. Menggabungkan hasil prediksi dan peramalan kedua model LSTM dengan proses operasi penjumlahan.
17. Menghitung nilai akurasi model terbaik



Gambar 2. Diagram Alir Metode Penelitian

## V. KESIMPULAN

Berdasarkan hasil dan pembahasan yang diperoleh, penelitian tentang peramalan harga saham *apple* dengan metode *hybrid* VARIMA-LSTM memberikan kesimpulan sebagai berikut :

1. Model *hybrid* VARIMA-LSTM terdiri dari dua model utama yang kemudian digabungkan dengan metode penjumlahan. Model pertama merupakan model yang dibangun dengan data prediksi dari VARIMA sedangkan model kedua dibangun dengan data residual dari VARIMA.
2. Metode *hybrid* VARIMA-LSTM dengan *splitting* data 80% data *training* dan 20% data *testing* menghasilkan RMSE sebesar 34.0735, dan MAPE sebesar 0.0297, serta akurasi sebesar 99.9703%. Sedangkan evaluasi model yang dihasilkan dengan *splitting* data 70% data *training* dan 30% data *testing* yaitu RMSE sebesar 49.0976, dan MAPE sebesar 0.0361, serta akurasi sebesar 99.9639%. Sehingga metode *hybrid* VARIMA-LSTM dengan *splitting* data 80% data *training* dan 20% data *testing* menghasilkan evaluasi model yang lebih baik dibandingkan dengan *splitting* data 70% data *training* dan 30% data *testing*.
3. Peramalan menggunakan metode *hybrid* VARIMA-LSTM menggunakan *splitting* data 80% data *training* dan 20% data *testing* menghasilkan peramalan yang sudah mengikuti pola data terbaru dari harga saham *apple*. Sedangkan untuk *splitting* data 70% data *training* dan 30% data *testing* menghasilkan pola data yang cukup berbeda dari data terbaru dari harga saham *apple*.

## DAFTAR PUSTAKA

- Ahmad, F. 2020. Penentuan Metode Peramalan pada Produksi Part New Granada BowlSt di Pt X. *JISI: Jurnal Integrasi Sistem Industri*. **7**(1): 31-39.
- Akbar, R., Santoso, R., & Warsito, B. 2022. Prediksi Tingkat Temperatur Kota Semarang Menggunakan Metode Long Short-Term Memory (LSTM). *Jurnal Gaussian*. **11**(4): 572-579.
- Aldi, M. W. P., Jondri, J., & Aditsaia, A. 2018. Analisis dan Implementasi Long Short Term Memory Neural Network untuk Prediksi Harga Bitcoin. *e-Proceeding of Engineering*. **5**(2): 3548-3555.
- Ambarwari, A., Adrian, Q.J., & Herdiyeni. 2020. Analysis of the Effect of Data Scaling on the Performance of the Machine Learning Algoritm for Plant Identification. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*. **4**(1): 117-122.
- Aulia, T., Rohmawati, A.W., & Indwiarti. 2021. Prediksi Nilai Ekstrem Jumlah Kasus Positif Covid-19 di Provinsi Jawa Timur dengan Model Vector Autoregressive Moving Average (VARMA). *e-Proceeding of Engineering*. **8**(5): 11137-11149.
- Caliwag, A. C., & Lim, W. 2019. Hybrid VARMA and LSTM Method for Lithium-ion Battery State-of-Charge and Output Voltage Forecasting in Electric Motorcycle Applications. *IEEE Acces*. **7**(10): 59680-59689.
- Chung, H., & Shin, K.S. 2018. Genetic Algoritm-Optimized Long Short-Term Memory Network for Stock Market Prediction. *Sustainability*. **10**(10).



- Dave, E., Leonardo, A., Jeanice, M., & Hanafiah, N. 2021. Forecasting Indonesia Export using a Hybrid Model ARIMA-LSTM. *Elsevier B.V.* **179**(8): 480-487.
- Dissanayake, B., Hemachandra, O., Lakshita, N., Haputhanthri, D., & Wijayasiri, A. 2021. A Comparison of ARIMAX, VAR and LSTM on Multivariate Short-Term Traffic Volume Forecasting. *Conference of Open Innovations Association.* 564-570.
- Hikmaturokhman, A., Nafi'ah, H., Larasati, S., Wahyudin, A., Ariprawira, G., & Pramono, S. 2020. Deep Learning Algorithm Models for Spam Identification on Cellular Short Message Service. *Journal of Communications.* **17**(9): 769-776.
- Hochreiter, S., & Schmidhuber, J. 1997. Long Short-Term Memory. *Neural Computation.* **9**(8): 1735-1780.
- Ivan, E., & Purnomo, H. D. 2022. Forecasting Prices of Fertilizer Raw Materials Using Long Short Term Memory. *Jenal Teknik Informatika (Jutif).* **3**(6): 1663-1673.
- Jarque, C.M., & Bera, A.K. 1987. A Test for Normality of Observations and Regression Residuals. *International Statistical Institute (ISI).* **55**(2): 163-172
- Kaur, H., Alam, M.A., Mariyam, S., Alankar, B., Chauhan, R., Adnan, A.M., & Kisi, O. 2021. Predicting Water Availability in Water Bodies under the Influence of Precipitation and Water Management Actions Using VAR/VECM/LSTM. *MDPI: Climate.* **9**(9).
- Lusia, D. A., & Ambarwati, A. 2018. Perbandingan Peramalan Univariat dan Multivariat Arima pada Indeks Harga Saham Gabungan. *Jurnal Statistika Universitas Muhamadiyah Semarang.* **6**(2): 78-84.
- Namini, S.S., Tavakoli, N., & Namin, A.S. 2018. A Comparison of ARIMA and LSTM in Forecasting Time Series. *IEEE International Conference on Machine Learning and Applications.* 1394-1401.

- Ningtiyas, S.R., 2018. Aplikasi Metode Double Exponential Smoothing Holt dan ARIMA untuk Meramalkan Voluntary Counseling and Testing (VCT) Odha di Povinsi Jawa Timur. *IJPH*. **3** : 158-168.
- Pasaribu, D. J. M., Kusriani, K., & Sudarman, S. 2020. Peningkatan Akurasi Klasifikasi Sentimen Ulasan Makanan Amazon dengan Bidirectional LSTM dan BERT Embedding. *Jurnal Teknologi Informasi dan Komunikasi*. **10**(1): 9-20.
- Pertiwi, A., Dewi, L.F., Toharudin, T. & Ruchjana, B.N. 2021. Penerapan Model Vector Autoregressive Integrated Moving Average (VARIMA) untuk Prakiraan Indeks Harga Saham Gabungan dan Kurs Rupiah Terhadap USD. *Pattimura Proceeding: Conference of Science and Technology*. 431-442.
- Pradnyandita, I.N.D., Rohmawati, A.A., & Indwiarti. 2022. Electronic Money Transactions Forecasting with Support Vector Regression (SVR) and Vector Autoregressive Moving Average (VARMA). *Intl. Journal on ICT*. **8**(17): 69-85.
- Prasetya, B. D., Pamungkas, F. S., & Khairisudin, I. 2020. Pemodelan dan Peramalan Data Saham dengan Analisis Time Series menggunakan Python. hlm. 714-718. *Prosiding Seminar Nasional Matematika*.
- Pratama, E. G., Purbasari, I. Y., & Saputra, W.S.J. 2022. Prediksi Harga Saham Menggunakan Metode Long Short Term Memory Saat Kondisi Pandemi. *Jurnal Informatika dan Sistem Informasi (JIFoSI)*. **3**(2): 86-95.
- Qiu, J., Wang, B., & Zhou, C. 2020. Forecasting Stock Prices with Long-Short Term Memory Neural Network based on attention Mechanism. *PLOS ONE*. 1-15.
- Rosyidah, H., Rahmawati, R., & Prahutama, A. 2017. Pemodelan Vector Autoregressive X (VARX) untuk Meramalkan Jumlah Uang Beredar di Indonesia. *Jurnal Gaussian*. **6**(3): 333-343.
- Rusyana, A., Tatsara, N., Balqis, R., & Rahmi, S. 2020. Application of Clustering and VARIMA for Rainfall Prediction. *IOP Conference Series: Materials Science and Engineering*. 796.

- Saputra, W., & Prabowo, Y.D. 2020. Pengembangan Aplikasi Klasifikasi Gambar menggunakan Library Tensorflow yang Menerapkan Algoritma Convolutional Neural Network Studi Kasus: Galeri Foto Kegiatan Ibadah Gereja Shoot Fellowship. *KALBISIANA Jurnal Sains, Bisnis dan Teknologi*. **8**(3): 2892-2901.
- Susilawati, R., & Sunendiari, S. 2022. Peramalan Jumlah Penumpang Kereta Api Menggunakan Metode ARIMA dan Grey System Theory. *Jurnal Riset Statistika (JRS)*. **2**(1): 1-12.
- Temur, A.S., Akgun, M., & Temur, G. 2019. Predicting Housing Sales in Turkey Using ARIMA, LSTM, and Hybrid Models. *Journal of Business Economics and Management*. **20**(5): 920-928.
- Tian, C., Ma, J., Zhang, C., & Zhan, P. 2018. A Deep Neural Network Model for Short-Term Load Forecast Based on Long Short-Term Memory Network and Convolutional Neural Network. *Energies*. **11**(12).
- Wei, W. W. S. 2006. *Time Series Analysis Univariate and Multivariate Methods*. Second Edition. Pearson Education, Inc. Boston.
- Yin, C., Zhu, Y., Fei, J., & He, X. 2017. A Deep Learning Approach for Intrusion Detection Using Recurrent Neural Networks. *IEEE Access*. **5**: 21954-21961.
- Yu, Y., Si, X., Hu, C., & Zhang, J. 2019. A Review of Recurrent Neural Networks: LSTM Cells and Network Architectures. *Massachusetts Institute of Technology*. **31**(7): 1236-1270.
- Zhang, G.P. 2003. Time Series Forecasting Using a Hybrid ARIMA and Neural Network Model. *Neurocomputing*. **50**(2003): 159-175.