

**IMPLEMENTASI METODE *HYBRID VECTOR AUTOREGRESSIVE  
INTEGRATED MOVING AVERAGE (VARIMA) – GATED  
RECURRENT UNIT (GRU)* DALAM PERAMALAN  
HARGA SAHAM PT BANK RAKYAT INDONESIA**

**(Skripsi)**

**Oleh**

**SALSABILA ADELIA FITHRI**

**(2017031062)**



**JURUSAN MATEMATIKA  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS LAMPUNG  
BANDAR LAMPUNG  
2024**

## **ABSTRACT**

### **IMPLEMENTATION OF THE HYBRID VECTOR AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (VARIMA) – GATED RECURRENT UNIT (GRU) METHOD IN FORECASTING THE STOCK PRICE OF PT BANK RAKYAT INDONESIA**

**By**

**Salsabila Adelia Fithri**

The capital market plays a crucial role in the national economic growth, including in Indonesia. PT Bank Rakyat Indonesia is one of the companies with the largest market capitalization, primarily due to its success in the UMKM segment. However, the fluctuations in BRI's stock prices demand in-depth analysis to assist potential investors in making informed investment decisions. This research aims to forecast the stock prices of PT Bank Rakyat Indonesia using the hybrid VARIMA-GRU method. This method is chosen because it combines VARIMA's ability to capture linear patterns with GRU's advantages in capturing nonlinear patterns. The results of the 30-day forward forecasting indicate that the hybrid VARIMA-GRU model performs exceptionally well in the 90% data training and 10% data testing scheme, with an RMSE value of 91.75 and MAPE of 0.0142%, compared to the 80% data training and 20% data testing scheme.

**Keywords** : Capital Market, Stocks, Prediction, Forecasting, VARIMA, GRU  
hybrid model.

## ABSTRAK

### **IMPLEMENTASI METODE *HYBRID VECTOR AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (VARIMA) – GATED RECURRENT UNIT (GRU)* DALAM PERAMALAN HARGA SAHAM PT BANK RAKYAT INDONESIA**

Oleh

**Salsabila Adelia Fithri**

Pasar modal berperan penting dalam pertumbuhan ekonomi nasional, termasuk di Indonesia. PT Bank Rakyat Indonesia adalah salah satu perusahaan dengan kapitalisasi pasar terbesar, terutama karena keberhasilannya dalam segmen UMKM. Meskipun demikian, fluktuasi harga saham BRI menuntut adanya analisis mendalam untuk membantu calon investor dalam membuat keputusan investasi yang tepat. Penelitian ini bertujuan meramalkan harga saham PT Bank Rakyat Indonesia dengan menggunakan metode *hybrid* VARIMA-GRU. Metode ini dipilih karena menggabungkan kemampuan VARIMA dalam menangkap pola linier dengan keunggulan GRU dalam menangkap pola nonlinier. Hasil peramalan 30 hari ke depan menunjukkan bahwa model *hybrid* VARIMA-GRU memiliki kinerja yang sangat baik pada skema 90% data *training* dan 10% data *testing* dengan nilai RMSE sebesar 91,75 dan MAPE sebesar 0,0142%, dibandingkan skema 80% data *training* dan 20% data *testing*.

**Kata Kunci** : Pasar Modal, Saham, Prediksi, Peramalan, VARIMA, GRU, *hybrid model*.

**IMPLEMENTASI METODE *HYBRID VECTOR AUTOREGRESSIVE  
INTEGRATED MOVING AVERAGE (VARIMA) – GATED  
RECURRENT UNIT (GRU)* DALAM PERAMALAN  
HARGA SAHAM PT BANK RAKYAT INDONESIA**

Oleh  
**Salsabila Adelia Fithri**

**Skripsi**

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

**Pada**

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



**JURUSAN MATEMATIKA  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS LAMPUNG  
BANDAR LAMPUNG  
2024**

Judul Skripsi

: **IMPLEMENTASI METODE *HYBRID VECTOR AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (VARIMA) – GATED RECURRENT UNIT (GRU)* DALAM PERAMALAN HARGA SAHAM PT BANK RAKYAT INDONESIA**

Nama Mahasiswa

: **Salsabila Adelia Fithri**

Nomor Pokok Mahasiswa

: **2017031062**

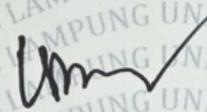
Jurusan

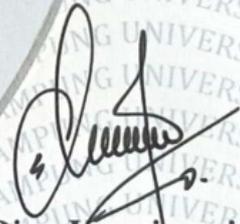
: **Matematika**

Fakultas

: **Matematika dan Ilmu Pengetahuan Alam**



  
**Ir. Warsono, M.S., Ph.D.**  
NIP. 196302161987031003

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

**2. Ketua Jurusan Matematika**

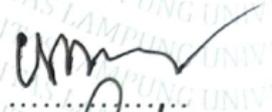
  
**Dr. Aang Nuryaman, S.Si., M.Si.**  
NIP. 197403162005011001

**MENGESAHKAN**

1. **Tim Penguji**

**Ketua**

**: Ir. Warsono, M.S., Ph.D.**



**Sekretaris**

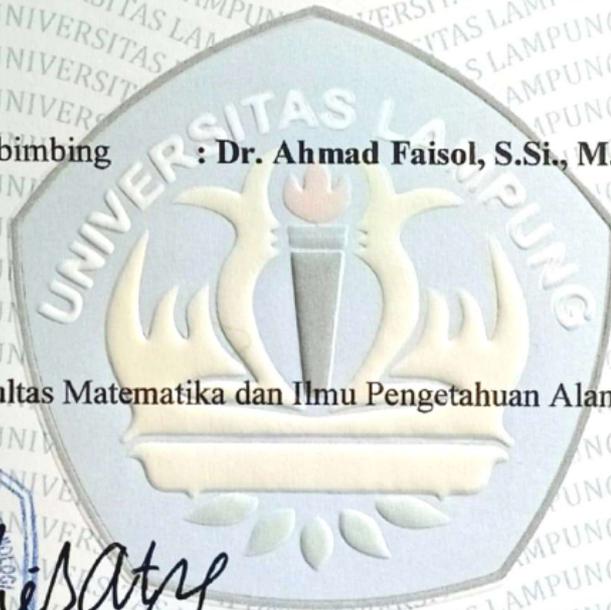
**: Dr. Dian Kurniasari, S.Si., M.Sc.**



**Penguji**

**Bukan Pembimbing**

**: Dr. Ahmad Faisol, S.Si., M.Sc.**



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



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

**NIP. 197110012005011002**



**Tanggal Lulus Ujian Skripsi : 09 Juli 2024**

## PERNYATAAN

Yang bertandatangan di bawah ini:

Nama : **Salsabila Adelia Fithri**  
Nomor Pokok : **2017031062**  
Jurusan : **Matematika**  
Judul Skripsi : **Implementasi Metode *Hybrid Vector Autoregressive Integrated Moving Average (VARIMA) – Gated Recurrent Unit (GRU)* dalam Peramalan Harga Saham PT Bank Rakyat Indonesia**

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, 17 Juli 2024

Yang menyatakan,



Salsabila Adelia Fithri

NPM. 2017031062

## **RIWAYAT HIDUP**

Penulis bernama lengkap Salsabila Adelia Fithri yang lahir di Kedondong pada 9 Desember 2002 dan merupakan anak kedua dari Bapak Adam Malik S.Ag. dan Eliza Rina S. Ag. Penulis memiliki 2 saudara kandung, yakni Kakak yang bernama Adly Ihza Mahesta dan Adik yang bernama Daffa Muhammad Alfarisi.

Pendidikan penulis dimulai dari jenjang taman kanak-kanak di RA Qurrota A'yun Kedondong pada tahun 2007 sampai 2008, dilanjutkan dengan Pendidikan sekolah dasar di MI Negeri 1 Kedondong pada 2008 sampai 2014, lalu Pendidikan Menengah Pertama di MTs Negeri 1 Pesawaran pada 2014 sampai 2017, hingga Pendidikan menengah atas di MA Negeri 1 Bandar Lampung pada tahun 2017 sampai 2020.

Pada tahun 2020, penulis terdaftar sebagai mahasiswa program studi S1 Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lampung melalui jalur Seleksi Bersama Masuk Perguruan Tinggi Negeri (SBMPTN). Selama menjadi mahasiswa, penulis aktif di beberapa organisasi, yaitu Himpunan Mahasiswa Jurusan Matematika (HIMATIKA) sebagai Anggota Bidang Eksternal tahun 2021 sampai 2022, Badan Eksekutif Mahasiswa Fakultas MIPA (BEM FMIPA) sebagai Staf Ahli Dinas Media dan Informasi pada 2021 sampai 2022, hingga menjadi Sekretaris Dinas Media dan Informasi BEM FMIPA pada tahun 2023. Penulis juga aktif di berbagai kepanitiaan, seperti Dies Natalis Jurusan Matematika (DINAMIKA) dan Karya Wisata Ilmiah (KWI) XXXII sebagai Sekretaris Divisi Publikasi, Desain, dan Dokumentasi.

Pada awal tahun 2023, penulis melaksanakan Kerja Praktik (KP) di PT PLN Unit Induk Distribusi Lampung. Selain itu, penulis juga mengikuti salah satu program Kampus Merdeka selama 6 bulan, yakni Studi Independen *Data Analyst* di mitra My Edu Solve, PT Dwi Inti Putra pada semester 7. Penulis juga melaksanakan magang secara mandiri di Bakrie Center Foundation pada bidang Media dan Komunikasi pada semester tersebut. Selain itu, penulis juga melaksanakan Kuliah Kerja Nyata sebagai bentuk pengabdian masyarakat di Desa Sidoluhur, Kecamatan Bangun Rejo, Kabupaten Lampung Tengah selama lebih dari 1 bulan.

Selama menjadi mahasiswa, penulis juga aktif pada beberapa kompetisi dan meraih pencapaian yang membanggakan seperti, menjadi *Top 10* finalis pada *Statistics Infographic Competition Satria Data 2022*, Juara 2 Lomba Poster Infografis Hari Gizi Nasional oleh Himpunan Mahasiswa Gizi Poltekkes Kalimantan Timur 2023, Juara Harapan 3 Kompetisi Infografis Statistika ANAVA #18 oleh Himpunan Mahasiswa Statistika Universitas Gadjah Mada 2024, *Top 5* finalis pada *Visualistics Competition Gammafest* oleh Himpunan Mahasiswa Profesi Gamma Sigma Beta IPB University 2024, Juara 3 Tempelاند *Infographic Challenge* oleh Tempelاند Indonesia 2024, dan Juara 2 Lomba Infografis Nasional Saintek Expo oleh UKM Saintek Universitas Lampung 2024.

## **KATA INSPIRASI**

“Barang siapa yang bersabar, maka Allah akan memberikan jalan keluar.”

**(QS. Al-Baqarah: 155)**

“Dan barang -siapa yang bertakwa kepada Allah, niscaya Allah menjadikan baginya kemudahan dalam urusannya.”

**(QS. At-Talaq: 4)**

“Karena sesungguhnya sesudah kesulitan itu ada kemudahan.”

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

“Ketahuilah bahwa kemenangan bersama kesabaran, kelapangan bersama kesempitan, dan kesulitan bersama kemudahan.”

**(HR Tirmidzi)**

“Jangan bersedih. Sesungguhnya pertolongan akan datang bersama kesabaran.”

**(HR. Ahmad)**

## **PERSEMBAHAN**

Rasa syukur kepada Allah SWT, karena atas limpahan taufik dan hidayah-Nya, skripsi ini dapat diselesaikan dengan baik. Dengan penuh kebahagiaan dan ketulusan hati, karya ini saya dedikasikan kepada:

### **Abi, Ibu, dan Keluarga**

Yang telah bekerja keras tanpa kenal lelah untuk memastikan penulis bisa menempuh pendidikan dan meraih gelar sarjana, yang selalu mendoakan, memberikan dukungan, nasihat, dan cinta yang tidak bisa tergantikan oleh apapun.

### **Dosen Pembimbing dan Penguji**

Yang dengan penuh kesabaran meluangkan waktu untuk memberikan bimbingan, saran, dan motivasi.

### **Sahabat-Sahabatku**

Yang terus memberikan bantuan, dukungan, doa, motivasi, serta canda tawa dalam segala situasi dan langkah perjalanan ini.

**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 “Implementasi Metode *Hybrid Vector Autoregressive Integrated Moving Average* (VARIMA) – *Gated Recurrent Unit* (GRU) dalam Peramalan Harga Saham PT Bank Rakyat Indonesia”.

Skripsi ini dibuat sebagai salah satu syarat memperoleh gelar Sarjana Matematika di Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung. Proses penyusunan skripsi ini dibimbing, didukung, dan dibantu oleh banyak pihak. Oleh karena itu, penulis ingin mengucapkan terima kasih kepada:

1. Bapak Ir. Warsono, M.S., Ph.D., selaku Pembimbing I yang selalu memberikan waktu, arahan, bimbingan, saran, serta dukungan kepada penulis sehingga dapat menyelesaikan skripsi ini.
2. Ibu Dr. Dian Kurniasari, S.Si., M.Sc., selaku Pembimbing II yang juga selalu memberikan waktu, arahan, bimbingan, saran, serta dukungan kepada penulis sehingga dapat menyelesaikan skripsi ini.
3. Bapak Dr. Ahmad Faisol, S.Si., M.Sc., selaku Penguji yang telah bersedia memberikan kritik, saran, dan evaluasi kepada penulis.
4. Ibu Dra. Dorrah Aziz, M.Si., selaku Pembimbing Akademik yang selalu bersedia membimbing, dan mendukung penulis pada hal yang berkaitan dengan akademik selama masa perkuliahan.
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, staf, dan karyawan Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
8. Abi, Ibu, dan keluarga yang telah memberikan doa dan dukungan tiada hentinya kepada penulis.
9. Diriku sendiri, yang telah dengan gigih dan penuh ketabahan menjalani setiap langkah dalam penyusunan skripsi ini.
10. Mawar, Callista, Fegy, Yulian, Adel, dan teman-teman seperbimbingan yang selalu menyamangati, memberikan bantuan, dukungan, dan motivasi kepada penulis.
11. Happy, Meryam, Cindy, Gustina, Maya, dan Nada, teman-teman seperjuangan Matematika Angkatan 2020 yang telah kebersamai perjalanan penulis selama perkuliahan.
12. 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 para pembaca.

Bandar Lampung, 17 Juli 2024  
Penulis,

Salsabila Adelia Fithri

## DAFTAR ISI

halaman

### DAFTAR TABEL

### DAFTAR GAMBAR

<b>I. PENDAHULUAN .....</b>	<b>1</b>
1.1. Latar Belakang dan Masalah .....	1
1.2. Rumusan Masalah .....	5
1.3. Tujuan Penelitian.....	5
1.4. Manfaat Penelitian.....	5
<b>II. TINJAUAN PUSTAKA.....</b>	<b>6</b>
2.1. Penelitian Terkait .....	6
2.2. Saham .....	10
2.3. Peramalan .....	10
2.4. Data Deret waktu.....	12
2.5. Analisis Deret Waktu Multivariat .....	12
2.5.1. <i>Vector Autoregressive</i> (VAR).....	13
2.5.2. <i>Vector Moving Average</i> (VMA) .....	13
2.5.3. <i>Vector Autoregressive Moving Average</i> (VARMA).....	14
2.5.4. <i>Vector Autoregressive Integrated Moving Average</i> (VARIMA).....	14
2.6. Stasioneritas Data .....	15
2.7. Uji Asumsi Residual.....	17
2.8. <i>Lag Optimum</i> .....	18
2.9. Normalisasi dan Denormalisasi Data .....	18
2.10. <i>Machine learning</i> .....	20
2.10. Fungsi Aktivasi .....	21
2.11. <i>Deep learning</i> .....	23
2.12. <i>Gated Recurrent Unit</i> (GRU).....	25
2.13. <i>Hyperparameter</i> .....	29
2.14. Model <i>Hybrid VARIMA-GRU</i> .....	30
2.15. Evaluasi Model.....	31

2.16. Uji <i>T-Hotelling</i> .....	33
<b>III. METODOLOGI PENELITIAN .....</b>	<b>34</b>
3.1. Tempat dan Waktu Penelitian .....	34
3.1.1. Tempat Penelitian .....	34
3.1.2. Waktu Penelitian.....	34
3.2. Data dan Alat Penelitian .....	35
3.2.1. Data penelitian .....	35
3.2.2. Alat Penelitian .....	35
3.3. Metode Penelitian .....	36
<b>IV. HASIL DAN PEMBAHASAN .....</b>	<b>41</b>
4.1. Proses <i>Input Data</i> .....	41
4.2. Visualisasi Data .....	41
4.3. <i>Preprocessing</i> Data .....	42
4.4. Prediksi dengan Model VARIMA.....	43
4.4.1. Uji Stasioneritas Data .....	43
4.4.2. Pembentukan Model VARIMA .....	45
4.4.3. <i>Splitting</i> Data .....	47
4.4.4. Estimasi Parameter .....	48
4.4.5. Prediksi Model VARIMA (2,1,1).....	49
4.4.6. Peramalan Model VARIMA.....	53
4.4.7. Residual Model VARIMA.....	56
4.4.8. Uji Asumsi Residual Model VARIMA (2,1,1).....	60
4.5. Prediksi Model <i>Hybrid</i> VARIMA-GRU .....	60
4.5.1. <i>Scaling Data</i> .....	61
4.5.2. <i>Hyperparameter Tuning</i> Model GRU Pertama .....	61
4.5.3. Membangun Model GRU Pertama .....	62
4.5.4. <i>Hyperparameter Tuning</i> Model GRU Kedua .....	62
4.5.5. Membangun Model GRU Kedua.....	63
4.5.6. Prediksi Model GRU Pertama .....	63
4.5.7. Prediksi Model GRU Kedua.....	67
4.5.8. Prediksi Model <i>Hybrid</i> VARIMA-GRU.....	71
4.7. Peramalan Model <i>Hybrid</i> VARIMA-GRU .....	73
4.7.1. Hasil Peramalan Model GRU Pertama .....	74
4.7.2. Hasil Peramalan Model GRU Kedua.....	75
4.7.3. Hasil Peramalan Model <i>Hybrid</i> VARIMA-GRU .....	76
4.8. Perbandingan Data Terbaru dan Peramalan .....	78
<b>V. KESIMPULAN DAN SARAN .....</b>	<b>79</b>

5.1. Kesimpulan.....	79
5.2. Saran.....	80
<b>DAFTAR PUSTAKA .....</b>	<b>81</b>

## DAFTAR TABEL

	halaman
1. Penelitian Terkait .....	6
2. Data Harga Saham PT Bank Rakyat Indonesia.....	35
3. <i>Missing Value</i> .....	42
4 Uji Stasioner Awal .....	44
5. Uji Stasioner Setelah <i>Differencing</i> Pertama.....	44
6. <i>Splitting Data</i> .....	47
7. Model-Model VARIMA .....	48
8. Evaluasi Prediksi Model VARIMA (2,1,1).....	52
9. Uji Ljung-Box .....	60
10. <i>Hyperparameter Tuning</i> Model Pertama .....	61
11. Hasil <i>Hyperparameter Tuning</i> Model GRU Pertama .....	62
12. <i>Hyperparameter Tuning</i> Model Kedua.....	62
13. Hasil <i>Hyperparameter Tuning</i> Model GRU Kedua.....	63
14. Evaluasi Prediksi Model GRU Pertama.....	67
15. Evaluasi Prediksi Model GRU Kedua.....	71
16. Evaluasi Prediksi Model <i>Hybrid</i> VARIMA-GRU .....	73
16. Uji T- <i>Hotelling</i> .....	738

## DAFTAR GAMBAR

	halaman
1. Fungsi <i>Sigmoid</i> .....	22
2. Fungsi <i>Tanh</i> .....	23
3. Diagram Arsitektur Model GRU.....	26
4. Gerbang <i>Reset</i> dan Gerbang <i>Update</i> .....	27
5. <i>Flowchart</i> Metode <i>Hybrid</i> VARIMA-GRU. ....	37
6. Plot Data Harga Saham PT Bank Rakyat Indonesia .....	42
7. Plot Data Stasioner.....	45
8. Plot ACF dan PACF Variabel <i>Open</i> . ....	45
9. Plot ACF dan PACF Variabel <i>High</i> .....	46
10. Plot ACF dan PACF Variabel <i>Low</i> .....	46
11. Plot ACF dan PACF Variabel <i>Close</i> .....	47
12. Plot Prediksi Model VARIMA (2,1,1) dengan Skema 80% <i>Data Training</i> dan 20% <i>Data Testing</i> . ....	50
13. Plot Prediksi Model VARIMA (2,1,1) dengan Skema 90% <i>Data Training</i> dan 10% <i>Data Testing</i> .....	52
14. Plot Peramalan VARIMA (2,1,1) dengan Skema 80% <i>Data Training</i> dan 20% <i>Data Testing</i> . ....	54
15. Plot Peramalan VARIMA (2,1,1) dengan Skema 90% <i>Data Training</i> dan 10% <i>Data Testing</i> . ....	56
16. Plot Residual Model VARIMA (2,1,1) dengan Skema 80% <i>Data Training</i> dan 20% <i>Data Testing</i> .....	58
17. Plot Residual Model VARIMA dengan Skema 90% <i>Data Training</i> dan 10% <i>Data Testing</i> . ....	59
18. Plot Prediksi Model GRU Pertama dengan Skema 80% <i>Data Training</i> dan 20% <i>Data Testing</i> . ....	65
19. Plot Prediksi Model GRU Pertama dengan Skema 90% <i>Data Training</i> dan 10% <i>Data Testing</i> .....	66

20. Plot Prediksi Model GRU Kedua dengan Skema 80% Data <i>Training</i> dan 20% Data <i>Testing</i> . .....	69
21. Plot Prediksi Model GRU Kedua dengan Skema 90% Data <i>Training</i> dan 10% Data <i>Testing</i> . .....	70
22. Plot Prediksi <i>Hybrid</i> VARIMA-GRU dengan Skema 80% Data <i>Training</i> dan 20% Data <i>Testing</i> . .....	72
23. Plot Prediksi Model <i>Hybrid</i> VARIMA-GRU dengan Skema 90% Data <i>Training</i> dan 10% Data <i>Testing</i> . .....	72
24. Plot Hasil Peramalan Model GRU Pertama dengan Skema 80% Data <i>Training</i> dan 20% Data <i>Testing</i> . .....	74
25. Plot Hasil Peramalan Model GRU Pertama dengan Skema 90% Data <i>Training</i> dan 10% Data <i>Testing</i> . .....	75
26. Plot Hasil Peramalan Model GRU Kedua dengan Skema 80% Data <i>Training</i> dan 20% Data <i>Testing</i> . .....	75
27. Plot Hasil Peramalan Model GRU Kedua dengan Skema 90% Data <i>Training</i> dan 10% Data <i>Testing</i> . .....	76
28. Plot Hasil Peramalan Model <i>Hybrid</i> VARIMA-GRU dengan Skema 80% Data <i>Training</i> dan 20% Data <i>Testing</i> . .....	77
29. Plot Hasil Peramalan Model <i>Hybrid</i> VARIMA-GRU dengan Skema 90% Data <i>Training</i> dan 10% Data <i>Testing</i> . .....	77

## **I. PENDAHULUAN**

### **1.1. Latar Belakang dan Masalah**

Perkembangan pasar modal memiliki keterkaitan yang sangat erat dengan kemajuan ekonomi suatu negara karena perannya dalam memajukan pendapatan nasional, sehingga memberikan dampak positif pada pertumbuhan pendapatan nasional (Kant, 2021). Menurut Dantes (2019), pasar modal merupakan rangkaian seluruh pasar yang diorganisir dan lembaga-lembaga yang terlibat dalam perdagangan berbagai segmen keuangan, seperti saham, obligasi, pinjaman berjangka, hipotek, tabungan, dan deposito berjangka. Salah satu segmen pasar modal yang paling diminati oleh investor adalah saham, khususnya sektor keuangan, yakni pada perusahaan perbankan. Berdasarkan informasi pada situs resmi Bursa Efek Indonesia, terdapat daftar 50 perusahaan dengan kapitalisasi pasar terbesar setiap bulannya. Perusahaan-perusahaan di sektor perbankan mendominasi posisi teratas dalam daftar tersebut, tak terkecuali PT Bank Rakyat Indonesia. Selama 11 tahun terakhir, PT Bank Rakyat Indonesia secara konsisten tercatat sebagai salah satu perusahaan dengan kapitalisasi pasar terbesar. Pada bulan Oktober 2023, perusahaan tersebut berhasil menduduki peringkat kedua.

Perusahaan PT Bank Rakyat Indonesia dianggap sebagai opsi investasi yang tepat mengingat faktor unik yang membedakannya dari bank lain di pasar, yakni memiliki keunggulan signifikan dalam menangani segmen pasar Usaha Mikro, Kecil, dan Menengah (UMKM). Berdasarkan laporan tahunan terakhir, proporsi kredit UMKM dari PT Bank Rakyat Indonesia terus meningkat, mencapai 83,86% dari total penyaluran kredit yang disalurkan oleh bank kepada segmen UMKM.

Perusahaan PT Bank Rakyat Indonesia menunjukkan pertumbuhan laba yang konsisten pada sektor ini. Sektor UMKM terus menjadi tulang punggung ekonomi Indonesia, sehingga keterlibatan PT Bank Rakyat Indonesia dalam mendukung sektor ini menjadi salah satu poin penentu yang menarik. Namun, berdasarkan Laporan Profil Industri Perbankan untuk triwulan III tahun 2023 yang diakses dari situs Otoritas Jasa Keuangan, kinerja total aset di bank umum Indonesia menunjukkan perlambatan dibandingkan tahun sebelumnya. Fluktuasi nilai saham yang terus-menerus menuntut calon investor untuk melakukan pemeriksaan dan analisis menyeluruh terhadap data historis perusahaan sebelum membuat keputusan investasi. Analisis ini sangat membantu investor dalam memahami potensi pergerakan harga saham perusahaan di masa depan. Oleh karena itu, peramalan terhadap perubahan harga saham untuk beberapa periode mendatang menjadi langkah krusial dalam menentukan keputusan investasi di pasar saham.

Data historis harga saham PT Bank Rakyat Indonesia merupakan data deret waktu karena mencakup informasi tentang harga saham yang terjadi pada interval waktu tertentu. Salah satu metode deret waktu yang dapat diterapkan dalam peramalan harga saham adalah metode *Autoregressive Integrated Moving Average* (ARIMA). Namun, metode ARIMA kurang tepat diterapkan untuk lebih dari satu pergerakan harga saham karena metode ini bersifat univariat yang hanya memodelkan satu variabel. Oleh karena itu, untuk memprediksi lebih dari satu pergerakan harga saham, dapat digunakan metode perluasan dari ARIMA, yakni metode *Vector Autoregressive Integrated Moving Average* (VARIMA). Dibandingkan dengan metode multivariat lainnya, seperti VAR hingga VARMA, VARIMA dapat diterapkan pada data yang tidak stasioner, yang kemudian dapat dilakukan *differencing* agar data menjadi stasioner.

Jusmawati dkk. (2020) melakukan penelitian untuk meramalkan laju inflasi dan suku bunga Indonesia pada periode Januari hingga Desember 2016 menggunakan metode VARIMA, dan hasil penelitian menunjukkan bahwa peramalan yang diperoleh memiliki nilai akurasi yang cukup baik. Nilai *Mean Absolute Percentage Error* (MAPE) yang diperoleh adalah sebesar 6,04% untuk variabel laju inflasi dan

1,84% untuk variabel suku bunga. Meskipun metode multivariat menunjukkan tingkat akurasi yang tinggi, penting untuk diingat bahwa tidak semua data deret waktu finansial bersifat linier (Tsay, 2005). Sejumlah besar data deret waktu seperti data saham, memiliki kecenderungan untuk bersifat nonlinier. Oleh karena itu, penggunaan metode klasik seperti metode VARIMA mungkin kurang sesuai karena metode tersebut didasarkan pada asumsi keberlakuan linier. Untuk mengatasi keterbatasan tersebut, dibutuhkan pendekatan yang lebih baru dalam menangani data yang memiliki sifat nonlinier.

f

*Machine learning* merupakan salah satu metode yang juga dapat diterapkan untuk meramalkan data deret waktu. Salah satu pendekatan *machine learning* yang belakangan ini populer dan berkembang dengan baik adalah metode *deep learning*. *Deep learning* merupakan metode pendekatan *machine learning* yang mengandalkan pengetahuan tentang otak manusia, statistik, dan matematika terapan (Goodfellow dkk., 2016). Penggunaan *deep learning* menghasilkan akurasi yang sangat baik, seperti penelitian oleh Halim dkk., (2022) yang melakukan prediksi pada zat pencemar udara di Daerah Khusus Ibukota (DKI) Jakarta menggunakan metode *Gated Recurrent Unit* (GRU). Penelitian tersebut memperoleh nilai akurasi terbaik pada zat SO<sub>2</sub>, yaitu MAPE sebesar 7,50% dan *Root Mean Square Error* (RMSE) sebesar 3,42.

Metode GRU merupakan salah satu metode *deep learning* sebagai varian dari *Recurrent Neural Network* (RNN), di mana GRU lebih baik dalam mengatasi masalah *vanishing gradient* menggunakan mekanisme gerbang untuk memilih informasi yang penting dan mengabaikan informasi yang tidak relevan dalam sel rekuren (Noh, 2021). Penggunaan metode *deep learning* GRU lebih unggul dibandingkan metode statistika klasik, seperti penelitian oleh Bousnguar dkk. (2023) yang melakukan peramalan pendidikan tinggi Universitas Alabama menggunakan metode GRU. Penelitian tersebut menyatakan bahwa model peramalan deret waktu statistika klasik memiliki keterbatasan dalam menangani hubungan yang kompleks dan tren nonlinier dalam data, sehingga dilakukan penelitian menggunakan metode GRU yang dapat menangani tren nonlinier dalam

data. Hasilnya menunjukkan bahwa GRU mengungguli model statistika klasik dan dapat membantu universitas merencanakan sumber daya mereka dengan lebih baik. Namun, pada kenyataannya pendekatan *deep learning* cenderung hanya membangun hubungan nonlinier.

Menurut Zhang (2003), salah satu pilihan dalam metode statistik untuk peramalan adalah dengan mengintegrasikan model linier dan nonlinier, atau yang dikenal sebagai model *hybrid*. Data deret waktu jarang menunjukkan karakteristik linier atau nonlinier saja, seringkali terjadi kombinasi antara keduanya. Dengan demikian, kombinasi model linier dan nonlinier dapat digunakan untuk menangani kasus deret waktu yang melibatkan elemen linier maupun nonlinier. Peneliti terdahulu kerap melakukan peramalan menggunakan metode *hybrid* antara statistika klasik dan *deep learning*. Salah satu penerapan yang digunakan adalah peramalan yang menggabungkan antara metode multivariat *Vector Autoregressive* (VAR) dan *Long Short Term Memory* (LSTM), serta VAR dan GRU yang dilakukan oleh Aji dan Surjandari (2020). Peramalan tersebut menghasilkan akurasi yang cukup baik, yakni memperoleh MAPE sebesar 1,02% dan RMSE sebesar 81,52 untuk metode VAR-LSTM. Sedangkan, untuk metode VAR-GRU memperoleh MAPE sebesar 0,60% dan RMSE sebesar 62,53. Selain itu, Munkhdalai dkk. (2020) menggunakan metode *hybrid* VAR-GRU dalam meramalkan data *financial market* yang menghasilkan nilai kesalahan yang sangat rendah, yaitu RMSE sebesar 17,24 dan MAPE sebesar 0,603%, yang mengindikasikan bahwa metode sudah cukup baik diterapkan pada data. Berdasarkan penjelasan di atas, untuk membantu para investor dalam membuat strategi dan memutuskan investasi yang tepat agar tidak terjadi kerugian, penelitian ini bermaksud melakukan implementasi metode *hybrid* VARIMA-GRU dalam meramalkan harga saham PT Bank Rakyat Indonesia.

## 1.2. Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan di atas, maka rumusan masalah dalam penelitian ini adalah:

1. Bagaimana cara mengetahui metode *hybrid* VARIMA-GRU sebagai metode dengan performa yang baik digunakan dalam peramalan harga saham di PT Bank Rakyat Indonesia?
2. Bagaimana hasil peramalan harga saham PT Bank Rakyat Indonesia menggunakan metode *hybrid* VARIMA-GRU?

## 1.3. Tujuan Penelitian

Tujuan adanya penelitian ini adalah sebagai berikut:

1. Mengevaluasi performa model dari metode *hybrid* VARIMA-GRU terhadap data harga saham PT Bank Rakyat Indonesia.
2. Mengetahui peramalan harga saham PT Bank Rakyat Indonesia menggunakan metode *hybrid* VARIMA-GRU.

## 1.4. Manfaat Penelitian

Manfaat yang dapat diperoleh dengan adanya penelitian ini adalah sebagai berikut:

1. Menjadi sumber ilmu dan referensi terkait metode *hybrid* VARIMA-GRU pada peramalan harga saham PT Bank Rakyat Indonesia.
2. Menjadi acuan bagi para investor dalam membuat keputusan investasi yang lebih cerdas.

## II. TINJAUAN PUSTAKA

### 2.1. Penelitian Terkait

Penelitian ini memanfaatkan literatur terkait sebagai referensi dalam analisis. Acuan yang dimanfaatkan yakni metode *Vector Autoregressive Integrated Moving Average* (VARIMA), metode *Gated Recurrent Unit* (GRU), metode *hybrid Vector Autoregressive-Long Short Term Memory* (VAR-LSTM), dan metode *hybrid Vector Autoregressive-Gated Recurrent Unit* (VAR-GRU). Kinerja metode-metode tersebut dinilai menggunakan dua metrik evaluasi utama, yaitu *Root Mean Square Error* (RMSE) dan *Mean Absolute Percentage Error* (MAPE). Rangkuman beberapa penelitian terdahulu yang terkait dengan penelitian ini disajikan dalam Tabel 1 sebagai berikut:

Tabel 1. Penelitian Terkait

No	Penelitian	Data	Metode	Hasil	
				RMSE	MAPE
1	Penerapan Model <i>Vector Autoregressive Integrated Moving Average</i> dalam Peramalan Laju Inflasi dan Suku Bunga di Indonesia (Jusmawati dkk., 2020)	<b>Data bulanan laju inflasi dan suku bunga Indonesia</b> Januari - Desember 2016 Sumber: Situs resmi Bank Sentral Indonesia	VARIMA	-	6,04% dan 1,84%

2	Penerapan <i>Gated Recurrent Unit</i> untuk Prediksi Zat Pencemar Udara (Halim dkk., 2022)	<b>Data pencemaran udara DKI Jakarta 2012-2021</b> Sumber: <i>data.jakarta.go.id</i>	GRU	3,42	7,50%
3	<i>Gated Recurrent units (GRU) for Time Series Forecasting in Higher Education</i> (Bousnguar dkk., 2023)	<b>Data pendaftaran di Universitas Alabama tahun 1831-2019</b> Sumber : Universitas Alabama	GRU	970,99	-
4	<i>Hybrid Vector Autoregression–Recurrent Neural Networks to Forecast Multivariate Time Series Jet Fuel Transaction Price</i> (Aji dan Surjandari, 2020)	<b>Data harga transaksi bahan bakar pesawat di Bandara CGK Tangerang Januari 2017-April 2020</b> Sumber: Data terbuka, <i>website platts, website pertamina's, website Bank Indonesia, maskapai penerbangan Indonesia</i>	VAR-LSTM	81,52	1,02%
			VAR-GRU	62,53	0,60%
5	<i>VAR-GRU: A Hybrid Model for Multivariate Financial Time Series Prediction</i> (Munkhdalai dkk., 2020)	<b>Data financial market</b> Jumlah: 3,608 observasi Sumber: <i>The Federal Reserve Bank of St. Louis dan Yahoo Finance</i>	VAR-GRU	17,24	0,603%

Ringkasan dari penelitian-penelitian yang tertera pada Tabel 1, yaitu sebagai berikut:

1. Penelitian Pertama (Jusmawati dkk., 2020)

Jusmawati dkk. (2020) melakukan penelitian dengan menerapkan metode VARIMA untuk menganalisis peramalan laju inflasi dan suku bunga Indonesia. Penelitian ini bertujuan untuk merancang model dari data laju inflasi

dan suku bunga dari Januari 2009 hingga Desember 2016, serta meramalkan kedua data deret waktu tersebut dengan menggunakan metode VARIMA. Model yang diperoleh adalah VARIMA (0,2,2) yang dipilih berdasarkan nilai terendah *Akaike Information Criterion* (AIC), yaitu -4,2891. Model tersebut dievaluasi menggunakan MAPE yang menunjukkan hasil peramalan yang cukup akurat, yakni sebesar 6,04% untuk peramalan laju inflasi dan 1,84% untuk peramalan suku bunga.

2. Penelitian Kedua (Halim dkk., 2022)

Halim dkk., 2022 melakukan penelitian dengan menerapkan metode GRU untuk memprediksi zat pencemar udara di Jakarta. Data penelitian yang digunakan diperoleh dari Dinas Lingkungan Hidup Provinsi DKI Jakarta yang mencakup lima variabel, yaitu *Partikulat* (PM10), *Sulfur Dioksida* (SO<sub>2</sub>), *Karbon Monoksida* (CO), *Ozon permukaan* (O<sub>3</sub>), dan *Oksida Nitrogen* (NO<sub>2</sub>). Hasil pengujian menggunakan metode GRU menunjukkan nilai prediksi yang bervariasi. Penilaian MAPE menunjukkan prediksi yang baik untuk pencemar udara jenis PM10, SO<sub>2</sub>, CO, dan O<sub>3</sub> dengan nilai MAPE rata-rata di bawah 50%. Namun, hasil yang cukup tidak memuaskan diperoleh untuk pencemar udara jenis NO<sub>2</sub>, dengan nilai MAPE lebih dari 50%. Sebaliknya, pada evaluasi RMSE, semua jenis pencemar udara menghasilkan nilai rata-rata di bawah 20%. Zat SO<sub>2</sub> menunjukkan nilai MAPE dan RMSE terendah, yakni sebesar 7,50% dan 3,42. Oleh karena itu, dapat dikatakan bahwa metode GRU memperoleh hasil yang cukup akurat untuk memprediksi pencemar udara di wilayah DKI Jakarta.

3. Penelitian Ketiga (Bousnguar dkk., 2023)

Bousnguar dkk., 2023 melakukan peramalan menggunakan model *Recurrent Neural Network* pada jumlah pendaftaran mahasiswa di universitas. Data yang digunakan diekstraksi dari pendaftaran di Universitas Alabama antara tahun 1831 dan 2019. Hasil studi kasus menunjukkan bahwa *Recurrent Neural Network* unggul dibandingkan model statistika klasik dalam hal akurasi prediksi dengan model GRU memiliki nilai RMSE sebesar 970,99.

4. Penelitian Keempat (Aji dan Surjandari, 2020)

Aji dan Surjandari (2020) melakukan peramalan pada data harga transaksi bahan bakar pesawat di Bandara CGK (Bandara Internasional Soekarno-Hatta) Tangerang periode Januari 2017 hingga April 2020. Data penelitian ini mencakup 8 variabel dengan memperhitungkan harga minyak mentah global serta harga bahan bakar pesawat global dan lokal di setiap bandara. Variabel tambahan juga dipertimbangkan untuk aspek ekonomi yang berlaku berbeda untuk setiap lokasi bandara. Proses dalam penelitian ini mencakup analisis model dua algoritma *Recurrent Neural Network* (RNN), yaitu LSTM dan GRU. Penelitian ini menggabungkan masing-masing model dengan model VAR untuk meminimalkan kelemahan LSTM dan GRU. Hasil dari penelitian ini menunjukkan bahwa metode *hybrid* memberikan kinerja prediksi terbaik dengan MAPE sebesar 1,02% dan RMSE sebesar 81,52 untuk metode *hybrid* VAR-LSTM. Sedangkan, untuk metode *hybrid* VAR-GRU memperoleh MAPE sebesar 0,60% dan RMSE sebesar 62,53.

5. Penelitian Kelima (Munkhdalai dkk., 2020)

Penelitian yang dilakukan Munkhdalai dkk. (2020) adalah penelitian dalam meramalkan data *financial market* dengan menggunakan metode *hybrid* VAR-GRU. Data yang digunakan memiliki 3.608 observasi dengan 10 variabel yang bersumber dari *The Federal Reserve Bank of St. Louis* dan *Yahoo Finance*. Model ini terdiri dari dua lapisan, yaitu VAR untuk pemilihan variabel dan *lag*, serta GRU untuk prediksi. Penggunaan AIC dalam lapisan VAR dan uji kausalitas *Granger* memungkinkan pemilihan variabel dan *lag* optimal. Hasil eksperimen menunjukkan bahwa model *hybrid* lebih unggul dibandingkan dengan model tunggal dalam kinerja prediksi, di mana nilai evaluasi model dengan performa terbaik terletak pada variabel *National Association of Securities Dealers Automated Quotations* (NASDAQ), yakni memperoleh nilai evaluasi MAPE sebesar 0,603% dan RMSE sebesar 17,24.

## 2.2. Saham

Saham adalah instrumen keuangan yang menandakan kepemilikan atas suatu perusahaan atau dapat dikatakan bahwa pemegang saham memiliki bagian kepemilikan pada perusahaan tersebut. Semakin banyak saham yang dimiliki oleh seseorang, semakin besar pula pengaruh atau kekuasaannya di dalam perusahaan tersebut (Ismawati dkk., 2022). Pemantauan harga saham menjadi krusial bagi para investor karena hal ini mencerminkan performa perusahaan yang menerbitkan saham. Harga saham di pasar modal dapat dibagi ke dalam tiga kategori, yakni *Open price*, *High price*, *Low price*, dan *Close price*. *Open price* adalah harga saham atau aset keuangan lainnya pada saat pembukaan pasar. *High price* dan *Low price* mencerminkan harga paling tinggi dan harga paling rendah yang dicapai oleh saham selama satu hari perdagangan di bursa saham. Sedangkan, *Close price* merujuk pada harga terakhir suatu saham yang terjadi menjelang penutupan bursa saham (Rahman dkk., 2023).

Demi mengurangi risiko kerugian dalam perdagangan saham, investor perlu melakukan analisis terhadap data harga saham yang terjadi sepanjang waktu dan cenderung memiliki volatilitas tinggi. Harga saham tidak selalu menunjukkan tren naik atau turun yang berkelanjutan. Sebaliknya, terjadi fluktuasi di mana kenaikan dan penurunan bergantian seiring berjalannya waktu dan membentuk pola tertentu. Oleh karena itu, peramalan menjadi krusial untuk mengoptimalkan keuntungan, sesuai dengan tujuan utama investor di pasar modal.

## 2.3. Peramalan

Kegiatan memprediksi jumlah yang dibutuhkan di masa depan dikenal sebagai peramalan. Menurut Chopra dan Meindl (2016), peramalan adalah upaya untuk mengestimasi peristiwa atau kondisi yang akan terjadi di masa depan dan berada di luar kendali perusahaan. Pada konteks ini, peramalan menjadi unsur integral dari

manajemen strategis. Tujuan utama perencanaan strategis adalah memutuskan tindakan yang perusahaan seharusnya ambil di masa depan, termasuk bagaimana bersaing di pasar tertentu, menentukan jenis produk yang akan ditawarkan, dan mengarahkan pertumbuhan perusahaan. Perusahaan membutuhkan peramalan untuk mengelola tantangan seperti perubahan musim, fluktuasi dalam permintaan, strategi pemotongan harga dari pesaing, atau bahkan fluktuasi ekonomi yang signifikan. Alasan yang menyebabkan peramalan menjadi penting adalah karena prediksi kejadian di masa mendatang merupakan suatu kontribusi penting dalam proses perencanaan dan pengambilan keputusan yang dapat diimplementasikan pada berbagai bidang, seperti manajemen operasi, pemasaran, keuangan dan manajemen risiko, ekonomi, proses kontrol industri, dan demografi (Montgomery dkk., 2015).

Meskipun terdapat berbagai situasi yang membutuhkan peramalan, secara umum hanya ada dua kategori teknik peramalan, yakni metode kualitatif dan metode kuantitatif (Montgomery dkk., 2015).

#### 1. Metode Kualitatif

Penggunaan teknik peramalan kualitatif umumnya terjadi ketika terdapat keterbatasan atau bahkan ketiadaan data historis yang dapat dijadikan dasar peramalan. Sebagai ilustrasi, ketika memasarkan produk baru yang tidak memiliki catatan sejarah yang relevan. Pada konteks ini, perusahaan dapat mengandalkan pandangan ahli dari tim penjualan dan pemasaran untuk mengestimasi dengan pendekatan subyektif terhadap penjualan produk selama tahap pengenalan produk baru dalam siklus hidupnya.

#### 2. Metode Kuantitatif

Metode peramalan kuantitatif secara formal memanfaatkan data historis dan model peramalan untuk menggambarkan dan menyatakan hubungan statistik antara nilai-nilai sebelumnya dan saat ini dari variabel yang bersangkutan. Kemudian, model tersebut digunakan untuk meramalkan pola dalam data menuju masa depan, menjadikan model peramalan sebagai alat untuk meneruskan tren perilaku masa lalu dan saat ini ke masa depan. Terdapat beberapa jenis model peramalan yang umum digunakan, di antaranya adalah model *regression*, model *smoothing*, dan model deret waktu.

## 2.4. Data Deret waktu

Data deret waktu adalah data yang dirangkai berdasarkan waktu yang beruntun atau berurut yang dikumpulkan dari waktu ke waktu. Waktu yang dapat diterapkan dapat berupa minggu, bulan, tahun, dan lain sebagainya. Data deret waktu merupakan serangkaian data pengamatan yang bersumber dari sumber konsisten dan terjadi karena adanya indeks waktu yang berurutan dengan jarak yang konsisten (Cryer dan Chan, 2008).

Menurut Makridakis dkk. (1999), ada empat pola data yang umumnya muncul dalam analisis deret waktu.

1. Pola Horizontal: Pola data horizontal muncul ketika data deret waktu cenderung berfluktuasi di sekitar nilai konstan atau rata-rata tertentu.
2. Pola Musiman: Pola data musiman terjadi ketika data deret waktu dipengaruhi oleh faktor-faktor musiman yang menyebabkan fluktuasi berulang dalam data tersebut.
3. Pola Siklis: Pola data siklis terjadi ketika data deret waktu dipengaruhi oleh fluktuasi ekonomi atau pergerakan jangka panjang yang bersifat siklis.
4. Pola Tren: Pola data tren muncul ketika data deret waktu menunjukkan tren kenaikan atau penurunan jangka panjang yang dapat diidentifikasi.

## 2.5. Analisis Deret Waktu Multivariat

Analisis deret waktu multivariat adalah suatu rangkaian metode statistik yang secara bersamaan menilai beberapa pengukuran pada individu atau objek yang sedang diteliti. Oleh karena itu, setiap analisis yang melibatkan simultan lebih dari dua variabel dapat disebut sebagai analisis multivariat (Hair dkk., 2009). Analisis deret waktu multivariat dibagi menjadi beberapa bagian diantaranya adalah metode VAR, *Vector Moving Average* (VMA), *Vector Autoregressive Moving Average* (VARMA), dan VARIMA.

### 2.5.1. *Vector Autoregressive (VAR)*

Menurut Enders (1995), VAR merupakan pengembangan dari *Autoregressive (AR)* yang melibatkan beberapa data deret waktu secara paralel. Model VAR dianggap sebagai suatu sistem persamaan dinamis, di mana estimasi pada setiap periode waktu diubah berdasarkan perubahan variabel dan variabel lainnya yang terlibat dalam sistem pada periode sebelumnya. Model VAR menggunakan data yang stasioner. Secara umum, model VAR dapat dikomputasikan menggunakan persamaan (2.1) sebagai berikut (Wei, 2006):

$$Y_t = \Phi_0 + \sum_{i=1}^p \Phi_i Y_{t-i} + a_t \quad (2.1)$$

dengan,

- $Y_t$  : vektor deret waktu stasioner ukuran  $(n \times 1)$
- $\Phi_0$  : rata-rata vektor deret waktu  $(n \times 1)$
- $\Phi_p$  : matriks koefisien  $(n \times n)$ , dengan  $i = 1, 2, 3, \dots, p$
- $a_t$  : vektor *error* ukuran  $(n \times 1)$

### 2.5.2. *Vector Moving Average (VMA)*

Menurut Wei (2006), VMA adalah perluasan dari Metode MA. Model VMA beroperasi sebagai suatu sistem persamaan dinamis, di mana perkiraan pada suatu periode waktu diubah berdasarkan kesalahan variabel dan kesalahan variabel lain yang terlibat dalam sistem pada periode sebelumnya. Model untuk VMA dengan *order q* atau VMA (*q*) dapat dirumuskan pada persamaan (2.2) sebagai berikut:

$$Y_t = \Theta_0 + \sum_{i=1}^q \Theta_i a_{t-i} + a_t \quad (2.2)$$

dengan,

- $Y_t$  : vektor deret waktu stasioner ukuran  $(n \times 1)$
- $\Theta_0$  : rata-rata vektor deret waktu  $(n \times 1)$
- $\Theta_q$  : matriks koefisien VMA  $(n \times n)$
- $a_t$  : vektor *error* ukuran  $(n \times 1)$

### 2.5.3. *Vector Autoregressive Moving Average (VARMA)*

Model VARMA merupakan pengembangan dari *Autoregressive Moving Average (ARMA)*. Model VARMA adalah sistem persamaan dinamis, di mana estimasi periode waktu diubah berdasarkan pergerakan variabel, kesalahan variabel, serta kesalahan variabel lainnya yang terlibat dalam sistem pada periode sebelumnya. Rumus untuk VARMA ( $p, q$ ) diuraikan pada persamaan (2.3) sebagai berikut (Wei, 2006):

$$Y_t = \Phi_0 + \Phi_1 Y_{t-1} + \dots + \Phi_p Y_{t-p} + \alpha_t - \Theta_1 \alpha_{t-1} - \dots - \Theta_q \alpha_{t-q} \quad (2.3)$$

dengan,

$Y_t$  : vektor deret waktu stasioner ukuran ( $n \times 1$ )

$\Phi_0$  : rata-rata vektor deret waktu ( $n \times 1$ )

$\Phi_p$  : matriks koefisien VAR ( $n \times n$ )

$\Theta_0$  : rata-rata vektor deret waktu ( $n \times 1$ )

$\Theta_q$  : matriks koefisien ( $n \times n$ ), dengan  $i = 1, 2, 3 \dots q$

$\alpha_t$  : vektor kesalahan ukuran ( $n \times 1$ )

### 2.5.4. *Vector Autoregressive Integrated Moving Average (VARIMA)*

Model VARIMA adalah model VARMA yang melibatkan proses *differencing*. Model VARIMA digunakan ketika data tidak bersifat stasioner, sehingga perlu dilakukan proses *differencing* untuk membuat data tersebut menjadi stasioner. Berikut adalah persamaan (2.4) yang digunakan dalam komputasi model VARIMA (Wei, 2006):

$$\Phi_p(B)(I - B)^d Y_t = \Phi_0 + \Theta_q(B)\alpha_t \quad (2.4)$$

dengan,

$Y_t$  : vektor deret waktu stasioner ukuran ( $n \times 1$ )

$\Phi_p$  : matriks koefisien VAR ( $n \times n$ )

$\Theta_q$  : matriks koefisien VMA ( $n \times n$ )

- $(I - B)^d$ : komponen *differencing*  
 $a_t$  : vektor kesalahan ukuran ( $n \times 1$ )  
 $B$  : *backshift operator*

## 2.6. Stasioneritas Data

Kestasioneran dari sebuah deret waktu merujuk pada sifat di mana deret tersebut tetap stabil selama periode waktu tertentu dengan rata-rata yang konstan. Namun, dalam beberapa kasus nyata, keberadaan tren tertentu dapat mempengaruhi perilaku deret waktu tersebut, yang pada akhirnya berdampak pada kestasionerannya. Sebuah deret waktu dapat stasioner untuk periode pendek atau panjang. Hal tersebut menandakan adanya perubahan dalam tren deret tersebut (Silva dkk., 2021).

Salah satu uji yang dapat digunakan dalam penentuan stasioneritas adalah uji akar unit. Uji akar unit awalnya dikembangkan oleh Dickey dan Fuller (1979), sehingga sering dikenal sebagai uji akar unit *Dickey-Fuller*. Uji akar unit *Dickey-Fuller* mengasumsikan bahwa residual  $e_t$  adalah residual independen dengan rata-rata nol, varians konstan, dan tidak memiliki autokorelasi. Namun, dalam banyak kasus, residual  $e_t$  sering saling terhubung atau mengandung elemen-elemen autokorelasi. Oleh karena itu, perlu mengembangkan uji akar unit untuk data yang mengandung autokorelasi dalam residual  $e_t$ , di mana berdasarkan hasil uji ini akan dilihat apakah terdapat akar unit dalam model *Autoregressive* dengan adanya uji hipotesis nol ( $H_0$ ). Pada evolusinya, metode yang paling umum digunakan untuk menguji data stasioner saat ini adalah Uji *Augmented Dickey-Fuller* (ADF). Metode ini merupakan pengembangan dari uji akar unit *Dickey-Fuller* dengan dasar bahwa uji ADF mempertimbangkan potensi adanya autokorelasi dalam istilah kesalahan jika seri yang digunakan tidak bersifat stasioner. Proses pengujian akar unit dengan metode uji ADF melibatkan persamaan (2.5) berikut (Roza dkk., 2022):

$$Y_t = \rho Y_{t-1} + \mu_t \quad (2.5)$$

dengan,

$\rho$  : koefisien AR

$\mu_t$ : kesalahan *white noise*, yang memiliki rata-rata 0, varians konstan, dan tidak mengandung autokorelasi.

$Y_t$ : nilai deret waktu pada waktu ke- $t$

Pada persamaan di atas, jika  $\rho = 1$ , maka dapat diputuskan bahwa variabel  $t$  memiliki akar unit, di mana hipotesis yang digunakan dalam uji ADF yaitu sebagai berikut:

$H_0 : \rho = 0$ , yang artinya data tidak stasioner karena terdapat akar unit dalam data

$H_1 : \rho \neq 0$  yang artinya data stasioner karena tidak terdapat akar unit dalam data.

Pada uji ini, jika suatu data *input* tidak stasioner, maka diperlukan penyesuaian untuk menciptakan data yang stasioner. Salah satu pendekatan umum adalah menggunakan metode *differencing*. *Differencing* adalah teknik dalam analisis deret waktu yang melibatkan mengubah deret waktu menjadi deret yang terdiri dari selisih antara nilai pada waktu tertentu dan nilai pada waktu yang diambil  $n$  langkah ke depan. Hal ini dilakukan untuk membuat deret waktu menjadi stasioner, yaitu deret yang memiliki rata-rata, varians, dan kovariansi yang konstan sepanjang waktu (Beard dkk., 2019). Persamaan (2.5) selanjutnya disusun ulang untuk mendapatkan suatu persamaan (2.6) yang menggunakan operasi *differencing*, yaitu sebagai berikut:

$$\Delta Y_t = (\rho - 1)Y_{t-1} + \mu_t \quad (2.6)$$

dengan,

$\Delta Y_t$  : *differencing* pertama

$\rho - 1$  : koefisien AR pada *differencing* pertama

## 2.7. Uji Asumsi Residual

Uji asumsi residual yang dilakukan pada penelitian ini adalah uji asumsi *white noise*. Menurut Cryer dan Chan (2008), *white noise* didefinisikan sebagai rangkaian variabel acak independen dan terdistribusi secara identik. Uji asumsi residual *white noise* dilakukan dengan tujuan untuk memeriksa adanya korelasi antara vektor residual yang berasal dari model yang telah dibangun. Salah satu uji yang dapat digunakan dalam asumsi residual *white noise* adalah uji Ljung-Box. Uji ini dijelaskan secara eksplisit oleh Ljung dan Box (1978) yang dapat diuraikan dengan tahapan berikut:

1. Hipotesis

$H_0$ : Model memenuhi syarat *white noise*

$H_1$ : Model tidak memenuhi syarat *white noise*

2. Tingkat signifikansi ( $\alpha$ )

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

3. Daerah kritis

Tolak  $H_0$  jika nilai  $Q_{hitung} > x^2$  atau p-value  $< \alpha$

Tidak tolak  $H_0$  jika nilai  $Q_{hitung} < x^2$  atau p-value  $> \alpha$

4. Statistik uji

Statistik uji yang digunakan adalah statistik uji *Ljung-Box* yang diuraikan dalam persamaan (2.7) sebagai berikut:

$$Q_{hitung} = n(n + 2) \sum_{i=1}^m \frac{\rho_i^2}{(n-i)} \quad (2.7)$$

dengan,

$n$  : banyaknya data pengamatan

$\rho_i$  : autokorelasi *lag* ke-  $i$

$m$  : jumlah *lag*

5. Keputusan dan kesimpulan

## 2.8. Lag Optimum

Pemilihan *lag* optimum diterapkan pada data deret waktu untuk menentukan model VARIMA terbaik dan untuk menentukan berapa banyak *lag* yang harus digunakan dalam persamaan. Salah satu cara untuk memilih *lag* optimum adalah dengan mempertimbangkan jumlah parameter dalam model menggunakan *Akaike Information Criteria* (AIC). AIC adalah salah satu kriteria pemilihan model yang pertama kali diperhatikan secara luas dalam komunitas statistik, dan tetap menjadi salah satu alat yang paling terkenal dan sering dipakai dalam praktek statistik. Kriteria ini diperkenalkan oleh Hirotugu Akaike pada tahun 1973. Model dengan kinerja yang optimal cenderung memiliki nilai AIC yang rendah (Cavanaugh dan Neath, 2019). Persamaan yang digunakan dalam komputasi AIC dapat diuraikan pada persamaan (2.8) sebagai berikut:

$$AIC(p, q) = \ln |\Sigma| + \frac{2k^2(p+q)}{T} \quad (2.8)$$

dengan,

- $p$  : *lag* VAR
- $q$  : *lag* VMA
- $\ln$  : Natural log
- $k$  : Jumlah variabel tak bebas
- $T$  : Banyak observasi
- $\Sigma$  : Matriks kovarian residual

## 2.9. Normalisasi dan Denormalisasi Data

Normalisasi merupakan langkah pada data mentah yang bertujuan untuk mereskal atau mengubah data sehingga setiap fitur memberikan kontribusi secara seragam. Salah satu tujuannya adalah mengatasi dua masalah utama pada data yang bisa menghambat proses pembelajaran algoritma *machine learning*, yakni keberadaan fitur yang dominan dan nilai-nilai *outlier* (Singh dan Singh, 2019). Terdapat berbagai metode untuk menormalisasi data, di antaranya adalah normalisasi *min-*

*max* dan normalisasi *standard*.

Menurut Koduru (2022), normalisasi *min-max* adalah teknik umum yang digunakan untuk mengubah data sehingga atribut sensitifnya tetap terjaga. Nilai dalam kumpulan data asli disesuaikan menggunakan metode normalisasi *min-max*, dengan mempertimbangkan nilai minimum dan maksimum dalam kumpulan data. Proses standarisasi *min-max* melibatkan transformasi linier terhadap data asli. Data diubah kembali dalam rentang antara 0 hingga 1 atau -1 hingga 1. Pendekatan normalisasi ini bermanfaat ketika nilai minimum dan maksimum aktual dari data tidak diketahui, atau ketika terdapat nilai *outlier* yang mempengaruhi dominasi normalisasi *min-max*. Persamaan yang digunakan dalam komputasi normalisasi *min-max* dapat diuraikan dengan persamaan (2.9) sebagai berikut:

$$x' = \frac{(x-x_{min})}{(x_{max}-x_{min})} \quad (2.9)$$

dengan,

- $x'$  : Nilai baru hasil normalisasi
- $x$  : Nilai pada variabel  $x$  yang akan dinormalisasi
- $x_{max}$  : Nilai maksimal variabel  $x$
- $x_{min}$  : Nilai minimum variabel  $x$

Berdasarkan persamaan di atas, diperoleh nilai hasil peramalan. Nilai ini selanjutnya dilakukan proses denormalisasi untuk mengembalikan nilai hasil peramalan ke nilai yang sebenarnya. Denormalisasi merupakan langkah yang melibatkan proses terbalik dari normalisasi, di mana data dikembalikan ke bentuk asli sebelum mengalami proses normalisasi (Dewi dkk., 2018). Menurut Tasyurek (2022), Denormalisasi dapat dijelaskan sebagai proses menambahkan baris atau kolom yang tidak diperlukan untuk meningkatkan kinerja pembacaan *database* ke dalam sistem *database* yang telah dinormalisasi. Komputasi yang digunakan dalam denormalisasi *min-max* dapat diuraikan dengan persamaan (2.10) sebagai berikut:

$$x = x_{min} + (x'(x_{max} - x_{min})) \quad (2.10)$$

Sedangkan, normalisasi *standard* adalah alat dalam *machine learning* yang mengharapakan data terdistribusi secara normal dalam setiap fitur dan menskalakannya sehingga distribusi data berpusat di sekitar 0 dengan standar deviasi 1. Rata-rata dan standar deviasi dihitung untuk setiap fitur, dan kemudian fitur tersebut diskalakan berdasarkan nilai-nilai tersebut. Persamaan yang digunakan untuk komputasi normalisasi *standard* dapat diuraikan dengan persamaan (2.11) sebagai berikut (Raju dkk., 2020):

$$x' = \frac{(x-\mu)}{\sigma} \quad (2.11)$$

dengan,

- $x'$  : Nilai baru hasil normalisasi
- $x$  : Nilai pada variabel  $x$  yang akan dinormalisasi
- $\mu$  : Nilai rata-rata
- $\sigma$  : Standar Deviasi

Komputasi yang digunakan dalam denormalisasi *standard* dapat diuraikan dengan persamaan (2.12) sebagai berikut:

$$x = \mu + (x'\sigma) \quad (2.12)$$

## 2.10. *Machine learning*

Menurut Nichols dkk. (2018), *machine learning* merujuk pada kumpulan algoritma yang cerdas dalam melakukan prediksi berdasarkan data besar dan kompleks, mampu mencapai tingkat pemahaman semantik dan ekstraksi informasi yang menyerupai kemampuan manusia, bahkan kadang-kadang dapat mendeteksi pola abstrak dengan akurasi yang melebihi kemampuan manusia. Perluasan dari teknik klasik dalam pemodelan statistik, *machine learning* modern menjadi alat yang sangat efektif berkat peningkatan besar dalam volume data, pertumbuhan eksponensial daya komputasi, dan perkembangan algoritma yang didorong oleh kebutuhan industri web.

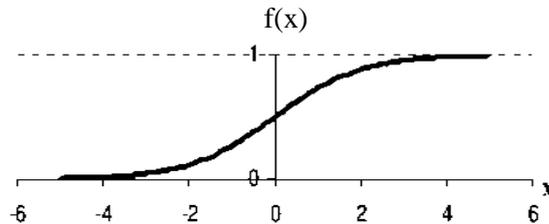
*Machine learning* dan *data mining* memiliki hubungan yang erat, karena keduanya kerap memanfaatkan metode serupa dan saling tumpang tindih. Meskipun demikian, *machine learning* lebih difokuskan pada prediksi berdasarkan properti yang sudah diketahui, sementara *data mining* lebih menitikberatkan pada penemuan properti yang masih belum teridentifikasi. Proses *data mining* sendiri melibatkan penerapan metode *machine learning*, begitu pula sebaliknya, di mana *machine learning* juga menggunakan metode *data mining*. Meskipun demikian, tujuan keduanya berbeda atau digunakan untuk meningkatkan akurasi pembelajaran. Dengan demikian, keduanya saling melengkapi dan dapat saling diperkaya dalam analisis data.

## 2.10. Fungsi Aktivasi

Pada jaringan saraf, jika setiap aktivasi *neuron* dihitung sebagai jumlah diboboti dari aktivasinya, maka neuron tersebut akan menjadi fungsi linier. Fungsi aktivasi bertujuan untuk memodelkan *output* dari seluruh jaringan dengan memanfaatkan fungsi linier. Pada *machine learning*, fungsi aktivasi digunakan untuk memasukkan nonlinieritas ke dalam jaringan saraf (Pomerat dkk., 2019). Menurut Feng dan Lu (2019), fungsi aktivasi terbagi menjadi dua kategori dasar yaitu fungsi aktivasi linier, yang mempertahankan kestabilan nilai, dan fungsi aktivasi nonlinier, yang menciptakan lebih banyak variasi dengan memanfaatkan struktur jaringan saraf. Fungsi aktivasi linier memiliki persamaan yang menyerupai garis lurus, di mana aktivasinya sebanding dengan *input*-nya. Jika dibandingkan dengan fungsi aktivasi linier, fungsi aktivasi nonlinier lebih umum digunakan. Fungsi memudahkan model untuk melakukan generalisasi atau menyesuaikan diri dengan berbagai jenis data dan membedakan antara keluaran. Terdapat berbagai jenis fungsi aktivasi nonlinier yang umumnya diklasifikasikan berdasarkan rentang atau kurva karakteristiknya, yaitu seperti fungsi aktivasi *sigmoid*, *tanh*, dan *ReLU*. Dua fungsi aktivasi yang akan digunakan pada penelitian ini adalah sebagai berikut:

### 1. Fungsi *Sigmoid*

Fungsi *sigmoid* merupakan fungsi nonlinier dengan nilai berkisar antara 0 hingga 1. *Sigmoid* merujuk pada bentuk 's', dan bentuk logistik dari *sigmoid* memetakan rentang nilai dari  $(-\infty, \infty)$  ke  $(0, 1)$  seperti yang diperlihatkan dalam Gambar 1 berikut (Karlik dan Olgac, 2019):



Gambar 1. Fungsi *Sigmoid*  
(Sumber : Karlik dan Olgac, 2019).

Fungsi *sigmoid* dapat dikomputasikan menggunakan persamaan (2.13) sebagai berikut:

$$f(x) = \frac{1}{1 + e^{-x}} \quad (2.13)$$

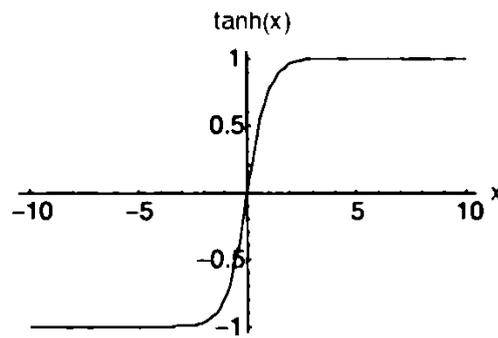
dengan,

$f(x)$  : Hasil *output* fungsi *tanh* berupa nilai dalam rentang 0 sampai 1

$e^{-x}$  : Nilai eksponensial negatif dari nilai data *input*

### 2. Fungsi *Tanh*

Fungsi aktivasi *tanh* adalah fungsi simetris berbentuk s yang menghasilkan *output* pada rentang nilai -1 hingga +1. Fungsi ini dengan mudah dapat dijelaskan sebagai rasio antara sinus hiperbolik dan kosinus, atau dapat diperluas sebagai rasio antara setengah selisih dan setengah jumlah dari dua fungsi eksponensial di titik  $x$  dan  $-x$ . Pada representasinya, fungsi *tanh* dapat dijelaskan melalui Gambar 2 sebagai berikut (Karlik dan Olgac, 2019):



Gambar 2. Fungsi *Tanh*  
(Sumber : Karlik dan Olgac, 2019).

Fungsi *tanh* dikomputasikan menggunakan persamaan (2.14) berikut:

$$f(x) = \frac{\sinh x_i}{\cosh x_i} = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (2.14)$$

dengan,

$f(x)$  : Hasil *output* fungsi *tanh* berupa nilai dalam rentang -1 sampai 1

$e^{-x}$  : Nilai eksponensial negatif dari nilai data *input*

$e^x$  : Nilai eksponensial positif dari nilai data *input*

### 2.11. *Deep learning*

*Deep learning* merupakan metode pendekatan *machine learning* yang mengandalkan pengetahuan tentang otak manusia, statistik, dan matematika terapan yang telah berkembang selama beberapa dekade terakhir (Goodfellow dkk., 2016). Menurut Chahal dan Gulia (2019), *deep learning* adalah serangkaian algoritma *machine learning* yang menggunakan berbagai lapisan untuk merepresentasikan tingkat abstraksi berbeda, terdiri dari lapisan *input*, lapisan *output*, dan beberapa *hidden layer*. *Deep learning* digunakan untuk sintesis suara, pemrosesan gambar, pengenalan tulisan tangan, deteksi objek, analisis prediksi, dan pengambilan keputusan. Belakangan ini, popularitas dan kegunaannya mengalami peningkatan signifikan, terutama karena adanya komputer yang lebih canggih, kumpulan data yang lebih besar, dan teknik pelatihan jaringan yang lebih kompleks.

Menurut Alzubaidi dkk. (2021), teknik *deep learning* dapat dikelompokkan menjadi tiga kategori utama, yaitu *unsupervised*, *supervised*, dan *semi-supervised learning*. Selain itu, *deep reinforcement learning* adalah jenis teknik pembelajaran lain yang sering dianggap sebagai bagian dari kategori *semi-supervised learning*.

#### 1. *Deep supervised learning*

Teknik ini berhubungan dengan data yang sudah diberi label. Keuntungan utama dari teknik ini adalah kemampuannya untuk mengumpulkan data atau menghasilkan *output* data dari pengetahuan sebelumnya. Namun, teknik ini memiliki kekurangan bahwa batas keputusan bisa menjadi terlalu ketat ketika data pelatihan tidak memiliki contoh yang seharusnya ada dalam sebuah kelas. Secara keseluruhan, teknik ini lebih sederhana daripada teknik lain dalam hal pembelajaran yang berkualitas tinggi

#### 2. *Deep semi-supervised learning*

Pembelajaran pada teknik ini dilakukan dengan menggunakan sebagian data yang telah diberi label. Hal ini membantu meminimalkan kebutuhan akan data yang berlabel secara keseluruhan. Namun, kekurangannya adalah jika terdapat fitur *input* yang tidak relevan dalam data pelatihan, maka keputusan yang dihasilkan dapat menjadi tidak akurat. Salah satu contoh penggunaannya adalah dalam klasifikasi dokumen teks, yang mana sulit untuk mendapatkan banyak dokumen yang sudah diberi label. Oleh karena itu, teknik ini sangat cocok untuk tugas klasifikasi dokumen teks.

#### 3. *Deep un-supervised learning*

Teknik ini memungkinkan proses pembelajaran tanpa memerlukan data berlabel. Pada teknik ini, agen belajar fitur-fitur penting atau representasi internal yang diperlukan untuk menemukan struktur atau hubungan yang tidak teridentifikasi dalam data *input*. Salah satu kekurangan utama pembelajaran tanpa pengawasan adalah ketidakmampuan dalam memberikan informasi yang akurat tentang pengurutan data dan kompleksitas komputasinya. Salah satu pendekatan pembelajaran tanpa pengawasan yang paling populer adalah pengelompokan.

#### 4. *Deep Reinforcement learning*

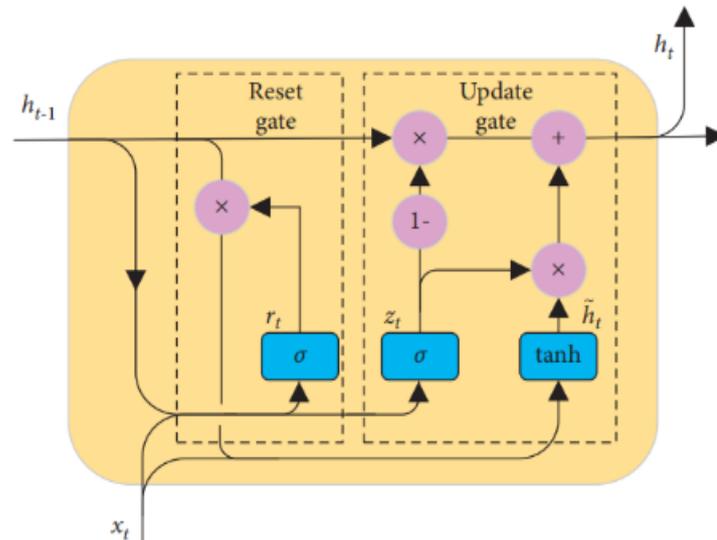
Pembelajaran *reinforced learning* melibatkan interaksi dengan lingkungan. Pembelajaran ini lebih sulit daripada metode *supervised* tradisional karena tidak memiliki fungsi kerugian langsung. Motivasi utama menggunakan pembelajaran ini adalah untuk membantu mengidentifikasi tindakan terbaik untuk mendapatkan *reward* tertinggi, menemukan situasi yang membutuhkan tindakan, dan menemukan pendekatan terbaik untuk mencapai *reward* besar. *Reinforcement learning* juga memberikan fungsi *reward* kepada agen pembelajaran. Namun, teknik ini tidak cocok digunakan dalam semua situasi, terutama jika data yang cukup sudah tersedia untuk menggunakan teknik *supervised*. Selain itu, pembelajaran ini membutuhkan komputasi yang berat dan memakan waktu. Kekurangan utama pembelajaran ini adalah bahwa parameter dapat mempengaruhi kecepatan pembelajaran.

#### 2.12. *Gated Recurrent Unit (GRU)*

Sebagai varian dari *Recurrent Neural Network (RNN)*, GRU memiliki struktur yang mirip dengan RNN dan berfungsi untuk memproses data deret waktu. Metode GRU membantu mengatasi masalah hilangnya gradien dan lonjakan yang terjadi selama pelatihan RNN, sehingga efektif dalam menyelesaikan masalah memori jangka panjang. Dibandingkan dengan varian RNN lainnya, seperti LSTM, GRU memiliki struktur yang lebih sederhana yang dapat mengurangi komputasi dan meningkatkan efisiensi pelatihan (Teng dkk., 2021).

Menurut Saha dkk. (2021), GRU merupakan bentuk yang lebih mutakhir dari LSTM. Struktur GRU mencakup unit GRU yang bertanggung jawab atas pengaturan fungsi *reset* dan *update* memori. Keseluruhan fungsionalitas GRU serupa dengan algoritma LSTM. Perbedaan utama antara GRU dan LSTM terletak pada fakta bahwa GRU memerlukan waktu komputasi yang lebih singkat dibandingkan LSTM dan memiliki jumlah parameter yang lebih sedikit. Hal ini memungkinkan GRU untuk melatih model dengan lebih efisien. Pada GRU,

terdapat dua gerbang yang berguna sebagai pengatur alur informasi yaitu, gerbang *update* dan gerbang *reset*. Diagram arsitektur model GRU dapat dilihat pada Gambar 3 sebagai berikut:



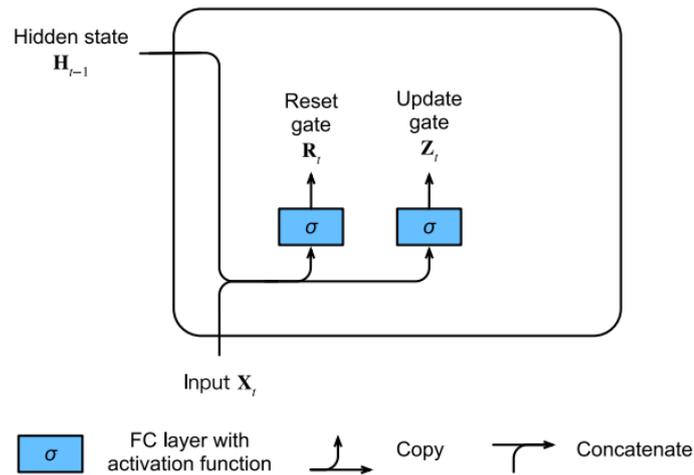
Gambar 3. Diagram Arsitektur Model GRU  
(Sumber : Teng dkk., 2021).

#### 1. Gerbang *reset* dan gerbang *update*

Menurut Zhang dkk. (2022), gerbang *reset* mengontrol seberapa banyak informasi dari langkah sebelumnya yang tetap diingat, sedangkan gerbang *update* berfungsi untuk mengendalikan seberapa banyak informasi baru yang akan dimasukkan. Pada model GRU, gambaran *input* untuk kedua gerbang ini diperlihatkan dalam Gambar 4 yang memperhitungkan data dari langkah waktu saat ini serta keadaan tersembunyi dari langkah waktu sebelumnya. *Output* dari kedua gerbang ini diperoleh dari dua lapisan terhubung penuh dengan fungsi aktivasi *sigmoid*.

*Output* dari gerbang *reset* memiliki rentang nilai antara 0 dan 1. Ketika nilai keluaran mendekati 0, ini mengindikasikan bahwa informasi dari langkah waktu sebelumnya memiliki dampak yang minim dan cenderung dihapus. Sebaliknya, jika nilai keluaran mendekati 1, itu mengindikasikan bahwa informasi dari langkah waktu terdahulu memiliki pengaruh yang signifikan dan akan dipertahankan. Sedangkan, jika *output* dari gerbang *update* bernilai 1,

maka hal ini menunjukkan bahwa informasi dari langkah waktu sebelumnya memiliki pengaruh pada *output* saat ini, sementara jika *output* bernilai 0, informasi dari langkah waktu sebelumnya tidak memengaruhi *output* saat ini. Arsitektur gerbang *update* dan gerbang *reset* dapat dilihat pada Gambar 4 sebagai berikut:



Gambar 4. Gerbang *Reset* dan Gerbang *Update*  
(Sumber : Zhang dkk., 2022)

Berikut persamaan (2.15) dan (2.16) yang digunakan untuk komputasi gerbang *reset* dan gerbang *update* (Zhang dkk., 2022).

$$R_t = \sigma(X_t W_{xr} + H_{t-1} W_{hr} + b_r) \quad (2.15)$$

$$Z_t = \sigma(X_t W_{xz} + H_{t-1} W_{hz} + b_z) \quad (2.16)$$

dengan,

- $R_t$  : gerbang *reset*
- $Z_t$  : gerbang *update*
- $\sigma$  : fungsi aktivasi *sigmoid*
- $X_t$  : nilai data *input* pada orde ke- $t$
- $W_{xr}, W_{hr}$  : nilai parameter bobot untuk gerbang *reset*
- $W_{xz}, W_{hz}$  : nilai parameter bobot untuk gerbang *update*
- $b_r, b_z$  : nilai parameter bias
- $H_{t-1}$  : nilai *output* sebelum orde ke- $t$

## 2. Kandidat *state* tersembunyi

Tahapan berikutnya melibatkan penentuan kandidat *state* tersembunyi atau *output* implisit pada langkah waktu saat ini ( $t$ ). Hal ini diperoleh dari informasi yang relevan pada langkah waktu sebelumnya ( $t - 1$ ) melalui penggunaan fungsi aktivasi *tanh*. Kandidat *state* tersembunyi ini bukanlah *output* akhir dari unit, melainkan merupakan memori konten pada langkah waktu saat ini. Nilai kandidat *state* tersembunyi dipengaruhi oleh *output* dari gerbang *reset*. Persamaan (2.17) berikut merupakan komputasi dalam penentuan nilai kandidat *state* tersembunyi (Zhang dkk., 2022).

$$\tilde{H}_t = \tanh(X_t W_{xh} + [R_t \odot H_{t-1}] W_{hh} + b_h) \quad (2.17)$$

dengan,

- $\tilde{H}_t$  : kandidat *state* tersembunyi
- $R_t$  : *output* pada gerbang *reset*
- $\odot$  : operator produk secara komponen
- $W_{xh}, W_{hh}$  : nilai parameter bobot
- $H_{t-1}$  : nilai *output* sebelum orde ke- $t$
- $X_t$  : nilai *input* pada orde ke- $t$
- $b_h$  : nilai parameter bias

## 3. *State* tersembunyi

Langkah terakhir melibatkan perhitungan *output* akhir dari unit saat ini, yang kemudian diteruskan ke langkah waktu berikutnya sebagai *state* tersembunyi. *State* tersembunyi yang dikirim ke langkah waktu berikutnya akan berfungsi sebagai dasar untuk menghitung *output* unit pada langkah waktu tersebut. Proses ini akan berulang dengan cara yang serupa pada setiap langkah waktu, hanya dengan nilai *input* yang berbeda. Perhitungan *output* akhir ini dipengaruhi oleh nilai kandidat *state* tersembunyi, nilai *state* tersembunyi dari langkah waktu sebelumnya, dan *output* dari gerbang *update*. Proses perhitungan nilai *state* tersembunyi dapat diamati dalam persamaan (2.18) (Zhang dkk., 2022).

$$H_t = Z_t \odot H_{t-1} + (1 - Z_t) \odot \tilde{H}_t \quad (2.18)$$

dengan,

$H_t$  : *output*

$\tilde{H}_t$  : kandidat *state* tersembunyi

$h_{t-1}$  : nilai *output* sebelum orde ke- $t$

$Z_h$  : *output* pada gerbang *update*

### 2.13. *Hyperparameter*

*Hyperparameter* adalah parameter-parameter yang digunakan untuk mengkonfigurasi model *machine learning* atau untuk menentukan algoritma yang digunakan dalam meminimalkan *loss function*, seperti fungsi aktivasi dan jenis *optimizer* dalam jaringan saraf (Yang dan Shami, 2020). Eksplorasi terhadap berbagai kemungkinan diperlukan untuk menciptakan model *machine learning* yang optimal. Proses merancang arsitektur model yang ideal dengan konfigurasi *hyperparameter* yang optimal dikenal sebagai *hyperparameter tuning*. *Hyperparameter tuning* dianggap sebagai bagian penting dalam membangun model *machine learning* yang efektif, terutama untuk model *machine learning* berbasis pohon dan *deep neural networks* yang memiliki banyak *hyperparameter* (Hutter dkk., 2019). Proses *hyperparameter tuning* memiliki beberapa variasi di antara berbagai algoritma *machine learning* karena adanya berbagai jenis *hyperparameter*, termasuk *hyperparameter* kategorikal, diskrit, dan kontinu (DeCastro-Garcia dkk., 2019).

*Hyperparameter tuning* dilakukan secara manual, tetapi terdapat faktor-faktor yang membuatnya tidak efektif, seperti jumlah *hyperparameter* yang besar, kompleksitas model, evaluasi model yang memakan waktu, dan interaksi *hyperparameter* yang nonlinier. Faktor-faktor ini telah mendorong pengembangan teknik baru dalam optimasi *hyperparameter* yang umumnya dikenal sebagai *hyperparameter optimization* (Yang dan Shami, 2020). Menurut Hutter dkk. (2019), beberapa alasan penting untuk menerapkan teknik *hyperparameter optimization* pada model *machine learning* adalah sebagai berikut:

1. Mengurangi upaya manusia yang diperlukan, karena banyak pengembang *machine learning* menghabiskan waktu yang signifikan untuk mengkonfigurasi *hyperparameter*, terutama untuk kumpulan data besar atau algoritma *machine learning* yang kompleks dengan banyak *hyperparameter*.
2. Meningkatkan kinerja model. Banyak *hyperparameter* memiliki optimum yang berbeda untuk mencapai kinerja terbaik dalam kumpulan data atau masalah yang berbeda.
3. Membuat model dan penelitian lebih dapat direproduksi. Hanya dengan menerapkan tingkat yang sama dari proses konfigurasi *hyperparameter*, berbagai algoritma *machine learning* dapat dibandingkan secara adil. Oleh karena itu, menggunakan metode *hyperparameter optimization* yang sama pada berbagai algoritma *machine learning* juga membantu menentukan model *machine learning* yang paling cocok untuk masalah tertentu.

*Hyperparameter Optimization* melibatkan berbagai metode, seperti *trial and error*, *grid search*, *random search*, dan lain sebagainya. Metode yang digunakan dalam penelitian ini adalah *grid search*. *Grid search* adalah metode yang paling umum digunakan untuk menjelajahi ruang konfigurasi *hyperparameter*. Metode ini bekerja dengan mengevaluasi produk *Cartesian* dari sekumpulan nilai yang telah ditentukan sebelumnya oleh pengguna (Hutter dkk., 2019). *Grid search* dapat dengan mudah diimplementasikan dan diparalelkan. Namun, kelemahan utama metode ini adalah ketidakefisienannya dalam ruang konfigurasi *hyperparameter* berdimensi tinggi, karena jumlah evaluasi meningkat secara eksponensial seiring dengan bertambahnya jumlah *hyperparameter* (Yang dan Shami, 2020).

#### **2.14. Model Hybrid VARIMA-GRU**

Model *hybrid* adalah hasil kombinasi dua metode yang dianggap memiliki tingkat akurasi yang tinggi. Pemanfaatan metode *machine learning* bertujuan untuk meraih hasil prediksi dengan tingkat akurasi yang tinggi dan pemahaman mendalam dalam setiap pengambilan keputusan. Oleh karena itu, hasil prediksi

yang diperoleh dari penerapan metode ini dapat diandalkan. Upaya untuk meningkatkan nilai akurasi tersebut dapat dilakukan dengan menggabungkan model *hybrid* (Azhima dkk., 2022). Secara umum, diasumsikan bahwa penggabungan model *hybrid* deret waktu memiliki dua komponen yaitu, linier dan nonlinier. Komponen-komponen tersebut dapat dikomputasikan dengan persamaan (2.19) berikut (Zhang, 2003):

$$y_t = L_t + N_t \quad (2.19)$$

dengan,

- $y_t$  : Nilai aktual ke- $t$
- $L_t$  : Komponen linier ke- $t$
- $N_t$  : Komponen tidak linier ke- $t$
- $t$  : Indeks waktu

Metode *hybrid* VARIMA-GRU terdiri dari dua tahap, dimulai dengan penerapan hasil metode VARIMA pada data deret waktu sebagai komponen linier, diikuti oleh penerapan metode GRU pada data residual untuk menangani pemodelan tidak linier. Setelah itu, kedua model, baik yang bersifat linier maupun nonlinier, digabungkan. Berikut merupakan persamaan (2.20) yang digunakan dalam komputasi metode *hybrid* VARIMA-GRU :

$$y'_t = L'_t + N'_t \quad (2.20)$$

dengan,

- $y'_t$  : Prediksi dari model *hybrid* ke- $t$
- $L'_t$  : Prediksi komponen linier dari GRU ke- $t$
- $N'_t$  : Prediksi komponen residual dari GRU ke- $t$

### 2.15. Evaluasi Model

Ketetapan atau akurasi mencerminkan sejauh mana suatu model peramalan mampu menghasilkan ulang data yang telah diketahui (Makridakis dkk., 1999). Untuk menilai sejauh mana kesalahan dihasilkan dan mengukur akurasi model, digunakan

metode-metode tertentu seperti sebagai berikut.

1. *Root Mean Square Error* (RMSE)

*Root Mean Square Error* adalah akar kuadrat dari *Mean Square Error* (MSE). Saat mengambil akar dari MSE, hal tersebut tidak mengubah peringkat relatif model, tetapi menghasilkan metrik dengan unit yang sama dengan variabel yang diprediksi. Hal ini memberikan gambaran tentang kesalahan rata-rata atau standar untuk kesalahan yang terdistribusi secara normal. Metode RMSE memberikan ukuran yang mudah dipahami tentang seberapa jauh nilai prediksi dari nilai sebenarnya, dengan mempertimbangkan variabilitas dalam data. Metode perhitungan dengan RMSE diperlihatkan pada persamaan (2.21) sebagai berikut (Hodson, 2022).

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{t=1}^n (X_t - F_t)^2} \quad (2.21)$$

dengan,

$X_t$  : Data aktual pada periode  $t$

$F_t$  : Nilai peramalan pada periode  $t$

$n$  : Jumlah data

2. *Mean Absolute Percentage Error* (MAPE)

*Mean Absolute Percentage Error* adalah metrik evaluasi yang digunakan dalam analisis peramalan untuk mengukur tingkat kesalahan relatif antara nilai prediksi dan nilai aktual dalam sebuah seri data. Metrik evaluasi MAPE dihitung dengan mengambil rata-rata dari nilai absolut dari selisih antara nilai prediksi dan nilai aktual, kemudian membaginya dengan nilai aktual, dan mengalikan hasilnya dengan 100% untuk mengungkapkan kesalahan dalam bentuk persentase. MAPE memungkinkan perbandingan akurasi antara metode peramalan yang berbeda atau dalam konteks seri data yang berbeda, serta memberikan gambaran tentang akurasi model dengan mengekspresikannya dalam rata-rata kesalahan persentase absolut (Prayudani dkk., 2019). Apabila nilai MAPE berada di bawah 10%, berarti suatu metode memiliki kemampuan yang sangat baik. Sedangkan, apabila nilai MAPE di antara 10% dan 20%, berarti kemampuan metode tersebut baik. Berikut persamaan (2.22) yang digunakan dalam mencari nilai MAPE :

$$\text{MAPE} = \frac{1}{n} \sum_{t=1}^n \frac{|X_t - F_t|}{|X_t|} \quad (2.22)$$

dengan,

$X_t$  : Data aktual pada periode  $t$

$F_t$  : Nilai peramalan pada periode  $t$

$n$  : Jumlah data

## 2.16. Uji *T-Hotelling*

Uji *t-Hotelling* adalah perluasan dalam analisis multivariat dari uji *t-Student* yang digunakan untuk menguji apakah terdapat perbedaan antara rata-rata dari vektor-vektor yang berbeda (Adubisi dkk., 2019). Pada penelitian ini perbedaan antara rata-rata hasil peramalan metode VARIMA dan model *hybrid* VARIMA-GRU akan diuji menggunakan uji *t-Hotelling*. Komputasi uji *t-Hotelling* dapat dilakukan menggunakan persamaan (2.23) sebagai berikut:

$$T^2 = \frac{n_1 n_2}{n_1 + n_2} (\bar{x}_1 - \bar{x}_2)' S_p^{-1} (\bar{x}_1 - \bar{x}_2) \quad (2.23)$$

dengan,

$T^2$ : Statistik uji *t-Hotelling*

$n_1$  : Populasi pertama

$n_2$  : Populasi kedua

$\bar{x}_1$  : Rata-rata sampel vector populasi pertama

$\bar{x}_2$  : Rata-rata sampel vector populasi kedua

$S_p^{-1}$ : Kovariansi gabungan kedua populasi

$p$  : Jumlah variabel

### **III. METODOLOGI PENELITIAN**

#### **3.1. Tempat dan Waktu Penelitian**

##### **3.1.1. Tempat Penelitian**

Penelitian ini dilakukan dengan studi literatur di Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lampung yang terletak di Jalan Prof. Sumantri Brojonegoro No. 1, Gedong Meneng, Bandar Lampung.

##### **3.1.2. Waktu Penelitian**

Penelitian ini dilakukan pada Semester Ganjil tahun akademik 2023/2024, dimulai pada bulan September 2023. Penelitian ini terbagi menjadi tiga tahap. Tahap pertama dimulai dengan studi literatur, pengumpulan jurnal, dan buku yang relevan dengan topik penelitian sebagai referensi ini akan digunakan dalam penyusunan proposal penelitian hingga tahap akhir. Tahap selanjutnya melibatkan pengumpulan data sebagai bahan untuk penelitian dan penyusunan draf proposal penelitian. Tahap kedua mencakup pengerjaan program, termasuk *preprocessing* data, pembangunan model VARIMA, pembagian data, pembangunan model *hybrid* VARIMA-GRU, dan evaluasi kinerja model. Tahap terakhir melibatkan penyusunan hasil pengujian dan analisis dalam bentuk draf laporan, yang akan disampaikan dalam seminar hasil dan sidang komprehensif.

## 3.2. Data dan Alat Penelitian

### 3.2.1. Data penelitian

Data yang digunakan dalam penelitian ini adalah data sekunder yaitu data harga saham PT Bank Rakyat Indonesia yang diperoleh dari situs <https://finance.yahoo.com/quote/BBRI.JK> yang mencakup rentang waktu dari 1 Januari 2019 hingga 30 November 2023, dengan jumlah data sebanyak 1213 observasi. Data yang digunakan terdiri dari 4 variabel yaitu, *Open*, *High*, *Low*, dan *Close* yang ditampilkan pada Tabel 2 sebagai berikut:

Tabel 2. Data Harga Saham PT Bank Rakyat Indonesia

<i>Date</i>	<i>Open</i>	<i>High</i>	<i>Low</i>	<i>Close</i>
2019-01-01	3327,215332	3327.215332	3327.215332	3327.215332
2019-01-02	3281,761475	3309.033691	3263.580078	3281.761475
2019-01-03	3254,489258	3309.033691	3254.489258	3290.852295
2019-01-04	3290,852295	3327.215332	3281.761475	3327.215332
2019-01-07	3363,578369	3372.668945	3327.215332	3327.215332
...	...	...	...	...
2023-11-24	5400,000000	5450,000000	5375,000000	5400,000000
2023-11-27	5425,000000	5450,000000	5350,000000	5350,000000
2023-11-28	5325,000000	5425,000000	5325,000000	5350,000000
2023-11-29	5350,000000	5375,000000	5300,000000	5300,000000
2023-11-30	5375,000000	5375,000000	5275,000000	5275,000000

### 3.2.2. Alat Penelitian

Adapun alat yang digunakan dalam penelitian ini adalah sebagai berikut:

#### 1. Perangkat keras (*Hardware*)

Perangkat keras yang digunakan dalam penelitian ini yaitu:

- *Processor* : Intel(R) Core(TM) i3-6006U CPU @ 2.00GHz (4 CPUs), ~2.0GHz
- *Installed RAM* : 4096MB

## 2. Perangkat lunak (*Software*)

Perangkat lunak yang digunakan dalam penelitian ini adalah sebagai berikut:

- Sistem Operasi *Windows 10 Pro 64-bit*
- *Google Colaboratory 3.10.12*

Adapun *package* yang digunakan dalam penelitian ini adalah sebagai berikut:

- *TensorFlow 2.12.0*

*TensorFlow* memiliki berbagai kegunaan dalam pemodelan *hybrid*, termasuk implementasi model *machine learning* modern, pengolahan data, dan pelatihan *neural networks*.

- *Scikit-learn 1.2.2*

Digunakan untuk pemrosesan data, evaluasi model, dan pemilihan *hyperparameter*.

- *NumPy 1.23.5*

*NumPy* digunakan untuk manipulasi dan operasi data numerik.

- *Pandas 1.5.3*

Digunakan untuk manipulasi dan analisis data.

- *Matplotlib 3.7.1*

*Matplotlib* digunakan untuk membuat berbagai jenis grafik dan visualisasi data, termasuk grafik garis, *scatter plot*, histogram, *bar chart*, dan banyak lagi.

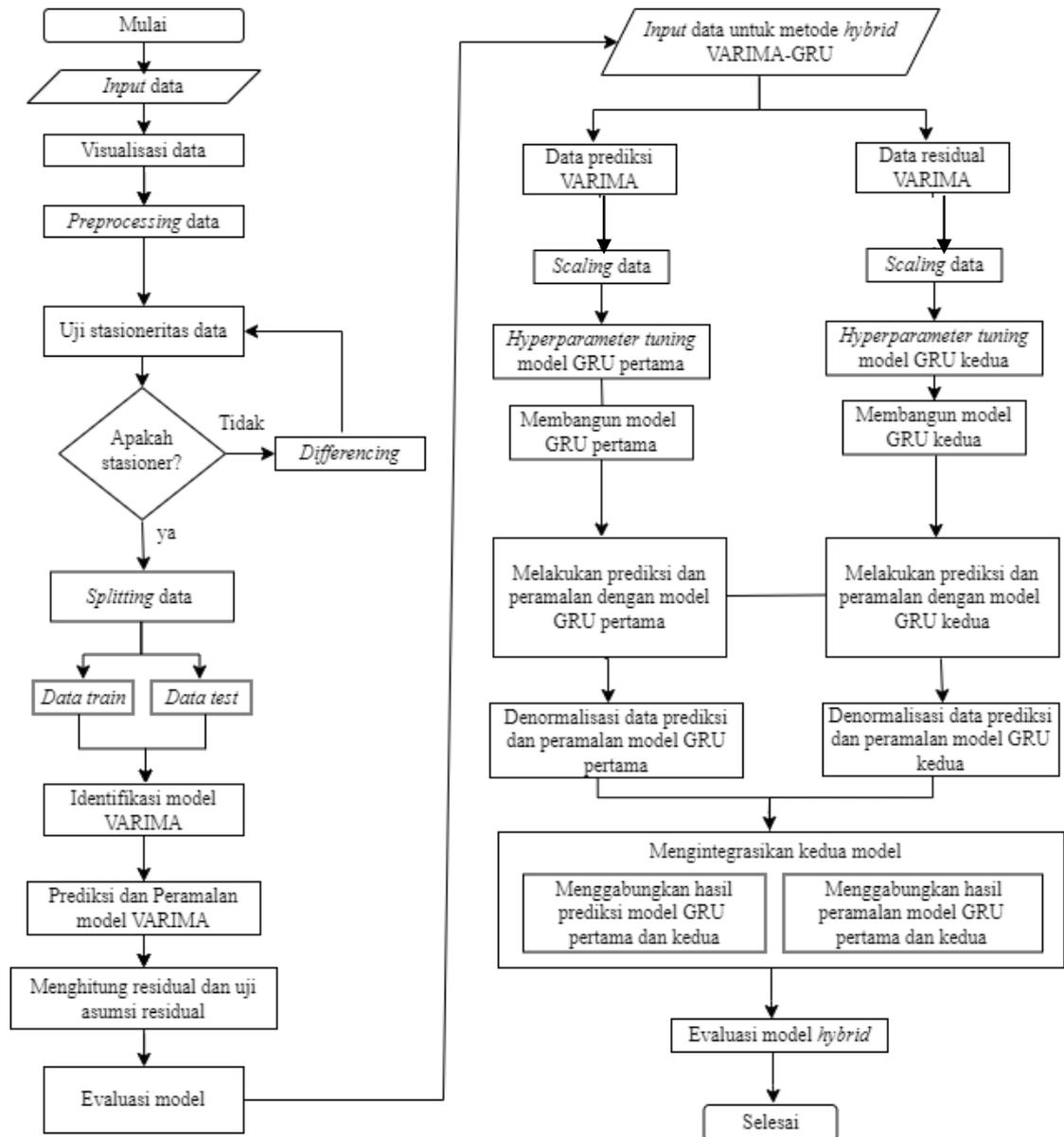
- *Statsmodels 0.14.0*

*Statsmodels* adalah *package* yang menyediakan alat untuk pemodelan statistik dan pengujian hipotesis.

### 3.3. Metode Penelitian

Penelitian ini diawali dengan studi literatur yang melibatkan eksplorasi jurnal, buku, dan artikel yang diperoleh penulis dari berbagai sumber, termasuk internet. Tujuan dari studi literatur ini adalah untuk mengumpulkan data dan informasi yang

relevan guna mendukung penyusunan penelitian ini. Selanjutnya, data yang terkumpul dianalisis melalui simulasi sebagai bentuk implementasi untuk menjelaskan konsep teori yang telah dikaji. Proses penelitian ini dijelaskan dalam bentuk *flowchart* pada Gambar 5 sebagai berikut:



Gambar 5. *Flowchart* Metode Hybrid VARIMA-GRU.

Berdasarkan *flowchart* di atas, tahap yang dilakukan dalam penelitian dapat dijelaskan sebagai berikut:

1. Melakukan *Input Data*  
Tahap pertama yang dilakukan adalah memasukkan data, yaitu data saham PT Bank Rakyat Indonesia yang akan digunakan dalam proses peramalan dengan metode *hybrid* VARIMA-GRU.
2. Melakukan Visualisasi Data  
Visualisasi data dilakukan untuk melihat tren, pola, dan hubungan antar variabel pada data.
3. Melakukan *Preprocessing Data*  
Hal ini dilakukan untuk melihat ada atau tidaknya *missing value* pada data yang akan digunakan.
4. Melakukan Uji Stasioneritas Data  
Pada tahap ini, uji stasioneritas yang dilakukan adalah uji stasioneritas rata-rata, yaitu menggunakan *Dickey-Fuller test*, dimana jika data tidak stasioner, maka harus dilakukan *differencing* data.
5. Identifikasi Model VARIMA  
Setelah data stasioner terpenuhi, dilakukan identifikasi model dengan melihat plot ACF dan PACF.
6. *Splitting Data Differencing*  
Sebelum membentuk model VARIMA, dilakukan *splitting* data terlebih dahulu. *Splitting* dilakukan dengan dua skema. Skema pertama dengan pembagian 80% data *training*, 20% data *testing*. Skema kedua sebesar 90% data *training* dan 10% data *testing*.
7. Membentuk Model VARIMA  
Model VARIMA dibentuk dari proses identifikasi model yang dilihat dari  $p$  dengan plot PACF,  $d$  dengan *differencing*, dan  $q$  dengan plot ACF melalui nilai *Akaike Information Criterion* (AIC) terkecil.
8. Melakukan Prediksi dengan Model VARIMA  
Pada tahap ini, data diprediksi menggunakan model VARIMA terbaik yang sudah ditentukan.
9. Menghitung Nilai Residual

Nilai residual diperoleh dari hasil pengurangan nilai pada data aktual dengan hasil prediksi.

10. Melakukan Uji Asumsi Residual  
Uji ini dilakukan menggunakan uji *Ljung-Box* untuk asumsi *white noise*.
11. Evaluasi Model  
Model yang diperoleh dievaluasi menggunakan *Root Mean Square Error* (RMSE) dan *Mean Absolute Percentage Error* (MAPE).
12. *Input* Data Hasil Prediksi dan Residual Model VARIMA  
*Input* data hasil prediksi dan data residual yang diperoleh sebelumnya untuk proses pembentukan model *hybrid* VARIMA-GRU.
13. Melakukan *Scaling* Data  
*Scaling* data pada data prediksi dan residual dari model VARIMA dilakukan menggunakan normalisasi *min-max* dan normalisasi *standard*.
14. Melakukan *Hyperparameter Tuning*  
Pada langkah ini, dua proses *hyperparameter* dilakukan dengan menggunakan metode *grid search*. *Hyperparameter* untuk Model GRU pertama melibatkan data prediksi dari model VARIMA, sementara *hyperparameter* untuk Model GRU kedua melibatkan data residual dari model VARIMA.
15. Membangun model GRU Pertama dan Model GRU Kedua  
Model GRU pertama dan model GRU kedua dibangun dengan memanfaatkan parameter terbaik yang diperoleh melalui proses *hyperparameter tuning*.
16. Melakukan Prediksi dan Peramalan  
Terdapat dua langkah dalam proses prediksi dan peramalan. Pertama, prediksi dan peramalan akan dilakukan menggunakan Model GRU pertama dengan menggunakan data prediksi dari model VARIMA. Langkah berikutnya akan melibatkan prediksi dan peramalan menggunakan Model GRU pertama dengan menggunakan data residual dari model VARIMA.
17. Menggabungkan Hasil Prediksi  
Penggabungan hasil prediksi kedua model GRU dilakukan dengan menggunakan proses operasi penjumlahan.
18. Evaluasi Model *Hybrid* VARIMA-GRU

Pada langkah ini, model-model GRU yang telah dibangun akan dievaluasi dengan menggunakan metrik RMSE dan MAPE.

19. Menggabungkan Hasil Peramalan

Pada tahap terakhir, hasil peramalan kedua model GRU digabungkan dengan menggunakan proses operasi penjumlahan.

## V. KESIMPULAN DAN SARAN

### 5.1. Kesimpulan

Berdasarkan hasil dan analisis yang telah diuraikan, penelitian tentang peramalan harga saham PT Bank Rakyat Indonesia menggunakan metode *hybrid* VARIMA-GRU memberikan beberapa kesimpulan sebagai berikut:

1. Model *hybrid* VARIMA-GRU merupakan integrasi antara model pertama yang dibangun dengan nilai prediksi model VARIMA dan model kedua yang dibangun dengan nilai residual model VARIMA. Model ini menghasilkan nilai RMSE dan MAPE yang rendah pada dua skema percobaan, menandakan kualitas model yang sangat baik dalam melakukan prediksi. Penggunaan skema 90% data *training* dan 10% data *testing* menghasilkan nilai RMSE sebesar 91,75 dan MAPE sebesar 0,0142%, sedangkan penggunaan skema 80% data *training* dan 20% data *testing* memperoleh RMSE sebesar 94,76 dan MAPE sebesar 0,0146%. Evaluasi menunjukkan bahwa skema 90% data *training* dan 10% data *testing* menghasilkan performa model yang lebih optimal. Hal ini terlihat dari nilai evaluasi yang lebih rendah, yang merepresentasikan tingkat kesalahan yang paling sedikit.
2. Peramalan harga saham PT Bank Rakyat Indonesia untuk 30 hari ke depan telah dilakukan menggunakan model *hybrid* VARIMA-GRU. Hasil peramalan model *hybrid* tersebut memiliki perbedaan signifikan dengan hasil peramalan model VARIMA. Model VARIMA memperoleh hasil evaluasi yang sangat baik, tetapi hasil peramalannya cenderung tidak berfluktuasi mengikuti pola data terbaru. Sedangkan, hasil peramalan model *hybrid* telah mengikuti pola data terbaru dengan evaluasi yang sangat baik. Hal ini menunjukkan bahwa

model *hybrid* VARIMA-GRU lebih efektif dalam meramalkan pergerakan harga saham PT Bank Rakyat Indonesia.

## 5.2. Saran

Berdasarkan kesimpulan telah diuraikan, penelitian tentang peramalan harga saham PT Bank Rakyat Indonesia menggunakan metode *hybrid* VARIMA-GRU menghasilkan beberapa saran untuk penelitian serupa di masa depan, yaitu:

1. Kinerja model *hybrid* VARIMA-GRU dapat dibandingkan dengan model *hybrid* lain, seperti VARIMA-RNN, VARIMA-LSTM, atau model *hybrid* lainnya.
2. Model *hybrid* VARIMA-GRU dapat diaplikasikan pada data saham perusahaan lain di berbagai sektor industri untuk mengetahui efektivitasnya dalam memprediksi harga saham di berbagai konteks.

## DAFTAR PUSTAKA

- Adubisi, O. D., Adubisi, C. E., David, I. J., Eleke, C. C., & Njoku, N. 2019. Multivariate test based on Hotelling's trace with application to crime rates. *Scientific African*, **6**: 1-9. <https://doi.org/10.1016/j.sciaf.2019.e00166>.
- Aji, B. A., dan Surjandari, I. 2020. *Hybrid Vector Autoregression-Recurrent Neural Networks to Forecast Multivariate Time Series Jet Fuel Transaction Price*. *IOP Conference Series: Materials Science and Engineering*. **909**(1): 1-10.
- Alzubaidi, L., Zhang, J., Humaidi, A. J., Dujaili, A. Al, Duan, Y., Shamma, O. Al, Santamaría, J., Fadhel, M. A., Amidie, M. Al, dan Farhan, L. 2021. Review of Deep Learning : Concepts , CNN Architectures, Challenges, Applications, Future Directions. *Journal of Big Data*. **8**(53): 1-74.
- Azhima, S. A. T. Al, Darmawan, D., Hakim, N. F. A., Kustiawan, I., Qibtiya, M. Al, dan Syafei, N. S. 2022. *Hybrid Machine learning Model untuk Memprediksi Penyakit Jantung dengan Metode Logistic Regression Dan Random Forest*. *Jurnal Teknologi Terpadu*. **8**(1): 40–46.
- Beard, E., Marsden, J., Brown, J., Tombor, I., Stapleton, J., Michie, S., dan West, R. 2019. Understanding and using time series analyses in addiction research. *Addiction*. **114**(10): 1866–1884.
- Bousnguar, H., Nadji, L., & Battou, A. 2023. Gated Recurrent units ( GRU ) for Time Series Forecasting in *Higher Education*. *International Journal of Engineering Research & Technology*, **12**(03): 152–154.
- Cavanaugh, J. E., & Neath, A. A. 2019. The Akaike information criterion: Background, derivation, properties, application, interpretation, and refinements. *Wiley Interdisciplinary Reviews: Computational Statistics*: **11**(3), 1–11.
- Chahal, A., dan Gulia, P. 2019. *Machine learning and deep learning*. *International Journal of Innovative Technology and Exploring Engineering*. **8**(12): 4910–4914.
- Chopra, S., dan Meindl, P. 2016. *Supply Chain Management: Strategy, Planning dan Operations 6th edition*. Pearson Education Limited, Boston.

- Cryer, J. D dan Chan, K. 2008. *Time Series analysis with Application R, 2nd Edition*. Springer, New York.
- Dantes, R. 2019. *Wawasan Pasar Modal Syariah*. Wade Group, Jawa Timur.
- Decastro-garcia, N., Luis, Á., Castañeda, M., Escudero-garcia, D., dan Carriegos, M. V. 2019. Effect of the Sampling of a Dataset in the Hyperparameter Optimization Phase over the Efficiency of a Machine Learning Algorithm. *Complexity*. **2019**(1) : 1-16.
- Dewi, K., Adikara, P. P., dan Adinugroho, S. 2018. Prediksi Indeks Harga Konsumen (IHK) Kelompok Perumahan, Air, Listrik, Gas dan Bahan Bakar Menggunakan Metode Support Vector Regression. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*. **2**(10): 3856–3862.
- Dickey, D. A., dan Fuller, W. A. 1979. Distribution of the Estimators for Autoregressive Time Series With a Unit Root. *Journal of the American Statistical Association*. **74**(366): 427-431.
- Enders, W. 1995. *Applied Econometric Time Series*. Iowa State University, New York.
- Feng, J., dan Lu, S. 2019. Performance Analysis of Various Activation Functions in Artificial Neural Networks. *Journal of Physics: Conference Series*. **1237**(2), doi: 10.1088/1742-6596/1237/2/022030.
- Goodfellow, I., Bengio, Y., dan Courville, A. 2016. *Deep learning*. MIT press, Cambridge.
- Hair, J. F., Black, W. C., Babin, B. R., dan Anderson, R. E. 2009. *Multivariate data analysis, 7th edition*. Pearson, London.
- Halim, J. K., Herwindiati, D. E., Hendryli, J. 2022. Penerapan Gated Recurrent Unit untuk Prediksi Zat Pencemar Udara. *Jurnal Ilmu Komputer dan Sistem Informasi*. **10**(2): 1-6.
- Hodson, T. O. 2022. Root-Mean-Square Error (RMSE) or Mean Absolute Error (MAE): When to Use Them or Not. *Geoscientific Model Development*. **15**(14): 5481–5487.
- Hutter, F., Kotthoff, L., dan Vanschoren, J. 2019. *Automated Machine Learning, Methods, Systems, Challenges*. Springer, Switzerland.
- Ismawati, K., Darmawati, Fuad, M., Dewi, Gst. A. K. R. S., Fatchurrohman, M., Ristiyana, R., Umar, A. U. A. Al., Nurriqli, A., Selasi, D., Widaryanti, Rafsanjani, H., Sofyanty, D., Samosir, M. S., Widiniarsih, D. M., dan Abdurohim. 2022. *Pengetahuan Dasar Pasar Modal Dan Investasi*. CV Eureka Media Aksara, Jawa Tengah.

- Jarque, C. M., dan Bera, A. K. 1987. A *Test for Normality of Observations and Regression Residuals*. *International Statistical Review / Revue Internationale de Statistique*. **55**(2): 163–172.
- Jusmawati, Hadijati, M., dan Fitriyani, N. 2020. Penerapan Model *Vector Autoregressive Integrated Moving Average* dalam Peramalan Laju Inflasi dan Suku Bunga di Indonesia. *Eigen Mathematics Journal*. **3**(2): 74-82.
- Kant, L. D. N. 2021. Global Investment Policy Issues During The Covid-19 Pandemic And Policy Employment On Investment. *Proceeding International Conference Faculty of Law*. **1**(1): 12-20.
- Karlik, B., dan Olgac, A. V. 2019. Performance Analysis of Various Activation Functions in Generalized MLP Architectures of Neural Networks. *International Journal of Artificial Intelligence And Expert Systems (IJAE)*. **1**(4): 111–122.
- Koduru, S. K. R. 2022. A Comprehensive Analysis Of Normalization Approaches For Privacy Protection In Data Mining. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*. **8**(5): 144–157.
- Ljung, G. M., dan Box, G. E. P. 1978. On a Measure of a Lack of Fit in Time Series Models. *Biometrika*. **65**(2): 297-303
- Makridakis, S., Wheelwright, S., dan Hyndman, R. J. 1999. *Metode dan Aplikasi Peramalan, jilid 1*. Erlangga, Jakarta.
- Montgomery, D. C., Jennings, C. L., dan Kulahci, M. 2015. *Introduction to time series analysis and forecasting*. John Wiley dan Sons, Canada.
- Munkhdalai, L., Li, M., Theera-Umpon, N., Auephanwiriyaikul, S., dan Ryu, K. H. 2020. VAR-GRU: A *hybrid* model for multivariate financial time series prediction. *Asian Conference on Intelligent Information and Database Systems*: 322-332.
- Noh, S. H. 2021. Analysis of Gradient Vanishing of RNNs and Performance Comparison. *Information*. **12**(11), 442-453.
- Nichols, J. A., Chan, H. W. H., dan Baker, M. A. B. 2018. *Machine learning: applications of artificial intelligence to imaging and diagnosis*. *Biophysical Reviews*. **11**(1): 111–118.
- Prayudani, S., Hizriadi, A., Lase, Y. Y., Fatmi, Y., dan Al-Khowarizmi. 2019. Analysis Accuracy of Forecasting Measurement Technique on Random K-Nearest Neighbor (RKNN) Using MAPE and MSE. *Journal of Physics Conference Series*. **1361**(1): 012089.

- Pomerat, J., Segev, A., dan Datta, R. 2019. On Neural Network Activation Functions and Optimizers in Relation to Polynomial Regression. *2019 IEEE International Conference on Big Data (Big Data)*: 6183–6185.
- Raju, V. N. G., Lakshmi, K. P., Jain, V. M., Kalidindi, A., & Padma, V. 2020. Study the Influence of Normalization/Transformation process on the Accuracy of Supervised Classification. *Proceedings of the 3rd International Conference on Smart Systems and Inventive Technology, ICSSIT 2020*: 729–735. <https://doi.org/10.1109/ICSSIT48917.2020.9214160>.
- Rahman, A. A., Permatasari, N., Yudhanto, W., dan Sijabat, Y. P. 2023. Pengaruh Likuiditas dan Solvabilitas Terhadap Harga Saham Pada Perusahaan Sektor Transportasi dan Logistik Periode 2019-2021. *GEMILANG: Jurnal Manajemen Dan Akuntansi*. **3**(1): 182–196.
- Ridwan, M., Sadik, K., dan Afendi, F. M. 2023. Comparison of ARIMA and GRU Models for *High-Frequency* Time Series Forecasting. *Scientific Journal of Informatics*. **10**(3): 389-400.
- Roza, A., Violita, E. S., dan Aktivani, S. 2022. Article Study of Inflation using Stationary *Test* with Augmented Dickey Fuller dan Phillips-Peron Unit Root *Test* (Case in Bukittinggi City Inflation for 2014-2019). *Eksakta : Berkala Ilmiah Bidang MIPA*. **23**(2): 106–116.
- Saha, S., Singh, N., Mohan, B. R., dan Naik, N. 2021. A Combined Model of ARIMA-GRU to Forecast Stock Price. *Proceedings of the International Conference on Paradigms of Computing, Communication and Data Sciences Algorithms for Intelligent Systems*: 987–998.
- Silva, R. P., Zarpelão, B. B., Cano, A., dan Barbon Junior, S. 2021. Time series segmentation based on stationarity analysis to improve new samples prediction. *Sensors*. **21**(21): 1–22.
- Singh, D., dan Singh, B. 2019. Investigating the impact of data normalization on classification performance. *Applied Soft Computing Journal*. **97**(2020), <https://doi.org/10.1016/J.ASOC.2019.105524>.
- Tasyurek, M. 2022. A Novel Approach to Improve the Performance of the Database Storing Big Data with Time Information. *Balkan Journal of Electrical and Computer Engineering*. **10**(4): 388–396.
- Teng, F., Song, Y., Wang, G., Zhang, P., Wang, L., dan Zhang, Z. 2021. A GRU-Based Method for Predicting Intention of Aerial Targets. *Computational Intelligence and Neuroscience*. **2021**(1): 1-13.
- Tsay, R. S. 2005. *Analysis of Financial Time Series*. John Wiley dan Sons, Canada.

- Wei, W.W.S. 2006. *Time Series Analysis: Univariate and Multivariate Methods, 2nd edition*. Addison Wesley, New York.
- Yang, L., dan Shami, A. 2020. Hyperparameter Optimization of Machine Learning Algorithms : Theory and Practice. *Neurocomputing*: **415**(1), 295–316.
- Zhang, A., Lipton, Z. C., Li, M., dan Smola, A. J. 2022. *Dive into deep learning*. Cambridge University Press, Cambridge.
- Zhang, P. G. 2003. Time Series Forecasting Using a *Hybrid* ARIMA and Neural Network Model. *Neurocomputing*. **50**(17): 159–175.