

**IMPLEMENTASI MODEL *HYBRID VECTOR AUTOREGRESSIVE -
LONG SHORT TERM MEMORY (VAR-LSTM)* PADA PERAMALAN
INDEKS HARGA SAHAM GABUNGAN DAN KURS RUPIAH**

(Skripsi)

Oleh

A. GILANG ALEYUSATA SAVADA

NPM 2015061017



**FAKULTAS TEKNIK
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG**

2024

**IMPLEMENTASI MODEL *HYBRID VECTOR AUTOREGRESSIVE -
LONG SHORT TERM MEMORY (VAR-LSTM)* PADA PERAMALAN
INDEKS HARGA SAHAM GABUNGAN DAN KURS RUPIAH**

Oleh

A. GILANG ALEYUSATA SAVADA

Skripsi

Sebagai Salah Satu Syarat untuk Mencapai Gelar Sarjana Teknik

Pada

**Program Studi S1 Teknik Informatika
Fakultas Teknik Universitas Lampung**



**FAKULTAS TEKNIK
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG**

2024

ABSTRAK

IMPLEMENTASI MODEL *HYBRID VECTOR AUTOREGRESSIVE - LONG SHORT TERM MEMORY* (VAR-LSTM) PADA PERAMALAN INDEKS HARGA SAHAM GABUNGAN DAN KURS RUPIAH

Oleh

A. Gilang Aleyusta Savada

Fluktuasi signifikan pada harga saham dan nilai tukar Rupiah menimbulkan ketidakpastian bagi investor dalam pengambilan keputusan investasi. Selanjutnya, model peramalan harga tradisional seperti *Vector Autoregressive* (VAR) hanya mampu menangkap pola linier, sementara model *deep learning* seperti *Long Short Term Memory* (LSTM) lebih efektif dalam menangani pola nonlinier namun memerlukan data yang besar untuk performa optimal. Oleh karena itu, penelitian ini bertujuan untuk mengembangkan model peramalan *hybrid* dengan mengintegrasikan pendekatan VAR dan LSTM guna memprediksi Indeks Harga Saham Gabungan (IHSG) dan Kurs Rupiah terhadap Dollar Amerika (USD/IDR). Model *hybrid* ini memanfaatkan keunggulan VAR dalam menganalisis pola linier dan LSTM dalam menangkap pola nonlinier pada data deret waktu. Penelitian ini menggunakan kerangka kerja OSEMN (*Obtain, Scrub, Explore, Model, iNterpret*) untuk memastikan proses analisis data dilakukan secara sistematis dan komprehensif. Data historis dari Januari 2004 hingga Desember 2023 digunakan untuk membangun model, sementara data dari Januari hingga Juli 2024 digunakan untuk validasi. Evaluasi kinerja model dilakukan menggunakan *Mean Absolute Error* (MAE) untuk mengukur tingkat kesalahan prediksi serta visualisasi pola pasar yang kompleks. Hasil penelitian mengungkapkan bahwa model *Hybrid* VAR-LSTM berhasil meningkatkan akurasi prediksi pada data validasi dibandingkan dengan penggunaan model VAR atau LSTM secara individu. Peningkatan ini terlihat dari nilai MAE yang lebih rendah pada model *Hybrid* VAR-LSTM saat melakukan peramalan terhadap data validasi variabel IHSG dan Kurs Rupiah.

Kata Kunci: *Hybrid* VAR-LSTM, Peramalan Deret Waktu, Indeks Harga Saham Gabungan, Kurs Rupiah, Kerangka Kerja OSEMN.

ABSTRACT

IMPLEMENTATION OF HYBRID VECTOR AUTOREGRESSIVE - LONG SHORT TERM MEMORY (VAR-LSTM) MODEL IN FORECASTING THE COMPOSITE STOCK PRICE INDEX AND RUPIAH EXCHANGE RATE

By

A. Gilang Aleyusta Savada

Significant fluctuations in stock prices and the Rupiah exchange rate create uncertainty for investors in making investment decisions. Furthermore, traditional forecasting models like Vector Autoregressive (VAR) are only capable of capturing linear patterns, whereas deep learning models such as Long Short Term Memory (LSTM) are more effective in handling non-linear patterns but require large datasets for optimal performance. Therefore, this study aims to develop a hybrid forecasting model by integrating the VAR and LSTM approaches to predict the Composite Stock Price Index (IHSG) and Rupiah exchange rate against the US Dollar (USD/IDR). The hybrid model leverages the strength of VAR in analyzing linear patterns and LSTM in capturing non-linear patterns in time series data. This study employs the OSEM framework (Obtain, Scrub, Explore, Model, iNterpret) to ensure that the data analysis process is conducted systematically and comprehensively. Historical data from January 2004 to December 2023 is used to build the model, while data from January to July 2024 is used for validation. Model performance evaluation is conducted using Mean Absolute Error (MAE) to measure prediction error rates and visualize complex market patterns. The results reveal that the Hybrid VAR-LSTM model successfully improves prediction accuracy on the validation data compared to using either the VAR or LSTM models individually. This improvement is evident from the lower MAE values achieved by the Hybrid VAR-LSTM model when forecasting validation data for the IHSG and Rupiah exchange rate variables.

Keywords: *Hybrid VAR-LSTM, Time Series Forecasting, Composite Stock Price Index, Rupiah Exchange Rate, OSEM Framework.*

Judul Skripsi : **IMPLEMENTASI MODEL *HYBRID VECTOR AUTOREGRESSIVE - LONG SHORT TERM MEMORY (VAR-LSTM)* PADA PERAMALAN INDEKS HARGA SAHAM GABUNGAN DAN KURS RUPIAH**

Nama Mahasiswa : **A. Gilang Aleyusta Savada**

Nomor Pokok Mahasiswa : 2015061017

Program Studi : Teknik Informatika

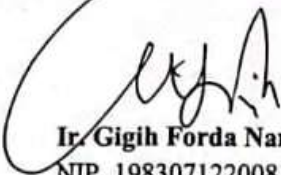
Jurusan : Teknik Elektro


Fakultas : Teknik



Pembimbing Utama

Pembimbing Pendamping

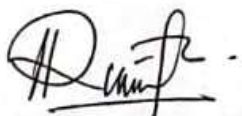

Ir. Gigih Forda Nama, S.T., M.T.I, I.P.M.
NIP. 19830712200812100



Ir. Titin Yulianti, S.T., M.Eng.
NIP. 198807092109032015

2. Mengetahui

Ketua Jurusan Teknik Elektro

Ketua Program Studi
Teknik Informatika

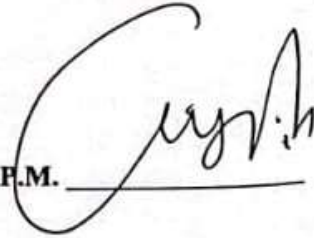

Herlinawati, S.T., M.T.
NIP 197103141999032001


Yessi Mulyani, S.T., M.T.
NIP 197312262000122001

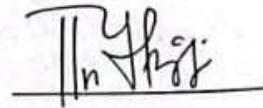
MENGESAHKAN

1. Tim Penguji

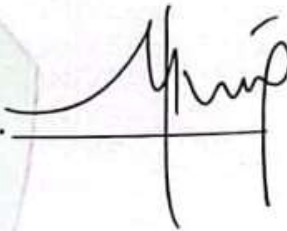
Ketua : Ir. Gigih Forda Nama, S.T., M.T.I., I.P.M.



Sekretaris : Ir. Titin Yulianti, S.T., M.Eng.



Penguji : Dr. Eng. Ir. Mardiana, S.T., M.T., I.P.M.



2. Dekan Fakultas Teknik



Dr. Eng. Ir. Helmy Fitriawan, S.T., M.Sc. }
NIP. 197509282001121002

Tanggal Lulus Ujian Skripsi: 20 September 2024

SURAT PERNYATAAN

Saya yang bertanda tangan dibawah ini :

Nama : A. Gilang Aleyusata Savada

NPM : 2015061017

Dengan ini saya menyatakan bahwa skripsi saya dengan judul “Implementasi Model *Hybrid Vector Autoregressive - Long Short Term Memory* (VAR-LSTM) pada Peramalan Indeks Harga Saham Gabungan dan Kurs Rupiah” dibuat oleh saya sendiri. Apabila di kemudian hari terbukti bahwa skripsi ini merupakan Salinan atau dibuat oleh orang lain, maka saya bersedia menerima sanksi sesuai dengan ketentuan hukum yang berlaku.

Bandar Lampung, 24 September 2024

Penulis,



A Gilang Aleyusta Savada

NPM. 2015061017

RIWAYAT HIDUP



Penulis lahir di Liwa pada tanggal 24 September 2002, sebagai anak pertama dari tiga bersaudara dari pasangan Bapak Yusri dan Ibu Rita Yustina. Penulis menyelesaikan pendidikan formalnya di SDN 03 Liwa (2014), SMPN Sekuting Terpadu (2017) dan SMAN 1 Liwa (2020). Pada tahun 2020 penulis terdaftar sebagai mahasiswa Program Studi Teknik Informatika, Jurusan Teknik Elektro, Fakultas Teknik Universitas Lampung melalui jalur undangan atau SNMPTN (Seleksi Nasional Masuk Perguruan Tinggi Negeri).

Selama berkuliah di Universitas Lampung, penulis terlibat aktif dalam berbagai kegiatan organisasi kemahasiswaan. Pada tahun 2021 dan 2022 sebagai pengurus Himpunan Mahasiswa Teknik Elektro Universitas Lampung. Pada tahun 2021 dan 2023 sebagai pengurus Badan Eksekutif Mahasiswa (BEM) Fakultas Teknik Universitas Lampung. Pada tahun 2022 dan 2023 sebagai pengurus Generasi Baru Indonesia (GenBI) Universitas Lampung.

Selain mengikuti berbagai kegiatan organisasi kemahasiswaan, penulis juga terlibat aktif dalam berbagai kegiatan akademis. Pada tahun 2022 dan 2023 sebagai asisten laboratorium Teknik Komputer Universitas Lampung. Pada tahun 2022 dan 2023 sebagai penerima beasiswa Bank Indonesia (BI). Pada tahun 2022 sebagai penerima *Best Capstone Project Awards* dalam program Studi Independen *Machine Learning and Front-End Web Developer* Dicoding X Kampus Merdeka. Pada tahun 2023 Rindang Tigasatu Pratama. Pada tahun 2024 sebagai *Data Analyst* dalam program Kampus Merdeka Bank Indonesia (KMBI) Angkatan VIII oleh Bank Indonesia.

MOTO

“Honesty, Responsible, Discipline”

(Gilang)

PERSEMBAHAN

Puji serta Syukur saya panjatkan kepada Allah Subhanahu Wa Ta'ala Tuhan Yang Maha Kuasa atas segala rahmat dan karunianya-Nya. Sholawat serta salam saya panjatkan kepada Nabi Muhammad Shallallahu Alaihi Wa Sallam yang selalu menjadi teladan umat manusia.

Kupersembahkan karya ini kepada:

Kedua Orang Tua dan Keluarga

“Yang tiada henti senantiasa memberikan dukungan serta doanya. Semoga saya diberikan kekuatan agar selalu dapat memberikan kebahagiaan bagi kedua orang tua saya. Serta dapat menjadi Abang yang lebih bijaksana dalam memberi teladan baik bagi Adik-adiknya.”

SANWACANA

Penulis panjatkan puji dan syukur kehadirat Allah SWT, Tuhan Yang Maha Esa, karena atas berkat dan hidayah-Nya, penulis dapat menyelesaikan skripsi yang berjudul “Implementasi Model *Hybrid Vector Autoregressive - Long Short Term Memory* (VAR-LSTM) pada Peramalan Indeks Harga Saham Gabungan dan Kurs Rupiah” sebagai salah satu syarat untuk memperoleh gelar sarjana pada Program Studi Teknik Informatika, Universitas Lampung. Dalam menyelesaikan penelitian ini, penulis mendapatkan banyak bantuan dari berbagai pihak. Oleh karena itu, dengan penuh rasa hormat penulis ingin mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Kedua orang tua dan keluarga yang senantiasa selalu memberikan dukungan dan doanya.
2. Bapak Dr. Eng. Helmy Fitriawan, S.T., M.Sc., selaku Dekan Fakultas Teknik Universitas Lampung.
3. Ibu Herlinawati, S.T., M.T., selaku Ketua Jurusan Teknik Elektro Universitas Lampung.
4. Ibu Yessi Mulyani, S.T., M.T. selaku Ketua Program Studi Teknik Informatika Universitas Lampung sekaligus dosen pembimbing akademik penulis.
5. Bapak Ir. Gigih Forda Nama, S.T., M.T.I, I.P.M. selaku Pembimbing Utama yang telah berkenan memberikan waktu serta pikiran dalam membimbing penulis selama mengerjakan penelitian dan penulisan skripsi.
6. Ibu Ir. Titin Yulianti, S. T., M.Eng. selaku Pembimbing Pendamping yang telah berkenan membimbing serta dengan sabar selalu memberikan masukan dan saran bagi penulis dalam menyelesaikan penelitian dan penulisan skripsi.
7. Ibu Dr. Eng. Ir. Mardiana, S.T., M.T., I.P.M. selaku Penguji yang telah memberikan kritik dan saran yang membangun untuk penulis dapat menyelesaikan skripsi dengan baik.

8. Kepada seluruh Dosen Program Studi Teknik Informatika Universitas Lampung yang telah banyak memberikan ilmu kepada penulis.
9. Kepada rekan-rekan seperjuangan yang telah menjadi teman dalam suka dan duka selama masa perkuliahan.
10. Serta seluruh pihak yang tidak dapat penulis sebutkan satu per satu.

Penulis menyadari bahwa skripsi ini masih jauh dari kata sempurna dan masih banyak kekurangan dalam penelitian ini, sehingga penulis sangat terbuka terhadap kritik dan saran yang membangun. Penulis memohon maaf atas seluruh kekurangan yang ada dan mengucapkan terima kasih atas seluruh bantuan yang telah diberikan. Penulis berharap skripsi ini dapat menjadi bermanfaat bagi peneliti lainnya atau bagi siapapun yang membacanya.

Bandar Lampung, 24 September 2024

Penulis,



A Gilang Aleyusta Savada

NPM. 2015061017

DAFTAR ISI

DAFTAR ISI	xii
DAFTAR GAMBAR	xv
DAFTAR TABEL	xvii
I. PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Tujuan Penelitian.....	4
1.4 Manfaat Penelitian.....	5
1.5 Batasan Masalah.....	5
1.6 Sistematika Penulisan.....	6
II. TINJAUAN PUSTAKA	7
2.1 Indeks Harga Saham Gabungan (IHSG)	7
2.2 Nilai Tukar (Kurs).....	7
2.3 <i>Artificial Intelligence</i> (AI).....	8
2.4 <i>Deep Learning</i>	8
2.5 Peramalan (<i>Forecasting</i>)	9
2.6 Deret Waktu (<i>Time Series</i>)	9
2.7 <i>Vector Autoregressive</i> (VAR)	10
2.8 <i>Long Short Term Memory</i> (LSTM)	11
2.9 <i>Hybrid VAR-LSTM</i>	12
2.10 Evaluasi Model.....	14
2.10.1 <i>Akaike's Information Criterion</i> (AIC)	14
2.10.2 <i>Mean Absolute Error</i> (MAE).....	14
2.11 <i>OSEMN Framework</i>	16
2.11.1 <i>Obtain</i> (Mendapatkan).....	17
2.11.2 <i>Scrub</i> (Membersihkan)	17

2.11.3 <i>Explore</i> (Eksplorasi)	17
2.11.4 <i>Modeling</i> (Model)	17
2.11.5 <i>Interpret</i> (Interpretasi)	18
2.12 Perangkat Lunak (<i>Software</i>)	14
2.12.1 Python	14
2.12.2 Microsoft Excel.....	15
2.12.3 Google Colab	15
2.13 Penelitian Terkait	18
III. METODOLOGI PENELITIAN.....	27
3.1 Waktu dan Tempat	27
3.2 Alat dan Bahan	27
3.2.1 Alat.....	27
3.2.3 Bahan	29
3.3 Tahapan Penelitian	30
3.3.1 <i>Obtain</i>	31
3.3.2 <i>Scrub</i>	31
3.3.3 <i>Explore</i>	32
3.3.4 <i>Modeling</i>	33
3.3.5 <i>Interpret</i>	33
IV. PEMBAHASAN.....	35
4.1 <i>Obtain</i>	35
4.2 <i>Scrub</i>	36
4.2.1 <i>Dataset</i> IHSG.....	37
4.2.2 <i>Dataset</i> Kurs Dollar (USD/ IDR)	38
4.2.3 <i>Validasi</i> Data.....	40
4.2.4 <i>Merge</i> <i>Data</i>	43
4.3 <i>Explore</i>	45
4.3.1 <i>Plot</i> <i>Time</i> <i>Series</i>	45
4.3.2 <i>Pemeriksaan</i> <i>Korelasi</i>	50
4.3.3 <i>Uji</i> <i>Kointegrasi</i> <i>Johansen</i>	53
4.3.4 <i>Augmented</i> <i>Dickey-Fuller</i>	54
4.3.5 <i>Granger</i> <i>Causality</i> <i>Test</i>	56

4.4 Model.....	57
4.4.1 Vector Autoregressive (VAR) – Linear	58
4.4.2 Long Short Term Memory (LSTM) – Nonlinear/ Residual	64
4.4.3 Hybrid VAR-LSTM.....	75
4.5 Interpret.....	78
4.5.1 Evaluasi Model	78
4.5.2 Perbandingan Model.....	82
V. KESIMPULAN DAN SARAN.....	86
5.1 Kesimpulan.....	86
5.2 Saran.....	86
DAFTAR PUSTAKA	88
LAMPIRAN.....	Error! Bookmark not defined.

DAFTAR GAMBAR

Gambar 2. 1 Konsep Kecerdasan Buatan	8
Gambar 2. 2 Arsitektur LSTM	11
Gambar 2. 3 Tahapan Kerangka Kerja OSEMN	16
Gambar 3. 1 Tahapan Penelitian	30
Gambar 4. 1 Data IHSG dari Portal Yahoo Finance.....	35
Gambar 4. 2 Data Kurs Dollar (USD/ IDR) dari Portal Bank Indonesia.....	36
Gambar 4. 3 <i>Plot Time Series Data Close: Daily VS Monthly</i>	41
Gambar 4. 4 <i>Plot Time Series Data Kurs Dollar: Daily VS Monthly</i>	42
Gambar 4. 5 Kode Python untuk <i>Merge Data Close</i> IHSG dan Kurs Dollar	43
Gambar 4. 6 Contoh Kode Python untuk Visualisasi Grafik dengan Matplotlib .	45
Gambar 4. 7 <i>Plot Time Series Data Close</i> IHSG	46
Gambar 4. 8 <i>Plot Time Series Data</i> Kurs Dollar (USD/ IDR).....	48
Gambar 4. 9 <i>Plot Time Series Trend Close</i> IHSG dan Kurs Dollar (USD/IDR)..	49
Gambar 4. 10 Matrik Korelasi antara <i>Close</i> IHSG dan Kurs Dollar (USD/ IDR .	51
Gambar 4. 11 <i>Scatter Plot Correlation Close</i> IHSG dan Kurs Dollar (USD/ IDR)	52
Gambar 4. 12 Data Setelah <i>Differencing</i> (Stasioner).....	56
Gambar 4. 13 Kode Python untuk Penentuan <i>Lag Order Optimum</i> Model VAR	58
Gambar 4. 14 Kode Python untuk <i>Fitted Model</i> VAR.....	60
Gambar 4. 15 Perbandingan Data Aktual VS Prediksi VAR pada Historis <i>Close</i>	60
Gambar 4. 16 Perbandingan Data Aktual VS Prediksi VAR pada Historis Kurs Dollar.....	61
Gambar 4. 17 Kode Python untuk Mendapatkan Nilai Residual VAR.....	62
Gambar 4. 18 <i>Plot</i> Residual VAR pada <i>Close</i> IHSG.....	63
Gambar 4. 19 <i>Plot</i> Residual VAR pada Kurs Dollar (USD/ IDR).....	63

Gambar 4. 20 Kode Python untuk <i>Scaling</i> dan <i>Splitting Data</i>	65
Gambar 4. 21 Kode Python untuk Definisi Arsitektur Model LSTM.....	67
Gambar 4. 22 Kode Python untuk Melatih Model LSTM	69
Gambar 4. 23 Kurva <i>Loss</i> pada Pelatihan LSTM untuk Residual <i>Close</i>	70
Gambar 4. 24 Hasil Pelatihan Model LSTM Data <i>Training</i> Residual <i>Close</i>	71
Gambar 4. 25 Hasil Pelatihan Model LSTM Data <i>Testing</i> Residual <i>Close</i>	72
Gambar 4. 26 Kurva <i>Loss</i> pada Pelatihan LSTM untuk Residual Kurs Dollar	73
Gambar 4. 27 Hasil Pelatihan Model LSTM Data <i>Training</i> Residual Kurs Dollar	74
Gambar 4. 28 Hasil Pelatihan Model LSTM Data <i>Testing</i> Residual Kurs Dollar	75
Gambar 4. 29 Kode Python untuk <i>Forecasting 7 Steps Ahead</i> pada Model VAR	76
Gambar 4. 30 Kode Python untuk <i>Forecasting 7 Steps Ahead</i> pada Model LSTM	77
Gambar 4. 31 <i>Plot Forecasting Hybrid VAR-LSTM</i> pada Variabel <i>Close</i>	80
Gambar 4. 32 <i>Plot Forecasting Hybrid VAR-LSTM</i> pada Variabel Kurs Dollar	81
Gambar 4. 33 <i>Plot Perbandingan 3 Model Forecasting</i> pada Variabel <i>Close</i>	83
Gambar 4. 34 <i>Plot Perbandingan 3 Model Forecasting</i> pada Variabel Kurs Dollar	84

DAFTAR TABEL

Tabel 2. 1 Peramalan LSTM Berbagai Horison.....	19
Tabel 2. 2 Peramalan LSTM Berbagai Iterasi.....	19
Tabel 2. 3 Penelitian Terkait	22
Tabel 2. 4 Perbandingan Berbagai <i>Framework Data Mining</i>	25
Tabel 3. 1 Jadwal dan Kegiatan Penelitian	27
Tabel 3. 2 Perangkat Lunak	28
Tabel 3. 3 Kolom Pada Data IHSB.....	29
Tabel 3. 4 Kolom Pada Data Kurs	29
Tabel 4. 1 <i>Head Dataset</i> IHSB (Sumber: Yahoo Finance).....	37
Tabel 4. 2 Statistik Deskriptif <i>Close</i> IHSB.....	37
Tabel 4. 3 <i>Dataset Akhir Close</i> IHSB	38
Tabel 4. 4 <i>Head Dataset</i> Kurs Dollar (Sumber: Bank Indonesia)	39
Tabel 4. 5 Statistik Deskriptif Kurs Dollar (USD/ IDR).....	39
Tabel 4. 6 <i>Dataset Akhir</i> Kurs Dollar	40
Tabel 4. 7 Hasil Uji Korelasi Antara Data Harian dan Bulanan	42
Tabel 4. 8 <i>Dataset Final</i> sebagai <i>Data In Sample Model</i>	44
Tabel 4. 9 <i>Dataset Update</i> sebagai <i>Data Out Sample Model</i>	45
Tabel 4. 10 Hasil Uji Kointegrasi Johansen.....	53
Tabel 4. 11 Hasil Uji <i>Augmented Dickey-Fuller</i> (ADF).....	55
Tabel 4. 12 Hasil <i>Granger Causality Test</i>	56
Tabel 4. 13 Hasil Perhitungan <i>Lag Order Optimum</i> dengan AIC	59
Tabel 4. 14 Data Residual VAR sebagai <i>Data Input</i> LSTM.....	64
Tabel 4. 15 Skala <i>Splitting Data</i> Model LSTM.....	66
Tabel 4. 16 Hasil <i>Hypertuning Parameters</i>	68
Tabel 4. 17 Ringkasan Arsitektur Model LSTM untuk Residual <i>Close</i> IHSB.....	70

Tabel 4. 18 Pelatihan Residual Kurs Dollar (USD/ IDR)	72
Tabel 4. 19 Hasil <i>Forecasting</i> VAR (7 Steps Ahead)	76
Tabel 4. 20 Hasil <i>Forecasting</i> Residual VAR dengan LSTM (7 Steps Ahead)....	77
Tabel 4. 21 Hasil Akhir <i>Forecasting Hybrid</i> VAR-LSTM (7 Steps Ahead)	78
Tabel 4. 22 Perbandingan Nilai Aktual dan <i>Forecasting Hybrid</i> VAR-LSTM....	79
Tabel 4. 23 Perbandingan Nilai MAE dari 3 Model <i>Forecasting</i>	82

I. PENDAHULUAN

1.1 Latar Belakang

Perkembangan pasar modal di Indonesia menunjukkan pertumbuhan yang signifikan dan konsisten. Berdasarkan data statistik dari Kustodian Sentral Efek Indonesia (KSEI) per Desember 2023, jumlah investor pasar modal Indonesia mencapai 12,13 juta *Single Investor Identification* (SID), mengalami peningkatan sebesar 17,60 persen *year to date* (ytd) dibandingkan dengan 10,31 juta SID per Desember 2022 [1]. Pada prinsipnya, pasar modal sendiri merupakan sarana pembentuk modal serta akumulasi dana jangka panjang yang ditujukan untuk meningkatkan partisipasi masyarakat dalam mendukung pembiayaan pembangunan nasional [2]. Peningkatan jumlah investor yang terjadi di pasar modal Indonesia mencerminkan tingginya minat masyarakat terhadap investasi sehingga sangat berdampak positif terhadap pertumbuhan ekonomi nasional.

Bagi investor, pasar modal berfungsi sebagai alternatif berinvestasi dengan keuntungan dan sejumlah risiko tertentu. Investasi saham dapat menjadi opsi pilihan bagi para investor yang ingin berpartisipasi dalam pertumbuhan perusahaan hingga memperoleh keuntungan dari imbal hasil atau *return* [3]. Di Indonesia, terdapat Indeks Harga Saham Gabungan (IHSG) yang merupakan indikator kinerja pergerakan harga saham dari berbagai sektor di Bursa Efek Indonesia (BEI). Pergerakan IHSG sebagai salah satu pengukuran kesehatan ekonomi negara sekaligus landasan analisis statistik atas kondisi pasar saham sekarang. Kondisi pasar modal Indonesia yang mengalami peningkatan (*bullish*) atau mengalami penurunan (*bearish*) dapat terlihat dari fluktuasi harga yang tercatat pada IHSG [4].

Kemudian, saham sebagai produk perdagangan utama di pasar modal cenderung mengalami fluktuasi dan tidak stabil. Ketidakpastian pasar memaksa para investor untuk memantau dengan cermat informasi dari berbagai faktor yang berpengaruh

terhadap harga saham, termasuk fundamental ekonomi makro berupa kurs atau nilai tukar mata uang [5]. Kurs merupakan perbandingan antara mata uang suatu negara dengan mata uang negara lain yang mencerminkan banyak unit mata uang yang dibutuhkan untuk memperoleh satu unit mata uang lainnya. Perubahan Kurs Rupiah dapat berdampak terhadap keputusan investasi para investor Indonesia. Ketika Kurs Rupiah melemah maka dapat mempengaruhi keputusan investor melakukan penjualan saham karena kemungkinan penurunan keuntungan dari biaya operasional perusahaan yang semakin meningkat. Sebaliknya, jika Kurs Rupiah menguat, maka dapat mendorong investor melakukan pembelian saham karena anggapan kondisi ekonomi sedang cerah [6].

Proyeksi melalui peramalan atau *forecasting* pada pergerakan IHSG dan Kurs Rupiah dapat menjadi panduan kritis bagi investor dalam mengambil keputusan investasi berdasarkan strategi sesuai dengan prediksi pasar. Peramalan sendiri merupakan suatu teknik untuk memperkirakan suatu nilai masa depan dengan memperhatikan data masa lalu maupun data masa kini. Aspek penting dalam peramalan merupakan analisis deret waktu atau *time series* yang melibatkan pengamatan serta pemodelan terhadap pola, *trend*, dan variasi data berurutan pada interval waktu tertentu [7]. Berdasarkan banyaknya variabel yang sedang diobservasi, metode peramalan deret waktu dapat digolongkan menjadi dua jenis, yaitu univariat dan multivariat. Metode peramalan univariat merupakan pendekatan yang hanya memiliki satu variabel pengamatan, sedangkan multivariat memiliki lebih dari satu variabel pengamatan yang saling mempengaruhi [8].

Sebenarnya, terdapat beberapa pendekatan klasik statistika konvensional yang dapat digunakan dalam melakukan peramalan data deret waktu, salah satunya *Vector Autoregressive* (VAR). Model VAR merupakan metode peramalan deret waktu multivariat yang sederhana tanpa harus membedakan variabel endogen (Y) dan variabel eksogen (X). Model VAR mengasumsikan hubungan antara variabel dengan cara yang simultan dan linear, sehingga memiliki kekurangan dalam menangkap kompleksitas hubungan nonlinear [9]. Jika diperhatikan lebih lanjut bahwa riwayat IHSG dan Kurs Rupiah merupakan representasi dari data *multivariate time series* yang berpola campuran, yaitu linear dan nonlinear. Dalam

implementasinya, penggunaan suatu model dalam peramalan perlu disesuaikan pada sifat data dan hubungan antara variabel yang diharapkan [2]. Oleh karena itu, penggunaan model linear saja tidak cukup menggambarkan hubungan campuran dalam peramalan IHSG dan Kurs Rupiah yang berpola kompleks.

Saat ini peramalan tidak hanya mengandalkan metode statistika konvensional, tetapi dapat juga memanfaatkan teknik *deep learning* yang muncul akibat perkembangan signifikan *Artificial Intelligence* (AI), salah satunya yaitu penggunaan metode *Long Short Term Memory* (LSTM). Model LSTM merupakan pengembangan *Recurrent Neural Network* (RNN) yang mampu mengekstraksi informasi jangka panjang sehingga menjadikannya alat yang kuat untuk meramalkan *trend* dan perubahan panjang data *time series* atau *sequential* [10]. Model LSTM dapat mempelajari data dengan komputasi fleksibel melalui peran beberapa gerbang yang mengatur memori pada setiap neuron. Berbeda dengan model VAR, model LSTM mampu memodelkan data berpola nonlinier. Namun, diperlukan sejumlah data besar untuk mempelajari pola data yang kompleks sehingga penggunaan LSTM secara independen pada data terbatas tidak mampu bersamaan menangani pola linier dan nonlinier dengan baik [11].

Melalui beberapa sumber empiris, ditemukan bahwa terdapat suatu pendekatan peramalan deret waktu yang disebut model *hybrid*, yaitu menggabungkan beberapa model berbeda untuk meningkatkan performa dibandingkan model individualnya. Penelitian model *hybrid* telah dilakukan oleh Zhang (2003) mengenai peramalan data bintik matahari, kejernihan air, dan kurs uang menggunakan model *hybrid Autoregressive Integrated Moving Average - Artificial Neural Network* (ARIMA-ANN). Hasilnya bahwa model *hybrid* memiliki akurasi prediksi lebih tinggi dibandingkan ARIMA atau ANN [12]. Penelitian model *hybrid* lainnya dilakukan oleh Caliwag dan Lim (2019) mengenai keamanan baterai *lithium-ion* dan peramalan tegangan listrik pada aplikasi motor listrik menggunakan model *hybrid Vector Autoregressive Moving Average - Long Short Term Memory* (VARMA-LSTM). Hasilnya bahwa model *hybrid* memiliki *Root Mean Square Error* (RMSE) lebih kecil dibandingkan penggunaan VARMA atau LSTM secara mandiri [11].

Pada dasarnya, prediksi deret waktu multivariat data IHSG dan Kurs Rupiah dengan

model VAR dan LSTM dapat dikembangkan menjadi model *Hybrid* VAR-LSTM. Konsep dasar dari kombinasi model dalam peramalan tersebut dengan menggabungkan keunggulan serta saling melengkapi kelemahan masing-masing model untuk menangkap berbagai komponen data. Model VAR digunakan untuk menganalisis dan meramalkan data dalam bentuk komponen linear. Selanjutnya, hasil residual yang diperoleh dari model VAR diintegrasikan kembali sebagai data *input* melatih model LSTM guna mengekstraksi elemen nonlinier dalam analisis. Oleh karena itu, pada penelitian ini akan dilakukan analisis implementasi model *Hybrid* VAR-LSTM pada peramalan harga IHSG dan Kurs Rupiah. Model *hybrid* ini akan dikembangkan berdasarkan kerangka kerja OSEMN (*Obtain, Scrub, Explore, Model, iNterpret*), dengan menggunakan bahasa pemrograman Python di Google Colab. Untuk performa model akan dievaluasi menggunakan matriks *Mean Absolute Error* (MAE) sebagai rata-rata kesalahan prediksi dalam satuan yang sama dengan data asli. Diharapkan bahwa pendekatan *Hybrid* VAR-LSTM yang dikembangkan mampu mengoptimalkan peramalan harga IHSG dan Kurs Rupiah di masa depan.

1.2 Rumusan Masalah

Adapun rumusan masalah dari penelitian ini adalah sebagai berikut:

1. Bagaimana membangun peramalan dengan model *Hybrid* VAR-LSTM pada data IHSG dan Kurs Rupiah?
2. Bagaimana performa model *Hybrid* VAR-LSTM dalam meramalkan data deret waktu multivariat pada data IHSG dan Kurs Rupiah?
3. Bagaimana perbandingan peramalan antara model VAR, LSTM, dan *Hybrid* VAR-LSTM pada data IHSG dan Kurs Rupiah?

1.3 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah sebagai berikut:

1. Membangun model peramalan *hybrid* melalui kombinasi keunggulan pendekatan VAR dan LSTM pada data IHSG dan Kurs Rupiah agar dapat memberikan akurasi prediksi yang lebih akurat.

2. Menganalisis kinerja model *Hybrid* VAR-LSTM dalam meramalkan deret waktu multivariat pada data IHSG dan Kurs Rupiah menggunakan matrik evaluasi *Mean Absolute Error* (MAE).
3. Menganalisis perbandingan peramalan antara model VAR, LSTM, dan *Hybrid* VAR-LSTM pada data IHSG dan Kurs Rupiah.

1.4 Manfaat Penelitian

Adapun manfaat dari penelitian ini adalah sebagai berikut:

1. Dapat memperluas pengetahuan bagi pembaca tentang pengembangan metodologi peramalan multivariat pada dinamika IHSG dan Kurs Rupiah menggunakan model *Hybrid* VAR-LSTM.
2. Dapat membantu investor dalam memahami dinamika IHSG dan Kurs Rupiah, dengan tujuan untuk memperbaiki kinerja portofolio, mengurangi risiko, dan meningkatkan keuntungan dalam konteks investasi.
3. Dapat memberikan dukungan bagi pemangku kepentingan dalam merancang kebijakan yang lebih baik, terutama dalam konteks pasar keuangan Indonesia.

1.5 Batasan Masalah

Penelitian ini membatasi peramalan menggunakan pendekatan model *Hybrid* VAR-LSTM terhadap variabel IHSG dan Kurs Rupiah dalam periode 01 Januari 2004 sampai 31 Juli 2024. Sumber data yang digunakan berupa nilai harian penutupan IHSG dari portal resmi Yahoo Finance dan nilai harian transaksi tengah Kurs Rupiah terhadap Dollar Amerika dari portal resmi Bank Indonesia (BI). Untuk kepentingan analisis, data harian tersebut akan dikonversi menjadi data bulanan dengan menggunakan metode rata-rata (*mean*) guna mencerminkan tren jangka panjang yang lebih stabil. Dari data harian yang tersedia, akan dihasilkan 247 data (setara 247 bulan) yang mewakili ukuran tiap bulannya, dengan berupa 240 data (*in-sample*) dari Januari 2004 hingga Desember 2023 akan digunakan untuk membangun model, sedangkan 7 data terakhir (*out-sample*) dari Januari hingga Juli 2024 akan digunakan sebagai validasi. Penilaian kinerja model hanya terbatas pada

matriks evaluasi *Mean Absolute Error* (MAE) karena relevansinya dalam konteks *forecasting* yang dapat memberikan rata-rata absolut kesalahan prediksi sesuai skala data asli. Hasil penelitian ini akan menampilkan hasil peramalan model *Hybrid VAR-LSTM*, serta melihat perbandingan performanya dengan model VAR dan LSTM secara individu.

1.6 Sistematika Penulisan

- | | |
|---------|---|
| BAB I | <p>PENDAHULUAN</p> <p>Pada bab ini berisi latar belakang, rumusan masalah, tujuan, manfaat, batasan masalah, dan sistematika penyusunan laporan penelitian skripsi.</p> |
| BAB II | <p>TINJAUAN PUSTAKA</p> <p>Pada bab ini berisi definisi konsep, tinjauan literatur, penelitian terdahulu, dan kerangka kerja yang digunakan sebagai pedoman dan acuan dalam pemecahan masalah.</p> |
| BAB III | <p>METODOLOGI PENELITIAN</p> <p>Pada bab ini berisi jadwal dan waktu, alat dan bahan, serta langkah sistematis dari <i>OSEMN framework</i> sebagai tahapan logis yang menjelaskan secara detail proses yang dilakukan selama penelitian.</p> |
| BAB IV | <p>HASIL DAN PEMBAHASAN</p> <p>Pada bab ini akan berisi pembahasan mengenai hasil peramalan dari model <i>Hybrid VAR-LSTM</i> pada data IHSG dan Kurs Rupiah, serta membandingkannya dengan model VAR dan LSTM secara independen.</p> |
| BAB V | <p>PENUTUP</p> <p>Pada bab ini akan berisi kesimpulan berdasarkan hasil pembahasan serta saran dengan harapan dapat meningkatkan pengetahuan dan kemajuan bersama.</p> |

II. TINJAUAN PUSTAKA

2.1 Indeks Harga Saham Gabungan (IHSG)

Indeks Harga Saham Gabungan (IHSG) atau *Jakarta Composite Index* (JSX *Composite*) merupakan nilai kinerja seluruh saham yang tercatat di Bursa Efek Indonesia (BEI) sebagai komponen penghitungan indeks. Nilai IHSG digunakan untuk mengetahui pergerakan harga saham secara umum yang dapat menjadi acuan tentang perkembangan kegiatan di pasar modal. Nilai IHSG dapat berubah setiap hari karena perubahan harga pasar dan tambahan saham baru. Pertambahan jumlah saham berasal dari emisi baru, yaitu masuknya emiten baru yang tercatat atau terjadi tindakan *corporate action* berupa *split*, *right*, waran, dividen saham, saham bonus, dan saham konversi [13]. Adapun persamaan umum menghitung IHSG adalah sebagai berikut:

$$IHSG = \frac{\text{total harga semua saham pada waktu berlaku}}{\text{total harga semua saham pada waktu dasar}} \times 100\% \quad (2.1)$$

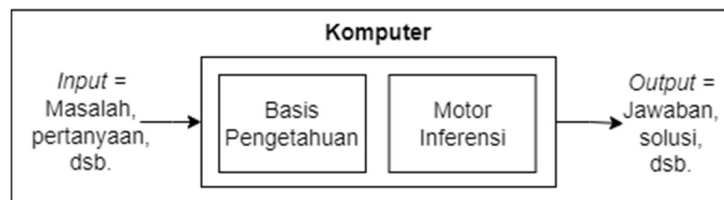
2.2 Nilai Tukar (Kurs)

Kurs atau nilai tukar uang merupakan alat pengukuran yang digunakan untuk menilai harga mata uang domestik terhadap mata uang negara lain. Kurs antara dua negara terjadi melalui harga yang disepakati penduduk dua negara tersebut ketika terjadi elemen krusial dalam konteks ekonomi dan perdagangan. Nilai tukar termasuk kedalam salah satu harga terpenting bagi suatu negara dengan sistem perekonomian terbuka karena berbagai pengaruhnya dalam kegiatan transaksi internasional. Nilai Kurs Dollar (US\$/Rp) berarti harga satu Dollar (US\$) yang ditunjukkan dalam mata uang Rupiah. Nilai tukar berdasarkan kekuatan pasar akan selalu berubah setiap kali harga salah satu dari dua komponen mata uang berubah. Sebuah mata uang akan cenderung menjadi menguat bila permintaan menjadi lebih besar dari ketersediaan (apresiasi) dan melemah bila permintaan kurang dari

ketersediaan (depresiasi). Pergerakan ekstrim nilai tukar yang tidak terkendali dapat menyulitkan pasar, khususnya dalam perilaku ekspor dan impor. Oleh karena itu, pengelolaan nilai mata uang yang stabil menjadi faktor moneter penting dalam perekonomian negara [14].

2.3 Artificial Intelligence (AI)

Artificial Intelligence (AI) atau kecerdasan buatan merupakan cabang ilmu komputer yang mampu menciptakan mesin atau program baru, sehingga dapat melakukan tugas yang biasanya memerlukan kecerdasan manusia. Jika dilihat dari sudut pandang pemrograman, maka kecerdasan buatan meliputi studi mengenai pemrograman simbolik, penyelesaian masalah (*problem solving*), dan pencarian (*searching*) [15].



Gambar 2. 1 Konsep Kecerdasan Buatan [16]

Berdasarkan konsep kecerdasan buatan pada gambar 2. 1, setidaknya terdapat dua bagian yang sangat penting dalam implementasi kecerdasan buatan yaitu:

1. Basis pengetahuan (*knowledge base*) berupa fakta, teori, pemikiran, dan hubungan antara komponen.
2. Motor inferensi (*inference engine*) sebagai proses menarik kesimpulan berdasarkan pengalaman.

2.4 Deep Learning

Deep learning merupakan suatu paradigma dalam pembelajaran mesin atau *machine learning* dengan menekankan penggunaan jaringan syaraf tiruan (*neural network*) yang terdiri dari sejumlah besar lapisan untuk mengekstraksi informasi. Peningkatan signifikan dalam kapasitas komputasi telah *memungkinkan deep*

learning untuk secara otomatis mempelajari representasi fitur yang semakin kompleks dan mengurangi ketergantungan ekstraksi fitur manual. Keunggulan utama *deep learning* terletak pada kemampuannya menangani data yang rumit secara terperinci, seperti kasus pengenalan gambar, pengolahan bahasa alami, dan pembuatan konten baru [17].

2.5 Peramalan (*Forecasting*)

Peramalan atau sering disebut *forecasting* merujuk pada proses analisis dan estimasi nilai masa depan berdasarkan data historis dengan pola yang telah diamati. Dalam konteks bisnis serta ekonomi, peramalan memiliki peran penting bagi perusahaan, organisasi atau pemerintahan dalam pengambilan strategi, perencanaan sumber daya, dan manajemen risiko. Metode peramalan sendiri terdiri atas pendekatan kualitatif, kuantitatif, atau campuran yang disesuaikan dengan karakteristik serta kompleksitas data. Pentingnya peramalan terletak pada kemampuannya memberikan informasi berharga berdasarkan prediksi nilai masa depan, sehingga bisa dijadikan sebagai alat penunjang keputusan yang kuat [18].

2.6 Deret Waktu (*Time Series*)

Deret waktu atau *time series* merupakan representasi data yang diatur secara kronologis untuk memahami suatu perilaku berdasarkan interval waktu tertentu, seperti harian, bulanan, atau tahunan. Deret waktu terbentuk atas kumpulan nilai pengamatan yang terjadi melalui titik waktu berbeda pada selang waktu yang sama, dengan asumsi terdapat korelasi pada variabelnya [19]. Jika sebuah data deret waktu dianalisis lebih terperinci, maka dapat memperlihatkan *trend* atau pola perubahan dari peristiwa tertentu. *Trend* atau pola tersebut akan membentuk gerakan teratur sehingga memungkinkan untuk dilakukan prediksi nilai mendatang berdasarkan nilai masa lalu. Teknik analisis pada kumpulan data deret waktu sangat penting dalam berbagai konteks, termasuk peramalan (*forecasting*) masa depan.

2.7 Vector Autoregressive (VAR)

Model *Vector Autoregressive* (VAR) merupakan pendekatan statistika klasik pada data deret waktu multivariat berdasarkan hubungan beberapa variabel endogen yang cenderung berkaitan menurut lokasi terdekat [20]. Persamaan model VAR(p) secara umum dapat ditulis sebagai berikut:

$$y_t = \alpha + \phi y_{t-1} + \phi y_{t-2} + \dots + \phi y_{t-p} + \varepsilon_t \quad (2.2)$$

dengan:

y_t = Vektor berukuran $m \times 1$ berisi m variabel yang masuk dalam model VAR pada waktu t dan $t - i$, dimana $i = 1, 2, \dots, p$

α = Vektor berukuran $m \times 1$ berisi konstanta

ϕ_i = Matriks parameter variabel endogen berukuran $m \times m$ untuk setiap $i = 1, 2, \dots, p$

ε_t = Vektor galat berukuran $m \times 1$

Model VAR(1) yang terdiri dari 2 variabel dapat ditulis:

$$y_{1,t} = \alpha_1 + \phi_{11}y_{1,t} + \phi_{12}y_{2,t-1} + \varepsilon_{1,t} \quad (2.3)$$

$$y_{2,t} = \alpha_2 + \phi_{21}y_{2,t} + \phi_{22}y_{2,t-1} + \varepsilon_{2,t} \quad (2.4)$$

Persamaan (2.3) dan (2.4) dapat ditulis menjadi:

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} + \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix} \quad (2.5)$$

Dengan:

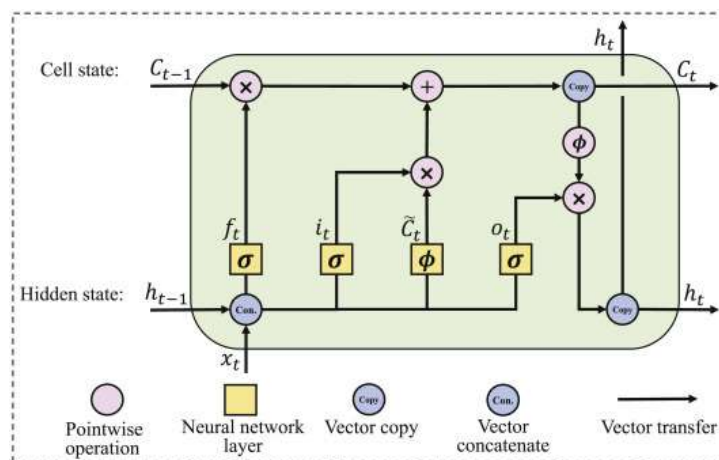
$$y_t = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}, \alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}, \phi = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix}, \varepsilon_t = \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}$$

Sehingga didapatkan model VAR orde p atau VAR(1), yaitu:

$$y_t = \alpha + \phi y_{t-1} + \varepsilon_t \quad (2.6)$$

2.8 Long Short Term Memory (LSTM)

Long Short Term Memory (LSTM) merupakan jaringan saraf yang dirancang untuk memproses data berurutan dengan kemampuan mengingat informasi kompleks dalam jangka waktu panjang. Metode LSTM sangat baik digunakan dalam kasus penyelesaian masalah yang berkaitan dengan klasifikasi, pengolahan, dan prediksi berdasarkan data *time series*. Struktur LSTM terdiri dari empat lapisan yang terhubung, yaitu tiga lapisan gerbang (*forget*, *input*, dan *output*) serta satu lapisan aktivasi *tanh*. Adapun arsitektur dari LSTM dapat dilihat pada Gambar 2.2 berikut.



Gambar 2. 2 Arsitektur LSTM [21].

Forget gate berfungsi untuk mengontrol ketentuan penghapusan atau penambahan informasi dari setiap data masukan yang akan disimpan atau dihapus dalam *memory cell* menggunakan fungsi aktivasi *sigmoid*. Persamaan yang digunakan pada *forget gate* yaitu:

$$f_t = \sigma (W_f \cdot [h_{t-1}, x_t] + b_f) \quad (2.7)$$

dengan:

- f_t = Vektor *output* dari *forget gate* pada waktu t
- σ = Fungsi aktivasi *sigmoid*
- h_{t-1} = *Hidden state* dari waktu sebelumnya
- x_t = Nilai *input* pada waktu t
- b = Nilai bias pada *gate*
- W = Matriks bobot pada *gate*

Input gate berfungsi untuk mengontrol informasi baru yang masuk ke dalam sel

memori menggunakan fungsi aktivasi *sigmoid* dalam menentukan nilai yang diperbarui dan fungsi aktivasi *tanh* dalam menghasilkan nilai baru. Persamaan yang digunakan pada *input gate* yaitu:

$$i_t = \sigma (W_i.[h_{t-1}, x_t]+b_i) \quad (2.8)$$

dengan:

i_t = Vektor *output* dari *input gate* pada waktu t

Cell gate berfungsi untuk mengontrol nilai pada sel memori sebelumnya agar digantikan oleh nilai sel memori baru dengan menggabungkan nilai-nilai dari *forget gate* dan *input gate*. Persamaan yang digunakan pada *cell gate* yaitu:

$$\hat{C}_t = \tanh (W_c.[h_{t-1}, x_t]+b_c) \quad (2.9)$$

dengan:

\hat{C}_t = Vektor kandidat *cell state* baru pada waktu t

tanh = Fungsi aktivasi *tanh*

Output gate berfungsi untuk mengontrol informasi dalam sel memori yang akan dikeluarkan melalui operasi pada lapisan *sigmoid* dan *tanh* sebelumnya. Persamaan yang digunakan pada *output gate* yaitu:

$$O_t = \sigma (W_o.[h_{t-1}, x_t] + b_o) \quad (2.10)$$

$$h_t = O_t \times \tanh (C_t) \quad (2.11)$$

dengan:

O_t = Vektor *output* dari *output gate* pada waktu t

h_t = *Cell state update* pada waktu t

2.9 Hybrid VAR-LSTM

Model *hybrid* merupakan pengembangan dengan konsep dasar kombinasi beberapa metode yang bertujuan untuk memberikan hasil lebih akurat dibanding model individual [22]. Dapat diasumsikan bahwa pada model deret waktu memiliki dua komponen dasar, yaitu linier dan nonlinier dengan persamaan berikut [12]:

$$Y_t = L_t + NL_t + \alpha_t \quad (2.12)$$

dengan:

- Y_t = Data pengamatan deret waktu ke-t
- L_t = Komponen linear ke-t
- NL_t = Komponen nonlinear ke-t
- α_t = *Error* ke-t
- t = Indeks waktu

Pendekatan VAR merupakan model yang cukup baik dalam memodelkan data linear, tetapi kesulitan dalam menangani pola data nonlinear. Sedangkan model LSTM memungkinkan memodelkan data berpola linear dan nonlinier, tetapi diperlukan data yang cukup handal, sehingga penggunaan LSTM dalam data terbatas secara langsung kurang optimal dalam menangani pola campuran sekaligus [11]. Penggunaan model *Hybrid* VAR-LSTM dapat memisahkan penanganan aspek linier dengan VAR dan aspek nonlinear dengan LSTM sehingga model *hybrid* dapat menangkap pola campuran dalam sebuah data dengan baik pada setiap tahapannya. Terdapat tiga tahapan utama dalam proses pembangunan model *Hybrid* VAR-LSTM yaitu:

1. Model VAR dibangun menggunakan data aktual untuk memperkirakan komponen linier dan mendapatkan nilai residual
2. Data residual dari model VAR kemudian digunakan sebagai masukan data pada model LSTM untuk memperkirakan komponen nonlinear
3. Menggabungkan kedua hasil prediksi pada langkah sebelumnya untuk mendapatkan hasil melalui model *Hybrid* VAR-LSTM dengan persamaan:

$$Y_t' = L_t' + N_t' \quad (2.13)$$

dengan:

- Y_t' = Nilai peramalan dari model *hybrid* ke-t
- L_t' = Nilai peramalan dari komponen linear ke-t (VAR)
- N_t' = Nilai peramalan dari komponen nonlinear ke-t (LSTM)
- t = Indeks waktu

2.10 Evaluasi Model

2.10.1 Akaike's Information Criterion (AIC)

Akaike's Information Criterion (AIC) merupakan metode yang digunakan dalam pemodelan statistik untuk membandingkan sejumlah model berbeda berdasarkan kualitas relatifnya. Metode AIC biasanya digunakan dalam konteks pemilihan berbagai model statistik dengan ketentuan nilai AIC paling rendah dianggap sebagai model terbaik. Berikut merupakan persamaan dari perhitungan nilai AIC [20]:

$$AIC(p) = \ln(\det(\Sigma_p)) \frac{2kP}{T} \quad (2.14)$$

dengan:

- k = Jumlah variabel dalam model VAR
- T = Jumlah observasi dalam data
- Σ_p = Matriks kovarians dari residual dengan order lag p
- p = Order lag yang sedang dievaluasi
- $\det(\Sigma_p)$ = Determinan dari matriks kovarians residual dengan lag p

2.11 Perangkat Lunak (*Software*)

2.11.1 Python

Python merupakan bahasa pemrograman tingkat tinggi yang memiliki sintaksis bersih dan mudah dibaca. Pada awalnya Python dikembangkan oleh Guido Van Rossum pada akhir 1980-an dan dirilis pertama kali pada tahun 1991. Kelebihan Python terletak pada sintaksis yang sederhana dan mudah dimengerti. Dukungan komunitas yang kuat juga menjadikan Python sebagai bahasa yang sangat diminati dan sering dipilih untuk berbagai jenis proyek pengembangan perangkat lunak. Saat ini, Python telah menjadi bahasa pemrograman yang mendominasi bidang ilmu data dan kecerdasan buatan. Pustaka-pustaka Python seperti NumPy, Pandas, Matplotlib, TensorFlow, dan PyTorch merupakan alat penting untuk manipulasi data, analisis, visualisasi, serta pengembangan solusi perangkat lunak di berbagai bidang.

2.11.2 Microsoft Excel

Microsoft Excel merupakan aplikasi *spreadsheet* yang dikembangkan oleh Microsoft Corporation. Microsoft Excel memiliki antarmuka pengguna yang mudah digunakan untuk membuat, mengelola, dan menganalisis data dalam bentuk tabel mirip seperti kertas kerja konvensional. Microsoft Excel memiliki fungsi perhitungan dan rumus yang beragam, sehingga memungkinkan pengguna membuat perhitungan matematis, statistik, keuangan, dan logika. Selain itu, Microsoft Excel menawarkan kemampuan visualisasi data melalui grafik dan diagram untuk mempermudah memahami serta menyajikan sebuah data.

2.11.3 Google Colab

Google Colab atau Google Colaboratory merupakan *platform* berbasis *cloud* yang memungkinkan pengguna untuk menulis dan mengeksekusi kode Python secara daring. Google Colab dikembangkan oleh Google Corporation untuk mendukung pengembangan proyek dan analisis data menggunakan Python dalam lingkungan *cloud*. Salah satu keunggulan utama dari Google Colab merupakan akses daringnya yang mudah dan kolaboratif. Dengan mengaksesnya melalui *browser website* memungkinkan pengguna dapat bekerja dari berbagai perangkat dan berkolaborasi secara *real-time* dengan rekan tim. Integrasi dengan Google Drive mempermudah penyimpanan dan akses proyek, sementara dukungan dari berbagai pustaka populer, seperti TensorFlow dan PyTorch membuatnya sangat relevan dalam pengembangan kecerdasan buatan. Pada Google Colab *import* dan *export* data dapat dilakukan dari berbagai sumber, seperti Google Drive, Google Sheets, dan GitHub melalui integrasi berformat *notebook* Jupyter sehingga dapat memfasilitasi kolaborasi pengguna lintas *platform*.

2.11.4 Mean Absolute Error (MAE)

Mean Absolute Error (MAE) merupakan salah satu matriks evaluasi yang dapat digunakan untuk mengukur akurasi dari model yang melakukan prediksi. Besaran MAE menghitung rata-rata dari selisih absolut antara nilai aktual yang terobservasi dengan nilai yang diprediksi oleh model. Dalam analisis *forecasting*, MAE

memberikan gambaran mengenai seberapa besar kesalahan rata-rata yang dihasilkan oleh model dalam satuan yang sama dengan data aktualnya. Hal ini menjadikan MAE mudah diinterpretasikan karena memberikan informasi langsung tentang rata-rata deviasi antara hasil prediksi dan nilai nyata tanpa melihat arah kesalahan (*overestimation* atau *underestimation*). Nilai MAE yang lebih rendah mengindikasikan bahwa model *forecasting* memiliki kemampuan prediksi yang lebih akurat, sehingga dapat diandalkan dalam membuat keputusan berbasis data. Berikut merupakan persamaan dari matriks MAE:

$$MAE = \frac{1}{n} \sum_{i=1}^t |Z_t - \hat{Z}_t| \quad (2.15)$$

dengan:

t	=	Banyak data pengamatan
Z_t	=	Data pengamatan ke-t
\hat{Z}_t	=	Nilai duga dari data pengamatan ke-t

2.12 OSEM N Framework

OSEMN *framework* merupakan kerangka kerja sistematis dalam bidang ilmu data populer untuk mengorganisir dan mengarahkan proses pengolahan data dari awal hingga akhir. Kerangka kerja OSEM N memungkinkan melakukan analisis akurat dan memudahkan dalam kembali ke langkah tertentu dari pemrosesan data mengikuti metodologi tahapan berupa *Obtain, Scrub, Explore, Model, Interpret* [23], [24].



Gambar 2. 3 Tahapan Kerangka Kerja OSEM N [25]

2.12.1 Obtain (Mendapatkan)

Langkah awal dalam kerangka kerja OSEMN yaitu proses pengumpulan data dengan jumlah yang sesuai dan berkualitas. Seorang praktisi atau ilmuwan harus dapat mengumpulkan data yang dibutuhkan selama analisis. Pada proses *obtain* diperlukan keterampilan atau koneksi untuk mendapatkan sebuah data. Adapun sumber data terpercaya bersumber dari *website*, *database*, atau pengembangan survei dapat digunakan pada tahap ini.

2.12.2 Scrub (Membersihkan)

Langkah kedua dalam kerangka kerja OSEMN setelah data diperoleh yaitu membersihkan data dari potensi kesalahan. Proses membersihkan data menjadi sangat esensial untuk memastikan keakuratan dan konsistensi data sebelum dilanjutkan ke proses berikutnya. Pada proses *scrub* akan melibatkan identifikasi dan penanganan nilai yang hilang, *outliers*, atau format tidak sesuai. Adapun teknik seperti imputasi nilai yang hilang, transformasi data, atau penghapusan *outliers* dapat diterapkan pada tahap ini.

2.12.3 Explore (Eksplorasi)

Langkah ketiga dalam kerangka kerja OSEMN setelah data dibersihkan yaitu eksplorasi data lebih lanjut untuk mendapatkan wawasan mendalam tentang karakteristiknya. Eksplorasi membantu dalam merumuskan pertanyaan atau hipotesis yang lebih spesifik untuk diuji diproses berikutnya. Pada tahap *explore* akan dilakukan identifikasi pola, *trend*, dan hubungan dalam data. Adapun teknik seperti visualisasi data, statistik deskriptif, atau bahkan eksplorasi dengan metode analisis yang lebih kompleks dapat diterapkan pada tahap ini.

2.12.4 Modeling (Model)

Langkah keempat dalam kerangka kerja OSEMN setelah pemahaman mendalam terhadap data yaitu mengembangkan model atau algoritma sesuai dengan sifat data dan tujuan analisisnya. Pada tahap *modeling* akan dilakukan pengujian, penilaian, serta penyesuaian untuk memastikan sebuah model memiliki hasil konsisten dan dapat diandalkan. Adapun teknik pendekatan seperti statistik atau kecerdasan

buatan dapat diterapkan pada tahap ini.

2.12.5 Interpret (Interpretasi)

Langkah terakhir dalam kerangka kerja OSEMN setelah mengembangkan model yaitu penerjemahan hasil dari model ke dalam pemahaman yang dapat dimengerti oleh pemangku kepentingan. Pada tahap *interpret* akan dilakukan penarikan kesimpulan untuk mendukung keputusan atau rekomendasi berdasarkan temuan penting dari pola data yang sudah dianalisis. Adapun teknik penyusunan laporan atau presentasi hasil analisis komprehensif dapat diterapkan pada tahap ini.

2.13 Penelitian Terkait

Dalam penelitian ini, beberapa kajian literatur dijadikan referensi dengan tujuan memperluas pengetahuan dari hasil penelitian yang telah dilakukan sebelumnya.

Pengembangan Model

Penelitian mengenai peramalan oleh Hakim, dkk. (2024) menggunakan model statistik *Seasonal Autoregressive Integrated Moving Average* (SARIMA) pada data historis harga telur ayam ras di pasar tradisional Kabupaten Manokwari selama periode enam tahun [26]. Data yang digunakan dalam penelitian ini diresampling dari data harian menjadi data bulanan berdasarkan rata-rata (*mean*) untuk memudahkan analisis pola musiman. Pada penelitian tersebut, model SARIMA(0,1,0)(0,1,1)₁₂ terbukti menjadi alat yang efektif dalam peramalan harga telur ayam ras dengan nilai RMSE sebesar 1491,30 dan MAPE sebesar 3,40%, yang menunjukkan tingkat akurasi yang sangat baik. Penelitian lain mengenai peramalan oleh Khan dan Alghulaiakh (2020) menggunakan model statistik univariat *Auto Regressive Integrated Moving Average* (ARIMA) pada data historis saham Netflix selama periode lima tahun [27]. Pada penelitian tersebut model ARIMA menjadi alat efektif dalam peramalan harga saham Netflix dengan tingkat akurasi mencapai 99,75%. Penelitian lain mengenai peramalan oleh Putri, dkk. (2020) menggunakan model statistika multivariat *Vector Autoregressive* pada data historis harga penutupan saham ICBP dan INDF (endogen) terhadap Kurs Rupiah (eksogen) selama periode 3 tahun [28]. Pada penelitian tersebut model VAR dapat dapat

memprediksi *trend* data saham 30 hari ke depan dengan interval kepercayaan mencapai 95%. Penelitian lain mengenai peramalan oleh Putri dan Setiawan (2015) yang membandingkan model univariat ARIMA dan multivariat VAR pada data historis harga penutupan saham BBRI, BMRI, dan BBCA selama periode 5 tahun [2]. Pada penelitian tersebut model multivariat lebih unggul daripada univariat sebab model VAR yang berhasil meramalkan harga saham BMRI dan BBCA dengan baik, dibandingkan model ARIMA yang hanya berhasil meramalkan harga saham BBRI.

Selain metode statistika tradisional, peramalan juga dapat mengadopsi teknik *deep learning* seperti penggunaan metode populer *Long Short Term Memory* (LSTM). Penelitian mengenai peramalan beban listrik oleh Muzaffar dan Afshari (2019) menggunakan model LSTM pada ketiga variabel eksogen yang direkam setiap jam (24 pembacaan/hari) selama periode 13 bulan [29]. Pada penelitian tersebut model terbaik dari LSTM memiliki potensi untuk meningkatkan akurasi peramalan beban listrik dibandingkan model serupa dengan hasil seperti pada tabel 2. 1 berikut.

Tabel 2. 1 Peramalan LSTM Berbagai Horison

LSTM Forecast	RMSE	MAPE
24 Jam	89,40	1,522
48 Jam	124,3	2,16
7 Hari	374	5,97
30 Hari	554,9	9,75

Penelitian lain mengenai peramalan oleh Moghar dan Hamiche (2020) menggunakan model LSTM pada data historis harga pembukaan saham GOOGL (periode 15 tahun) dan NKE (periode 10 tahun) [30]. Pada penelitian tersebut menunjukkan bahwa saat melakukan pelatihan melalui *dataset* lebih kecil dengan jumlah iterasi (*epochs*) lebih banyak ternyata dapat meningkatkan hasil pengujian sekaligus memungkinkan peningkatan kualitas prediksi dan ramalan. Sebagai perbandingan akurasi pelatihan dan pengujian pada semua iterasi peramalan harga aset NKE dan GOOGL akan diperlihatkan melalui tabel 2. 2 berikut.

Tabel 2. 2 Peramalan LSTM Berbagai Iterasi

Epochs	GOOGL		NKE	
	waktu proses/detik	Loss	waktu proses/detik	Loss
12 <i>epochs</i>	264	0,0011	132	0,0019

25 epochs	550	0,001	275	0,0016
50 epochs	1100	6,57E-04	550	0,001
100 epochs	2200	4.97E-04	1100	8,74E-04

Penelitian lain mengenai peramalan indeks harga penutupan saham S&P 500 oleh Bhandari, dkk. (2022) menggunakan model LSTM pada data pasar fundamental, makro ekonomi dengan indikator teknis selama periode 15 tahun [31]. Pada penelitian tersebut menunjukkan bahwa model LSTM satu lapisan memberikan hasil lebih baik dan akurasi prediksi lebih tinggi daripada model LSTM banyak lapisan. Model terbaik dari LSTM satu lapisan didapat dengan parameter 150 *neurons* yang menghasilkan nilai RMSE sebesar 40,4574, MAPE sebesar 0,7989, dan R sebesar 0,9976.

Selanjutnya, peramalan juga dapat menggunakan model *hybrid* melalui pengembangan dengan konsep dasar menggabungkan beberapa metode dengan tujuan memberikan hasil lebih akurat dibanding model individual. Seperti penelitian mengenai peramalan oleh Zhang (2003) yang menggunakan model *Hybrid* ARIMA-ANN (*Artificial Neural Network*) pada data bintang matahari, kejernihan air, dan kurs uang [12]. Pada penelitian model *hybrid* tersebut, ARIMA digunakan untuk menganalisis bagian linear dan ANN untuk memodelkan bagian nonlinear dari data. Hasil penelitian menunjukkan bahwa penggunaan model *Hybrid* ARIMA-LSTM dapat meningkatkan akurasi peramalan dibandingkan dengan menggunakan ARIMA atau ANN secara terpisah. Penelitian lain mengenai peramalan oleh Dave, dkk. (2021) menggunakan model *Hybrid* ARIMA-LSTM pada data ekspor Indonesia per bulan selama periode 20 tahun [22]. Pada penelitian model *hybrid* tersebut, LSTM diterapkan pada komponen nonlinear dari data sementara ARIMA diterapkan pada komponen linear dari data. Hasil penelitian menunjukkan bahwa model *Hybrid* ARIMA-LSTM memiliki matrik kesalahan paling rendah di antara semua model yang dipakai. Model *hybrid* berhasil mengungguli kinerja model tunggal lainnya dengan nilai MAPE sebesar 7,38% dan RMSE sebesar $1,66 \times 10^{13}$.

Penelitian lainnya mengenai peramalan pengisian daya kendaraan listrik oleh Caliwag dan Lim (2019) menggunakan model *Hybrid* VARMA (*Vector*

Autoregressive Moving Average)-LSTM pada beberapa variabel independen terkait konsumsi baterai motor listrik [11]. Pada penelitian model *Hybrid* VARMA-LSTM diasumsikan dapat menangkap fitur linier serta nonlinier dari data. Hasil penelitian menunjukkan bahwa model *Hybrid* VARMA-LSTM memiliki nilai RMSE yang lebih rendah dibandingkan dengan menggunakan VARMA atau LSTM secara individual ketika meramalkan proses pengisian daya kendaraan listrik. Penelitian lain mengenai peramalan biaya operasional maskapai oleh Aji dan Surjandari (2023) menggunakan model *Hybrid* VAR-LSTM dan *Hybrid* VAR-GRU (*Gated Recurrent Units*) pada tujuh variabel independen terkait biaya operasional maskapai [32]. Pada penelitian model tersebut, VAR akan digunakan sebagai estimasi untuk parameter dan *input* sebelum proses pada dua algoritma selanjutnya yaitu LSTM dan GRU. Hasil analisis menunjukkan bahwa akurasi peramalan masing-masing metode *hybrid* sangat baik, yaitu VAR-LSTM sebesar 98,98% dan VAR-GRU sebesar 99,40%.

Pada dasarnya, kombinasi pendekatan statistik dan pembelajaran mesin pada model *hybrid* sangat mampu mengatasi hubungan kompleks dalam data, seperti dalam kasus prediksi ekonomi atau peramalan keuangan. Hal tersebut sejalan dengan penelitian oleh Sina, dkk. (2023) yang menyajikan tinjauan literatur sistematis tentang metode *hybrid* dengan konsep dasar menggabungkan keuntungan dari pendekatan statistik dan jaringan saraf dalam peramalan deret waktu [33]. Penelitian ini juga membahas kemungkinan aplikasi pendekatan prediksi *hybrid* dalam sistem analitik visual untuk meningkatkan proses pengambilan keputusan. Metode yang digunakan dalam tinjauan literatur sistematis tersebut merupakan *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* (PRISMA) dalam menyelidiki pendekatan peramalan model *hybrid*. Berbagai langkah penyaringan dan kriteria pengecualian diterapkan untuk mempersempit hasil, sehingga dimasukkannya dua puluh satu karya dalam tinjauan. Karya-karya tersebut diperiksa secara rinci dan kualitas hasil prediksi dikumpulkan serta diringkas dalam sebuah tabel. Hasil analisis menunjukkan bahwa semua metode peramalan model *hybrid* yang diselidiki berkinerja lebih baik dan mengarah kepada prediksi yang lebih akurat daripada model secara individual.

Tabel 2. 3 Penelitian Terkait

No	Peneliti	Data	Model	Hasil
1	Hakim, dkk. (2024) [26]	Harga telur ayam ras di Manokwari selama 6 tahun (dilakukan resampling data)	SARIMA	RMSE = 1491,30, MAPE = 3,40%
2	Khan dan Alghulaiakh (2020) [27]	Data historis saham Netflix selama 5 tahun	ARIMA	Akurasi = 99,75%
3	Putri, dkk. (2020) [28]	Harga penutupan saham ICBP, INDF, dan Kurs Rupiah selama 3 tahun	VAR	Prediksi 30 hari ke depan, Interval kepercayaan = 95%
4	Putri dan Setiawan (2015) [2]	Harga penutupan saham BBRI, BMRI, dan BBCA selama 5 tahun	ARIMA/ VAR	VAR (<i>multivariate</i>) lebih unggul dalam meramalkan saham BMRI dan BBCA
5	Muzaffar dan Afshari (2019) [29]	Beban listrik selama 13 bulan dengan 24 pembacaan per hari	LSTM	RMSE 24 jam = 89,40, MAPE 24 jam = 1,522%
6	Moghar dan Hamiche (2020) [30]	Harga pembukaan saham GOOGL (15 tahun) dan NKE (10 tahun)	LSTM	Peningkatan kualitas prediksi dengan iterasi lebih banyak
7	Bhandari, dkk. (2022) [31]	Indeks harga penutupan saham S&P 500 selama 15 tahun	LSTM	RMSE = 40,4574, MAPE = 0,7989, R = 0,9976
8	Zhang (2003) [12]	Data bintang matahari, kejernihan air, dan kurs uang	<i>Hybrid</i> ARIMA-ANN	Peningkatan akurasi peramalan dibanding ARIMA atau ANN individual
9	Dave, dkk. (2021) [22]	Data ekspor Indonesia per bulan selama 20 tahun	<i>Hybrid</i> ARIMA-LSTM	MAPE = 7,38%, RMSE = 1,66 x 10 ¹³

10	Caliwag dan Lim (2019) [11]	Data pengisian daya kendaraan listrik terkait konsumsi baterai motor listrik	<i>Hybrid</i> VARMA-LSTM	RMSE lebih rendah dibandingkan VARMA atau LSTM individual
11	Aji dan Surjandari (2023) [32]	Data biaya operasional maskapai dengan 7 variabel independen	<i>Hybrid</i> VAR-LSTM/ Hybrid VAR-GRU	Akurasi VAR-LSTM = 98,98%, Akurasi VAR-GRU = 99,40%

Tabel 2. 3 merupakan beberapa penelitian yang berkaitan dengan peramalan data menggunakan berbagai pendekatan, seperti statistika, pembelajaran mesin, dan *hybrid*. Namun, belum terdapat penelitian yang secara khusus melakukan peramalan pada kasus Indeks Harga Saham Gabungan (IHSG) dan Kurs Rupiah dengan menggunakan model *Hybrid* VAR-LSTM mengacu pada yang telah diubah dari harian menjadi bulanan melalui proses resampling. Oleh karena itu, akan dilakukan penelitian yang berfokus pada penerapan model *Hybrid* VAR-LSTM dengan harapan dapat memberikan hasil peramalan lebih akurat dibandingkan metode individual, khususnya dalam konteks peramalan IHSG dan Kurs Rupiah. Implementasi model *Hybrid* VAR-LSTM akan menggabungkan kekuatan analisis linier dari VAR dan kemampuan LSTM dalam menangani pola non-linier, sehingga dapat memberikan solusi peramalan yang lebih komprehensif dan tepat guna.

Kerangka Kerja

Selanjutnya, dalam proyek bidang data tentunya diperlukan suatu pendekatan kerangka kerja yang sistematis dan dapat terukur agar mendapatkan hasil yang optimal. Seperti penelitian mengenai penggunaan kerangka kerja proyek bidang data oleh Purnama, dkk. (2012) yang menggunakan metodologi *Cross-Industry Standard Process for Data Mining* (CRISP-DM) dalam menganalisis potensi produksi perikanan [34]. Penelitian tersebut berhasil mengorganisir berbagai tahapan melalui metode CRISP-DM yang terdiri dari enam tahapan, yaitu pemahaman bisnis (*business understanding*), pemahaman data (*data understanding*), pengolahan data (*data preparation*), pemodelan (*modeling*), evaluasi (*evaluation*), dan penyebaran hasil (*deployment*). Pada pemrosesan data dilakukan pemilihan dan pengolahan data yang diperlukan untuk mendukung tahap

pemodelan, termasuk proses klasifikasi hasil produksi yang relevan dengan tujuan penelitian. Evaluasi komprehensif dilakukan untuk memvalidasi hasil pemodelan dengan mengambil sampel data acak dan membandingkannya dengan hasil pemodelan. Hasil penelitian menunjukkan bahwa penggunaan CRISP-DM menghasilkan pendekatan data mining yang sistem dalam pengembangan sistem informasi dengan potensi memberikan wawasan penting bagi pengelolaan sumber daya perikanan.

Penelitian lain mengenai kerangka kerja proyek bidang data oleh Dineva dan Atanasova (2018) dengan menggunakan metodologi OSEMN (*Obtaining, Scrubbing, Exploring, Modeling, Interpreting*) untuk memproses data dari sensor *Internet of Things* (IoT) di sarang lebah [24]. Penelitian tersebut berhasil mengorganisir proses penelitian menjadi sistematis, mulai dari penerimaan data hingga interpretasi hasil. Proses OSEMN melibatkan tahapan pengumpulan data dari sensor di sarang lebah, pembersihan data, eksplorasi data, pemodelan, dan interpretasi hasil untuk memahami perilaku koloni lebah. Hasil penelitian menunjukkan bahwa penggunaan OSEMN berhasil mengidentifikasi korelasi antara berbagai peristiwa dalam sarang lebah sehingga dapat membantu dalam memperoleh pengetahuan lebih mendalam tentang faktor-faktor yang mempengaruhi kesehatan dan produktivitas koloni lebah.

Penelitian lain mengenai kerangka kerja proyek bidang data oleh Suwitno dan Kaunang (2022) dengan menggunakan metodologi SEMMA (*Sample, Explore, Modify, Model, Assess*) untuk mengklasifikasikan data gambar daun buah-buahan [35]. Penelitian tersebut membuktikan penggunaan metodologi SEMMA memudahkan peneliti dalam setiap proses dengan terstruktur dalam. Proses SEMMA diawali dengan mencari data yang valid digunakan, mengeksplorasi ide bersangkutan dengan data yang didapatkan, *pre-processing image* daun menjadi, membuat sebuah model klasifikasi *multi-class* dan evaluasi model untuk proses terakhir. Hasil penelitian menunjukkan bahwa penggunaan SEMMA berhasil mengimplementasikan model CNN (*Convolutional Neural Network*) dalam mengklasifikasikan sembilan kelas daun dengan akurasi mencapai 98%.

Tabel 2. 4 Perbandingan Berbagai *Framework Data Mining* [36], [37]

Nama Proses	OSEMN	CRISP-DM	SEMMA
Jumlah Langkah	5	6	5
<i>Understand</i> (Memahami)	----	<i>Business Understanding</i>	----
<i>Acquire</i> (Mendapatkan)	<i>Obtain</i>	<i>Data Understanding</i>	<i>Sample</i> <i>Explore</i>
<i>Clean</i> (Membersihkan)	<i>Scrub</i>	<i>Data Preparation</i>	<i>Modify</i>
<i>Explore</i> (Eksplorasi)	<i>Explore</i>	----	----
<i>Model</i> (Modelisasi)	<i>Modeling</i>	<i>Modeling</i>	<i>Model</i>
<i>Evaluate</i> (Evaluasi)	<i>Interpret</i>	<i>Evaluation</i>	<i>Assessment</i>
<i>Cumminate</i> (Merangkum)	----	----	----
<i>Deploy</i> (Implementasi)	----	<i>Deployment</i>	----
<i>Monitor</i> (Memantau)	----	----	----

Tabel 2. 4 merupakan perbandingan beberapa framework data mining, yaitu OSEMN, CRISP-DM, dan SEMMA. Berdasarkan analisis ketiga kerangka kerja proyek bidang data, masing-masing metodologi menawarkan pendekatan yang terstruktur dan sistematis. Metodologi SEMMA menonjolkan kesederhanaan dan fokus pada pemrosesan data, dimulai dari pemilihan sampel yang representatif hingga eksplorasi data serta modifikasi sebelum membangun model. Metodologi CRISP-DM melalui enam tahap yang menawarkan pendekatan terstruktur serta modular, dengan penekanan pada evaluasi model dan implementasi solusi dalam praktik bisnis. Kemudian, metodologi OSEMN menawarkan siklus iteratif yang fleksibel sehingga memungkinkan kembali ke tahap sebelumnya jika diperlukan dengan penekanan khusus pada interpretasi hasil analisis atau temuan. Keunggulan masing-masing metodologi terletak pada karakteristiknya serta kemampuannya dalam menangani berbagai proyek. Pada dasarnya, pemilihan metodologi yang tepat harus memperhitungkan faktor-faktor seperti kompleksitas proyek, sumber daya, dan preferensi tim proyek. Oleh karenanya, dengan mempertimbangkan kebutuhan analisis dan kemampuan pengembang, maka pada penelitian mengenai peramalan IHSG dan Kurs Rupiah menggunakan model *Hybrid VAR-LSTM* akan memakai kerangka kerja OSEMN. Kerangka kerja OSEMN dipilih karena menyediakan pendekatan sistematis, fleksibel, dan efisien dalam dalam proyek data berskala menengah, sehingga sangat cocok dengan kebutuhan analisis peramalan IHSG dan Kurs Rupiah. Fokus OSEMN dalam eksplorasi

data memungkinkan identifikasi pemanfaatan informasi penting yang mendukung keakuratan model *Hybrid* VAR-LSTM. Dengan demikian, OSEMN memberikan kerangka kerja ideal untuk menghasilkan peramalan IHSG dan Kurs Rupiah yang dapat diandalkan sesuai tujuan.

III. METODOLOGI PENELITIAN

3.1 Waktu dan Tempat

Penelitian dan pembuatan skripsi ini akan dilakukan pada:

Waktu : Mei 2024 – Agustus 2024.

Tempat : Program Studi Teknik Informatika Universitas Lampung.

Jadwal kegiatan yang akan dilakukan terdapat pada Tabel 3. 1 berikut.

Tabel 3. 1 Jadwal dan Kegiatan Penelitian

Aktivitas	Mei				Juni				Juli				Agustus			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Perumusan Masalah	■	■	■	■												
Studi Pustaka					■	■	■	■								
Persiapan Penelitian									■	■						
<i>Obtain dan Scrub</i>										■	■					
<i>Explore</i>											■	■				
<i>Modeling</i>												■	■			
<i>Interpret</i>													■	■		
Penyusunan Laporan							■	■	■	■	■	■	■	■	■	■

3.2 Alat dan Bahan

3.2.1 Alat

Spesifikasi laptop sebagai perangkat keras yang digunakan selama penelitian adalah sebagai berikut:

Processor : 12th Gen Intel(R) Core (TM) i3-1215U (8 CPUs), ~1.2GHz

Memory : 12288MB RAM

Selanjutnya, terdapat beberapa aplikasi sebagai perangkat lunak pendukung selama penelitian yang ditunjukkan pada Tabel 3. 2 berikut.

Tabel 3. 2 Perangkat Lunak

No	Nama	Versi	Keterangan
1	Python	3.10.12	Bahasa pemrograman Python digunakan untuk membangun algoritma <i>forecasting</i> secara keseluruhan dengan bantuan beberapa pustaka.
	a. Numpy	1.23.5	Pustaka Python yang menyediakan operasi matematika serta <i>array</i> multidimensi untuk manipulasi data numerik dan komputasi selama penelitian.
	b. Matplotlib	3.7.1	Pustaka Python yang digunakan untuk membuat grafik seperti <i>line plots</i> dan <i>scatter plots</i> selama penelitian.
	c. Pandas	1.5.3	Pustaka Python yang digunakan untuk melakukan pembersihan, pengelompokan, dan analisis data selama penelitian.
	d. Statsmodels	0.14.1	Pustaka Python yang digunakan untuk pembangunan model VAR, uji kointegrasi, uji stasioneritas, uji kausalitas, dan perhitungan AIC selama penelitian.
	e. TensorFlow Keras	2.15.0	Pustaka Python yang digunakan untuk pengembangan algoritma pembelajaran mesin, khususnya dalam membangun, melatih, dan mengevaluasi model LSTM selama penelitian.
2	Microsoft Excel	2016	Perangkat lunak pengolahan angka yang akan digunakan pada tahap awal pengumpulan data.
3	Google Colab	0.0.1a2	Kode editor program python yang digunakan selama proses pengembangan <i>forecasting</i> .

3.2.3 Bahan

Bahan pada penelitian ini menggunakan data IHSG dan Kurs Rupiah terhadap Dollar Amerika dalam periode waktu 1 Januari 2004 sampai dengan 31 Juli 2024 dengan rincian yaitu:

1. Data Indeks Harga Saham Gabungan (IHSG)

Data diperoleh dari situs resmi Yahoo Finance (finance.yahoo.com) dengan fokus pada data nilai penutupan Indeks Harga Saham Gabungan (IHSG). Data terdiri atas 5093 baris serta dua kolom yang akan dipakai, yaitu *Date* dan *Close* dengan rincian seperti Tabel 3.3 berikut.

Tabel 3. 3 Kolom Pada Data IHSG

No	Nama Kolom	Keterangan
1	<i>Date</i>	Informasi dengan format <i>Datetime</i> (Tahun-Bulan-Tanggal) yang menyatakan waktu pencatatan harian satu baris data tertentu.
2	<i>Close</i>	Harga penutupan IHSG dalam bentuk IDR (Rupiah Indonesia) berdasarkan waktu sesuai kolom <i>Date</i> .

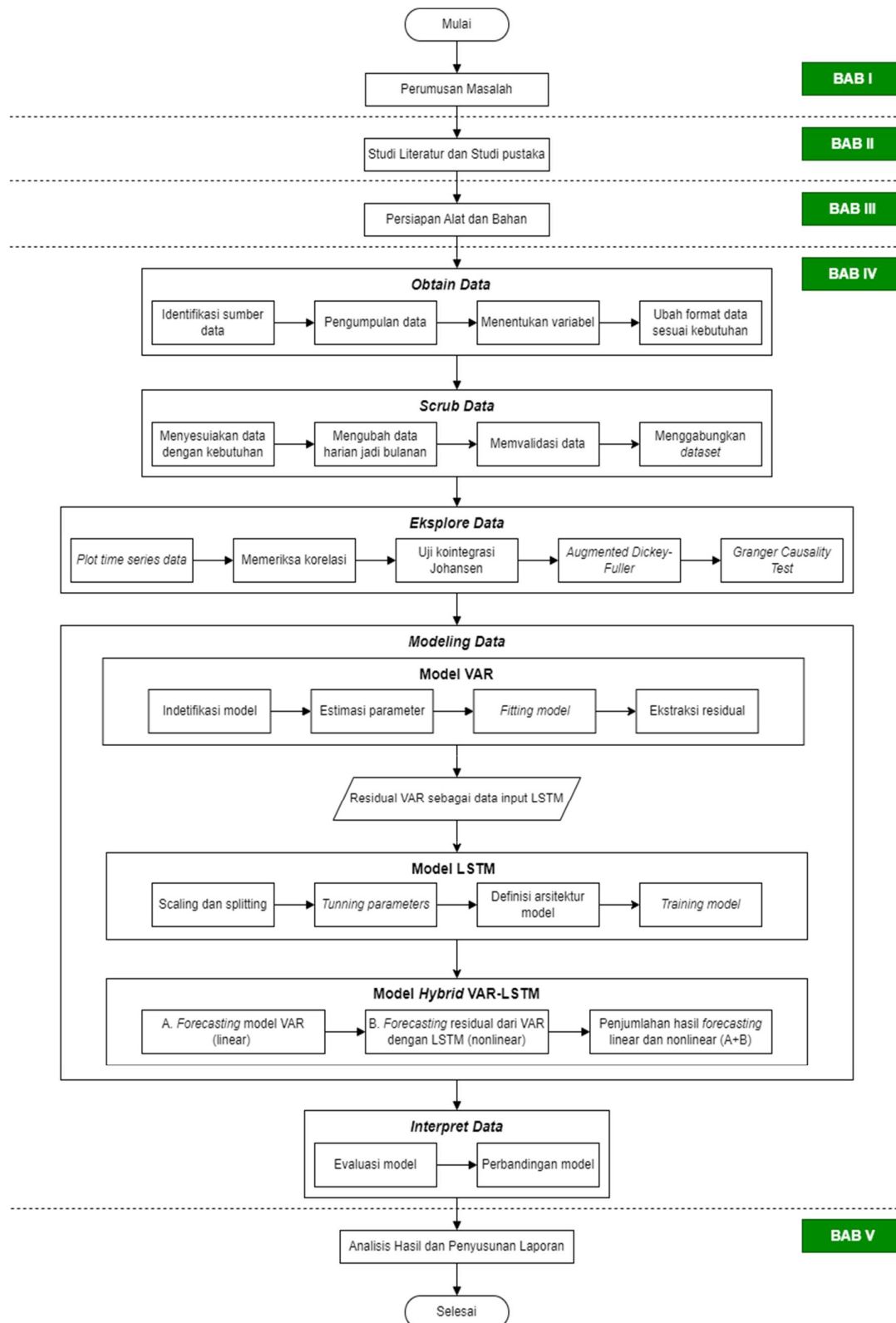
2. Data Kurs Rupiah terhadap Dollar Amerika

Data diperoleh dari situs resmi Bank Indonesia (BI) (bi.go.id) dengan fokus pada data nilai kurs transaksi tengah antara Rupiah dan Dollar Amerika. Data terdiri atas 5038 baris serta dua kolom yang akan dipakai, yaitu *Date* dan Kurs dengan rincian seperti Tabel 3. 4 berikut.

Tabel 3. 4 Kolom Pada Data Kurs

No	Nama Kolom	Keterangan
1	<i>Date</i>	Informasi dengan format <i>Datetime</i> (Tahun-Bulan-Tanggal) yang menyatakan waktu pencatatan harian satu baris data tertentu.
2	Kurs Dollar	Nilai konversi dari 1 USD (Dolar Amerika) menjadi IDR (Rupiah Indonesia) berdasarkan waktu sesuai kolom <i>Date</i> .

3.3 Tahapan Penelitian



Gambar 3. 1 Tahapan Penelitian

Gambar 3. 1 merupakan alur atau tahapan yang akan dilakukan selama penelitian berlangsung. Tahapan tersebut disusun berdasarkan metodologi OSEMN *framework* yang dipakai untuk mendukung pengembangan proyek *data science* secara sistematis dan terstruktur.

3.3.1 *Obtain*

Pada tahap ini dilakukan pengumpulan data sesuai dengan tujuan, yaitu data Indeks Harga Saham Gabungan (IHSG) dari situs resmi Yahoo Finance dan data Kurs Rupiah terhadap Dolar Amerika dari situs resmi Bank Indonesia (BI). Kedua *dataset* berbentuk *file excel* terpisah yang berisi berbagai kolom informasi *time series* selama periode 1 Januari 2004 – 31 Juli 2024.

3.3.2 *Scrub*

Pada tahap ini *dataset* IHSG dan Kurs Rupiah yang sebelumnya telah dikumpulkan di tahap obtain akan diolah menggunakan program Python pada aplikasi Google Colab. Kedua *dataset* tersebut dimasukkan ke dalam satu proyek Google Colab, tetapi dengan bentuk dua *dataset* yang masih terpisah. Kemudian, dilakukan penetapan tipe data serta menghapus *missing value* pada masing-masing variabel *dataset*. Langkah berikutnya merupakan mengubah data yang semula berbentuk perhitungan harian menjadi hitungan bulanan dengan mengambil rata-rata atau total dalam sebulan. Proses perubahan data tersebut merupakan salah satu metode konversi data yang dapat digunakan untuk mengubah level data dalam bentuk yang sama sesuai kebutuhan. Dalam mengonversi data harian menjadi data bulanan, langkah pertama dengan mengelompokkan data berdasarkan bulan di setiap tahun. Setelah itu, dilakukan perhitungan rata-rata nilai harian untuk setiap bulan dengan menjumlahkan semua nilai harian dalam bulan tersebut, lalu membaginya dengan jumlah hari dalam bulan tersebut, hasilnya berupa *mean* yang merepresentasikan data bulanan. Hasil konversi dari data harian menjadi bulanan nantinya akan divalidasi dengan melihat korelasi antara data harian dan bulanan untuk memastikan informasi yang terkandung masih tetap relevan digunakan dalam penelitian. Selanjutnya, *dataset* IHSG dan Kurs Rupiah digabungkan (*merge data*) berdasarkan kolom *Date* pada masing-masing *dataset*. Penggabungan data

dilakukan dari dua *dataset* menjadi satu *dataset* tunggal tanpa mengubah makna data antara kolom nilai dan kolom waktu pada masing-masing *dataset* sebelumnya. Dari penggabungan tersebut akan didapatkan sebuah *dataset* dengan jumlah 247 baris data bulanan dari kolom *Close* IHSG dan Kurs Dollar (USD/ IDR). Adapun 240 baris data pertama (Januari 2004 hingga Desember 2023) dari *dataset* tersebut akan digunakan dalam pembangunan model (*data in sample*). Sedangkan 7 baris data terakhir (Januari 2024 hingga Juli 2024) merupakan data aktual (*data out sample*/ tidak dipakai ketika melatih model) yang akan dibandingkan dengan hasil *forecasting* guna melihat performa model dalam memprediksi data baru yang tidak pernah dilihat sama sekali sebelumnya.

3.3.3 Explore

Pada tahap ini dilakukan analisis data secara deskriptif untuk memahami karakteristik dasar dari *dataset*. Dalam penelitian ini, tahap eksplorasi data melibatkan lima proses, yaitu:

1. *Plot Time Series Data*

Proses ini bertujuan untuk menganalisis perubahan IHSG dan Kurs Rupiah sepanjang periode waktu yang diamati melalui sebuah *plot*. Visualisasi ini dapat membantu dalam identifikasi *trend*, pola musiman, serta fluktuasi dalam kedua variabel tersebut.

2. Pemeriksaan Korelasi

Proses ini bertujuan untuk menganalisis apakah terdapat hubungan antara variabel IHSG dan Kurs Rupiah sepanjang periode waktu yang diamati. Jika korelasi yang ditemukan tinggi, maka menunjukkan adanya hubungan yang kuat antara variabel tersebut.

3. Uji Kointegrasi Johansen

Proses ini bertujuan untuk mengevaluasi apakah terdapat hubungan jangka panjang antara variabel IHSG dan Kurs Rupiah. Hasil dari uji ini membantu dalam memahami hubungan jangka panjang antara kedua variabel tersebut.

4. Uji Stasioneritas *Augmented Dickey-Fuller*

Proses ini bertujuan untuk mengevaluasi apakah terdapat unit akar dalam data yang menunjukkan tingkatan variasi dari properti data. Hasil uji ini membantu

untuk dilakukannya transformasi atau diferensiasi pada data agar menjadi data yang stabil atau stasioner.

5. Uji *Granger Causality*

Proses ini bertujuan untuk mengevaluasi apakah terdapat perubahan dalam variabel Kurs yang berpotensi meramalkan perubahan dalam variabel IHSG atau sebaliknya. Hasil dari uji ini membantu dalam memahami hubungan sebab-akibat antara kedua variabel tersebut.

3.3.4 *Modeling*

Pada tahap ini dilakukan pemodelan algoritma peramalan berdasarkan data yang telah dieksplorasi di tahap sebelumnya. Awalnya dilakukan pembentukan model *Vector Autoregressive* (VAR) dengan menentukan *lag order* yang tepat berdasarkan estimasi parameter model terbaik melalui nilai *Akaike Information Criterion* (AIC) terkecil. Setelah model VAR terbentuk, akan dilakukan *fitting* model dan estimasi residual dari model. Data residual yang didapatkan dari model VAR disiapkan sebagai *input* untuk dua model *Long Short Term Memory* (LSTM) secara terpisah berdasarkan variabel yang sedang diobservasi. Sebelum melakukan pelatihan model LSTM perlu dilakukan *scaling* dan *splitting* atau pengkondisian skala serta pembagian data menjadi *dataset* pelatihan dan pengujian. Proses pembuatan kedua model LSTM dilakukan dengan mencari parameter terbaik melalui proses *tuning parameter*. Kemudian, dilakukan pelatihan model menggunakan algoritma LSTM dengan *input* data berupa residual dari sisa hitung *fitting* model VAR dalam data historis. Seluruh proses tersebut diprogram menggunakan bahasa Python dengan menggunakan pustaka seperti *Statsmodels* untuk VAR serta *Tensorflow Keras* untuk LSTM. Hasil akhir merupakan model *Hybrid VAR-LSTM* yang diperoleh melalui penjumlahan hasil *forecasting* VAR (linear) dengan *forecasting* LSTM (nonlinear) yang sebelumnya telah dibangun.

3.3.5 *Interpret*

Pada tahap ini dilakukan interpretasi hasil pengujian model yang telah dibangun dengan tujuan agar mendapatkan *insight* dari membangun model *Hybrid VAR-LSTM*. Hasil pengujian model berupa visualisasi grafik prediksi dengan aktualnya.

Selain itu, terdapat juga perbandingan performa *forecasting* antara model *hybrid* dengan model dasar secara mandiri, yaitu VAR dan LSTM *multivariate*. Adapun matriks *Mean Absolute Error* (MAE) akan digunakan untuk mengukur kesalahan prediksi yang dilakukan oleh model. Besaran MAE yang semakin kecil menyatakan bahwa model memiliki performa *forecasting* yang lebih baik.

V. KESIMPULAN DAN SARAN

5.1 Kesimpulan

Adapun kesimpulan dari hasil penelitian pada skripsi ini adalah sebagai berikut:

1. Penelitian ini berhasil menerapkan model *Hybrid* VAR-LSTM dengan mengkombinasikan pendekatan linear (VAR) dan nonlinear (LSTM) dalam meramalkan harga *Close* IHSG dan Kurs Dollar (USD/IDR). Model dibangun menggunakan data historis bulanan dari Januari 2004 hingga Desember 2023, serta validasi dilakukan dengan data dari Januari 2004 hingga Juli 2024.
2. Fluktuasi yang lebih luas pada Kurs Dollar (USD/IDR) dibandingkan *Close* IHSG menunjukkan volatilitas pasar yang tinggi, dengan perubahan yang cepat dan tidak terduga. Kondisi ini menyulitkan model untuk menangkap pola yang konsisten, terlihat dari nilai MAE yang lebih besar pada prediksi Kurs Dollar (USD/IDR) dibandingkan prediksi *Close* IHSG, yang lebih stabil.
3. Berdasarkan perbandingan MAE yang dihasilkan pada data validasi, model *Hybrid* VAR-LSTM secara terbukti lebih unggul dibandingkan model VAR atau LSTM yang digunakan secara independen. Kombinasi pendekatan linear dan nonlinear ini mampu menangkap dinamika pasar yang kompleks dengan akurasi prediksi yang lebih tinggi dalam peramalan *Close* IHSG dan Kurs Dollar (USD/IDR).

5.2 Saran

Adapun saran dari hasil penelitian pada skripsi ini adalah sebagai berikut:

1. Model Hybrid VAR-LSTM efektif digunakan untuk *forecasting data time series* dengan kompleksitas tinggi, seperti pada variabel saham dan nilai tukar yang dipengaruhi banyak faktor ekonomi eksternal. Model VAR lebih cocok untuk data yang stabil dan linear, namun kurang optimal untuk data dinamis dengan pola campuran. Sementara, penggunaan LSTM *multivariate* sebaiknya

dioptimalkan atau digabungkan dengan teknik lain untuk meningkatkan akurasi prediksi pada data dengan pola campuran berfluktuasi tinggi.

2. Pada proses analisis (*explore*), ditemukan korelasi dengan hubungan positif yang kuat antara variabel. Namun, hasil uji *granger causality* mengungkapkan bahwa meskipun terdapat korelasi kuat, hanya perubahan pada Kurs Dollar (USD/ IDR) yang dapat digunakan untuk memprediksi perubahan *Close* IHSG. Hasil ini menegaskan pentingnya analisis lebih mendalam untuk memahami hubungan yang sebenarnya terjadi antara variabel, mengingat *forecasting* hanya alat bantu pengambilan keputusan dalam bentuk prediksi secara ilmiah.
3. Untuk aplikasi praktis dari model *forecasting* yang dikembangkan, disarankan untuk melakukan *deploy* model menggunakan *platform* berbasis *cloud*. Selain itu, penting untuk membangun API yang memungkinkan akses *real-time* terhadap hasil prediksi, serta membuat *dashboard* visualisasi yang interaktif untuk membantu pengguna dalam memahami hasil prediksi dengan lebih baik. Pemantauan secara berkala terhadap performa model juga penting dilakukan guna mengantisipasi model *drift* akibat perubahan data di pasar saham dan nilai tukar.

DAFTAR PUSTAKA

- [1] Kustodian Sentral Efek Indonesia (KSEI), “Statistik Pasar Modal Indonesia Desember 2023,” Jakarta, 2023.
- [2] R. N. Putri dan S. Setiawan, “Peramalan Indeks Harga Saham Perusahaan Finansial LQ45 Menggunakan Metode *Autoregressive Integrated Moving Average* (ARIMA) dan *Vector Autoregressive* (VAR),” *Jurnal Sains dan Seni ITS*, vol. 4, no. 2, hlm. 266–271, 2015, doi: 10.12962/j23373520.v4i2.11162.
- [3] N. P. A. K. Almira dan N. L. P. Wiagustini, “*Return on Asset, Return on Equity*, dan *Earning per Share* Berpengaruh terhadap *Return Saham*,” *E-Jurnal Manajemen*, vol. 9, no. 3, hlm. 1069–1088, 2020, doi: 10.24843/EJMUNUD.2020.v09.i03.p13.
- [4] E. Hartayu dan R. A. S. Paramita, “Pengaruh Indikator Makroekonomi, Harga Minyak Dunia Dan Indeks Bursa Asia Terhadap IHSG Periode 2015-2019,” *Sinomika Journal: Publikasi Ilmiah Bidang Ekonomi dan Akuntansi*, vol. 1, no. 5, hlm. 1075–1092, 2023, doi: 10.54443/sinomika.v1i5.547.
- [5] Y. Maurina, R. R. Hidayat, dan S. Sulasmiyati, “Pengaruh Tingkat Inflasi, Kurs Rupiah dan Tingkat Suku Bunga BI Rate Terhadap IHSG,” *Jurnal Administrasi Bisnis SI Universitas Brawijaya*, vol. 27, no. 2, 2015.
- [6] S. Wulandari, S. A. Hutabarat, T. Sihombing, M. Simanjuntak, dan R. Khairani, “Pengaruh Inflasi, BI Rate Dan Nilai Kurs Dollar As Terhadap Indeks Harga Saham Gabungan (IHSG) Yang Terdaftar Di Bursa Efek Indonesia (BEI),” *Journal of Economic, Bussines and Accounting (COSTING)*, vol. 4, no. 2, hlm. 779–786, 2021, doi: 10.31539/costing.v4i2.2049.
- [7] F. Wang, M. Li, Y. Mei, dan W. Li, “*Time Series Data Mining: A Case Study with Big Data Analytics Approach*,” *IEEE Access*, vol. 8, hlm. 14322–14328, 2020, doi: 10.1109/ACCESS.2020.2966553.
- [8] J. L. Chen, G. Li, D. C. Wu, dan S. Shen, “*Forecasting Seasonal Tourism Demand Using a Multiseries Structural Time Series Method*,” *J Travel Res*,

vol. 58, no. 1, hlm. 92–103, 2019, doi: 10.1177/0047287517737191.

- [9] J. H. Stock dan M. W. Watson, “*Vector Autoregressions*,” *Journal of Economic Perspectives*, vol. 15, no. 4, hlm. 101–115, 2001, doi: 10.1257/jep.15.4.101.
- [10] R. Julian dan M. R. Pribadi, “Peramalan Harga Saham Pertambangan pada Bursa Efek Indonesia (BEI) menggunakan *Long Short Term Memory (LSTM)*,” *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, vol. 8, no. 3, hlm. 1570–1580, 2021, doi: 10.35957/jatisi.v8i3.1159.
- [11] A. C. Caliwag dan W. Lim, “*Hybrid VARMA and LSTM Method for Lithium-ion Battery State-of-Charge and Output Voltage Forecasting in Electric Motorcycle Applications*,” *IEEE Access*, vol. 7, hlm. 59680–59689, 2019, doi: 10.1109/ACCESS.2019.2914188.
- [12] G. P. Zhang, “*Time Series Forecasting Using a Hybrid ARIMA and Neural Network Model*,” *Neurocomputing*, vol. 50, hlm. 159–175, 2003, doi: 10.1016/S0925-2312(01)00702-0.
- [13] M. Samsul, *Pasar Modal dan Manajemen Portofolio*. Jakarta: Erlangga, 2006.
- [14] A. Pohan, *Kerangka Kebijakan Moneter dan Implementasinya di Indonesia*. Jakarta: RajaGrafindo Persada, 2008.
- [15] I. M. A. Wirawan, *Metode Penalaran Dalam Kecerdasan Buatan*. Jakarta: RajaGrafindo Persada, 2023.
- [16] S. Kusumadewi, “*Artificial Intelligence (Teknik dan Aplikasinya)*,” *Yogyakarta: Graha Ilmu*, 2003.
- [17] A. Mathew, A. Arul, dan S. Sivakumari, “*Deep Learning Techniques: An Overview*,” 2021, hlm. 599–608. doi: 10.1007/978-981-15-3383-9_54.
- [18] B. Abraham dan J. Ledolter, *Statistical Methods for Forecasting*. New Jersey: John Wiley & Sons, 2009.
- [19] G. E. P. Box dan G. M. Jenkins, *Time Series Analysis Forecasting and Control 5th Edition*, 5 ed. New Jersey: John Wiley & Sons, 2015.
- [20] W. W. S. Wei, *Time Series Analysis: Univariate and Multivariate Methods*, 2 ed. California: Pearson Addison Wesley, 2006.
- [21] X. Luo, D. Zhang, dan X. Zhu, “*Deep Learning Based Forecasting of*

- Photovoltaic Power Generation by Incorporating Domain Knowledge,” Energy*, vol. 225, hlm. 120240, 2021, doi: 10.1016/j.energy.2021.120240.
- [22] E. Dave, A. Leonardo, M. Jeanice, dan N. Hanafiah, “*Forecasting Indonesia Exports Using a Hybrid Model ARIMA-LSTM*,” *Procedia Comput Sci*, vol. 179, hlm. 480–487, 2021, doi: 10.1016/j.procs.2021.01.031.
- [23] H. Mason dan C. Wiggins, “*A Taxonomy of Data Science*.” Diakses: 5 Juni 2024. [Daring]. Tersedia pada: web.archive.org/web/20160220042455/dataists.com/2010/09/a-taxonomy-of-data-science/
- [24] K. Dineva dan T. Atanasova, “*OSEMN Process for Working Over Data Acquired by IoT Devices Mounted in Beehives*,” *Current Trends in Natural Sciences*, vol. 7, no. 13, hlm. 47–53, 2018.
- [25] N. Hotz, “*OSEMN Data Science Life Cycle*,” Data Science Process Alliance. Diakses: 7 Februari 2024. [Daring]. Tersedia pada: <https://www.datascience-pm.com/osemn/>
- [26] I. L. Hakim, M. Sanglise, dan C. D. Suhendra, “*Analisis Peramalan Harga Telur Ayam Ras Dengan Menggunakan Metode SARIMA*,” *Jurnal Media Informatika Budidarma*, vol. 8, no. 2, hlm. 966–977, 2024, doi: 10.30865/mib.v8i2.7610.
- [27] S. Khan dan H. Alghulaiakh, “*ARIMA Model for Accurate Time Series Stocks Forecasting*,” *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 7, hlm. 524–528, 2020, doi: 10.14569/ijacsa.2020.0110765.
- [28] A. R. Putri, M. Usman, Warsono, Widiarti, dan Virginia, “*Application of Vector Autoregressive with Exogenous Variable: Case Study of Closing Stock Price of PT INDF. Tbk And PT ICBP. Tbk*,” dalam *Journal of Physics: Conference Series*, 2021, hlm. 12012. doi: 10.1088/1742-6596/1751/1/012012.
- [29] S. Muzaffar dan A. Afshari, “*Short-Term Load Forecasts Using LSTM Networks*,” *Energy Procedia*, vol. 158, hlm. 2922–2927, 2019, doi: 10.1016/j.egypro.2019.01.952.
- [30] A. Moghar dan M. Hamiche, “*Stock Market Prediction Using LSTM Recurrent Neural Network*,” *Procedia Comput Sci*, vol. 170, hlm. 1168–1173, 2020, doi: 10.1016/j.procs.2020.03.049.

- [31] H. N. Bhandari, B. Rimal, N. R. Pokhrel, R. Rimal, K. R. Dahal, dan R. K. C. Khatri, “Predicting Stock Market Index using LSTM,” *Machine Learning with Applications*, vol. 9, hlm. 100320, 2022, doi: <https://doi.org/10.1016/j.mlwa.2022.100320>.
- [32] A. B. Aji dan I. Surjandari, “Hybrid Vector Autoregression–Recurrent Neural Networks to Forecast Multivariate Time Series Jet Fuel Transaction Price,” dalam *IOP Conference Series: Materials Science and Engineering*, 2020, hlm. 12079. doi: 10.1088/1757-899X/909/1/012079.
- [33] L. B. Sina, C. A. Secco, M. Blazevic, dan K. Nazemi, “Hybrid Forecasting Methods—A Systematic Review,” *Electronics (Basel)*, vol. 12, no. 9, hlm. 2019, 2023, doi: 10.3390/electronics12092019.
- [34] I. Purnama, R. Saputra, dan A. Wibowo, “Implementasi Data Mining Menggunakan CRISP-DM pada Sistem Informasi Eksekutif Dinas Kelautan dan Perikanan Provinsi Jawa Tengah,” dalam *Prosiding Seminar Nasional Ilmu Komputer UNDIP 2012*, 2012.
- [35] Y. A. Suwitono dan F. J. Kaunang, “Implementasi Algoritma Convolutional Neural Network (CNN) Untuk Klasifikasi Daun Dengan Metode Data Mining SEMMA Menggunakan Keras,” *Jurnal Komtika (Komputasi dan Informatika)*, vol. 6, no. 2, hlm. 109–121, 2022, doi: 10.31603/komtika.v6i2.8054.
- [36] J. Saltz, “What is a Data Science Workflow?,” Data Science Process Alliance. Diakses: 23 Februari 2024. [Daring]. Tersedia pada: <https://www.datascience-pm.com/data-science-workflow/>
- [37] U. Shafique dan H. Qaiser, “A Comparative Study of Data Mining Process Models (KDD, CRISP-DM and SEMMA),” *International Journal of Innovation and Scientific Research*, vol. 12, no. 1, hlm. 217–222, 2014.
- [38] A. K. Fatmawati, L. S. Badriah, dan D. S. Gunawan, “Perkembangan Indeks Harga Saham Gabungan (IHSG) Sebelum dan Sesudah Krisis Ekonomi Global Tahun 2008,” *Journal of Economic, Management, Accounting and Technology*, vol. 7, no. 1, hlm. 31–39, 2024, doi: 10.32500/jematech.v7i1.5233.
- [39] V. Nugroho dan I. Ramli, “Krisis Ekonomi Krisis Politik Dunia dan IHSG,” *Jurnal Ekonomi*, vol. 21, no. 1, hlm. 61–75, 2016, doi: 10.24912/je.v21i1.382.
- [40] A. Kurnia, L. Marlinah, dan R. Mointi, “The Influence Of Rupiah Exchange

Course In IHSG In Index Stock Exchange (IDX), Case Study Of Covid-19 Effects In Indonesia,” Economy Deposit Journal (E-DJ), vol. 3, no. 2, hlm. 106–112, 2021, doi: 10.36090/e-dj.v3i2.1088.