

**PREDIKSI HARGA SAHAM PADA INDEKS KOMPAS100  
MENGUNAKAN *BIDIRECTIONAL LONG SHORT-TERM  
MEMORY* DENGAN FAKTOR *FOREIGN FLOW***

Skripsi



oleh:  
**NATHANAEL IVAN PRATAMA**  
**71190448**

PROGRAM STUDI INFORMATIKA FAKULTAS TEKNOLOGI INFORMASI  
UNIVERSITAS KRISTEN DUTA WACANA

2023

**PREDIKSI HARGA SAHAM PADA INDEKS KOMPAS100  
MENGUNAKAN *BIDIRECTIONAL LONG SHORT-TERM  
MEMORY* DENGAN FAKTOR *FOREIGN FLOW***

Skripsi



Diajukan kepada Program Studi Informatika Fakultas Teknologi Informasi  
Universitas Kristen Duta Wacana  
Sebagai Salah Satu Syarat dalam Memperoleh Gelar  
Sarjana Komputer

Disusun Oleh :

**NATHANAEL IVAN PRATAMA**  
**71190448**

**PROGRAM STUDI INFORMATIKA FAKULTAS TEKNOLOGI INFORMASI  
UNIVERSITAS KRISTEN DUTA WACANA**

2023

## PERNYATAAN KEASLIAN SKRIPSI

Saya menyatakan dengan sesungguhnya bahwa skripsi dengan judul:

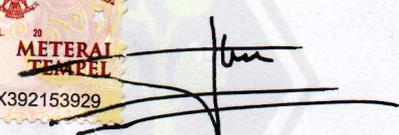
**PREDIKSI HARGA SAHAM PADA INDEKS KOMPAS100  
MENGUNAKAN *BIDIRECTIONAL LONG SHORT-TERM MEMORY*  
DENGAN FAKTOR *FOREIGN FLOW***

yang saya kerjakan untuk melengkapi sebagian persyaratan menjadi Sarjana Komputer pada pendidikan Sarjana Program Studi Informatika Fakultas Teknologi Informasi Universitas Kristen Duta Wacana, bukan merupakan tiruan atau duplikasi dari skripsi kesarjanaan di lingkungan Universitas Kristen Duta Wacana maupun di Perguruan Tinggi atau instansi manapun, kecuali bagian yang sumber informasinya dicantumkan sebagaimana mestinya.

Jika dikemudian hari didapati bahwa hasil skripsi ini adalah hasil plagiasi atau tiruan dari skripsi lain, saya bersedia dikenai sanksi yakni pencabutan gelar kesarjanaan saya.

Yogyakarta, 8 Juni 2023



  
NATHANAEL IVAN PRATAMA  
71190448

## HALAMAN PERSETUJUAN

Judul Skripsi : PREDIKSI HARGA SAHAM PADA INDEKS  
KOMPAS100 MENGGUNAKAN BIDIRECTIONAL  
LONG SHORT-TERM MEMORY DENGAN  
FAKTOR FOREIGN FLOW

Nama Mahasiswa : NATHANAEL IVAN PRATAMA

N I M : 71190448

Matakuliah : Skripsi (Tugas Akhir)

Kode : TI0366

Semester : Genap

Tahun Akademik : 2022/2023

Telah diperiksa dan disetujui di  
Yogyakarta,  
Pada tanggal 8 Juni 2023

Dosen Pembimbing I

Dosen Pembimbing II



Yuan Lukito, S.Kom., M.Cs.



Antonius Rachmat C., S.Kom.,M.Cs.

**HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI  
TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS  
SECARA ONLINE  
UNIVERSITAS KRISTEN DUTA WACANA YOGYAKARTA**

Saya yang bertanda tangan di bawah ini:

NIM : 71190448  
Nama : Nathanael Ivan Pratama  
Prodi / Fakultas : Teknologi Informasi / Informatika  
Judul Tugas Akhir : PREDIKSI HARGA SAHAM PADA INDEKS  
KOMPAS100 MENGGUNAKAN  
*BIDIRECTIONAL LONG SHORT-TERM  
MEMORY* DENGAN FAKTOR *FOREIGN  
FLOW*

bersedia menyerahkan Tugas Akhir kepada Universitas melalui Perpustakaan untuk keperluan akademis dan memberikan **Hak Bebas Royalti Non Eksklusif** (*Non-exclusive Royalty-free Right*) serta bersedia Tugas Akhirnya dipublikasikan secara online dan dapat diakses secara lengkap (*full access*).

Dengan Hak Bebas Royalti Noneklusif ini Perpustakaan Universitas Kristen Duta Wacana berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk *database*, merawat, dan mempublikasikan Tugas Akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta. Demikian pernyataan ini saya buat dengan sebenar-benarnya.

Yogyakarta, 8 Juni 2023

Yang menyatakan,



(71190448 – Nathanael Ivan Pratama)

## HALAMAN PENGESAHAN

### PREDIKSI HARGA SAHAM PADA INDEKS KOMPAS100 MENGUNAKAN BIDIRECTIONAL LONG SHORT-TERM MEMORY DENGAN FAKTOR FOREIGN FLOW

Oleh: NATHANAEL IVAN PRATAMA / 71190448

Dipertahankan di depan Dewan Penguji Skripsi  
Program Studi Informatika Fakultas Teknologi Informasi  
Universitas Kristen Duta Wacana - Yogyakarta  
Dan dinyatakan diterima untuk memenuhi salah satu syarat memperoleh gelar  
Sarjana Komputer  
pada tanggal 19 Juni 2023

Yogyakarta, 3 Juli 2023  
Mengesahkan,

Dewan Penguji:

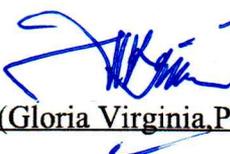
1. Yuan Lukito, S.Kom., M.Cs.
2. Antonius Rachmat C., S.Kom.,M.Cs.
3. Nugroho Agus Haryono, M.Si
4. R. Gunawan Santosa, Drs. M.Si.



Dekan

(Restyandito, S.Kom.,MSIS.,Ph.D.)

Ketua Program Studi



(Gloria Virginia, Ph.D.)

**HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI  
TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS  
SECARA ONLINE  
UNIVERSITAS KRISTEN DUTA WACANA YOGYAKARTA**

Saya yang bertanda tangan di bawah ini:

NIM : 71190448  
Nama : Nathanael Ivan Pratama  
Prodi / Fakultas : Teknologi Informasi / Informatika  
Judul Tugas Akhir : PREDIKSI HARGA SAHAM PADA INDEKS  
KOMPAS100 MENGGUNAKAN  
*BIDIRECTIONAL LONG SHORT-TERM  
MEMORY* DENGAN FAKTOR *FOREIGN  
FLOW*

bersedia menyerahkan Tugas Akhir kepada Universitas melalui Perpustakaan untuk keperluan akademis dan memberikan **Hak Bebas Royalti Non Eksklusif** (*Non-exclusive Royalty-free Right*) serta bersedia Tugas Akhirnya dipublikasikan secara online dan dapat diakses secara lengkap (*full access*).

Dengan Hak Bebas Royalti Noneklusif ini Perpustakaan Universitas Kristen Duta Wacana berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk *database*, merawat, dan mempublikasikan Tugas Akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta. Demikian pernyataan ini saya buat dengan sebenar-benarnya.

Yogyakarta, 8 Juni 2023

Yang menyatakan,



(71190448 – Nathanael Ivan Pratama)



Karya sederhana ini dipersembahkan

Kepada Papa dan Mama



*Be not overcome of evil, but overcome evil with good.*

Paul the Apostle

## KATA PENGANTAR

Segala puji, hormat, dan syukur kepada Tuhan Sang Kasih, karena atas segala rahmat, bimbingan, dan pertolongan-Nya maka pada akhirnya Skripsi dengan judul PREDIKSI HARGA SAHAM PADA INDEKS KOMPAS100 MENGGUNAKAN BIDIRECTIONAL LONG SHORT-TERM MEMORY DENGAN FAKTOR FOREIGN FLOW ini dapat diselesaikan.

Penulis memperoleh banyak bantuan dari kerja sama baik secara moral maupun spiritual dalam penulisan Skripsi ini, untuk itu tak lupa penulis ucapkan terima kasih yang sebesar-besarnya kepada:

1. Tuhan Yesus yang maha kasih,
2. Orang tua yang selama ini telah sabar membimbing dan mendoakan penulis,
3. Jessica Amanda Benu, selaku bunga hati yang tidak pernah berhenti memberikan nyawa api untuk bersemuka dengan hari-hari pengerjaan tugas akhir,
4. Bapak Restyandito, S.Kom., MSIS., Ph.D. selaku Dekan FTI,
5. Ibu Gloria Virginia, S.Kom., MAI, Ph.D. selaku Kaprodi Informatika,
6. Bapak Yuan Lukito, S.Kom., M.Cs. selaku Dosen Pembimbing 1, yang telah memberikan ilmunya dan dengan penuh kesabaran membimbing penulis,
7. Bapak Antonius Rachmat C., S.Kom., M.Cs. selaku Dosen Pembimbing 2 yang telah memberikan ilmu dan kesabaran dalam membimbing penulis
8. Keluarga tercinta di Sleman, Tata Anik, Mbah Kakung, dan Mbah Uti
9. Lain-lain yang telah mendukung moral, spiritual, dan dana untuk belajar selama ini.

Skripsi ini tentunya tidak lepas dari segala kekurangan dan kelemahan, untuk itu segala kritikan dan saran yang bersifat membangun guna kesempurnaan skripsi ini sangat diharapkan. Semoga skripsi ini dapat bermanfaat bagi pembaca semua dan lebih khusus lagi bagi pengembangan ilmu komputer dan teknologi informasi.

Yogyakarta, 08 Juni 2023

Penulis

## DAFTAR ISI

PERNYATAAN KEASLIAN SKRIPSI.....	iii
HALAMAN PERSETUJUAN.....	iv
HALAMAN PENGESAHAN.....	v
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS SECARA ONLINE.....	vi
UNIVERSITAS KRISTEN DUTA WACANA YOGYAKARTA .....	vi
KATA PENGANTAR .....	ix
DAFTAR ISI.....	x
DAFTAR TABEL.....	xii
DAFTAR GAMBAR .....	xiii
INTISARI.....	xiv
ABSTRACT.....	xvi
BAB I.....	1
PENDAHULUAN .....	1
1.1. Latar Belakang Masalah.....	1
1.2. Perumusan Masalah.....	2
1.3. Batasan Masalah.....	2
1.4. Tujuan Penelitian.....	3
1.5. Manfaat Penelitian.....	3
1.6. Sistematika Penulisan.....	3
BAB II.....	5
TINJAUAN PUSTAKA DAN DASAR TEORI .....	5
2.1 Tinjauan Pustaka .....	5
2.2 Landasan Teori .....	6
BAB III .....	19
METODOLOGI PENELITIAN.....	19
3.1 Analisis Kebutuhan Sistem .....	19
3.1.1 Kebutuhan Fungsional .....	19
3.1.2 Kebutuhan Non Fungsional.....	19
3.2 Perancangan Penelitian.....	20
3.2.1 Pengumpulan Data .....	20
3.2.2 Perancangan Pra-Pemrosesan .....	21
3.2.3 Perancangan Model.....	21
3.2.4 Perancangan Evaluasi.....	24
3.2.5 Perancangan Antar Muka.....	24
3.3 Diagram Alir.....	25
BAB IV .....	29
IMPLEMENTASI DAN PEMBAHASAN.....	29

4.1	Implementasi Awal.....	29
4.1.1	Pengumpulan Data .....	29
4.1.2	Pra-pemrosesan Data.....	30
4.1.3	Pembuatan Input Fitur.....	32
4.2	Implementasi Sistem .....	33
4.2.1	Pembuatan Model Terbaik .....	33
4.2.2	Perhitungan Pearson Correlation Fitur Close dan Netforeign.....	41
4.2.3	Prediksi.....	43
4.2.4	Pembangunan Antarmuka .....	47
4.3	Analisis Hasil .....	48
BAB V.....		52
KESIMPULAN DAN SARAN.....		52
5.1	Kesimpulan.....	52
5.2	Saran .....	52
DAFTAR PUSTAKA .....		54
LAMPIRAN A.1 .....		56
DATA PENELITIAN .....		56
LAMPIRAN A.2 .....		70
HASIL MODEL TERBAIK .....		70
LAMPIRAN A.3 .....		73
KODE SUMBER PROGRAM .....		73
LAMPIRAN B .....		79
KARTU KONSULTASI DOSEN 1.....		79
LAMPIRAN C .....		80
KARTU KONSULTASI DOSEN 2.....		80
LAMPIRAN D.....		81
LAMPIRAN LAIN-LAIN .....		81



## DAFTAR TABEL

Tabel 2.1 Daftar Emiten saham indeks KOMPAS100 periode Februari 2023 s.d Juli 2023 .....	8
Tabel 2. 2 Data contoh perhitungan MAE .....	15
Tabel 2. 3 Data hasil pertama dalam perhitungan MAE.....	15
Tabel 2. 4 Data hasil perhitungan pertama dari RMSE .....	16
Tabel 2. 5 Kriteria Pearson Correlation .....	18
Tabel 4. 1 Hasil rata-rata RMSPE perbandingan BiLSTM dengan LSTM	34
Tabel 4. 2 Hasil rata-rata RMSPE perbandingan Foreign Flow .....	37
Tabel 4. 3 Hasil rata-rata RMSPE perbandingan Stacking .....	39
Tabel 4. 4 Emiten saham dengan model terbaik .....	41
Tabel 4. 5 Hasil perhitungan Pearson Correlation .....	42
Tabel 4. 6 Foreign flow terhadap Pearson Correlation .....	49
Tabel 4. 7 Hasil prediksi data testing pada saham terbaik .....	50
Tabel 4. 8 Hasil prediksi data real pada saham terbaik.....	50
Tabel A. 1 Hasil validasi perbandingan BiLSTM vs LSTM bagian BiLSTM.....	56
Tabel A. 2 Hasil validasi perbandingan BiLSTM vs LSTM bagian LSTM .....	58
Tabel A. 3 Hasil validasi Perbandingan Foreign Flow saat menggunakan Foreign Flow .....	60
Tabel A. 4 Hasil validasi Perbandingan Foreign Flow saat Tanpa Foreign Flow	62
Tabel A. 5 Hasil validasi Perbandingan Stacking bagian Single .....	64
Tabel A. 6 Hasil validasi Perbandingan Stacking bagian Stacking .....	67
Tabel A. 7 Model Terbaik.....	70

## DAFTAR GAMBAR

Gambar 2. 1 Arsitektur LSTM.....	10
Gambar 2. 2 Forget gate.....	11
Gambar 2. 3 Input gate.....	12
Gambar 2. 4 Output gate.....	13
Gambar 2. 5 Arsitektur BiLSTM.....	14
Gambar 3. 1 Total parameter menggunakan <i>single</i> BiLSTM.....	22
Gambar 3. 2 Total parameter menggunakan stacked BiLSTM.....	23
Gambar 3. 3 Diagram Alir Penelitian.....	25
Gambar 4. 1 Contoh data historis saham yang telah dilakukan filtering.....	30
Gambar 4. 2 Contoh data yang telah dilakukan perubahan indeks.....	31
Gambar 4. 3 Data statistik pada emiten BBKA.....	31
Gambar 4. 4 Contoh output prediksi.....	45
Gambar 4. 5 Dataframe gabungan Prediction dan Harga Real.....	46
Gambar 4. 6 Grafik prediksi BBKA.....	47
Gambar 4. 7 Tampilan antarmuka utama.....	47
Gambar 4. 8 Tampilan antarmuka hasil prediksi.....	48



## INTISARI

### **PREDIKSI HARGA SAHAM PADA INDEKS KOMPAS100 MENGUNAKAN *BIDIRECTIONAL LONG SHORT-TERM MEMORY* DENGAN FAKTOR *FOREIGN FLOW***

Oleh:

NATHANAEL IVAN PRATAMA

71190448

Saham merupakan instrumen investasi yang memiliki peluang keuntungan besar jika dilakukan analisis yang baik dan benar. *Long Short-Term Memory (LSTM)* menjadi salah satu algoritma yang dipakai untuk menganalisis harga saham. Penggunaan LSTM yang bekerja dua arah (*BiLSTM*) untuk memprediksi harga closing saham di beberapa penelitian lain memiliki hasil evaluasi yang baik. Ditambahnya faktor lain, dalam hal ini Arus Asing (*Foreign Flow*) diharapkan mampu menghasilkan model *BiLSTM* yang lebih baik lagi.

Hasil penelitian pada tahap pengembangan model terbaik menggunakan saham dari berbagai sektor dalam Indeks KOMPAS100 menunjukkan bahwa tidak semua emiten saham berhasil memperoleh performa yang baik dengan metode *BiLSTM*. Hanya 41,67% dari total emiten saham yang menghasilkan prediksi yang lebih baik saat menggunakan *BiLSTM*. Saat perbandingan faktor *Foreign Flow*, 25% emiten saham menunjukkan performa yang lebih baik, tetapi beberapa juga menunjukkan hasil yang sebaliknya. Korelasi positif dalam *Pearson Correlation* tidak selalu terjadi pada saham dengan hasil baik saat menggunakan *Foreign Flow* meskipun 75% emiten saham menunjukkan korelasi positif. Mayoritas emiten saham (64,28%) lebih baik dalam prediksi saat menggunakan *single BiLSTM*,

mengindikasikan bahwa prediksi saham yang akurat tidak selalu memerlukan model yang kompleks. Beberapa saham seperti BBCA, INDF, ICBP, SIDO, dan MYOR menunjukkan hasil terbaik dalam penelitian ini. Secara keseluruhan, penggunaan metode BiLSTM dengan faktor Foreign Flow tidak selalu menghasilkan prediksi yang baik untuk semua saham yang diperjualbelikan oleh asing setiap harinya, dan terkadang prediksi yang akurat dapat dicapai dengan model yang lebih sederhana.

**Kata-kata kunci :** Saham, BiLSTM, Bursa Efek Indonesia, Foreign Flow, KOMPAS100, Pearson Correlation, RMSPE



## **ABSTRACT**

### **STOCK PRICE PREDICTION ON THE KOMPAS100 INDEX USING BIDIRECTIONAL LONG SHORT-TERM MEMORY WITH THE FOREIGN FLOW FACTOR**

By:

NATHANAEL IVAN PRATAMA

71190448

Stocks are investment instruments that have big profit opportunities if done a good and correct analysis. Long Short-Term Memory (LSTM) is one of the algorithms used to analyze stock prices. The use of LSTM that works in two directions (BiLSTM) to predict stock closing prices in several other studies has had good evaluation results. The addition of other factors, in this case Foreign Flows (Foreign Flow) is expected to be able to produce an even better BiLSTM model.

The results of research at the best model development stage using stocks from various sectors in the KOMPAS100 Index show that not all stock issuers have succeeded in obtaining good performance with the BiLSTM method. Only 41.67% of total stock issuers produce better predictions when using BiLSTM. When comparing Foreign Flow factors, 25% of stock issuers show better performance, but some also show the opposite result. The positive correlation in the Pearson Correlation does not always occur in stocks with good results when using Foreign Flow even though 75% of stock issuers show a positive correlation. The majority of stock issuers (64.28%) are better at predicting when using a single BiLSTM, indicating that accurate stock prediction does not always require a complex model. Several stocks such as BBKA, INDF, ICBP, SIDO, and MYOR showed the best results in this study. Overall, the use of the BiLSTM method with the Foreign Flow factor does not always produce good predictions for all shares traded by foreigners

every day, and sometimes accurate predictions can be achieved with simpler models.

**Keywords:** Stock, Indonesian Stock Exchange, Foreign Flow, KOMPAS100, Pearson Correlation, RMSPE



# BAB I

## PENDAHULUAN

### 1.1. Latar Belakang Masalah

Saham merupakan salah satu instrumen dalam berinvestasi yang dinilai menarik karena memiliki peluang keuntungan yang besar. Berbagai pendekatan analisis fundamental dan analisis teknikal kerap kali menjadi senjata utama untuk meramal atau memprediksi naik dan turunnya harga saham agar meminimalkan risiko serta mendapatkan lebih banyak profit (Patel, Choudhary, Saxena, & Singh, 2021). Selain analisis fundamental dan teknikal, *Deep Learning* menjadi salah satu pilihan untuk membuat ramalan saham yang akurat.

Dalam penggunaan *Deep Learning*, analisis biasanya menggunakan bermacam-macam model serta sumber data. Satu dari beberapa model yang populer digunakan untuk meramalkan harga saham adalah *Long Short-Term Memory (LSTM)*. LSTM sendiri memiliki beberapa tipe lainnya yang arsitekturnya telah disesuaikan. Salah satunya yaitu *Bidirectional LSTM* atau LSTM yang bekerja secara dua arah.

Di lain sisi, untuk memperoleh hasil yang lebih andal, analisis memakai faktor-faktor yang dapat menyebabkan perubahan dari harga saham, salah satunya adalah *Foreign Flow* atau Aliran Asing. *Foreign flow* adalah indikator momentum yang mengukur kekuatan aktivitas pembelian saham asing. Faktor aliran asing telah terbukti memiliki dapat memengaruhi pergerakan harga saham (Zulfan, Hashfi, & Ary, 2021).

Dengan memanfaatkan model *Bidirectional LSTM* yang digunakan untuk memperhitungkan tren yang mungkin menjadi indikasi harga saham di masa depan serta faktor *foreign flow* yang digunakan untuk memperhitungkan dampak investasi asing di pasar saham Indonesia, maka gambaran harga saham masa depan yang diharapkan akurat dapat dihasilkan.

Pada akhirnya, penelitian ini akan membahas penggunaan *Bidirectional Long Short-Term Memory* dengan faktor *foreign flow* dalam peramalan saham serta mengevaluasi keefektifan model ini dalam meramalkan tren pasar saham.

## 1.2. Perumusan Masalah

Sesuai dengan latar belakang yang telah diterangkan di atas, permasalahan yang akan diteliti dalam skripsi ini adalah menguji kemampuan *Bidirectional LSTM* dalam memprediksi saham indeks KOMPAS100 dengan menggunakan faktor *Foreign Flow*. Penelitian ini juga Mencakup perhitungan evaluasi seperti *Mean Absolute Error (MAE)*, dan *Root Mean Square Error (RMSE)*, serta *Root Mean Square Percentage Error (RMSPE)* sebagai parameter keberhasilan model.

## 1.3. Batasan Masalah

Pada penelitian ini, mempresentasikan beberapa batasan masalah yang telah ditentukan, seperti misalnya:

1. *Dataset* didapat dari Dosen Pembimbing I yang bersumber dari situs Bursa Efek Indonesia ([idx.co.id](http://idx.co.id)).
2. Dalam penelitian ini, emiten saham yang digunakan adalah seluruh saham yang terdapat pada daftar indeks KOMPAS100 dari bulan Februari 2023 hingga Juli 2023, dengan periode data dari awal pembukaan bursa saham tahun 2016 pada 04 Januari 2016 sampai dengan penutupan bursa saham tahun 2022 pada 30 Desember 2022.
3. Data dari tahun 2016 sampai dengan 2021 akan digunakan untuk training dan validasi dengan detail sebagai berikut, 70% data akan dipakai untuk *training* dan 30% sisanya digunakan untuk data *validasi*.
4. Penelitian ini hanya sampai pada tahap eksperimental pengembangan model, serta pembuatan antarmuka untuk memudahkan menampilkan data dan informasi, belum sampai kepada tahap pembangunan sistem rekomendasi jual beli saham

#### **1.4. Tujuan Penelitian**

Pada penelitian ini, tujuan yang harapannya dapat dicapai adalah menggunakan *BiLSTM* dengan faktor *Foreign Flow* untuk meramalkan harga saham pada masa mendatang. Dengan menggabungkan model dan faktor tersebut, penulis berharap dapat membuat prediksi saham yang lebih akurat dan andal daripada tanpa menggunakan faktor *Foreign Flow*. Oleh sebab itu dalam penelitian ini juga akan membandingkan prediksi menggunakan *BiLSTM* dan *LSTM*, perbandingan penggunaan *Foreign Flow* sebagai fitur dan tidak sebagai fitur. Serta korelasi dari harga closing saham dengan arus asing menggunakan *Pearson Correlation*. Tujuan lain yang ingin dicapai dalam penelitian ini adalah menyediakan pemahaman lebih tentang pasar saham dan memberikan opini lain kepada para investor dalam upaya untuk mengambil keputusan investasi yang lebih bijak.

#### **1.5. Manfaat Penelitian**

Di dalam penelitian ini besar harapan penulis agar dapat bermanfaat bagi:

1. Investor mau pun *trader* saham yang hendak terjun ke dalam dunia saham dan memulai dari saham yang terdapat pada indeks *KOMPAS100* agar mendapat gambaran prediksi ke depannya sebagai bahan pertimbangan dalam mengambil keputusan.
2. *Deep learning enthusiast* di Indonesia yang ingin menggunakan *Bidirectional Long Short-Term Memory* ditambah faktor lain dalam eksperimennya.
3. Mahasiswa dan masyarakat dalam pengembangan keilmuan dalam prediksi saham khususnya di Indonesia

#### **1.6. Sistematika Penulisan**

Supaya gambaran penelitian yang akan dilakukan dapat dipahami dengan jelas, penulis membagi laporan ini dalam 5 bab: Bab 1 Pendahuluan yang berisi

latar belakang masalah, lalu perumusan masalah, batasan masalah, tujuan penelitian, dan juga manfaat penelitian. Bab yang kedua yaitu Tinjauan Pustaka dan Dasar Teori yang berisi tinjauan pustaka di mana berisi tentang penelitian-penelitian terkait dengan topik yang serupa, dan landasan teori. Bab 3 yaitu Metodologi Penelitian yang terdiri atas yang pertama analisis kebutuhan sistem yang terdiri atas dua bagian yaitu kebutuhan fungsional dan kebutuhan non fungsional. untuk kebutuhan non fungsional terdiri dari analisis kebutuhan perangkat keras serta kebutuhan perangkat lunak. yang kedua adalah perancangan penelitian yang terdiri atas pengumpulan data, perancangan pra-pemrosesan, perancangan model *BiLSTM*, dan perancangan pengujian. yang terakhir adalah diagram alir. Bab 4 berisi Implementasi awal, Implementasi Sistem, Analisis Hasil. Ditutup dengan Bab 5 yaitu kesimpulan serta saran yang dapat dilakukan oleh penelitian di masa mendatang.



## **BAB II**

### **TINJAUAN PUSTAKA DAN DASAR TEORI**

#### **2.1 Tinjauan Pustaka**

Penelitian tentang prediksi saham tentu sudah tidak terhitung lagi jumlahnya di dunia khususnya Indonesia. Terutama penggunaan LSTM konvensional untuk prediksi saham, seperti penelitian yang dilakukan oleh salah satu lulusan informatika UKDW pada tahun 2022 yang berjudul “Prediksi Pergerakan Harga Saham Dengan Metode Long Short-Term Memory (LSTM)” (Santoso, 2022). Namun, para peneliti mulai mencoba metode yang dianggap lebih baik dari LSTM konvensional. *Bidirectional Long Short-Term Memory (BiLSTM)* adalah sebuah variasi model konvensional dari *Long Short-Term Memory* yang dapat memproses *input* secara dua arah, yaitu dengan memproses *input* dari periode yang lampau ke masa mendatang (*forward*) dan dari masa mendatang ke masa lampau (*backward*). BiLSTM umumnya digunakan dalam aplikasi *Natural Language Processing (NLP)* karena kemampuannya untuk menangkap lebih banyak informasi dan konteks dari *input* yang sedang diproses. Berikut ini adalah tinjauan pustaka untuk penggunaan BiLSTM dalam dunia regresi atau timeseries untuk prediksi saham:

Yang pertama adalah penelitian berjudul “*Analysis and Research on Stock Price of LSTM and Bidirectional LSTM Neural Network*” (Mingzhu Jia, 2019). Penelitian tersebut menghasilkan tingkat prediksi yang lebih akurat daripada menggunakan LSTM. Kriteria penilaian yang digunakan termasuk *Root Mean Square Error (RMSE)*, *Mean Absolute Error (MAE)*, dan *Loss*.

Yang kedua adalah penelitian yang dilakukan (Chenyu Han, 2023) berjudul “*Challenge and Opportunity: Deep Learning-Based Stock Price Prediction by Using Bi-Directional LSTM Model*”. Dalam penelitian mereka, BiLSTM diterapkan pada data historis saham perusahaan Apple dan mengevaluasi kinerjanya menggunakan MSE. Penelitian tersebut menunjukkan bahwa model Bi-LSTM mampu membuat prediksi yang akurat pada data pengujian dan mampu menangkap beberapa tren dan pola dalam data, meskipun mungkin mengalami kesulitan saat terjadi perubahan tiba-tiba di pasar. Secara keseluruhan, hasil yang didapat

menunjukkan bahwa model Bi-LSTM adalah alat yang menjanjikan untuk prediksi harga saham dan memiliki banyak aplikasi potensial dalam bidang keuangan dan investasi.

Adapun tinjauan pustaka yang menjadi cikal-bakal dipilihnya topik ini adalah penelitian yang dilakukan oleh (Nurdi Afrianto, 2021) dengan judul “Prediksi Harga Saham Menggunakan BiLSTM dengan Faktor Sentimen Publik”. Dalam penelitiannya, Afrianto menggabungkan faktor sentimen publik, yaitu opini dan pandangan masyarakat tentang saham, sebagai input pada model BiLSTM. Dengan memanfaatkan data historis harga saham dan sentimen publik, penelitian ini bertujuan untuk mengembangkan sebuah model prediksi harga saham yang lebih akurat. Hasil penelitian menunjukkan bahwa model BiLSTM dengan faktor sentimen publik mampu memberikan hasil prediksi harga saham yang lebih akurat dibandingkan dengan model yang hanya mengandalkan data historis harga saham.

Berdasarkan penelitian-penelitian terdahulu di atas, penggunaan BiLSTM beberapa kali mengalahkan penggunaan LSTM konvensional dalam prediksi, hal ini tentu membuat BiLSTM lebih unggul dari LSTM konvensional. Ditambah dengan penambahan faktor selain harga saham sebagai fitur dalam penelitian menambah akurasi dalam prediksi harga saham. Oleh sebab itu dalam penelitian ini, penulis memilih untuk menggunakan BiLSTM ditambah dengan faktor selain harga dalam hal ini adalah *Foreign Flow*, dengan harapan akan mendapatkan hasil prediksi yang jauh akurat.

Hal yang menjadi pembeda penelitian ini dengan penelitian sebelumnya di atas adalah dari segi model, di mana model akan dibuat per emiten saham tidak satu model untuk semua emiten saham, yang kedua dari segi obyek penelitian yaitu menggunakan foreign flow sehingga akan didapat hasil korelasi antara foreign flow dengan harga saham, masih dari segi obyek penelitian, penelitian ini menggunakan indeks KOMPAS100 yang merupakan pembobotannya berdasar *Market Capitalization*.

## **2.2 Landasan Teori**

### **2.2.1 Saham**

Saham merupakan sebuah instrumen keuangan yang menyatakan klaim atas sebagian dari aset dan laba suatu perusahaan, bisa juga diartikan sebagai tanda kepemilikan seseorang maupun badan dalam suatu perusahaan, dapat juga diartikan sebagai surat bukti penyertaan atas suatu perusahaan yang mengeluarkan emiten (Asia, 2020). Saat seseorang membeli saham dari suatu perusahaan, ia secara otomatis menjadi pemegang saham atau pemilik perusahaan tersebut sebesar persentase kepemilikan saham yang dia miliki.

Harga saham ditentukan oleh permintaan serta penawaran yang terjadi di pasar saham. Naik turunnya harga saham dapat berubah dalam waktu yang relatif cepat, bisa dalam menit, bahkan detik. Jika permintaan saham lebih tinggi daripada penawaran, harga saham akan meningkat, dan sebaliknya jika penawaran saham lebih tinggi daripada permintaan, harga saham akan menurun. Faktor-faktor yang dapat memengaruhi permintaan dan penawaran saham antara lain kondisi ekonomi, prospek keuangan perusahaan, dan sentimen pasar (Linanda & Arfiyenis, 2018).

Membeli saham berarti berinvestasi dalam perusahaan tertentu dan berharap bahwa perusahaan tersebut akan menghasilkan laba yang baik di masa depan sehingga harga sahamnya akan naik. Namun seperti investasi lainnya, risiko dalam berinvestasi saham cukup tinggi karena harga saham dapat berfluktuasi secara signifikan dalam waktu singkat. Oleh karena itu, sebelum berinvestasi saham sangat penting untuk melakukan riset dan menganalisis kondisi perusahaan dan pasar secara menyeluruh.

### **2.2.2 KOMPAS100**

Kompas100 merupakan sekelompok saham yang memiliki fundamental dan kinerja yang baik, terutama dalam hal kapitalisasi pasar yang diterbitkan oleh Grup Kompas Gramedia. Sesuai dengan namanya, indeks ini terdiri dari 100 saham perusahaan publik. Indeks ini secara resmi diterbitkan pada tanggal 10 Agustus 2007. Kompas100 memuat saham-saham yang diperdagangkan di BEI dari berbagai sektor seperti perbankan, infrastruktur, konsumen dan lain-lainnya. Saham-saham yang termasuk dalam indeks ini dipilih berdasarkan kriteria seperti

likuiditas, kapitalisasi pasar, dan kinerja keuangan (Alfath & Dewi, 2022). Adapun kriteria KOMPAS100 dilansir dari situs Bursa Efek Indonesia (idx.co.id) adalah:

1. Terdapat pada list saham Bursa Efek Indonesia sekurang-kurangnya 3 bulan,
2. Masuk dalam perhitungan Indeks Harga Saham Gabungan,
3. BEI dapat memasukkan saham tersebut ke dalam perhitungan indeks harga 100 saham berdasarkan evaluasi faktor-faktor fundamental perusahaan dan pola perdagangan pasar,
4. Selama periode 12 bulan terakhir, saham tersebut telah berada di posisi 150 teratas berdasarkan nilai transaksi, frekuensi transaksi, dan kapitalisasi pasar di Pasar Reguler,
5. Dari total 150 saham, jumlahnya kemudian dikurangi menjadi 60 saham dengan mempertimbangkan nilai transaksi terbesar. Selanjutnya, dari sisa 90 saham tersebut, dipilih 40 saham berdasarkan kinerja mereka dalam hal hari transaksi, frekuensi transaksi, dan nilai kapitalisasi pasar di pasar reguler. Setelah itu, dari sisa 90 saham yang tersisa, dipilih 75 saham berdasarkan hari transaksi di pasar reguler. Dari 75 saham tersebut, dipilih 60 saham berdasarkan frekuensi transaksi di pasar reguler. Lalu, dari 60 saham tersebut, dipilih 40 saham berdasarkan kapitalisasi pasar. Akhirnya, daftar 100 saham diperoleh dengan menggabungkan daftar saham dari perhitungan sebelumnya, yaitu hasil dari langkah 5 dan langkah 6.

Adapun daftar emiten saham yang tergabung dalam indeks KOMPAS100 terdapat pada Tabel 2.1

Tabel 2.1 Daftar Emiten saham indeks KOMPAS100 periode Februari 2023 s.d Juli 2023

AALI	ABMM	ACES	ADHI	ADMR	ADRO	AGII	AGRO
AKRA	AMRT	ANTM	APIC	ARTO	ASII	ASSA	AVIA
BABP	BBCA	BBHI	BBKP	BBNI	BBRI	BBTN	BFIN

BMRI	BMTR	BRIS	BRMS	BRPT	BSDE	BSSR	BTPS
BUKA	CPIN	CTRA	DOID	DSNG	ELSA	EMTK	ENRG
ERAA	ESSA	EXCL	GGRM	GOTO	HEAL	HMSP	HRUM
ICBP	INCO	INDF	INDY	INKP	INTP	IPPE	ISAT
ITMG	JPFA	JSMR	KLBF	LPPF	LSIP	MAPI	MDKA
MEDC	MIKA	MLPL	MNCN	MPMX	MTEL	MTMH	MYOR
PGAS	PNBN	PNLF	PTBA	PTPP	PWON	RAJA	RMKE
SCMA	SGER	SIDO	SMDR	SMGR	SMRA	SRTG	TAPG
TBIG	TINS	TKIM	TLKM	TOBA	TOWR	TPIA	UNTR
UNVR	WIFI	WIKA	WSKT				

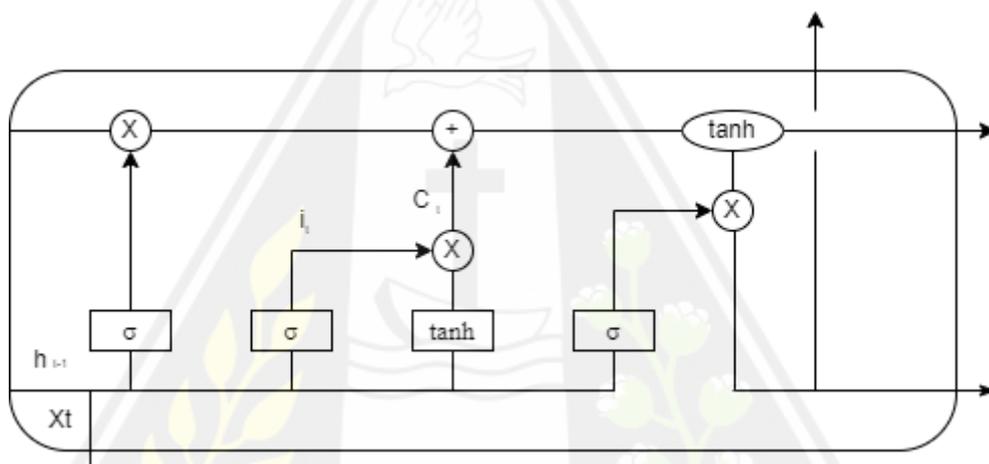
### 2.2.3 *Deep Learning*

Deep learning mengacu pada teknik machine learning baik itu supervised learning maupun unsupervised learning untuk secara otomatis mempelajari representasi hierarkis dalam arsitektur mendalam untuk klasifikasi. Terinspirasi oleh pengamatan biologis pada mekanisme otak manusia untuk memproses sinyal alami, deep learning telah menarik banyak perhatian dari komunitas akademik dalam beberapa tahun terakhir karena kinerjanya yang canggih di banyak domain penelitian seperti pengenalan ucapan, penyaringan kolaboratif, dan visi komputer. (Chandrashekar Mishra, 2017). Perbedaan *deep learning* dengan *machine learning* adalah terdapat pada tipe data dan metode belajar. *Deep learning* membuat komputer untuk belajar berbagai model dan berdasarkan data yang telah dikumpulkan akan dilakukan klasifikasi tugas-tugasnya. Pengklasifikasian data dapat berbentuk teks, gambar, bahkan suara. Tingkat akurasi tentu lebih tinggi dalam mengolah data yang berjumlah banyak

### 2.2.6 *Long Short-Term Memory*

*Long Short-Term Memory* atau LSTM merupakan algoritma lanjutan dari *Recurrent Neural Network* yang dapat menangani ketergantungan jangka panjang yang ada dalam teks (Sreelakshmi Ka, 2018). Secara definisi, LSTM merupakan algoritma yang telah dilakukan modifikasi dengan menambah *memory cell*. Penambahan ini membuat LSTM mampu menangani masalah yang tidak bisa ditangani oleh RNN yaitu *vanishing gradient* pada saat memproses data *sequential* (Wati & Pramatha, 2022). Karena penambahan sel memori ini, LSTM dapat menyimpan informasi untuk jangka waktu yang lama.

Untuk gambaran arsitektur LSTM dapat dilihat pada Gambar 2.1.

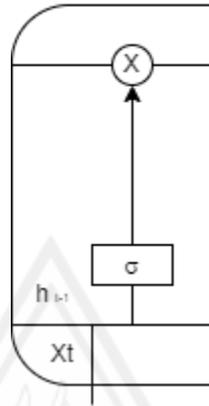


Gambar 2. 1 Arsitektur LSTM

LSTM memiliki struktur yang terdiri atas jaringan saraf dan beberapa blok memori (*cell*). Informasi yang telah dikumpulkan oleh LSTM lalu disimpan oleh blok memori dan *gate* atau gerbang akan melakukan manipulasi. Terdapat tiga jenis *gate* pada LSTM yaitu *forget gate*, *input gate*, dan *output gate*.

1. *Forget Gate*

Dapat dilihat pada Gambar 2.2, yang merupakan sebuah gerbang pertama dalam LSTM yang sesuai namanya memiliki tugas untuk “*forget*” atau melupakan beberapa informasi yang sudah tidak diperlukan atau kurang relevan dalam sistem (Wati & Pramatha, 2022). Oleh sebab itu imbasnya adalah LSTM dapat menyajikan informasi yang lengkap namun aktual sesuai dengan kebutuhan.



Gambar 2. 2 Forget gate

Luaran biner akan dievaluasi dari dua masukan yaitu  $x(t)$  dan *output cell* sebelumnya  $h(t-1)$  lalu dikalikan dengan matriks bobot kemudian ditambahkan dengan nilai bias seperti yang dapat dilihat pada persamaan 2.1.

$$f_t = \sigma(X_t * U_f + H_{t-1} * W_f) \quad [2.1]$$

Hasil tersebut lalu akan melewati fungsi aktivasi dan menghasilkan luaran biner.

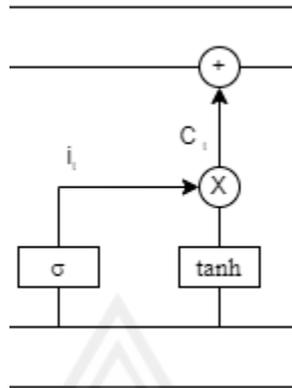
$$C_{t-1} * f_t = 0 \dots \text{if } f_t = 0 \text{ (hapus atau lupakan)} \quad [2.2]$$

$$C_{t-1} * f_t = C_{t-1} \dots \text{if } f_t = 1 \text{ (tidak dihapus)} \quad [2.3]$$

persamaan 2.2 dan 2.3 menjelaskan jika luarannya adalah nol (0) maka dianggap tidak berguna dan akan dihapus. Namun jika luarannya adalah satu (1) maka informasi tersebut akan disimpan untuk penggunaan di masa mendatang

## 2. Input Gate

Dapat dilihat pada Gambar 2.3 merupakan Gerbang lanjutan dari *forget gate* yaitu *input gate* yang memiliki tugas untuk “*input*” atau memasukkan informasi yang berguna untuk mendukung akurasi data di mana cara kerjanya adalah dengan menambahkan informasi yang didapat dari *forget gate* yang telah diseleksi (Wati & Pramatha, 2022). Gerbang ini adalah gerbang yang tidak dimiliki oleh RNN karena pada RNN hanya memungkinkan satu masukan untuk satu luaran.



Gambar 2. 3 Input gate

Informasi yang didapat dari *forget gate* akan diatur memakai fungsi *sigmoid* dan nilai yang akan disimpan harus disaring terlebih dahulu. Sama halnya dengan proses yang terdapat di *forget gate* yaitu menggunakan masukan  $h(t-1)$  dan  $x(t)$ .

$$i_t = \sigma(X_t * U_i + H_{t-1} * W_i) \quad [2.4]$$

Lalu pada persamaan 2.5, akan dibuat sebuah vektor menggunakan fungsi *tanh* dengan luaran -1 hingga +1 yang berisi probabilitas nilai  $h(t-1)$  dan  $x(t)$ . Jika hasil  $N_t$  adalah bernilai negatif, informasi dikurangkan dari blok memori namun jika bernilai positif maka informasi akan ditambahkan ke blok memori pada *timestamp* saat ini.

$$N_t = \tanh(x_i * U_c + H_{t-1} * W_c) \text{ (informasi baru)} \quad [2.5]$$

Bagian terakhir dari *input gate* ini adalah melakukan perbaruan karena  $N_t$  tidak bisa ditambah secara langsung ke dalam blok memori.

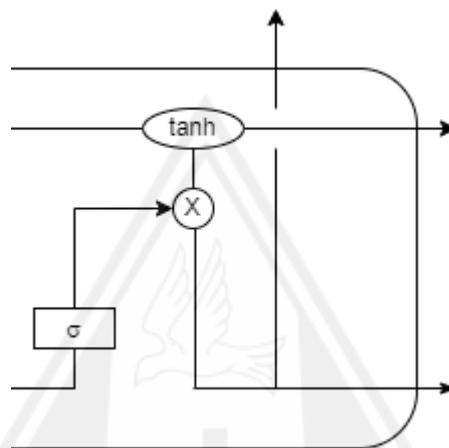
$$C_t = f_t * C_{t-1} + i_t * N_t \text{ (update blok memori)} \quad [2.6]$$

Nilai-nilai vektor dikalikan dengan nilai-nilai yang diatur untuk menghasilkan informasi yang dapat digunakan. Dengan  $C_{t-1}$  adalah blok memori yang terdapat pada *timestamp* saat ini dan nilai-nilai lainnya merupakan penghitungan sebelumnya.

### 3. Output Gate

Setelah melewati *forget gate* dan *input gate*, gerbang yang terakhir adalah *Output gate* yang bisa dilihat pada Gambar 2.4. Di mana gerbang ini akan menghasilkan informasi yang aktual dan lengkap. Gerbang ini dapat menjadi

gerbang terakhir untuk menghasilkan informasi atau bisa saja menjadi bagian dari suatu tahap, sebelum akhirnya akan diproses kembali lewat *input gate* pada blok memori berikutnya.



Gambar 2. 4 Output gate

Rumus dari *output gate* relatif mirip dengan dua gerbang sebelumnya. Karena menggunakan fungsi *sigmoid* maka akan pula bernilai 0 dan 1. Dapat kita lihat pada persamaan 2.7.

$$o_t = \sigma(x_t * U_t + H_{t-1} * W_o) \quad [2.7]$$

Lalu untuk menghitung blok memori tersembunyi,  $O_t$  akan dikalikan dengan fungsi *tanh* dari blok memori yang telah diperbarui (persamaan 2.8).

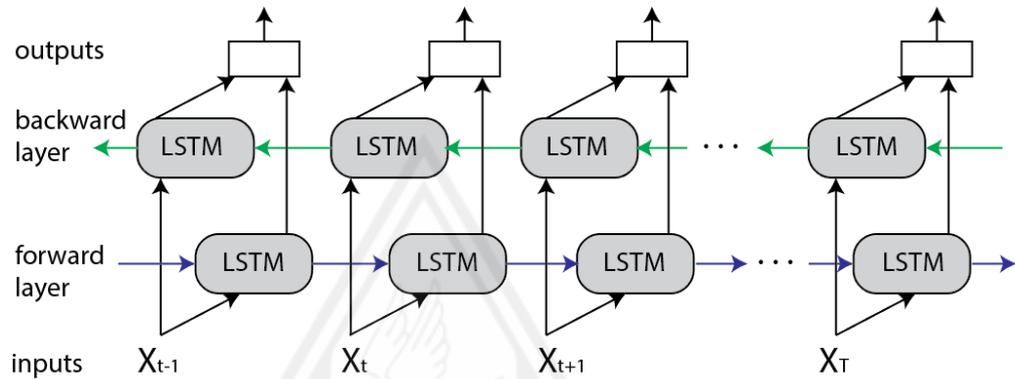
$$H_t = o_t * \tanh(C_t) \quad [2.8]$$

Informasi tersebut akan disaring dan nilai-nilai tersebut akan disimpan memakai *input*  $h(t-1)$  dan  $x(t)$ .

### 2.2.7 Bidirectional Long Short-Term Memory

*Bidirectional Long Short-Term Memory* (BiLSTM) merupakan sebuah variasi dari model LSTM konvensional yang mampu memproses input di kedua arah, maju (*forward*) dan mundur (*backward*). Dengan demikian BiLSTM memungkinkan untuk memasukkan lebih banyak konteks dengan menggunakan data dari masa lalu dan masa depan. Ini akan membantu menemukan representasi

fitur yang lebih baik dari input. Penumpukan unit LSTM Bi-Directional di setiap lapisan membantu menangkap representasi fitur non-linear (Sreelakshmi Ka, 2018)



Gambar 2. 5 Arsitektur BiLSTM (sumber:

<https://www.baeldung.com/cs/bidirectional-vs-unidirectional-lstm>)

Dapat dilihat pada Gambar 2.5, terdapat dua arah yang terdiri dari model *Forward LSTM* dan *Backward LSTM*, kemudian hasil dari kedua arah tersebut digabungkan menjadi satu *output*. Hal ini memungkinkan BiLSTM untuk menangkap lebih banyak informasi dan konteks dari *input* yang sedang diproses.

### 2.2.8 Mean Absolute Error (MAE)

*Mean Absolute Error* (MAE) merupakan rata-rata dari semua kesalahan yang absolut atau dengan kata lain MAE merupakan selisih (mutlak) dari nilai aktual dengan nilai prediksi (Suryanto & Muqtadir, 2019)

$$MAE = \frac{1}{n} \sum_{i=1}^n |x_i - \hat{x}_i| \quad [2.9]$$

Keterangan:

$n$  = jumlah data

$x_i$  = nilai hasil prediksi

$\hat{x}_i$  = nilai sebenarnya

Langkah-langkah dalam menggunakan rumus 2.9 tersebut adalah dengan yang pertama menemukan terlebih dahulu semua kesalahan absolut,  $|x_i - \hat{x}_i|$ . Lalu tambahkan seluruhnya. Yang terakhir adalah dengan membagi angka yang sudah ditambahkan tadi dengan jumlah *error*. Misalnya jika ditemukan 10 pengukuran, maka 10 dibagi dengan 10. Contoh perhitungan menggunakan data pada Tabel 2.2:

Tabel 2. 2 Data contoh perhitungan MAE

Aktual	Prediksi
1760	1759
1760	1756
1805	1755
1795	1797

Perhitungan dari MAE ini cukup sederhana. Cukup dengan nilai prediksi dikurangi nilai aktual lalu dicari nilai mutlaknya. Dalam hal ini, kita memakai contoh data baris pertama,  $|1759-1760| = 1$ . Detail dapat dilihat pada Tabel 2.3

Tabel 2. 3 Data hasil pertama dalam perhitungan MAE

Aktual	Prediksi	$ PT-At $
1760	1759	1
1760	1756	4
1805	1755	50
1795	1797	2

Setelah ditemukan masing-masing nilai mutlak, kita hanya perlu menjumlahkan seluruh nilai mutlak yang telah ditemukan yaitu 57. Lalu membaginya dengan jumlah data prediksi yaitu 57 dibagi dengan 4. Jadi, Didapati hasil yaitu 14,25. Dengan demikian MAE dari contoh perhitungan di atas adalah 14,25.

### 2.2.9 Root Mean Square Error (RMSE)

*Root Mean Square Error* atau RMSE ini adalah standar deviasi dari kesalahan prediksi. RMSE adalah hasil pengakaran dari MSE yang telah dihitung sebelumnya. RMSE berguna dalam mengevaluasi tingkat keakuratan peramalan berdasarkan data historis. Semakin kecil nilai RMSE, maka semakin akurat metode yang digunakan dalam prediksi. (Septiawan & Astuti, 2016)

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (P_t - A_t)^2}{n}} \quad [2.10]$$

Keterangan:

$n$  = jumlah data

$P_t$  = nilai prediksi pada waktu  $t$

$A_t$  = nilai aktual pada waktu  $t$

Seperti yang dapat dilihat di persamaan 2.10, cara menghitung RMSE ini adalah dengan mengurangi nilai prediksi ( $P_t$ ) dengan nilai aktual ( $A_t$ ) lalu dikuadratkan dan dijumlahkan seluruh hasilnya lalu dibagi dengan data keseluruhan. Langkah terakhir mendapatkan RMSE adalah dengan mencari nilai dari akar kuadratnya. Contoh perhitungan seperti di bawah ini:

Tetap memakai data pada tabel 2.2 Yang pertama dilakukan untuk mendapatkan nilai RMSE adalah mengurangi prediksi dengan angka aktual. Setelah itu dilakukan kuadrat. Sebagai contoh pada baris pertama  $(1759-1760)^2 = 1$ .

Tabel 2. 4 Data hasil perhitungan pertama dari RMSE

Aktual	Prediksi	$(P_t - A_t)^2$
1760	1759	1
1760	1756	16
1805	1755	2500
1795	1797	4

Lalu didapati hasilnya seperti di tabel 2.4 Setelah itu dilakukan penjumlahan atas seluruh data yang ada pada kolom  $(P_t - A_t)^2$ . Lalu didapati hasil 2521. Untuk mendapatkan hasil RMSE kita perlu mengakarkan hasil penjumlahan tadi.  $\sqrt{2521}$  adalah 50,20956084. Dengan demikian RMSE dari contoh kasus ini adalah 50,20956084.

### 2.2.10 Root Mean Square Percentage Error (RMSPE)

Jika RMSE mengukur kesalahan rata-rata antara nilai aktual dan nilai yang diprediksi atau simulasi. Evaluasi model yang menggunakan Root Mean Square Percentage Error (RMPSE) adalah untuk melihat besaran error yang terjadi.

RMSPE dapat mengukur rata-rata persentase perbedaan antara data aktual dengan hasil prediksi (Nelly Marlina, 2015) seperti rumus 2.12

$$RMSPE = \sqrt{\frac{1}{n} \left[ \sum_{t=1}^n \left( \frac{P_t - A_t}{A_t} \right)^2 \right]} \times 100 \quad [2.12]$$

Keterangan:

$P_t$  = nilai prediksi pada waktu  $t$

$A_t$  = nilai aktual pada waktu  $t$

$n$  = jumlah titik pengukuran ( $t = 1, 2, \dots, n$ )

Semakin kecil nilai RMSPE yang diperoleh maka semakin baik model yang digunakan. Menurut Reed (dalam (Aprilia, 2023)), nilai RMSPE kurang dari 25% dapat dikategorikan “acceptable” (diterima) dan nilai RMSPE kurang dari 10% dapat diartikan bahwa model berada dalam kategori “good” (baik) .

### 2.2.11 Streamlit

Streamlit merupakan sebuah *library open-source* untuk bahasa pemrograman Python yang memudahkan menyebarkan model dan visualisasi dengan mudah (Hastomo, Aini, Karno, & Rere, 2022) dalam rangka pembuatan aplikasi web interaktif untuk aplikasi *machine learning*. *Library* ini memungkinkan para peneliti untuk menulis kode Python standar dan menampilkan hasilnya sebagai aplikasi web yang interaktif. Peneliti dapat menggunakan *Streamlit* untuk membuat antarmuka antara lain prediksi saham, visualisasi data, atau aplikasi *machine learning* lainnya dengan sangat mudah tanpa harus menulis kode HTML, CSS, atau JavaScript.

Streamlit memungkinkan pengoptimalan data, penerapan, dan analisis statistik yang mudah dengan jumlah kode yang minimal. Streamlit bisa sangat berguna dalam membangun dasbor data terutama ketika tim sebagian besar terdiri dari anggota non-teknis. Streamlit mudah digunakan karena menggunakan perintah yang telah ditentukan sebelumnya untuk membangun aplikasi web berbasis data interaktif. Perintah sederhana seperti *st.write()* untuk mengimplementasikan

berbagai macam objek mulai dari teks sederhana hingga visualisasi dataframe atau matplotlib pandas (Saxena, Dhadwal, & Kowsigan, 2021).

### 2.2.11 Pearson Correlation

Pearson Correlation merupakan sebuah metode statistika yang digunakan untuk mengukur seberapa kuat hubungan antara dua variabel, tanpa mempertimbangkan apakah suatu variabel bergantung pada variabel lainnya. Semakin kuat derajat hubungan antara kedua variabel maka semakin nyata pula hubungan linier (Miftahuddin, Pratama, & Setiawan, 2021). Rumus dapat dilihat pada persamaan 2.13

$$r_{xy} = \frac{n \sum xy - (\sum x)(\sum y)}{\sqrt{[n \sum x^2 - (\sum x)^2][n \sum y^2 - (\sum y)^2]}} \quad [2.13]$$

Di mana:

- x = variabel pertama
- y = variabel kedua
- n = banyaknya pengamatan

Tabel 2. 5 Kriteria Pearson Correlation

No.	Nilai r	Interpretasi
1	0.00 – 0.19	Sangat Rendah
2	0.20 – 0.39	Rendah
3	0.40 – 0.59	Sedang
4	0.60 – 0.79	Kuat
5	0.80 – 1.00	Sangat Kuat

Pada table 2.5 terdapat kriteria pearson correlation dengan contoh pembacaan adalah untuk nilai 0.00 sampai 0.19 menunjukkan interpretasi yang sangat rendah, jika nilai positif maka Korelasi Positif yang sangat rendah, begitu juga negatif maka dapat diinterpretasikan sebagai Korelasi Negatif Sangat Rendah. Sampai dengan tertinggi adalah interpretasi sangat kuat dengan nilai r 0.80 sampai dengan 1.00

## **BAB III**

### **METODOLOGI PENELITIAN**

#### **3.1 Analisis Kebutuhan Sistem**

##### **3.1.1 Kebutuhan Fungsional**

Untuk membuat prediksi saham menggunakan model BiLSTM dengan Faktor *Foreign Flow*, proses yang dilakukan oleh sistem dan kebutuhan sistem, di antaranya:

1. *Input data*: data harga saham yang terdiri dari harga open, harga *closing*, dan *foreign flow*
2. *Preprocessing data*: Termasuk perubahan data menjadi format yang sesuai dengan kebutuhan model, dan juga mengisi *missing values* yang terdapat pada data.
3. Pelatihan model BiLSTM
4. Prediksi dari model yang telah dibentuk
5. Evaluasi hasil untuk mengetahui akurasi dari prediksi tersebut.
6. Optimasi model: Jika hasil prediksi harga saham yang diperoleh masih belum sesuai dengan yang diinginkan.

##### **3.1.2 Kebutuhan Non Fungsional**

###### **3.1.2.1 Analisis Kebutuhan Perangkat Keras**

Adapun pada penelitian ini dibutuhkan perangkat keras sebagai berikut:

- Processor Intel(R) Core(TM) i7-8750H CPU @ 2.20GHz, 2208 Mhz, 6 Core(s), 12 Logical Processor(s)
- Kartu Grafis Nvidia Geforce GTX 1050
- RAM 8GB
- SSD 256GB

### 3.1.2.2 Kebutuhan Perangkat Lunak

Kebutuhan perangkat lunak dalam adalah sebagai berikut:

- OS Name Microsoft Windows 11 Pro Version 10.0.22621
- *Phpmyadmin* dalam XAMPP yang diunduh dari situs <https://www.apachefriends.org/download.html>
- Anaconda v4.12.0 yang diunduh dari situs <https://www.anaconda.com/download>
- Jupiter Notebook v6.0.1 yang sudah satu paket dalam Anaconda
- Streamlit v1.22.0
- Visual Studio Code v1.78.2 yang diunduh dari situs <https://code.visualstudio.com/>

## 3.2 Perancangan Penelitian

### 3.2.1 Pengumpulan Data

Pengumpulan data pada penelitian ini didapat dari Dosen Pembimbing I yang diambil dari situs Bursa Efek Indonesia ([idx.co.id](http://idx.co.id)). Data merupakan riwayat saham yang berbentuk angka. Fitur yang diambil adalah Harga *Open*, Harga *Closing*, dan *Foreign Flow*. Pengumpulan data tersebut dilakukan untuk tiap emiten yang berada pada indeks KOMPAS100 periode Februari 2023 hingga Juli 2023 dalam rentang waktu 2016-2022. Pada tahap ini juga dilakukan pengubahan format *dataset* dari *sql* (Gambar 3.1) menjadi *csv* (Gambar 3.2).

date	ticker	open	high	low	close	volume	average	frequency	netforeign
2016-01-04	AALI	15850	15950	15600	15700	902000	15729	1267	6796530000
2016-01-04	ABBA	50	50	50	50	0	0	0	0
2016-01-04	ABDA	7975	7500	7500	7500	100	7500	1	0
2016-01-04	ABMM	2900	2700	2650	2650	10000	2675	2	0
2016-01-04	ACES	825	820	800	800	3396800	802	365	-590283000
2016-01-04	ACST	3020	3010	2850	2900	1053100	2947	48	2889810000
2016-01-04	ADES	1015	1025	1015	1020	3100	1018	8	0
2016-01-04	ADHI	2135	2155	2070	2070	19552400	2103	2486	-7091830000
2016-01-04	ADMF	3470	3470	3470	3470	3200	3470	6	0
2016-01-04	ADMG	89	90	88	89	36200	89	8	0
2016-01-04	ADRO	510	515	494	494	20864700	501	1170	1940960000
2016-01-04	AGRO	97	97	88	95	121800	93	16	0
2016-01-04	AGRS	85	88	88	88	200	88	1	0

Gambar 3. 1 Contoh data asli sql

	A	B	C	D
1	date	open	close	netforeign
2	1/4/2016	15850	15700	6.8E+09
3	1/5/2016	15700	16675	1.95E+10
4	1/6/2016	16725	16725	6.39E+09
5	1/7/2016	16400	16400	3.19E+09
6	1/8/2016	16400	16250	-1.4E+09
7	1/11/2016	16250	16000	-7.1E+08
8	1/12/2016	16050	16550	2.05E+10
9	1/13/2016	16575	16375	-1.8E+09
10	1/14/2016	16200	16450	-9.5E+09
11	1/15/2016	16500	16600	5.96E+09
12	1/18/2016	16300	17100	1.74E+09
13	1/19/2016	17250	17275	-1.9E+09
14	1/20/2016	17275	16700	3.26E+09
15	1/21/2016	16700	16375	-4.8E+09
16	1/22/2016	16600	16550	-1.8E+09
17	1/25/2016	16850	16800	5.61E+09

Gambar 3. 2 Contoh data yang sudah diubah menjadi bentuk CSV

### 3.2.2 Perancangan Pra-Pemrosesan

Pada tahap ini yang pertama memastikan seluruh data yang diambil tidak ada yang kosong. Data lalu dibagi ke dalam dua set yaitu set *training* dan set *test*. Tahap terakhir dari pra pemrosesan ini adalah *training*-set akan di-*scaling* menggunakan *min-max scaler*. Tentu maksud dilakukannya *scaling* setelah pemisahan data adalah untuk menghindari *data leakage* yang tentu menjadi masalah besar dalam *machine learning* saat mengembangkan model prediksi (Bronwlee, 2016).

### 3.2.3 Perancangan Model

Pada tahap perancangan model akan dilakukan penentuan ukuran *batch*, panjang *timestep / sequence*, dan jumlah fitur untuk model BiLSTM. Untuk perancangan model menggunakan *library keras tensorflow*. Di mana sekarang *tensorflow* telah menyediakan kelas *bidirectional*, sehingga tidak perlu menggunakan dua LSTM yang dibolak-balik arahnya.

#### A. Model Single BiLSTM

Model yang dipakai pada single BiLSTM akan dibuat dengan beberapa konfigurasi, namun untuk penggunaan single BiLSTM akan memiliki pola arsitektur yang sama yaitu seperti yang bisa dilihat pada Gambar 3.1. Model ini menggunakan satu BiLSTM dengan jumlah 128 unit, lalu menggunakan dua dense layer, yang pertama sejumlah 64 dan yang kedua adalah 1. Untuk timestep

dilakukan secara kondisional, meskipun pada akhirnya untuk masing-masing model akan memiliki satu timestep tergantung keandalannya. Dari model *single* BiLSTM ini, untuk total parameter dari setiap *layer*-nya adalah berjumlah 151,681. Angka ini didapat dari perhitungan seperti yang ada pada persamaan 3.1 sampai dengan 3.6

```

Model: "sequential_2"
-----
Layer (type)                Output Shape                Param #
-----
bidirectional_2 (Bidirectio  (None, 256)                135168
nal)
dense_4 (Dense)             (None, 64)                 16448
dense_5 (Dense)             (None, 1)                  65
-----
Total params: 151,681
Trainable params: 151,681
Non-trainable params: 0

```

Gambar 3. 3 Total parameter menggunakan *single* BiLSTM

$$bidirectional\_2 = 2 \times 4 \times (u \times (u + f) + u) \quad [3.1]$$

$$bidirectional\_2 = 2 \times 4 \times (128 \times (128 + 3) + 128) = 135.168 \quad [3.2]$$

$$dense\_4 = (a + 1) \times b \quad [3.3]$$

$$dense\_4 = (256 + 1) \times 64 = 16448 \quad [3.4]$$

$$dense\_5 = (a + 1) \times b \quad [3.5]$$

$$dense\_5 = (64 + 1) \times 1 = 65 \quad [3.6]$$

Persamaan 3.1 dan 3.2 merupakan perhitungan parameter Bidirectional LSTM , di mana  $u$  merupakan jumlah unit pada BiLSTM dan  $f$  merupakan jumlah fitur. Lalu pada perhitungan *dense\_4* yaitu persamaan 3.3 dan 3.4,  $a$  merupakan output shape dari proses sebelumnya dan  $b$  merupakan output shape dari proses berjalan begitu juga dengan *dense\_5* pada persamaan 3.5 dan 3.6.

## B. Model Stacked BiLSTM

Model yang dipakai pada *stacked* BiLSTM tidak terlalu berbeda dari *single* BiLSTM. Jika dilihat pada Gambar 3.2, yang membedakan hanya pada jumlah BiLSTM yaitu sejumlah dua, dengan konfigurasi BiLSTM pertama sebanyak 128

unit, dan BiLSTM kedua adalah 64 unit, lalu fungsi `flatten()` digunakan untuk membentuk ulang tensor agar memiliki bentuk yang sama dengan jumlah elemen yang terdapat dalam. Dense layer sama seperti model single BiLSTM yaitu berjumlah 64 dan 1. Timestep pun demikian, bergantung pada keandalannya saat prediksi. Dari model single BiLSTM ini, untuk total parameter dari setiap layernya adalah berjumlah 308.865. Angka ini didapat dari perhitungan seperti yang ada pada persamaan 3.7 sampai dengan 3.15

```

Model: "sequential_4"
-----
Layer (type)                Output Shape         Param #
-----
bidirectional_6 (Bidirectio  (None, 50, 256)     136192
nal)

bidirectional_7 (Bidirectio  (None, 128)         164352
nal)

flatten_1 (Flatten)         (None, 128)         0

dense_4 (Dense)             (None, 64)          8256

dense_5 (Dense)             (None, 1)           65
-----
Total params: 308,865
Trainable params: 308,865
Non-trainable params: 0

```

Gambar 3. 4 Total parameter menggunakan stacked BiLSTM

$$bidirectional\_6 = 2 \times 4 \times (u \times (u + f) + 2 \times u) \quad [3.7]$$

$$bidirectional\_6 = 8 \times (128 \times (128 + 3) + 2 \times 128) = 136.192 \quad [3.8]$$

$$bidirectional\_7 = 2 \times 4 \times (u \times (u + f) + 2 \times u) \quad [3.9]$$

$$bidirectional\_7 = 8 \times (64 \times (64 + 256) + 64) = 164.352 \quad [3.10]$$

$$flatten\_1 = 0 \quad [3.11]$$

$$dense\_4 = (a + 1) \times b \quad [3.12]$$

$$dense\_4 = (128 + 1) \times 64 = 8256 \quad [3.13]$$

$$dense\_5 = (a + 1) \times b \quad [3.14]$$

$$dense\_5 = (64 + 1) \times 1 = 65 \quad [3.15]$$

Persamaan 3.7 dan 3.8 merupakan perhitungan parameter Bidirectional LSTM yang pertama, di mana  $u$  merupakan jumlah unit pada BiLSTM dan  $f$  merupakan jumlah

fitur. Selanjutnya masuk pada Bidirectional LSTM kedua pada persamaan 3.9 sampai dengan 3.10,  $u$  adalah jumlah unit LSTM pada BiLSTM kedua, dan  $f$  adalah jumlah fitur. Lalu pada perhitungan `dense_4` yaitu persamaan 3.12 dan 3.13,  $a$  merupakan output shape dari proses sebelumnya dan  $b$  merupakan output shape dari proses berjalan begitu juga dengan `dense_5` pada persamaan 3.14 dan 3.15.

#### 3.2.4 Perancangan Evaluasi

Setelah didapatkan hasil prediksi harga saham dari model BiLSTM, perlu dilakukan evaluasi hasil untuk mengetahui seberapa akurat prediksi tersebut. Dalam penelitian ini menggunakan beberapa metrik evaluasi seperti *Mean Absolute Error* (MAE), *Root Mean Squared Error* (RMSE), dan *Root Mean Square Percentage Error* (RMSPE) untuk mengevaluasi hasil tersebut. dengan *output* hasil adalah grafik sebagai pembandingan prediksi dengan data historis yang sebenarnya (asli). Jika hasil evaluasi kurang baik atau dirasa belum puas, maka akan dilakukan kembali perancangan model.

#### 3.2.5 Perancangan Antar Muka

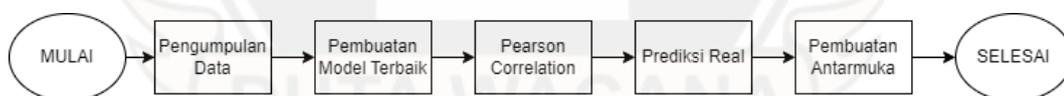
Tahap terakhir dalam proses ini adalah perancangan antarmuka menggunakan Streamlit. Antarmuka ini dirancang dengan tujuan untuk memberikan kemudahan dalam membaca data atau hasil prediksi tanpa perlu bingung dengan coding yang ada di dalam workspace. Dengan menggunakan Streamlit, pengguna dapat dengan mudah mengakses informasi yang relevan, menavigasi antarmuka secara intuitif, dan melihat visualisasi data secara langsung. Hal ini memungkinkan para pengguna, terutama yang tidak memiliki latar belakang teknis yang kuat, untuk tetap terlibat dalam proses analisis data dan memahami hasil prediksi tanpa kesulitan. Dalam perancangan antarmuka ini, pengaturan tata letak, navigasi, dan tampilan data yang jelas dan responsif sangat diperhatikan agar pengalaman pengguna menjadi lebih nyaman dan efisien.



Gambar 3. 5 Rancangan UI Streamlit

Gambar 3.3 merupakan rancangan UI Streamlit, tampilan dibuat sesederhana mungkin. Terdiri dari judul dari web, disclaimer, drop down untuk memilih emiten saham yang hendak diprediksi yang mana user juga bisa mengetik tanpa perlu scrolling untuk mencari emiten saham yang hendak diprediksi, lalu ada slider yang berguna untuk menyesuaikan jumlah hari yang hendak diprediksi, di bawahnya terdapat tombol prediksi, lalu outputnya adalah grafik dan tabel harga prediksi.

### 3.3 Diagram Alir



Gambar 3. 6 Diagram Alir Penelitian

Gambar 3.3 adalah merupakan diagram alir penelitian mengenai tahap-tahapan yang akan dilalui dalam penelitian ini. Yang pertama dilakukan adalah pengumpulan data historis harga saham yang terdiri dari harga open, close, dan juga netforeign. Untuk data historis saham yang diambil merupakan seluruh emiten saham yang telah masuk ke dalam indeks KOMPAS100 periode Februari 2023 sampai dengan Juli 2023 dengan rentang data dari tahun 2016 hingga tahun 2022.

Setelah data historis telah berhasil dikumpulkan maka tahap selanjutnya adalah pembuatan model terbaik, di mana pada bagian ini akan melalui beberapa tahap yaitu yang pertama Perbandingan BiLSTM dengan LSTM. Untuk alur dalam tahap ini secara rinci dijelaskan pada poin 3.3.1. Setelah dilakukan perbandingan antara BiLSTM dengan LSTM, akan didapati emiten saham apa saja yang ternyata lebih baik menggunakan BiLSTM dan lebih baik menggunakan LSTM. Setelah proses perbandingan antara BiLSTM dengan LSTM, masuk ke tahap selanjutnya yaitu Perbandingan Foreign Flow. Untuk alur dalam tahap ini secara rinci dijelaskan pada poin 3.3.2. Di tahap ini, emiten saham akan dibandingkan dalam prediksi menggunakan fitur Netforeign, tentu ada jenis saham yang berpengaruh terhadap penggunaan fitur Netforeign dan ada pula emiten saham yang tidak atau kurang berpengaruh jika menggunakan fitur Netforeign. Tahap selanjutnya adalah Perbandingan Stacking. Di tahap ini akan dites emiten saham apa saja yang cocok jika menggunakan model yang semakin dalam (*deeper / complex*) dan model yang sederhana atau memiliki jumlah hyperparameter yang lebih sedikit. Selanjutnya, dilakukan penyesuaian timestep yang melalui beberapa proses pembuktian menggunakan beberapa timestep untuk mendapatkan timestep terbaik untuk masing-masing emiten saham.

Setelah berhasil menciptakan model terbaik, langkah selanjutnya adalah memasuki tahap prediksi real atau yang sebenarnya. Dalam tahap ini, model akan diterapkan pada data yang sudah disiapkan untuk proses testing untuk menghasilkan prediksi yang relevan dan berguna. Setelah tahap prediksi selesai, langkah terakhir adalah pembuatan antarmuka yang memudahkan pengguna dalam mengakses dan memahami hasil prediksi. Antarmuka ini dirancang untuk menampilkan hasil prediksi dengan cara yang intuitif dan mudah dimengerti, sehingga pengguna dapat dengan cepat melihat dan memanfaatkan informasi yang diberikan oleh model. Dengan antarmuka yang efektif, pengguna dapat dengan mudah mengeksplorasi dan memanfaatkan hasil prediksi dengan lebih baik, tanpa perlu terlibat dalam detail teknis yang kompleks.

### **3.3.1 Perbandingan BiLSTM dengan LSTM**

Pada tahap perbandingan BiLSTM dengan LSTM ini, data-data dari masing-masing emiten saham akan dilatih untuk melakukan prediksi menggunakan BiLSTM dan LSTM di mana target akhirnya adalah nilai RMSPEnya untuk membandingkan hasil dalam penggunaan dua algoritma tersebut. Pada setiap emiten saham dilakukan Pra pemrosesan Supaya tidak ada data yang tidak berisi, selanjutnya dilakukan *checking* data dan pembersihan data. Lalu data akan dibagi menjadi 2 set yaitu *training* set 70% dan *validation* set 30%. Dengan fitur yang diambil adalah harga *open*, harga *closing*, dan *foreign flow*. Setelah data siap, selanjutnya adalah pembangunan dan pelatihan model dilanjutkan dengan pengujian hasil prediksi dengan *test*-set. Langkah selanjutnya adalah evaluasi hasil menggunakan *RMSPE*. Langkah terakhir tahap ini adalah membandingkan emiten saham mana saja yang lebih unggul ketika menggunakan BiLSTM. Perbandingan dilakukan dengan menggunakan timestep 30 dan 60.

### 3.3.2 Perbandingan Foreign Flow

Pada tahap perbandingan Foreign flow, data-data dari masing-masing emiten saham tersebut akan dilatih untuk melakukan prediksi menggunakan dengan dan tanpa Netforeign di mana target akhirnya juga sama yaitu nilai RMSPE untuk membandingkan hasil dalam penggunaan dan tanpa penggunaan Netforeign. Untuk langkahnya kurang lebih sama seperti pada tahap perbandingan BiLSTM dengan LSTM di mana setiap emiten saham adalah dilakukan terlebih dahulu Pra pemrosesan, beda terdapat pada bagian ini yaitu fitur yang diambil terbagi menjadi dua yaitu yang pertama menggunakan harga *open*, harga *closing*, dan *Netforeign*, lalu yang kedua adalah tanpa Netforeign, jadi hanya harga *open* dan *closing*. Supaya tidak terdapat data yang *null*, maka dilakukan *checking* dan *cleaning* data. Lalu data dibagi menjadi 2 set yaitu *training* set 70% dan *validation* set 30%. Selanjutnya adalah pembangunan dan pelatihan model di mana timestep sudah disesuaikan menggunakan timestep terbaik dari masing-masing emiten saham antara 30 dan 60. Lalu dilanjutkan dengan pengujian hasil prediksi dengan *test*-set. Langkah selanjutnya adalah evaluasi hasil menggunakan *RMSPE*. Terakhir,

membandingkan emiten saham apa saja yang lebih unggul ketika menggunakan fitur Netforeign.



## BAB IV

### IMPLEMENTASI DAN PEMBAHASAN

#### 4.1 Implementasi Awal

##### 4.1.1 Pengumpulan Data

Penelitian ini menggunakan Indeks KOMPAS100 terbaru yaitu periode Februari 2023 sampai dengan Juli 2023 disebabkan karena Indeks KOMPAS100 secara berkala diperbarui. Data awalnya berisi seluruh saham yang diperdagangkan di Bursa Efek Indonesia dalam format SQL dengan jumlah 1.146.458 (satu juta seratus empat puluh enam ribu empat ratus lima puluh delapan) baris terbagi dalam 3(tiga) tabel. Data lalu diimpor ke dalam phpMyAdmin. Pada tahap ini, dilakukan *filtering* untuk setiap emiten saham yang tergabung dalam indeks KOMPAS100. Dikarenakan phpMyAdmin adalah *web-based application* pengelola basis data MySQL, Filtering dilakukan menggunakan syntax SQL dengan contoh seperti pada listing 4.1.

```
SELECT date, open, close, netforeign FROM `quote` WHERE ticker =  
'AALI' AND date BETWEEN '2016-01-04' and '2022-12-30'
```

Listing 4.1 Kode SQL untuk memperoleh data penelitian

Dapat dilihat pada Listing 4.1, merupakan contoh *filtering* pada emiten saham AALI di mana data yang diambil adalah *date*, *open*, *close*, dan *netforeign* dari tanggal pembukaan Bursa Efek Indonesia tahun 2016 yaitu pada 4 Januari 2016 hingga penutupan Bursa Efek Indonesia tahun 2022 pada 30 Desember 2022. Hasil dari *filtering* tersebut lalu diekspor menjadi format *Comma Separated Value* (CSV) untuk memudahkan penelitian. Pada prosesnya, tidak semua emiten saham datanya tersedia dari 4 Januari 2016 dengan kata lain terdapat beberapa emiten yang IPO lebih dari tahun 2016 antara lain ADMR, AGII, ARTO, AVIA, BRIS, BTPS, BUKA, GOTO, HEAL, IPPE, MTEL, MTMH, TAPG, RMKE, SGER, dan WIFI. Pada akhirnya terdapat 84 emiten saham yang masing-masing berjumlah 2700 baris data (termasuk header).

### 4.1.2 Pra-pemrosesan Data

Tahap pertama pada bagian pra-pemrosesan data ini adalah melakukan beberapa penyesuaian terhadap data yang telah dilakukan filtering.

	date	open	close	netforeign
0	2016-01-04	2635	2645	-24324600000
1	2016-01-05	2640	2675	-9945780000
2	2016-01-06	2665	2640	-120409000000
3	2016-01-07	2620	2600	-198300000000
4	2016-01-08	2590	2600	-179818000000
...	...	...	...	...
1695	2022-12-26	8500	8575	-11802900000
1696	2022-12-27	8600	8600	67042500000
1697	2022-12-28	8650	8650	-39759100000
1698	2022-12-29	8525	8575	-141717000000
1699	2022-12-30	8575	8550	-71641200000

Gambar 4. 1 Contoh data historis saham yang telah dilakukan filtering

Seperti yang tertera pada Gambar 4.1 kolom date tidak berfungsi sebagai indeks. Sedangkan pada penelitian ini memerlukan date sebagai indeks untuk memudahkan proses analisis.

```
df = pd.read_csv('E:\Skripsi\Dataset\BBCA.csv',  
parse_dates=['date'], index_col=['date'])  
df.dropna(axis="rows", how="any", inplace=True)
```

Listing 4.2 Kode untuk mengubah kolom menjadi indeks

	open	close	netforeign
date			
2016-01-04	2635	2645	-24324600000
2016-01-05	2640	2675	-9945780000
2016-01-06	2665	2640	-120409000000
2016-01-07	2620	2600	-198300000000
2016-01-08	2590	2600	-179818000000
...	...	...	...
2022-12-26	8500	8575	-11802900000
2022-12-27	8600	8600	67042500000
2022-12-28	8650	8650	-39759100000
2022-12-29	8525	8575	-141717000000
2022-12-30	8575	8550	-71641200000

Gambar 4. 2 Contoh data yang telah dilakukan perubahan indeks

Menggunakan kode Listing 4.2, data yang terdapat pada kolom date dapat diubah menjadi indeks seperti dapat dilihat pada Gambar 4.2. Serta penggunaan fungsi dropna() untuk menghapus nilai yang hilang atau null. Tentu langkah penghapusan null ini sangat penting dilakukan karena nilai null dapat memengaruhi kinerja model LSTM dan dapat menyebabkan kesalahan pada proses pelatihan dan prediksi. Dataset tersebut perlu dilakukan proses pemotongan, di mana data dari tahun 2016 hingga 2021 akan dipakai untuk proses training dan validasi. Sedangkan data tahun 2022 akan disimpan dan dipakai nantinya pada saat proses testing.

	open	close	netforeign
count	1700.000000	1700.000000	1.700000e+03
mean	5403.461765	5403.785294	7.105663e+09
std	1717.974421	1719.830906	1.763488e+11
min	2530.000000	2550.000000	-1.366290e+12
25%	4047.500000	4063.750000	-6.373192e+10
50%	5560.000000	5545.000000	4.191810e+09
75%	6591.250000	6600.000000	8.106325e+10
max	9275.000000	9300.000000	9.638280e+11

Gambar 4. 3 Data statistik pada emiten BCA

Sedikit melihat ke data statistik yang ada pada dataset, secara kasat mata pada Gambar 4.3, tipe angka yang ada pada kolom open dan close berbeda dengan kolom netforeign, di mana open dan close memiliki angka ribuan sedangkan netforeign memiliki angka positif hingga minus. Oleh sebab itu perlu dilakukan yang namanya scaling. Proses scaling juga merupakan salah satu teknik untuk mengoptimalkan kinerja model. Scaling dapat meningkatkan kecepatan pelatihan dan stabilitas numerik. Pada penelitian kali ini menggunakan MinMax Scaler. Setelah data discaling, maka selanjutnya data dibagi menjadi training-set sebesar 70% dan 30% sisanya digunakan sebagai validation-set.

#### 4.1.3 Pembuatan Input Fitur

Langkah selanjutnya adalah membuat fitur yang akan menjadi input dalam pelatihan data. Pembuatan input fitur ini perlu dilakukan karena jika tidak demikian, proses pelatihan tidak akan berjalan karena input fitur harus memiliki format yang cocok dengan ketentuan pelatihan. Proses pembuatan input fitur pada data timeseries berbeda dengan clustering dan classification karena timeseries berbentuk regression yang menghasilkan prediksi berdasarkan pola berkelanjutan.

```
def create_dataset(timestep, data):  
    x, y = [], []  
    for i in range(timestep, data.shape[0]):  
        x.append(data[i- timestep:i,:])  
        y.append(data[i, close_sliced])  
    x = np.array(x)  
    y = np.array(y)  
    return x, y
```

Listing 4.3 Fungsi untuk membuat input fitur

Listing 4.3 merupakan fungsi yang dipakai untuk membuat input fitur. Fungsi `create_dataset` memiliki dua parameter, yaitu `timestep` dan `data`. Fungsi ini bertujuan untuk membuat dataset yang akan digunakan untuk pelatihan model. Pada bagian pertama fungsi, variabel `x` dan `y` dideklarasikan sebagai list kosong. Variabel `x` akan digunakan untuk menyimpan sekumpulan data input (sering disebut juga

sebagai features), sedangkan variabel  $y$  akan digunakan untuk menyimpan data output (sering disebut juga sebagai target). Selanjutnya, dilakukan perulangan for dari timestep hingga sebanyak `data.shape[0]`. Perulangan ini akan mengiterasi setiap baris data, dimulai dari baris ke- $i$ . Hal ini berguna untuk membuat input dan output terkait dengan timestep yang telah ditentukan. Di dalam perulangan, baris data dari  $i$  dikurangi timestep hingga  $i$  (sejumlah timestep baris) diambil dari data dan ditambahkan ke dalam list  $x$ . Baris data pada indeks  $i$  untuk kolom `close_sliced` diambil dan ditambahkan ke dalam list  $y$ . Dengan demikian, setiap iterasi perulangan akan menambahkan input dan output yang sesuai ke dalam list  $x$  dan  $y$ . Setelah selesai perulangan, list  $x$  dan  $y$  diubah menjadi array numpy menggunakan `np.array(x)` dan `np.array(y)`. Hal ini penting untuk mempermudah manipulasi dan pelatihan model menggunakan library numpy. Terakhir, fungsi mengembalikan nilai  $x$  dan  $y$  dalam bentuk array numpy sebagai hasil dari proses pembuatan dataset.

## 4.2 Implementasi Sistem

### 4.2.1 Pembuatan Model Terbaik

#### 4.2.1.1 Perbandingan BiLSTM dengan LSTM

Dalam praktiknya, terdapat jenis saham yang berhasil (memiliki hasil pelatihan model yang memuaskan) terhadap jenis algoritma A namun kurang atau bahkan tidak berhasil di algoritma B, ada pula emiten saham yang berhasil di algoritma B namun tidak berhasil di algoritma A. Oleh sebab itu, tahap pertama pembuatan model ini adalah memisahkan antara emiten saham yang baik saat memakai BiLSTM dengan emiten saham yang baik saat memakai LSTM.

```
model = Sequential()  
model.add(Bidirectional(LSTM(epoch, input_shape=(X_train.shape[1],  
X_train.shape[2]))))  
model.add(Dense(64))  
model.add(Dense(1))  
model.compile(loss='mean_squared_error', optimizer='adam')
```

Listing 4.4 Model pertama BiLSTM

```

model = Sequential()
model.add(LSTM(epoch, input_shape=(X_train.shape[1],
X_train.shape[2])))
model.add(Dense(64))
model.add(Dense(1))
model.compile(loss='mean_squared_error', optimizer='adam')

```

Listing 4.5 Model pertama LSTM

Adapun untuk menghasilkan perbandingan yang adil, model yang dipakai memiliki jumlah *hyperparameter* yang sama persis seperti yang bisa dilihat pada Listing 4.4 dan Listing 4.5, perbedaan hanya terletak pada algoritma yang dipakai. Batch size dibarkan default yaitu 128. Dengan kata lain, proses perbandingan menggunakan model yang sama, namun perbandingan tidak hanya dilakukan sekali jadi, melainkan dua kali dengan pembeda adalah timestep. Ini juga dilakukan untuk melihat saham mana saja yang berkorelasi jangka panjang dan jangka menengah. Dalam hal ini, timestep yang dipakai adalah 30, 60, dan 90. Pemilihan timestep salah satu hal yang tidak kalah penting (BERGSTRÖM & HJELM, 2019) karena timestep mengacu pada jumlah langkah waktu atau urutan yang diberikan sebagai input ke LSTM pada setiap iterasi. Untuk metrik akurasi yang digunakan adalah RMSPE. Penggunaan RMSPE ini didasari karena RMSPE adalah merupakan ukuran akurasi yang baik untuk membandingkan kesalahan peramalan dari model yang berbeda atau konfigurasi model untuk variabel tertentu karena bergantung pada skala serta untuk alasan praktis adalah agar mudah membandingkannya karena dalam bentuk persen.

Tabel 4. 1 Hasil rata-rata RMSPE perbandingan BiLSTM dengan LSTM

No.	Kode Emiten	Rata-rata BiLSTM	Rata-rata LSTM	No.	Kode Emiten	Rata-rata BiLSTM	Rata-rata LSTM
1	AALI	3.00	3.06	43	INDY	4.40	4.49
2	ABMM	5.96	4.87	44	INKP	4.22	3.92
3	ACES	2.58	2.72	45	INTP	3.44	2.82
4	ADHI	5.88	5.36	46	ISAT	4.27	4.11
5	ADRO	3.00	3.30	47	ITMG	3.82	3.33
6	AGRO	6.09	6.50	48	JPFA	3.41	2.88
7	AKRA	2.91	2.89	49	JSMR	3.03	2.93

No.	Kode Emiten	Rata-rata BiLSTM	Rata-rata LSTM	No.	Kode Emiten	Rata-rata BiLSTM	Rata-rata LSTM
8	AMRT	3.36	3.00	50	KLBF	2.15	2.33
9	ANTM	4.18	4.22	51	LPPF	6.13	12.82
10	APIC	2.27	2.62	52	LSIP	2.67	2.90
11	ASII	2.62	2.50	53	MAPI	3.27	3.02
12	ASSA	5.88	5.25	54	MDKA	4.95	4.16
13	BABP	7.70	7.24	55	MEDC	7.60	5.06
14	BBCA	1.87	1.76	56	MIKA	2.36	2.49
15	BBHI	15.30	15.53	57	MLPL	6.59	6.72
16	BBKP	6.69	6.50	58	MNCN	2.93	2.88
17	BBNI	2.76	2.83	59	MPMX	3.43	3.52
18	BBRI	2.36	2.32	60	MYOR	2.04	2.06
19	BBTN	3.70	3.22	61	PGAS	3.73	3.46
20	BFIN	4.40	4.28	62	PNBN	2.39	2.39
21	BMRI	2.46	2.30	63	PNLF	2.19	2.10
22	BMTR	3.93	4.95	64	PTBA	2.72	2.82
23	BRMS	4.34	4.15	65	PTPP	4.52	5.40
24	BRPT	4.59	5.85	66	PWON	3.39	2.84
25	BSDE	3.09	3.10	67	RAJA	7.22	5.02
26	BSSR	4.39	4.32	68	SCMA	5.19	3.84
27	CPIN	3.02	2.79	69	SIDO	2.40	2.07
28	CTRA	3.14	3.29	70	SMDR	4.95	5.15
29	DOID	5.31	5.41	71	SMGR	2.71	2.65
30	DSNG	3.15	3.00	72	SMRA	3.72	3.43
31	ELSA	3.33	3.33	73	SRTG	5.16	5.40
32	EMTK	45.21	47.28	74	TBIG	3.44	3.21
33	ENRG	4.38	5.03	75	TINS	3.87	3.93
34	ERAA	2.93	3.33	76	TKIM	4.61	4.43
35	ESSA	9.35	10.90	77	TLKM	2.67	2.44
36	EXCL	2.71	2.59	78	TOBA	5.55	5.54
37	GGRM	2.94	2.62	79	TOWR	2.74	2.54
38	HMSP	3.84	4.55	80	TPIA	3.82	3.75
39	HRUM	5.98	5.83	81	UNTR	2.82	2.86
40	ICBP	1.78	1.77	82	UNVR	5.75	4.40
41	INCO	3.02	3.34	83	WIKA	3.91	3.77
42	INDF	1.88	1.73	84	WSKT	6.58	5.62

Perbandingan BiLSTM dengan LSTM pada percobaan pertama menghasilkan rata-rata RMSPE untuk LSTM adalah 4,54 sedangkan menggunakan BiLSTM sebesar 4,56. Lalu pada percobaan tahap 2, menghasilkan rata-rata

RMSPE untuk LSTM adalah 4,56 sedangkan BiLSTM sebesar 4,55. Lalu pada percobaan ketiga nilai RMSPE pada LSTM adalah 4,57 sedangkan dari BiLSTM 4,60. Dari ketiga percobaan tersebut dicari rata-rata tiap emiten untuk penggunaan BiLSTM dan LSTM, seperti yang bisa dilihat pada tabel 4.1, didapati bahwa dari perhitungan rata-rata, sebanyak 35 emiten saham dinilai lebih baik saat menggunakan BiLSTM dan 49 emiten saham dinilai lebih baik saat menggunakan LSTM. Adapun dari ketiga percobaan tersebut dicari emiten yang lebih baik saat menggunakan 30 timestep, 60 timestep, dan 90 timestep. Sebanyak 23 emiten saham lebih unggul saat menggunakan 30 timestep, 32 emiten saham lebih baik saat menggunakan timestep 60, dan 29 emiten saham lebih unggul saat menggunakan 90 timestep. Secara keseluruhan perbandingan RMSPE dari BiLSTM adalah 4,57 dan LSTM di angka 4,56.

#### 4.2.1.2 Perbandingan Foreign Flow

Pada kenyataannya tentu terdapat pula emiten saham yang secara rutin diperjualbelikan oleh asing baik itu dalam frekuensi sering maupun jarang, namun ada pula yang sama sekali tidak dilirik oleh asing. Oleh sebab itu tahap selanjutnya adalah melihat dan memisahkan antara emiten saham yang berelasi secara langsung dengan Foreign Flow dan yang tidak berelasi secara langsung dengan Foreign Flow sekaligus menemukan apakah saham yang tidak pernah absen dijualbelikan asing menghasilkan model yang lebih akurat dalam prediksi. Model yang dipakai kurang lebih sama dengan model yang digunakan pada tahap sebelumnya, hal yang membedakan terletak pada tahap awal yaitu preprocessing data di mana fitur yang diambil dibagi menjadi dua bagian, percobaan pertama memakai fitur Netforeign dan percobaan kedua tidak memakai fitur Netforeign (dapat dilihat pada Listing 4.6).

```
df = pd.read_csv('E:\Skripsi\Dataset\BBCA.csv',
parse_dates=['date'], index_col=['date'])
df.dropna(axis="rows", how="any", inplace=True)
df = df[['open', 'close']]
```

Listing 4.6 Pelepasan Netforeign dari Dataset

Tabel 4. 2 Hasil rata-rata RMSPE perbandingan Foreign Flow

No.	Kode Emiten	Foreign Flow	Rata-rata tanpa FF	No.	Kode Emiten	Tanpa Foreign Flow	Rata-rata tanpa FF
1	AALI	2.95	1.17	43	INDY	4.75	2.64
2	ABMM	5.53	2.36	44	INKP	3.69	1.41
3	ACES	2.91	1.74	45	INTP	3.21	0.91
4	ADHI	5.84	2.65	46	ISAT	4.12	2.28
5	ADRO	2.93	1.33	47	ITMG	3.52	1.35
6	AGRO	6.55	3.43	48	JPFA	3.38	1.22
7	AKRA	3.16	1.54	49	JSMR	2.66	1.65
8	AMRT	3.17	1.34	50	KLBF	2.12	1.03
9	ANTM	4.13	1.86	51	LPPF	5.94	2.74
10	APIC	2.19	1.21	52	LSIP	2.72	1.27
11	ASII	2.52	1.19	53	MAPI	2.99	1.38
12	ASSA	5.17	1.96	54	MDKA	3.92	1.08
13	BABP	6.87	2.73	55	MEDC	5.95	3.54
14	BBCA	2.19	1.00	56	MIKA	2.41	1.61
15	BBHI	13.14	4.65	57	MLPL	6.36	2.82
16	BBKP	6.62	3.20	58	MNCN	2.78	1.67
17	BBNI	2.66	1.12	59	MPMX	3.13	1.60
18	BBRI	2.37	1.13	60	MYOR	1.89	0.93
19	BBTN	3.38	1.42	61	PGAS	3.40	1.27
20	BFIN	4.40	2.34	62	PNBN	2.53	1.21
21	BMRI	2.27	0.99	63	PNLF	2.10	1.06
22	BMTR	3.72	2.14	64	PTBA	2.58	1.64
23	BRMS	4.22	2.13	65	PTPP	3.78	1.65
24	BRPT	5.93	2.35	66	PWON	3.14	1.11
25	BSDE	2.97	1.34	67	RAJA	5.05	2.69
26	BSSR	3.91	1.30	68	SCMA	4.95	1.02
27	CPIN	2.52	1.01	69	SIDO	2.29	1.16
28	CTRA	3.17	1.47	70	SMDR	4.59	1.63
29	DOID	4.15	1.28	71	SMGR	2.80	1.34
30	DSNG	3.00	1.39	72	SMRA	3.28	1.74
31	ELSA	3.30	1.47	73	SRTG	4.50	1.47
32	EMTK	45.51	22.73	74	TBIG	3.23	1.62
33	ENRG	4.92	2.35	75	TINS	4.08	2.39
34	ERAA	3.45	3.07	76	TKIM	4.23	2.14
35	ESSA	7.17	1.95	77	TLKM	2.55	0.89
36	EXCL	2.65	1.63	78	TOBA	5.60	2.78
37	GGRM	2.78	1.30	79	TOWR	2.23	1.46
38	HMSP	4.27	2.20	80	TPIA	3.22	0.74

No.	Kode Emiten	Foreign Flow	Rata-rata tanpa FF	No.	Kode Emiten	Tanpa Foreign Flow	Rata-rata tanpa FF
39	HRUM	4.77	3.13	81	UNTR	2.79	1.55
40	ICBP	1.71	0.92	82	UNVR	4.07	6.35
41	INCO	3.36	1.97	83	WIKA	3.57	2.11
42	INDF	1.73	0.77	84	WSKT	5.23	2.06

Dari hasil rata-rata percobaan yang dilakukan tiga kali menggunakan Foreign Flow dan tiga kali tanpa menggunakan Foreign Flow yang dapat dilihat pada Tabel 4.2, ditemukan bahwa 21 emiten saham hasilnya lebih baik saat menggunakan Foreign Flow.

#### 4.2.1.3 Perbandingan Stacking

Perbandingan stacking dilakukan untuk memisahkan emiten saham mana yang menghasilkan hasil prediksi lebih baik saat menggunakan model yang rumit dan model yang sederhana. Perbandingan stacking ini menggunakan model yang mirip seperti pada percobaan Perbandingan BiLSTM dengan LSTM sebelumnya. Jika dilihat pada Listing 4.6, yang menjadi pembeda dari model BiLSTM sebelumnya terletak pada jumlah tumpukan BiLSTM, di mana pada percobaan stacking ini menggunakan 2 BiLSTM yang bertumpuk.

```

model = Sequential()
model.add(Bidirectional(LSTM(128, return_sequences=True,
input_shape=(X_train.shape[1], X_train.shape[2])))
model.add(Bidirectional(LSTM(64, input_shape=(X_train.shape[1],
X_train.shape[2])))
model.add(Flatten())
model.add(Dense(64))
model.add(Dense(1, activation='sigmoid'))
model.compile(loss='mean_squared_error', optimizer='adam')

```

Listing 4.6 Model stacked BiLSTM

Pada tahap ini, untuk hasil evaluasi RMSPE pada penggunaan 1(satu) BiLSTM diambil dari hasil perbandingan BiLSTM dan LSTM pada tahap perbandingan foreign flow. Sedangkan untuk percobaan stacking dilakukan percobaan baru dalam menggunakan 2 (dua) BiLSTM. Karena cara model belajar

tidak multak, dengan kata lain bisa saja model belajar baik saat evaluasi pertama namun lebih buruk saat evaluasi kedua, oleh sebab itu percobaan dilakukan tiga kali.

Tabel 4. 3 Hasil rata-rata RMSPE perbandingan Stacking

No.	Kode Emiten	Single	Stacked	No.	Kode Emiten	Single	Stacked
1	AALI	3.06	3.89	43	INDY	4.73	4.80
2	ABMM	5.36	8.02	44	INKP	5.49	3.85
3	ACES	2.58	2.96	45	INTP	2.77	3.01
4	ADHI	5.77	6.84	46	ISAT	4.23	4.69
5	ADRO	2.91	3.11	47	ITMG	3.45	3.42
6	AGRO	7.21	6.92	48	JPFA	3.19	3.20
7	AKRA	2.81	3.58	49	JSMR	3.15	2.97
8	AMRT	3.07	3.57	50	KLBF	2.18	2.74
9	ANTM	4.53	4.34	51	LPPF	10.46	5.73
10	APIC	2.19	3.00	52	LSIP	2.78	2.87
11	ASII	2.62	2.79	53	MAPI	3.14	3.02
12	ASSA	5.81	6.51	54	MDKA	4.37	5.86
13	BABP	7.21	7.94	55	MEDC	6.66	9.73
14	BBCA	1.67	2.02	56	MIKA	2.38	2.36
15	BBHI	14.93	18.15	57	MLPL	6.02	7.63
16	BBKP	5.99	7.02	58	MNCN	2.91	2.96
17	BBNI	3.07	2.88	59	MPMX	3.72	4.11
18	BBRI	2.41	2.34	60	MYOR	2.21	2.16
19	BBTN	3.35	3.61	61	PGAS	5.86	6.32
20	BFIN	4.24	4.54	62	PNBN	2.44	2.43
21	BMRI	2.21	2.36	63	PNLF	2.73	2.35
22	BMTR	4.04	4.88	64	PTBA	2.57	3.70
23	BRMS	4.22	4.45	65	PTPP	4.59	4.45
24	BRPT	6.99	4.72	66	PWON	3.41	3.05
25	BSDE	2.97	3.51	67	RAJA	6.95	5.05
26	BSSR	4.41	4.88	68	SCMA	4.14	4.06
27	CPIN	2.92	2.66	69	SIDO	1.97	2.13
28	CTRA	3.21	3.24	70	SMDR	5.61	5.16
29	DOID	4.15	6.47	71	SMGR	2.86	2.72
30	DSNG	3.07	3.86	72	SMRA	3.35	3.73
31	ELSA	3.66	3.09	73	SRTG	5.95	6.87
32	EMTK	43.67	47.59	74	TBIG	3.65	3.68
33	ENRG	4.22	5.56	75	TINS	4.07	4.64
34	ERAA	3.06	4.10	76	TKIM	4.17	3.80
35	ESSA	8.69	10.41	77	TLKM	2.44	2.71

No.	Kode Emiten	Single	Stacked	No.	Kode Emiten	Single	Stacked
36	EXCL	3.08	3.27	78	TOBA	5.25	5.67
37	GGRM	2.75	3.29	79	TOWR	4.24	2.59
38	HMSP	4.71	5.64	80	TPIA	3.19	3.66
39	HRUM	5.58	6.79	81	UNTR	2.99	3.18
40	ICBP	1.84	2.12	82	UNVR	3.40	5.46
41	INCO	3.27	3.91	83	WIKA	4.32	3.76
42	INDF	1.72	1.83	84	WSKT	7.76	10.34

Hasil rata-rata dari percobaan Stacked BiLSTM yang bisa dilihat pada Tabel 4.3, penggunaan yang ditumpuk menghasilkan hanya 25 emiten saham yang lebih baik menggunakan Stacking, ini menunjukkan bahwa mayoritas emiten saham lebih baik saat menggunakan single BiLSTM. Adapun perbandingan single BiLSTM dengan stacked BiLSTM dapat menentukan emiten saham yang lebih cocok menggunakan model yang lebih rumit atau model yang lebih sederhana. Hal ini dikarenakan bila kita melihat pada jumlah parameter, stacked BiLSTM tentu memiliki jauh lebih banyak parameter dalam modelnya.

#### 4.2.1.4 Penyimpanan Model

```
import joblib
filename = "BBCA.joblib"
joblib.dump(model, filename)
```

Listing 4.7 Code untuk menyimpan model

Setelah model berhasil dibuat, tahap selanjutnya adalah menyimpan model agar model dapat digunakan kembali untuk prediksi. Penyimpanan model ini menggunakan library Python bernama Joblib. Dibandingkan dengan teknik lain untuk menyimpan dan memuat model Machine Learning / Deep Learning, ada beberapa keuntungan dari penggunaan Joblib. Data dapat disimpan dengan cepat dan mudah dalam jumlah ruang penyimpanan(besar file) yang lebih kecil daripada penyimpanan konvensional, ini dikarenakan data disimpan sebagai string byte daripada objek. Selain itu, Joblib secara otomatis mengoreksi kesalahan saat membaca atau menulis file (Sharma, 2023). Pada Listing 4.7 bisa dilihat teknis

pemakaian joblib dalam penyimpanan model. Diawali dengan mengimpor library joblib, menyimpan nama file dalam variabel, lalu terakhir adalah menggunakan fungsi joblib.dump() untuk menyimpan model yang sudah disimpan sebelumnya pada variabel model.

#### 4.2.1.5 Model Terbaik

Pada akhir proses pembuatan model, didapati model terbaik untuk masing-masing emiten saham, untuk selebihnya dapat dilihat pada lampiran A.3. Pada prosesnya juga ditemukan emiten saham terbaik bila ditinjau dari modelnya menggunakan metode evaluasi RMSPE. 5 saham terbaik dari segi model dapat dilihat pada Tabel 4.4.

Tabel 4. 4 Emiten saham dengan model terbaik

Kode Emiten	MAE	RMSE	RMSPE
BBCA	78.04	105.63	1.62
INDF	86.17	118.27	1.74
ICBP	115.78	167.39	1.83
SIDO	9.87	14.22	1.83
MYOR	34.77	46.52	1.90

#### 4.2.2 Perhitungan Pearson Correlation Fitur Close dan Netforeign

```
import pandas as pd
df = pd.read_csv(f, parse_dates=['date'], index_col=['date'])
df.dropna(axis="rows", how="any", inplace=True)
dataset = pd.DataFrame(df)
x, y = dataset['close'], dataset['netforeign'].cumsum()
print('Pearson Correlation = ', x.corr(y))
```

Listing 4.8 Code untuk perhitungan Pearson Correlation

Pada Listing 4.8, terlihat bahwa tahap ini membutuhkan modul Pandas. Perhitungannya sangat sederhana, hanya dengan mengambil fitur-fitur yang ingin dibandingkan, dalam hal ini fitur "close" dan "netforeign". Untuk Netforeign menggunakan fungsi cumsum() untuk mendapatkan akumulasi dari fitur Netforeign. Setelah fitur-fitur tersebut diambil, langkah selanjutnya adalah menggunakan fungsi "x.corr(y)", di mana dalam percobaan ini, fungsi tersebut

digunakan untuk mencari korelasi antara "close" dan "netforeign". Dengan melakukan perhitungan ini, dapat diketahui sejauh mana kedua fitur tersebut saling berhubungan atau berkorelasi. Korelasi ini dapat memberikan wawasan tentang bagaimana perubahan pada fitur "close" dapat berhubungan dengan perubahan pada fitur "netforeign".

Tabel 4. 5 Hasil perhitungan Pearson Correlation

No.	Kode Emiten	Pearson Correlation	No.	Kode Emiten	Pearson Correlation
1	AALI	0.553862501905	43	INDY	0.676039827316
2	ABMM	-0.285401646109	44	INKP	0.673643885195
3	ACES	0.688670225534	45	INTP	0.829400054750
4	ADHI	0.746236676358	46	ISAT	-0.261839562091
5	ADRO	0.187961709957	47	ITMG	-0.214982372649
6	AGRO	0.597579818605	48	JPFA	0.117356579172
7	AKRA	0.836612893829	49	JSMR	0.653740771670
8	AMRT	0.425739667202	50	KLBF	0.423145108047
9	ANTM	0.848065312142	51	LPPF	0.748813787504
10	APIC	0.922090505994	52	LSIP	0.192551756985
11	ASII	0.702224157676	53	MAPI	0.689382892556
12	ASSA	0.113576280968	54	MDKA	0.405130217228
13	BABP	0.685833286495	55	MEDC	0.674668275279
14	BBCA	0.680945744184	56	MIKA	-0.485544341821
15	BBHI	0.596145916049	57	MLPL	-0.129008591256
16	BBKP	0.206447303625	58	MNCN	0.657237747153
17	BBNI	0.303568237057	59	MPMX	-0.207841718059
18	BBRI	0.442019088144	60	MYOR	-0.291867184118
19	BBTN	0.719448971642	61	PGAS	0.776024554984
20	BFIN	-0.791198197079	62	PNBN	-0.344729878070
21	BMRI	0.538564342591	63	PNLF	-0.370643765844
22	BMTR	0.673224499987	64	PTBA	0.675231133423
23	BRMS	0.683913365277	65	PTPP	0.779212264294
24	BRPT	-0.404911760534	66	PWON	0.618031872933
25	BSDE	0.580398624498	67	RAJA	0.547082301893
26	BSSR	0.765123415488	68	SCMA	0.773073413763
27	CPIN	-0.547992053872	69	SIDO	0.593023135212
28	CTRA	-0.408565848401	70	SMDR	0.916229199269
29	DOID	0.288476138658	71	SMGR	0.548759443833
30	DSNG	0.344775835357	72	SMRA	0.696397256024
31	ELSA	0.106285163563	73	SRTG	-0.707810570608

No.	Kode Emiten	Pearson Correlation	No.	Kode Emiten	Pearson Correlation
32	EMTK	-0.562017843709	74	TBIG	-0.499985392340
33	ENRG	-0.144288604448	75	TINS	-0.332015204733
34	ERAA	-0.735021777135	76	TKIM	0.465302249150
35	ESSA	0.230741228598	77	TLKM	0.599907088802
36	EXCL	0.392576948110	78	TOBA	0.222082490402
37	GGRM	0.885166810173	79	TOWR	-0.795835880853
38	HMSP	0.840755242535	80	TPIA	0.027816061467
39	HRUM	-0.285647474620	81	UNTR	0.237400229419
40	ICBP	0.418340502709	82	UNVR	0.733458545843
41	INCO	0.517116036910	83	WIKA	0.502097481170
42	INDF	0.735485173699	84	WSKT	0.834721604186

Berdasarkan hasil perhitungan Pearson Correlation yang dapat dilihat pada Tabel 4.4, sebanyak 8 emiten saham memiliki korelasi Positif sangat kuat, 24 emiten saham dengan korelasi positif kuat, 17 emiten saham korelasi positif sedang, 8 emiten saham korelasi positif rendah, dan 6 emiten saham korelasi positif sangat rendah. Sedangkan emiten saham dengan korelasi negatif antara lain memiliki interpretasi negatif sangat rendah sebanyak 2, lalu 9 emiten saham dengan korelasi negatif rendah, 6 emiten saham korelasi negatif sedang, serta 4 emiten dengan korelasi negatif kuat.

#### 4.2.3 Prediksi

Bagian yang paling penting di dalam penelitian ini adalah prediksi. Di mana output yang diharapkan adalah prediksi harga closing saham selama  $n$  hari. Bisa dilihat pada Listing 4.9, Bagian ini dimulai dengan menginisiasi saham yang hendak diprediksi, menyesuaikan timestep dengan timestep yang telah ditentukan, `pred_day` adalah variabel untuk menyimpan data berapa hari akan memprediksi harga saham, lalu `get_data` untuk menambah timestep dan `pred_day` di mana `get_data` diperlukan untuk mengambil seberapa banyak data yang akan diprediksi. Penambahan yang terjadi pada `get_data` juga disebabkan oleh proses prediksi, data akan dipotong sebanyak timestep, jadi untuk menyiasati tersebut, supaya jumlah outputnya seperti yang diinginkan, jumlah data untuk prediksi ditambah sebanyak

jumlah timestep. Selanjutnya adalah impor data yang dipakai sebagai input, dalam hal ini, data yang dipakai adalah data saham selama tahun 2022. Lalu dataset dikonversi menjadi bentuk pandas lalu data diambil sebanyak `get_data` dari indeks terakhir. Baru selanjutnya dilakukan scaling menggunakan `MinMaxScaler`, data yang sudah discaling lalu dibuat menjadi input fitur `X_pred` dan `y_pred`. Model yang sudah dibuat lalu diimpor menggunakan `joblib`. Lalu model akan memprediksi `X_pred` sebanyak data yang masuk. Untuk melihat hasil prediksi perlu dilakukan unscaling dengan membuat scaler baru yaitu `pred_scaler`, scaler ini lalu di `fit_transform` dengan fitur `close` supaya hasil prediksi dapat di unscaling menggunakan `pred_scaler`.

```
stock_selected = "BBCA"
timestep = 60
pred_day = 90
get_data = timestep+pred_day

loaded_model = joblib.load(f"{stock_selected}.joblib")
df = pd.read_csv(f'E:\Skripsi\Dataset\{stock_selected}.csv',
parse_dates=['date'], index_col=['date'])
df.dropna(axis="rows", how="any", inplace=True)
df = df['20220103':'20221230']
data_test = pd.DataFrame(df)
data_test = data_test[-get_data:]

close_sliced = data_test.columns.get_loc("close")
scaler=MinMaxScaler()
pred_scaler=MinMaxScaler()
data_scaled = scaler.fit_transform(data_test)
X_pred, y_pred = create_dataset(timestep, data_scaled)

prediction = loaded_model.predict(X_pred)
test_close = pd.DataFrame(data_test['close'])
np_test_close_scaled = pred_scaler.fit_transform(test_close)
prediction_scaled = pred_scaler.inverse_transform(prediction)
```

Listing 4.9 Code untuk prediksi harga close saham

```
In [14]: prediction_unscaled
```

```
Out[14]: array([[4443.4946],  
                [4441.8687],  
                [4443.7646],  
                [4430.1953],  
                [4441.578 ],  
                [4477.108 ],  
                [4469.4473],  
                [4485.279 ],  
                [4461.8037],  
                [4511.163 ],  
                [4609.567 ],  
                [4606.5405],  
                [4664.5283],  
                [4607.8135],  
                [4539.9565],  
                [4568.9116],
```

Gambar 4. 4 Contoh output prediksi

Hasil dari prediksi berbentuk numpy array seperti bisa dilihat pada Gambar 4.4. Untuk kebutuhan analisis hasil dari prediksi, dilakukan plotting dari data close asli yang diambil dari yahoo finance. Karena data terakhir yang diinputkan adalah 2022-12-30, berarti data yang akan diprediksi adalah dari tanggal bursa saham 2023 dibuka yaitu 2023-01-02. Pada Listing 4.10 terdapat code yang digunakan untuk mengimport yahoo finance dataset, dalam hal ini mengimport dataset BBKA dari 2023-01-02. Lalu dari prediction yang sudah diunscaled diambil sebanyak hari yang diprediksi untuk dijadikan format pandas dataframe. Baru selanjutnya data prediksi dan data real digabung menjadi satu tabel (lihat Gambar 4.5) untuk selanjutnya diplot ke dalam grafik. Listing 4.11 merupakan code untuk menampilkan hasil prediksi yang telah digabungkan dengan menggunakan dataframe yang diambil dari yahoo finance. Dapat dilihat hasil plotting pada Gambar 4.6 di mana hasil prediksi berwarna biru dan harga asli adalah garis berwarna hijau.

```
import yfinance as yf  
stock_data = yf.download(f"{stock_selected}.JK", start='2023-01-02')  
stock_data = stock_data[:pred_day]  
pd_result = pd.DataFrame(prediction_scaled, columns = ['Close'])  
real = stock_data["Close"]  
close_pred = pd_result["Close"]  
frames = [real, close_pred]
```

```

apred = pd.DataFrame(close_pred, columns = ["Close"])
real = real.reset_index()
result = pd.concat([real, apred], axis=1, ignore_index=True)
result.columns = ["date", "Real", "Prediction"]
result = result.set_index('date')

```

Listing 4.10 Code untuk mengambil data dari Yahoo Finance dan concat dengan prediction

date	Real	Prediction
2023-01-02	8550.0	7797.561035
2023-01-03	8550.0	7826.495605
2023-01-04	8350.0	7884.707031
2023-01-05	8250.0	7892.868652
2023-01-06	8300.0	7923.431641
2023-01-09	8450.0	7946.144531
2023-01-10	8175.0	7945.602051
2023-01-11	8125.0	7992.076660
2023-01-12	8175.0	8035.535645

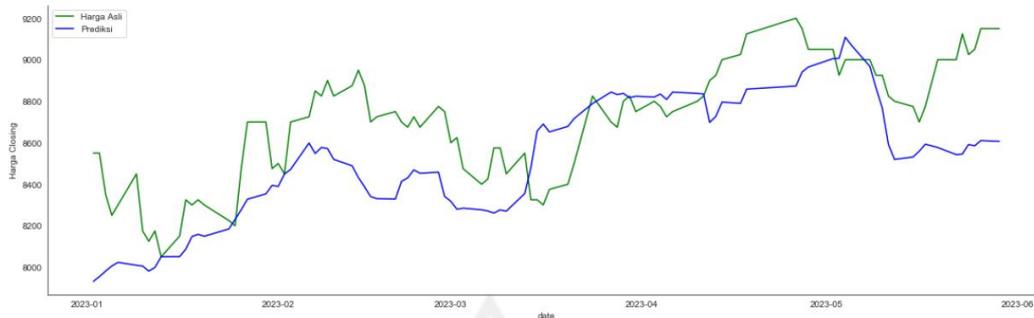
Gambar 4. 5 Dataframe gabungan Prediction dan Harga Real

```

fig = plt.figure(figsize = (20,6))
plt.plot(result['Real'], 'g', label = 'Harga Asli')
plt.plot(result['Prediction'], 'b', label = 'Prediksi')
plt.xlabel('date')
plt.ylabel('Harga Closing')
plt.legend()

```

Listing 4.11 Code untuk menampilkan grafik



Gambar 4. 6 Grafik prediksi BBCA

#### 4.2.4 Pembangunan Antarmuka

Pembangunan Antarmuka sederhana menggunakan Streamlit, di mana pengguna tinggal memilih emiten saham apa yang akan diprediksi lalu dengan menggunakan slider, pengguna memilih ingin memprediksi dalam berapa hari ke depan. Selanjutnya pengguna hanya tinggal klik tombol MULAI PREDIKSI. Sistem lalu memunculkan informasi yang akan ditampilkan dalam bentuk tabel yang menunjukkan hari ke-n dan prediksi harga close hari n tersebut serta grafik dari prediksi. Bisa dilihat untuk tampilan utama saat membuka file streamlit akan menampilkan Gambar 4.9. Lalu setelah pengguna memilih emiten saham, memilih jumlah hari yang akan diprediksi dan diakhiri dengan klik MULAI PREDIKSI, maka akan menampilkan Gambar 4.10

## Prediksi

oleh Nathanael Ivan Pratama

**⚠️ DISCLAIMER ON!** Setiap hasil dalam prediksi bersifat sebagai referensi / bahan pertimbangan. Setiap keuntungan dan kerugian menjadi tanggung jawab dari pelaku pasar.

Pilih emiten saham

BBCA

Pick a number

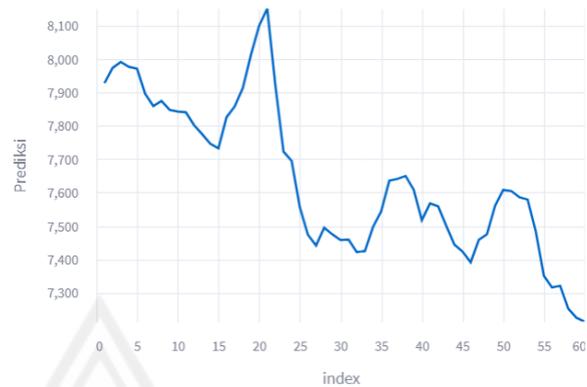


Anda akan memprediksi BBCA selama 60 hari ke depan

MULAI PREDIKSI

Gambar 4. 7 Tampilan antarmuka utama

	Prediksi
1	7,927.6172
2	7,973.1235
3	7,990.9829
4	7,976.2632
5	7,970.4736
6	7,895.415
7	7,858.9468
8	7,874.1108
9	7,847.5923
10	7,842.7139



Gambar 4. 8 Tampilan antarmuka hasil prediksi

Untuk dapat menghasilkan tampilan seperti pada Gambar 4.10 dan Gambar 4.11, diperlukan komponen Streamlit antara lain:

1. `st.write()` untuk menampilkan tulisan
2. `st.warning()` untuk menampilkan warning
3. `st.slider()` untuk menghasilkan slider guna memilih angka
4. `st.container()`, `st.columns()` untuk menghasilkan container yang menampung kolom untuk tabel di sebelah kiri dan grafik di sebelah kanan
5. `st.dataframe()` untuk menampilkan tabel prediksi
6. `st.altair_chart()` untuk menampilkan grafik

### 4.3 Analisis Hasil

Hasil eksperimen menunjukkan bahwa meskipun BiLSTM memiliki keunggulan dalam mengakses informasi dari kedua arah, perbandingan antara BiLSTM dan LSTM dalam prediksi saham menunjukkan bahwa ada saham yang lebih baik diprediksi menggunakan BiLSTM dan ada juga saham yang lebih baik diprediksi menggunakan LSTM. Dibuktikan dengan hanya **41,67%** emiten saham yang hasilnya lebih baik saat menggunakan BiLSTM. Hal ini mungkin terjadi jika ketergantungan dalam urutan data lebih bermanfaat ketika hanya mengacu pada satu konteks, yaitu masa lalu. Namun, penting untuk dicatat bahwa kinerja model dipengaruhi oleh faktor-faktor lain seperti volume data pelatihan dan pengaturan hiperparameter.

Saat membandingkan penggunaan Foreign Flow, ditemukan saham dengan RMSPE yang baik ketika fitur ini dimasukkan, tetapi ada juga yang sebaliknya. Di mana hanya sebesar **25%** atau sebanyak 21 emiten saham yang hasilnya lebih baik saat menggunakan Foreign Flow. Selain itu, bila dilihat pada tabel 4.6, saat dilihat nilai Pearson Correlation, tidak semua saham dengan hasil baik saat menggunakan Foreign Flow memiliki korelasi positif dalam Pearson Correlation.

Tabel 4. 6 Foreign flow terhadap Pearson Correlation

No	Kode Emiten	Nilai Pearson Correlation	Interpretasi
1	ACES	0.688670226	Positif Kuat
2	AGRO	0.597579819	Positif Sedang
3	AKRA	0.836612894	Positif Sangat Kuat
4	APIC	0.922090506	Positif Sangat Kuat
5	BBCA	0.680945744	Positif Kuat
6	BRPT	-0.40491176	Negatif Sedang
7	ENRG	-0.1442886	Negatif Sangat Rendah
8	ERAA	-0.73502178	Negatif Kuat
9	EXCL	0.392576948	Positif Rendah
10	HMSP	0.840755243	Positif Sangat Kuat
11	INCO	0.517116037	Positif Sedang
12	INDY	0.676039827	Positif Kuat
13	ISAT	-0.26183956	Negatif Rendah
14	LSIP	0.192551757	Positif Sangat Rendah
15	MIKA	-0.48554434	Negatif Sedang
16	PNBN	-0.34472988	Negatif Rendah
17	PTBA	0.675231133	Positif Kuat
18	SMGR	0.548759444	Positif Sedang
19	TINS	-0.3320152	Negatif Rendah
20	UNTR	0.237400229	Positif Rendah
21	UNVR	0.733458546	Positif Kuat

Korelasi negatif antara arus asing dan harga closing saham tidak berarti bahwa tidak ada korelasi antara dua variabel melainkan mengindikasikan hubungan terbalik antara kedua variabel tersebut. Artinya, ketika arus asing meningkat, harga closing saham cenderung menurun, dan sebaliknya. Hal ini menunjukkan adanya kecenderungan bahwa aktivitas arus asing dapat mempengaruhi harga saham.

Pergerakan arus asing dapat memberikan sinyal kepada investor dan pedagang tentang kekuatan pasar dan sentimen investor global yang mungkin mempengaruhi harga saham secara keseluruhan. Korelasi negatif ini dapat digunakan dalam analisis dan prediksi pergerakan harga saham untuk mengidentifikasi tren atau pola dalam pasar. Penting untuk diingat bahwa korelasi tidak menunjukkan hubungan sebab-akibat, melainkan hanya mengukur sejauh mana dua variabel berhubungan secara linier atau sejalan satu sama lain.

Pada bagian perbandingan stacking, mayoritas emiten saham atau sebanyak **70,23%** memiliki hasil yang lebih baik saat menggunakan single BiLSTM, ini mungkin disebabkan setiap saham memiliki karakteristik unik dan berbeda-beda. Beberapa saham mungkin lebih cocok diprediksi menggunakan model yang lebih sederhana seperti single BiLSTM, sementara yang lain mungkin memerlukan model yang lebih kompleks atau pendekatan yang berbeda. Ketepatan model dalam memprediksi tren saham sangat tergantung pada sifat dan dinamika masing-masing saham. Pada akhir dari proses pembuatan model hingga prediksi, ditemukan saham terbaik dalam penggunaan metode dari penelitian ini, antara lain adalah BBCA, INDF, ICBP, SIDO, dan MYOR. Dari saham terbaik tersebut dilakukan Prediksi Real, hasil dapat dilihat pada Tabel 4.8

Tabel 4. 7 Hasil prediksi data testing pada saham terbaik

Kode Emiten	Prediksi Validasi		
	MAE	RMSE	RMSPE
BBCA	78.04	105.63	1.62
INDF	86.17	118.27	1.74
ICBP	115.78	167.39	1.83
SIDO	9.87	14.22	1.83
MYOR	34.77	46.52	1.90

Tabel 4. 8 Hasil prediksi data real pada saham terbaik

Kode Emiten	Prediksi Real								
	30 Hari			60 Hari			90 Hari		
	MAE	RMSE	RMSPE	MAE	RMSE	RMSPE	MAE	RMSE	RMSPE
BBCA	400.34	489.81	5.48	202.20	241.63	2.76	284.65	325.90	3.83

Kode Emiten	Prediksi Real								
	30 Hari			60 Hari			90 Hari		
	MAE	RMSE	RMSPE	MAE	RMSE	RMSPE	MAE	RMSE	RMSPE
INDF	210.42	258.40	3.96	468.79	523.48	8.05	345.61	409.36	6.57
ICBP	312.22	343.50	3.48	675.58	794.97	8.65	935.24	1146.38	13.10
SIDO	33.26	38.97	5.10	73.84	83.69	11.19	73.26	96.99	13.79
MYOR	124.18	144.21	5.84	171.73	220.12	10.06	364.68	420.30	21.67

Dapat dilihat pada tabel 4.7 dan 4.8, bila dibandingkan hasilnya jika dirata-rata menunjukkan hasil RMSPE prediksi validasi yang baik sedangkan prediksi real yang buruk, ini bukan semata-mata karena overfitting melainkan hal-hal yang tidak bisa diduga oleh model terkait perilaku pasar. Jika dilihat pada metrik MAE emiten SIDO, pada saat prediksi menggunakan validasi atau data testing, hasilnya 9,87. Sedangkan pada ICBP, nilai MAE sangat besar yaitu 115,78 padahal kedua emiten tersebut memiliki nilai RMSPE yang tidak terpaut jauh. Ini adalah akibat dari Fraksi harga saham, di mana ini mengacu pada satuan pembagian harga saham. Dalam Bursa Efek Indonesia, harga saham biasanya dinyatakan dalam pecahan tertentu, yang merupakan unit pembagian terkecil dari harga saham yang dapat diperdagangkan. Dalam hal ini, SIDO memiliki fraksi harga 5 rupiah dan ICBP dengan fraksi harga 25 rupiah, oleh sebab itu meskipun memiliki nilai RMSPE yang tidak terlalu jauh, nilai MAE bisa saja jauh karena perbedaan **fraksi harga**, oleh sebab itu MAE tidak dapat digunakan untuk membandingkan antar emiten saham, melainkan melihat perbedaannya pada satu emiten saham tersebut jika menggunakan konfigurasi yang berbeda dalam prediksi untuk melihat berapa selisih angka prediksi dengan angka aktual.

Pada kelima saham terbaik tersebut dilihat pada prediksi real 4 dari 5 selalu lebih baik saat memprediksi jangka pendek atau dalam hal ini adalah 30 hari. Berbeda halnya dengan BBCA yang memiliki kecenderungan hasil prediksi yang baik pada jangka menengah yaitu 60 hari. Namun dari kelima emiten tersebut bisa dikatakan bahwa prediksi 90 hari adalah yang paling jelek karena tidak ada dari kelima saham tersebut yang memiliki hasil terbaik di 90 hari.

## **BAB V**

### **KESIMPULAN DAN SARAN**

#### **5.1 Kesimpulan**

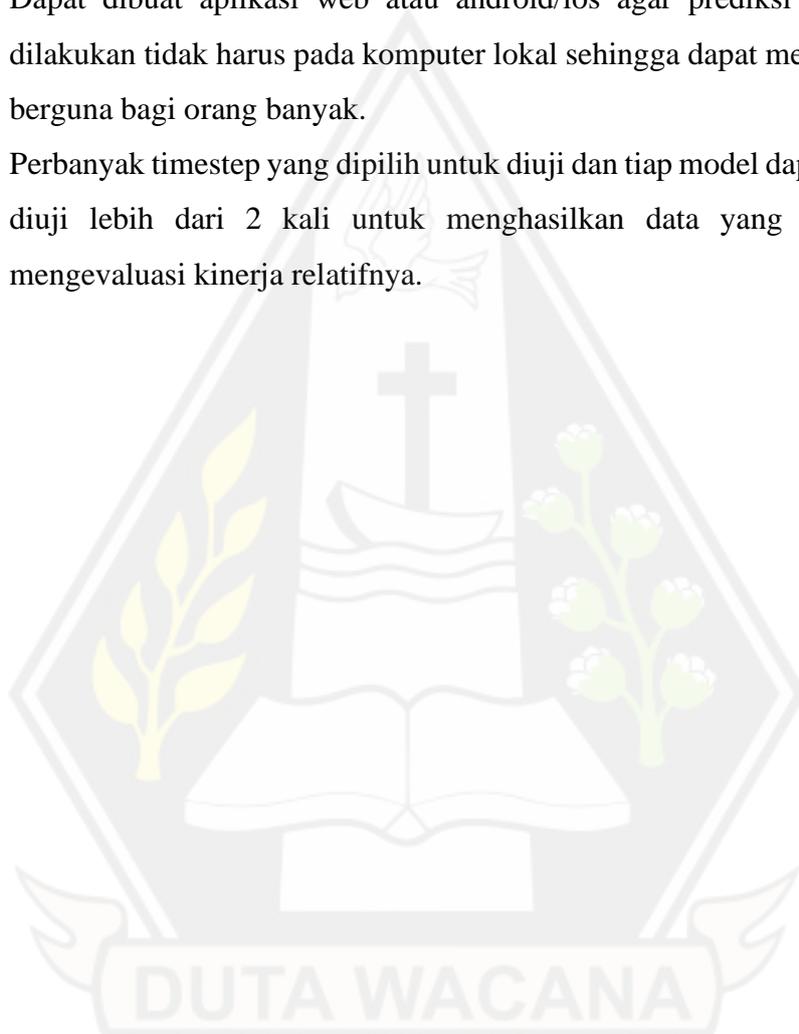
Berdasarkan tugas akhir yang telah dikerjakan pada penelitian Prediksi Harga Saham Pada Indeks KOMPAS100 Menggunakan Bidirectional Long Short Term Memory dengan Faktor Foreign Flow, kesimpulan yang didapat adalah sebagai berikut:

1. Mayoritas emiten saham masih lebih baik saat memakai LSTM, hanya 35 emiten saham yang lebih baik saat menggunakan BiLSTM. Dengan kata lain, 49 emiten saham memiliki kecenderungan kurang bisa dipelajari oleh model BiLSTM yang bekerja dua arah dari masa depan dan masa lalu.
2. Terdapat 21 emiten saham yang hasilnya lebih baik saat menggunakan Foreign Flow. 63 emiten saham yang lain ketika digunakan fitur tersebut hanya merusak akurasi dari model.
3. Berdasarkan metode yang dipakai pada penelitian ini ditemukan saham terbaik yaitu BBCA, INDF, ICBP, SIDO, dan MYOR dengan tolok ukur yang dipakai adalah dari nilai RMSPE.
4. Secara spesifik dilihat dari hasil evaluasi, saham terbaik adalah BBCA dengan nilai RMSPE 1.62% pada validasi dan 2.76% saat prediksi real. Perbedaan yang tidak jauh ini bisa diartikan sebagai saham yang memiliki korelasi kuat dengan fitur Foreign Flow baik itu dilihat dari permodelan maupun pearson correlation. Tentunya tren saham yang relatif stabil dan konsisten juga menjadi alasan dibalik hasil prediksi.
5. Model lebih mudah menghasilkan prediksi yang baik ketika grafik saham terus menerus naik atau terus menerus turun, seperti halnya BBCA yang dalam beberapa tahun terakhir selalu naik.

#### **5.2 Saran**

Adapun saran yang kiranya dapat dilakukan pada penelitian selanjutnya adalah sebagai berikut:

1. Karena emiten saham memiliki kecocokan tertentu untuk prediksi menggunakan variasi LSTM, penelitian selanjutnya bisa melakukan penelitian untuk variasi LSTM yang lain dengan harapan mendapatkan kecocokan dengan suatu emiten saham. Dan lebih baik lagi untuk meneliti salah satu emiten saja agar analisis lebih mendalam
2. Dapat dibuat aplikasi web atau android/ios agar prediksi saham dapat dilakukan tidak harus pada komputer lokal sehingga dapat menjangkau dan berguna bagi orang banyak.
3. Perbanyak timestep yang dipilih untuk diuji dan tiap model dapat dilatih dan diuji lebih dari 2 kali untuk menghasilkan data yang cukup untuk mengevaluasi kinerja relatifnya.



## DAFTAR PUSTAKA

- Alfath, M. R., & Dewi, F. R. (2022). Periodic Rebalancing Portofolio Saham Indeks Kompas100 Di Masa Pandemi Covid-19. *Jurnal Ilmiah Manajemen Fakultas Ekonomi*, 109-120.
- Aprilia, A. (2023). *Analisis Sebaran Emisi So2 Dan No2 Dari Cerobong Boiler Industri Pengolahan Kelapa Sawit Dengan Model Aermod Di Pt Perkebunan Nusantara Vi Unit Usaha Pinang Tinggi Kabupaten Muaro Jambi*. Jambi: Fakultas Sains Dan Teknologi Universitas Jambi.
- Asia, N. (2020). Faktor – Faktor Yang Mempengaruhi Return Saham Pada Perusahaan Property Tahun 2014 – 2016 Di Bursa Efek Indonesia. *FORECASTING: Jurnal Ilmiah Ilmu Manajemen*.
- BERGSTRÖM, C., & HJELM, O. (2019). *Impact of Time Steps on Stock Market Prediction with LSTM*. Stockholm: KTH Royal Institute of Technology.
- Bronwlee, J. (2016, Agustus 2). *Data Leakage in Machine Learning*. Diambil kembali dari Machine Learning Mastery: <https://machinelearningmastery.com/data-leakage-machine-learning/>
- Chandahas Mishra, D. L. (2017). Deep Machine Learning and Neural Networks: An Overview. *IAES International Journal of Artificial Intelligence (IJ-AI)*, 66-73.
- Chenyu Han, X. F. (2023). Challenge and Opportunity: Deep Learning-Based Stock Price Prediction by Using Bi-Directional LSTM Model. Dalam *Frontiers in Business, Economics and Management*. New Brunswick: Darcy & Roy Press.
- Hastomo, W., Aini, N., Karno, A. S., & Rere, L. R. (2022). Metode Pembelajaran Mesin untuk Memprediksi Emisi Manure Management. *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*.
- Linanda, R., & Arfiyenis, W. (2018). Pengaruh Struktur Modal Dan Profitabilitas. *JEBI (Jurnal Ekonomi dan Bisnis Islam)*, 135-144.
- Miftahuddin, Pratama, A., & Setiawan, I. (2021). Analisis Hubungan Antara Kelembaban Relatif Dengan Beberapa Variabel Iklim Dengan Pendekatan Korelasi Pearson Di Samudera Hindia. *Siger Matematika*, 25-33.
- Mingzhu Jia, J. H. (2019). Analysis and Research on Stock Price of LSTM and Bidirectional LSTM Neural Network. *3rd International Conference on Computer Engineering, Information Science & Application Technology (ICCIA 2019)*.
- Nelly Marlina, K. d. (2015). Evaluasi Daya Tampung Terhadap Beban Pencemar Menggunakan Model Kualitas Air (Studi Kasus: Sungai Winongo). *Jurnal Inovasi dan Kewirausahaan*, 78-86.
- Nurdi Afrianto, D. H. (2021). Prediksi Harga Saham Menggunakan BiLSTM dengan Faktor Sentimen Publik. *Rekayasa Sistem dan Teknologi Informasi (RESTI)*, 41-46.

- Patel, R., Choudhary, V., Saxena, D., & Singh, A. K. (2021). Review Of Stock Prediction Using Machine Learning Techniques. *2021 5th International Conference on Trends in Electronics and Informatics (ICOEI)*, 840-846.
- Santoso, V. A. (2022). *Prediksi Pergerakan Harga Saham Dengan Metode Long Short-term Memory (Lstm)*. Yogyakarta: Universitas Kristen Duta Wacana.
- Saxena, A., Dhadwal, M., & Kowsigan, M. (2021). Indian Crop Production: Prediction And Model Deployment Using MI And. *Turkish Journal of Physiotherapy and Rehabilitation*, 1874-1886.
- Septiawan, R., & Astuti, E. Z. (2016). Perbandingan Metode Setengah Rata-rata Dan Metode Kuadrat Terkecil Untuk Peramalan Pendapatan Perusahaan Di Blu Uptd Terminal Mangkang Semarang. *Techno.COM*.
- Sreelakshmi Ka, R. P. (2018). Deep Bi-Directional LSTM Network for Query Intent Detection. *8th International Conference on Advances in Computing and Communication (ICACC-2018)*.
- Suryanto, A. A., & Muqtadir, A. (2019). Penerapan Metode Mean Absolute Error (MEA) Dalam Algoritma Regresi Linear Untuk Prediksi Produksi Padi. *SAINTEKBU: Jurnal Sains dan Teknologi*, 78-83.
- Wati, N. P., & Pramatha, C. (2022). Penerapan Long Short Term Memory dalam Mengklasifikasi Jenis Ujaran Kebencian pada Tweet Bahasa Indonesia. *Jurnal Nasional Teknologi Informasi dan Aplikasinya (JNATIA)*, 755-762.
- Zulfan, Hashfi, R. U., & Ary, W. W. (2021). Interaksi dan Hubungan Kausalitas antara Net Foreign Flow dan Return Saham Syariah (Studi Empiris Saham Syariah Jakarta Islamic Index Periode 2012-2018). *JURNAL MANAJEMEN DAN KEUANGAN*, 253-279.