****

# UJI PERFORMA PREDIKSI METODE *AUTO REGRESSIVE FRACTIONALLY INTEGRATED MOVING AVERAGES* DAN *LONG SHORT-TERM MEMORY* DENGAN DATA SAHAM DUA PERUSAHAAN BANK

**PROPOSAL TUGAS AKHIR**

**Berli Suharmanto**

**1910511033**

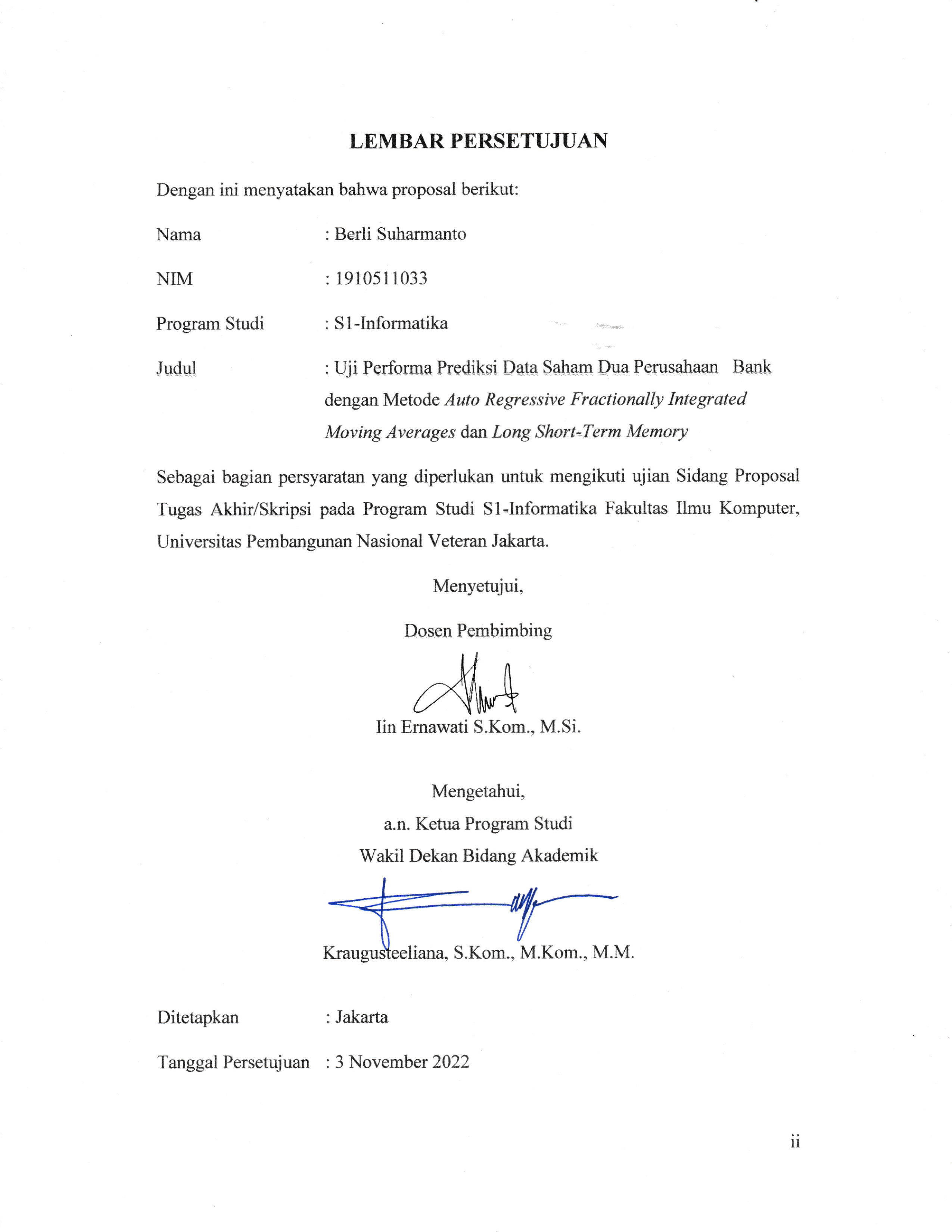
**UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAKARTA**

**FAKULTAS ILMU KOMPUTER**

**PROGRAM STUDI INFORMATIKA**

**2022**

# LEMBAR PERSETUJUAN



UJI PERFORMA PREDIKSI METODE *AUTO REGRESSIVE FRACTIONALLY INTEGRATED MOVING AVERAGES* DAN *LONG SHORT-TERM MEMORY* DENGAN DATA SAHAM DUA PERUSAHAAN BANK

Berli Suharmanto

# ABSTRAK

Investasi saham adalah salah satu cara untuk menghasilkan keuntungan bagi investor pada zaman sekarang. Namun dibalik keuntungan tersebut terdapat juga resiko mengalami kerugian karena tidak dapat diprediksi dengan mudah. Dalam analisis data *time series* menggunakan data saham adalah salah satu bentuk analisis yang sulit. Terdapat beberapa metode *deep learning* yang cocok untuk melakukan analisis prediksi data *time series*, dua di antaranya adalah metode *Auto Regressive Fractionally Integrated Moving Averages* (ARFIMA) dan *Long Short-Term Memory* (LSTM). Data saham yang digunakan adalah data saham dari IDX dengan data saham Bank Raya Indonesia Tbk. dan Bank IBK Indonesia Tbk. dalam periode September 2019 sampai Desember 2022. Hasil pengolahan dari model ARFIMA terbaik adalah model ARFIMA(8,0.5,0) dari data BRI dengan SMAPE sebesar 5.57% dan model ARFIMA(4,0.5,0) dari data Bank IBK Indonesia dengan SMAPE sebesar 23.31%. Untuk hasil terbaik LSTM BRI adalah 1.61% dengan 150 *epoch* dan LSTM Bank IBK Indonesia adalah 2.22% dengan 300 *epoch*. Model terbaik yang dihasilkan adalah model LSTM dengan 150 *epoch* untuk BRI dan model LSTM dengan 300 *epoch*.

**Kata Kunci:** Saham, *Deep Learning*, ARFIMA, LSTM, Prediksi

# *ABSTRACT*

*Stock investment is one way to generate profits for investors today. But behind these advantages there is also a risk of experiencing losses because they cannot be predicted easily. In the analysis of time series data using stock data is a form of analysis that is difficult. There are several deep learning methods that are suitable for predicting time series data, two of which are the Auto Regressive Fractionally Integrated Moving Averages (ARFIMA) and Long Short-Term Memory (LSTM) methods. The stock data used is stock data from IDX with Bank Raya Indonesia Tbk stock data. and Bank IBK Indonesia Tbk. in the period September 2019 to December 2022. The processing results of the best ARFIMA model are the ARFIMA(8,0.5,0) from BRI data with SMAPE of 5.57% and the ARFIMA(4,0.5,0) from Bank IBK Indonesia data with SMAPE of 23.31%. For the best results, BRI's LSTM is 1.61% with 150 epochs and Bank IBK Indonesia's LSTM is 2.22% with 300 epochs. The best model produced is the LSTM model with 150 epochs for BRI and the LSTM model with 300 epochs.*

**Kata Kunci:** *Stock*, *Deep Learning*, ARFIMA, LSTM, *Prediction*

# DAFTAR ISI

Halaman

[HALAMAN JUDUL i](#_Toc118289135)

LEMBAR PERSETUJUAN [**Error! Bookmark not defined.**](#_Toc118289136)

[ABSTRAK ii](#_Toc118289137)

[*ABSTRACT* ii](#_Toc118289138)

[DAFTAR ISI ii](#_Toc118289139)

[DAFTAR TABEL ii](#_Toc118289140)

[DAFTAR GAMBAR ii](#_Toc118289141)

[BAB 1 PENDAHULUAN 2](#_Toc118289142)

[1.1. Latar Belakang 2](#_Toc118289143)

[1.2. Rumusan Masalah 2](#_Toc118289144)

[1.3. Tujuan Penelitian 2](#_Toc118289145)

[1.4. Batasan Masalah 2](#_Toc118289146)

[1.5. Luaran yang Diharapkan 2](#_Toc118289147)

[1.6. Manfaat Penelitian 2](#_Toc118289148)

[1.7. Sistematika Penulisan 2](#_Toc118289149)

[BAB 2 TINJAUAN PUSTAKA 2](#_Toc118289150)

[2.1. Saham 2](#_Toc118289151)

[2.2. Bank 2](#_Toc118289152)

[2.3. *Python* 2](#_Toc118289153)

[2.4. *Data Mining* 2](#_Toc118289154)

[2.5. Jaringan Saraf Tiruan 2](#_Toc118289155)

[2.6. *Deep Learning* 2](#_Toc118289156)

[2.7. *Recurrent Neural Network* 2](#_Toc118289157)

[2.8. *Long Short-Term Memory* 2](#_Toc118289158)

[2.8.1. Tahap *Long Short-Term Memory* 2](#_Toc118289159)

[2.9. *Auto Regressive Fractionally Integrated Moving Averages* (ARFIMA) 2](#_Toc118289160)

[2.10. Normalisasi 2](#_Toc118289161)

[2.11. *Root Mean Square Error* **Error! Bookmark not defined.**](#_Toc118289162)

[2.12. *Mean Absolute Percentage Error* 2](#_Toc118289163)

[2.13. *Symmetric Mean Absolute Percentage Error* 2](#_Toc118289164)

[2.14. *R-Squared* 2](#_Toc118289165)

[2.15. Penelitian Relevan 2](#_Toc118289166)

[BAB 3 METODOLOGI PENELITIAN 2](#_Toc118289167)

[3.1. Tahapan Penelitian 2](#_Toc118289168)

[3.2. Identifikasi Masalah 2](#_Toc118289169)

[3.3. Studi Literatur 2](#_Toc118289170)

[3.4. Pengumpulan Data 2](#_Toc118289171)

[3.5. Praproses Data 2](#_Toc118289172)

[3.5.1. *Attribute Removal* 2](#_Toc118289173)

[3.5.2. Cek *Missing Value* **Error! Bookmark not defined.**](#_Toc118289174)

[3.5.3. Uji stasioner 2](#_Toc118289175)

[3.5.4. Stasionerisasi 2](#_Toc118289176)

[3.6. Pembagian Data 2](#_Toc118289177)

[3.7. Penentuan Parameter 2](#_Toc118289178)

[3.8. *Training* ARFIMA **Error! Bookmark not defined.**](#_Toc118289179)

[3.9. *Training* LSTM 2](#_Toc118289180)

[3.10. Prediksi Kedua Model 2](#_Toc118289181)

[3.11. Evaluasi 2](#_Toc118289182)

[3.12. Kesimpulan Uji Performa Algoritma 2](#_Toc118289183)

[3.13. Waktu dan Tempat Penelitian 2](#_Toc118289184)

[3.14. Alat Bantu Penelitian 2](#_Toc118289185)

[3.14.1. Perangkat keras 2](#_Toc118289186)

[3.14.2. Perangkat Lunak 2](#_Toc118289187)

[3.15. Jadwal Penelitian 2](#_Toc118289188)

[DAFTAR PUSTAKA 2](#_Toc118289189)

# DAFTAR TABEL

[Tabel 3.1. Jadwal Penelitian 2](#_Toc117951913)

# DAFTAR GAMBAR

[Gambar 2.1. Proses Knowledge Discovery in Database (Chiu and Tavella, 2020) 2](#_Toc118285550)

[Gambar 2.2. Recurrent Neural Network (Manaswi, 2018) 2](#_Toc118285551)

[Gambar 2.3. Struktur model RNN memiliki satu layer (Agusta, 2021) 2](#_Toc118285552)

[Gambar 2.4. Struktur LSTM dengan empat layer (Agusta, 2021) 2](#_Toc118285553)

[Gambar 2.5. Cell State pada LSTM (Agusta, 2021) 2](#_Toc118285554)

[Gambar 2.6. Sigmoid layer pada LSTM (Agusta, 2021) 2](#_Toc118285555)

[Gambar 2. 7. Forget gate layer pada LSTM (Agusta, 2021) 2](#_Toc118285556)

[Gambar 2.8. Input gate layer dan tanh layer pada LSTM (Agusta, 2021) 2](#_Toc118285557)

[Gambar 2.9. Cell state terbaru pada LSTM (Agusta, 2021) 2](#_Toc118285558)

[Gambar 2.10. Output layer pada LSTM (Agusta, 2021) 2](#_Toc118285559)

[Gambar 3.1. Tahapan Alur Metodologi Penelitian 2](#_Toc117951668)

# BAB I

**PENDAHULUAN**

## 1.1. Latar Belakang

Investasi saham adalah cara untuk memasukkan uang ke dalam sesuatu dengan harapan menghasilkan keuntungan di kemudian hari. Ada kemungkinan bagi investor untuk mengalami kerugian atau kehilangan uangnya, karena imbalan yang diprediksi bukan tanpa risiko dan ukurannya tidak dapat ditentukan (Mustafa, 2020).

Tindakan berinvestasi dalam instrumen keuangan (surat berharga), seperti saham, obligasi, reksa dana, derivatif, berjangka, dan opsi, dikenal sebagai pasar modal. Mengingat beragamnya investasi yang tersedia, sangat penting bagi setiap investor untuk menyadari bahaya saat ini dan potensi masa depan. Tabungan harus diganti dengan investasi agar setiap investor pasar saham dapat menggunakan kekayaannya sebagai jaminan untuk masa depan (Nuzula Agustin dan Lysion, 2021).

Investor saham harus memiliki pertimbangan yang matang dalam bertransaksi saham di pasar modal. Membuat keputusan investasi adalah tugas yang membutuhkan dua pertimbangan signifikan bagi setiap investor, serta eksekusi yang cermat. Untuk membuat pilihan terbaik dalam investasi saham, perlu untuk sepenuhnya menganalisis semua elemen yang relevan. Investor harus selalu mewaspadai elemen kunci yang mempengaruhi investasi pasar modal. Behavioral finance adalah salah satu ilmu yang dapat digunakan investor untuk membantu mereka memilih pasar saham yang baik dengan bijak (Fauzi, 2020).

Memprediksi saham adalah salah satu masalah yang paling sulit dalam analisis data *time series*. Dari sektor keuangan suatu perusahaan bagaimana cara memperkirakan perubahan kinerja saham secara akura. Karena kompleksitas pasar dan sifatnya yang dinamis, serta banyak elemen implisit yang saling berhubungan yang ikut berperan, memprediksi kinerja saham adalah tugas yang menantang (Gu *et al.*, 2020).

Deep learning banyak digunakan di berbagai domain akademik dan situasi dunia nyata seperti prediksi saham sebagai hasil dari kemajuan berkelanjutan di bidang kecerdasan buatan. Karena popularitas *neural network* seperti *deep learning* yang digunakan dalam pembuatan model predikis saham juga telah berkembang dan ditingkatkan (Hu, Zhao dan Khushi, 2021).

Dalam penelitian terdahulu yang dilakukan oleh Fajar Akbardipura yang menghasilkan akurasi prediksi data saham ICBP dengan menggunakan ARIMA(0,1,0) dan LSTM sebesar 0,8% (Akbardipura, 2021). Penelitian lainnya adalah prediksi saham farmasi dengan LSTM oleh Ardiyan Agusta dengan RMSE sebesar 1,9879 (Agusta, 2021). Penelitian yang dilakukan oleh Puspita Kartikasari dengan data Bank Indonesia algoritma ARFIMA (0,0.499,5) memiliki nilai SMAPE 1,23% (Kartikasari, 2020).

## 1.2. Rumusan Masalah

Berdasarkan permasalahan yang terdapat pada latar belakang, maka rumusan masalah yang akan dibahas dalam penelitian ini sebagai berikut:

1. Bagaimana penerapan metode ARFIMA dalam prediksi nilai saham?
2. Bagaiman penerapan metode LSTM dalam prediksi nilai saham?
3. Bagaimana hasil perbandingan prediksi harga saham dari model yang dibuat?

## 1.3. Tujuan Penelitian

Berdasarkan rumusan masalah yang telah dijelaskan, maka tujuan penelitian adalah:

1. Mengetahui penggunaan ARFIMA dalam memprediksi nilai saham dengan memanfaatkan data-data saham tahun ini.
2. Mengetahui penggunaan LSTM dalam memprediksi nilai saham dengan memanfaatkan data-data saham tahun ini.
3. Mengetahui model terbaik dari hasil prediksi yang telah dilakukan.

## 1.4. Batasan Masalah

Batasan masalah dalam penelitian ini adalah sebagai berikut:

1. Dataset yang digunakan merupakan data dari saham Bank Rakyat Indonesia (Persero) Tbk. dan Bank IBK Indonesia Tbk. periode September 2019 hingga Desember 2022.
2. Prediksi nilai saham menggunakan metode ARFIMA dan LSTM.
3. Menggunakan bahasa pemrograman *python*.
4. Atribut prediksi yang digunakan adalah Penutupan.

## 1.5. Luaran yang Diharapkan

Luaran yang diharapkan berupa model prediksi saham terbaik menggunakan antara metode ARFIMA dan LSTM periode data saham September 2019 hingga Desember 2022 yang dapat membantu keputusan pembelian saham.

## 1.6. Manfaat Penelitian

Dari penelitian ini diharapkan dapat memberikan manfaat antara lain:

1. Mengetahui metode terbaik untuk tiap data Bank untuk digunakan investor.
2. Mengetahui hasil dari penggunaan ARFIMA dan LSTM dalam prediksi saham.

## 1.7. Sistematika Penulisan

BAB 1 PENDAHULUAN

Pada bab ini dijelaskan latar belakang dilakukannya penelitian, rumusan permasalahan, tujuan penelitian, manfaat penelitian, ruang lingkup yang digunakan dalam penelitian, luaran yang diharapkan dari penelitian, dan sistematika penulisan proposal penelitian.

BAB 2 TINJAUAN PUSTAKA

Pada bab ini dijelaskan landasan teori yang digunakan pada penelitian dan membahas penelitian-penelitian yang sudah ada kemudian dibandingkan satu antar yang lainnya.

BAB 3 METODOLOGI PENELITIAN

Pada bab ini dijelaskan proses penelitian yang dilakukan mulai dari alur penelitian, tahapan penelitian, waktu dan tempat penelitian, dan jadwal penelitian untuk dilaksanakan.

DAFTAR PUSTAKA

Daftar pustaka berisi semua referensi yang digunakan pada penelitian yang dilakukan.

LAMPIRAN

Lampiran berisi mengenai informasi atau tautan tambahan yang dilakukan dalam penelitian ini.

# BAB II

**TINJAUAN PUSTAKA**

## 2.1. Saham

Saham adalah suatu yang membuktikan kepemilikan daripada nilai suatu perusahaan. Hal ini membuat perusahaan yang membutuhkan dana jangka panjang untuk menjual kepentingan dalam bisnis-saham atau efek ekuitas dengan uang investasi tunai. Metode ini menjadi metode yang utama dalam meningkatkan modal bisnis selain diperoleh dari obligasi. Secara tidak langsung, seorang investor menjadi pemilik perusahaan dan memiliki haknya dalam perusahaan. Akan tetapi, porsi yang dimilikinya hanya tergantung dengan jumlah presentase kepemilikan modal yang dibeli.

Saham sendiri memiliki perhitungan-perhitungan tersendiri untuk setiap unsurnya. Unsur tersebut adalah empat rasio saham dan dua unsur laba, yaitu:

### 2.1.1. *Price to Book Value Ratio*

` *Price to book value ratio* merupakan rasio penentuan harga suatu saham yang sedang beredar. Rasio ini membandingkan antara harga saham perusahaan di pasar dan harga yang sesuai buku perusahaan itu sendiri yang tercantum dalam laporan keuangan dengan mengurangkan nilai aktiva (kekayaan perusahaan) dengan kewajiban perusahaan sehingga menampakan perkiraan banyaknya pemegang saham yang memegang tanggung jawab pembiayaan aset bersih perusahaan. Berikut merupakan perumusan *price to book value ratio*:

(1)

Keterangan:

Harga saham dipasar

Harga pada buku

### 2.1.2. *Price to Earnings Ratio*

*Price per earnings ratio* merupakan rasio yang berfungsi dalam penilaian perusahaan berdasarkan harga saham saat ini dengan pendapatan per-sahamnya. Rasio ini biasa digunakan para investor dalam menganalisis nilai relatif saham perusahaan di dalam perbandingan *apple to apple*. Rasio ini juga memiliki manfaat lain yaitu untuk membandingkan tiap perusahaan berdasarkan sejarah maupun membandingkan pasar agregat dalam rentetan waktu tertentu. Berikut merupakan perumusan *price to earnings ratio*:

(2)

### 2.1.3. *Price Earning Growth Ratio*

*Price earning growth ratio* merupakan rasio yang berfungsi dalam perhitungan nilai saham berdasarkan dari pendapatan dan potensi kedepannya suatu perusahaan. Rasio ini biasa digunakan para investor dalam menganalisis tinggi rendahnya suatu saham pada waktu tertentu dengan mempertimbangkan pertumbuhannya di masa yang akan datang. Berikut merupakan perumusan *price earning growth ratio*:

(3)

Keterangan:

rasio harga ke pendapatanlaba per lembar saham

### 2.1.4. *Dividen Yield*

*Dividen yield* merupakan suatu metode yang dapat digunakan untuk mengukur jumlah arus kas yang didapat dari seluruh modal yang diinvestasikan oleh investor dalam pasar saham. Artinya *dividen yield* berfungsi untuk menghitung presentase laba yang didapatkan oleh investor dari *dividen* atau pembagian keuntungan dari perusahaan kepada investornya. Untuk menghitung *dividen yield* dapat dirumuskan:

(4)

### 2.1.5. *Return of Equity*

*Return of equity* merupakan suatu metode untuk mengukur penghasilan laba dari suatu perusahaan dari nilai investasi saham. *Return of equity* dapat dirumuskan:

(5)

### 2.1.6. *Earnings per Share*

*Earnings per share* merupakan pendapatan bersih perusahaan dalam satu tahun dikurangi *dividen preferen* atau pembagian hasil kepada pemegang saham yang lebih prioritas. Hasil dari pengurangan tersebut kemudian dibagi dengan jumlah saham yang berada pada peredaran pasar. Berikut merupakan rumusan *earnings per share*:

(6)

Unsur-unsur saham di atas dapat digunakan untuk menghitung pendapatan maksimal dari investasi. Hal ini dapat memudahkan investor dalam memperhitungkan investasi yang dikeluarkan dan *dividen* yang dihasilkan (K, 2020).

## 2.2. Bank

Bank adalah perusahaan yang bertindak sebagai perantara untuk transaksi keuangan dan biasanya dimotivasi oleh keinginan untuk mendapatkan wesel, mencetak uang, dan menerbitkan uang kertas. Frase "tempat penukaran uang" merupakan istilah bank yang berasal dari bahasa Italia “*banca*”. Selain itu, bank dapat disebut sebagai usaha yang mengambil uang dari masyarakat dalam bentuk simpanan dan memberikannya dalam bentuk kredit atau bentuk pembayaran lainnya dengan tujuan untuk meningkatkan taraf hidup mereka (Gortsos, 2020).

Penelitian ini menggunakan dua bank sebagai objek penelitiannya. Bank-bank tersebut memiliki sejarah yang panjang dalam perjalanan bisnis yang dilakukan. Sejarah bank tersebut antara lain:

### 2.2.1. Sejarah Bank Rakyat Indonesia

Bank Rakyat Indonesia merupakan bank yang dimiliki oleh pemerintah yang terbesar di Indonesia. BRI didirikan pada tanggal 16 Desember 1895 oleh Raden Bei Aria Wirjaatmadja di Purwokerto, Jawa Tengah.

Setelah kemerdekaan Republik Indonesia, pemerintah mengeluarkan Peraturan Pemerintah No. 1 tahun 1946 Pasal 1 yang berisi bahwa BRI adalah sebagai Bank Pemerintah pertama di Republik Indonesia. Pada saat perang mempertahankan kemerdekaan pada tahun 1948, seluruh aktifitas BRI terhenti sementara dan kembali aktif ketika perjanjian Renville ditanda tangani pada tahun 1949 dan berubah nama menjadi Bank Rakyat Indonesia Serikat.

Kemudian melalui Perpu No. 41 tahun 1960, Bank Koperasi Tani dan Nelayan (BKTN) terbentuk dari peleburan BRI, Bank Tani Nelayan, dan Nederlandsche Maatschappij (NHM). Setelah itu melalui Penetapan Presiden No. 9 tahun 1965, BKTN diintegrasikan ke dalam Bank Indonesia dengan nama Bank Indonesia Urusan Koperasi Tani dan Nelayan. Satu bulan kemudian dikeluarkan Penpres No. 17 tahun 1965 tentang pembentukan bank tunggal dengan nama Bank Negara Indonesia. Pada ketentuan itulah diintegrasikannya Bank Indonesia Urusan Koperasi, Tani, dan Nelayan dengan nama Bank Negara Indonesia unit II bidang Rural dan NHM menjadi Bank Negara Indonesia unit II bidang Ekspor Impor.

Pada 1 Agustus 1992 dalam Undang-Undang Perbankan No. 7 tahun 1992 dan Peraturan Pemerintah RI No. 21 tahun 1992 status BRI mendapatkan perubahan status menjari perseroan terbatas. Untuk kepemilikannya sendiri, BRI masih 100% dimiliki oleh Pemerintah Republik Indonesia. Di tahun 2003, Pemerintah Indonesia menjual 30% saham BRI yang membuat BRI menjadi perusahaan publik dengan mengganti nama resminya menjadi PT. Bank Rakyat Indonesia (Persero) Tbk., sampai sekarang (*Info Perusahaan - Bank BRI | Melayani Dengan Setulus Hati*, 2022).

### 2.2.2. Sejarah Bank IBK (*Industrial Bank of Korea*) Indonesia

Berdasarkan Keputusan Menteri Keuangan Republik Indonesia No. Kep. 792/MK/IV/12/1970 tanggal 7 Desember 1970 dan berdasarkan Akta Perseroan Terbatas No. 85 tanggal 13 November 1973 juncto Akta Perubahan No. 315 tanggal 29 Maret 1974, Bank IBK Indonesia didirikan pada Jakarta dengan nama PT Finconesia. Bank Negara Indonesia, Nomura Securities, Barclays Bank International, Manufactures Hanover International Finance Corporation, Mitsui Bank, Banque Francaise Du Commerce Exterieur, dan Commerzbank semuanya memiliki saham di organisasi keuangan yang dikenal sebagai Finconesia.

Berdasarkan Surat Keputusan Menteri Keuangan Republik Indonesia No. 442/KMK.017/1993 tanggal 9 Maret 1993, Finconesia berubah dari lembaga keuangan menjadi Bank Umum pada tahun 1993, mengubah namanya menjadi PT Bank Finconesia (Bank Finconesia).

Berdasarkan Akta Perseroan Terbatas No. 146 tanggal 18 Juli 2008 yang dibuat dihadapan Sutjipto, SH, M.Kn, Notaris di Jakarta dan diumumkan dalam Surat Keputusan Menkumham No. AHU-45703. AH.01.02 Tahun 2008, tanggal 29 Juli 2008 (Akta No. 146/2008), yang juga merupakan perubahan atas Undang-Undang No. 40 Tahun 2007 (UUPT), nama Bank Finconesia resmi berubah menjadi PT Bank Agris Tbk.

Penawaran Umum Perdana (Initial Public Offering/IPO) diselenggarakan oleh Bank Agris pada tahun 2014, dan selanjutnya perusahaan mencatatkan sahamnya di Bursa Efek Indonesia. Alhasil, Bank Agris kini dikenal sebagai PT Bank Agris Tbk dan merupakan perusahaan publik. Kemudian perjanjian Industrial Bank of Korea (IBK) untuk membeli saham PT Bank Agris Tbk (AGRS) dari PT Dian Intan Perkasa ditandatangani pada 15 Januari Saat Transaksi pada tahun 2019.

Berdasarkan Pernyataan Keputusan Rapat nomor 250 tanggal 22 Agustus 2019, ditandatangani di hadapan Notaris Christina Dwi Utami, SH, M.Hum, MK.N, dan telah mendapat pengesahan dari Kementerian Hukum dan Hak Asasi Manusia Republik Indonesia dengan Surat Keputusan No. AHU-00560, Perseroan menerima penggabungan usaha dengan PT Bank Mitraniaga Tbk dan mengubah nama Perseroan dari PT Bank Agris Tbk menjadi PT Bank IBK Indonesia Tbk pada 4 September 2019 (*Bank IBK Indonesia - Wikipedia bahasa Indonesia, ensiklopedia bebas*, 2022).

## 2.3. *Python*

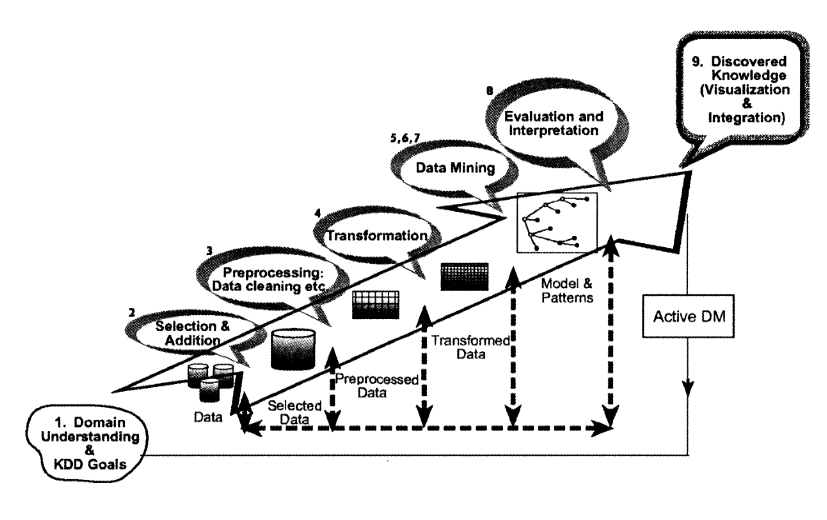
Berbeda dengan bahasa yang dikompilasi seperti C atau FORTRAN, *Python* adalah bahasa yang ditafsirkan. Terlepas dari kenyataan bahwa kedua situasi dimulai dengan file kode sumber, kompiler memeriksa kode secara menyeluruh sebelum membuat *executable* yang ditautkan ke file perpustakaan khusus sistem. Kompiler tidak diperlukan setelah *executable* diproduksi. Di sistem, Anda cukup menjalankan yang dapat dieksekusi. Di sisi lain, untuk menjalankan kode dalam bahasa yang ditafsirkan seperti *Python*, Anda harus selalu menjalankan proses *Python*. Ini karena proses *Python*, yang beroperasi sebagai abstraksi pada *platform*, harus menerjemahkan perintah dalam kode sumber untuk menjalankannya di sistem. *Interpreter* *Python* bertanggung jawab untuk masalah khusus platform karena bertindak sebagai jembatan antara kode sumber *platform* dan pengguna. Ini memiliki manfaat bahwa kode sumber dapat berjalan di berbagai *platform* selama juru bahasa *Python* fungsional tersedia di setiap *platform*.

Karena *interpreter* *Python* yang sesuai menangani detail spesifik *platform*, ini membuat kode sumber *Python* lintas *platform*. Kemampuan *Python* untuk berjalan pada *platform* yang berbeda merupakan nilai tambah yang besar, terutama di tahun-tahun awalnya. Kembali ke bahasa yang dikompilasi, *executable* terikat ke *platform* tertentu dan ke *library* tertentu yang telah ditautkan ke *executable* karena informasi spesifik *platform* disematkan di *executable*. Karena kompiler dapat menautkan ke *platform* tertentu, ia memiliki opsi untuk memanfaatkan optimasi dan pustaka level khusus *platform*, yang membuat kode ini kurang portabel daripada *Python*. Selain itu, kompiler dapat memeriksa file kode sumber dan menerapkan pengoptimalan tingkat kompiler yang mempercepat hasil yang dapat dieksekusi. Ini adalah perbedaan utama antara bahasa yang ditafsirkan dan dikompilasi secara luas. Nanti, kita akan melihat bahwa ada banyak kompromi antara dua ekstrem ini yang membantu kode *Python* berjalan lebih cepat (Unpingco, 2021).

## 2.4. *Data Mining*

Dalam proses data mining, kumpulan data yang besar dikumpulkan, data historis digunakan untuk menghasilkan pola yang konsisten atau teratur, dan hubungan terbentuk. Data mining adalah proses yang digunakan untuk mengekstrak nilai lebih dari kumpulan data yang besar sehingga dapat menghasilkan informasi yang berguna. Informasi yang dihasilkan berlaku untuk berbagai bidang. Data mining adalah teknik yang digunakan untuk mengungkap pengetahuan dan informasi dari data yang ada sehingga dapat digunakan (Zhou, 2020).

Kumpulan data besar diurutkan dalam penambangan data untuk menemukan pola dan hubungan yang dapat digunakan dalam analisis data untuk membantu memecahkan masalah bisnis. Perusahaan dapat memperkirakan tren masa depan dan membuat keputusan bisnis yang lebih tepat berkat teknik dan alat data mining. Data mining adalah komponen penting dari analitik data secara keseluruhan dan salah satu bidang fundamental dalam ilmu data, yang memanfaatkan metode analitik mutakhir untuk menggali informasi berharga dalam kumpulan data. Data mining, pada tingkat yang lebih rinci, adalah langkah dalam prosedur penemuan pengetahuan dalam database (KDD), metodologi ilmu data untuk mengumpulkan, memproses, dan menganalisis data. Meskipun kadang-kadang digunakan secara bergantian, penambangan data dan KDD lebih sering dipahami sebagai konsep yang terpisah (Alhajj dan Rokne, 2018). Untuk KDD dari *data mining* itu sendiri dapat dilihat pada Gambar 2.1.



Gambar 2.1. Proses Knowledge Discovery in Database (Chiu and Tavella, 2020)

Proses "*start*" dengan menetapkan tujuan KDD dan "*end*" dengan mempraktikkan pengetahuan yang dipelajari. Perulangan kemudian selesai, dan fase *mining* Data Aktif dimulai. Akibatnya, penyesuaian perlu dilakukan pada data yang digunakan.

### 2.4.1. *Domain Understanding and KDD Goals*

Langkah pertama dalam proses persiapan adalah memperoleh pemahaman tentang domain aplikasi atau data. Hal ini bertujuan mengatur panggung untuk menentukan cara menangani banyak keputusan (tentang transformasi, algoritma, representasi, dan lain-lain.). Individu yang bertanggung jawab atas proyek KDD harus memahami dan menentukan tujuan akhir pengguna serta lingkungan yang terkait lainnya (termasuk pengetahuan sebelumnya yang relevan) di mana proses pemahaman pengetahuan akan berlangsung. Bahkan tahap ini dapat direvisi seiring dengan kemajuan proses KDD.

### 2.4.2. *Selection and Addition*

Setelah itu memilih dataset yang pencarian pemahamannya akan dilakukan. Setelah menentukan tujuan dari KDD ini, dataset akan digunakan juga harus ditentukan. Mencari data-data lain yang terkait juga dibutuhkan dalam proses ini termasuk penentuan atribut yang akan digunakan di dalam proses.

### 2.4.3. *Preprocessing, Data Cleaning, etc*

Kemudian dilakukan praproses dan pembersihan data untuk keandalan data yang lebih efektif. Hal tersebut termasuk menghilangkan *noise* dan *outlier* pada data, serta menangani data-data yang kosong yang biasanya disebabkan dari pengisian data atau pendataan yang keliru.

### 2.4.4. *Transformation*

Dilakukan transformasi data untuk menghasilkan *data mining* yang lebih disiapkan dan lebih baik. Proses ini dilakukan dengan reduksi dimensi dan transformasi atribut dengan mendiskritkan atribut numerik dan kategori.

### 2.4.5. *Data Mining*

Memilih jenis *data mining* yang akan dilakukan juga termasuk bagian penting dari KDD. Hal ini mulai berupa menentukan apakah akan menjadi proses klasifikasi, regresi, ataupun klastering. Kemudian memilih algoritma yang digunakan setelah penyesuasian proses yang akan dilakukan, seperti menentukan *supervised* atau *unsupervised*.

Setelah itu melakukan implementasi terhadap algoritma yang telah ditentukan sebelumnya. Proses ini membutuhkan sekiranya beberapa kali percobaan sampai hasil yang memuaskan dapat dicapai. Hal ini karena terdapat beberapa hal yang mempengaruhi hasil dari proses latih data. Yang mempengaruhi hasil dapat berupa parameter yang digunakan pada tiap kali proses latih dilakukan.

### 2.4.6. *Evaluation and Interpretation*

Kemudian dievaluasi hasil yang didapat dapat dengan *plot* grafik hingga nilai *scoring* dari model *machine learning* yang dibuat. Lalu tahapan evaluasi ini dapat dilanjutkan ke tahapan *knowledge discovery*.

### 2.4.7. *Discovery Knowledge*

*Knowledge discovery* dapat dilakukan setelah proses evaluasi dilakukan untuk mendapatkan hasil dan informasi yang diperoleh untuk melihat efek dari perubahan sistem yang dibangun. Keberhasilan tahapan ini bergantung pada seluruh proses KDD yang dilakukan. Apabila terdapat proses yang kurang diperhatikan atau terdapat kekurangan, maka akan berpengaruh ke *discovery* untuk setiap informasinya (Chiu dan Tavella, 2020).

## 2.5. Jaringan Saraf Tiruan

Jaringan saraf tiruan adalah suatu sistem yang melakukan simulasi apa yang dilakukan otak manusia dalam sistem syaraf untuk menjalankan suatu pekerjaan. Sistem ini menggunakan konsep neuron sebagai konsep dasar penggerak syaraf sebagaimana sistem sel syaraf pada manusia bekerja. Pada sistem jaringan sistem syaraf tiruan neuron-neuron tersebar pada lapisan-lapisan yang di antaranya adalah *input layer*, *hidden layer*, dan *output layer* (Fitriadini, Pramiyati dan Pangaribuan, 2020).

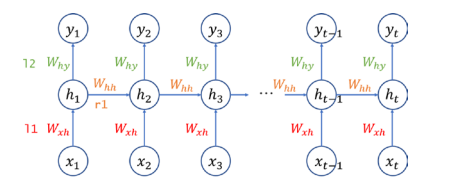
## 2.6. *Deep Learning*

*Deep learning* adalah cabang *machine learning* yang berfokus pada pengembangan algoritme yang menangani berbagai masalah kecerdasan buatan kontemporer dengan sangat mirip dengan versi otak manusia yang terlalu disederhanakan. Deteksi wajah di kamera, koreksi otomatis dan teks prediktif di *keyboard*, aplikasi kecantikan yang disempurnakan dengan AI, asisten pintar seperti *Siri/Alexa/Google Assistant*, *Face*-*ID* (*face unlock* di *iPhone*), saran video di *YouTube*, saran teman di *Facebook* , dan filter kucing di *Snapchat* hanyalah beberapa contoh produk yang dibuat khusus untuk pembelajaran mendalam. *Deep learning* pada dasarnya menembus setiap aspek kehidupan digital kita (Ketkar dan Moolayil, 2021).

*Deep learning* juga pengembangan AI terbaru. Meskipun idenya telah ada untuk sementara waktu, *deep learning* telah mendapatkan popularitas dalam beberapa tahun terakhir karena kemajuan luar biasa yang terus dibuatnya. Sebelumnya fiksi ilmiah, sekarang menjadi kenyataan. *Deep learning* telah membuat teknologi AI lebih meresap dalam kehidupan kita sehari-hari. Saat ini, kecerdasan buatan hadir di sebagian besar produk dan layanan konsumen kami. Mungkin sudah waktunya bagi Anda untuk bergabung dengan revolusi. Anda dapat mulai memberikan kontribusi untuk inisiatif AI ini (Amaratunga, 2021).

## 2.7. *Recurrent Neural Network*

*Recurrent Neural Network* (RNN) adalah bentuk evolusi dari jaringan saraf tiruan yang dapat mengontrol urutan data input. Model RNN menampilkan beberapa gerbang yang dapat mengelola data sekuensial dan menyimpan input sebelumnya. RNN mampu menangani data sekuensial sebelumnya dalam teori, tetapi dalam praktiknya, mereka memiliki masalah dengan memori jangka panjang karena mereka hanya dapat melihat ke belakang sejumlah langkah dan tidak dapat menangani data sekuensial yang panjang. Di antara masalah yang dihadapi RNN adalah gradien yang menghilang dan gradien yang meledak. RNN standar berjuang dengan data pelatihan ulang untuk interval yang panjang dalam gradien yang hilang. Dalam gradien yang meledak, proses pelatihan secara sewenang-wenang memberikan nilai yang lebih besar ke bobot matriks. Model RNN telah dibuat untuk mencegah masalah ini, dan model dengan daya pemeliharaan jangka panjang dan pendek telah disajikan Algoritma ini memberikan nilai yang lebih tinggi ke bobot matriks.



Gambar 2.2. Recurrent Neural Network (Manaswi, 2018)

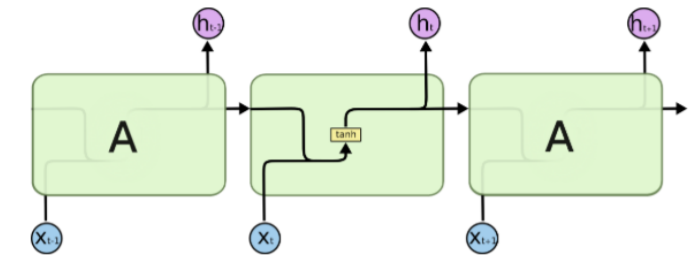
RNN dapat dianggap sebagai jaringan MLP dengan *loop* ditambahkan ke arsitektur. Ada lapisan *input* (dengan node seperti x1, x2, dan seterusnya), *hidden layer* (dengan node seperti h1, h2, dan seterusnya), dan lapisan output, seperti dapat dilihat pada Gambar 2.1. (dengan node seperti y1, y2, dan seterusnya). Arsitektur MLP dan ini sebanding. Node *hidden layer* terhubung, yang membuat perbedaan. Node dalam RNN/LSTM biasa (dasar) terhubung dalam satu arah. Dengan demikian, h2 bergantung pada h1 (dan x2), dan h3 bergantung pada h2 (dan x3). Node sebelumnya di *hidden layer* menentukan node di *hidden layer* (Manaswi, 2018).

## 2.8. *Long Short-Term Memory*

Menurut (Agusta, 2021)Model RNN telah dikembangkan untuk membantu menghindari masalah ini, dan model LSTM baru dengan daya pemeliharaan jangka panjang dan pendek telah diperkenalkan. Tiga *gate* dapat ditambahkan ke model RNN untuk meningkatkan kinerja model LSTM. Ketiga *gate* ini memungkinkan model LSTM untuk mempelajari dependensi jangka panjang sekuensial.

dan merupakan bobot matriks, dan merupakan vektor bias, adalah nilai dari sel memori pada saat t, sebagai nilai dari forget gate layer, menunjukan keadaan sel, menunjukan nilai input *gate*, dan merupakan *output gate layer* (Rezaei, Faaljou dan Mansourfar, 2021).

RNN memiliki satu *layer* sederhana yang digunakan sebagai pengulangan jaringan bernama *layer* tanh pada Gambar 2.2.



Gambar 2.3. Struktur model RNN memiliki satu layer (Agusta, 2021)

*Layer* tanh pada RNN ini dapat dirumuskan dengan:

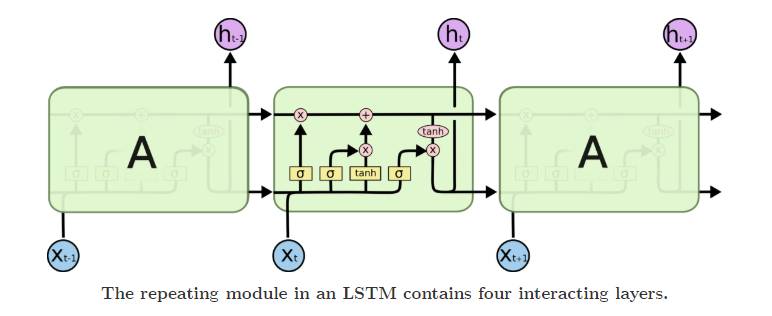
(7)

Keterangan:

Sigmoid

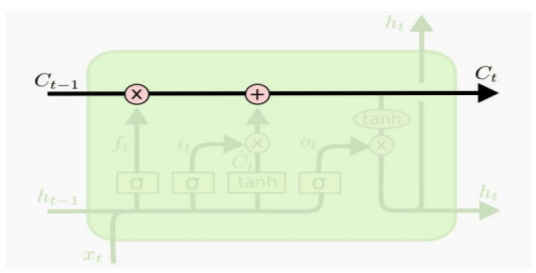
*Input*

Di lain sisi, algoritma LSTM memiliki empat *layer* yang berintraksi ketika melakukan *looping* tiap-tiap model *machine learning*-nya seperti yang tertera pada Gambar 2.3.



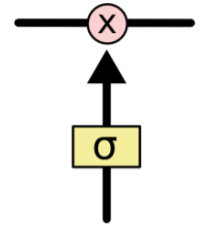
Gambar 2.4. Struktur LSTM dengan empat layer (Agusta, 2021)

Poin terpenting pada LSTM adalah *cell state* yang berupa garis lurus secara horizontal pada tiap *layer* seperti pada Gambar 2.4.



Gambar 2.5. Cell State pada LSTM (Agusta, 2021)

Fungsi dari *gates* pada LSTM adalah sebagai penambah informasi-informasi yang penting dan menghapus informasi-informasi yang sudah tidak digunakan pada *cell state*. Dalam hal ini, yang mengatur informasi akan dilanjutkan ke *state* selanjutnya atau dihentikan adalah salah satu kinerja dari *gates*. Tersusun atas *sigmoid layer*, *gates* memiliki fungsi untuk melakukan pekerjaan tersebut seperti pada Gambar 2.5.



Gambar 2.6. Sigmoid layer pada LSTM (Agusta, 2021)

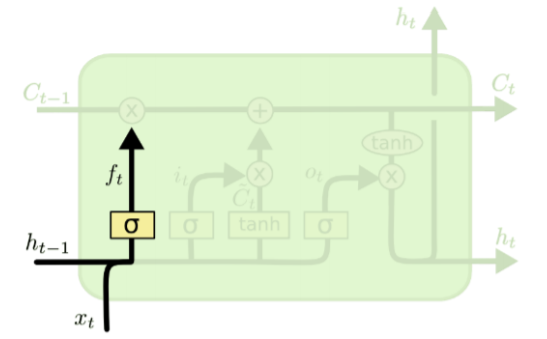
Hasil dari *sigmoid layer* berupa 0 dan 1, dari kedua angka ini memberikan informasi bahwa nilai 0 tidak berisi informasi dan 1 mengirim informasi yang akan diteruskan.

*Gates* terdiri dari tiga pada LSTM, antara lain:

1. *Forget gate* yang berfungsi sebagai penghapus informasi-informasi dari *cell*.
2. *Input gate* yang berfungsi sebagai penentu dan tempat perbaharuan *input*  *memory state*.
3. *Output gate* yang berfungsi sebagai penerima hasil dari *input memory cell*.

### 2.8.1. Tahap *Long Short-Term Memory*

Pertama, LSTM akan masuk ke *forget gate layer* untuk menentukan informasi yang akan dihapus dalam *cell state*. Setelah itu akan diolah dan sebagai nilai *input* yang akan menghasilkan *output* berupa 0 dan 1 sesuai Gambar 2.6.



Gambar 2. 7. Forget gate layer pada LSTM (Agusta, 2021)

*Forget gate* dapat dirumuskan dengan:

(8)

Dengan:

= *forget gate*

= sigmoid

= bobot

= *input* ke-t

= *output* sebelum ke-t

= nilai *bias*

Selain itu, bobot dapat dirumuskan dengan:

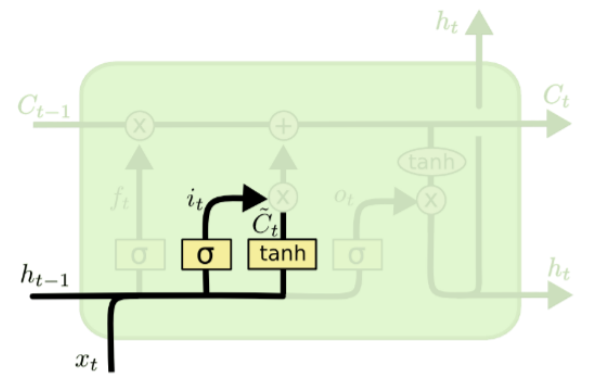
(9)

Dengan:

= bobot

= jumlah data

Langkah kedua dalam LSTM adalah memahami informasi apa yang disimpan di setiap *cell state*. Bagian ini memiliki dua lankah. Pertama, *input gate layer* mengidentifikasi lokasi sinyal baru dengan fungsi lapisan atas. Selanjutnya, *cell state* menambahkan sinyal baru. Komponen kedua adalah mengatur keadaan sel berdasarkan hasil dari *input gate layer* dan *tanh layer*. Langkah ini sesuai dengan Gambar 2.7.



Gambar 2.8. Input gate layer dan tanh layer pada LSTM (Agusta, 2021)

*Input gate* dapat dirumuskan dengan:

(10)

Dengan:

= *input gate*

𝜎 = sigmoid

= bobot

= *output score* sebelum ke-t

= *input score* ke-t

= nilai bias

Nilai baru yang didapatkan dapat dirumuskan dengan:

(11)

Dengan:

= nilai baru

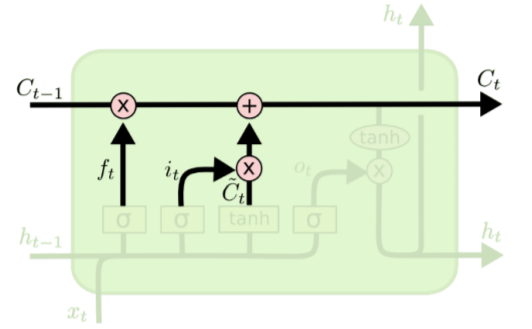
= bobot

= *output score* sebelum ke-t

= *input score* pada ke-t

= nilai bias

Langkah keempat dalam LSTM adalah mengubah *cell state* menjadi *cell state* baru. Setelah itu, *cell state* akan dikali dengan *cell state* sebelumnya menggunakan informasi yang telah ditentukan oleh *forget gate layer*. Hasil akan ditambahkan dengan nilai paling baru sesuai dengan Gambar 2.8.



Gambar 2.9. Cell state terbaru pada LSTM (Agusta, 2021)

*Cell state* dapat dirumuskan dengan:

(12)

Dengan:

= *cell state*

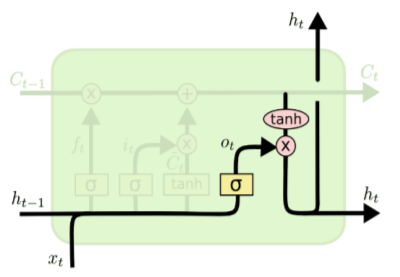
= *forget gate*

= *cell state* sebelum ke-t

= *input gate*

= nilai baru

Langkah terakhir dari keempat tahapan LSTM adalah menghitung hasil *output*. Berdasarkan *cell state* sebelumnya, diperoleh hasil. Lapisan sigmoid awal menentukan *cell state* untuk *output*; hasil keadaan sel kemudian ditransfer ke lapisan berikutnya dan digabungkan dengan gerbang sigmoid untuk menghasilkan *output* yang konsisten dengan putusan sebelumnya sesuai dengan Gambar 2.9.



Gambar 2.10. Output layer pada LSTM (Agusta, 2021)

Rumusan *output gate* sebagai berikut:

(13)

Dengan:

= *output gate*

𝜎 = sigmoid

= nilai bobot untuk *output gate*

= *output score* sebelum ke-t

= *input score* pada ke-t

= nilai bias

Rumusan untuk mendapatkan nilai *output* ke-t seperti pada persamaan (8).

(14)

## 2.9. *Auto Regressive Fractionally Integrated Moving Averages* (ARFIMA)

Granger dan Joyeuk pertama kali memperkenalkan model ARFIMA pada tahun 1980 untuk mengatasi kekurangan dari pendekatan *Autoregressive Integrated Moving Averages* (ARIMA). ARFIMA dapat menjelaskan data deret waktu jangka panjang sedangkan ARIMA hanya dapat menjelaskan data deret waktu jangka pendek (*short memory*). Metode ARFIMA terkadang menghasilkan hasil yang lebih unggul daripada teknik peramalan deret waktu lainnya (Kondo Lembang, Sinay dan Irfanullah, 2021).

Menurut (Shalalfeh, Bogdan and Jonckheere, 2019), ARFIMA itu sendiri sebenarnya hanya berbeda pada *differencing operator*-nya yang dapat berupa nilai non integer. Dari penggunaan nilai non integer inilah yang membuat metode ini untuk *long time series*. ARFIMA dapat dirumuskan sebagai:

(15)

Dengan:

= *backshift operator*

= *auto regressive parameter*

= *moving average parameter*

= ordo *auto regressive*

= ordo *moving average*

= data *time series*

= error pada waktu ke *t*

### 2.9.1. *Time Series*

Berdasarkan (Ambach and Ambach, 2018), dalam data *time series* perlu dilakukan analisis skalar pada jangka waktu tertentu untuk dipelajari polanya. Pola-pola tersebut dapat dipecah menjadi empat bagian, antara lain:

1. Konstan yang merupakan data mean yang membentuk garis lurus dalam jangka waktu tertentu.
2. *Trend* yang merupakan data jangka panjang yang biasanya dalam bentuk setahun dan terdapat naik turunnya data dalam waktu tertentu.
3. *Seasonal* yang merupakan pola berulang-ulang yang konstan dalam periode tertentu (biasanya per tiga sampai enam bulan) yang biasanya dipengaruhi oleh faktor lingkungan maupun kebutuhan pada waktu tertentu.
4. *Residual* merupakan suatu titik data yang biasanya disebabkan oleh sesuatu yang tidak jelas dan dari *plot* tersebut tidak dapat dilihat pola yang dapat dipelajari karena sangat acak.

#### 2.9.1.1. *Resampling Mean*

Untuk mendapatkan pola-pola *time series* yang telah disebutkan, perlu dilakukan *resampling* data dari perharian transaksi menjadi perbulan. Untuk *resampling mean* biasanya digunakan untuk atribut penutupan karena tujuannya adalah menngetahui laju besarnya harga transaksi saham (Lu *et al.*, 2021). Dalam (Borges and Neves, 2021), untuk *resampling mean* dapat menggunakan persamaan.

(16)

Keterangan:

= data mean

= data asli pada waktu *t*

= jumlah total harian transaksi bulan ke *i*

### 2.9.2. *Differencing*

Proses *differencing* dilakukan dengan tujuan menstasionerkan data yang belum stasioner. Salah satu yang tepat untuk menggambarkan proses *differencing* adalah Operator *shift* (*backward shift*). Untuk penggunaan *backward shift* adalah sebagai berikut:

(17)

Keterangan:

= nilai X pada saat *t*

= nilai X pada waktu *t-*1

= *backward shift*

*Backshift operator* ini memiliki pengaruh yang dapat menggeser sebuah data dalam satu waktu sebelumnya. Misal, ketika sebuah data *time series* tidak stasioner, maka data perlu dibuat stasioner dengan cara *differencing*. Berikut merupakan rumus *differencing*, yaitu:

(18)

Keterangan:

= nilai X setelah dilakukan *differencing*

Atau dapat menggunakan *backward shift*, persamaan (11) dapat menjadi

(19)

*Differencing* pada akhirnya dapat dilakukan dengan (Sharma *et al.*, 2021).

### 2.9.3. *Auto Correlation Function*

Bedasarkan (Trierweiler Ribeiro *et al.*, 2021), perhitungan ini merupakan bagian yang penting dalam pemodelan *time series* menggunakan model ARMA yang pada penelitian ini adalah ARFIMA. Perhitungan ini berfungsi untuk menentukan nilai *q* sebagai parameter ketika data sudah stasioner. Persamaan ACF dapat dirumuskan dengan

(20)

Keterangan:

= koefisien auto korelasi pada *lag* ke-*k*

= data pada waktu ke *t*

= nilai rata-rata data

= jumlah data

*Plotting* pada ACF ataupun PACF juga memerlukan batasan dengan *convidence interval* sebagai area yang menjadi batasan regresi tiap *lag*. Untuk menghitung nilai *convidence interval* dapat menggunakan persamaan.

(21)

Keterangan:

= *convidence interval*

= nilai rata-rata

= *standard error*

= nilai Z-score

Untuk mencari nilai *standard* *error* yang merupakan ukuran variansi sampel statistik dari rata-rata sampel. Ini menunjukkan seberapa dekat rata-rata sampel dari populasi yang mewakili. Semakin kecil nilai *standard error*, semakin baik rata-rata sampel mencerminkan nilai rata-rata populasi. Untuk menghitungnya dapat menggunakan persamaan deviasi standar dibagi dengan akar kuadrat dari jumlah sampel.

(22)

Keterangan:

= deviasi standar

= jumlah sampel

Untuk mencari nilai deviasi standar yang merupakan ukuran variansi atau spread dari sekumpulan data. Ini mengukur seberapa jauh setiap data dari nilai rata-ratanya. Semakin besar deviasi standar, semakin besar variansi dalam data, dan sebaliknya. Deviasi standar juga digunakan dalam statistik untuk menentukan seberapa baik suatu model memprediksi data. Untuk menghitungnya dapat dilakukan dengan persamaan.

(23)

Keterangan:

= deviasi standar

= data sampel

Untuk mencari nilai *z-score* yang merupakan nilai yang mengukur seberapa jauh nilai individu dari rata-rata populasi dalam satuan deviasi standar. Ini membantu untuk membandingkan nilai individu dengan rata-rata populasi dan menentukan seberapa unik atau biasa nilai tersebut. Nilai *z-score* positif menunjukkan bahwa nilai individu lebih besar dari rata-rata populasi, sedangkan nilai *z-score* negatif menunjukkan bahwa nilai individu lebih kecil dari rata-rata populasi (Heng, Longfu and Kaiyou, 2020). Untuk menghitung *z-score* dapat dilakukan dengan tabel standar normal yang dapat digunakan persamaan.

(24)

### 2.9.4. *Partial Auto Correlation Function*

Berdasarkan (Trierweiler Ribeiro *et al.*, 2021), perhitungan PACF memiliki fungsi untuk menandakan besar korelasi parsial tiap-tiap pengamatan *lag* pada waktu tertentu dan juga sebagai penentu parameter *p* pada model ARMA. Untuk mencari koefisien korelasi parsial dapat digunakan persamaan

(25)

Keterangan:

= koefisien korelasi parsial pada *lag* parsial ke-*j*

= parameter regresi ke-*k*

### 2.9.5. *Augmented Dickey-Fuller Test*

Menurut (Shalalfeh, Bogdan and Jonckheere, 2019), untuk mengecek stasioneritas dapat menggunakan *Augmented Dickey-Fuller Test*. Uji ADF ini dilakukan untuk mengetahui apakah data perlu di-*differencing* atau tidak. ADF dapat digunakan persamaan.

(26)

Keterangan:

= nilai *different* pertama dari Y

= nilai konstan

= koefisien *trend*

= koefisien *lag* Y

= *error*

### 2.9.6. *Fractional Differencing*

Menurut (Saleh, Grivel and Omar, 2018), *Fractional differencing* adalah proses mengubah seri waktu agar menjadi stasioner (memiliki mean dan varians yang konstan dalam waktu). Ini dilakukan dengan membuat differensi sebagian (fractional difference) dari seri waktu. Proses ini sangat berguna dalam analisis time series karena seri waktu yang stasioner lebih mudah diprediksi dan dianalisis.

Untuk menghitung *fractional differencing*, pertama-tama harus ditentukan nilai *fractional differencing* yang akan digunakan. Nilai ini biasanya didapat dengan menggunakan metode seperti *Auto-Regressive Fractional Integration Moving Average* (ARFIMA) atau penentuan parameter melalui *grid search*. Setelah nilai *fractional differencing* ditentukan, seri waktu dapat diferensiasi dengan menggunakan persamaan.

(27)

Keterangan:

= data setelah *fractional differencing*

= data pada waktu ke *t*

*=* operator *lag* yang diambil dari *shifting* data

= nilai *fractional differencing*

## 2.10. Normalisasi

Normalisasi sangat penting selama pemrosesan data untuk mengubah beberapa variabel sehingga memiliki rentang nilai yang konsisten dengan tujuan membuat analisis statistik menjadi lebih mudah. Teknik normalisasi paling sederhana disebut MinMax Scaling, dan digunakan untuk mengubah data menjadi informasi dengan rentang nilai antara 0 dan 1. Metode ini menciptakan hubungan antara data sebelumnya dan proses saat ini (Agrawal, 2021).

(28)

Dimana:

= data baru

= data lama

= data terendah

= data tertinggi

Setelah normalisasi dilakukan data dapat dimasukan ke proses latih. Untuk mengembalikan nilai datanya dapat digunakan persamaan (29).

(29)

## 2.11. *Mean Absolute Percentage Error*

Menurut (Swanjaya dan Putra Pamungkas, 2021), Indikator kinerja lain untuk model regresi disebut MAPE, dan memiliki definisi yang relatif sederhana dalam hal *relative error*. Karena itu, penerapannya disarankan untuk tugas-tugas di mana lebih penting untuk peka terhadap fluktuasi relatif daripada variasi absolut. MAPE juga memiliki sejumlah kelemahan, salah satunya adalah bahwa MAPE hanya dapat digunakan dengan data yang benar-benar positif secara definisi dan bias terhadap prakiraan rendah, sehingga tidak sesuai untuk model prediktif di mana *error* yang besar diantisipasi. MAPE dapat dirumuskan dengan:

(30)

Dengan:

= *Mean Absolute Percentage Error*

= jumlah iterasi

= nilai asli

= nilai perkiraan

## 2.12. Symmetric Mean Absolute Percentage Error

Menurut (Chicco, Warrens dan Jurman, 2021), SMAPE diajukan untuk mengatasi kekurangan metrik MAPE. SMAPE pertama kali didefinisikan oleh Armstrong (1985) dan kemudian diperbaiki dalam bentuknya yang sekarang oleh Flores (1986) dan Makridakis (1993). Namun, ada sedikit kesepakatan tentang formula SMAPE definitif, dan berbagai penulis terus menggunakan versi yang sedikit berbeda. Nilai maksimum dalam rumus SMAPE asli adalah 200%, yang secara matematis setara dengan 2. SMAPE dapat dirumuskan dengan:

(31)

Dengan:

= *Symmetric Mean Absolute Percentage Error*

= jumlah iterasi

= nilai asli

= nilai perkiraan

## 2.13. *R-Squared*

Menurut (Plevris *et al.*, 2022), Koefisien determinasi atau *r-squared* adalah proporsi varians variabel dependen yang diprediksi oleh variabel independen. *R-squared* atau koefisien determinasi dapat dirumuskan sesuai RSS dan TSS dengan:

(32)

Dengan:

= nilai ke-i variabel yang diprediksi

= nilai yang terprediksi

= batas perhitungan

= nilai rata-rata

= *Residual Sum of Squares*

= *Total Sum of Squares*

Atau dapat sesuai dengan ESS and TSS:

(33)

Dengan:

= nilai ke-i variabel yang diprediksi

= nilai yang terprediksi

= batas perhitungan

= nilai rata-rata

= *Explained Sum of Squares*

= *Total Sum of Squares*

Cara ESS adalah *squared error* yang digunakan dalam model linear dan TSS adalah keseluruhan dari *squared error*. Dari persamaan (33) dapat disimpulkan bahwa koefisien determinasi atau *r-squared* merupakan rasio dari varians dari model liner ke jumlah total varians. Semakin besar nilai dari *r-squared*, semakin baik model yang dibuat.

## 2.14. Penelitian Relevan

Dalam penelitian sebelumnya yang telah dilakukan oleh peneliti-peneliti terdahulu dengan menggunakan metode-metode yang berbeda sebagai acuan penelitian ini. Berikut merupakan beberapa penelitian terdahulu yang relevan dengan penelitian ini.

Pertama, Fajar Akbardipura (2021) dengan penelitiannya yang berjudul “Perbandingan Performa Peramalan Harga Saham 5 Perusahaan Pada Indeks Lq45 Menggunakan *Autoregressive Integrated Moving Average* dan *Long Short – Term Memory*”. Penelitian ini menggunakan *Autoregressive Integrated Moving Average* dan *Long Short – Term Memory* yang membandingkan model prediksi saham ANTM, ADRO, ICBP, KLBF, dan TLKM dari bulan Januari 2021 sampai 2022. Model pertama yang menggunakan ARIMA parameter ARIMA(0,1,0) dengan data saham ICBP menhasilkan MAPE sebesar 0,0090. Model kedua yang menggunakan LSTM dengan data saham ICBP yang berparameter *epoch* dan *batch size* menghasilkan MAPE sebesar 0,00812. Dari hasil penelitian ini dapat disimpulkan bahwa model LSTM yang dibuat menghasilkan model yang lebih baik karena tingkat kesalahan lebih sedikit dari model ARIMA. Penelitian ini juga memberikan kesimpulan bahwa LSTM dapat menangani data *time series* dengan lebih baik dari model ARIMA (Akbardipura, 2021).

Kedua, Ardiyan Agusta (2021) dengan penelitiannya yang berjudul ”Prediksi Pergerakan Harga Saham Pada Sektor Farmasi Menggunakan Algoritma *Long Short-Term Memory*”. Penelitian ini menggunakan metode LSTM yang bertujuan untuk mencari parameter terbaik dalam pembuatan model prediksinya dengan menggunakan data saham PT Kalbe Farma Tbk periode Januari 2018 sampai Desember 2020. Model LSTM yang dibuat dengan menggunakan parameter *epoch* saja menghasilkan RMSE sebesar 3,46314 dengan waktu komputasi selama 108,88 detik. Model kedua LSTM dengan menggunakan parameter *batch size* dan *epoch* menghasilkan RMSE sebesar 1,9879 dengan waktu komputasi selama 176,93 detik. Dari penelitian ini dapat disimpulkan bahwa penggunaan *batch size* dan *epoch* dengan lebih spesifik 16 dan 200 menghasilkan model yang paling optimal dalam diterapkannya pada data saham PT Kalbe Farma Tbk. meskipun memerlukan waktu komputasi yang lebih lama (Agusta, 2021).

Ketiga, Puspita Kartikasari (2020) dengan penelitiannya yang berjudul “Prediksi Harga Saham PT Bank Negara Indonesia Dengan Menggunakan Model *Autoregressive Fractional Integrated Moving Average* (ARFIMA)”. Penelitian ini menggunakan metode ARFIMA yang bertujuan untuk mencari model prediksi terbaik dari data saham PT Bank Negara Indonesia periode Februari 2016 sampai Februari 2020. Model yang dibuat menggunakan ARFIMA(0,0.499,5) menggunakan data saham yang telah dilakukan stasioneritas dalam varians menghasilkan AIC sebesar 15997,52. Karena residual model tidak memenuhi asumsi distribusi normal dan *white noise* dengan banyaknya *outlier* pada data tersebut. Namun dari peramalan 12 periode ke depan dengan model ARFIMA(0,0.499,5) dengan data asli menghasilkan SMAPE sebesar 1,23%. Dari penelitian ini dapat disimpulkan bahwa metode ARFIMA dengan parameter ARFIMA(0,0.499,5) dan menggunakna data asli menghasilkan hasil yang optimal dalam peramalan saham (Kartikasari, 2020).

Keempat, Arjun Singh Saud dan Subarna Shakya (2019) dengan penelitiannya yang berjudul “*Analysis of Look Back Period for Stock Price Prediction with RNN Variants: A Case Study on Banking Sector of NEPSE*”. Penelitian ini menggunakan LSTM dan GRU dengan data saham *Nepal Stock Exchange* dari *Nepal Investment Bank* (NIB) dan *Nabil Bank Limited* (NABIL). Model pertama yang menggunakan GRU dari NIB dan NABIL menghasilkan MAPE sebesar 4,74 dan 4,71. Model kedua yang menggunakan LSTM dari NIB dan NABIL menghasilkan MAPE sebesar 5,58 dan 5,06. Dari penelitian ini dapat disimpulkan bahwa model yang menggunakan metode GRU lebih optimal dari model yang menggunakan metode LSTM dalam data saham bank NIB dan NABIL (Saud and Shakya, 2020).

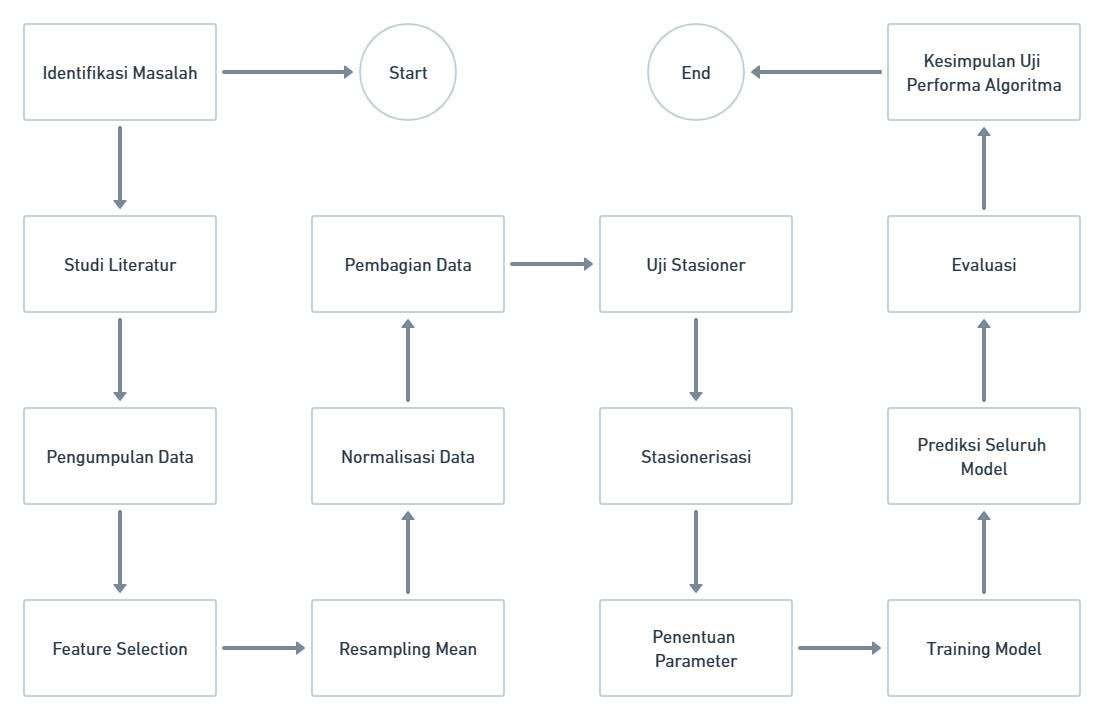
Kelima, Dehua Zhang dan Sha Lou (2020) dengan judul “*The Application Research of Neural Network and BP Algorithm in Stock Price Pattern Classification and Prediction*”. Penelitian ini menggunakan metode *Backpropagation* dari data pola saham. Model pertama yang menggunakan *deep learning* *fuzzy* menghasilkan akurasi prediksi sebesar 62,12%. Model kedua yang menggunakan *backpropagation* menghasilkan akurasi prediksi 73,29%. Dari penelitian ini dapat disimpulkan bahwa model yang menggunakan *backpropagation* lebih akurat dan optimal dibandingkan dengan metode *deep learning fuzzy* dalam data pola saham (Zhang and Lou, 2021).

# BAB III

**METODOLOGI PENELITIAN**

## 3.1. Tahapan Penelitian

Berdasarkan penelitian-penelitian yang terkait sebelumnya, penelitian ini disusun dengan prosedur penelitian yang berupa alur penelitian ini akan dilaksanakan. Alur penelitian ini berisi gambaran umum dari penelitian mulai dari identifikasi masalah sampai dengan evaluasi penelitian. Tahapan penelitian ini sesuai dengan yang tertera pada Gambar 3.1.



Gambar 3.1. Tahapan Alur Metodologi Penelitian

## 3.2. Identifikasi Masalah

Pada tahapan penelitian ini adalah melakukan pencarian terhadap permasalah yang akan menjadi subjek dari penelitian. Dari hal ini akan dapat dirumuskan permasalahan yang akan ditelaah lebih lanjut dalam penelitian yang akan dilakukan. Setelah masalah ditemukan, penelitian ini akan dibatasi dengan mempersempin ruang lingkup lebih spesifiknya karena subjek dari penelitian ini adalah data saham, maka data saham yang digunakan hanya data saham dari Bank Raya Indonesia Tbk. dan Bank IBK Indonesia Tbk. dengan membandingkan hasil dari evaluasi menggunakan metode ARFIMA dan LSTM.

## 3.3. Studi Literatur

Tahapan penelitian ini mencari teori-teori berdasarkan buku dan jurnal yang terkait mengenai saham, *deep learning*, LSTM, ARFIMA, sampai dengan SMAPE dan RMSE mengenai informasi dan rumusan dari teori yang telah dijabarkan untuk melakukan perhitungan sebagai hasil dari evaluasi dari metode yang digunakan dalam pembuatan model *deep learning*. Tahapan ini merupakan tahapan yang penting sebagai acuan penelitian mulai dari mengetahui bagaimana suatu metode bekerja dan bagaimana metode tersebut dan cara evaluasinya itu dilakukan komputasi untuk mendapatkan angka yang menjadi acuan dari proses prediksi suatu model (Jing, Wu dan Wang, 2021). Dari studi literatur ini didapatkan informasi untuk menopang metode penelitian.

## 3.4. Pengumpulan Data

Proses pengumpulan data adalah komponen penting pertama dari sistem yang akan dibangun. Data dapat diperoleh dari sumber luar, melalui penyedia data, atau melalui penelitian internal. Struktur data keuangan secara luas dapat dibagi menjadi dua kategori: data *cross-sectional* dan data *time series*. Kegiatan utama untuk menyiapkan sumber data yang dapat dipercaya dalam penyimpanan ke *database* adalah pembersihan dan persiapan data. Membangun prediksi yang tepat, analisis statistik, dan model optimasi adalah fokus inti dari proses pemodelan data *machine learning* (Ta, Liu dan Tadesse, 2020).

Tahapan penelitian ini adalah melakukan pencarian terhadap data saham yang akan dilakukan pembuatan model prediksi. Pada tahapan ini digunakan data *time series* menggunakan data saham dari IDX untuk mendapatkan data saham dari perusahaan-perusahaan yang ada di Indonesia. Data saham yang diambil dari IDX ini adalah data saham perusahaan yang terdapat di Indonesia pada Periode Januari 2022 sampai dengan Oktober 2022 berdasarkan harian transaksi saham.

Dalam pengumpulan data ini, data yang akan digunakan adalah data dari perusahaan Bank Rakyat Indonesia (Persero) Tbk. dan Bank IBK Indonesia Tbk. dengan kolom-kolom antara lain Kode Saham, Nama Perusahaan, *Remarks*, Sebelumnya, Open Price, *First Trade*, Tertinggi, Terendah, Penutupan, Selisih, Volume, Nilai, Frekuensi, Index Individual, *Listed Shares*, *Offer*, *Offer Volume*, *Bid*, *Bid Volume*, *Last Trading Date*, *Tradeble Shares*, *Weight for Index*, *Foreign Sell*, *Foreign Buy*, *Non Regular Volume*, *Non Regular Value*, dan *Non Regular Frequency*.

## 3.5. Praproses Data

Tahapan penelitian ini melakukan beberapa proses data sebelum dilakukan pembuatan model dengan melakukan *resampling mean*, uji stasioner, stasionerisasi, dan normalisasi data dengan *MinMaxScaler*. Tahapan ini akan memudahkan proses latih data serta dapat mengurangi waktu komputasi dari model.

### 3.5.1. *Feature Selection*

Tahapan ini dilakukan dengan cara menghilangkan atau menghapus atribut kolom yang sekiranya tidak akan mempengaruhi dengan melihat korelasi dengan kolom targetnya. Atribut yang memiliki korelasi rendah ini hanya akan menambah waktu komputasi dari proses latih data yang seharusnya dapat lebih singkat. Pemilihan atribut juga diperlukan dengan penyesuaian dan kebutuhan dari data dengan memperhatikan metode ARFIMA dan LSTM. Setelah ditelaah dari jurnal penelitian terdahulu, atribut yang akan digunakan pada penelitian ini adalah Penutupan karena merupakan harga terakhir pada pasar di hari tersebut (Lu *et al.*, 2021).

### 3.5.2. *Resampling Mean*

Tahapan ini dilakukan dengna mengubah data yang perharian transaksi jual beli saham menjadi rata-rata transaksi perbulan. Diubahnya data menjadi mean bertujuan untuk mendapatkan *insight* laju penjualan saham pada bulan tertentu karena data yang digunakan merupakan atribut penutup (Borges and Neves, 2021).

### 3.5.3. Normalisasi Data

Untuk mengubah beberapa variabel selama pemrosesan data sehingga memiliki rentang nilai yang disesuaikan untuk tujuan penyederhanaan analisis statistik, normalisasi sangat penting. *MinMax Scaling* adalah metode normalisasi yang paling sederhana, dan digunakan untuk mengubah data menjadi informasi dengan rentang nilai antara 0 dan 1. Teknik ini membuat hubungan antara informasi historis dan prosedur yang sedang berjalan (Agrawal, 2021).

### 3.5.4. Pembagian Data

Pada tahapan ini dilakukan pembagian data antara data latih dan data uji dengan penentuan target dan atributnya. Pembagian data akan menggunakan metode *hold out* cara memisahkan data uji dan data latih berdasarkan tanggal sebagai index (Chen and Zhou, 2021). Pembagian data ini juga akan dilakukan tiga kali dengan pembagian data pertama untuk data latihnya adalah 80%, yaitu dari September 2019 hingga April 2022 dan data ujinya adalah 20% dari Agustus 2022 hingga Desember 2022. Pembagian data kedua untuk data latihnya adalah 60%, yaitu dari September 2019 hingga Agustus 2021 dan data ujinya adalah 40%, yaitu dari Agustus 2021 hingga Desember 2022. Pembagian data ketiga untuk data latihnya adalah 20%, yaitu September 2019 hingga April 2020 dan data ujinya adalah 80%, yaitu dari Mei 2020 hingga Desember 2022.

### 3.5.5. Uji stasioner

Tahapan ini dilakukan dengan tujuan mengetahui apakah data saham yang digunakan sudah stasioner atau belum stasioner. Tahapan ini dilakukan untuk penggunaan metode ARFIMA karena metode ini perlu data yang stasioner. Stasioner yang dimaksud adalah bentuk garis lurus pada *plot* grafik dataset saham. Apabila grafik tidak berada di garis lurus, maka perlu dilakukan stasionerisasi. Jika sudah ada pada garis lurus, maka data saham sudah stasioner. Perlu diketahui juga, untuk uji stasioner hanya dilakukan pada metode ARFIMA saja karena uji stasioner bertujuan untuk menentukan parameter ARFIMA yang digunakan (Kartikasari, 2020).

### 3.5.6. Stasionerisasi

Dalam pembuatan model ARFIMA, data *time series* yang tidak stasioner harus dilakukan stasionerisasi dengan cara *differencing*. Tahapan ini menghitung nilai selisih dari nilai observasi dari data *time series*. Selisih ini perlu dilihat dengan *augmented dickey fuller test* pada tahapan uji stasioner untuk melihat apakan sudah kurang dari 0.05. Apabila masih belum stasioner, maka masih perlu dilakukan *differencing* kembali. Apabila sudah stasioner, maka bisa langsung dilakukan pembagian data untuk dilatih pada model yang akan dibuat (Ambach and Ambach, 2018).

## 3.6. Penentuan Parameter

Tahapan ini dilakukan dengan menetukan parameter yang sekiranya sesuai untuk mendapatkan hasil yang optimat dari metode ARFIMA dan LSTM. Metode ARFIMA dapat menggunakan parameter nilai desimal dan . Parameter ARFIMA ini sendiri didapatkan berdasarkan *plot* PACF dan ACF untuk dilihat *lag* yang signifikan pada saat stasioner untuk *p* (jumlah autoregresif) dan *q* (jumlah differensi non-seasonal) (Kartikasari, 2020). Apabila dilakukan *differencing* sekali untuk stasionerisasi, maka parameter *d* adalah 1 dan begitu juga seterusnya. Metode LSTM dapat menggunakan parameter fungsi LSTM *epoch*. Berdasarkan penelitian-penelitian terkait sebelumnya, parameter yang digunakan akan mempengaruhi hasil dan waktu komputasi dalam proses latih data dari pembuatan suatu model (Jing, Wu and Wang, 2021).

## 3.7. *Training* *Model*

Untuk pemodelan ARFIMA sendiri, hanya perlu melakukan *fractional differencing* pada data yang akan digunakan kemudian di latih dengan metode SARIMAX dengan parameter sesuai yang telah didapatkan pada penetuan parameter. Untuk pemodelan LSTM dilakukan dengan inisialisasi parameter berupa *epoch*. Pada penelitian ini fungsi aktivasi uang digunakan adalah tanh karena akan memaksimalkan target dengan interval -1 sampai 1. Model LSTM akan berjalan di atas *gates*, yaitu *forget gate layer*, *input layer*, dan *tanh layer*. Dari pengujian ini data akan berhenti pada *epoch* yang ditentukan pada parameter sebelumnya. Pembuatan model akan dilakukan untuk melihat hasil penggunaan parameter yang paling optimal dan pencatatan waktu komputasi yang diperlukan pada model yang menggunakan LSTM.

## 3.8. Prediksi Seluruh Model

Setelah proses latih, dilakukan prediksi dengan menggunakan kedua model terbaik yang sudah dibuat. Prediksi dilakukan dengan melihat *plot* grafik data saham uji dibandingkan hasil *forecasting* dari model prediksi yang dibuat (Hu, Zhao and Khushi, 2021).

## 3.9. Evaluasi

Tahapan ini bertujuan melihat perbandingan antara kedua model yang dibuat karena dari *plot* grafik saja tidak akan mendapat *insight* dari informasi hasil prediksi dari model. Tahapan ini dilakukan dengan melihat metrik SMAPE dari kedua model (Kartikasari, 2020).

## 3.10. Kesimpulan Uji Performa Algoritma

Tahapan terakhir ini, yaitu melihat hasil dari evaluasi yang telah dilakukan dan mempertimbangkan model terbaik berdasarkan metrik SMAPE dengan memperhatikan juga *plot* grafik yang membandingkan pertumbuhan sebenarnya dengan pertumbuhan yang diramal oleh model. Metode yang memiliki SMAPE yang lebih kecil dari metode yang satunya akan menjadi metode yang paling optimal untuk menyelesaikan masalah prediksi saham dengan data *time series* jangka panjang.

## 3.11. Waktu dan Tempat Penelitian

Adapun waktu dan tempat pelaksanaan dalam melakukan penilitian ini adalah sebagai berikut:

Tempat : Bank Rakyat Indonesia Tbk. dan Bank IBK Indonesia Tbk.

Waktu : 13 Oktober 2022 - 31 Maret 2023

## 3.12. Alat Bantu Penelitian

Penelitian ini memerlukan beberapa perangkat keras dan perangkat lunak untuk mendukung penelitian yang dilakukan. Berikut merupakan perangkat-perangkat pendukung yang digunakan pada penelitian ini:

### 3.12.1. Perangkat keras

Berikut merupakan spesifikasi perangkat keras yang digunakan berupa komputer personal:

1. *Motherboard* : *Z730 Tomahawk (MS-7B47)*
2. *Processor* : *Intel(R) Core(TM) i7-8700CPU@3.20GHz (12 CPUs)*
3. *Memory* : 16384MB RAM
4. *Hard Disk* : 2 TB + 500 GB
5. *GPU* : *Galax RTX 3060 8 GB 1777 MHz*

### 3.12.2. Perangkat Lunak

Berikut ini merupakan perangkat lunak yang digunakan dalam pengerjaan penelitian:

1. Sistem Operasi : *Windows 10 Pro 64-bit*
2. *Text Editor* : *Visual Studio Code*
3. *Data Studio* : *Microsoft Excel*
4. *Online IDE* : *Google Colab Jupyter Notebook*
5. *CUDA Version* : *Cuda v11.2.0.460.89*
6. *Cudnn Version* : *Cudnn v10.2*
7. *Keras Version* : *Keras v2.10*
8. *Tensoflow* : *Tensorflow v2.10 Windows CUDA GPU Supported*

## 3.13. Jadwal Penelitian

Penelitian ini dilakukan secara teratur dan dengan susunan penjadwalan penelitian yang sistematis dari kegiatan pertama pengumpulan data sampai selesai evaluasi dan didokumentasinya hasil berupa laporan penelitian dihitung dari Oktober 2023 hingga Maret 2023 dengan penjadwalan sebagai berikut.

Tabel 3.1. Jadwal Penelitian

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Kegiatan | Bulan Ke | | | | | | | | | | | | | | | | | | | | | | | |
| I | | | | II | | | | III | | | | IV | | | | V | | | | VI | | | |
| 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 |
| Identifikasi Masalah |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Studi Literatur |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Pengumpulan Data |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| *Feature Selection* |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| *Resampling Mean* |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Normalisasi Data |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Pembagian Data |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Uji Stasioner |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Stasioner |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Penentuan Parameter |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| *Training Model* |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Prediksi Seluruh Model |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Evaluasi |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Kesimpulan Uji Performa Algoritma |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Dokumentasi |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

Kegiatan terkait sedang dijalankan

Kegiatan istirahat

# BAB IV

**HASIL DAN PEMBAHASAN**

## 4.1. Data

Data yang digunakan dalam penelitian ini adalah data saham yang diperoleh dari *website* IDX. *Website* ini merupakan *website* resmi regulator dan pengurus perdagangan pada Pasar Modal Indonesia yang dimiliki oleh Bursa Efek Indonesia (*IDX Data Services Portal*, 2022). Penelitian ini menggunakan 2 data saham perusahaan bank yang digunakan dalam 3 tahun terakhir yaitu data Bank Rakyat Indonesia Tbk. dan Bank IBK Indonesia Tbk. Data kedua bank tersebut memiliki 27 kolom, yaitu Kode Saham, Nama Perusahaan, *Remarks*, Sebelumnya, Open Price, *First Trade*, Tertinggi, Terendah, Penutupan, Selisih, Volume, Nilai, Frekuensi, Index Individual, *Listed Shares*, *Offer*, *Offer Volume*, *Bid*, *Bid Volume*, *Last Trading Date*, *Tradeble Shares*, *Weight for Index*, *Foreign Sell*, *Foreign Buy*, *Non Regular Volume*, *Non Regular Value*, dan *Non Regular Frequency*. Data berjumlah masing-masing 1029 dan 862.

Atribut-atribut yang disebutkan memiliki definisi dan fungsinya masing-masing, yaitu:

1. Kode Saham, yang merupakan kode unik yang mengidentifikasi saham tertentu yang terdaftar di bursa saham.
2. Nama Perusahaan, yang merupakan perusahaan yang menerbitkan saham.
3. Remarks, yang dapat digunakan untuk memberikan catatan atau keterangan tambahan tentang saham tersebut.
4. *Open Price*, yang merupakan harga pembuka pada hari saham sebuah perusahaan mulai diperdagangkan setiap hari.
5. *First Trade*, Ini adalah harga pertama yang diterima dalam transaksi saham pada hari itu.
6. Tertinggi, yang merupakan harga tertinggi saham pada perdagangan saham dalam satu hari.
7. Terendah, yang merupakan harga terendah saham pada perdagangan saham dalam satu hari.
8. Penutupan, yang merupakan harga penutupan saham pada hari perdagangan saham dalam suatu hari. Pada penelitian ini menggunakan atribut Penutupan karena merupakan harga kesepakatan terakhir pada hari saham diperdagangkan dan biasa diramal dalam peramalan saham (Lu *et al.*, 2021).
9. Selisih, yang merupakan perbedaan antara harga penutupan dan harga sebelumnya.
10. Volume, yang merupakan Ini adalah jumlah saham yang diperdagangkan pada hari itu.
11. Nilai, yang merupakan nilai total dari transaksi saham pada hari itu.
12. Frekuensi, yang merupakan jumlah transaksi saham pada hari itu.
13. Index Individual, yang merupakan indeks saham tertentu yang saham tersebut terdaftar.
14. *Listed Shares*, yang merupakan jumlah saham yang terdaftar pada bursa saham.
15. *Offer*, yang merupakan harga tawaran saham oleh pembeli.
16. *Offer Volume*, yang merupakan jumlah saham yang diperdagangkan dalam periode satu hari pasar saham berlangsung.
17. *Bid*, yaitu harga beli saham pada hari perdagangan saat ini.
18. *Bid Volume*, yang merupakan jumlah saham yang tersedia untuk dibeli pada hari perdagangan saat ini.
19. *Last Trading Date*, yang merupakan tanggal terakhir saham tersebut diperdagangkan.
20. *Tradeable Shares*, yaitu jumlah saham yang dapat diperdagangkan oleh investor.
21. *Weight for Index*, yaitu bobot saham tersebut dalam indeks, jika ada.
22. *Foreign Sell*, yaitu jumlah saham yang dijual oleh investor asing.
23. *Foreign Buy*, yaitu jumlah saham yang dibeli oleh investor asing.

Untuk *Non Regular Volume*, *Non Regular Value*, dan *Non Regular Frequency* adalah kategori data yang dapat digunakan untuk mengukur aktivitas perdagangan saham yang tidak teratur.

1. *Non Regular Volume* mengacu pada jumlah saham yang diperdagangkan pada transaksi tidak teratur.
2. *Non Regular Value* mengacu pada nilai total dari saham yang diperdagangkan pada transaksi tidak teratur.
3. *Non Regular Frequency* mengacu pada frekuensi transaksi saham yang tidak teratur.

Data ini dapat digunakan untuk memahami bagaimana saham tersebut diperdagangkan dalam konteks pasar dan dapat membantu dalam menganalisis aktivitas perdagangan dan pergerakan harga saham. Namun, karena data ini mengacu pada transaksi tidak teratur, hasil analisis yang didapat mungkin tidak representatif dari aktivitas perdagangan saham secara keseluruhan.

Berikut merupakan data saham BRI dan Bank IBK Indonesia tanpa beberapa kolom yang sekiranya tidak akan digunakan dalam proses peramalan di jurnal penelitian yang sudah ada (Agusta, 2021; Akbardipura, 2021; Maruddani and Astuti, 2021; Zhang and Lou, 2021) karena tidak dapat masuk kertas jurnal ini. Data yang digunakan pada penelitian ini periode September 2019 hingga Desember 2022 pada Tabel 4.1-4.2.

Tabel 4.1. Data Saham Bank Rakyat Indonesia (Persero) Tbk.

1. Kolom Kode Saham-*First Trade*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Kode Saham | Nama Perusahaan | Remarks | Sebelumnya | Open Price | Tanggal Perdagangan Terakhir | First Trade |
| BBRI | Bank Rakyat Indonesia (Persero) Tbk. | #NAME? | 3610 | 3580 | 1/3/2019 | 3590 |
| BBRI | Bank Rakyat Indonesia (Persero) Tbk. | #NAME? | 3610 | 3580 | 1/3/2019 | 3590 |
| BBRI | Bank Rakyat Indonesia (Persero) Tbk. | #NAME? | 3620 | 3620 | 1/4/2019 | 3610 |
| BBRI | Bank Rakyat Indonesia (Persero) Tbk. | #NAME? | 3660 | 3660 | 1/8/2019 | 3660 |
| BBRI | Bank Rakyat Indonesia (Persero) Tbk. | #NAME? | 3660 | 3660 | 1/8/2019 | 3660 |
| … | … | … | … | … | … | … |
| BBRI | Bank Rakyat Indonesia (Persero) Tbk. | --MO1UQSPBH000G111------------ | 4960 | 4920 | 12/23/2022 | 4940 |
| BBRI | Bank Rakyat Indonesia (Persero) Tbk. | --MO1UQSPBH000G111------------ | 4880 | 4900 | 12/26/2022 | 4900 |
| BBRI | Bank Rakyat Indonesia (Persero) Tbk. | --MO1UQSPBH000G111------------ | 4930 | 4930 | 12/27/2022 | 4930 |
| BBRI | Bank Rakyat Indonesia (Persero) Tbk. | --MO1UQSPBH000G111------------ | 4870 | 4870 | 12/28/2022 | 4880 |
| BBRI | Bank Rakyat Indonesia (Persero) Tbk. | --MO1UQSPBH000G111------------ | 4820 | 4820 | 12/29/2022 | 4830 |

1. Kolom Tertinggi-*Bid*

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Tertinggi | Terendah | Penutupan | Selisih | Volume | Nilai | Frekuensi | Index Individual | Offer | Offer Volume | Bid |
| 3640 | 3580 | 3620 | 10 | 93458300 | 3.37E+11 | 7797 | 4137.1 | 3620 | 265100 | 3610 |
| 3640 | 3580 | 3620 | 10 | 93458300 | 3.37E+11 | 7797 | 4137.1 | 3620 | 265100 | 3610 |
| 3660 | 3610 | 3660 | 40 | 114459600 | 4.17E+11 | 6677 | 4182.9 | 3660 | 2879500 | 3640 |
| 3680 | 3630 | 3680 | 20 | 81361900 | 2.98E+11 | 6863 | 4205.7 | 3680 | 11653700 | 3670 |
| 3680 | 3630 | 3680 | 20 | 81361900 | 2.98E+11 | 6863 | 4205.7 | 3680 | 11653700 | 3670 |
| … | … | … | … | … | … | … | … | … | … | … |
| 4940 | 4880 | 4880 | -80 | 53350600 | 2.61E+11 | 8132 | 5723.5 | 4890 | 79100 | 4880 |
| 4930 | 4890 | 4930 | 50 | 43590300 | 2.14E+11 | 5802 | 5782.2 | 4930 | 2968200 | 4920 |
| 4950 | 4860 | 4870 | -60 | 81817500 | 4.00E+11 | 13744 | 5711.8 | 4880 | 107800 | 4870 |
| 4880 | 4810 | 4820 | -50 | 134315900 | 6.48E+11 | 17964 | 5653.2 | 4830 | 118200 | 4820 |
| 4870 | 4800 | 4870 | 50 | 106690700 | 5.17E+11 | 7736 | 5711.8 | 4870 | 1044300 | 4860 |

1. Kolom *Bid Volume*-*Non Regular Frequency*

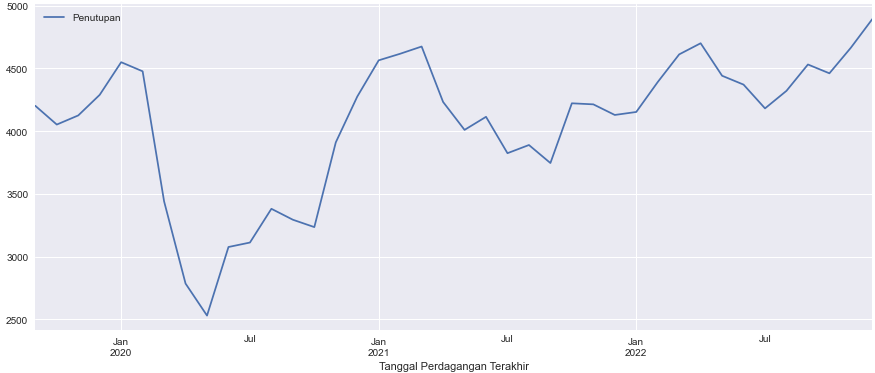
|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Bid Volume | Listed Shares | Tradeble Shares | Weight for Index | Foreign Sell | Foreign Buy | Non Regular Volume | Non Regular Value | Non Regular Frequency |
| 4000 | 1.22E+11 | 1.22E+11 |  | 69049100 | 68444000 | 33439022 | 1.21E+11 | 15 |
| 4000 | 1.22E+11 | 1.22E+11 |  | 69049100 | 68444000 | 33439022 | 1.21E+11 | 15 |
| 25600 | 1.22E+11 | 1.22E+11 |  | 80301300 | 94353800 | 2849973 | 10346055912 | 8 |
| 2009100 | 1.22E+11 | 1.22E+11 |  | 51060000 | 52577300 | 38938096 | 1.41E+11 | 17 |
| 2009100 | 1.22E+11 | 1.22E+11 |  | 51060000 | 52577300 | 38938096 | 1.41E+11 | 17 |
| … | … | … | … | … | … | … | … | … |
| 1010000 | 1.50E+11 | 1.50E+11 |  | 40557100 | 20421400 | 1191432 | 5578631387 | 12 |
| 10000 | 1.50E+11 | 1.50E+11 |  | 5914300 | 11799100 | 9151481 | 45753819940 | 30 |
| 130300 | 1.50E+11 | 1.50E+11 |  | 41696800 | 16977100 | 1999083 | 10665668105 | 10 |
| 945900 | 1.50E+11 | 1.50E+11 |  | 86326500 | 60839000 | 1535169 | 8648238140 | 23 |
| 52500 | 1.50E+11 | 1.50E+11 |  | 61337400 | 44636400 | 2706521 | 12777650560 | 14 |

Dari Tabel 4.1 di atas dapat dilihat pada data saham Bank Rakyat Indonesia (Persero) Tbk. untuk harga penutupannya mengalami kenaikan dari yang harga mulanya Rp 3620 menjadi Rp 4870 pada akhir Desember 2022. Hal ini menunjukan adanya keuntungan bagi para investor BRI dalam kurun waktu 2 tahun.

Untuk kenaikannya sendiri dapat dihitung presentase keuntungan yang didapat dari saham BRI. Dalam kurun waktu 2 tahun, saham BRI naik sebesar 34,53%. Keuntungan ini dapat dinilai sebagai keuntungan yang cukup besar jika melihat 2 tahun tersebut terdapat pandemi yang mempengaruhi harga pasar secara global (Trierweiler Ribeiro *et al.*, 2021). Presentase keuntungan yang didapat bisa dihitung dengan cara yang paling mudah dengan:

(34)

Untuk pertumbuhan data penutupan saham dari September 2019 hingga akhir Desember 2022 dapat dilihat pada Gambar 4.1.



Gambar 4.1. Grafik Pertumbuhan Saham Bank Rakyat Indonesia (Persero) Tbk.

Tabel 4.2. Data Saham Bank IBK Indonesia Tbk.

1. Kolom Kode Saham-*First Trade*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Kode Saham | Nama Perusahaan | Remarks | Sebelumnya | Open Price | Tanggal Perdagangan Terakhir | First Trade |
| AGRS | Bank IBK Indonesia Tbk. | 589896798 | 200 | 0 | 9/10/2019 | 0 |
| AGRS | Bank IBK Indonesia Tbk. | 589896798 | 202 | 0 | 9/12/2019 | 0 |
| AGRS | Bank IBK Indonesia Tbk. | 589896798 | 200 | 0 | 9/13/2019 | 0 |
| AGRS | Bank IBK Indonesia Tbk. | 589896798 | 200 | 0 | 9/13/2019 | 0 |
| AGRS | Bank IBK Indonesia Tbk. | 589896798 | 197 | 0 | 9/17/2019 | 0 |
| … | … | … | … | … | … | … |
| AGRS | Bank IBK Indonesia Tbk. | --U-2105000000G111------------ | 89 | 0 | 12/23/2022 | 0 |
| AGRS | Bank IBK Indonesia Tbk. | --U-2105000000G111------------ | 88 | 0 | 12/26/2022 | 0 |
| AGRS | Bank IBK Indonesia Tbk. | --U-2105000000G111------------ | 87 | 0 | 12/27/2022 | 0 |
| AGRS | Bank IBK Indonesia Tbk. | --U-2105000000G111------------ | 86 | 0 | 12/28/2022 | 0 |
| AGRS | Bank IBK Indonesia Tbk. | --U-2105000000G111------------ | 87 | 0 | 12/29/2022 | 0 |

1. Kolom Tertinggi-*Bid*

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Tertinggi | Terendah | Penutupan | Selisih | Volume | Nilai | Frekuensi | Index Individual | Offer | Offer Volume | Bid |
| 200 | 198 | 200 | 0 | 39000 | 7747400 | 23 | 181.8 | 200 | 74300 | 198 |
| 204 | 198 | 200 | -2 | 150300 | 29990100 | 36 | 181.8 | 200 | 6000 | 199 |
| 198 | 196 | 198 | -2 | 52000 | 10241600 | 25 | 180 | 198 | 4000 | 196 |
| 198 | 196 | 198 | -2 | 52000 | 10241600 | 25 | 180 | 198 | 4000 | 196 |
| 197 | 193 | 196 | -1 | 69500 | 13550900 | 26 | 178.2 | 197 | 20900 | 193 |
| … | … | … | … | … | … | … | … | … | … | … |
| 90 | 88 | 88 | -1 | 3828100 | 337994900 | 305 | 113.9 | 89 | 800900 | 88 |
| 89 | 83 | 87 | -1 | 2665500 | 233654100 | 262 | 112.6 | 88 | 592000 | 87 |
| 88 | 84 | 86 | -1 | 7347700 | 630599700 | 427 | 111.3 | 86 | 401700 | 85 |
| 90 | 86 | 87 | 1 | 7227900 | 639124700 | 405 | 112.6 | 88 | 50700 | 87 |
| 91 | 87 | 90 | 3 | 3206000 | 286193200 | 316 | 116.5 | 90 | 231600 | 89 |

1. Kolom *Bid Volume*-*Non Regular Frequency*

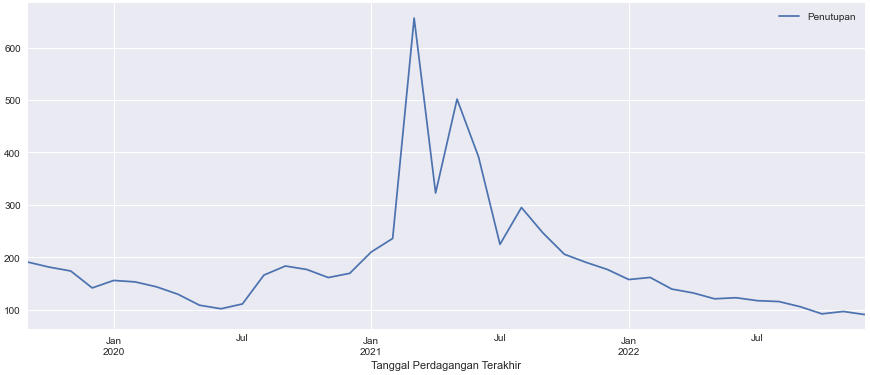
|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Bid Volume | Listed Shares | Tradeble Shares | Weight for Index | Foreign Sell | Foreign Buy | Non Regular Volume | Non Regular Value | Non Regular Frequency |
| 19100 | 7037266148 | 7037266148 |  | 0 | 0 | 0 | 0 | 0 |
| 31300 | 7037266148 | 7037266148 |  | 0 | 0 | 0 | 0 | 0 |
| 4900 | 7037266148 | 7037266148 |  | 0 | 0 | 0 | 0 | 0 |
| 4900 | 7037266148 | 7037266148 |  | 0 | 0 | 0 | 0 | 0 |
| 35200 | 7037266148 | 7037266148 |  | 0 | 0 | 0 | 0 | 0 |
| … | … | … | … | … | … | … | … | … |
| 550800 | 27357208711 | 27357208711 |  | 100 | 10000 | 0 | 0 | 0 |
| 989200 | 27357208711 | 27357208711 |  | 100 | 0 | 0 | 0 | 0 |
| 907700 | 27357208711 | 27357208711 |  | 0 | 214800 | 0 | 0 | 0 |
| 181400 | 27357208711 | 27357208711 |  | 0 | 0 | 0 | 0 | 0 |
| 1241100 | 27357208711 | 27357208711 |  | 279100 | 0 | 0 | 0 | 0 |

Dari Tabel 4.2 di atas, penelitian ini mendapatkan informasi bahwa dalam kurun waktu 2 tahun harga penutupan untuk Bank IBK Indonesia Tbk. mengalami penurunan yang signifikan dari awalnya Rp 200 menjadi Rp 90 saja. Penurunan ini menyebabkan kerugian bagi pemegang investasi saham perusahaan ini.

Untuk kerugiannya sendiri dapat dihitung dengan cara yang hampir sama dengan metode perhitungan presentase keuntungan BRI. Presentase kerugiannya itu sendiri adalah 55%, angka yang cukup besar untuk kerugian investasi saham. Perhitungannya itu sendiri dapat dengan cara sebagai berikut:

(35)

Untuk penurunan data penutupan saham dari September 2019 sampai dengan Desember 2022 dapat dilihat pada Gambar 4.2.



Gambar 4.2. Grafik Pertumbuhan Saham Bank IBK Indonesia Tbk.

Pada Gambar 4.2 dapat terlihat pada September 2019 pada tahun hingga July di tahun 2020 dan 2022 memiliki tingkat penurunan yang sama. Pada awal tahun hingga pertengahan tahun 2021, harga penutupan saham Bank IBK Indonesia sempat mencapai angka tertingginya dibandingkan tahun sebelum dan sesudahnya. Setelah mencapai harga penutupan tertingginya, harganya mulai turun kembali seperti dilihat dalam grafik tahun sebelumnya dan hanya saja lebih rendah.

## 4.2. Praproses Data

Setelah data didapatkan, dilakukan praproses data dengan beberapa langkah-langkah. Untuk pengerjaan praproses data ini berupa rentetan proses yang berurutan berupa antara lain:

### 4.2.1. *Feature Selection*

Dalam *feature selection* ini, penelitian ini perlu memproses atribut yang akan digunakan dalam pembuatan model. Dalam hal ini menggunakan variabel penutupan sebagaimana pada penelitian (Fitriadini, Pramiyati and Pangaribuan, 2020; Kartikasari, 2020; Agusta, 2021) sebelumnya. Hal ini bertujuan untuk mempersempit unsur saham apa yang akan diramal. Selain itu, proses ini juga bertujuan untuk melihat *plot* pertumbuhan saham dan hasil ramalan model menjadi lebih mudah.

Untuk atribut yang digunakan pada penelitian ini adalah Penutupan. Penelitian ini menggunakan atribut ini karena penutupan merupakan harga kesepakatan terakhir pada hari perdagangan saham. Dalam satu hari dapat disimpulkan bahwa harga penutupan juga merupakan harga acuan pada hari sebelumnya untuk diperdagangkan di hari berikutnya (Lu *et al.*, 2021).

Tabel 5 teratas untuk data penutupan saham BRI dan Bank IBK Indonesia karena atribut yang akan dipakai adalah penutupan dapat dilihat pada tabel 4.1 dan 4.2.

Tabel 4.1. Data Saham BRI

|  |  |
| --- | --- |
| **Tanggal Perdagangan Terakhir** | **Penutupan** |
| 9/10/2019 | 4150 |
| 9/12/2019 | 4150 |
| 9/13/2019 | 4160 |
| 9/13/2019 | 4270 |
| 9/17/2019 | 4270 |

Tabel 4.2. Data Saham Bank IBK Indonesia

|  |  |
| --- | --- |
| **Tanggal Perdagangan Terakhir** | **Penutupan** |
| 9/10/2019 | 191.067 |
| 9/12/2019 | 181.435 |
| 9/13/2019 | 174.095 |
| 9/13/2019 | 141.789 |
| 9/17/2019 | 156.045 |

### 4.2.2. Normalisasi Data

Sebelum data dibagi, data yang digunakan harus dilakukan perubahan terlebih dahulu. Perubahan ini adalah berupa mengubah data yang berdasarkan satu hari menjadi satu bulan. Data diubah menjadi data penutup perbulan dengan isian rata-rata karena yang akan diobservasi adalah laju pendapatan bukan jumlah pendapatan (Saleh, Grivel and Omar, 2018). Berikut merupakan 5 data pertama yang telah diubah ke dalam bentuk perbulan terlihat pada Tabel 4.3-4.4.

Tabel 4.3. Data Bank Rakyat Indonesia (Persero) Tbk. Perbulan.

|  |  |
| --- | --- |
| **Tanggal Perdagangan Terakhir** | **Penutupan** |
| 9/30/2019 | 4202.380952 |
| 10/31/2019 | 4053.043478 |
| 11/30/2019 | 4125.714286 |
| 12/31/2019 | 4290.000000 |
| 1/31/2020 | 4550.000000 |

Tabel 4.4. Data Bank IBK Indonesia Tbk. Perbulan

|  |  |
| --- | --- |
| **Tanggal Perdagangan Terakhir** | **Penutupan** |
| 9/30/2019 | 191.067 |
| 10/31/2019 | 181.435 |
| 11/30/2019 | 174.095 |
| 12/31/2019 | 141.789 |
| 1/31/2020 | 156.045 |

Untuk mendapatkan nilai data perbulan di atas dapat dilakukan dengan cara *resampling mean* sesuai dengan persamaan (16). Sebagai contoh perhitungannya dapat digunakan bulan September 2019.

Diketahui:

= 4150, 4150, 4160, 4270, 4270, 4170, 4190, 4280, 4310, 4310, 4250, 4240, 4240, 4220, 4160, 4180, 4110, 4080, 4210, 4180, 4120

= 21

= 88250

4202.380952

Setelah data diubah menjadi format perbulan, data memiliki jumlah total 40 data tiap dataset dan akan dilakukan normalisasi dengan *MinMax Scaler*. Normalisasi ini dilakukan dengan cara menghitung selisih data dengan data terendah dibagi selisih data tertinggi dengan data terendah (Agrawal, 2021). Berikut merupakan hasil normalisasi data dari kedua data bank dapat dilihat pada Tabel 4.5-4.6.

Tabel 4.5. Data Bank Rakyat Indonesia (Persero) Tbk. normalisasi

|  |  |
| --- | --- |
| **Tanggal Perdagangan Terakhir** | **Close** |
| 2019-09-30 | 0.707417 |
| 2019-10-31 | 0.644223 |
| 2019-11-30 | 0.674974 |
| 2019-12-31 | 0.744493 |
| 2020-01-31 | 0.854514 |
| 2020-02-29 | 0.823835 |
| 2020-03-31 | 0.385212 |
| 2020-04-30 | 0.107943 |
| 2020-05-31 | 0.000000 |
| 2020-06-30 | 0.231263 |
| 2020-07-31 | 0.246321 |
| 2020-08-31 | 0.360125 |

Tabel 4.6. Data Bank IBK Indonesia Tbk. normalisasi

|  |  |
| --- | --- |
| **Tanggal Perdagangan Terakhir** | **Close** |
| 9/30/2019 | 0.17704 |
| 10/31/2019 | 0.16 |
| 11/30/2019 | 0.14701 |
| 12/31/2019 | 0.08986 |
| 1/31/2020 | 0.11508 |
| 2/29/2020 | 0.11022 |
| 3/31/2020 | 0.09352 |
| 4/30/2020 | 0.06841 |
| 5/31/2020 | 0.03151 |
| 6/30/2020 | 0.01963 |
| 7/31/2020 | 0.03587 |
| 8/31/2020 | 0.13308 |

Data yang telah dinormalisasikan di atas dapat diperoleh dengan persamaan (28), berikut merupakan salah satu contoh perhitungannya yang diambil dari data September 2019 BRI.

Diketahui:

= 4202.380952

= 2530.625000

= 4893.809524

### 4.2.3. Pembagian Data

Dalam tahapan ini, penelitian akan membagi antara data latih dan data uji. Pembagian data ini dilakukan dua kali yang bertujuan untuk percobaan pengaruh penambahan data uji untuk dibuat model *forecasting*. Pembagian ini dilakukan dengan cara memilih data dalam rentang watku tertentu dari index berupa tanggal (Chen and Zhou, 2021). Untuk pembagian data pertama, data latih dari data penutupan saham dari September 2019 hingga April 2022 dan data ujinya dari Agustus 2022 hingga Desember 2022 berdasarkan 80% pebagian data latih dan 20% data uji. Untuk pembagian data kedua data latihnya dikurangin menjadi 60%, yaitu dari September 2019 hingga Agustus 2021 dan data ujinya 40%, yaitu September 2021 hingga Desember 2022. Untuk pembagian data ketiga untuk data latihnya adalah 20%, yaitu September 2019 hingga April 2020 dan data ujinya adalah 80%, yaitu dari Mei 2020 hingga Desember 2022.

### 4.2.4. Uji Stasioner dan Stasionerisasi

Tahapan ini hanya akan digunakan untuk metode ARFIMA saja. Hal ini disebabkan karena tahapan ini digunakan untuk menentukan parameter ARFIMA. Tahapan ini dilakukan dengan melakukan *Augmented Dickey Fuller Test* dan mengecek apakah nilai *p-value* sudah lebih kecil dari 0.05. Apabila tidak, maka perlu dilakukan *differencing* pada data latihnya. Berikut merupakan tabel *p-value* kedua data bank dapat dilihat pada tabel 4.3-4.5.

Tabel 4.3. *P-value* metode pembagian data pertama.

|  |  |  |
| --- | --- | --- |
| **Diffensiasi ke-N** | **BRI** | **Bank IBK Indonesia** |
| 0 | 0.5710386815654017 | 0.4894135557466029 |
| 1 | 0.023722727478323914 | 0.001439254938209262 |

Tabel 4.4. *P-value* metode pembagian data kedua.

|  |  |  |
| --- | --- | --- |
| **Diffensiasi ke-N** | **BRI** | **Bank IBK Indonesia** |
| 0 | 0.006325424864391016 | 0.2963657640402307 |
| 1 | Sudah stasioner | 0.9935911006407631 |
| 2 | Sudah stasioner | 0.04419979997246791 |

Tabel 4.5. *P-value* metode pembagian data ketiga

|  |  |  |
| --- | --- | --- |
| **Diffensiasi ke-N** | **BRI** | **Bank IBK Indonesia** |
| 0 | 0.21881147863093942 | 0.734207518448737 |
| 1 | 0.15511181849993977 | 0.037153189746917115 |
| 2 | 0.31111546600090656 | Sudah stasioner |

Dapat dilihat pada Tabel 4.5 menunjukan bahwa nilai *p-value* dari pembagian data ketiga BRI pada *differencing* keduanya naik. Hal ini sangat umum karena tidak semua data *real* adalah stasioner, maka digunakan ordo *differencing* yang mendekati stasioner Perlu diketahui bahwa untuk mendapatkan nilai *p-value* perlu mendapatkan nilai *differencing* per baris data. Untuk mendapatkan hasil dari *differencing* dapat digunakan persamaan (17), sebagai contohnya.

Diketahui:

= 0.707417

= 0.64422328

Setelah tahapan *differencing*, maka digunakan fungsi ADF *test* untuk memastikan data sudah stasioner. Hasil ADF *test* yang perlu diperhatikan adalah nilai *p-value* dari hasil uji ADF-nya sebagaimana Tabel 4.3-4.5.

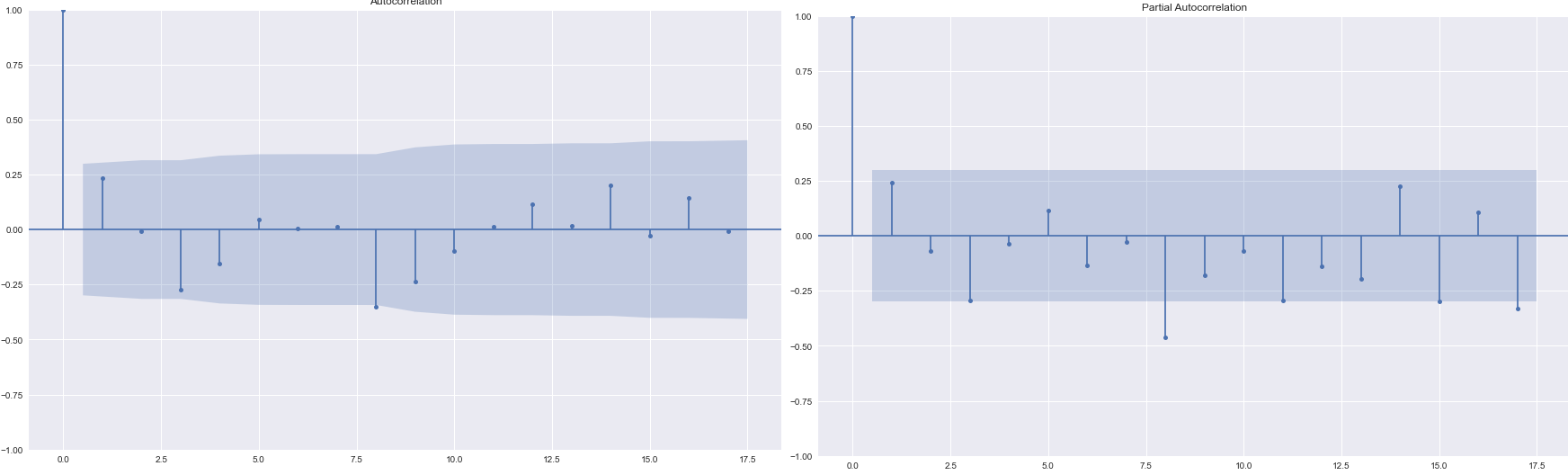
## 4.3. Penentuan Parameter

Pada tahapan ini untuk setiap metode berbeda cara menentukan parameternya. Untuk tahapan penentuan parameter dapat menjadi antara lain:

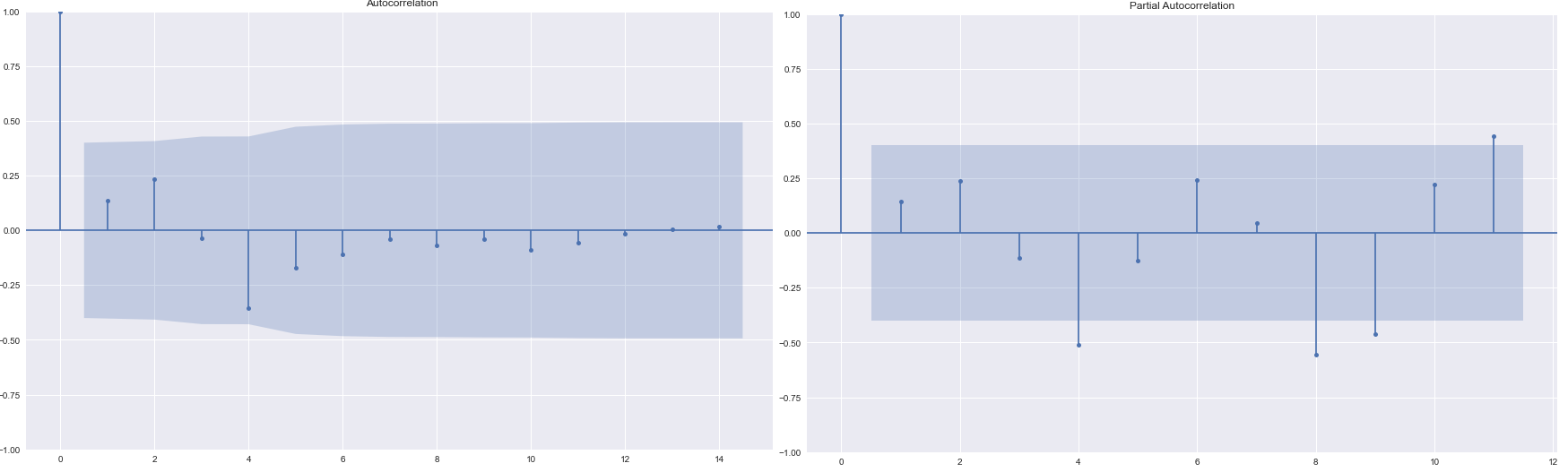
### 4.3.1 Parameter ARFIMA

Setelah dilakukan uji stasioneritas, dapat dilihat bahwa kedua data tidak ada yang stasioner dan perlu dilakukan proses *differencing*. Pada tahapan stasioneritas inilah penelitian ini mendapatkan parameter *d* untuk parameter ARFIMA-nya. Nilai parameter *d* dapat ditentukan dengan menghitung berapa kali data di-*differencing* hingga menjadi stasioner.

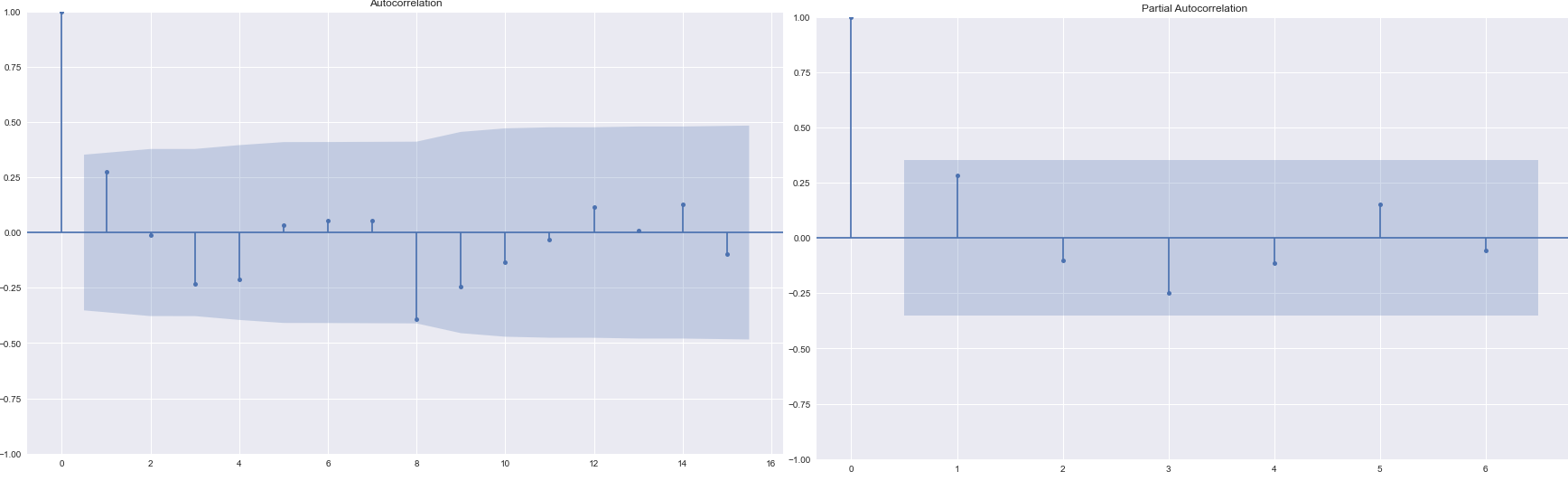
Setelah itu, melakukan *plotting* PACF dan ACF untuk melihat *lag* pada data latih. Berikut merupakan *plot* PACF dan ACF Kedua data latih Bank dapat dilihat pada gambar 4.1-4.6.



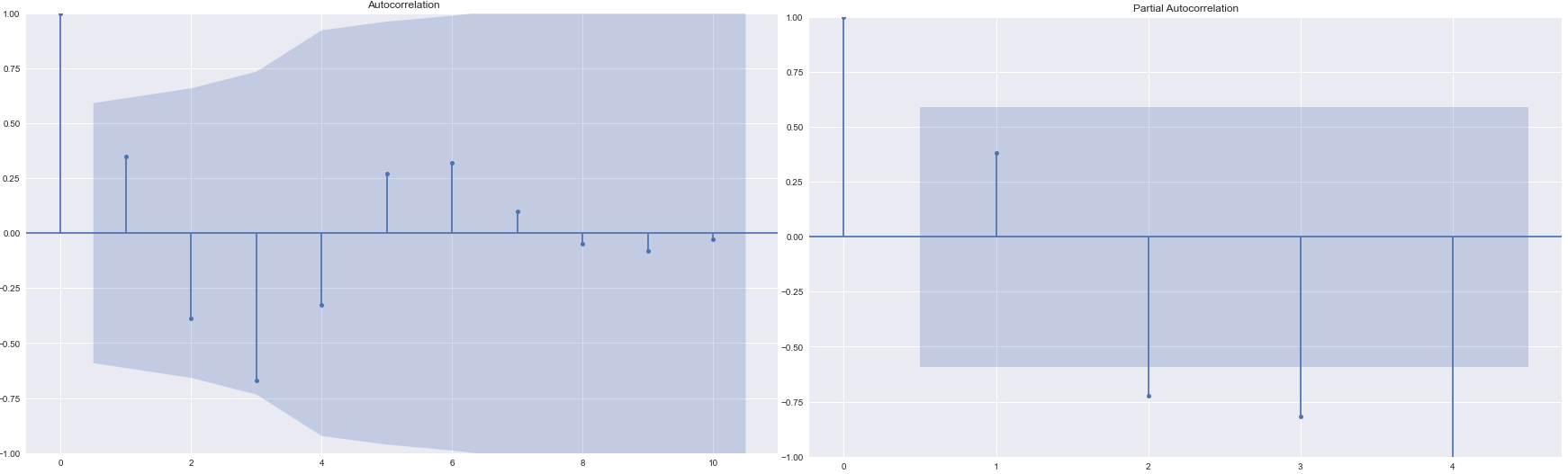
Gambar 4.1. ACF dan PACF BRI pembagian pertama



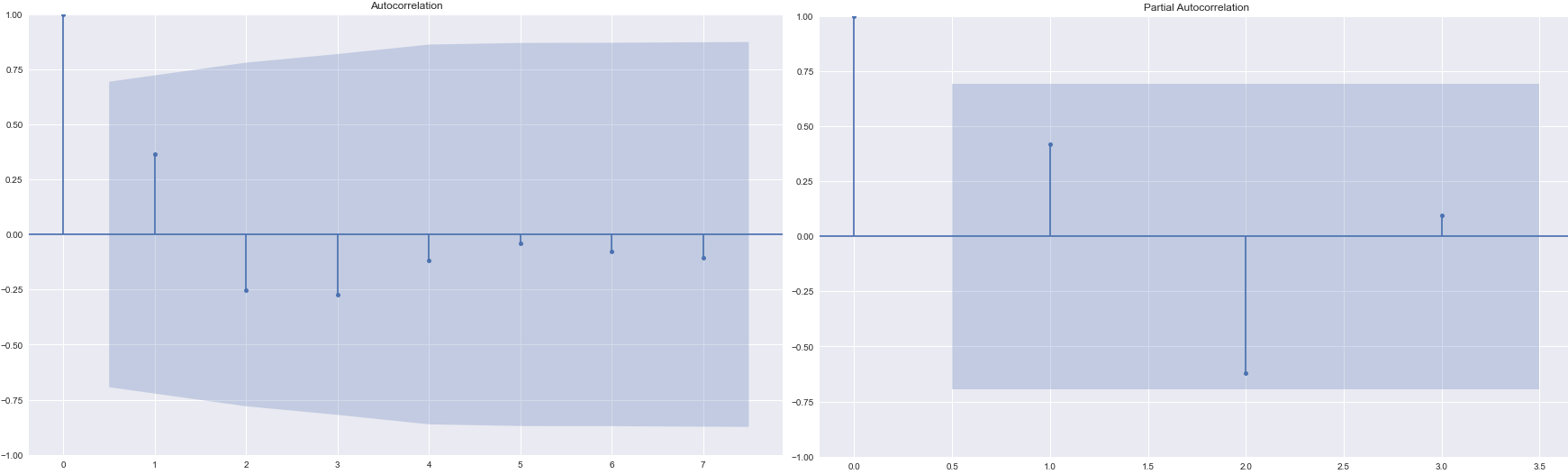
Gambar 4.2. ACF dan PACF Bank IBK Indonesia pembagian pertama



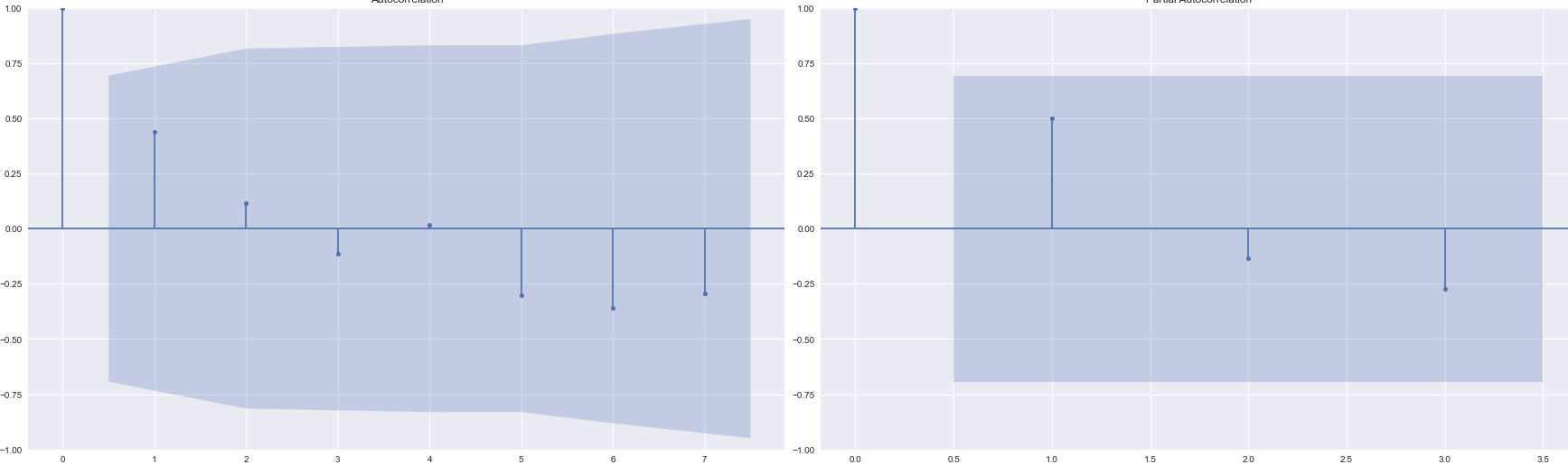
Gambar 4.3. ACF dan PACF BRI pembagian kedua



Gambar 4.4. ACF dan PACF Bank IBK pembagian kedua



Gambar 4.5. ACF dan PACF BRI pembagian ketiga



Gambar 4.4. ACF dan PACF Bank IBK pembagian ketiga

Pada Gambar 4.1 dapat dilihat bahwa tidak terdapat *lag* yang signifikan keluar dari area biru pada *plot* ACF-nya. Hal ini mengindikasikan bahwa parameter *q* yang dapat digunakan adalah 0. Sedangkan untuk *plot* PACF-nya terdapat *lag* yang signifikan di angka 3. Dari sini didapatkan parameter *p* adalah 3.

Setelah itu pada Gambar 4.2 juga tidak terlihat *lag* yang signifikan pada *plot* ACF sehingga parameter *q* adalah 0. Namun pada *plot* PACF terlihat *lag* yang signifikan pada *lag* ke 4. Untuk data ini menggunakan parameter *p*-nya adalah 4.

Gambar 4.3 tidak terlihat *lag* yang melambung secara signifikan baik pada *plot* ACF dan juga PACF. Hal ini mengindikasikan bahwa parameter *p* dan parameter *q* adalah 0.

Kemudian untuk Gambar 4.4 tidak memperlihatkan *lag* yang melambung secara signifikan pada *plot* ACF sehingga parameter *q* yang digunakan adalah 0. Dari *plot* PACF dapat dilihat bahwa terdapat *lag* yang signifikan pada *lag* ke 2. Dari sini didapatkan bahwa nilai parameter *p* yang digunakan adalah 2.

Pada *plot* terakhir Gambar 4.5-4.6 tidak memperlihatkan *lag* yang melambung secara signifikan baik pada *plot* ACF maupun PACF. Dari hal ini dapat digunakan parameter *p* dan *q* untuk data BRI dan Bank IBK adalah 0.

Untuk mendapatkan *plot* di atas digunakan persamaan (20). Sebagai contoh penggunaannya dapat digunakan data latih BRI yang telah dilakukan *differencing* dan nilai rata-ratanya 0.006807 pada *lag* ke 1.

Diketahui:

= -0.063193

= 0.030751

= 0.006807

= 31

= 1

Dari hasil tersebut dapat diketahui nilai ACF pada *lag* pertama. Kemudian dapat digunakan untuk mencari nilai PACF-nya. Untuk mencari nilai PACF dapat digunakan persamaan (25), sebagai contohnya.

Diketahui:

= 0.3373

= = 0.3373

Maka,

Kemudian untuk mendapatkan *convidence intervalnya* (area biru pada *plot*) dapat menggunakan persamaan (21), sebagai contoh.

Diketahui:

= 0.6200317426051204

=

= -0.4949642926300218

= 0.006807

Maka,

### 4.3.2. Parameter LSTM

Untuk parameter yang digunakan pada metode LSTM adalah menggunakan parameter yang direkomendasikan oleh buku (Ikeuchi, 2021) sebagai parameter yang biasa digunakan, yaitu 64 unit LSTM, 2 x 32 *Dense Layer*, dan 64 *batch size* karena data yang digunakan tidak terlalu besar hanya 40 data dari 40 bulan setelah data dilakukan *resampling*.

Pada metode LSTM, penelitian ini akan membagi model dari tiap pembagian data berapa *epoch* yang digunakan dan pada seberapa besar pengaruh *epoch* atau iterasi pelatihan model. Untuk penggunaan *epoch* terkecil adalah 50 karena menyesuaikan ukuran data dan mempermudah dalam pengamatan pengaruh penambahan *epoch*. Pembagian parameter LSTM dapat dilihat sesuai Tabel 4.6.

Tabel 4.6. Parameter LSTM

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *Parameter* | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 |
| *unit* | 64 | 64 | 64 | 64 | 64 |
| *dense* | 32 | 32 | 32 | 32 | 32 |
| *dense* | 32 | 32 | 32 | 32 | 32 |
| *dense* | 1 | 1 | 1 | 1 | 1 |
| *optimizer* | adam | adam | adam | adam | adam |
| *learning rate* | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 |
| *loss* | *mse* | *mse* | *mse* | *mse* | *mse* |
| *epoch* | 50 | 100 | 150 | 200 | 300 |
| *batch size* | 64 | 64 | 64 | 64 | 64 |

## 4.4. *Training Model*

Model yang dibuat adalah model menggunakan dua metode yaitu metode ARFIMA dan LSTM. Pemodelan dibuat dengan antara lain:

### 4.4.1. Model Pembagian Data Latih 80%

#### 4.4.1.1. BRI ARFIMA(8,0.5,0)

Setelah menyelesaikan semua prosedur sebelum pemodelan ARFIMA yang diperlukan, kita akan menggunakan banyak model ARFIMA dengan berbagai urutan berdasarkan penentuan parameter dari *plot* ACF dan PACF. Penelitian akan memeriksa *output* dari model ARFIMA, yang akan diuji dan konsisten dengan dataset yang digunakan, terlebih dahulu dalam prosedur pemodelan ini.

Metode ARFIMA perlu melakukan *fractional differencing* pada data latih yang digunakan. Untuk model ini menggunakan parameter *p* adalah 3 dan *q* dengan nilai 0 sesuai pada proses penentuan ACF dan PACF. Berikut merupakan contoh perhitungan Model ARFIMA(8,0.5,0).

Diketahui:

= 0.707417

= = 0.644223

= 0.5

Maka,

Setelah didapatkan nilai *fractional difference*-nya, data dapat dilakukan *training* menggunakan ARIMA(8,1,0) menggunakan persamaan (15).

= 8

= 0

Dalam persamaan di atas didapatkan hasil data prediksi dari model ARFIMA(8,0.5,0). Untuk hasil 15 teratas dan 4 terbawah dari prediksi dengan model tersebut dapat dilihat pada Tabel 4.7.

Tabel 4.7. Data prediksi BRI ARFIMA(8,0.5,0)

|  |  |
| --- | --- |
| Tanggal Perdagangan Terakhir | Prediksi Latih |
| 9/30/2019 | 0.000000 |
| 10/31/2019 | 0.841081 |
| 11/30/2019 | 0.798581 |
| 12/31/2019 | 0.823946 |
| 1/31/2020 | 0.879030 |
| 2/29/2020 | 0.927038 |
| 3/31/2020 | 0.890921 |
| 4/30/2020 | 0.569654 |
| 5/31/2020 | 0.288299 |
| 6/30/2020 | 0.080833 |
| 7/31/2020 | 0.590430 |
| 8/31/2020 | 0.622407 |
| 9/30/2020 | 0.509263 |
| 10/31/2020 | 0.573280 |
| 11/30/2020 | 0.741604 |
| … | … |
| 1/31/2022 | 0.812551 |
| 2/28/2022 | 0.809661 |
| 3/31/2022 | 0.932172 |
| 4/30/2022 | 0.925235 |

Dapat dilihat pada Tabel 4.7, nilai prediksi pada bulan Juli 2020 mengalami penurunan yang signifikan. Hal ini karena hasil latih tersebut merupakan hasil latih dari data yang telah dilakukan *fractional differentiation* yang membuat data *shifting* dari data sebelumnya. Dapat dilihat pada Tabel 4.5, bahwa pada bulan sebelumnya, yaitu bulan Juni 2020 memang terjadi penurunan yang signifikan dan membuat data hasil prediksi pada ­ yang di kasus ini adalah bulan setelah Juni 2020, yaitu bulan Juli 2020 terjadi penurunan yang signifikan sebagaimana pada Juni 2020 pada data asli.

Setelah didapatkan hasil modelnya, dilakukan denormalisasi untuk mengembalikan nilainya. Denormalisasi dapat menggunakan persamaan (29), sebagai contoh.

= 0

= 2530.625000

= 4893.809524

Dari perhitungan di atas didapatkan data prediksi yang telah dikembalikan nilainya. Hasil perhitungannya dapat dilihat pada Tabel 4.8.

Tabel 4.8. Data Prediksi Denormalisasi BRI ARFIMA(8,0.5,0)

|  |  |
| --- | --- |
| Tanggal Perdagangan Terakhir | Prediksi Latih |
| 9/30/2019 | 2530.63 |
| 10/31/2019 | 4202.38 |
| 11/30/2019 | 4037.70 |
| 12/31/2019 | 4134.96 |
| 1/31/2020 | 4356.64 |
| 2/29/2020 | 4561.55 |
| 3/31/2020 | 4406.38 |
| 4/30/2020 | 3297.49 |
| 5/31/2020 | 2727.04 |
| 6/30/2020 | 2546.07 |
| 7/31/2020 | 3354.45 |
| 8/31/2020 | 3446.10 |
| 9/30/2020 | 3143.51 |
| 10/31/2020 | 3307.29 |
| 11/30/2020 | 3830.32 |
| … | … |
| 1/31/2022 | 4090.89 |
| 2/28/2022 | 4079.81 |
| 3/31/2022 | 4584.10 |
| 4/30/2022 | 4553.65 |

Kemudian untuk hasil prediksi data latih dapat dilihat sesuai Tabel 4.9 berikut.

Tabel 4.9. Data Ramal Denormalisasi BRI ARFIMA(8,0.5,0)

|  |  |
| --- | --- |
| Tanggal Perdagangan Terakhir | Prediksi Uji |
| 5/31/2022 | 4556.446392 |
| 6/30/2022 | 4313.166302 |
| 7/31/2022 | 4279.967827 |
| 8/31/2022 | 4297.277255 |
| 9/30/2022 | 4310.176050 |
| 10/31/2022 | 4225.886777 |
| 11/30/2022 | 4153.649852 |
| 12/31/2022 | 4179.530388 |

#### 4.4.1.2. Bank IBK ARFIMA(4,0.5,0)

Pada model ini menggunakan data Bank IBK Indonesia dengan parameter yang didapatkan berdasarkan *plot* pada Gambar 4.2, yaitu parameter *p* adalah 4 dan *q* adalah 0. Sebagaimana model BRI sebelumnya, data dilakukan *fractional* *differencing* dan langsung dilatih menggunakan ARIMA dengan parameter *d* sebagaimana pada Tabel 4.3 adalah 1 karena hanya perlu satu kali dilakukan *differencing* hingga dapat menjadi data yang stasioner. Untuk data yang telah dilatih dapat dilihat pada Tabel 4.10.

Tabel 4.10. Data prediksi Bank IBK ARFIMA(4,0.5,0)

|  |  |
| --- | --- |
| Tanggal Perdagangan Terakhir | Prediksi Latih |
| 9/30/2019 | 0.000000 |
| 10/31/2019 | 0.420759 |
| 11/30/2019 | 0.407593 |
| 12/31/2019 | 0.387747 |
| 1/31/2020 | 0.323565 |
| 2/29/2020 | 0.317461 |
| 3/31/2020 | 0.336837 |
| 4/30/2020 | 0.336932 |
| 5/31/2020 | 0.260811 |
| 6/30/2020 | 0.196009 |
| 7/31/2020 | 0.140908 |
| 8/31/2020 | 0.171414 |
| 9/30/2020 | 0.335356 |
| 10/31/2020 | 0.433485 |
| 11/30/2020 | 0.404707 |
| … | … |
| 1/31/2022 | 0.406427 |
| 2/28/2022 | 0.368810 |
| 3/31/2022 | 0.348363 |
| 4/30/2022 | 0.316836 |

Dari data dalam rentang 0 sampai 1 dapat dijadikan data kembali sebagaimana data aslinya dapat dilihat pada Tabel 4.1.

Tabel 4.11. Data prediksi denormalisasi Bank IBK ARFIMA(4,0.5,0)

|  |  |
| --- | --- |
| Tanggal Perdagangan Terakhir | Prediksi |
| 9/30/2019 | 91.000000 |
| 10/31/2019 | 191.066665 |
| 11/30/2019 | 184.902513 |
| 12/31/2019 | 175.980434 |
| 1/31/2020 | 150.176169 |
| 2/29/2020 | 147.964386 |
| 3/31/2020 | 155.130153 |
| 4/30/2020 | 155.166533 |
| 5/31/2020 | 129.448122 |
| 6/30/2020 | 112.715690 |
| 7/31/2020 | 102.222688 |
| 8/31/2020 | 107.607855 |
| 9/30/2020 | 154.567601 |
| 10/31/2020 | 197.211566 |
| 11/30/2020 | 183.577141 |
| … | … |
| 1/31/2022 | 184.366006 |
| 2/28/2022 | 167.882687 |
| 3/31/2022 | 159.594088 |
| 4/30/2022 | 147.740401 |

Kemudian untuk hasil prediksi data ujinya dapat dilihat sesuai Tabel 4.12 berikut.

Tabel 4.12. Data ramal denormalisasi Bank IBK ARFIMA(4,0.5,0)

|  |  |
| --- | --- |
| Tanggal Perdagangan Terakhir | Prediksi Uji |
| 5/31/2022 | 135.390 |
| 6/30/2022 | 130.837 |
| 7/31/2022 | 136.692 |
| 8/31/2022 | 136.205 |
| 9/30/2022 | 136.079 |
| 10/31/2022 | 137.765 |
| 11/30/2022 | 135.692 |
| 12/31/2022 | 136.701 |

#### 4.4.1.3. LSTM BRI

Dalam menggunakan model LSTM, proses latih dilakukan berkali-kali hingga model mendapatkan hasil yang optimal. Terdapat dua model yang akan dibuat menggunakan LSTM BRI dengan data latih 80% dan data uji 20% sebagaimana pembagian parameter pada Tabel 4.6.

Dari parameter-parameter tersebut , akan dilakukan latih model menggunakan LSTM. Cara kerjanya LSTM adalah sebagai berikut.

1. Pada tahapan LSTM pertama, ditentukan informasi yang akan dihilangkan dalam *cell state* dengan *forget gate* sesuai dengan persamaan (8). Perlu diketahui bahwa nilai bobot dan *bias* akan ditentukan secara acak setiap inisialisasi model dilakukan. Pada model ini menggunakan optimasi *adam* untuk menentukan bobot dan *bias*-nya. Parameter optimizer *adam* merupakan sebuah parameter fungsi yang menghitung bobot dan *bias* secara acak. Maka dari itu perlu dilakukan inisialisasi lebih dari satu kali untuk mendapatkan nilai yang paling optimal.
2. Berdasarkan (Chandriah and Naraganahalli, 2021), untuk optimasi *adam* itu sendiri bekerja pada perhitungan bobot dan *bias* sebelum dilakukan perhitungan pada *cell state*. Dalam menentukan bobot dan *bias*, optimasi ini menentukan terlebih dahulu nilai dari vektor momentum dan *squared gradient* yang didapatkan menggunakan persamaan(36).

(36)

Keterangan:

= estimasi rata-rata kuadrat gradien

= faktor pengurangan

= pada waktu sebelumnya

= gradien

(37)

Keterangan:

= estimasi rata-rata gradien

= pada waktu sebelumnya

(38)

Keterangan:

= nilai yang diperbaiki

= faktor pengurangan pangkat *t*

(39)

Keterangan:

= nilai yang diperbaiki

(40)

Keterangan:

= nilai bobot

= nilai bobot sebelumnya

= *learning rate*

= nilai konstanta yang mencegah pembagian nol

1. Setelah informasi yang dihilangkan telah dilakukan, *input gate* dapat menentukan nilai *input* dengan persamaan (10).
2. Dari nilai *input* yang telah ditentukan, *cell state* akan diperaharui dengan menggunakan persamaan (12).
3. Kemudian pada *output gate* dapat menghitung nilai *output* menggunakan persamaan (13).
4. Dari sini selesai satu iterasi yang kemudian proses 1-5 akan diulang sebanyak *epoch* yang telah ditentukan.
5. Keluaran yang dihasilkan adalah nilai prediksi dari data yang digunakan.

Dari pemodelan LSTM di atas, berikut merupakan hasil informasi yang didapatkan yang dapat dilihat pada Tabel 4.13.

Tabel 4.13. Hasil LSTM BRI pembagian data pertama data latih 80%

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *Epoch* | *Batch Size* | *Loss* | MAPE | Waktu Kompilasi (detik) |
| 50 | 64 | 0.0559 | 19079446.0 | 1.7 |
| 100 | 64 | 0.0491 | 18432604.0 | 2.3 |
| 150 | 64 | 0.0182 | 5531986.0 | 2.9 |
| 200 | 64 | 0.0182 | 5841409.0 | 3.3 |
| 300 | 64 | 0.0180 | 5635091.5 | 4.4 |

Berdasarkan Tabel 4.13 pada *epoch* di atas 100 terjadi penurunan *loss* yang cukup signifikan dari 0,0491 menjadi 0,0182. Pada *epoch* di atas 150 nilai *loss* tidak berubah atau perubahannya sangat kecil dan memakan lebih banyak waktu sehingga yang dapat diambil pada beberapa percobaan ini adalah menggunakan 150 *epoch* saja.

#### 4.4.1.4. LSTM Bank IBK Indonesia

Dalam menggunakan model LSTM, proses latih dilakukan berkali-kali hingga model mendapatkan hasil yang optimal. Terdapat dua model yang akan dibuat menggunakan LSTM Bank IBK Indonesia dengan data latih 80% dan data uji 20% sebagaimana pembagian parameter pada Tabel 4.6. Hasil dari percobaan LSTM dengan pembagian data ini dapat dilihat pada Tabel 4.14.

Tabel 4.14. Hasil LSTM Bank IBK pembagian data pertama data latih 80%

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *Epoch* | *Batch Size* | *Loss* | MAPE | Waktu Kompilasi (detik) |
| 50 | 64 | 0.0393 | 109.0337 | 1.7 |
| 100 | 64 | 0.0344 | 103.7441 | 2.3 |
| 150 | 64 | 0.0311 | 88.6333 | 2.9 |
| 200 | 64 | 0.0286 | 71.4418 | 3.3 |
| 300 | 64 | 0.0236 | 47.1807 | 4.4 |

Berdasarkan Tabel 4.14 tiap penambahan 50 *epoch* terjadi penurunan *loss* yang stabil yaitu sekitar 0,003 sampai 0,005. Pada *epoch* 300 nilai *loss* berkurang signifikan dan perbedaan waktu yang sedikit dengan 200 *epoch* sehingga yang dapat diambil pada beberapa percobaan ini adalah menggunakan 300 *epoch*.

### 4.4.2. Model Pembagian Data Latih 60%

#### 4.4.2.1. BRI ARFIMA(0,0.5,0)

Pada model ini menggunakan data saham BRI dengan parameter berdasarkan *plot* ACF dan PACF pada Gambar 4.3 dengan nilai *p* dan *q* adalah 0. Untuk ordo *fractional differenciation*-nya adalah 0.5 yang kemudian datanya akan dilatih dengan ARIMA berparameter *d* adalah 0 hingga mendapatkan hasil sebagai mana Tabel 4.15 berikut.

Tabel 4.15. Data prediksi BRI ARFIMA(0,0.5,0)

|  |  |
| --- | --- |
| Tanggal Perdagangan Terakhir | Prediksi Latih |
| 9/30/2019 | 0.00000 |
| 10/31/2019 | 0.82289 |
| 11/30/2019 | 0.78004 |
| 12/31/2019 | 0.80444 |
| 1/31/2020 | 0.84628 |
| … | … |
| 4/30/2021 | 0.93139 |
| 5/31/2021 | 0.81896 |
| 6/30/2021 | 0.76833 |
| 7/31/2021 | 0.80255 |
| 8/31/2021 | 0.71493 |

Dari data dalam rentang 0 sampai 1 dapat dijadikan data kembali sebagaimana data aslinya dapat dilihat pada Tabel 4.16.

Tabel 4.16. Data prediksi dernormalisasi BRI ARFIMA(0,0.5,0)

|  |  |
| --- | --- |
| Tanggal Perdagangan Terakhir | Prediksi Latih |
| 9/30/2019 | 2530.625000 |
| 10/31/2019 | 4130.832383 |
| 11/30/2019 | 3968.529991 |
| 12/31/2019 | 4059.900280 |
| 1/31/2020 | 4223.121770 |
| … | … |
| 4/30/2021 | 4580.656973 |
| 5/31/2021 | 4115.615689 |
| 6/30/2021 | 3925.687811 |
| 7/31/2021 | 4052.709620 |
| 8/31/2021 | 3738.493749 |

Kemudian untuk hasil prediksi data ujinya dapat dilihat sesuai Tabel 4.17 berikut.

Tabel 4.17. Data ramal denormalisasi BRI ARFIMA(0,0.5,0)

|  |  |
| --- | --- |
| Tanggal Perdagangan Terakhir | Prediksi Uji |
| 9/30/2021 | 3835.67 |
| 10/31/2021 | 3769.36 |
| 11/30/2021 | 3706.43 |
| 12/31/2021 | 3646.69 |
| 1/31/2022 | 3589.98 |
| … | … |
| 8/31/2022 | 3266.03 |
| 9/30/2022 | 3228.66 |
| 10/31/2022 | 3193.20 |
| 11/30/2022 | 3159.53 |
| 12/31/2022 | 3127.58 |

#### 4.4.2.2. Bank IBK ARFIMA(2,0.5,0)

Pada model ini menggunakan data saham Bank IBK Indonesia dengan parameter berdasarkan *plot* ACF dan PACF pada Gambar 4.4 dengan nilai *p* adalah 2 dan *q* adalah 0. Untuk ordo *fractional differenciation*-nya adalah 0.5 yang kemudian datanya akan dilatih dengan ARIMA berparameter *d* adalah 2 hingga mendapatkan hasil sebagai mana Tabel 4.18 berikut.

Tabel 4.18. Data prediksi ARFIMA(2,0.5,0)

|  |  |
| --- | --- |
| Tanggal Perdagangan Terakhir | Prediksi Latih |
| 9/30/2019 | 0.000000 |
| 10/31/2019 | 0.631138 |
| 11/30/2019 | 0.379234 |
| 12/31/2019 | 0.363820 |
| 1/31/2020 | 0.286624 |
| … | … |
| 4/30/2021 | 1.030036 |
| 5/31/2021 | 0.981913 |
| 6/30/2021 | 0.870013 |
| 7/31/2021 | 0.684795 |
| 8/31/2021 | 0.540321 |

Dari data dalam rentang 0 sampai 1 dapat dijadikan data kembali sebagaimana data aslinya dapat dilihat pada Tabel 4.19.

Tabel 4.19. Data prediksi denormalisasi Bank IBK ARFIMA(2,0.5,0)

|  |  |
| --- | --- |
| Tanggal Perdagangan Terakhir | Prediksi Latih |
| 9/30/2019 | 91.000000 |
| 10/31/2019 | 316.149998 |
| 11/30/2019 | 172.290215 |
| 12/31/2019 | 165.816397 |
| 1/31/2020 | 137.435424 |
| … | … |
| 4/30/2021 | 690.691327 |
| 5/31/2021 | 635.965692 |
| 6/30/2021 | 518.833116 |
| 7/31/2021 | 356.059799 |
| 8/31/2021 | 256.016414 |

Kemudian untuk hasil prediksi data ujinya dapat dilihat sesuai Tabel 4.20 berikut.

Tabel 4.20. Data ramal denormalisasi Bank IBK ARFIMA(2,0.5,0)

|  |  |
| --- | --- |
| Tanggal Perdagangan Terakhir | Prediksi Uji |
| 9/30/2021 | 176.231918 |
| 10/31/2021 | 160.950809 |
| 11/30/2021 | 137.659634 |
| 12/31/2021 | 106.481679 |
| 1/31/2022 | 99.091380 |
| … | … |
| 8/31/2022 | 212.822241 |
| 9/30/2022 | 258.169741 |
| 10/31/2022 | 312.981514 |
| 11/30/2022 | 373.683963 |
| 12/31/2022 | 442.518810 |

#### 4.4.2.3. LSTM BRI

Dalam menggunakan model LSTM, proses latih dilakukan berkali-kali hingga model mendapatkan hasil yang optimal. Terdapat dua model yang akan dibuat menggunakan LSTM BRI dengan data latih 60% dan data uji 40% sebagaimana pembagian parameter pada Tabel 4.6. Hasil dari percobaan LSTM dengan pembagian data ini dapat dilihat pada Tabel 4.21.

Tabel 4.21. Hasil LSTM BRI pembagian data kedua data latih 60%

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *Epoch* | *Batch Size* | *Loss* | MAPE | Waktu Kompilasi (detik) |
| 50 | 64 | 0.0589 | 21187552.0 | 1.7 |
| 100 | 64 | 0.0441 | 19107288.0 | 2.3 |
| 150 | 64 | 0.0304 | 15015791.0 | 2.8 |
| 200 | 64 | 0.0208 | 7501390.0 | 3.3 |
| 300 | 64 | 0.0205 | 7281831.5 | 4.4 |

Berdasarkan Tabel 4.21 tiap penambahan 50 *epoch* terjadi penurunan *loss* yang stabil yaitu sekitar 0,01. Pada *epoch* di atas 200 nilai *loss* tidak berubah atau perubahannya sangat kecil dan memakan lebih banyak waktu sehingga yang dapat diambil pada beberapa percobaan ini adalah menggunakan 200 *epoch* saja.

#### 4.4.2.4. LSTM Bank IBK Indonesia

Dalam menggunakan model LSTM, proses latih dilakukan berkali-kali hingga model mendapatkan hasil yang optimal. Terdapat dua model yang akan dibuat menggunakan LSTM Bank IBK Indonesia dengan data latih 60% dan data uji 40% sebagaimana pembagian parameter pada Tabel 4.6. Hasil dari percobaan LSTM dengan pembagian data ini dapat dilihat pada Tabel 4.22.

Tabel 4.22. Hasil LSTM Bank IBK pembagian data kedua data latih 60%

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *Epoch* | *Batch Size* | *Loss* | MAPE | Waktu Kompilasi (detik) |
| 50 | 64 | 0.0488 | 152.7699 | 1.6 |
| 100 | 64 | 0.0477 | 144.5893 | 3.1 |
| 150 | 64 | 0.0391 | 110.4037 | 3.2 |
| 200 | 64 | 0.0372 | 95.6701 | 3.4 |
| 300 | 64 | 0.0293 | 63.7587 | 5.3 |

Berdasarkan Tabel 4.22 tiap penambahan 50 *epoch* terjadi penurunan *loss* yang stabil yaitu sekitar 0,001 sampai 0,008. Pada *epoch* 300 nilai *loss* berkurang signifikan dan perbedaan waktu yang sedikit dengan 200 *epoch* sehingga yang dapat diambil pada beberapa percobaan ini adalah menggunakan 300 *epoch*.

### 4.4.3. Model Pembagian Data Latih 20%

#### 4.4.3.1. BRI ARFIMA(0,0.5,0)

Pada model ini menggunakan data saham BRI dengan parameter berdasarkan *plot* ACF dan PACF pada Gambar 4.5 dengan nilai *p* dan *q* adalah 0. Untuk ordo *fractional differenciation*-nya adalah 0.5 yang kemudian datanya akan dilatih dengan ARIMA berparameter *d* adalah 1 hingga mendapatkan hasil sebagai mana Tabel 4.23 berikut.

Tabel. 4.23. Data prediksi BRI ARFIMA(0,0.5,0)

|  |  |
| --- | --- |
| Tanggal Perdagangan Terakhir | Prediksi Latih |
| 9/30/2019 | 0.00000 |
| 10/31/2019 | 0.84108 |
| 11/30/2019 | 0.80264 |
| 12/31/2019 | 0.82157 |
| 1/31/2020 | 0.86284 |
| 2/29/2020 | 0.92440 |
| 3/31/2020 | 0.90765 |
| 4/30/2020 | 0.62066 |

Dari data dalam rentang 0 sampai 1 dapat dijadikan data kembali sebagaimana data aslinya dapat dilihat pada Tabel 4.24.

Tabel 4.24. Data prediksi denormalisasi BRI ARFIMA(0,0.5,0)

|  |  |
| --- | --- |
| Tanggal Perdagangan Terakhir | Prediksi Latih |
| 9/30/2019 | 2530.625000 |
| 10/31/2019 | 4202.380952 |
| 11/30/2019 | 4053.043478 |
| 12/31/2019 | 4125.714285 |
| 1/31/2020 | 4290.000000 |
| 2/29/2020 | 4550.000000 |
| 3/31/2020 | 4477.500000 |
| 4/30/2020 | 3440.952381 |

Kemudian untuk hasil prediksi data ujinya dapat dilihat sesuai Tabel 4.25 berikut.

Tabel 4.25. Data ramal denormalisasi BRI ARFIMA(0,0.5,0)

|  |  |
| --- | --- |
| Tanggal Perdagangan Terakhir | Prediksi Latih |
| 5/31/2020 | 2785.71429 |
| 6/30/2020 | 2785.71429 |
| 7/31/2020 | 2785.71429 |
| 8/31/2020 | 2785.71429 |
| 9/30/2020 | 2785.71429 |
| … | … |
| 8/31/2022 | 2785.71429 |
| 9/30/2022 | 2785.71429 |
| 10/31/2022 | 2785.71429 |
| 11/30/2022 | 2785.71429 |
| 12/31/2022 | 2785.71429 |

#### 4.4.3.2. Bank IBK ARFIMA(0,0.5,0)

Pada model ini menggunakan data saham Bank IBK Indonesia dengan parameter berdasarkan *plot* ACF dan PACF pada Gambar 4.5 dengan nilai *p* dan *q* adalah 0. Untuk ordo *fractional differenciation*-nya adalah 0.5 yang kemudian datanya akan dilatih dengan ARIMA berparameter *d* adalah 1 hingga mendapatkan hasil sebagai mana Tabel 4.26 berikut.

Tabel 4.26. Data prediksi Bank IBK ARFIMA(0,0.5,0)

|  |  |
| --- | --- |
| Tanggal Perdagangan Terakhir | Prediksi Latih |
| 9/30/2019 | 0.000000 |
| 10/31/2019 | 0.420759 |
| 11/30/2019 | 0.399997 |
| 12/31/2019 | 0.383422 |
| 1/31/2020 | 0.299761 |
| 2/29/2020 | 0.339232 |
| 3/31/2020 | 0.331996 |
| 4/30/2020 | 0.305802 |

Dari data dalam rentang 0 sampai 1 dapat dijadikan data kembali sebagaimana data aslinya dapat dilihat pada Tabel 4.27.

Tabel 4.27. Data prediksi denormalisasi Bank IBK ARFIMA(0,0.5,0)

|  |  |
| --- | --- |
| Tanggal Perdagangan Terakhir | Prediksi Latih |
| 9/30/2019 | 91.000000 |
| 10/31/2019 | 191.066667 |
| 11/30/2019 | 181.434783 |
| 12/31/2019 | 174.095238 |
| 1/31/2020 | 141.789474 |
| 2/29/2020 | 156.045455 |
| 3/31/2020 | 153.300000 |
| 4/30/2020 | 143.857143 |

Kemudian untuk hasil prediksi data ujinya dapat dilihat sesuai Tabel 4.28 berikut.

Tabel 4.28. Data ramal denormalisasi Bank IBK ARFIMA(0,0.5,0)

|  |  |
| --- | --- |
| Tanggal Perdagangan Terakhir | Prediksi Latih |
| 5/31/2020 | 129.666667 |
| 6/30/2020 | 129.666667 |
| 7/31/2020 | 129.666667 |
| 8/31/2020 | 129.666667 |
| 9/30/2020 | 129.666667 |
| … | … |
| 8/31/2022 | 129.666667 |
| 9/30/2022 | 129.666667 |
| 10/31/2022 | 129.666667 |
| 11/30/2022 | 129.666667 |
| 12/31/2022 | 129.666667 |

#### 4.4.3.3. LSTM BRI

Dalam menggunakan model LSTM, proses latih dilakukan berkali-kali hingga model mendapatkan hasil yang optimal. Terdapat dua model yang akan dibuat menggunakan LSTM BRI dengan data latih 20% dan data uji 80% sebagaimana pembagian parameter pada Tabel 4.6. Hasil dari percobaan LSTM dengan pembagian data ini dapat dilihat pada Tabel 4.29.

Tabel 4.29. Hasil LSTM BRI pembagian data ketiga data latih 60%

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *Epoch* | *Batch Size* | *Loss* | MAPE | Waktu Kompilasi (detik) |
| 50 | 64 | 0.0866 | 61518744.000 | 1.8 |
| 100 | 64 | 0.0684 | 50704928.000 | 2.2 |
| 150 | 64 | 0.0523 | 38399592.000 | 2.8 |
| 200 | 64 | 0.0337 | 14321970.000 | 3.4 |
| 300 | 64 | 0.0297 | 952087.125 | 4.4 |

Berdasarkan Tabel 4.29 tiap penambahan 50 *epoch* terjadi penurunan *loss* yang stabil yaitu sekitar 0,01 sampai 0,02. Pada *epoch* 300 nilai *loss* berkurang signifikan sehingga yang dapat diambil pada beberapa percobaan ini adalah menggunakan 300 *epoch*.

#### 4.4.3.4. LSTM Bank IBK Indonesia

Dalam menggunakan model LSTM, proses latih dilakukan berkali-kali hingga model mendapatkan hasil yang optimal. Terdapat dua model yang akan dibuat menggunakan LSTM Bank IBK Indonesia dengan data latih 20% dan data uji 80% sebagaimana pembagian parameter pada Tabel 4.6. Hasil dari percobaan LSTM dengan pembagian data ini dapat dilihat pada Tabel 4.30.

Tabel 4.30. Hasil LSTM Bank IBK pembagian data ketiga data latih 60%

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *Epoch* | *Batch Size* | *Loss* | MAPE | Waktu Kompilasi (detik) |
| 50 | 64 | 0.0019 | 60.4006 | 1.5 |
| 100 | 64 | 0.0014 | 46.675 | 2.7 |
| 150 | 64 | 0.0013 | 45.1134 | 2.9 |
| 200 | 64 | 0.0013 | 45.1613 | 3.2 |
| 300 | 64 | 0.0012 | 44.1631 | 4.3 |

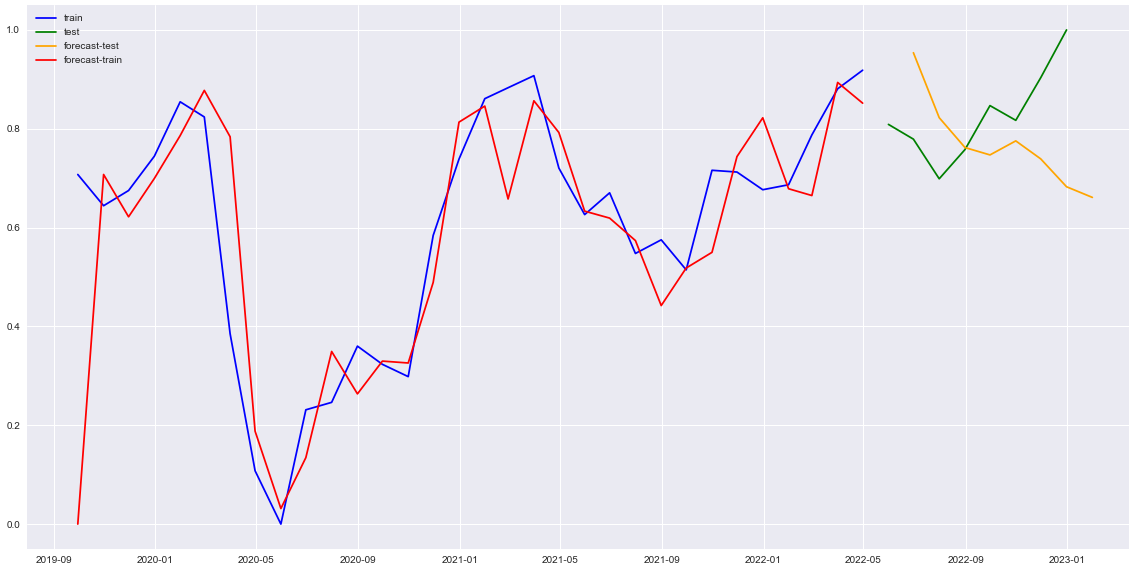
Berdasarkan Tabel 4.30 tiap penambahan 50 *epoch* terjadi penurunan *loss* yang stabil yaitu sekitar 0,0001 sampai 0,0005. Pada *epoch* 100 ke atas nilai *loss* tidak berkurang atau selisih sedikit saja membuat waktu kompilasinya tidak efisien sehingga yang dapat diambil pada beberapa percobaan ini adalah menggunakan 100 *epoch*.

## 4.5. Prediksi Seluruh Model

Semua data yang telah dilatih dengan algoritma ARFIMA dan LSTM dapat dilakukan *plotting* untuk melihat perbandingan antara grafik data asli dengan data terprediksi. Berikut merupakan *plot* dari hasil latih model.

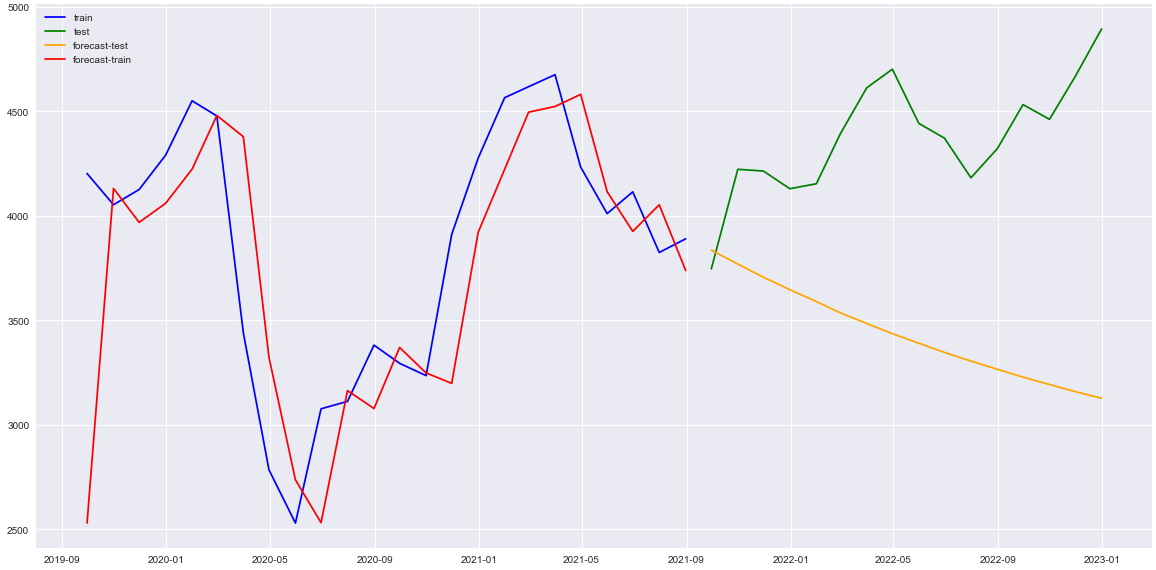
### 4.5.1. Prediksi ARFIMA

Pemodelan ARFIMA menghasilkan enam model yang dibuat berdasarkan pembagian data. Berikut merupakan hasil dari *plotting* dari model ARFIMA.



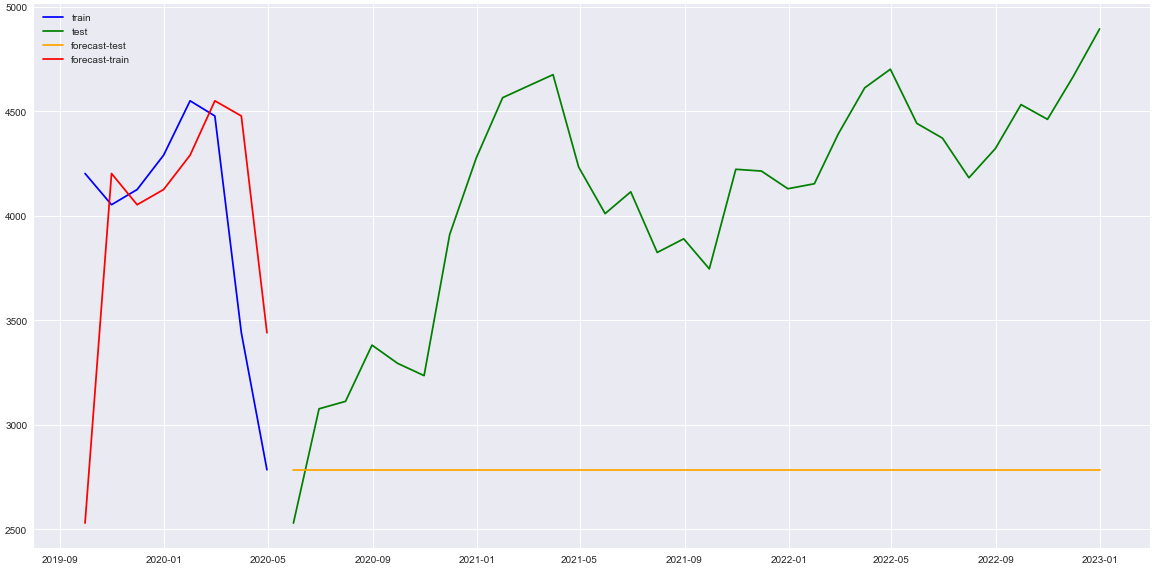
Gambar 4.5. Hasil prediksi ARFIMA BRI data latih 80%

Pada Gambar 4.5 dapat dilihat bahwa terdapat jarak antara grafik latih dan uji. Hal ini terjadi karena *plotting* garis di atas tersambung berdasarkan titik harga penutupan pada data yang sama. Karena data sudah dipisah menjadi data latih dan data uji, maka membuat keduanya tidak tersambung dalam satu garis. Untuk hasil pengujian menggunakan data latih (warna merah) cukup baik karena masih mengikuti sebagaimana pergerakan harga penutupan saham pada data latih (warna biru). Sedangkan untuk hasil uji menggunakan data ujinya (warna kuning) dapat dilihat bahwa hasil tersebut bertolak belakang dengan data ujinya (warna hijau) yang seharusnya harga penutupannya naik, tetapi hasil ujinya menunjukan penurunan pada harga penutupannya yang membuat hasil uji menggunakan data uji tidak sesuai dengan kondisi nyata.



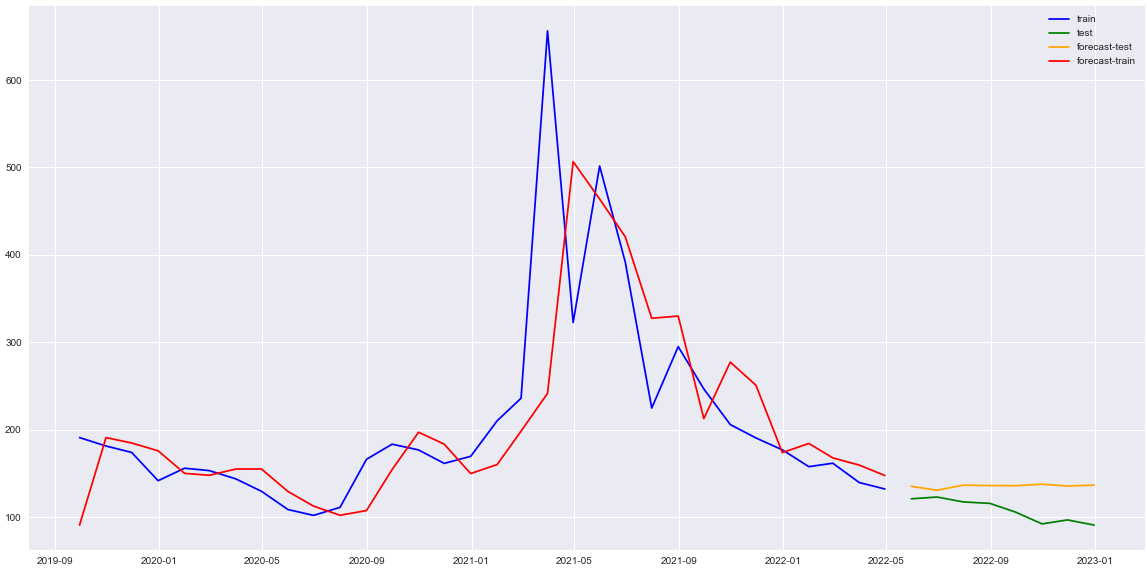
Gambar 4.6. Hasil prediksi ARFIMA BRI data latih 60%

Pada Gambar 4.6 dapat dilihat juga untuk hasil uji data latih (warna merah) masih mengikuti pergerakan data latih (warna biru) sebagaimana penggunaan data latih 80%. Namun untuk pengujian data uji (warna kuning), menghasilkan data ramal yang bertolak belakang dengan data uji (warna hijau).



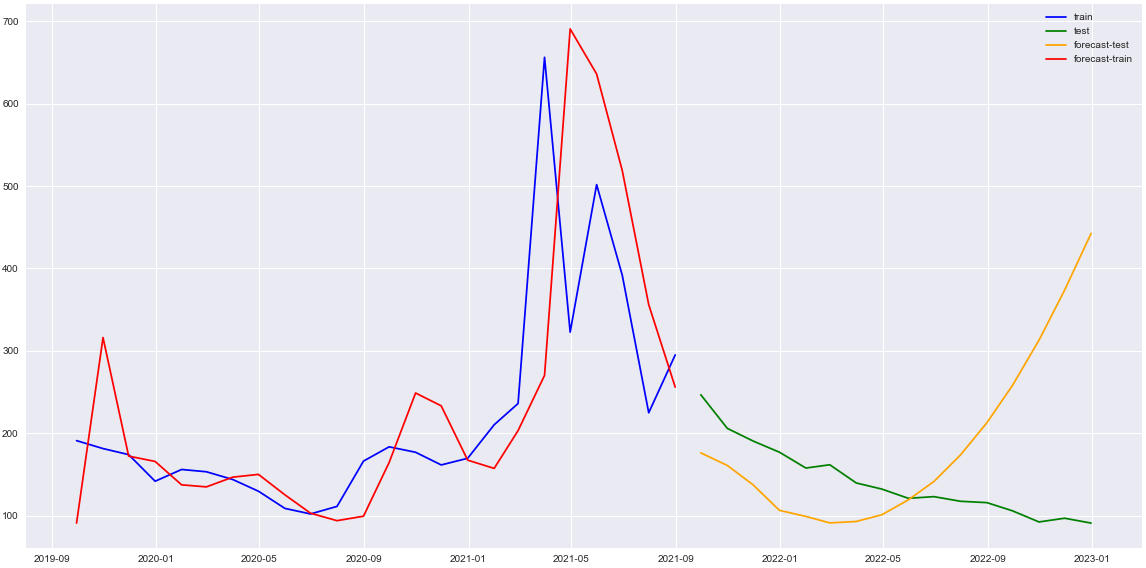
Gambar 4.7. Hasil prediksi ARFIMA BRI data latih 20%

Pada Gambar 4.7 dapat dilihat juga untuk hasil uji data latih (warna merah) masih mengikuti pergerakan data latih (warna biru) sebagaimana penggunaan data latih 80% dan 60%. Pada pengujian data uji (warna kuning), hasilnya sangat tidak sesuai dengan data uji (warna hijau) yang seharusnya terdapat naik turun harga penutupan pada jangka waktu tertentu, tetapi pada hasil uji harga tetap stabil sampai akhir periode. Hal ini terjadi karena kurangnya informasi data yang dapat dilatih dibanding dengan varian data ujinya (Nisbet, Miner and Yale, 2020).



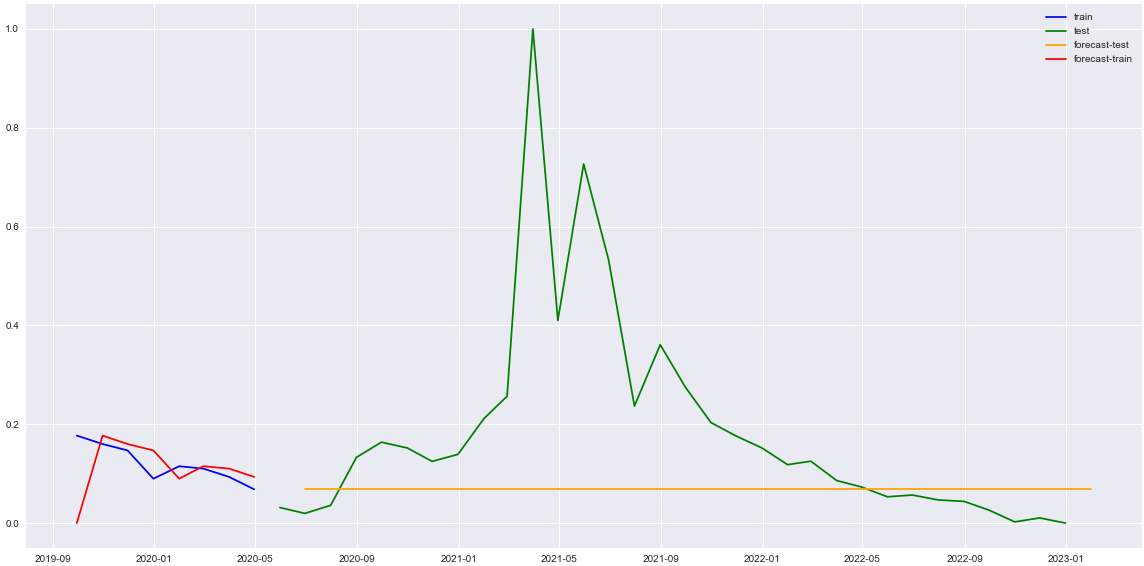
Gambar 4.8. Hasil prediksi ARFIMA Bank IBK data latih 80%

Pada Gambar 4.8 dapat dilihat bahwa hasil uji dengan data latih (warna merah) sudah cukup sesuai dengan data latih (warna biru) meskipun terdapat persimpangan pada bulan Juni 2021 hingga September 2021. Untuk hasil uji data uji (warna kuning) pada beberapa bulan awalnya sudah cukup sesuai meskipun jumlah pastinya berbeda namun untuk pertumbuhan sudah cukup sesuai, tetapi setelah September 2022 data uji (warna hijau) yang mengalami penurunan, sedangkan hasil ujinya stabil sehingga masih kurang sesuai.



Gambar 4.9. Hasil prediksi ARFIMA Bank IBK data latih 60%

Pada Gambar 4.9 dapat dilihat bahwa hasil uji dengan data latih (warna merah) sudah cukup sesuai dengan data latih (warna biru) sebagai mana penjelasan pada pembagian data 80%. Untuk hasil uji data uji (warna kuning) pada beberapa bulan pertama sudah cukup sesuai pertumbuhan harga penutupannya. Akan tetapi pada bulan Maret-April dan setelahnya yang seharusnya pada data uji (warna hijau) mengalami penurunan yang konstan, pada hasil ujinya mengalami kenaikan yang sangat tinggi secara konstan sehingga kurang sesuai.

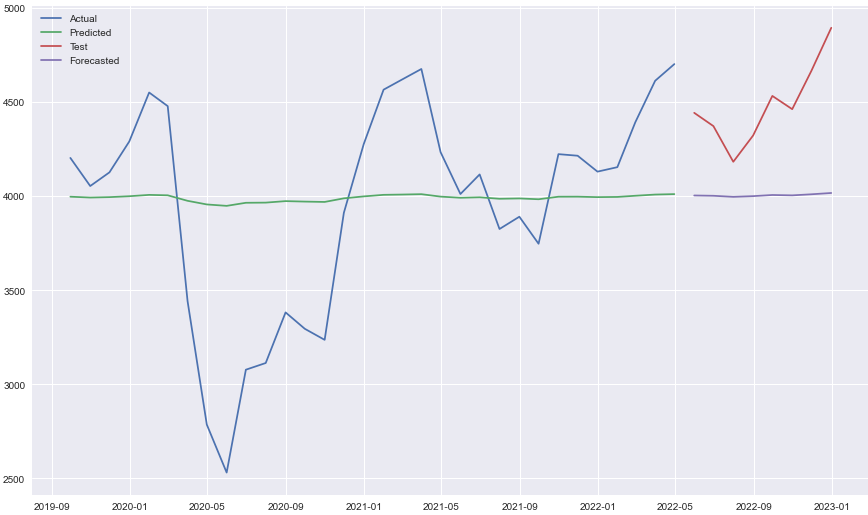


Gambar 4.10. Hasil prediksi ARFIMA Bank IBK data latih 20%

Pada Gambar 4.10 dapat dilihat bahwa hasil uji dengan data latih (warna merah) sudah cukup sesuai dengan data latih (warna biru). Untuk hasil uji data uji (warna kuning) terlihat sangat jauh dari data ujinya (warna hijau). Pada kondisi nyata mengalami kenaikan dan penurunan, tetapi pada hasil uji menunjukan kestabilan harga.

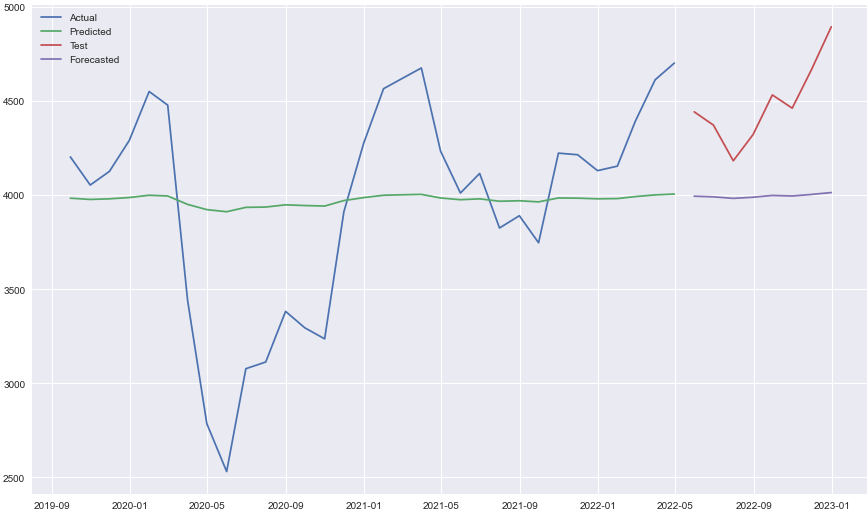
### 4.5.2. Prediksi LSTM

Pemodelan LSTM menghasilkan tiga puluh model yang dibuat berdasarkan pembagian data dan jumlah *epoch*. Berikut merupakan hasil dari *plotting* dari model LSTM.



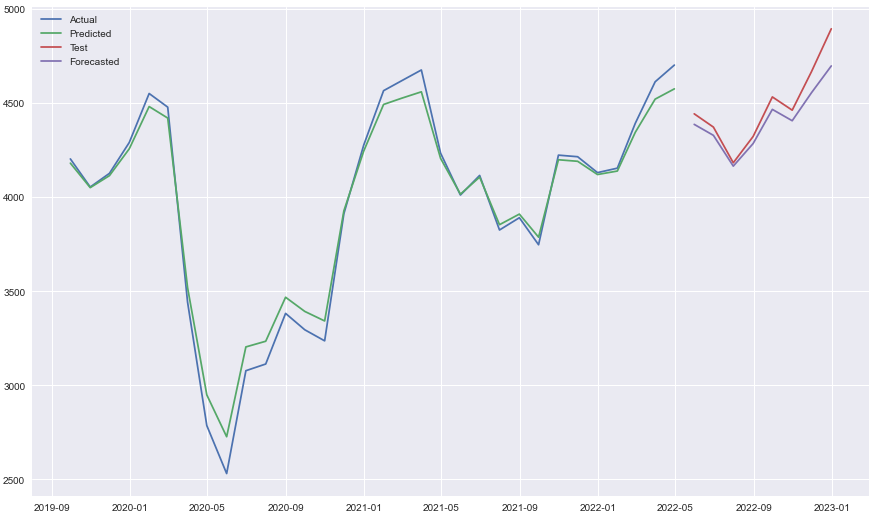
Gambar 4.11. Hasil prediksi LSTM BRI data latih 80% 50 *epoch*

Pada Gambar 4.11 dapat dilihat bahwa baik hasil uji data latih (warna hijau) maupun hasil uji data uji (warna ungu) memperlihatkan hasil yang tidak sesuai keadaan nyata. Pada grafik data latih (warna biru) menunjukan pola naik turunnya pertumbuhan harga penutupan yang beragam, sedangkan untuk hasil ujinya menunjukan pola harga penutupan yang stabil sehingga membuat hasil uji tidak relevan.



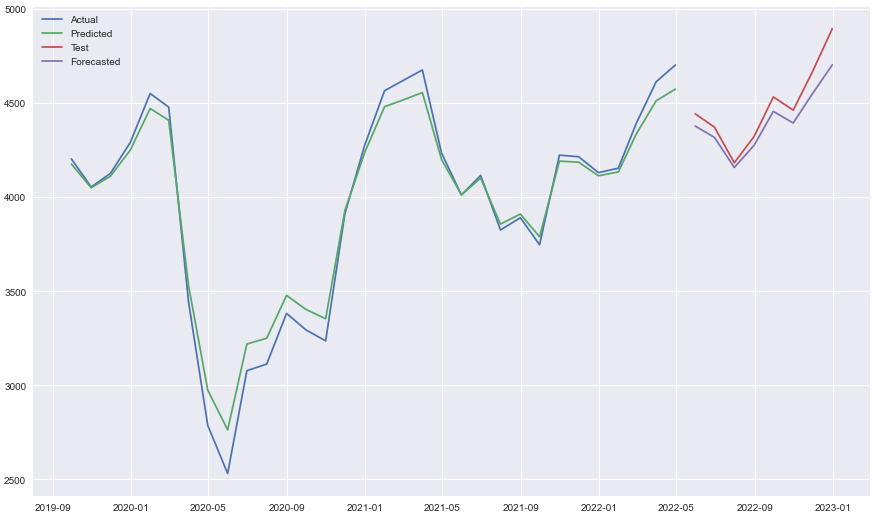
Gambar 4.12. Hasil prediksi LSTM BRI data latih 80% 100 *epoch*

Pada Gambar 4.12 dapat dilihat bahwa baik hasil uji data latih (warna hijau) maupun hasil uji data uji (warna ungu) memperlihatkan hasil yang tidak sesuai keadaan nyata sebagaimana pada 50 *epoch*. Pada grafik data latih (warna biru) menunjukan pola naik turunnya pertumbuhan harga penutupan yang beragam, sedangkan untuk hasil ujinya menunjukan pola harga penutupan yang stabil sehingga membuat hasil uji tidak relevan.



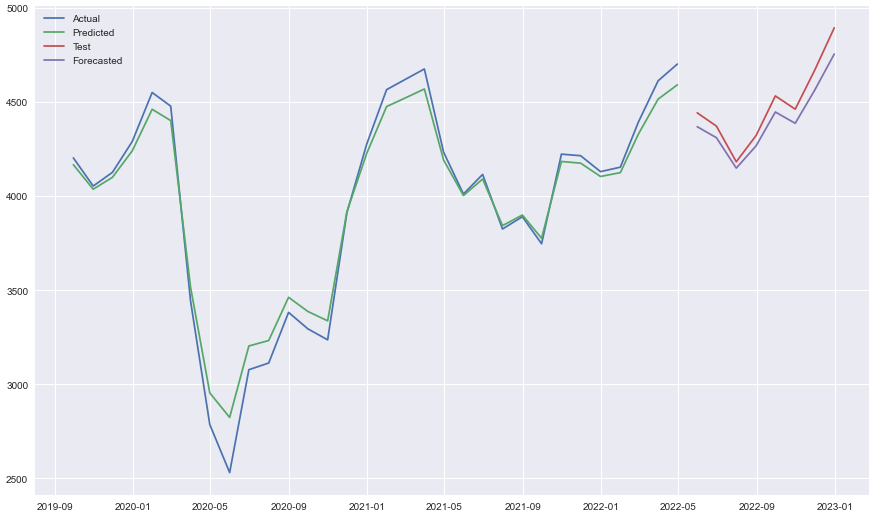
Gambar 4.13. Hasil prediksi LSTM BRI data latih 80% 150 *epoch*

Pada Gambar 4.13 dapat dilihat hasil uji data latih (warna hijau) dan hasil uji data uji (warna ungu) menunjukan pola yang hampir sama sebagaimana data yang digunakan sehingga sangat relevan dan sesuai pada kondisi nyata.



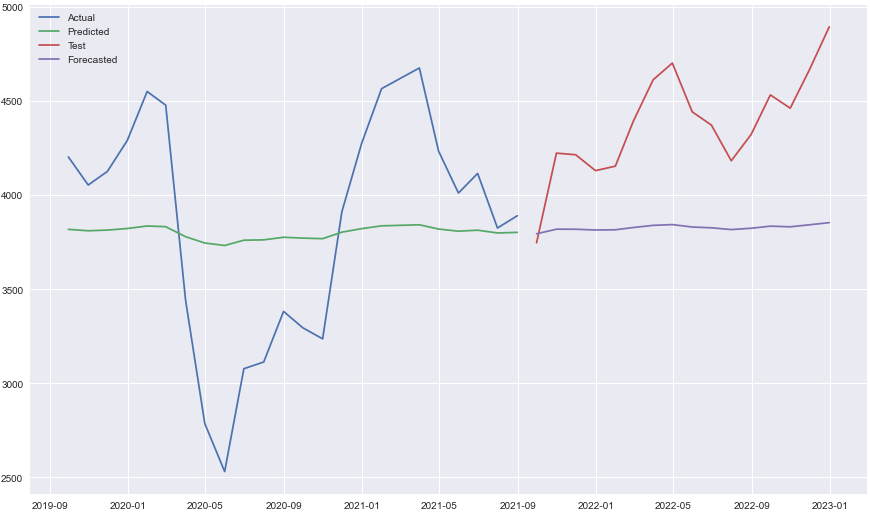
Gambar 4.14. Hasil prediksi LSTM BRI data latih 80% 200 *epoch*

Pada Gambar 4.14 dapat dilihat hasil uji data latih (warna hijau) dan hasil uji data uji (warna ungu) menunjukan pola yang hampir sama sebagaimana data yang digunakan sehingga sangat relevan dan sesuai pada kondisi nyata. Meski begitu hasil yang diperoleh tidak berbeda dari 150 *epoch*. Hal ini dapat membuat penelitian ini lebih memilih hasil 150 *epoch*.



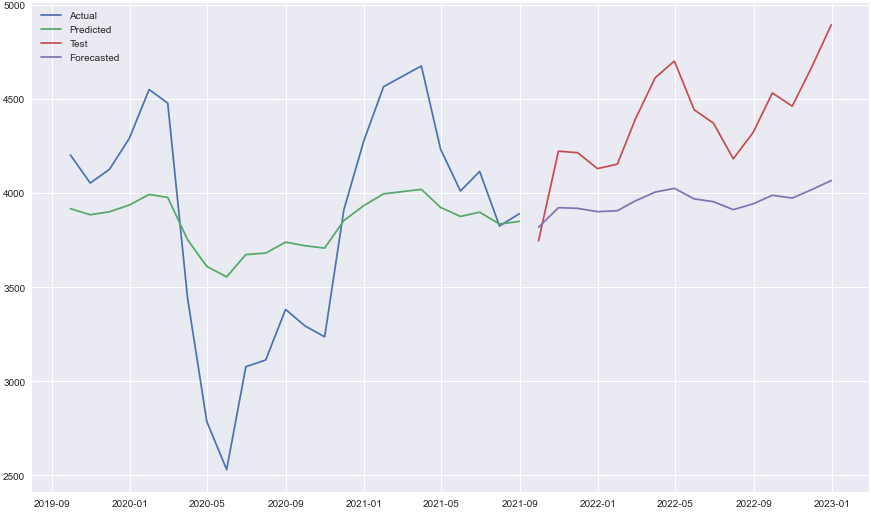
Gambar 4.15. Hasil prediksi LSTM BRI data latih 80% 300 *epoch*

Pada Gambar 4.14 dapat dilihat hasil uji data latih (warna hijau) dan hasil uji data uji (warna ungu) menunjukan pola yang hampir sama sebagaimana data yang digunakan sehingga sangat relevan dan sesuai pada kondisi nyata. Meski begitu hasil yang diperoleh tidak begitu jauh dan hanya sedikit lebih baik terlebih pada bulan Juni 2021 hingga sekitar September-Okober 2021 pada hasil uji data latih dibandingkan dari 150 *epoch* dan 200 *epoch*.



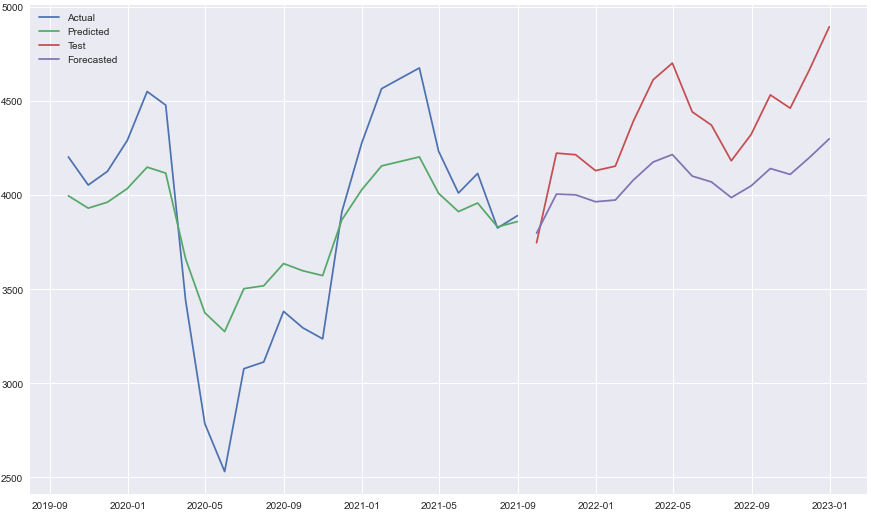
Gambar 4.16. Hasil prediksi LSTM BRI data latih 60% 50 *epoch*

Pada Gambar 4.16 dapat dilihat bahwa baik hasil uji data latih (warna hijau) maupun hasil uji data uji (warna ungu) memperlihatkan hasil yang tidak sesuai keadaan nyata. Pada grafik data latih (warna biru) menunjukan pola naik turunnya pertumbuhan harga penutupan yang beragam, sedangkan untuk hasil ujinya menunjukan pola harga penutupan yang stabil sehingga membuat hasil uji tidak relevan.



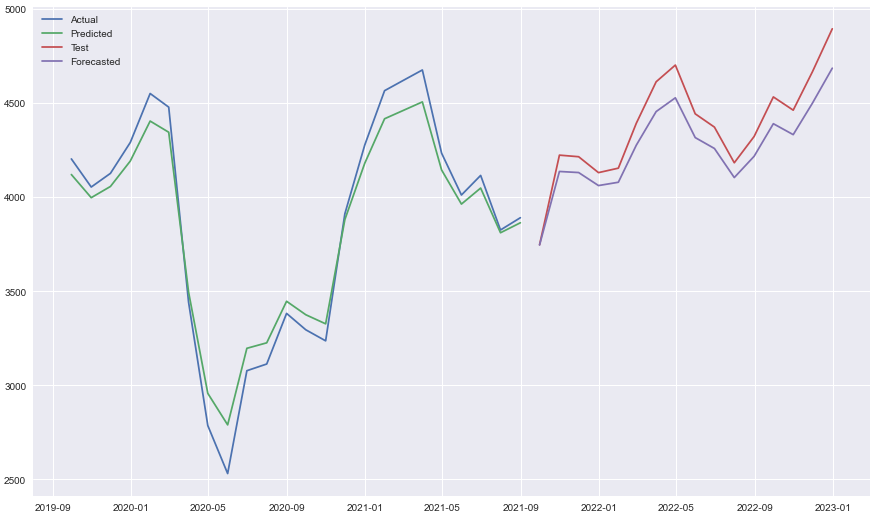
Gambar 4.17. Hasil prediksi LSTM BRI data latih 60% 100 *epoch*

Pada Gambar 4.17 dapat dilihat bahwa hasil uji sedikit lebih baik daripada menggunakan 50 *epoch*. Akan tetapi hasil prediksi yang didapatkan masih kurang memuaskan karena kurang sesuai pola pertumbuhannya pada kasus nyata.



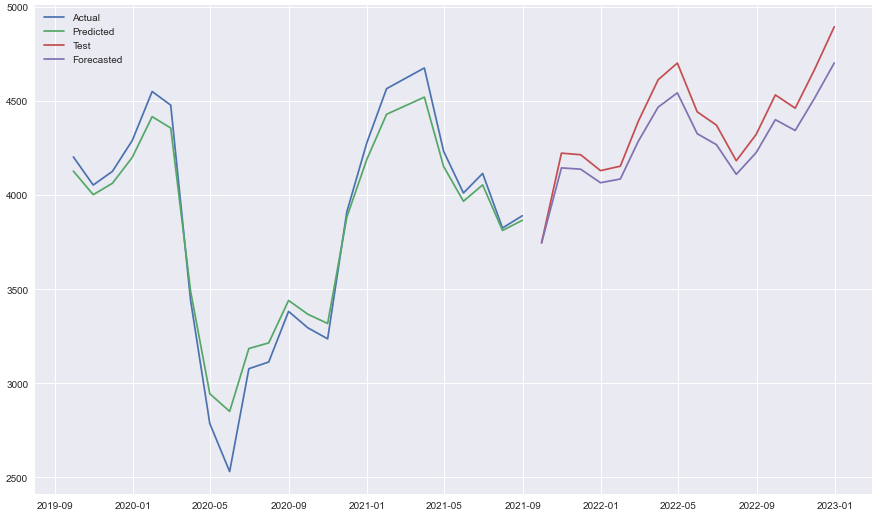
Gambar 4.18. Hasil prediksi LSTM BRI data latih 60% 150 *epoch*

Pada Gambar 4.18 dapat dilihat bahwa hasil uji juga masih sedikit lebih baik daripada menggunakan 100 *epoch*. Akan tetapi hasil prediksi yang didapatkan masih kurang memuaskan karena kurang sesuai pola pertumbuhannya pada kasus nyata.



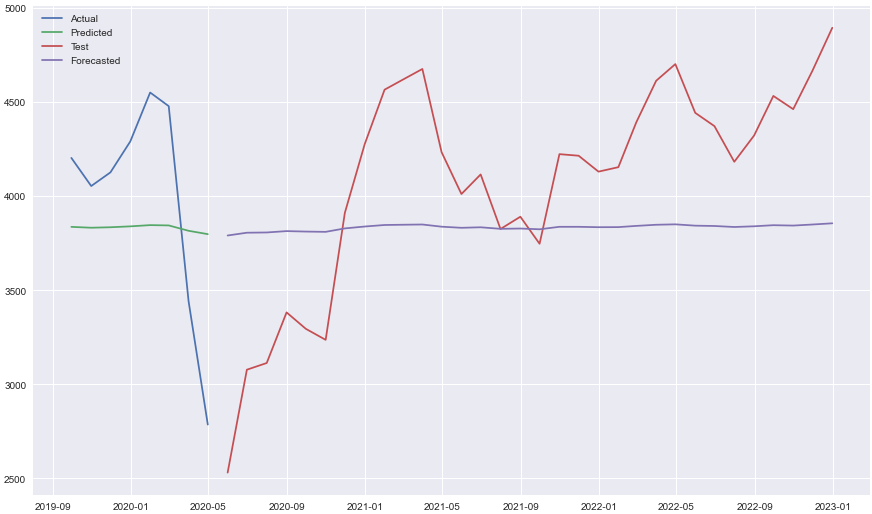
Gambar 4.19. Hasil prediksi LSTM BRI data latih 60% 200 *epoch*

Pada Gambar 4.19 dapat dilihat bahwa hasil uji data latih (warna hijau) dan hasil uji data uji (warna ungu) sudah cukup sesuai dengan data yang diuji.



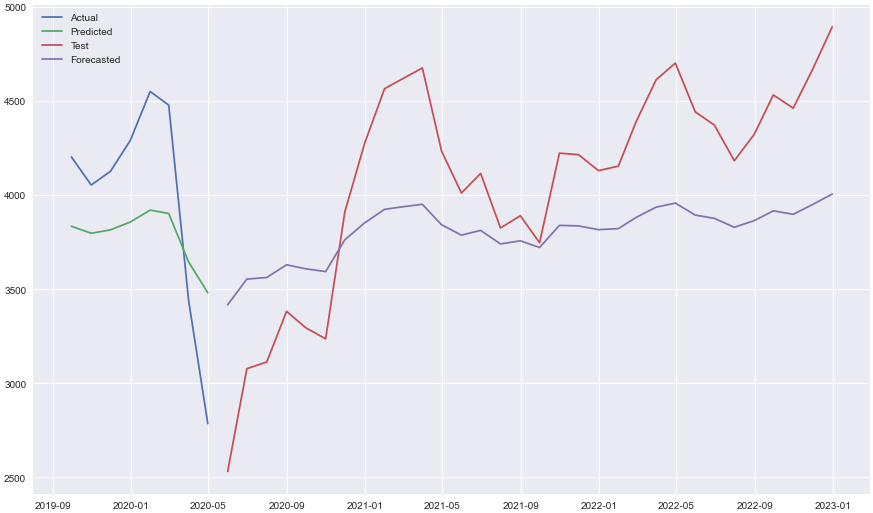
Gambar 4.20. Hasil prediksi LSTM BRI data latih 60% 300 *epoch*

Pada Gambar 4.20 dapat dilihat bahwa hasil uji data latih (warna hijau) dan hasil uji data uji (warna ungu) sudah cukup sesuai dengan data yang diuji. Namun dibandikan dengan 200 *epoch* hampir tidak ada perubahan sehingga lebih baik menggunakan 200 *epoch*.



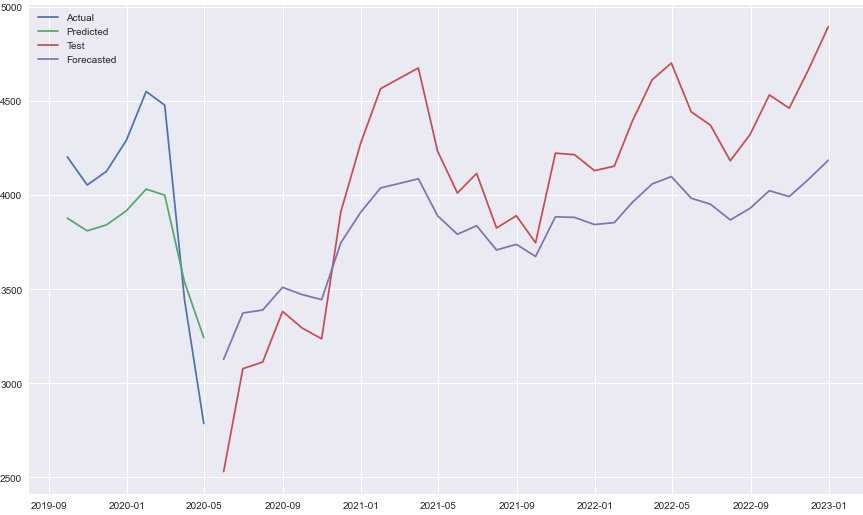
Gambar 4.21. Hasil prediksi LSTM BRI data latih 20% 50 *epoch*

Pada Gambar 4.21 dapat dilihat bahwa baik hasil uji data latih (warna hijau) maupun hasil uji data uji (warna ungu) memperlihatkan hasil yang tidak sesuai keadaan nyata. Pada grafik data latih (warna biru) menunjukan pola naik turunnya pertumbuhan harga penutupan yang beragam, sedangkan untuk hasil ujinya menunjukan pola harga penutupan yang stabil sehingga membuat hasil uji tidak relevan.



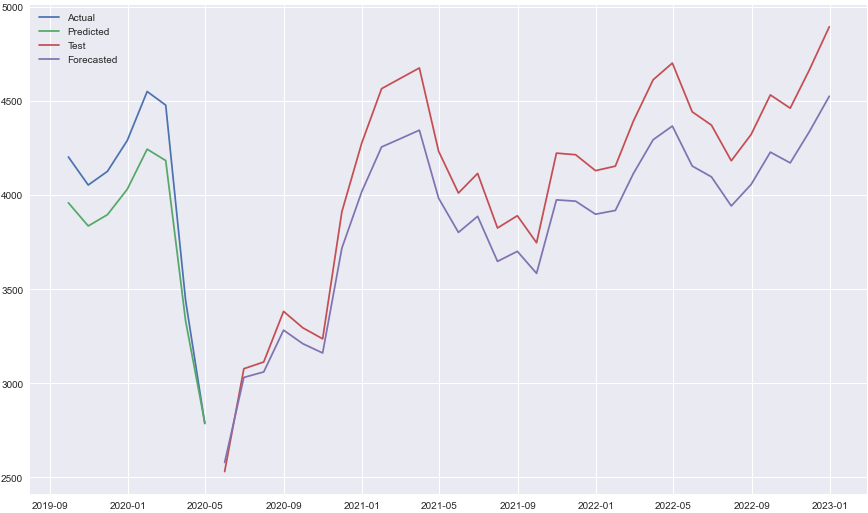
Gambar 4.22. Hasil prediksi LSTM BRI data latih 20% 100 *epoch*

Pada Gambar 4.23 dapat dilihat bahwa hasil uji sedikit lebih baik daripada menggunakan 50 *epoch*. Akan tetapi hasil prediksi yang didapatkan masih kurang memuaskan karena kurang sesuai pola pertumbuhannya pada kasus nyata.



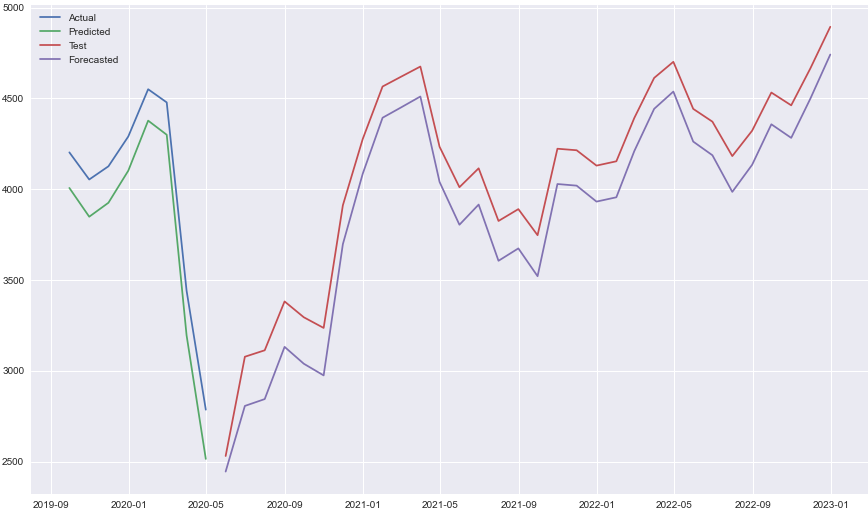
Gambar 4.24. Hasil prediksi LSTM BRI data latih 20% 150 *epoch*

Pada Gambar 4.24 dapat dilihat bahwa hasil uji juga masih sedikit lebih baik daripada menggunakan 100 *epoch*. Akan tetapi hasil prediksi yang didapatkan masih kurang memuaskan karena kurang sesuai pola pertumbuhannya pada kasus nyata.



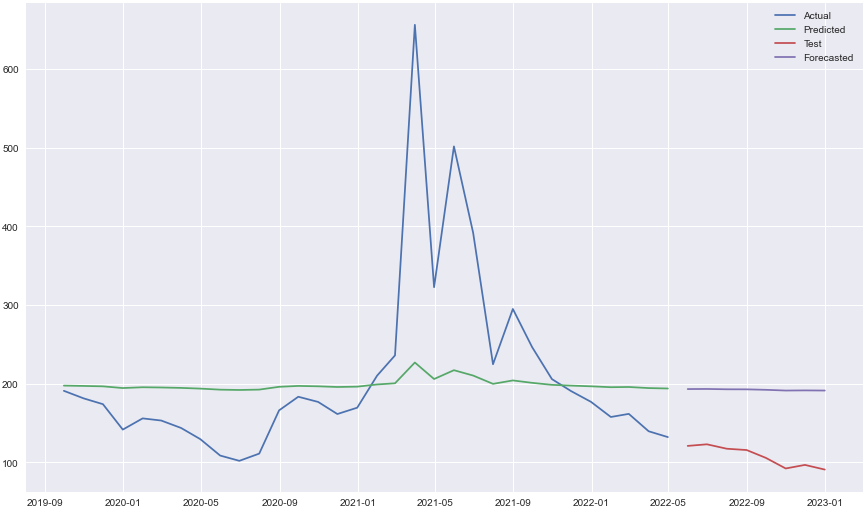
Gambar 4.25. Hasil prediksi LSTM BRI data latih 20% 200 *epoch*

Pada Gambar 4.25 dapat dilihat bahwa hasil uji data latih (warna hijau) dan hasil uji data uji (warna ungu) sudah cukup sesuai dengan data yang diuji.



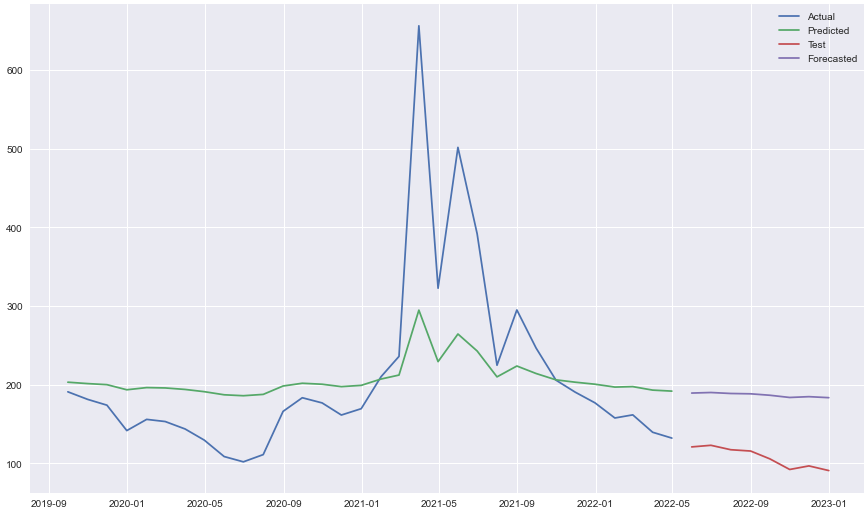
Gambar 4.26. Hasil prediksi LSTM BRI data latih 20% 300 *epoch*

Pada Gambar 4.30 dapat dilihat bahwa hasil uji data latih (warna hijau) dan hasil uji data uji (warna ungu) sudah cukup sesuai dengan data yang diuji. Namun dibandikan dengan 200 *epoch* hampir tidak ada perubahan sehingga lebih baik menggunakan 200 *epoch*.



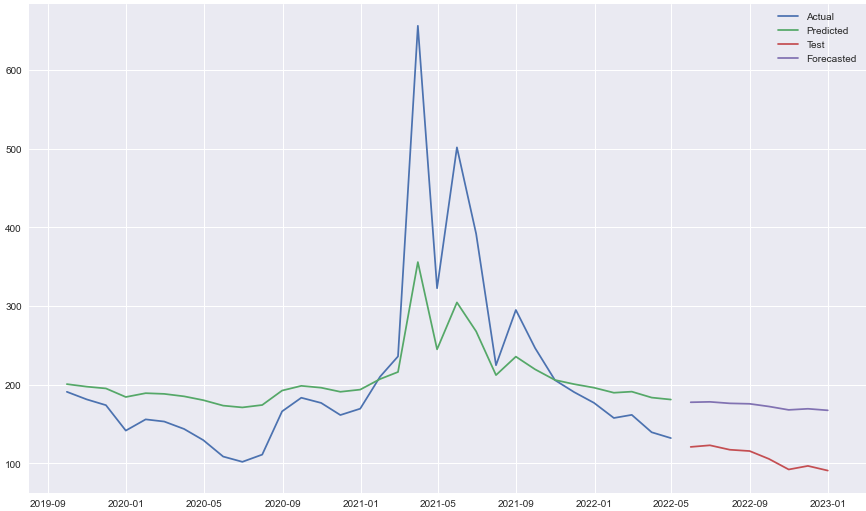
Gambar 4.27. Hasil prediksi LSTM Bank IBK data latih 80% 50 *epoch*

Pada Gambar 4.27 dapat dilihat bahwa baik hasil uji data latih (warna hijau) maupun hasil uji data uji (warna ungu) memperlihatkan hasil yang tidak sesuai keadaan nyata. Pada grafik data latih (warna biru) menunjukan pola naik turunnya pertumbuhan harga penutupan yang beragam, sedangkan untuk hasil ujinya menunjukan pola harga penutupan yang stabil sehingga membuat hasil uji tidak relevan.



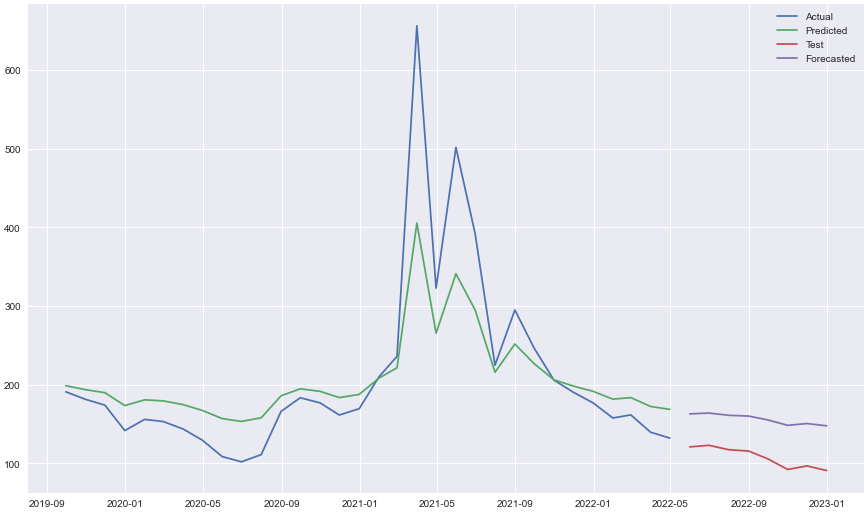
Gambar 4.28. Hasil prediksi LSTM Bank IBK data latih 80% 100 *epoch*

Pada Gambar 4.28 dapat dilihat bahwa hasil uji sedikit lebih baik daripada menggunakan 50 *epoch*. Akan tetapi hasil prediksi yang didapatkan masih kurang memuaskan karena kurang sesuai pola pertumbuhannya pada kasus nyata.



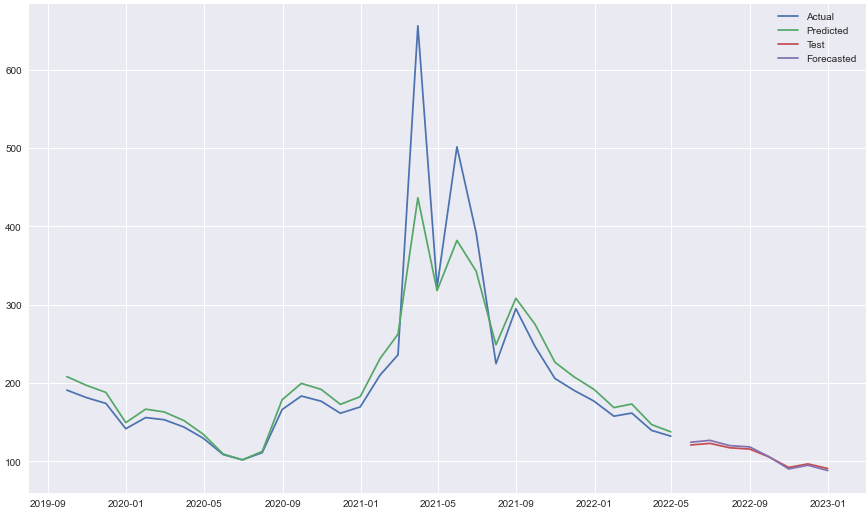
Gambar 4.29. Hasil prediksi LSTM Bank IBK data latih 80% 150 *epoch*

Pada Gambar 4.29 dapat dilihat bahwa hasil uji juga masih sedikit lebih baik daripada menggunakan 100 *epoch*. Akan tetapi hasil prediksi yang didapatkan masih kurang memuaskan karena kurang sesuai pola pertumbuhannya pada kasus nyata.



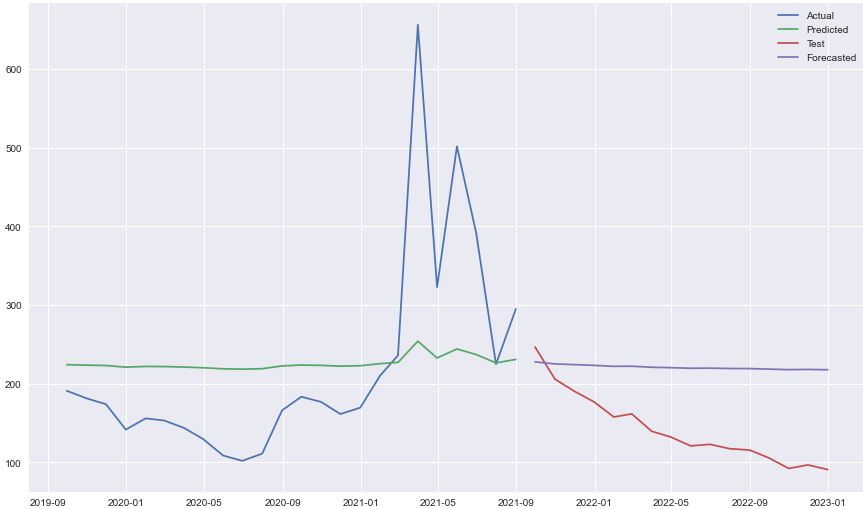
Gambar 4.30. Hasil prediksi LSTM Bank IBK data latih 80% 200 *epoch*

Pada Gambar 4.30 dapat dilihat bahwa hasil uji data latih (warna hijau) dan hasil uji data uji (warna ungu) sudah cukup sesuai dengan data yang diuji.



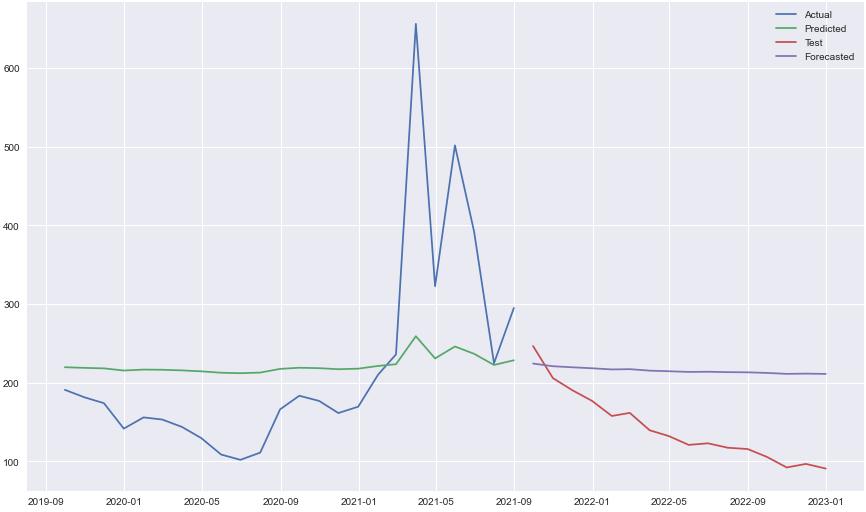
Gambar 4.31. Hasil prediksi LSTM Bank IBK data latih 80% 300 *epoch*

Pada Gambar 4.31 dapat dilihat bahwa tingkat keakuratan antara hasil uji data uji (warna ungu) dengan data uji (warna merah) cukup tinggi. Oleh karena itu penggunaan 300 *epoch* dapat digunakan untuk pelatihan data Bank IBK.



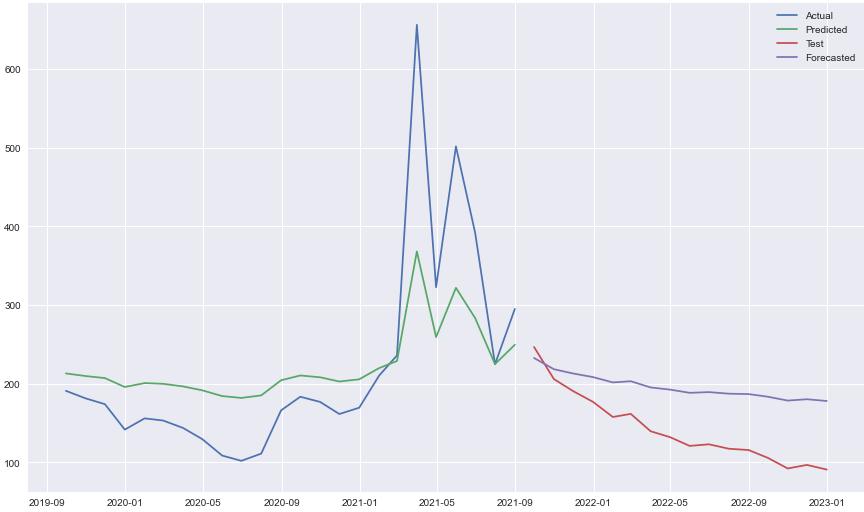
Gambar 4.32. Hasil prediksi LSTM Bank IBK data latih 60% 50 *epoch*

Pada Gambar 4.32 dapat dilihat bahwa baik hasil uji data latih (warna hijau) maupun hasil uji data uji (warna ungu) memperlihatkan hasil yang tidak sesuai keadaan nyata. Pada grafik data latih (warna biru) menunjukan pola naik turunnya pertumbuhan harga penutupan yang beragam, sedangkan untuk hasil ujinya menunjukan pola harga penutupan yang stabil sehingga membuat hasil uji tidak relevan.



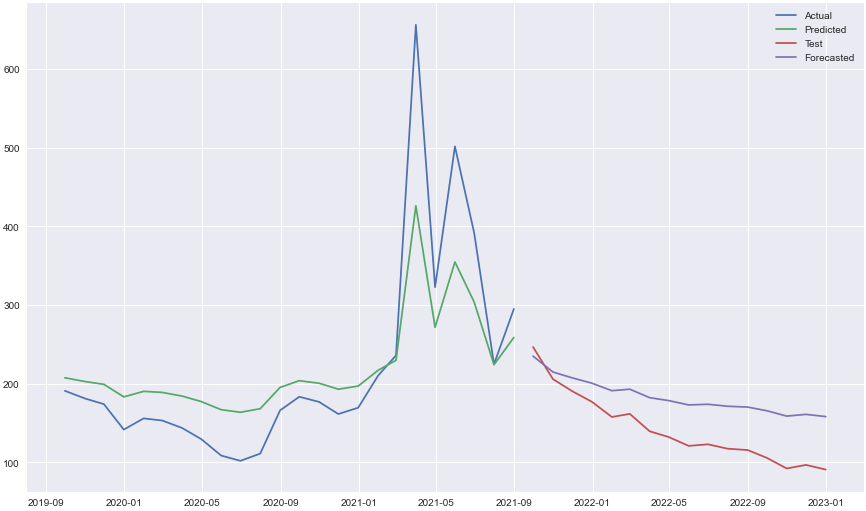
Gambar 4.33. Hasil prediksi LSTM Bank IBK data latih 60% 100 *epoch*

Pada 4.33 dapat dilihat bahwa hasil uji sedikit lebih baik daripada menggunakan 50 *epoch*. Akan tetapi hasil prediksi yang didapatkan masih kurang memuaskan karena kurang sesuai pola pertumbuhannya pada kasus nyata.



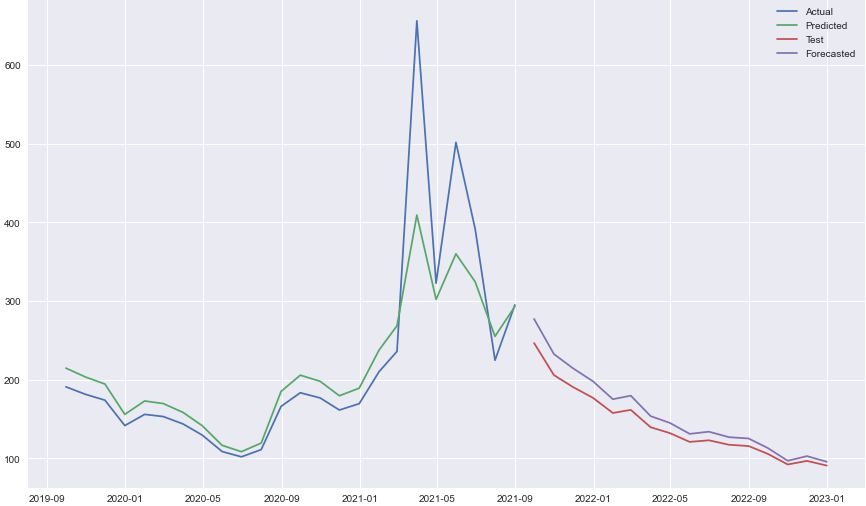
Gambar 4.34. Hasil prediksi LSTM Bank IBK data latih 60% 150 *epoch*

Pada Gambar 4.29 dapat dilihat bahwa hasil uji juga masih sedikit lebih baik daripada menggunakan 100 *epoch*. Akan tetapi hasil prediksi yang didapatkan masih kurang memuaskan karena kurang sesuai pola pertumbuhannya pada kasus nyata.



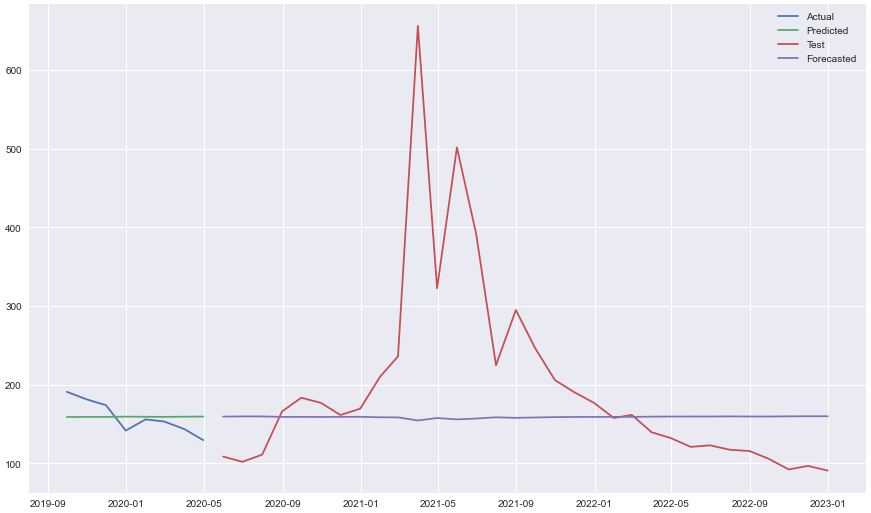
Gambar 4.35. Hasil prediksi LSTM Bank IBK data latih 60% 200 *epoch*

Pada Gambar 4.35 dapat dilihat bahwa hasil uji menggunakan data latih (warna hijau) masih sebaik penggunaan data latih sebesar 80%. Namun pada hasil uji data uji (warna ungu) keakuratannya sedikit menurun. Hal ini disebabkan karena kurangnya data yang dilatih sebagaimana dijelaskan pada hasil LSTM BRI dengan data latih 20%.



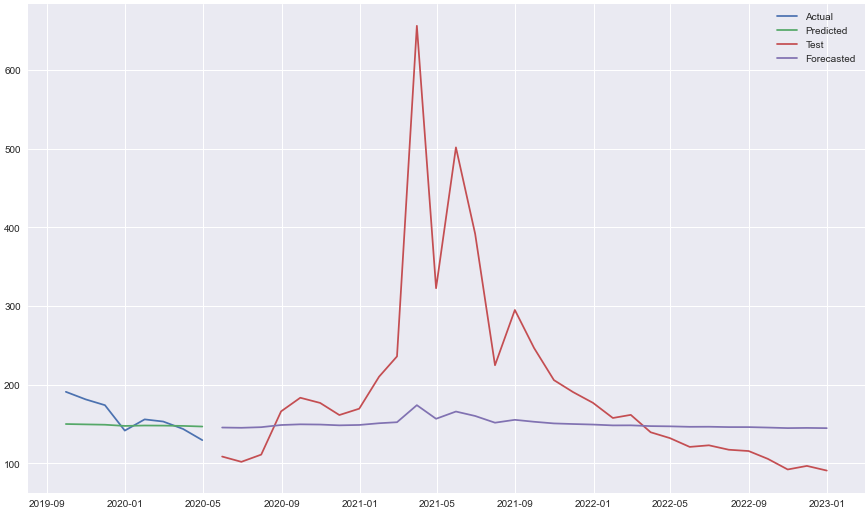
Gambar 4.36. Hasil prediksi LSTM Bank IBK data latih 60% 300 *epoch*

Pada Gambar 4.36 dapat dilihat bahwa hasil uji menggunakan data latih (warna hijau) hampir sama dengan penggunaan 200 *epoch*. Akan tetapi pada hasil uji menggunakan data uji (warna ungu) memperlihatkan hasil yang lebih akurat pada 300 *epoch*.



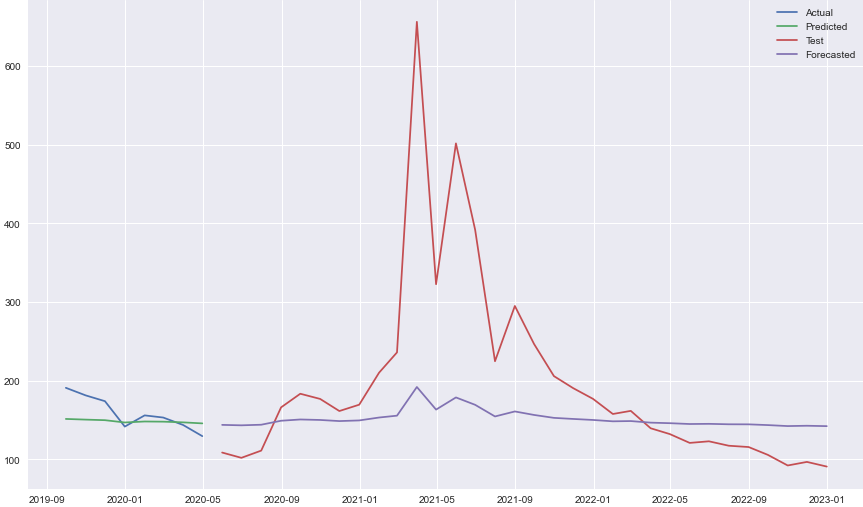
Gambar 4.37. Hasil prediksi Bank IBK data latih 20% 50 *epoch*

Pada Gambar 4.37 terlihat baik hasil uji data latih (warna hijau) dan hasil uji data uji (warna ungu) menunjukan ketidakakuratan yang jauh dari kondisi nyata.



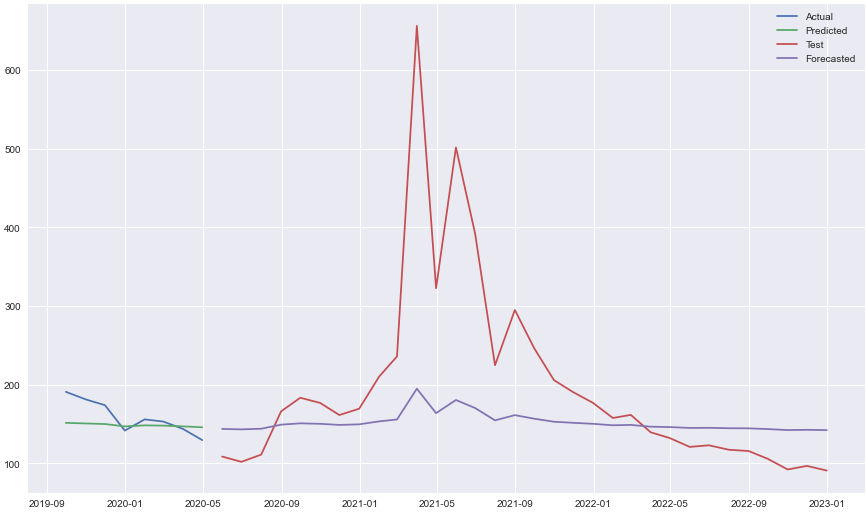
Gambar 4.38. Hasil prediksi LSTM Bank IBK data latih 20% 100 *epoch*

Pada Gambar 4.38 terlihat hasil pengujian sedikit lebih baik dari 50 *epoch*. Namun hasil yang diperoleh belum cukup untuk dibandingkan dengan keadaan nyata sehingga perlu dilakukan percobaan berikutnya.



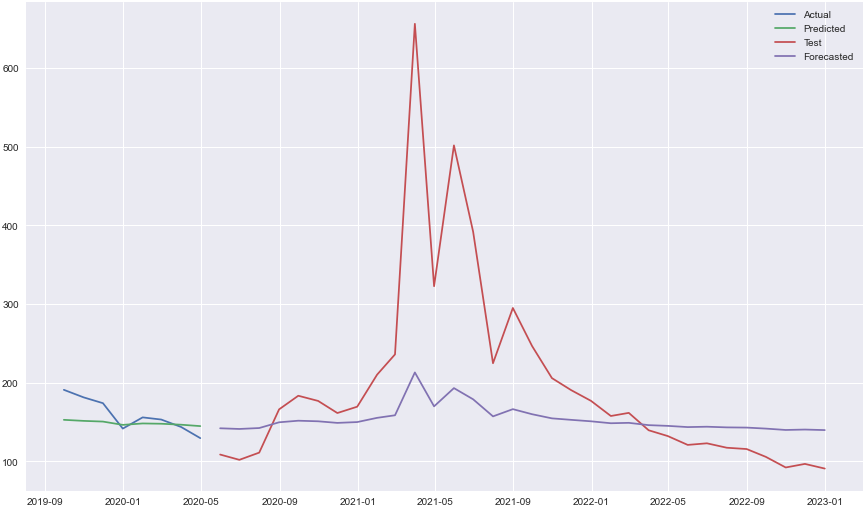
Gambar 4.39. Hasil prediksi LSTM Bank IBK data latih 20% 150 *epoch*

Pada Gambar 4.39 terlihat hasil pengujian masih sedikit lebih baik dari 100 *epoch*. Namun hasil yang diperoleh belum cukup untuk dibandingkan dengan keadaan nyata sehingga perlu dilakukan percobaan berikutnya.



Gambar 4.40. Hasil prediksi LSTM Bank IBK data latih 20% 200 *epoch*

Pada Gambar 4.40 terlihat hasil pengujian masih sedikit lebih baik dari 100 *epoch*. Namun hasil yang diperoleh belum cukup untuk dibandingkan dengan keadaan nyata sehingga perlu dilakukan percobaan berikutnya.



Gambar 4.41. Hasil prediksi LSTM Bank IBK data latih 20% 300 *epoch*

Pada Gambar 4.41 terlihat hasil masih sedikit lebih baik walaupun sudah menggunakan 300 *epoch* dan masih belum cukup untuk digunakan sebagai model yang relevan karena kurangnya data yang dilatih.

## 4.6. Evaluasi Model

Setelah proses-proses pemodelan telah dilakukan dengan algoritma ARFIMA dan LSTM, tahapan terakhir adalah mengevaluasi hasil dari pengujian masing-masing algoritma baik dataset penutupan BRI dan Bank IBK. Pada penelitian ini digunakan SMAPE untuk mengukur tingkat kesalahan dari setiap model yang dibuat. Perhitungan SMAPE dapat menggunakan persamaan (31), sebagai contoh yang digunakan adalah hasil dari ARFIMA BRI(8,0.5,0) pembagian data latih 80%.

Tabel 4.31. Perbandingan hasil SMAPE prediksi harga penutupan saham

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Model | Data Latih | SMAPE | SMAPE(%) | Waktu(detik) |
| BRI ARFIMA(8,0.5,0) | 80% | 0.0557 | 5.57% | 1.2 |
| Bank IBK ARFIMA(4,0.5,0) | 80% | 0.2331 | 23.31% | 1.2 |
| BRI LSTM 50 *epoch* | 80% | 0.1119 | 11.19% | 1.7 |
| BRI LSTM 100 *epoch* | 80% | 0.1141 | 11.41% | 2.3 |
| BRI LSTM 150 *epoch* | 80% | 0.0161 | 1.61% | 2.9 |
| BRI LSTM 200 *epoch* | 80% | 0.0178 | 1.78% | 3.3 |
| BRI LSTM 300 *epoch* | 80% | 0.0174 | 1.74% | 4.4 |
| Bank IBK LSTM 50 *epoch* | 80% | 0.5670 | 5.67% | 1.7 |
| Bank IBK LSTM 100 *epoch* | 80% | 0.5409 | 54.09% | 2.3 |
| Bank IBK LSTM 150 *epoch* | 80% | 0.4695 | 46.95% | 2.9 |
| Bank IBK LSTM 200 *epoch* | 80% | 0.3715 | 37.15% | 3.3 |
| Bank IBK LSTM 300 *epoch* | 80% | 0.0222 | 2.22% | 4.4 |
| BRI ARFIMA(1,0.5,1) | 60% | 0.2424 | 24.24% | 1.2 |
| Bank IBK ARFIMA(4,0.5,0) | 60% | 1.9438 | 194.38% | 1.2 |
| BRI LSTM 50 *epoch* | 60% | 0.1339 | 13.39% | 1.7 |
| BRI LSTM 100 *epoch* | 60% | 0.1020 | 10.20% | 2.3 |
| BRI LSTM 150 *epoch* | 60% | 0.0723 | 7.23% | 2.8 |
| BRI LSTM 200 *epoch* | 60% | 0.0260 | 2.60% | 3.3 |
| BRI LSTM 300 *epoch* | 60% | 0.0238 | 2.38% | 4.4 |
| Bank IBK LSTM 50 *epoch* | 60% | 0.4769 | 47.69% | 1.6 |
| Bank IBK LSTM 100 *epoch* | 60% | 0.4554 | 45.54% | 3.1 |
| Bank IBK LSTM 150 *epoch* | 60% | 0.3612 | 36.12% | 3.2 |
| Bank IBK LSTM 200 *epoch* | 60% | 0.2933 | 29.33% | 3.4 |
| Bank IBK LSTM 300 *epoch* | 60% | 0.0889 | 8.89% | 5.3 |
| BRI ARFIMA(1,0.5,1) | 20% | 0.3731 | 37.31% | 0.9 |
| Bank IBK ARFIMA(4,0.5,0) | 20% | 0.3902 | 39.02% | 0.9 |
| BRI LSTM 50 *epoch* | 20% | 0.1326 | 13.26% | 1.8 |
| BRI LSTM 100 epoch | 20% | 0.1144 | 11.44% | 2.2 |
| BRI LSTM 150 epoch | 20% | 0.0919 | 9.19% | 2.8 |
| BRI LSTM 200 epoch | 20% | 0.0551 | 5.51% | 3.4 |
| BRI LSTM 300 epoch | 20% | 0.0510 | 5.10% | 4.4 |
| Bank IBK LSTM 50 epoch | 20% | 0.3583 | 35.83% | 1.5 |
| Bank IBK LSTM 100 epoch | 20% | 0.3435 | 34.35% | 2.7 |
| Bank IBK LSTM 150 epoch | 20% | 0.3274 | 32.74% | 2.9 |
| Bank IBK LSTM 200 epoch | 20% | 0.3260 | 32.60% | 3.2 |
| Bank IBK LSTM 300 epoch | 20% | 0.3113 | 31.13% | 4.3 |

Berdasarkan Tabel 4.31 dapat dilihat setiap model memiliki nilai SMAPE yang berbeda-beda. Dapat dilihat juga berdasarkan hasil perhitungan SMAPE tersebut nilai SMAPE model ARFIMA lebih besar daripada model LSTM baik data penutupan BRI maupun Bank IBK. Hasil pengujian yang paling optimal adalah BRI LSTM 150 *epoch* pembagian data latih 80% dengan SMAPE sebesar 1,61% dan Bank IBK LSTM 300 *epoch* pembagian data latih 80% dengan SMAPE sebesar 2,22%.

# BAB V

**PENUTUP**

## 5.1. Kesimpulan

Berdasarkan hasil penelitian yang menggunakan data saham dari BRI dan Bank IBK Indonesia dapat ditarik kesimpulan bahwa:

1. Hasil terbaik model ARFIMA untuk BRI adalah model ARFIMA(8,0.5,0) pada pembagian data latih 80% dengan hasil SMAPE sebesar 0,0557 atau sekitar 5,57%. Hasil terbaik model ARFIMA untuk Bank IBK Indonesia adalah model ARFIMA(8,0.5,0) pada pembagian data latih 80% dengan hasil SMAPE sebesar 0,2331 atau sekitar 23,31%.
2. Hasil terbaik model LSTM untuk BRI adalah model LSTM menggunakan 150 *epoch* pada pembagian data latih 80% dengan hasil SMAPE sebesar 0,0161 atau sekitar 1,61%. Hasil terbaik model LSTM untuk Bank IBK Indonesia adalah model LSTM menggunakan 300 *epoch* pada pembagian data latih 80% dengan hasil SMAPE sebesar 0,0222 atau sekitar 2,22%.
3. Algoritma LSTM dengan pembagian data latih 80% menggunakan 150 *epoch* untuk dataset Bank BRI dan Algoritma LSTM dengan pembagian data 80% menggunakan 300 *epoch* untuk dataset Bank IBK Indonesia menghasilkan hasil yang terbaik pada data saham BRI maupun data saham Bank IBK Indonesia. Hal ini mengindikasikan algoritma LSTM merupakan model terbaik pada penelitian menggunakan data saham tersebut.

## 5.2. Saran

Ada beberapa saran yang dapat diterapkan jika ada penelitian dengan topik terkait prediksi saham yang menggunakan algoritma serupa, yaitu:

1. Prediksi data *time series* saham menggunakan algoritma ARFIMA sebaiknya untuk jangka pendek. Meskipun data latihnya sudah mencukupi ARFIMA masih tidak dapat memprediksi 3 sampai 5 deret waktu di masa depan dengan akurat walaupun pada pertumbuhan awalnya sudah cukup akurat.
2. Menggunakan variasi kombinasi parameter seperti menambahkan unit LSTM dan menambah *batch size* menyesuaikan bentuk data yang digunakan untuk menentukan model optimal yang dapat menghasilkan harga prediksi saham dan nilai SMAPE yang lebih rendah dan akurat dengan menggunakan LSTM.

# DAFTAR PUSTAKA

Agrawal, T. (2021) *Hyperparameter Optimization in Machine Learning: Make Your Machine Learning and Deep Learning Models More Efficient*, *Hyperparameter Optimization in Machine Learning*. Bangalore: Apress. doi:https://doi.org/10.1007/978-1-4842-6579-6 ISBN-13.

Agusta, A. (2021) ‘Prediksi Pergerakan Harga Saham Pada Sektor Farmasi Menggunakan Algoritma Long Short-Term Memory’, *Universitas Pembangunan Nasional Veteran Jakarta*, D(2), pp. 1–58.

Akbardipura, F. (2021) ‘Perbandingan Performa Peramalan Harga Saham 5 Perusahaan Pada Indeks LQ45 Menggunakan Autoregressive Integrated Moving Average Dan Long Short – Term Memory’, *Universitas Pembangunan Nasional Veteran Jakarta*, pp. 1–60.

Alhajj, R. and Rokne, J. (2018) *Encyclopedia of Social Network Analysis and Mining*, *Cooperative and Graph Signal Processing: Principles and Applications*. New York: Springer. doi:10.1016/B978-0-12-813677-5.00025-0.

Amaratunga, T. (2021) *Deep Learning on Windows*, *Deep Learning on Windows*. Nugegoda: Apress. doi:10.1007/978-1-4842-6431-7.

Ambach, D. and Ambach, O. (2018) ‘Forecasting the Oil Price with a Periodic Regression ARFIMA-GARCH Process’, *Proceedings of the 2018 IEEE 2nd International Conference on Data Stream Mining and Processing, DSMP 2018*, pp. 212–217. doi:10.1109/DSMP.2018.8478447.

*Bank IBK Indonesia - Wikipedia bahasa Indonesia, ensiklopedia bebas* (2022). Available at: https://id.wikipedia.org/wiki/Bank\_IBK\_Indonesia (Accessed: 23 January 2023).

Borges, T.A. and Neves, R. (2021) *Financial Data Resampling for Machine Learning Based Trading Application to Cryptocurrency Markets*. Edited by J. Kacprzyk. Warsaw: SpringerBriefs in Applied Sciences and Technology. Available at: https://doi.org/10.1007/978-3-030-68379-5%0A©.

Chandriah, K.K. and Naraganahalli, R. V. (2021) ‘RNN / LSTM with modified Adam optimizer in deep learning approach for automobile spare parts demand forecasting’, *Multimedia Tools and Applications*, 80(17), pp. 26145–26159. doi:10.1007/s11042-021-10913-0.

Chen, S. and Zhou, C. (2021) ‘Stock Prediction Based on Genetic Algorithm Feature Selection and Long Short-Term Memory Neural Network’, *IEEE Access*, 9, pp. 9066–9072. doi:10.1109/ACCESS.2020.3047109.

Chicco, D., Warrens, M.J. and Jurman, G. (2021) ‘The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation’, *PeerJ Computer Science*, 7, pp. 1–24. doi:10.7717/PEERJ-CS.623.

Chiu, S. and Tavella, D. (2020) ‘Introduction to Data Mining’, *Data Mining and Market Intelligence for Optimal Marketing Returns*, pp. 151–206. doi:10.4324/9780080878096-12.

Fauzi, D.A. (2020) ‘FAKTOR PENENTU PENGAMBILAN KEPUTUSAN INVESTASI SAHAM’, *Universitas Islam Indonesia*, 2507(1), pp. 1–9. Available at: https://doi.org/10.1016/j.solener.2019.02.027%0Ahttps://www.golder.com/insights/block-caving-a-viable-alternative/%0A???

Fitriadini, A., Pramiyati, T. and Pangaribuan, A.B. (2020) ‘Penerapan Backpropagation Neural Network Dalam Prediksi Harga Saham’, *Seminar Nasional Mahasiswa Ilmu Komputer dan Aplikasinya (SENAMIKA)*, pp. 561–573.

Gortsos, C. V. (2020) *European Central Banking Law*. Athens: Palgrave Macmillan. doi:https://doi.org/10.1007/978-3-030-34564-8.

Gu, Y. *et al.* (2020) ‘Prediction of stock performance using deep neural networks’, *Applied Sciences (Switzerland)*, 10(22), pp. 1–20. doi:10.3390/app10228142.

Heng, L., Longfu, Z. and Kaiyou, Y. (2020) ‘Statistic analysis of safety accidents in filling stations based on big data’, *ACM International Conference Proceeding Series* [Preprint]. doi:10.1145/3414274.3414280.

Hu, Z., Zhao, Y. and Khushi, M. (2021) ‘A survey of forex and stock price prediction using deep learning’, *Applied System Innovation*, 4(1), pp. 1–30. doi:10.3390/ASI4010009.

*IDX Data Services Portal* (2022). Available at: https://data.idx.co.id/ (Accessed: 24 January 2023).

Ikeuchi, K. (2021) *Computer Vision*. Redmond: Springer.

*Info Perusahaan - Bank BRI | Melayani Dengan Setulus Hati* (2022). Available at: https://bri.co.id/info-perusahaan (Accessed: 23 January 2023).

Jing, N., Wu, Z. and Wang, H. (2021) ‘A hybrid model integrating deep learning with investor sentiment analysis for stock price prediction’, *Expert Systems with Applications*, 178(May 2020), p. 115019. doi:10.1016/j.eswa.2021.115019.

K, A. (2020) *Dasar Investasi Saham*, *AK Pedia*. Jakarta: AK Pedia.

Kartikasari, P. (2020) ‘Prediksi Harga Saham PT. Bank Negara Indonesia dengan Menggunakan Model Autoregressive Fractional Integrated Moving Average (ARFIMA )’, *Jurnal Statistika*, 8(1), pp. 1–7.

Ketkar, N. and Moolayil, J. (2021) *Deep Learning with Python: Learn Best Practices of Deep Learning Models with PyTorch, 2nd Edition*, *First*.

Kondo Lembang, F., Sinay, L.J. and Irfanullah, A. (2021) ‘ARFIMA Modelling for Tectonic Earthquakes in The Maluku Region’, *Indonesian Journal of Statistics and Its Applications*, 5(1), pp. 39–49. doi:10.29244/ijsa.v5i1p39-49.

Lu, W. *et al.* (2021) ‘A CNN-BiLSTM-AM method for stock price prediction’, *Neural Computing and Applications*, 33(10), pp. 4741–4753. doi:10.1007/s00521-020-05532-z.

Manaswi, N.K. (2018) ‘Deep Learning with Applications Using Python’, *Deep Learning with Applications Using Python*, 1, pp. 115–126. doi:10.1007/978-1-4842-3516-4.

Maruddani, D.A.I. and Astuti, T.D. (2021) ‘Risiko dan Strategi Investasi Saham Second Liner dengan Global Minimum Variance Portfolio’, *Jurnal Riset Akuntansi Mercu Buana*, 7(1), pp. 11–19. Available at: http://ejurnal.mercubuana-yogya.ac.id/index.php/akuntansi/article/view/1559.

Mustafa, A. (2020) ‘ANALISIS KEPUTUSAN INVESTASI SAHAM DENGAN DENGAN PENDEKATAN PRICE EARNING RATIO ( PER ) PADA PT UNILEVER INDONESIA Tbk .’, *Universitas Negeri Makassar*, pp. 1–13.

Nisbet, R., Miner, G. and Yale, K. (2020) *HANDBOOK OF STATISTICAL ANALYSIS AND DATA MINING APPLICATIONS*. Edited by S. Ikeda. London: Elsevier Ltd. Available at: https://www.sciencedirect.com/book/9780124166325/handbook-of-statistical-analysis-and-data-mining-applications?via=ihub=.

Nuzula Agustin, I. and Lysion, F. (2021) ‘Isnaini Nuzula Agustin 1 Fiona Lysion 2’, *Universitas Internasional Batam*, 1(1). Available at: https://journal.uib.ac.id/index.php/combines.

Plevris, V. *et al.* (2022) ‘Investigation of performance metrics in regression analysis and machine learning-based prediction models’, *The 8th European Congress on Computational Methods in Applied Sciences and Engineering, 5–9 June 2022, Oslo, Norway* [Preprint], (June). Available at: https://www.eccomas2022.org/admin/files/fileabstract/a2088.pdf.

Rezaei, H., Faaljou, H. and Mansourfar, G. (2021) ‘Stock price prediction using deep learning and frequency decomposition’, *Expert Systems with Applications*, 169, p. 114332. doi:10.1016/j.eswa.2020.114332.

Saleh, M., Grivel, E. and Omar, S.M. (2018) ‘Jeffrey’s divergence between ARFIMA processes’, *Digital Signal Processing: A Review Journal*, 82, pp. 175–186. doi:10.1016/j.dsp.2018.06.013.

Saud, A.S. and Shakya, S. (2020) ‘Analysis of look back period for stock price prediction with RNN variants: A case study on banking sector of NEPSE’, *Procedia Computer Science*, 167(2019), pp. 788–798. doi:10.1016/j.procs.2020.03.419.

Shalalfeh, L., Bogdan, P. and Jonckheere, E. (2019) ‘Modeling of PMU Data Using ARFIMA Models’, *Clemson University Power Systems Conference, PSC 2018*, pp. 1–6. doi:10.1109/PSC.2018.8664019.

Sharma, R.R. *et al.* (2021) ‘EVDHM-ARIMA-Based Time Series Forecasting Model and Its Application for COVID-19 Cases’, *IEEE Transactions on Instrumentation and Measurement*, 70(July). doi:10.1109/TIM.2020.3041833.

Swanjaya, D. and Putra Pamungkas, D. (2021) ‘Analisa Hasil Prediksi Metode Least Square menggunakan Korelasi dan MAPE pada Toko PS’, *Generation Journal*, 5(1), pp. 11–18. doi:10.29407/gj.v5i1.15440.

Ta, V.D., Liu, C.M. and Tadesse, D.A. (2020) ‘Portfolio Optimization-Based Stock Prediction Using Long-Short Term Memory Network in Quantitative Trading Van-Dai’, *Applied Sciences (Switzerland)*, 10(2).

Trierweiler Ribeiro, G. *et al.* (2021) ‘Novel hybrid model based on echo state neural network applied to the prediction of stock price return volatility’, *Expert Systems with Applications*, 184(August 2020). doi:10.1016/j.eswa.2021.115490.

Unpingco, J. (2021) *Python Programming for Data Analysis*, *Python Programming for Data Analysis*. San Diego: Springer. doi:10.1007/978-3-030-68952-0.

Zhang, D. and Lou, S. (2021) ‘The application research of neural network and BP algorithm in stock price pattern classification and prediction’, *Future Generation Computer Systems*, 115, pp. 872–879. doi:10.1016/j.future.2020.10.009.

Zhou, H. (2020) *Learn Data Mining Through Excel*, *Learn Data Mining Through Excel*. Apress. doi:10.1007/978-1-4842-5982-5.