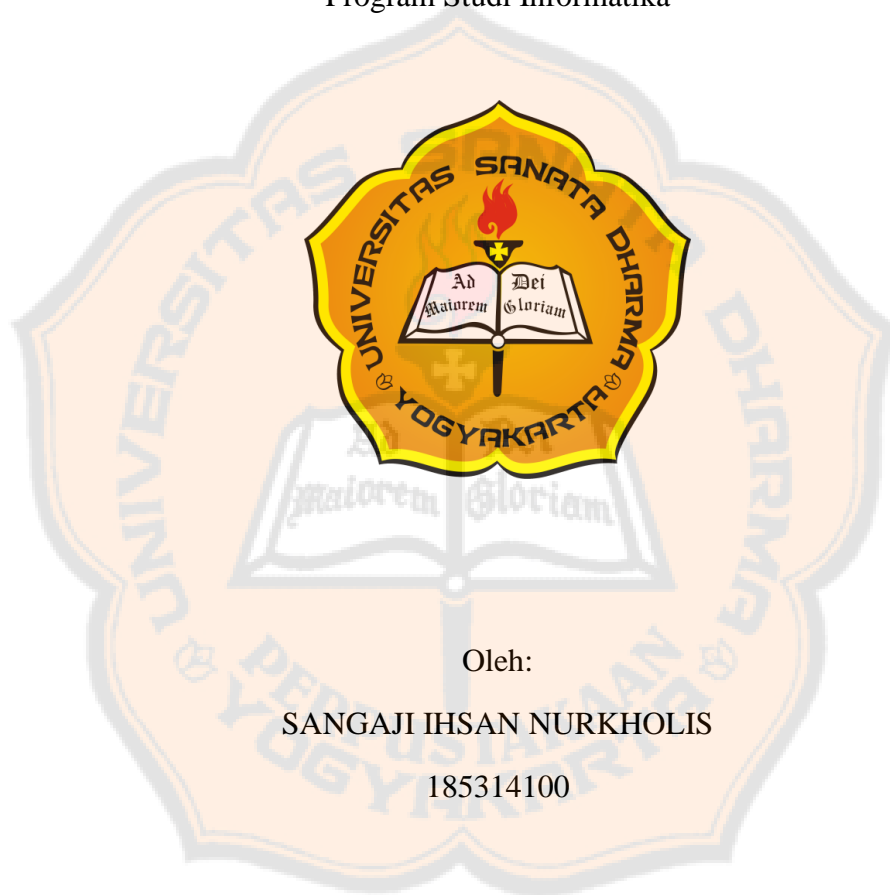


**ALGORITMA MULTIPLE LINEAR REGRESSION UNTUK PREDIKSI
HARGA PENUTUPAN SAHAM PT BANK CENTRAL ASIA**

SKRIPSI

Diajukan Untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer
Program Studi Informatika



Oleh:

SANGAJI IHSAN NURKHOLIS

185314100

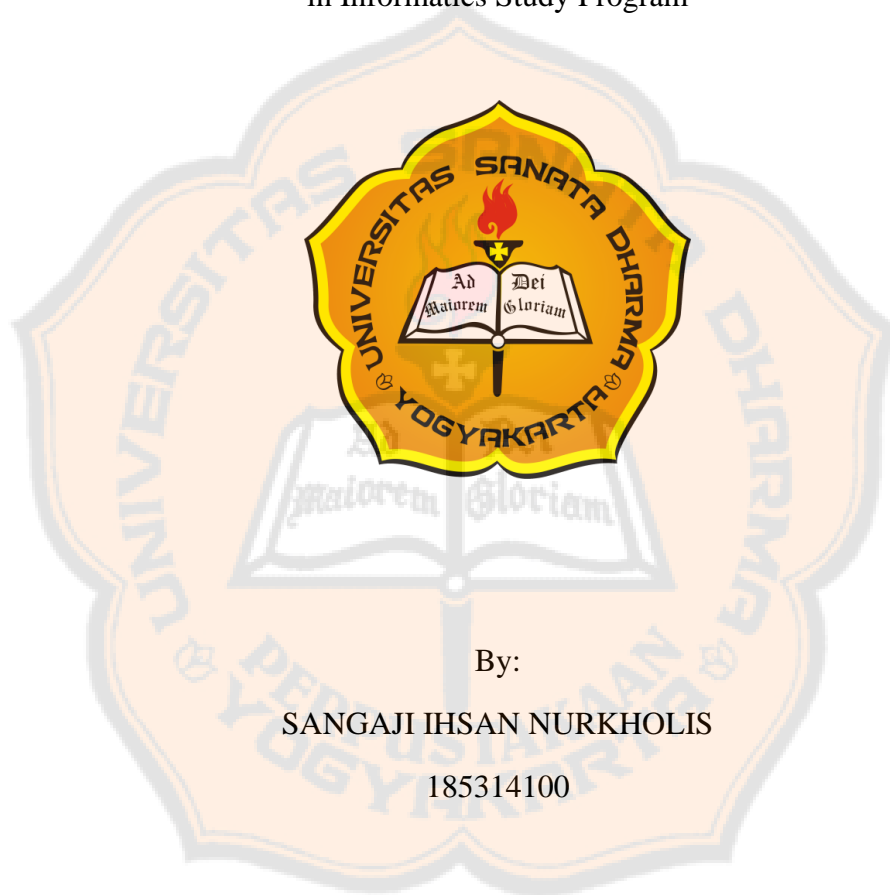
**PROGRAM STUDI INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS SANATA DHARMA
YOGYAKARTA**

2022

**MULTIPLE LINEAR REGRESSION ALGORITHM FOR PREDICTING
THE CLOSING PRICE OF PT BANK CENTRAL ASIA'S SHARE**

THESIS

Present of Partial Fulfillment of The Requirements
to Obtain *Sarjana Komputer* Degree
in Informatics Study Program



By:

SANGAJI IHSAN NURKHOLIS

185314100

**INFORMATICS STUDY PROGRAM
FACULTY OF SCIENCE AND TECHNOLOGY
SANATA DHARMA UNIVERSITY
YOGYAKARTA**

2022

HALAMAN PERSETUJUAN

SKRIPSI

**“ALGORITMA MULTIPLE LINEAR REGRESSION UNTUK PREDIKSI
HARGA PENUTUPAN SAHAM PT BANK CENTRAL ASIA”**


Oleh:

SANGAJI IHSAN NURKHOLIS

185314100

Telah disetujui Oleh:

Dosen Pembimbing


Drs. Hari Suparwito, S.I. M.App.IT

Tanggal: 30 Januari 2023

HALAMAN PENGESAHAN
SKRIPSI

**“ALGORITMA MULTIPLE LINEAR REGRESSION UNTUK PREDIKSI
HARGA PENUTUPAN SAHAM PT BANK CENTRAL ASIA”**

Dipersiapkan dan ditulis oleh:

SANGAJI IHSAN NURKHOLIS

185314100

Telah dipertahankan di depan panitia dan penguji

Pada tanggal 17 Januari 2023

Dan dinyatakan telah memenuhi syarat.

Susunan Panitia Penguji

Nama Lengkap

Tanda Tangan

Ketua : Drs. Johannes Eka Priyatma, M.Sc., Ph.D.

Sekretaris : Dr. Ir. Anastasia Rita Widiarti

Anggota : Drs. Hari Suparwito, S.J., M.App.IT

Yogyakarta, 31 Januari 2023

Fakultas Sains dan Teknologi

Universitas Sanata Dharma



Dekan

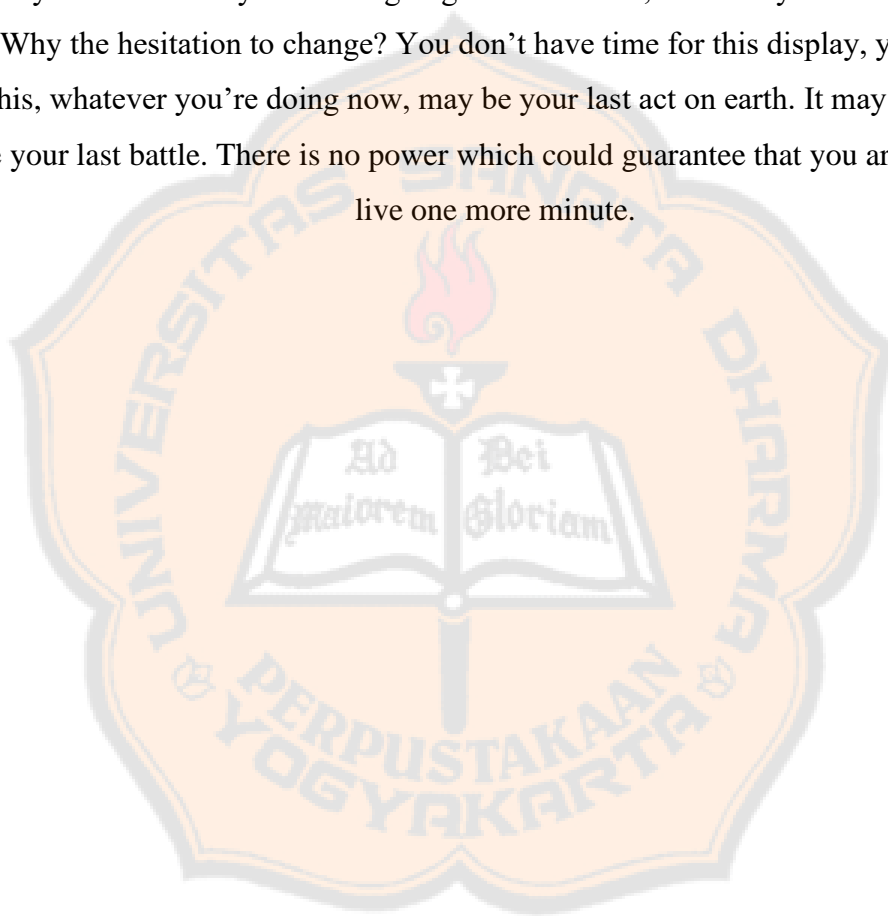
Drs. Haris Sriwindono, M.Kom., Ph.D.

HALAMAN PERSEMBAHAN

“The trouble is, you think you have time”

-Buddha-

There is one simple thing wrong with you – you think you have plenty of time
If you don't think your life is going to last forever, what are you waiting for?
Why the hesitation to change? You don't have time for this display, you fool.
This, whatever you're doing now, may be your last act on earth. It may very well
be your last battle. There is no power which could guarantee that you are going to
live one more minute.



PERNYATAAN KEASLIAN KARYA

Saya menyatakan dengan sesungguhnya bahwa tugas akhir yang saya tulis ini tidak memuat atau mengandung hasil karya orang lain, selain yang telah disebutkan dalam kutipan dan daftar pustaka sebagaimana layaknya sebuah karya ilmiah.

Yogyakarta, 27 Januari 2023

Penulis



Sangaji Ihsan Nurkholis



**LEMBAR PERNYATAAN PERSETUJUAN
PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN
AKADEMIS**

Yang bertanda tangan di bawah ini, saya mahasiswa Universitas Sanata Dharma:

Nama : Sangaji Ihsan Nurkholis

Nomor Mahasiswa : 185314100

Demi pengembangan ilmu pengetahuan, saya memberikan kepada Perpustakaan Universitas Sanata Dharma karya ilmiah saya yang berjudul:

**“ALGORITMA MULTIPLE LINEAR REGRESSION UNTUK PREDIKSI
HARGA PENUTUPAN SAHAM PT BANK CENTRAL ASIA”**

Dengan demikian saya memberikan kepada Perpustakaan Universitas Sanata Dharma hak untuk menyimpan, mengalihkan dalam bentuk media lain, mengelolanya dalam bentuk pangkalan data, mendistribusikan secara terbatas, dan mempublikasikannya di Internet atau media lain untuk kepentingan akademis tanpa perlu meminta izin dari saya maupun memberikan royalti kepada saya selama tetap mencantumkan nama saya sebagai penulis.

Atas kemajuan teknologi informasi, saya tidak keberatan jika nama, tanda tangan, gambar atau *image* yang ada di dalam karya ilmiah saya terindeks oleh mesin pencari (*search engine*), misalnya *google*.

Demikian pernyataan ini yang saya buat dengan sebenarnya.

Dibuat di Yogyakarta

Pada tanggal: 27 Januari 2023

Yang menyatakan



Sangaji Ihsan Nurkholis

ABSTRAK

Saham adalah salah satu bentuk investasi yang memperdagangkan surat berharga sebagai bukti sebagian kepemilikan seseorang atas suatu perusahaan. Harga saham yang bersifat fluktuatif, menjadi masalah bagi para investor dalam menentukan kapan saat yang tepat untuk membeli/menjual saham. Penelitian ini bertujuan untuk memprediksi harga penutupan saham PT Bank Central Asia Tbk dengan menerapkan metode *Multiple Linear Regression*. Prediksi harga saham akan membantu mengurangi resiko dan memberikan keuntungan yang lebih maksimal bagi para investor. *Technical indicators* berupa *Exponential Moving Average* dan *Relative Strength Index* akan diterapkan pada proses prediksi sebagai variabel independen. Data saham yang digunakan adalah data historis harian yang diambil dari website finance.yahoo.com selama 5 tahun terakhir. Pengujian pada penelitian ini berfokus pada model dengan nilai periode *technical indicators* yang menghasilkan nilai *residuals (error)* terkecil. Hasil dari penelitian ini mendapatkan model terbaik dengan nilai *Mean Absolute Percentage Error (MAPE)* yaitu sebesar 0.377% dengan nilai *R-Square* sebesar 0.992, nilai *Mean Absolute Error (MAE)* sebesar 29.619, dan *Root Mean Square Error (RMSE)* sebesar 45.774.

Kata Kunci: prediksi saham, technical indicators, multiple linear regression.

ABSTRACT

Stocks are a form of investment that trades securities as proof of someone's partial ownership of a company. Fluctuating stock prices are a problem for investors in determining when is the right time to buy/sell shares. This study aims to predict the closing price of shares of PT Bank Central Asia Tbk by applying the Multiple Linear Regression method. Stock price predictions will help reduce risk and provide maximum profits for investors. Technical indicators in the form of Exponential Moving Average and Relative Strength Index will be applied to the prediction process as independent variables. The stock data used is historical daily data taken from the website finance.yahoo.com for the last 5 years. Tests in this study focused on models with technical indicators period values that produced the smallest residuals (errors). The results of this study obtained the best model with a Mean Absolute Percentage Error (MAPE) value of 0.377% with an R-Square value of 0.992, a Mean Absolute Error (MAE) value of 29.619, and a Root Mean Square Error (RMSE) of 45.774.

Keyword: stock predictions, technical indicators, multiple linear regression.

KATA PENGANTAR

Puji dan syukur kepada Tuhan Yang Maha Esa, karena penulis menyelesaikan Tugas Akhir dengan judul “Algoritma Multiple Linear Regression Untuk Prediksi Harga Penutupan Saham PT Bank Central Asia” sebagai syarat memperoleh gelar sarjana komputer dari program studi Informatika Universitas Sanata Dharma Yogyakarta. Sehingga dalam kesempatan ini penulis ingin mengucapkan terimakasih kepada semua pihak yang telah hadir dalam memberikan dukungan. Ucapan terima kasih penulis sampaikan kepada:

1. Tuhan Yang Maha Esa yang selalu menyertai sehingga penulis dapat menyelesaikan tugas akhir.
2. Kedua orang tua serta keluarga yang selalu memberikan doa, semangat, dan dukungan untuk menyelesaikan tugas akhir.
3. Bapak Drs. Hari Suparwito, S.J., M.App.IT selaku dosen pembimbing yang selalu membimbing dan memberi saran selama penyusunan tugas akhir.
4. Bapak Ir.Drs. Haris Sriwindono, M.Kom., Ph.D. selaku dekan Fakultas Sains dan Teknologi.
5. Bapak Ir.Robertus Adi Nugroho, M.Eng. selaku Ketua Program Studi Informatika Universitas Sanata Dharma.
6. Bapak Bambang Soelistijanto, Ph.D. selaku dosen pembimbing akademik.
7. Seluruh jajaran dosen, sekretariat, dan laboran yang turut membantu penulis menyelesaikan tugas akhir.
8. Seseorang yang memberikan dorongan untuk menyelesaikan tugas akhir tepat waktu.
9. Anay, Aulia, Ayu, Banu, Bima, Diki, Recha, Rizky, Roberto, dan Salsa terimakasih atas bantuan, saran, dan diskusi selama penyusunan tugas akhir.
10. Seluruh teman-teman Informatika angkatan 2018 yang tidak bisa disebutkan satu-persatu yang telah memberikan dukungan serta kesan selama perkuliahan.

Penulis menyadari masih banyak kekurangan dan jauh dari kata sempurna, maka dari itu kritik dan saran yang bersifat membangun akan sangat diharapkan demi penyempurnaan Tugas Akhir ini. Akhir kata, penulis mengucapkan terima kasih, semoga Tugas Akhir ini dapat bermanfaat bagi pembaca.

Yogyakarta, 27 Januari 2023


Sangaji Ihsan Nurkholis



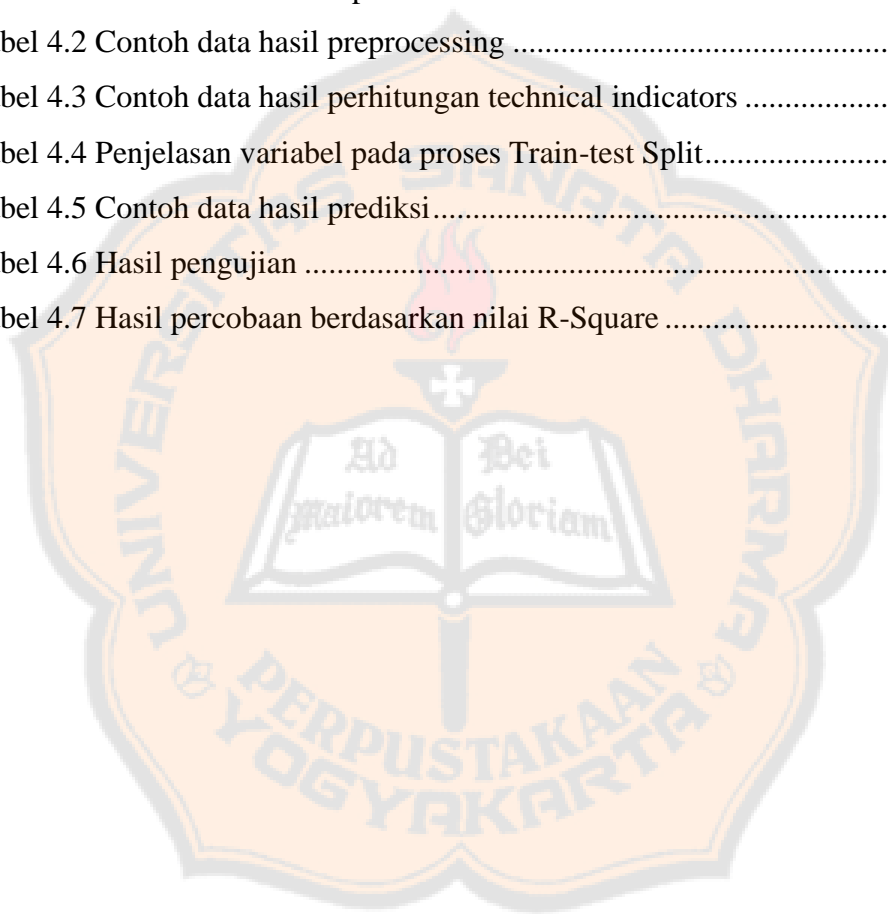
DAFTAR ISI

| | |
|--|------|
| HALAMAN PERSETUJUAN..... | iii |
| HALAMAN PENGESAHAN..... | iv |
| HALAMAN PERSEMBAHAN | v |
| PERNYATAAN KEASLIAN KARYA | vi |
| LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS | vii |
| ABSTRAK..... | viii |
| ABSTRACT..... | ix |
| KATA PENGANTAR | x |
| DAFTAR ISI..... | xii |
| DAFTAR TABEL..... | xiv |
| DAFTAR GAMBAR | xv |
| DAFTAR LAMPIRAN..... | xvi |
| BAB I PENDAHULUAN | 1 |
| 1.1 Latar Belakang | 1 |
| 1.2 Rumusan Masalah..... | 3 |
| 1.3 Tujuan | 4 |
| 1.4 Batasan Masalah..... | 4 |
| 1.5 Sistematika Penulisan | 4 |
| 1.6 Metodologi Penelitian | 5 |
| BAB II TINJAUAN PUSTAKA DAN LANDASAN TEORI | 7 |
| 2.1 Tinjauan Pustaka | 7 |
| 2.2 Landasan Teori..... | 10 |
| 2.2.1 Saham..... | 10 |
| 2.2.2 Technical Indicators | 11 |
| 2.2.2.1 Pengertian Technical Indicators..... | 11 |
| 2.2.2.2 Exponential Moving Average | 12 |
| 2.2.2.3 Relative Strength Index..... | 14 |
| 2.2.3 Multiple Linear Regression..... | 15 |

| | | |
|--------------------------------|---|----|
| 2.2.4 | Performance Evaluation | 17 |
| 2.2.4.1 | Koefisien Determinasi..... | 17 |
| 2.2.4.2 | Mean Absolute Error..... | 17 |
| 2.2.4.3 | Root Mean Square Error | 18 |
| 2.2.4.4 | Mean Absolute Percentage Error | 19 |
| BAB III METODE PENELITIAN..... | | 20 |
| 3.1 | Desain Alat Uji..... | 20 |
| 3.2 | Pengumpulan Data | 23 |
| 3.3 | Skenario Pengujian..... | 24 |
| 3.4 | Desain User Interface | 25 |
| BAB IV HASIL DAN ANALISA..... | | 29 |
| 4.1 | Hasil | 29 |
| 4.1.1 | Implementasi Program | 29 |
| 4.1.1.1 | Fungsi <code>__init__()</code> | 29 |
| 4.1.1.2 | Fungsi <code>check()</code> | 30 |
| 4.1.1.3 | Fungsi <code>forecast()</code> | 32 |
| 4.1.1.3.1 | Pre-processing | 32 |
| 4.1.1.3.2 | Perhitungan Technical Indicators..... | 34 |
| 4.1.1.3.3 | Train-test Split..... | 36 |
| 4.1.1.3.4 | Train Model Multiple Linear Regression..... | 37 |
| 4.1.1.3.5 | Prediction | 37 |
| 4.1.1.3.6 | Performance Evaluation | 38 |
| 4.1.1.3.7 | Update Hasil Prediksi Pada User Interface | 39 |
| 4.1.2 | Pengujian..... | 40 |
| 4.2 | Analisa..... | 41 |
| BAB V PENUTUP..... | | 45 |
| 5.1 | Kesimpulan | 45 |
| 5.2 | Saran..... | 45 |
| DAFTAR PUSTAKA | | 46 |
| LAMPIRAN | | 49 |

DAFTAR TABEL

| | |
|--|----|
| Tabel 2.1 Literature review | 9 |
| Tabel 3.1 Atribut data | 23 |
| Tabel 3.2 Contoh dataset saham..... | 24 |
| Tabel 3.3 Kombinasi periode pada pengujian..... | 25 |
| Tabel 4.1 Contoh dataset dari proses data retrieval | 32 |
| Tabel 4.2 Contoh data hasil preprocessing | 34 |
| Tabel 4.3 Contoh data hasil perhitungan technical indicators | 35 |
| Tabel 4.4 Penjelasan variabel pada proses Train-test Split..... | 36 |
| Tabel 4.5 Contoh data hasil prediksi..... | 38 |
| Tabel 4.6 Hasil pengujian | 40 |
| Tabel 4.7 Hasil percobaan berdasarkan nilai R-Square | 42 |



DAFTAR GAMBAR

| | |
|--|----|
| Gambar 2.1 Technical Indicators | 12 |
| Gambar 2.2 Exponential Moving Average (50)..... | 13 |
| Gambar 2.3 Relative Strength Index (9) | 15 |
| Gambar 3.1 Desain alat uji..... | 20 |
| Gambar 3.2 Desain User Interface | 25 |
| Gambar 3.3 Flowchart Penggunaan User Interface | 27 |
| Gambar 4.1 Tampilan user interface awal | 30 |
| Gambar 4.2 Hasil jumlah data kosong pada dataset. | 33 |
| Gambar 4.3 Jumlah data yang memiliki nilai volume 0 | 33 |
| Gambar 4.4 Jumlah data kosong dari perhitungan technical indicators..... | 35 |
| Gambar 4.5 Hasil nilai konstanta dan koefisien model MLR..... | 37 |
| Gambar 4.6 Hasil evaluasi | 39 |
| Gambar 4.7 Tampilan user interface setelah prediksi | 39 |
| Gambar 4.8 Hasil visualisasi percobaan EMA 50 dan RSI 5 | 43 |
| Gambar 4.9 Hasil visualisasi percobaan EMA 9 dan RSI 5 | 43 |

DAFTAR LAMPIRAN

| | |
|--|----|
| Lampiran 1 Source code fungsi <code>__init__()</code> | 49 |
| Lampiran 2 Source code fungsi <code>check()</code> | 51 |
| Lampiran 3 Source code untuk proses data retrieval | 52 |
| Lampiran 4 Contoh data dari proses data retrieval | 53 |
| Lampiran 5 Source code untuk proses Pre-processing | 53 |
| Lampiran 6 Contoh data dari proses Pre-processing..... | 54 |
| Lampiran 7 Source code untuk proses perhitungan technical indicators. | 55 |
| Lampiran 8 Contoh data dari proses perhitungan technical indicators. | 56 |
| Lampiran 9 Source code untuk proses Train-test Split. | 56 |
| Lampiran 10 Source code untuk proses pembentukan model MLR..... | 57 |
| Lampiran 11 Source code untuk proses prediction. | 57 |
| Lampiran 12 Contoh data hasil prediksi | 58 |
| Lampiran 13 Source code untuk proses performance evaluation. | 58 |
| Lampiran 14 Source code Update hasil prediksi pada User Interface. | 59 |
| Lampiran 15 Hasil percobaan periode EMA 9 dan periode RSI 5 | 60 |
| Lampiran 16 Hasil percobaan periode EMA 9 dan periode RSI 9 | 60 |
| Lampiran 17 Hasil percobaan periode EMA 9 dan periode RSI 14 | 61 |
| Lampiran 18 Hasil percobaan periode EMA 9 dan periode RSI 21 | 61 |
| Lampiran 19 Hasil percobaan periode EMA 9 dan periode RSI 25 | 62 |
| Lampiran 20 Hasil percobaan periode EMA 10 dan periode RSI 5 | 62 |
| Lampiran 21 Hasil percobaan periode EMA 10 dan periode RSI 9 | 63 |
| Lampiran 22 Hasil percobaan periode EMA 10 dan periode RSI 14 | 63 |
| Lampiran 23 Hasil percobaan periode EMA 10 dan periode RSI 21 | 64 |
| Lampiran 24 Hasil percobaan periode EMA 10 dan periode RSI 25 | 64 |
| Lampiran 25 Hasil percobaan periode EMA 21 dan periode RSI 5 | 65 |
| Lampiran 26 Hasil percobaan periode EMA 21 dan periode RSI 9 | 65 |
| Lampiran 27 Hasil percobaan periode EMA 21 dan periode RSI 14 | 66 |
| Lampiran 28 Hasil percobaan periode EMA 21 dan periode RSI 21 | 66 |
| Lampiran 29 Hasil percobaan periode EMA 21 dan periode RSI 25 | 67 |

| | |
|---|----|
| Lampiran 30 Hasil percobaan periode EMA 50 dan periode RSI 5 | 67 |
| Lampiran 31 Hasil percobaan periode EMA 50 dan periode RSI 9 | 68 |
| Lampiran 32 Hasil percobaan periode EMA 50 dan periode RSI 14 | 68 |
| Lampiran 33 Hasil percobaan periode EMA 50 dan periode RSI 21 | 69 |
| Lampiran 34 Hasil percobaan periode EMA 50 dan periode RSI 25 | 69 |



BAB I

PENDAHULUAN

1.1 Latar Belakang

Saham adalah salah satu bentuk investasi yang memperdagangkan surat berharga sebagai bukti sebagian kepemilikan seseorang atas suatu perusahaan. Harga saham yang ada di pasar modal ditentukan oleh kinerja perusahaan yang bisa dilihat dari laporan keuangannya. Jika kinerja perusahaan baik maka harga saham juga akan meningkat (Oktavia & Nugraha, 2018). Semakin baik kinerja keuangan suatu perusahaan akan berdampak pada laba yang diperoleh perusahaan dan keuntungan yang didapat oleh investor, sehingga akan mempengaruhi peningkatan harga saham. Selain kinerja perusahaan, ada faktor fundamental internal dan eksternal yang mempengaruhi harga saham. Fundamental yang terdiri dari *intern* meliputi keadaan para emiten seperti, pembagian deviden, perubahan strategi dalam rapat umum pemegang saham, akan menjadi informasi penting bagi para investor di pasar modal. Sedangkan faktor *ekstern* meliputi kebijakan pemerintah, perubahan nilai tukar rupiah terhadap dollar (kurs), keadaan inflasi, penetapan suku bunga oleh Bank Indonesia (Putri & Tjahhjono, 2011). Pergerakan harga saham yang tidak menentu menjadi masalah bagi para investor dalam menentukan kapan saat yang tepat untuk membeli/menjual saham. Kondisi inilah yang menjadi alasan diperlukannya sebuah sistem atau metode untuk memprediksi harga saham di bursa efek (*stock exchange*).

Prediksi atau *forecasting* merupakan sebuah proses untuk mengetahui data pada masa depan. Saat ini prediksi banyak dibutuhkan di banyak bidang termasuk saham. Sebelum era *Artificial Intelligence* (AI) dan *Machine Learning*, prediksi dilakukan secara manual oleh ahli statistik dengan membuat grafik dan menggunakan metode matematika untuk mengamati tren (Gururaj dkk, 2019). Dengan memprediksi harga saham, investor dapat menganalisis dan mengambil keputusan pembelian/

penjualan yang sesuai. Sehingga dapat meminimalkan resiko dan memberikan keuntungan yang lebih maksimal bagi investor.

Beberapa penelitian tentang prediksi harga telah dilakukan menggunakan berbagai metode, antara lain penelitian yang dilakukan untuk memprediksi harga penutupan saham dengan menggunakan metode *Artificial Neural Networks* (ANN) dan *Random Forest* (RF). Algoritma ANN mendapatkan nilai RMSE sebesar 0.42, MAPE sebesar 0.77% dan MBE sebesar -0.0156. Sedangkan algoritma RF mendapatkan nilai RMSE sebesar 0.43, MAPE sebesar 0.8% dan MBE sebesar -0.0155 (Vijh dkk, 2020). Penelitian lain yang dilakukan untuk memprediksi harga saham menggunakan metode *Multiple Linear Regression* dan *Moving Average* menghasilkan nilai MSE sebesar 15087.465, RMSE sebesar 122.831, dan MAPE sebesar 3.255 (Izzah & Widyastuti, 2017). Penelitian yang dilakukan dengan menggabungkan metode *Decision Tree* dan *Artificial Neural Networks* untuk memprediksi harga saham mendapatkan hasil akurasi sebesar 77% (Tsai & Wang, 2009).

Berdasarkan penjelasan dan beberapa penelitian yang telah dilakukan, tujuan dari penelitian ini adalah untuk menerapkan metode *Multiple Linear Regression* yang dapat memprediksi harga penutupan terakhir saham pada hari ini. Dalam melakukan prediksi harga saham dibutuhkan data secara berkala dengan satuan waktu yang dapat digolongkan sebagai *data time series* (Izzah & Widyastuti, 2017). Salah satu metode yang dapat digunakan untuk memprediksi data berbentuk *time series* adalah metode *Linear Regression*. Pemilihan metode *Linear Regression* sebagai metode prediksi pada penelitian ini didasari karena *Linear Regression* dapat melakukan generalisasi dan ekstraksi dari pola data tertentu, mampu mengakuisisi pengetahuan meskipun tidak ada kepastian, dan mampu melakukan perhitungan secara paralel sehingga proses lebih singkat (Gaol dkk, 2019). *Data time series* tidak langsung digunakan untuk memprediksi, melainkan dibutuhkan *technical indicators* sebagai variabel independen (Agrawal dkk, 2019).

Technical indicators adalah nilai yang dihitung untuk menggambarkan pergerakan harga historis sekuritas seperti saham yang berguna untuk memprediksi harga aset di masa depan karena dapat diintegrasikan ke dalam sistem perdagangan otomatis (Agrawal dkk, 2019). Beberapa indikator umum yang sering digunakan dalam memprediksi harga antara lain: *Moving Averages* (SMA, EMA, MACD), *Relative Strength Index* (RSI), *Bollinger Bands* (BBANDS) dan lainnya. Pada penelitian ini *technical indicators* EMA dan RSI akan digunakan sebagai variabel independen pada model *Linear Regression*. *Exponential Moving Average* adalah indikator yang membantu memperlancar tren harga dan digunakan untuk membantu mengidentifikasi pembalikan harga tren yang muncul dengan memberikan bobot pada rata-rata perubahan harga (Vargas dkk, 2018). Sedangkan *Relative Strength Index* merupakan *technical indicators* yang berperan sebagai parameter momentum, yakni mengukur pergerakan harga dengan membandingkan keuntungan rata-rata periode naik dengan kerugian rata-rata periode turun selama jangka waktu tertentu (Vergas dkk, 2018). Kedua *technical indicators* ini sama-sama memiliki nilai input berupa periode waktu, nilai ini akan sangat berpengaruh pada hasil akhir prediksi. Nilai periode waktu akan menjadi fokus utama dalam pengujian yang akan dilakukan pada penelitian ini untuk mendapatkan model dengan performa terbaik.

1.2 Rumusan Masalah

Pergerakan harga saham yang bersifat fluktuatif dan sulit diprediksi menjadi masalah utama bagi para investor. Permasalahan tersebut menjadi dasar penelitian ini untuk menjawab bagaimana cara memprediksi harga saham dengan pendekatan *machine learning*.

1.3 Tujuan

Berdasarkan rumusan masalah yang ada, terdapat beberapa tujuan yang ingin dicapai sebagai berikut:

1. Menerapkan metode *Multiple Linear Regression* yang dapat memprediksi harga penutupan saham dengan bantuan *technical indicators* sebagai atribut independen.
2. Menghitung hasil performa dari penerapan model *Multiple Linear Regression*.
3. Melakukan percobaan untuk menentukan model dengan performa terbaik dengan mengubah nilai periode pada *Exponential Moving Average* (EMA) dan *Relative Strength Index* (RSI).

1.4 Batasan Masalah

Agar pembahasan dalam penelitian ini tidak terlalu luas, maka dalam penelitian ini akan dibatasi oleh beberapa hal berikut:

1. Data yang digunakan pada penelitian ini adalah data historis harian dengan jangka waktu selama 5 tahun terakhir.
2. Data yang diprediksi adalah harga penutupan terakhir pada hari ini berdasarkan nilai *technical indicators* berupa *Exponential Moving Average* (EMA) dan *Relative Strength Index* (RSI).
3. Nilai periode yang digunakan untuk perhitungan *Exponential Moving Average* pada percobaan yang akan dilakukan adalah 9, 10, 21, dan 50 serta periode *Relative Strength Index* adalah 5, 9, 14, 21, dan 25.

1.5 Sistematika Penulisan

BAB I PENDAHULUAN

Bagian pendahuluan berisikan mengenai latar belakang, rumusan masalah, tujuan, batasan masalah, metodologi penelitian dan sistematika penulisan.

BAB II LANDASAN TEORI

Penelitian yang pernah dilakukan dan teori yang berkaitan dengan prediksi menggunakan metode *Multiple Linear Regression* yang akan digunakan pada penelitian ini akan dipaparkan pada bagian landasan teori.

BAB III METODE PENELITIAN

Metode penelitian akan menjelaskan mengenai rencana langkah-langkah yang akan dilakukan dalam penelitian yang meliputi desain alat uji, bagaimana cara mendapatkan data serta cara data tersebut diolah, proses pengujian yang akan dilakukan, dan desain *user interface* pada bagian akhir.

BAB IV HASIL DAN ANALISA

Hasil dan analisa akan menjelaskan hasil implementasi dan percobaan yang dilakukan dalam penelitian serta analisa dari hasil yang didapatkan.

BAB V PENUTUP

Kesimpulan atas penelitian yang telah dilakukan serta saran dari peneliti bagi penelitian berikutnya agar penelitian ini dapat berkembang lebih baik akan diberikan pada bagian penutup.

1.6 Metodologi Penelitian

Metodologi penelitian yang digunakan dalam penelitian ini adalah sebagai berikut:

1. Pengumpulan Data

Data yang digunakan adalah data historis harian selama 5 tahun terakhir yang diunduh langsung dari website finance.yahoo.com.

2. Implementasi Sistem

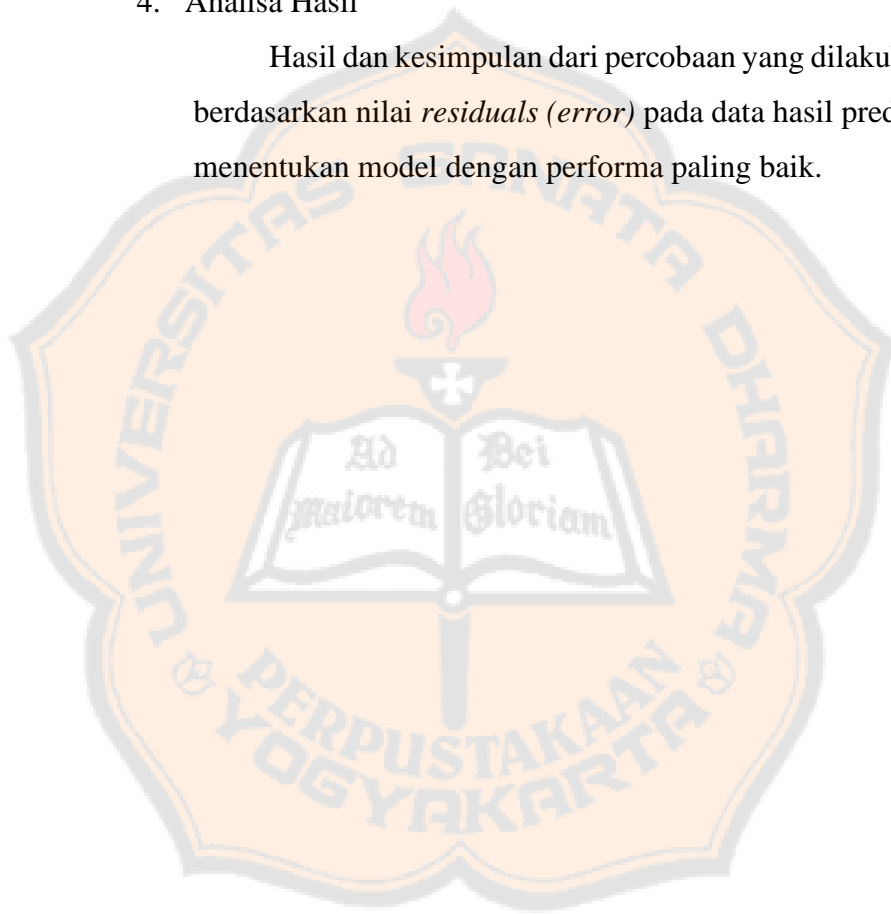
Pengimplementasian metode *Multiple Linear Regression* dengan menggunakan bahasa pemrograman Python 3.

3. Skenario Pengujian

Pengujian dilakukan untuk mendapatkan model paling optimal berdasarkan hasil prediksi yang memiliki nilai *error* terkecil. Pengujian dilakukan dengan mengubah nilai periode *Exponential Moving Average* dengan nilai 9, 10, 21, dan 50 dan periode *Relative Strength Index* dengan nilai 5, 9, 14, 21, dan 25.

4. Analisa Hasil

Hasil dan kesimpulan dari percobaan yang dilakukan dilihat berdasarkan nilai *residuals (error)* pada data hasil prediksi untuk menentukan model dengan performa paling baik.



BAB II

TINJAUAN PUSTAKA DAN LANDASAN TEORI

2.1 Tinjauan Pustaka

Beberapa penelitian sebelumnya yang melakukan prediksi dengan menggunakan metode *Multiple Linear Regression* antara lain adalah penelitian yang dilakukan untuk memprediksi perkiraan pengisian Bahan Bakar Mesin High Speed Diesel (BBM HSD) lokomotif yang akan berangkat menggunakan metode *Multiple Linear Regression*. Data yang digunakan pada penelitian ini adalah data kilometer tempuh atau jarak, data waktu tempuh atau durasi perjalanan dan berat setiap kereta api. Data diambil dari salah satu Dipo Lokomotif di Jakarta dengan kurun waktu satu tahun dari April 2019 sampai Maret 2020. Hasil rata-rata kesalahan prediksi saat dihitung menggunakan MSE dan MAPE adalah 7.874 dan 5%. Dapat dikatakan bahwa kesalahan prediksi volume pengisian BBM HSD adalah cukup rendah (Deni & Latifah, 2022). Penggunaan metode *Multiple Linear Regression* juga dilakukan dalam penelitian untuk memprediksi Indeks Prestasi Semester mahasiswa. Hasil prediksi dari penelitian ini menghasilkan tingkat error ME sebesar 0,12, tingkat error MAD sebesar 0,15, tingkat error MSE sebesar 1,53, tingkat error RMSE sebesar 1,24, dan tingkat error MAPE sebesar 4,05% (Faruqhy dkk, 2021).

Penelitian lain yang telah dilakukan untuk memprediksi harga penutupan saham dengan menggunakan berbagai metode antara lain, penelitian yang dilakukan untuk memprediksi harga penutupan saham dengan menggunakan metode *Artificial Neural Networks* (ANN) dan *Random Forest* (RF). Hasil terbaik yang didapatkan dengan menggunakan algoritma ANN mendapatkan nilai RMSE sebesar 0.42, MAPE sebesar 0.77% dan MBE sebesar -0.0156. Sedangkan algoritma RF mendapatkan nilai RMSE sebesar 0.43, MAPE sebesar 0.8% dan MBE sebesar -0.0155 (Vijh dkk, 2020).

Penelitian selanjutnya membandingkan metode *Linear Regression* dan *Support Vector Machine* untuk memprediksi harga penutupan saham dari perusahaan *Coca-Cola* selama satu tahun dari Januari 2017 sampai 2018. Model *Linear Regression* mendapatkan hasil RMSE 3.22 MAE 2.53 MSE 10.37 R-Squared 0.73. Model *Support Vector Machine* mendapatkan hasil RMSE 1.58 MAE 1.33 MSE 2.51 R-Squared 0.93. (Gururaj dkk, 2019).

Prediksi harga saham juga dilakukan pada saham “Jakarta Composite Index (JKSE)” dengan menggunakan metode *Multiple Linear Regression* (MLR) dibantu dengan teknik *K-Means* (KM) dan *Moving Average* (MA) yang berfungsi untuk mengatasi pengaruh *data outlier*. Hasil dari metode MLR mendapatkan nilai MSE sebesar 26849.807, RMSE sebesar 163.859, dan MAPE sebesar 5.754. Sedangkan metode MLR-KM mendapatkan hasil yang lebih buruk dengan nilai MSE sebesar 55143754.119, RMSE sebesar 7425.884, dan MAPE sebesar 12199.802, hal ini menunjukkan bahwa *K-Means* tidak berfungsi baik dalam hal menangani data outlier pada data saham. Hasil metode terakhir yakni MLR-MA menunjukkan hasil yang lebih baik dibandingkan kedua metode sebelumnya yakni nilai MSE sebesar 15087.465, RMSE sebesar 122.831, dan MAPE sebesar 3.255 (Izzah & Widyastuti, 2017).

Penelitian terakhir yang dilakukan untuk memprediksi harga saham dengan mengkombinasikan metode *Decision Tree* (DT) dan *Artificial Neural Networks* (ANN). Hasil akurasi dengan menggunakan DT mendapat nilai sebesar 65%, sedangkan nilai akurasi dari ANN adalah 59%. Sedangkan hasil kombinasi antara DT dan ANN mendapatkan nilai akurasi sebesar 77% dan kombinasi metode DT dan DT mendapatkan nilai akurasi sebesar 67%. Dari hasil ini dapat dilihat bahwa dengan mengkombinasikan metode DT dan ANN dapat meningkatkan nilai akurasi dengan cukup baik (Tsai & Wang, 2009). Perbandingan dari penelitian yang telah dilakukan sebelumnya dapat dilihat pada tabel 2.1.

Tabel 2.1 Literature review

| No | Peneliti | Algoritma | Hasil |
|----|---|---|--|
| 1 | Deni, M., & Latifah, R. (2022) | <i>Multiple Linear Regression</i> | MSE 7.874 MAPE 5% |
| 2 | Faruqhy, M. N., Andreswari, D., & Sari, J. P. (2021) | <i>Multiple Linear Regression</i> | ME 0,12 MAD 0,15 MSE 1,53 RMSE 1,24 MAPE 4,05% |
| 3 | Vijh, M., Chandola, D., Tikkiwal, V. A., & Kumar, A. (2020) | <i>Artificial Neural Networks</i> | RMSE 0.42 MAPE 0.77% MBE -0.0156. |
| | | <i>Random Forest</i> | RMSE 0.43 MAPE 0.8% MBE -0.0155. |
| 4 | Gururaj, V., Shriya, V. R., & Ashwini, K. (2019). | <i>Linear Regression</i> | RMSE 3.22 MAE 2.53 MSE 10.37 R-Squared 0.73 |
| | | <i>Support Vector Machine</i> | RMSE 1.58 MAE 1.33 MSE 2.51 R-Squared 0.93 |
| 5 | Izzah, A., & Widyastuti, R. (2017) | <i>Multiple Linear Regression</i> | MSE 26849.807 RMSE 163.859 MAPE 5.754 |
| | | <i>Multiple Linear Regression + K-Means</i> | MSE 55143754.119 RMSE 7425.884 MAPE 12199.802 % |
| | | <i>Multiple Linear</i> | MSE 15087.465 |

| | | | |
|---|-----------------------------------|---|------------------------------|
| | | <i>Regression + Moving Average</i> | RMSE 122.831 MAPE 3.255 % |
| 6 | Tsai, C. F., & Wang, S. P. (2009) | <i>Decision Tree</i> | Akurasi 65% |
| | | <i>Artificial Neural Networks</i> | Akurasi 59% |
| | | <i>Decision Tree + Artificial Neural Networks</i> | Akurasi 77% |
| | | <i>Decision Tree + Decision Tree</i> | Akurasi 67% |

2.2 Landasan Teori

2.2.1 Saham

Menanamkan sejumlah dana pada saat ini dengan harapan di masa datang dana tersebut akan meningkat nilainya disebut investasi. Salah satu bentuk investasi yang populer saat ini adalah saham. Saham adalah bentuk investasi yang memperdagangkan surat berharga sebagai bukti sebagian kepemilikan seseorang atas suatu perusahaan. Dengan membeli saham pada suatu perusahaan, maka investor menginvestasikan dana yang dimilikinya pada perusahaan tersebut. Salah satu tempat untuk memperjualbelikan saham dapat dilakukan melalui pasar modal. Keuntungan dalam investasi saham berupa capital gain dan dividen. Capital gain adalah selisih antara harga saham pada periode ini dengan harga saham pada periode sebelumnya. Sedangkan dividen merupakan pembagian keuntungan dari perusahaan untuk pemegang saham (Kustina & Safitri, 2019).

Harga saham yang bersifat fluktuatif, dapat naik dan turun kapan saja dan mudah dipengaruhi oleh sentimen publik. Beberapa penelitian mengaitkan harga saham dengan kinerja perusahaan yang bisa dilihat dari laporan keuangannya. Jika kinerja perusahaan baik maka harga saham juga akan meningkat. Perusahaan yang ada di pasar modal memiliki kewajiban

untuk mempublikasikan laporan keuangannya minimal setiap tiga bulan sekali (Oktavia & Nugraha, 2018). Selain kinerja perusahaan, ada faktor fundamental internal dan eksternal yang mempengaruhi harga saham. Fundamental yang terdiri dari intern meliputi keadaan para emiten seperti, pembagian deviden, perubahan strategi dalam rapat umum pemegang saham, akan menjadi informasi penting bagi para investor di pasar modal. Sedangkan faktor ekstern meliputi kebijakan pemerintah, perubahan nilai tukar rupiah terhadap dollar (kurs), keadaan inflasi, penetapan suku bunga oleh Bank Indonesia (Putri & Tjahjono, 2011).

2.2.2 *Technical Indicators*

2.2.2.1 *Pengertian Technical Indicators*

Pada dasarnya untuk menganalisa perubahan harga saham terdapat dua cara yaitu dengan analisis fundamental dan analisis teknikal. Analisis fundamental adalah analisis yang mempelajari kondisi keuangan suatu perusahaan dengan melihat laporan keuangan perusahaan. Sedangkan analisis teknikal adalah analisis yang mempelajari pergerakan pasar dengan melihat perubahan harga dan volume perdagangan (Asthi dkk, 2016). *Technical indicators* adalah nilai yang dihitung untuk menggambarkan pergerakan harga historis sekuritas seperti saham yang berguna untuk memprediksi harga aset di masa depan karena dapat diintegrasikan ke dalam sistem perdagangan otomatis (Agrawal dkk, 2019). Dalam analisis teknikal terdapat berbagai *technical indicators* yang dapat membantu dalam menganalisis pergerakan harga saham. *Technical indicators* yang umum digunakan untuk menganalisis harga saham antara lain *Moving Average (MA)*, *Exponential Moving Average (EMA)*, *Moving Average Convergence Divergence (MACD)*, dan *Relative Strength Index (RSI)* yang mana setiap *technical indicators* memiliki peran dan kegunaan yang berbeda-beda. Contoh penerapan *technical indicators* dapat dilihat pada gambar 2.1.



Sumber: Yahoo Finance

Gambar 2.1 Technical Indicators

2.2.2.2 Exponential Moving Average

Exponential Moving Average (EMA) adalah *technical indicators* yang membantu memperlancar tren harga dan digunakan untuk membantu mengidentifikasi pembalikan harga tren yang muncul dengan memberikan bobot pada rata-rata perubahan harga (Vargas dkk, 2018). Terdapat kesepakatan antara para peneliti analis teknikal mengenai rentang periode untuk moving average dimana periode tren jangka pendek kurang dari 20 hari, periode tren jangka menengah 20 hingga 100, dan periode tren jangka panjang adalah lebih dari 100 (El-Khodary, 2009). Periode yang terlalu pendek dalam harga sekuritas tidak memiliki efek signifikan pada rata-rata pergerakan yang lebih panjang. Namun, jika tren saham berubah secara tiba-tiba, nilai *Exponential Moving Average* akan membutuhkan waktu lebih lama untuk beradaptasi. Periode nilai periode terbaik untuk moving average harian adalah 9 atau 10 hari untuk jangka pendek, 21 hari untuk jangka menengah, dan 50 hari untuk jangka panjang (Rolf, 2018).



Sumber: Yahoo Finance

Gambar 2.2 Exponential Moving Average (50)

Untuk menghitung *Exponential Moving Average* dapat menggunakan persamaan (2. 1).

$$EMA = P(t) \times k + EMA(y) \times (1 - k) \quad (2.1)$$

Yang mana:

$$k = \frac{2}{n + 1} \quad (2.2)$$

Keterangan:

P = Price (Harga saham)

t = Today (Hari ini).

y = Yesterday (Kemarin).

n = Periode waktu.

k = Nilai bobot.

Karena untuk menghitung EMA membutuhkan nilai EMA sebelumnya maka untuk nilai EMA pertama dapat digantikan dengan *Simple Moving Average* (SMA). *Simple Moving Average* adalah rata-rata harga selama n hari. Rumus menghitung SMA sebagai berikut:

$$SMA = \frac{P_1 + P_2 + \dots + P_n}{n} \quad (2.3)$$

Keterangan:

P = Price (Harga saham)

n = Periode waktu.

2.2.2.3 *Relative Strength Index*

Relative Strength Index (RSI) merupakan *technical indicators* yang berperan sebagai parameter momentum, yakni mengukur pergerakan harga dengan membandingkan keuntungan rata-rata periode naik dengan kerugian rata-rata periode turun selama jangka waktu tertentu (Vergas dkk, 2018). RSI pertama kali diperkenalkan oleh Wilder dengan merekomendasikan periode 14 hari. Wilder berpendapat bahwa, semakin pendek jangka waktu atau periode perhitungan yang digunakan, maka *oscillator* akan semakin sensitif, sebaliknya, semakin besar jangka waktu atau periode yang digunakan, maka *oscillator* ini akan semakin tidak sensitif (*amplitude*). Beberapa *chartist* menggunakan periode 5 atau 9 hari, untuk menambah *volatility* garis RSI, juga ada yang menggunakan periode 21 atau 25 hari, dengan tujuan untuk menghaluskan RSI (Waldi, 2020).

Relative Strength Index (RSI) memiliki kisaran angka dari 0-100. Penilaian yang biasa digunakan adalah kisaran antara 30-70. Ketika RSI memotong garis 30 maka harga saham berada pada kondisi *oversold*, dan ketika memotong garis 70, harga saham berada pada kondisi *overbought*. Didalam RSI juga terdapat *centerline* yaitu garis yang menentukan tren kenaikan atau penurunan yang berada pada level 50. Pada saat RSI menembus level 50 dari bawah maka sedang terjadi tren kenaikan yang menandakan saham sedang dibeli secara aktif, sedangkan apabila

RSI menembus level 50 dari atas maka yang terjadi adalah tren penurunan yang mengindikasikan melemahnya tren saham.



Sumber: Yahoo Finance

Gambar 2.3 Relative Strength Index (9)

Untuk menghitung *Relative Strength Index* dapat menggunakan persamaan (2. 4).

$$RSI = 100 - \frac{100}{1 + RS} \quad (2.4)$$

yang mana:

$$RS = \frac{Avg. Gain}{Avg. Loss} \quad (2.5)$$

keterangan:

Avg. Gain = Rata-rata keuntungan.

Avg. Loss = Rata-rata kerugian.

2.2.3 *Multiple Linear Regression*

Metode Regresi, yaitu sebuah metode statistik juga yang melakukan peramalan dengan menggunakan pengembangan hubungan matematis antara variabel dependen (Y) dengan variabel independen (X) (Khotimah & Nindyasari, 2017). Metode *Linear Regression* dibedakan menjadi dua jenis yaitu *Simple Linear Regression* (SLR) dan *Multiple Linear Regression* (MLR). Secara sederhana, dapat dikatakan bahwa *Simple Linear Regression* hanya melibatkan satu variabel independen, sementara *Multiple Linear Regression* melibatkan lebih dari satu variabel independen (Syahra dkk, 2019). *Multiple Linear Regression* dapat dihitung menggunakan persamaan (2. 6).

$$\hat{y} = \alpha + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n \quad (2.6)$$

Keterangan:

\hat{y} = Hasil prediksi.

α = Konstanta (nilai \hat{y} apabila $X_1, X_2, \dots, X_n = 0$).

β = Nilai koefisien regresi.

X = Variabel independen.

Untuk memprediksi menggunakan *Multiple Linear Regression* dengan dua variabel independen dengan menggunakan persamaan (2. 7).

$$\hat{y} = \alpha + \beta_1 X_1 + \beta_2 X_2 \quad (2.7)$$

Nilai α , β_1 , dan β_2 dapat dihitung dengan rumus berikut:

$$\beta_1 = \frac{[(\sum x_2^2 \times \sum x_1 y) - (\sum x_2 y \times \sum x_1 x_2)]}{[(\sum x_1^2 \times \sum x_2^2) - (\sum x_1 x_2)^2]} \quad (2.8)$$

$$\beta_2 = \frac{[(\sum x_1^2 \times \sum x_2 y) - (\sum x_1 y \times \sum x_1 x_2)]}{[(\sum x_1^2 \times \sum x_2^2) - (\sum x_1 x_2)^2]} \quad (2.9)$$

$$\alpha = \frac{(\sum y) - (\beta_1 \times \sum x_1) - (\beta_2 \times \sum x_2)}{n} \quad (2.10)$$

yang mana:

$$\sum x_1^2 = \sum X_1^2 - \frac{(\sum X_1)^2}{n} \quad (2.11)$$

$$\sum x_2^2 = \sum X_2^2 - \frac{(\sum X_2)^2}{n} \quad (2.12)$$

$$\sum y^2 = \sum Y^2 - \frac{(\sum Y)^2}{n} \quad (2.13)$$

$$\sum x_1 y = \sum X_1 Y - \frac{\sum X_1 \sum Y}{n} \quad (2.14)$$

$$\sum x_2 y = \sum X_2 Y - \frac{\sum X_2 \sum Y}{n} \quad (2.15)$$

$$\sum x_1x_2 = \sum X_1X_2 - \frac{\sum X_1 \sum X_2}{n} \quad (2.16)$$

2.2.4 Performance Evaluation

2.2.4.1 Koefisien Determinasi

Koefisien determinasi atau dikenal juga dengan *R-Square* (r^2) merupakan salah satu metode yang dapat digunakan untuk mengetahui seberapa baik model regresi berdasarkan informasi yang dihasilkan oleh variabel independen pada saat memprediksi variabel respon. Nilai koefisien determinasi memiliki interval antara 0 sampai 1, apabila nilai koefisien determinasi semakin mendekati nilai 1 maka semakin kuat keterkaitan variabel independen dengan variabel respon pada model persamaan regresi yang terbentuk (Yuliara, 2016).

Untuk menghitung nilai koefisien determinasi digunakan persamaan (2. 17).

$$r^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y}_i)^2} \quad (2.17)$$

Keterangan:

y = Nilai aktual.

\hat{y} = Nilai prediksi y .

\bar{y} = Nilai rata-rata y .

2.2.4.2 Mean Absolute Error

Salah satu metode untuk melihat seberapa baik suatu metode dengan melihat dari kesalahan (*error*) dari hasil prediksi adalah metode *Mean Absolute Error* (MAE). MAE adalah rata-rata perbedaan absolut antara prediksi dan data aktual di mana semua perbedaan individu memiliki bobot yang sama (Sutoyo & Almaarif, 2020). Jika nilai MAE semakin mendekati 0 maka tingkat kesalahan

(*error*) dari hasil prediksi semakin kecil. *Mean Absolute Error* dapat dihitung dengan persamaan (2. 18).

$$MAE = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n} \quad (2.18)$$

Keterangan:

y = Nilai aktual.

\hat{y} = Nilai prediksi.

n = Jumlah data.

2.2.4.3 Root Mean Square Error

Metode lain untuk menilai hasil prediksi adalah *Root Mean Square Error* (RMSE). RMSE adalah aturan penilaian kuadrat yang mengukur besarnya rata-rata kesalahan (*error*) dari suatu model. RMSE menghitung *error* dengan akar kuadrat dari rata-rata perbedaan kuadrat antara data hasil prediksi dan data sebenarnya (Chai & Draxler, 2014). Jika nilai yang dihasilkan dari metode RMSE semakin mendekati 0 maka semakin baik hasil prediksi. RMSE dapat dihitung dengan persamaan (2. 19).

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (2.19)$$

Keterangan:

y = Nilai aktual.

\hat{y} = Nilai prediksi.

n = Jumlah data.

2.2.4.4 Mean Absolute Percentage Error

Mean Absolute Percentage Error (MAPE) adalah nilai rata-rata perbedaan absolut yang ada diantara nilai dari prediksi dan nilai aktual yang ditampilkan dalam hasil persentase. Penggunaan *Mean Absolute Percentage Error* (MAPE) pada evaluasi dari hasil peramalan dapat melihat tingkat akurasi terhadap angka peramalan dan angka realisasi (Nabillah & Ranggadara, 2020). MAPE dihitung dengan menggunakan kesalahan absolut pada setiap periode dibagi dengan menggunakan nilai observasi yang nyata pada periode tersebut. Kemudian, merata-rata kesalahan persentase absolut tersebut. Pendekatan ini berguna ketika ukuran atau besar variabel ramalan itu penting dalam mengevaluasi ketepatan ramalan. MAPE mengindikasi seberapa besar kesalahan dalam meramal yang dibandingkan dengan nilai nyata (Wulandari dkk, 2015). MAPE dapat dihitung dengan persamaan (2. 20).

$$MAPE = \frac{\sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i}}{n} \times 100\% \quad (2.20)$$

Keterangan:

y = Nilai aktual.

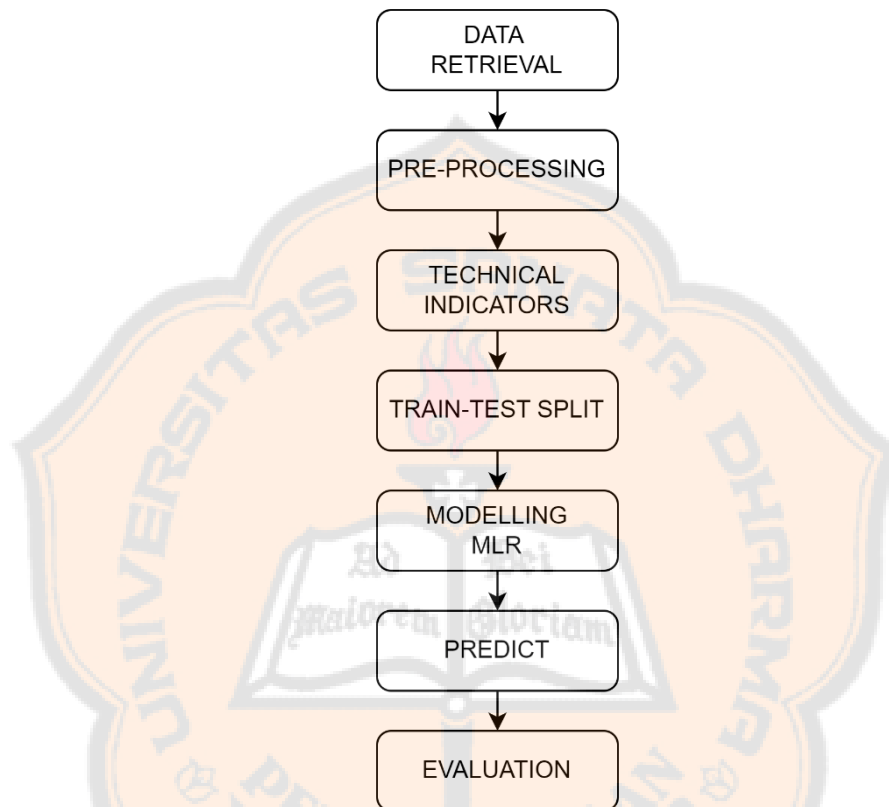
\hat{y} = Nilai prediksi.

n = Jumlah data.

BAB III

METODE PENELITIAN

3.1 Desain Alat Uji



Gambar 3.1 Desain alat uji

Gambar 3.1 merupakan alur proses prediksi harga saham yang akan dibangun dengan menggunakan algoritma *Multiple Linear Regression*. Berikut penjelasan mengenai setiap proses:

1. *Data Retrieval*

Proses *data retrieval* adalah proses mengunduh data dari website finance.yahoo.com yang akan digunakan menjadi data mentah dari proses prediksi.

2. *Pre-processing*

Pre-processing merupakan proses yang dilakukan sebelum proses data mining dilaksanakan. Terdapat dua buah proses yang akan dilakukan yakni proses *data cleaning* dan *data selection*. Proses *data cleaning* yang dilakukan adalah menghapus data kosong dan data yang memiliki nilai volume 0. Data yang memiliki volume 0 disebabkan karena tidak terjadi transaksi pada hari tersebut. Proses ini perlu dilakukan untuk menghindari perubahan harga yang tidak normal. Proses selanjutnya adalah melakukan *data selection* atau pemilihan data untuk memilih atribut yang akan digunakan dalam prediksi. Atribut yang digunakan adalah Close, karena proses perhitungan nilai *technical indicators* baik *Exponential Moving Average (EMA)* ataupun *Relative Strength Index (RSI)* membutuhkan data harga penutupan saham lihat rumus 2.1 dan 2.4. Sedangkan atribut lainnya akan dihapus karena tidak digunakan pada proses prediksi.

3. Perhitungan *Technical Indicators*

Setelah data melalui proses *pre-processing* dan *data selection* selanjutnya dilakukan perhitungan *Exponential Moving Average (EMA)* dan *Relative Strength Index (RSI)* yang akan digunakan sebagai atribut independen.

4. *Train-Test Split*

Train-test split merupakan proses untuk memisahkan data menjadi 2 bagian yaitu data *train* (latih) dan data *test* (uji). Meskipun perhitungan untuk mendapatkan nilai konstanta dan koefisien dapat dilakukan dengan keseluruhan data, tidak harus memisahkan data menjadi data latih dan data uji. Akan tetapi karena pada penelitian ini membutuhkan model yang dapat memprediksi data, maka data dapat dipisahkan menjadi 2 buah yakni data latih dan data uji. Data latih digunakan untuk membentuk model *Multiple Linear Regression* dengan menghitung nilai konstanta dan koefisien. Sedangkan data uji

adalah data yang akan diprediksi dengan menggunakan model yang terbentuk dari data training. Hasil prediksi dari data uji tersebut dapat digunakan untuk menilai seberapa baik model regresi yang terbentuk. Model tersebut juga dapat digunakan untuk memprediksi data lain tanpa harus melakukan perhitungan untuk menentukan nilai konstanta dan koefisien secara berulang. Pembagian data untuk *train-test split* adalah 80% data awal sebagai data latih dan 20% sisanya digunakan sebagai data uji.

5. *Multiple Linear Regression* (MLR)

Perhitungan *Multiple Linear Regression* (MLR) merupakan tahap pembentukan model, data latih yang telah dibuat pada proses *Train-Test Split* akan digunakan dalam perhitungan ini. Atribut yang menjadi variabel independen adalah *Exponential Moving Average* dan *Relative Strength Index* sedangkan *Close* menjadi target data yang akan diprediksi. Pada proses pembentukan model menghasilkan nilai konstanta dan koefisien yang digunakan untuk memprediksi data uji.

6. Prediksi

Setelah model *Multiple Linear Regression* terbentuk, kemudian dilanjutkan dengan memprediksi data uji. Setelah data uji diprediksi maka akan menghasilkan nilai prediksi untuk setiap data uji.

7. *Evaluation*

Tahap terakhir yang akan dilakukan adalah penilaian seberapa baik hasil prediksi yang dihasilkan dari model yang telah terbentuk. Metode penilaian yang digunakan adalah *Koefisien Determinasi* (r^2), *Mean Absolute Error* (MAE), dan *Root Mean Square Error* (RMSE) dan *Mean Absolute Percentage Error* (MAPE).

3.2 Pengumpulan Data

Dataset yang digunakan untuk memprediksi harga saham adalah data historis harian dari saham PT Bank Central Asia Tbk yang diambil dari website finance.yahoo.com. Data yang digunakan merupakan data selama 5 tahun terakhir. Data ini memiliki 6 atribut yaitu *Open*, *High*, *Low*, *Close*, *Adj Close*, dan *Volume* dengan index *Date*. Atribut *Close* akan menjadi variabel dependen (target data) yang akan diprediksi pada penelitian ini. Penjelasan setiap atribut dari dataset dapat dilihat pada tabel 3.1:

Tabel 3.1 Atribut data

| No | Atribut | Penjelasan |
|----|----------------|---|
| 1 | Date | Tanggal terjadinya transaksi dalam format bulan/hari/tahun sebagai index data. |
| 2 | Open | Harga pertama kali transaksi dilakukan pada hari itu. |
| 3 | High | Kisaran harga tertinggi pada pergerakan harian dari saham. |
| 4 | Low | Kisaran harga terendah pada pergerakan harian dari saham. |
| 5 | Close | Harga terakhir kali transaksi dilakukan pada hari itu. |
| 6 | Adjusted Close | Harga penutupan setelah penyesuaian untuk semua split dan pembagian dividen yang berlaku. |
| 7 | Volume | Jumlah transaksi yang terjadi dalam jumlah lembar saham. |

Tabel 3.2 Contoh dataset saham

| Date | Open | High | Low | Close | Adj Close | Volume |
|------------|--------|--------|--------|--------|-------------|-------------|
| 2017-11-30 | 4260.0 | 4260.0 | 4070.0 | 4070.0 | 3758.374756 | 205897500.0 |
| 2017-12-01 | 4070.0 | 4070.0 | 4070.0 | 4070.0 | 3758.374756 | 0.0 |
| 2017-12-04 | 4165.0 | 4230.0 | 4130.0 | 4160.0 | 3841.483643 | 128402500.0 |
| 2017-12-05 | 4210.0 | 4225.0 | 4180.0 | 4200.0 | 3878.421143 | 62670500.0 |
| 2017-12-06 | 4230.0 | 4275.0 | 4200.0 | 4260.0 | 3933.827393 | 106299500.0 |
| ... | ... | ... | ... | ... | ... | ... |
| 2022-11-24 | 8975.0 | 9000.0 | 8950.0 | 9000.0 | 8965.000000 | 53520800.0 |
| 2022-11-25 | 9000.0 | 9000.0 | 8900.0 | 8975.0 | 8940.096680 | 75048300.0 |
| 2022-11-28 | 9050.0 | 9100.0 | 8975.0 | 9025.0 | 8989.902344 | 96209700.0 |
| 2022-11-29 | 8975.0 | 9075.0 | 8925.0 | 8975.0 | 8940.096680 | 91562100.0 |
| 2022-11-30 | 8875.0 | 9400.0 | 8850.0 | 9300.0 | 9263.833008 | 644359600.0 |

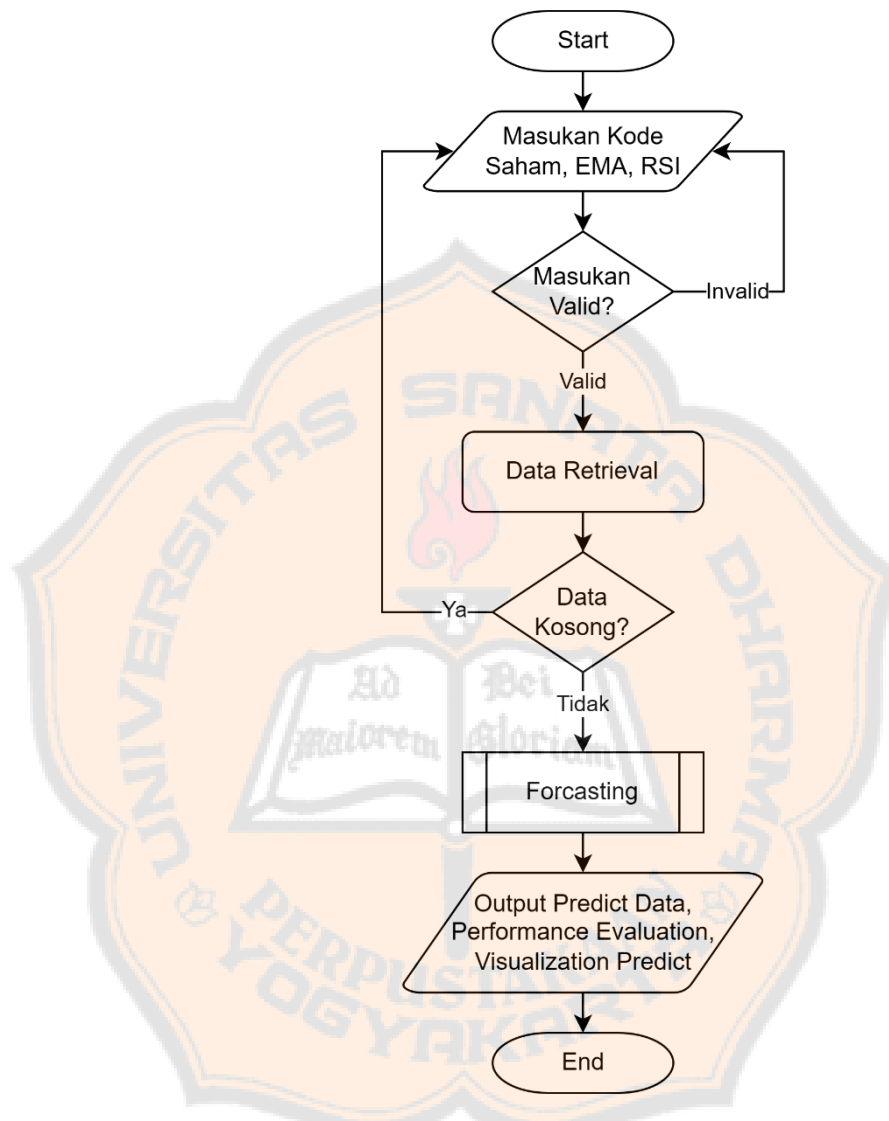
3.3 Skenario Pengujian

Skenario pengujian bertujuan untuk mengetahui nilai periode EMA dan RSI dengan hasil performa terbaik. Pada penelitian ini skenario pengujian dilakukan dengan mengubah nilai periode *Exponential Moving Average* dan *Relative Strength Index*. Nilai periode yang direkomendasikan untuk *Moving Average* harian yakni 9 atau 10 hari untuk jangka pendek, 21 hari untuk jangka menengah, dan 50 hari untuk jangka panjang (Rolf, 2018). Sedangkan untuk nilai periode *Relative Strength Index* akan menggunakan nilai 5 dan 9 hari untuk jangka pendek, nilai yang direkomendasikan oleh Wilder yakni 14 hari sebagai jangka menengah, dan 21 dan 25 hari sebagai periode jangka panjang. Dengan mengkombinasikan nilai periode EMA dan RSI yang direkomendasikan maka akan terdapat 20 kali percobaan yang dapat dilihat pada tabel 3.3.

Gambar 3.2 merupakan desain *user interface* dari penelitian prediksi harga saham, berikut penjelasan untuk komponen *interface*:

- a. *Entry SYMBOL* merupakan *entry* untuk pengguna memasukan data *code stock*.
- b. *Entry EMA* merupakan *entry* untuk pengguna memasukan nilai *Exponential Moving Average (EMA)*.
- c. *Entry RSI* merupakan *entry* untuk pengguna memasukan nilai *Relative Strength Index (RSI)*.
- d. *Button FORECAST* merupakan *button* yang berfungsi untuk memulai proses prediksi.
- e. Tabel pada bagian kiri berfungsi untuk menampilkan data hasil prediksi, yakni *date*, *actual close price*, dan *predict close price*.
- f. *Field R²* berfungsi untuk menampilkan nilai *coefficient of determination* dari hasil prediksi.
- g. *Field MAE* berfungsi untuk menampilkan nilai *Mean Absolute Error* dari hasil prediksi.
- h. *Field RMSE* berfungsi untuk menampilkan nilai *Root Mean Square Error* dari hasil prediksi.
- i. *Field MAPE* berfungsi untuk menampilkan nilai *Mean Absolute Percentage Error* dari hasil prediksi.
- j. *Area chart* pada bagian bawah berfungsi untuk menampilkan visualisasi dari hasil prediksi.

Alur pengguna memprediksi harga saham dengan menggunakan model *Multiple Linear Regression* dapat dilihat pada *flowchart* Gambar 3.3.



Gambar 3.3 Flowchart Penggunaan User Interface

1. User memasukkan kode saham (contoh kode saham untuk PT Bank Central Asia Tbk adalah “BBCA.JK”), nilai periode EMA, dan nilai periode RSI pada *entry* yang disediakan. Selanjutnya user menekan tombol FORECAST.
2. Saat user menekan tombol forecast maka program akan melihat apakah data yang masukan oleh user *valid* atau *invalid*. Kondisi data *valid* adalah:

- Nilai masukan kode saham tidak kosong, dan tidak mengandung tanda koma “,”.
- Nilai masukan periode EMA tidak kosong, periode EMA berupa angka bilangan bulat, dan periode EMA berupa bilangan positif.
- Nilai masukan periode RSI tidak kosong, periode RSI berupa angka bilangan bulat, dan periode RSI berupa bilangan positif.

Apabila data yang di masukan user *invalid* maka program akan menampilkan error dan user harus memasukan kembali data yang salah.

3. Apabila data yang dimasukan user *valid* maka program akan mengunduh data saham berdasarkan masukan dari user.
4. Selanjutnya program akan mengecek data yang berhasil di unduh. Apabila data kosong maka program akan menampilkan pesan error dan user harus memasukan kembali data kode saham, karena tidak ada data dengan kode saham yang dimasukan oleh user.
5. Apabila data tidak kosong maka program akan melakukan proses *forecasting* pada data yang berhasil di unduh. Proses *forecasting* adalah tahap prediksi data mulai dari tahap *preprocessing* sampai *evaluation* pada desain alat uji yang dapat dilihat pada gambar 3.1.
6. Setelah proses prediksi selesai maka akan ditampilkan hasil berupa:
 - a. Data hasil prediksi pada bagian tabel sebelah kiri.
 - b. Hasil evaluasi berupa nilai R^2 , MAE, RMSE, dan MAPE pada *field* kanan atas.
 - c. Hasil visualisasi pada bagian CHART.

BAB IV

HASIL DAN ANALISA

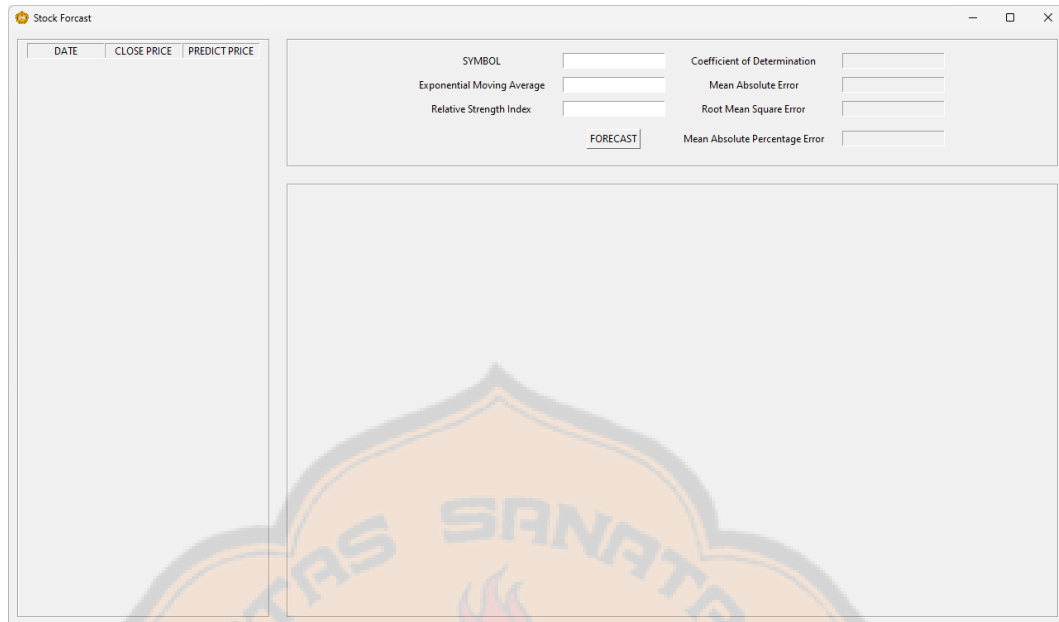
4.1 Hasil

4.1.1 Implementasi Program

Untuk mengimplementasikan menjadi sebuah program siap pakai, digunakan bahasa pemrograman Python versi 3.11.1 64-bit. Hasil data yang ditampilkan pada implementasi program ini, diuji dengan menggunakan nilai periode EMA 9 dan RSI 5 pada tanggal 30 November 2022. Pada implementasi program prediksi harga penutupan saham, semua perhitungan dan pembuatan *user interface* dilakukan di dalam *class* bernama *StockForecast*. Di dalam *class* *StockForecast* terdapat 3 fungsi yaitu `__init__()`, `check()`, dan `forecast()`. Penjelasan untuk setiap fungsi pada *class* *StockForecast* sebagai berikut:

4.1.1.1 Fungsi `__init__()`

Fungsi `__init__()` merupakan fungsi yang akan otomatis dipanggil saat *class* *StockForecast* diinisialisasikan, dimana fungsi ini akan membuat *user interface* dengan menginisialisasikan setiap komponen yang dibutuhkan dengan menggunakan *library Tkinter*. *Source code* fungsi `__init__()` dapat dilihat pada lampiran 1. Tampilan awal saat *user interface* dijalankan dapat dilihat pada gambar 4.1.



Gambar 4.1 Tampilan user interface awal

4.1.1.2 Fungsi check()

Fungsi check() merupakan fungsi yang dipanggil saat *button FORECAST* ditekan. Dimana fungsi ini akan melakukan prose pengecekan data yang dimasukan oleh user *valid* (benar) atau *invalid* (salah) dengan kondisi sebagai berikut:

- Jika nilai masukan kode saham kosong, atau mengandung tanda koma “,” maka akan menampilkan pesan error “*No Data Found, Symbol May Be Delisted!*”.
- Jika nilai masukan periode EMA kosong atau periode EMA bukan berupa bilangan bulat atau periode EMA bernilai kurang dari 1 maka akan menampilkan pesan error “*Data Is Invalid!*”.
- Jika nilai masukan periode RSI kosong atau periode RSI bukan berupa bilangan bulat atau periode RSI bernilai kurang dari 1 maka akan menampilkan pesan error “*Data Is Invalid!*”.

Apabila masukan dari user memenuhi salah satu kondisi *invalid*, maka user diminta untuk memasukan kembali data yang salah pada *user interface*. Jika data masukan dari user *valid* (benar) maka dilanjutkan

pada tahap *data retrieval*. *Source code* fungsi `check()` dapat dilihat pada lampiran 2.

Tahap *data retrieval* adalah proses mengunduh data dari website `finance.yahoo.com` dengan menggunakan *method download* dari *library yfinance* menjadi data mentah. Dalam *method download*, terdapat parameter *ticker* yaitu *code stock* dari data yang akan diprediksi dimana data diinputkan berdasarkan masukan dari user. Parameter kedua adalah *period* yaitu rentang waktu data yang akan digunakan, berdasarkan batasan masalah pada penelitian ini adalah 5y (5 tahun). Parameter ketiga yang digunakan adalah *interval* yaitu *interval* waktu pada dataset, karena data yang dibutuhkan adalah data harian maka interval yang digunakan adalah 1d (1 hari). *Source code* untuk proses *data retrieval* dapat dilihat pada lampiran 3.

Setelah data berhasil diunduh maka akan dilihat apakah data tersebut kosong atau tidak, jika data kosong maka program akan menampilkan pesan error “*No Data Found, Symbol May Be Delisted!*” dan user diminta untuk memasukan kembali data kode saham pada *user interface*. Jika data yang diunduh tidak kosong maka proses prediksi akan dilanjutkan dengan memanggil fungsi `forecast()`. Contoh data yang didapatkan dari proses *data retrieval* dapat dilihat pada tabel 4.1.

Tabel 4.1 Contoh dataset dari proses data retrieval

| Date | Open | High | Low | Close | Adj Close | Volume |
|------------|--------|--------|--------|--------|-------------|-------------|
| 2017-11-30 | 4260.0 | 4260.0 | 4070.0 | 4070.0 | 3758.374756 | 205897500.0 |
| 2017-12-01 | 4070.0 | 4070.0 | 4070.0 | 4070.0 | 3758.374756 | 0.0 |
| 2017-12-04 | 4165.0 | 4230.0 | 4130.0 | 4160.0 | 3841.483643 | 128402500.0 |
| 2017-12-05 | 4210.0 | 4225.0 | 4180.0 | 4200.0 | 3878.421143 | 62670500.0 |
| 2017-12-06 | 4230.0 | 4275.0 | 4200.0 | 4260.0 | 3933.827393 | 106299500.0 |
| ... | ... | ... | ... | ... | ... | ... |
| 2022-11-24 | 8975.0 | 9000.0 | 8950.0 | 9000.0 | 8965.000000 | 53520800.0 |
| 2022-11-25 | 9000.0 | 9000.0 | 8900.0 | 8975.0 | 8940.096680 | 75048300.0 |
| 2022-11-28 | 9050.0 | 9100.0 | 8975.0 | 9025.0 | 8989.902344 | 96209700.0 |
| 2022-11-29 | 8975.0 | 9075.0 | 8925.0 | 8975.0 | 8940.096680 | 91562100.0 |
| 2022-11-30 | 8875.0 | 9400.0 | 8850.0 | 9300.0 | 9263.833008 | 644359600.0 |

Data yang didapatkan dari proses *data retrieval* berjumlah 1254 data dengan 6 Atribut yang dapat dilihat pada lampiran 4 pada bagian akhir tulisan.

4.1.1.3 Fungsi forecast()

Pada fungsi *forecast()* akan dilakukan beberapa tahapan untuk memprediksi data saham yang sudah diunduh. Tahap-tahap yang dilakukan adalah sebagai berikut:

4.1.1.3.1 *Pre-processing*

Pada tahap *Pre-processing* terdapat dua buah proses yang akan dilakukan yakni proses *data cleaning* dan *data selection*. Sebelum melakukan proses *data cleaning*, peneliti ingin mendapatkan informasi tentang berapa banyak data yang perlu dihapus, maka peneliti akan mencoba mencari jumlah data yang memiliki nilai NaN dan data yang memiliki nilai 0 pada atribut Volume. Jumlah data kosong (NaN) dapat dilihat pada gambar 4.2.

```

=====
Jumlah Data Kosong

Open      0
High     0
Low      0
Close    0
Adj Close 0
Volume   0
dtype: int64
=====

```

Gambar 4.2 Hasil jumlah data kosong pada dataset.

Pada gambar 4.2 dapat dilihat bahwa terdapat 0 data untuk setiap atribut pada dataset yang berarti tidak ada data NaN pada dataset. Sedangkan jumlah data yang memiliki nilai Volume 0 dapat dilihat pada gambar 4.3.

```

=====
Jumlah Data Dengan Volume 0 : 46
=====

```

Gambar 4.3 Jumlah data yang memiliki nilai volume 0

Dari gambar 4.3 terlihat jumlah data yang memiliki nilai 0 pada atribut Volume adalah 46 data. Setelah mendapatkan informasi tentang jumlah data yang akan dihapus maka selanjutnya dilakukan proses untuk menghapus data tersebut menggunakan method dropna, kemudian dilanjutkan dengan menghilangkan data dengan nilai volume 0.

Pada proses data selection, atribut selain Close akan dihapus karena tidak digunakan pada proses selanjutnya. *Source code* untuk proses *Pre-processing* dapat dilihat pada lampiran 5. Setelah proses *Pre-processing* selesai maka dataset sudah siap digunakan untuk proses perhitungan selanjutnya. Contoh data dari hasil proses *Pre-processing* dapat dilihat pada tabel 4.2.

Tabel 4.2 Contoh data hasil preprocessing

| Date | Close |
|------------|--------|
| 2017-11-30 | 4070.0 |
| 2017-12-04 | 4160.0 |
| 2017-12-05 | 4200.0 |
| 2017-12-06 | 4260.0 |
| 2017-12-07 | 4195.0 |
| ... | ... |
| 2022-11-24 | 9000.0 |
| 2022-11-25 | 8975.0 |
| 2022-11-28 | 9025.0 |
| 2022-11-29 | 8975.0 |
| 2022-11-30 | 9300.0 |

Jumlah data yang tersisa dari hasil proses *Pre-processing* adalah sebanyak 1208 data, yang dapat dilihat lampiran 6. Hal ini menunjukkan bahwa terdapat 46 data yang dihapus dari proses *Pre-processing*.

4.1.1.3.2 Perhitungan *Technical Indicators*

Perhitungan *technical indicators* merupakan tahap menambahkan *technical indicators* yang berperan sebagai atribut independen dari proses prediksi harga saham. *Technical indicators* dihitung dengan menggunakan *library pandas_ta*. Parameter *length* berfungsi sebagai nilai periode waktu pada *technical indicators* EMA dan RSI, dimana nilai periode ini diinputkan dari user. *Source code* untuk proses perhitungan *technical indicators* dapat dilihat pada lampiran 7. Setelah perhitungan selesai, akan ada beberapa data pada awal entri pertama yang memiliki nilai NaN atau data kosong lihat Tabel 4.3.

Tabel 4.3 Contoh data hasil perhitungan *technical indicators*

| Date | Close | EMA | RSI |
|------------|--------|-------------|-----------|
| 2017-11-30 | 4070.0 | NaN | NaN |
| 2017-12-04 | 4160.0 | NaN | NaN |
| 2017-12-05 | 4200.0 | NaN | NaN |
| 2017-12-06 | 4260.0 | NaN | NaN |
| 2017-12-07 | 4195.0 | NaN | NaN |
| ... | ... | ... | ... |
| 2022-11-24 | 9000.0 | 8841.329870 | 69.050690 |
| 2022-11-25 | 8975.0 | 8868.063896 | 65.140170 |
| 2022-11-28 | 9025.0 | 8899.451117 | 69.463564 |
| 2022-11-29 | 8975.0 | 8914.560893 | 60.140168 |
| 2022-11-30 | 9300.0 | 8991.648715 | 80.933210 |

Dapat terlihat dari tabel 4.3 bahwa terdapat beberapa data yang kosong pada bagian awal entri. Hal ini disebabkan karena tidak ada nilai lanjutan dari mana EMA dan RSI dapat dihitung. Jumlah data kosong dapat dilihat pada gambar 4.4.

```
=====
```

```
JUMLAH DATA KOSONG
```

```
Close      0
EMA        8
RSI        5
dtype: int64
```

```
=====
```

Gambar 4.4 Jumlah data kosong dari perhitungan *technical indicators*

Dapat dilihat jumlah data kosong dari perhitungan *technical indicators* EMA 9 dan RSI 5 menghasilkan data NaN sebanyak 8 pada atribut EMA dan 5 pada atribut RSI. Hal perlu diatasi dengan menghapus data yang kosong dengan menggunakan method *dropna*

agar tidak terjadi error pada proses perhitungan selanjutnya. Setelah data dihapus maka jumlah data yang tersisa adalah 1200 data, yang dapat dilihat pada lampiran 8.

4.1.1.3.3 *Train-test Split*

Sebelum pembentukan model *Multiple Linear Regression*, dataset yang digunakan harus dipisahkan menjadi dua buah dataset yakni data train dan data testing. Pembagian data ini menggunakan 80% data awal sebagai data training dan 20% sisanya digunakan sebagai data testing. *Source code* untuk proses *Train-test Split* dapat dilihat pada lampiran 9.

Hasil dari proses *Train-test Split* adalah 4 variabel yakni X_{train} , y_{train} , X_{test} , dan y_{test} . Penjelasan untuk setiap variabel dapat dilihat pada tabel 4.4.

Tabel 4.4 Penjelasan variabel pada proses *Train-test Split*

| Variabel | Data Interval | Jumlah Data | Penjelasan |
|-------------|---------------------------|-------------|---|
| X_{train} | 12/13/2017 - 12/8/2021 | 960 | Variabel independen dari data train yang berisi atribut EMA dan RSI |
| y_{train} | 12/13/2017 - 12/8/2021 | 960 | Variabel target dari data train yang berisi atribut Close |
| X_{test} | 12/9/2021 - 11/30/2022 | 240 | Variabel independen dari data test yang berisi atribut EMA dan RSI |
| y_{test} | 12/9/2021 - 11/30/2022 | 240 | Variabel target dari data test yang berisi atribut Close |

Jumlah data train dan testing didapatkan dengan cara sebagai berikut:

$$\begin{aligned} \text{Data Train} &= 0.8 \times 1200 \\ &= 960 \end{aligned}$$

$$\begin{aligned} \text{Data Test} &= 0.2 \times 1200 \\ &= 240 \end{aligned}$$

4.1.1.3.4 *Train Model Multiple Linear Regression*

Proses pembentukan model *Multiple Linear Regression* (MLR) menggunakan kelas *LinearRegression* dari *library Scikit-learn*. Pembentukan model menggunakan variabel *X_train* dan *y_train* yang mana dari proses ini akan menghasilkan nilai konstanta dan koefisien untuk setiap atribut independen. *Source code* untuk proses pembentukan model *Multiple Linear Regression* dapat dilihat lampiran 10. Hasil nilai konstanta dari model yang telah terbentuk dapat dilihat dengan memanggil nilai *intercept_* sedangkan untuk koefisien setiap variabel independen dapat dilihat dengan nilai *coef_*. Nilai konstanta dan koefisien yang dihasilkan dari pembentukan model ini dapat dilihat pada gambar 4.5.

```
=====
```

```
NILAI KONSTANTA & KOEFISIEN DARI MODEL MLR
```

```
Konstanta      : -345.8158861315487
Koefisien 1    : 1.0085809439355629
Koefisien 2    : 5.752810029350752
```

```
=====
```

Gambar 4.5 Hasil nilai konstanta dan koefisien model MLR

4.1.1.3.5 *Prediction*

Proses prediksi dilakukan untuk memprediksi data dari variabel *X_test* dengan menggunakan nilai konstanta dan koefisien yang dihasilkan pada pembentukan model *Multiple Linear Regression* dengan menggunakan persamaan 2.7. Contoh perhitungan prediksi pada data 2022-11-30 dengan menggunakan persamaan 2.7 adalah sebagai berikut:

$$\hat{y} = \alpha + \beta_1 X_1 + \beta_2 X_2$$

$$\hat{y} = -345.8158861315487 + 1.0085809439355629 \times 8991.648715 \\ + 5.752810029350752 \times 80.933210$$

$$\hat{y} = -345.8158861315487 + 9068.80554851169 \\ + 465.5933821955506$$

$$\hat{y} = 9188.583044575691$$

Hasil dari prediksi dengan menggunakan model *Multiple Linear Regression* disimpan pada variabel *y_pred* yang selanjutnya akan digabungkan dengan nilai *Close* untuk mempermudah perhitungan evaluasi hasil prediksi. *Source code* untuk proses *prediction* dapat dilihat pada Lampiran 11. Contoh data hasil prediksi dapat dilihat pada tabel 4.5.

Tabel 4.5 Contoh data hasil prediksi

| Date | Close | Pred_Close |
|------------|--------|-------------|
| 2021-12-09 | 7350.0 | 7355.482036 |
| 2021-12-10 | 7375.0 | 7379.501294 |
| 2021-12-13 | 7300.0 | 7298.785886 |
| 2021-12-14 | 7300.0 | 7285.930598 |
| 2021-12-15 | 7300.0 | 7275.646367 |
| ... | ... | ... |
| 2022-11-24 | 9000.0 | 8968.616440 |
| 2022-11-25 | 8975.0 | 8973.083395 |
| 2022-11-28 | 9025.0 | 9029.611609 |
| 2022-11-29 | 8975.0 | 8991.215314 |
| 2022-11-30 | 9300.0 | 9188.583044 |

4.1.1.3.6 Performance Evaluation

Untuk mengetahui seberapa baik hasil prediksi dari model *Multiple Linear Regression*, maka perlu dilakukan proses *performance evaluation*. Terdapat empat nilai yang akan dihitung yakni *Koefisien Determinasi* (r^2), *Mean Absolute Error* (MAE), *Root Mean Square Error* (RMSE), dan *Mean Absolute Percentage Error* (MAPE). *Source code* untuk proses *performance evaluation* dapat dilihat pada lampiran 13. Hasil *performance evaluation* dari data yang telah diprediksi dapat dilihat pada gambar 4.6.

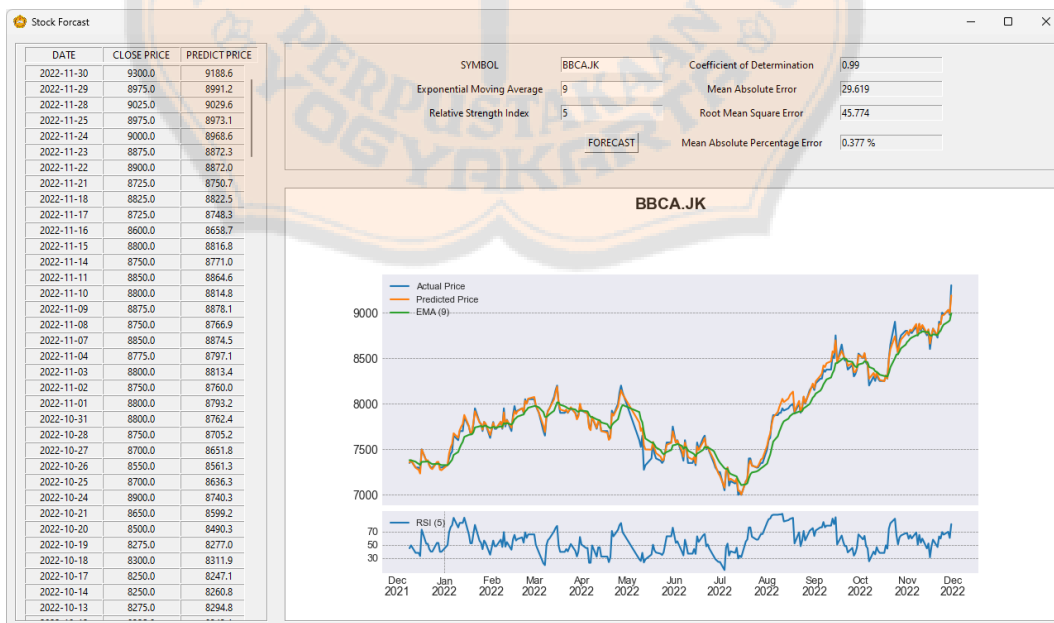
HASIL EVALUASI

Coefficient of Determination = 0.9916685242935429
 Mean Absolute Error = 29.618895641414234
 Root Mean Square Error = 45.77437850953243
 Mean Absolute Percentage Error = 0.0037656643880017064

Gambar 4.6 Hasil evaluasi

4.1.1.3.7 Update Hasil Prediksi Pada User Interface

Tahap terakhir yang akan dilakukan adalah menampilkan hasil prediksi data saham pada *user interface*. Hasil yang akan ditampilkan adalah hasil validasi (R^2 , MAE, RMSE, dan MAPE), data pada tabel (Date, Close Price, dan Predict Price), dan visualisasi data dari hasil prediksi. *Source code* untuk menampilkan data hasil prediksi berupa yang dilakukan dapat dilihat pada lampiran 14. Tampilan *User interface* saat selesai proses prediksi harga saham PT Bank Central Asia Tbk dengan menggunakan nilai periode EMA 9 dan periode RSI 5 dapat dilihat pada gambar 4.7.



Gambar 4.7 Tampilan user interface setelah prediksi

4.1.2 Pengujian

Pengujian ini dilakukan pada tanggal 30 November 2022. Dengan menggunakan kombinasi nilai periode EMA dan RSI yang direkomendasikan pada Tabel 3.3. Hasil percobaan yang telah dilakukan dapat dilihat pada Tabel 4.6.

Tabel 4.6 Hasil pengujian

| No | EMA | RSI | R-Square | MAE | RMSE | MAPE |
|----|-----|-----|----------|---------|---------|--------|
| 1 | 9 | 5 | 0.99 | 29.619 | 45.774 | 0.377% |
| 2 | 9 | 9 | 0.99 | 41.235 | 56.146 | 0.522% |
| 3 | 9 | 14 | 0.98 | 52.498 | 69.585 | 0.666% |
| 4 | 9 | 21 | 0.97 | 60.996 | 82.717 | 0.774% |
| 5 | 9 | 25 | 0.97 | 65.046 | 88.12 | 0.825% |
| 6 | 10 | 5 | 0.99 | 30.165 | 46.554 | 0.385% |
| 7 | 10 | 9 | 0.99 | 39.801 | 54.847 | 0.504% |
| 8 | 10 | 14 | 0.98 | 51.981 | 68.825 | 0.659% |
| 9 | 10 | 21 | 0.97 | 61.551 | 83.428 | 0.781% |
| 10 | 10 | 25 | 0.97 | 66.189 | 89.543 | 0.839% |
| 11 | 21 | 5 | 0.97 | 59.771 | 79.263 | 0.775% |
| 12 | 21 | 9 | 0.99 | 36.903 | 54.781 | 0.475% |
| 13 | 21 | 14 | 0.99 | 36.806 | 50.321 | 0.466% |
| 14 | 21 | 21 | 0.98 | 51.695 | 69.7 | 0.655% |
| 15 | 21 | 25 | 0.97 | 60.396 | 81.353 | 0.765% |
| 16 | 50 | 5 | 0.87 | 152.368 | 181.28 | 1.958% |
| 17 | 50 | 9 | 0.92 | 118.246 | 143.88 | 1.522% |
| 18 | 50 | 14 | 0.96 | 89.83 | 101.715 | 1.147% |
| 19 | 50 | 21 | 0.99 | 52.81 | 60.757 | 0.665% |
| 20 | 50 | 25 | 0.99 | 42.232 | 52.349 | 0.528% |

Pada tabel 4.6 dapat dilihat bahwa percobaan yang mendapatkan hasil performa terbaik adalah percobaan No 1 dengan nilai MAPE 0.377% yang menggunakan kombinasi nilai periode EMA 9 dan RSI 5. Sedangkan percobaan yang mendapatkan nilai terburuk adalah percobaan No 16, percobaan ini menghasilkan nilai MAPE sebesar 1.958% yang menggunakan nilai periode EMA 50 dan RSI 5. Hasil untuk setiap pengujian yang dilakukan dapat dilihat pada lampiran 14 sampai 33.

4.2 Analisa

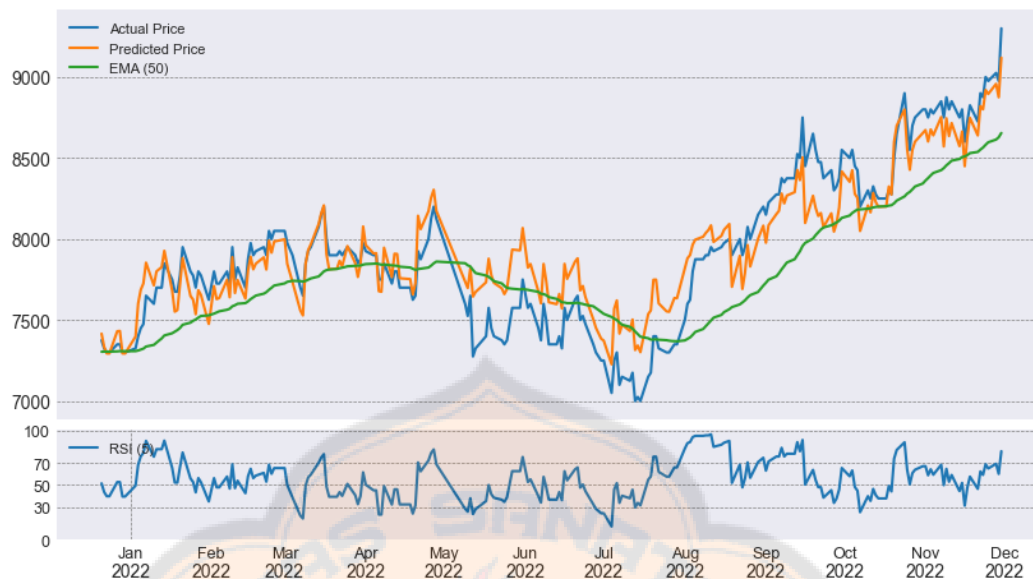
Dari 20 percobaan yang telah dilakukan menunjukkan akurasi yang sangat baik yakni nilai MAPE terendah dari keseluruhan pengujian adalah 1.958%, lihat Tabel 4.6. Untuk menentukan percobaan dengan performa terbaik maka digunakan nilai MAPE sebagai acuan utama karena nilai MAPE memberikan hasil dalam bentuk persentase terhadap nilai rata-rata kesalahan absolut dari hasil prediksi dan data sebenarnya. Sedangkan nilai MAE dan RMSE menjadi acuan seberapa besar nilai *residuals (error)* dari hasil prediksi pada percobaan yang dilakukan.

Nilai *R-Square* digunakan untuk mengetahui seberapa baik nilai EMA dan RSI saat memprediksi harga penutupan saham. Nilai *R-Square* sangat berpengaruh pada hasil nilai MAPE, semakin tinggi nilai *R-Square* maka nilai MAPE yang dihasilkan akan semakin baik. Lihat Tabel 4.7 yang memperlihatkan hasil percobaan yang diurutkan berdasarkan nilai *R-Square*.

Tabel 4.7 Hasil percobaan berdasarkan nilai *R-Square*

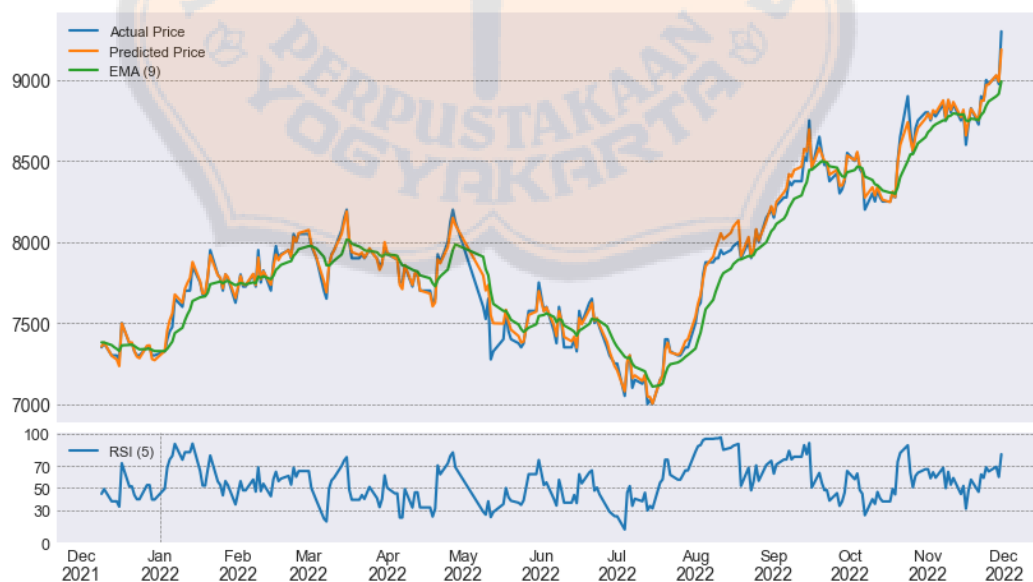
| No | EMA | RSI | R-Square | MAE | RMSE | MAPE |
|----|-----|-----|----------|---------|---------|--------|
| 1 | 9 | 5 | 0.99 | 29.619 | 45.774 | 0.377% |
| 6 | 10 | 5 | 0.99 | 30.165 | 46.554 | 0.385% |
| 13 | 21 | 14 | 0.99 | 36.806 | 50.321 | 0.466% |
| 12 | 21 | 9 | 0.99 | 36.903 | 54.781 | 0.475% |
| 7 | 10 | 9 | 0.99 | 39.801 | 54.847 | 0.504% |
| 2 | 9 | 9 | 0.99 | 41.235 | 56.146 | 0.522% |
| 20 | 50 | 25 | 0.99 | 42.232 | 52.349 | 0.528% |
| 19 | 50 | 21 | 0.99 | 52.81 | 60.757 | 0.665% |
| 14 | 21 | 21 | 0.98 | 51.695 | 69.7 | 0.655% |
| 8 | 10 | 14 | 0.98 | 51.981 | 68.825 | 0.659% |
| 3 | 9 | 14 | 0.98 | 52.498 | 69.585 | 0.666% |
| 11 | 21 | 5 | 0.97 | 59.771 | 79.263 | 0.775% |
| 15 | 21 | 25 | 0.97 | 60.396 | 81.353 | 0.765% |
| 4 | 9 | 21 | 0.97 | 60.996 | 82.717 | 0.774% |
| 9 | 10 | 21 | 0.97 | 61.551 | 83.428 | 0.781% |
| 5 | 9 | 25 | 0.97 | 65.046 | 88.12 | 0.825% |
| 10 | 10 | 25 | 0.97 | 66.189 | 89.543 | 0.839% |
| 18 | 50 | 14 | 0.96 | 89.83 | 101.715 | 1.147% |
| 17 | 50 | 9 | 0.92 | 118.246 | 143.88 | 1.522% |
| 16 | 50 | 5 | 0.87 | 152.368 | 181.28 | 1.958% |

Dapat dilihat dari Tabel 4.7 bahwa semakin tinggi nilai *R-Square* yang dihasilkan maka nilai MAPE akan semakin baik. Nilai MAPE menjadi penting karena tingkat presisi hasil prediksi yang dihasilkan dapat menentukan percobaan terbaik dari hasil penelitian ini. Untuk melihat perbedaan hasil prediksi pada percobaan yang menghasilkan *R-Square* 0.87 dan 0.99 dapat dilihat dari hasil visualisasi pada gambar 4.8 dan 4.9.



Gambar 4.8 Hasil visualisasi percobaan EMA 50 dan RSI 5

Dari gambar 4.8 dapat dilihat visualisasi *Actual Price* dengan *Predict Price* terdapat jarak yang cukup lebar. Hal ini disebabkan karena *residuals (error)* dari hasil prediksi yang cukup besar apabila dibandingkan dengan hasil dari visualisasi dari percobaan pada gambar 4.9 yang memiliki *residuals (error)* dari hasil prediksi yang lebih kecil.



Gambar 4.9 Hasil visualisasi percobaan EMA 9 dan RSI 5

Dari percobaan ini dapat disimpulkan bahwa model *Multiple Linear Regression* dengan tambahan *technical indicators* EMA dan RSI dapat memprediksi data saham PT Bank Central Asia Tbk dengan sangat baik. Hasil dari 20 percobaan yang dilakukan menunjukkan tingkat akurasi yang sangat baik, dilihat dari persentase kesalahan yang dihasilkan kurang dari 2%. Percobaan dengan performa terbaik pada penelitian ini adalah percobaan No 1 dengan nilai EMA 9 dan RSI 5. Percobaan No 1 menghasilkan nilai MAE sebesar 29.619, nilai RMSE sebesar 45.774, dan MAPE sebesar 0.377% dengan nilai *R-Square* sebesar 0.99.

Hasil percobaan yang dipaparkan pada penelitian ini dilakukan pada 30 November 2022 dimana hasil yang didapatkan bisa berbeda apabila dilakukan di lain waktu. Penulis menyarankan untuk menggunakan kombinasi nilai periode EMA dan RSI yang mendapatkan nilai *R-Square* sebesar 0.99 yang dapat dilihat pada tabel 4.7 untuk mendapatkan hasil maksimal apabila melakukan percobaan di lain waktu.

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan hasil penelitian prediksi harga penutupan saham dengan menggunakan metode *Multiple Linear Regression* (MLR), dapat diambil kesimpulan sebagai berikut.

1. Metode *Multiple Linear Regression* (MLR) dapat memprediksi harga penutupan saham PT Bank Central Asia Tbk dengan baik.
2. Model dengan performa terbaik menghasilkan nilai MAE sebesar 29.619, nilai RMSE sebesar 45.774, dan MAPE sebesar 0.377% dengan nilai *R-Square* sebesar 0.99.
3. Nilai periode dari *Exponential Moving Average* (EMA) dan *Relative Strength Index* (RSI) yang menghasilkan model dengan performa terbaik adalah EMA 9 dan RSI 5.

5.2 Saran

Agar penelitian ini dapat berkembang menjadi lebih baik lagi, maka saran yang dapat diberikan adalah:

1. Menambahkan fitur yang dapat mengetahui nilai periode terbaik dengan melakukan percobaan secara otomatis.
2. Menambahkan/mengubah *technical indicators* lain sebagai variabel independen.
3. Ditambahkan proses ranking apabila banyak menggunakan *technical indicators* untuk menentukan *technical indicators* terbaik.
4. Kembangkan metode yang dapat memprediksi harga penutupan dalam beberapa waktu kedepan (mingguan atau bulanan).

DAFTAR PUSTAKA

- Agrawal, M., Khan, A. U., & Shukla, P. K. (2019). Stock price prediction using technical indicators: A predictive model using optimal deep learning. *Learning*, 6(2), 7.
- Asthri, D. D. P., Topowijono, T., & Sulasmiyati, S. (2016). Analisis Teknikal Dengan Indikator Moving Average Convergence Divergence Untuk Menentukan Sinyal Membeli Dan Menjual Dalam Perdagangan Saham (Studi Pada Perusahaan Sub Sekto Makanan Dan Minuman Di Bei Tahun 2013-2015) (Doctoral dissertation, Brawijaya University).
- Chai, T. and Draxler, R.R. (2014) Root Mean Square Error (RMSE) or Mean Absolute Error (MAE)?—Arguments against Avoiding RMSE in the Literature. *Geoscientific Model Development*, 7, 1247-1250.
- Deni, M., & Latifah, R. (2022). PREDIKSI PENGISIAN BBM HSD DENGAN METODE MULTIPLE LINEAR REGRESSION. *JUST IT: Jurnal Sistem Informasi, Teknologi Informasi dan Komputer*, 11(3).
- El-Khodary, I. A. (2009). A decision support system for technical analysis of financial markets based on the moving average crossover. *World Applied Sciences Journal*, 6(11), 1457-1472.
- Faruqhy, M. N., Andreswari, D., & Sari, J. P. (2021). Prediksi Prestasi Nilai Akademik Mahasiswa Berdasarkan Jalur Masuk Perguruan Tinggi Menggunakan Metode Multiple Linear Regression (Studi Kasus: Fakultas Teknik Universitas Bengkulu). *Rekursif: Jurnal Informatika*, 9(2), 172-183.
- Gaol, I. L. L., Sinurat, S., & Siagian, E. R. (2019). Implementasi Data Mining Dengan Metode Regresi Linear Berganda Untuk Memprediksi Data Persediaan Buku Pada Pt. Yudhistira Ghalia Indonesia Area Sumatera Utara. *KOMIK (Konferensi Nasional Teknologi Informasi dan Komputer)*, 3(1).

- Gururaj, V., Shriya, V. R., & Ashwini, K. (2019). Stock market prediction using linear regression and support vector machines. *Int. J. Appl. Eng. Res.*, 14(8), 1931-1934.
- Izzah, A., & Widyastuti, R. (2017). Prediksi Harga Saham Menggunakan Improved Multiple Linear Regression Untuk Pencegahan Data Outlier. *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control*, 141-150.
- Khotimah, T., & Nindyasari, R. (2017). Forecasting Dengan Metode Regresi Linier Pada Sistem Penunjang Keputusan Untuk Memprediksi Jumlah Penjualan Batik (Studi Kasus Kub Sarwo Endah Batik Tulis Lasem). *Jurnal Mantik Penusa*, 1(1).
- Kustina, L., & Safitri, O. (2019). Kebijakan Deviden Dan Capital Gain: Pengaruhnya Terhadap Harga Saham: Devident Policy And Capital Gain: The Effect On Stock Prices. *Jurnal Investasi*, 5(1), 24-37.
- Nabillah, I., & Ranggadara, I. (2020). Mean Absolute Percentage Error untuk Evaluasi Hasil Prediksi Komoditas Laut. *Journal of Information System*, 5(2), 250-255.
- Nafi'iyah, N. (2016). Perbandingan regresi linear, backpropagation dan fuzzy mamdani dalam prediksi harga emas. *Prosiding SENIATI*, 291-B.
- Oktavia, I., & Nugraha, K. G. S. (2018). Faktor-Faktor Yang Mempengaruhi Harga Saham. *UNEJ e-Proceeding*.
- Putri, F. K., & Tjahjono, J. D. (2011). Pengaruh Analisis Fundamental Terhadap Pergerakan Harga Saham. *Jurnal Bisnis Indonesia Vol*, 2(1).
- Rolf. (2018) . How To Use Moving Averages – Moving Average Trading 101. <https://tradeciety.com/how-to-use-moving-averages>, Diakses pada 10 November 2022.

- Sutoyo, E., & Almaarif, A. (2020). Educational Data Mining for Predicting Student Graduation Using the Naïve Bayes Classifier Algorithm. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 4(1), 95-101.
- Syahra, Y., Santoso, I., & Kustini, R. (2019, August). Implementasi Data Mining Untuk Prediksi Angka Kelahiran Bayi Pada Desa Sibolangit Menggunakan Multi Regresi. In *Seminar Nasional Sains dan Teknologi Informasi (SENSASI)* (Vol. 2, No. 1).
- Tsai, C. F., & Wang, S. P. (2009, March). Stock price forecasting by hybrid machine learning techniques. In *Proceedings of the international multiconference of engineers and computer scientists* (Vol. 1, No. 755, p. 60).
- Vargas, M. R., Dos Anjos, C. E., Bichara, G. L., & Evsukoff, A. G. (2018, July). Deep leaning for stock market prediction using technical indicators and financial news articles. In *2018 international joint conference on neural networks (IJCNN)* (pp. 1-8). IEEE.
- Vijh, M., Chandola, D., Tikkiwal, V. A., & Kumar, A. (2020). Stock closing price prediction using machine learning techniques. *Procedia computer science*, 167, 599-606.
- Waldi, L. (2020). Analisis Teknikal Pergerakan Harga Saham Untuk Menentukan Keputusan Investasi (Studi Pada Perusahaan Bumn Sub Sektor Kontruksi Dan Bangunan Yang Terdaftar Di Bursa Efek Indonesia) (Doctoral dissertation, Universitas Islam Negeri Sultan Syarif Kasim Riau).
- Wulandari, N. L. P., Sarja, S. T., Yuniastari, N. L. A. K., Saryanti, S., & Desi, I. G. A. (2015). Prediksi jumlah pelanggan dan persediaan barang menggunakan metode regresi linier berganda pada bali orchid. *Jurnal Online Sistem Informasi*, 1(1), 243347.
- Yuliara, I. M. (2016). *Regresi Linier Berganda*. Denpasar: Universitas Udayana.

LAMPIRAN

Lampiran 1 Source code fungsi `__init__()`

```

26 def __init__(self):
27     # Build UI
28     topLevel = tk.Tk()
29
30     self.r2 = StringVar()
31     self.mae = StringVar()
32     self.rmse = StringVar()
33     self.mape = StringVar()
34
35     labelframe1 = tk.LabelFrame(topLevel)
36     frame1 = tk.Frame(labelframe1)
37     entry1 = tk.Entry(frame1)
38     entry1.configure(justify="center", width=15)
39     entry1.insert("0", "DATE")
40     entry1["state"] = "readonly"
41     entry1.grid(column=0, row=0)
42     entry2 = tk.Entry(frame1)
43     entry2.configure(justify="center", width=15)
44     entry2.insert("0", "CLOSE PRICE")
45     entry2["state"] = "readonly"
46     entry2.grid(column=1, row=0)
47     entry3 = tk.Entry(frame1)
48     entry3.configure(justify="center", width=15)
49     entry3.insert("0", "PREDICT PRICE")
50     entry3["state"] = "readonly"
51     entry3.grid(column=2, row=0)
52     frame1.configure(height=200, width=200)
53     frame1.pack(expand="false", fill="y", padx=10,
54                pady=3, side="top")
55     self.tkscrolledframe1 = TkScrolledFrame(labelframe1,
56                                           scrolltype="vertical")
57     self.tkscrolledframe1.configure(usemousewheel=True)
58     self.tkscrolledframe1.pack(expand="true", fill="both",
59                               padx=10, side="top")
60     labelframe1.configure(height=200, width=200)
61     labelframe1.pack(fill="y", padx=10, pady=10, side="left")
62
63     labelframe2 = tk.LabelFrame(topLevel)
64     frame2 = tk.Frame(labelframe2)
65     label1 = tk.Label(frame2)
66     label1.configure(padx=20, pady=5, text="SYMBOL")
67     label1.grid(column=0, row=0)
68     label2 = tk.Label(frame2)
69     label2.configure(padx=20, pady=5,
70                    text="Exponential Moving Average")
71     label2.grid(column=0, row=1)

```



```
72
73 label3 = tk.Label(frame2)
74 label3.configure(padx=20, pady=5,
75                 text="Relative Strength Index")
76 label3.grid(column=0, row=2)
77 self.entry4 = tk.Entry(frame2)
78 self.entry4.grid(column=1, row=0)
79 self.entry5 = tk.Entry(frame2)
80 self.entry5.grid(column=1, row=1)
81 self.entry6 = tk.Entry(frame2)
82 self.entry6.grid(column=1, row=2)
83 label4 = tk.Label(frame2)
84 label4.configure(padx=20, pady=5,
85                 text="Coefficient of Determination")
86 label4.grid(column=4, row=0)
87 label5 = tk.Label(frame2)
88 label5.configure(padx=20, pady=5,
89                 text="Mean Absolute Error")
90 label5.grid(column=4, row=1)
91 label6 = tk.Label(frame2)
92 label6.configure(padx=20, pady=5,
93                 text="Root Mean Square Error")
94 label6.grid(column=4, row=2)
95 label7 = tk.Label(frame2)
96 label7.configure(padx=20, pady=5,
97                 text="Mean Absolute Percentage Error")
98 label7.grid(column=4, row=3)
99 entry7 = tk.Entry(frame2, textvariable=self.r2)
100 entry7.configure(state="readonly")
101 entry7.grid(column=5, row=0)
102 entry8 = tk.Entry(frame2, textvariable=self.mae)
103 entry8.configure(state="readonly")
104 entry8.grid(column=5, row=1)
105 entry9 = tk.Entry(frame2, textvariable=self.rmse)
106 entry9.configure(state="readonly")
107 entry9.grid(column=5, row=2)
108 entry10 = tk.Entry(frame2, textvariable=self.mape)
109 entry10.configure(state="readonly")
110 entry10.grid(column=5, row=3)
111 button1 = tk.Button(frame2, command=self.check)
112 button1.configure(borderwidth=1, text="FORECAST")
113 button1.grid(column=1, pady=10, row=3, sticky="n")
114 frame2.configure(height=200, width=200)
115 frame2.pack(padx=10, pady=10, side="top")
116 labelframe2.configure(height=200, width=200)
117 labelframe2.pack(fill="x", padx=10, pady=10, side="top")
118
119 self.labelframe3 = tk.LabelFrame(topLevel)
120 self.labelframe3.configure(height=200, width=200)
121 self.labelframe3.pack(expand="true", fill="both",
122                      padx=10, pady=10, side="top")
```

```
123
124 topLevel.state('zoomed')
125 topLevel.iconbitmap('USD100.ico')
126 topLevel.resizable(True, True)
127 topLevel.title("Stock Forcast")
128
129 topLevel.mainloop()
```

Lampiran 2 Source code fungsi check()

```
131 def check(self):
132     symbolVal = self.entry4.get()
133     emaVal = self.entry5.get()
134     rsiVal = self.entry6.get()
135
136     if symbolVal == '' or symbolVal.find(',') != -1:
137         messagebox.showerror('Error',
138             f' -{symbolVal}: No Data Found, Symbol May Be Delisted!')
139         self.entry4.delete(0, 'end')
140     if not emaVal.isdigit() or emaVal == '' or int(emaVal) < 1:
141         self.entry5.delete(0, 'end')
142     if not rsiVal.isdigit() or rsiVal == '' or int(rsiVal) < 1:
143         self.entry6.delete(0, 'end')
144     else:
145         if not emaVal.isdigit() or emaVal == '' or int(emaVal) < 1:
146             messagebox.showerror('Error', f'{emaVal} Is Invalid!')
147             self.entry5.delete(0, 'end')
148             if not rsiVal.isdigit() or rsiVal == '' or int(rsiVal) < 1:
149                 self.entry6.delete(0, 'end')
150         elif not rsiVal.isdigit() or rsiVal == '' or int(rsiVal) < 1:
151             messagebox.showerror('Error', f'{rsiVal} Is Invalid!')
152             self.entry6.delete(0, 'end')
153     else:
```

Lampiran 3 Source code untuk proses data retrieval

```
154 # Data Retrieval
155     print("\n", '=' * 70, "\n")
156     data = yf.download(
157         tickers=symbolVal,
158         period="5y",
159         interval="1d")
160
161     if len(data) == 0:
162         messagebox.showerror('Error',
163             f'-{symbolVal}: No Data Found, Symbol May Be Delisted!')
164         self.entry4.delete(0, 'end')
165     else:
166         print(f"""\n{'=' * 70}
167         \nCONTOH DATA YANG BERHASIL DIUNDUH
168         \n{data}
169         \n{'=' * 70}""")
170
171         self.lenEma = int(emaVal)
172         self.lenRsi = int(rsiVal)
173         self.forcast(data)
```



Lampiran 4 Contoh data dari proses data retrieval

```
=====
[*****100%*****] 1 of 1 completed
=====

CONTOH DATA YANG BERHASIL DIUNDUH

      Open   High   Low   Close   Adj Close   Volume
Date
2017-11-30 4260.0 4260.0 4070.0 4070.0 3758.374268 205897500.0
2017-12-01 4070.0 4070.0 4070.0 4070.0 3758.374268      0.0
2017-12-04 4165.0 4230.0 4130.0 4160.0 3841.483887 128402500.0
2017-12-05 4210.0 4225.0 4180.0 4200.0 3878.421387 62670500.0
2017-12-06 4230.0 4275.0 4200.0 4260.0 3933.827148 106299500.0
...
2022-11-24 8975.0 9000.0 8950.0 9000.0 8965.000000 53520800.0
2022-11-25 9000.0 9000.0 8900.0 8975.0 8940.096680 75048300.0
2022-11-28 9050.0 9100.0 8975.0 9025.0 8989.902344 96209700.0
2022-11-29 8975.0 9075.0 8925.0 8975.0 8940.096680 91562100.0
2022-11-30 8875.0 9400.0 8850.0 9300.0 9263.833008 644359600.0

[1254 rows x 6 columns]
```

Lampiran 5 Source code untuk proses Pre-processing

```
175 def forecast(self, stock_data):
176     # Preprocessing
177     print(f"""\nJUMLAH DATA KOSONG
178     \n{stock_data.isnull().sum()}
179     \n{'=' * 70}
180     \nJUMLAH DATA DENGAN VOLUME 0 : {(stock_data['Volume'] == 0).sum()}
181     \n{'=' * 70}""")
182
183     # Data Cleaning
184     df = stock_data.dropna()
185     df = df[df['Volume'] != 0]
186     # Data Selection
187     df = df[['Close']]
188
189     print(f"""\nCONTOH DATA HASIL PRE-PROCESSING
190     \n{df}
191     \n{'=' * 70}""")
```

Lampiran 6 Contoh data dari proses Pre-processing

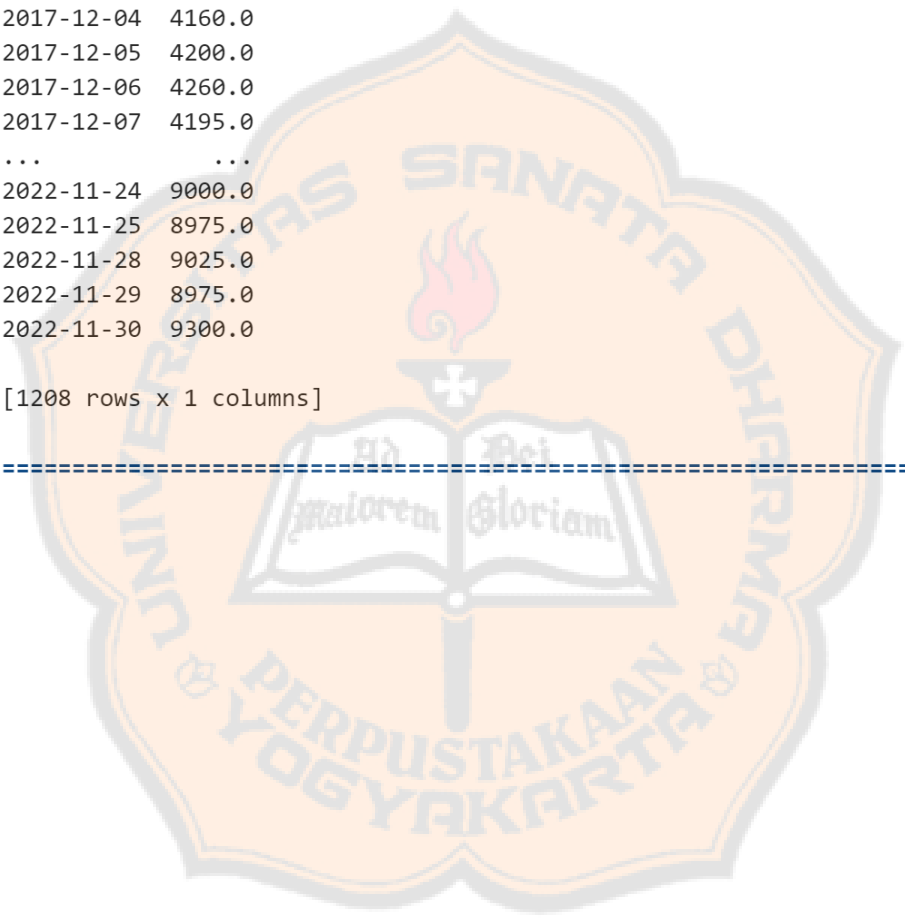
=====

CONTOH DATA HASIL PRE-PROCESSING

| Date | Close |
|------------|--------|
| 2017-11-30 | 4070.0 |
| 2017-12-04 | 4160.0 |
| 2017-12-05 | 4200.0 |
| 2017-12-06 | 4260.0 |
| 2017-12-07 | 4195.0 |
| ... | ... |
| 2022-11-24 | 9000.0 |
| 2022-11-25 | 8975.0 |
| 2022-11-28 | 9025.0 |
| 2022-11-29 | 8975.0 |
| 2022-11-30 | 9300.0 |

[1208 rows x 1 columns]

=====



Lampiran 7 Source code untuk proses perhitungan technical indicators.

```
193 # Technical Indicators
194 # EMA
195 df['EMA'] = df.ta.ema(close='Close', length=self.lenEma)
196 # RSI
197 df['RSI'] = df.ta.rsi(close='Close', length=self.lenRsi)
198
199 print(f"""\nCONTOH DATA HASIL PERHITUNGAN TECHNICAL INDICATORS
200 \n EMA : {self.lenEma} & RSI : {self.lenRsi}
201 \n{df}
202 \n{'=' * 70}""")
203 print(f"""\nJUMLAH DATA KOSONG
204 \n{df.isnull().sum()}
205 \n{'=' * 70}""")
206
207 df = df.dropna()
208
209 print(f"""\nCONTOH DATA HASIL PERHITUNGAN TECHNICAL INDICATORS
210 \nSETELAH DATA KOSONG DIHAPUS
211 \n EMA : {self.lenEma} & RSI : {self.lenRsi}
212 \n{df}
213 \n{'=' * 70}""")
```

Lampiran 8 Contoh data dari proses perhitungan technical indicators.

=====

CONTOH DATA HASIL PERHITUNGAN TECHNICAL INDICATORS

SETELAH DATA KOSONG DIHAPUS

EMA : 9

RSI : 5

| Date | Close | EMA | RSI |
|------------|--------|-------------|-----------|
| 2017-12-13 | 4185.0 | 4195.555556 | 48.518139 |
| 2017-12-14 | 4220.0 | 4200.444444 | 58.137119 |
| 2017-12-15 | 4220.0 | 4204.355556 | 58.137119 |
| 2017-12-18 | 4220.0 | 4207.484444 | 58.137119 |
| 2017-12-19 | 4220.0 | 4209.987556 | 58.137119 |
| ... | ... | ... | ... |
| 2022-11-24 | 9000.0 | 8841.329870 | 69.050690 |
| 2022-11-25 | 8975.0 | 8868.063896 | 65.140170 |
| 2022-11-28 | 9025.0 | 8899.451117 | 69.463564 |
| 2022-11-29 | 8975.0 | 8914.560893 | 60.140168 |
| 2022-11-30 | 9300.0 | 8991.648715 | 80.933210 |

[1200 rows x 3 columns]

=====

Lampiran 9 Source code untuk proses Train-test Split.

```

215 # Train-test Split
216 X = df[['EMA', 'RSI']]
217 y = df['Close']
218 t = int(0.8*len(df))
219 X_train, y_train, X_test, y_test = X[:t], y[:t], X[t:], y[t:]
220
221 print(f"""\nJUMLAH DATA TRAIN & TEST\n
222 Data Train : {len(X_train)}
223 Data Test  : {len(X_test)}
224 \n{'=' * 70}""")

```

Lampiran 10 Source code untuk proses pembentukan model MLR.

```
226 # Modelling MLR
227     model = LinearRegression()
228     model.fit(X_train, y_train)
229
230     print(f"""\nNILAI KONSTANTA & KOEFISIEN DARI MODEL MLR\n
231     Konstanta : {model.intercept_}
232     Koefisien 1 : {model.coef_[0]}
233     Koefisien 2 : {model.coef_[1]}
234     \n{'=' * 70}""")
```

Lampiran 11 Source code untuk proses prediction.

```
236 # Predict
237     y_pred = model.predict(X_test)
238
239     predict_df = pd.DataFrame(list(zip(y_test, y_pred)),
240                               index=y_test.index,
241                               columns=['Close', 'Pred_Close'])
242
243     print(f"""\nCONTOH DATA HASIL PREDIKSI
244     \n{predict_df}
245     \n{'=' * 70}""")
```


Lampiran 12 Contoh data hasil prediksi

=====

CONTOH DATA HASIL PREDIKSI

| Date | Close | Pred_Close |
|------------|--------|-------------|
| 2021-12-09 | 7350.0 | 7355.482036 |
| 2021-12-10 | 7375.0 | 7379.501294 |
| 2021-12-13 | 7300.0 | 7298.785886 |
| 2021-12-14 | 7300.0 | 7285.930598 |
| 2021-12-15 | 7300.0 | 7275.646367 |
| ... | ... | ... |
| 2022-11-24 | 9000.0 | 8968.616440 |
| 2022-11-25 | 8975.0 | 8973.083395 |
| 2022-11-28 | 9025.0 | 9029.611609 |
| 2022-11-29 | 8975.0 | 8991.215314 |
| 2022-11-30 | 9300.0 | 9188.583044 |

[240 rows x 2 columns]

=====

Lampiran 13 Source code untuk proses performance evaluation.

```

247 # Validation
248 # R-Square
249 sse = np.sum((predict_df['Pred_Close']-predict_df['Close'])**2)
250 sst = np.sum((predict_df['Close']-np.mean(predict_df['Close']))**2)
251 r2 = 1-sse/sst
252 # MAE
253 mae = np.mean(np.abs(predict_df['Close']-predict_df['Pred_Close']))
254 # RMSE
255 mse = np.mean((predict_df['Close']-predict_df['Pred_Close'])**2)
256 rmse = math.sqrt(mse)
257 # MAPE
258 mape = np.mean(np.abs((predict_df['Close']-predict_df['Pred_Close'])/predict_df['Close']))
259
260 print(f"""\nHASIL EVALUASI\n
261 Coefficient of Determination\t= {r2}
262 Mean Absolute Error\t\t= {mae}
263 Root Mean Square Error\t\t= {rmse}
264 Mean Absolute Percentage Error\t= {mape}
265 \n{'=' * 70}""")

```

Lampiran 14 Source code Update hasil prediksi pada User Interface.

```

267 # Update Predict Results in UI
268
269 # Update Validation
270 self.val1.set(round(r2, 2))
271 self.val2.set(round(mae, 3))
272 self.val3.set(round(rmse, 3))
273 self.val4.set(f'{round(mape*100, 3)} %')
274
275 # Update Table
276 for widgets in self.tkscrolledframe1.innerframe.winfo_children():
277     widgets.destroy()
278 predict = predict_df.sort_index(ascending=False)
279 for i in range(len(predict)):
280     entry = tk.Entry(self.tkscrolledframe1.innerframe)
281     entry.configure(justify="center", width=15)
282     entry.insert("0", predict.index[i].date())
283     entry["state"] = "readonly"
284     entry.grid(column=0, row=i+1)
285     for j in range(2):
286         entry = tk.Entry(self.tkscrolledframe1.innerframe)
287         entry.configure(justify="center", width=15)
288         entry.insert("0", round(predict.iloc[i][j], 1))
289         entry["state"] = "readonly"
290         entry.grid(column=j+1, row=i+1)
291
292 # Update Chart
293 for widgets in self.labelframe3.winfo_children():
294     widgets.destroy()
295 fmt_month = mdates.MonthLocator()
296 fmt_year = mdates.YearLocator()
297 fig = plt.figure()
298 ax1 = plt.subplot2grid((10, 1), (1, 0), rowspan=7, colspan=1)
299 ax2 = plt.subplot2grid((10, 1), (8, 0), rowspan=3, colspan=1)
300 ax1.plot(predict_df.index, predict_df['Close'])
301 ax1.plot(predict_df['Pred_Close'])
302 ax1.plot(X_test['EMA'])
303 ax1.legend(['Actual Price', 'Predicted Price', f'EMA ({self.lenEma})'],
304           loc='upper left', fontsize=8)
305 ax1.set_ylabel('', fontsize=9)
306 ax1.grid(color='grey', linestyle='--', linewidth=0.5)
307 ax1.get_xaxis().set_visible(False)
308 ax2.plot(X_test['RSI'])
309 ax2.legend([f'RSI ({self.lenRsi})'], loc='upper left', fontsize=8)
310 ax2.set_yticks([30, 50, 70])
311 ax2.grid(color='grey', linestyle='--', linewidth=0.5)
312 ax2.xaxis.set_minor_locator(fmt_month)
313 ax2.xaxis.set_minor_formatter(mdates.DateFormatter('%b'))
314 ax2.xaxis.set_major_locator(fmt_year)
315 ax2.xaxis.set_major_formatter(mdates.DateFormatter('%b'))
316 ax2.tick_params(labelsize=9, which='both')
317 sec_xaxis = ax2.secondary_xaxis(-0.25)
318 sec_xaxis.xaxis.set_major_locator(fmt_month)
319 sec_xaxis.xaxis.set_major_formatter(mdates.DateFormatter('%Y'))

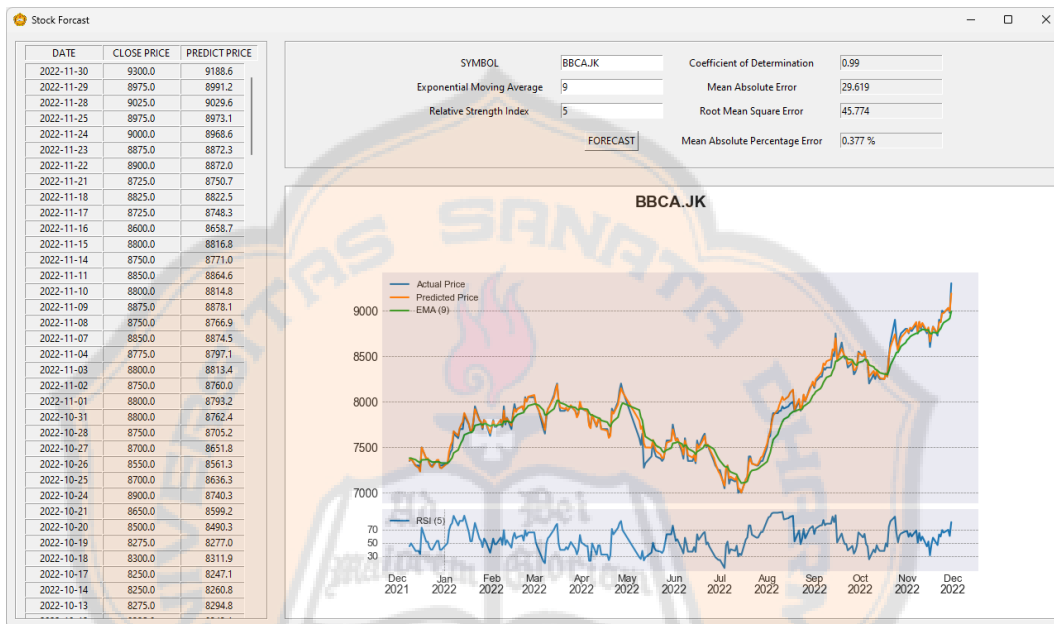
```

```

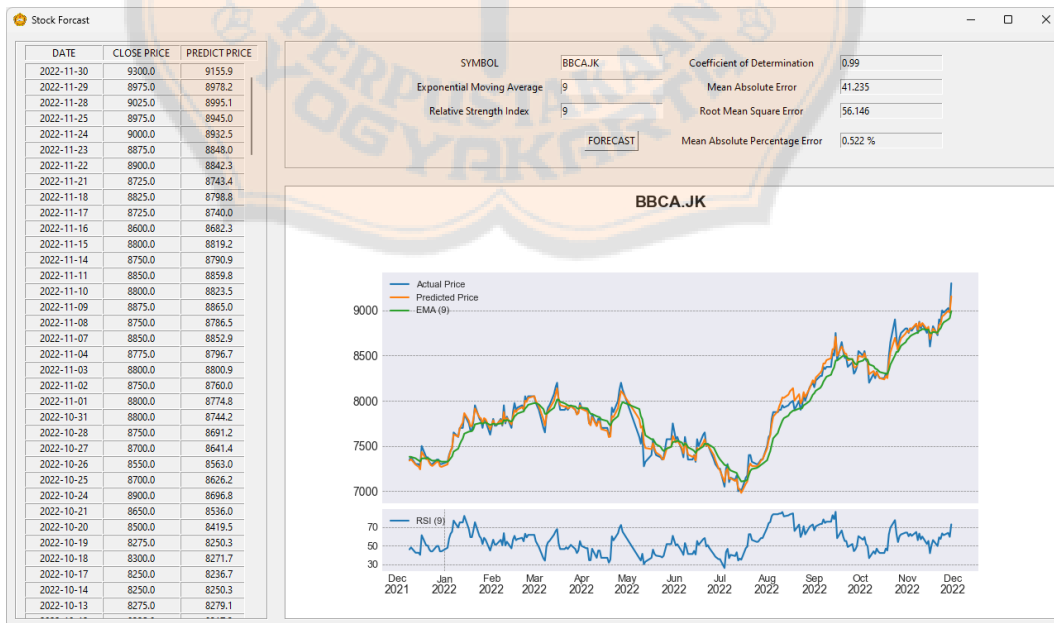
320     sec_axis.spines['bottom'].set_visible(False)
321     sec_axis.tick_params(length=0, labels=10)
322     fig.suptitle(self.entry4.get().upper(), fontsize=14, fontweight='bold')
323     canvas1 = FigureCanvasTkAgg(fig, master=self.labelframe3)
324     canvas1.draw()
325     canvas1.get_tk_widget().pack(fill="both", expand=True)

```

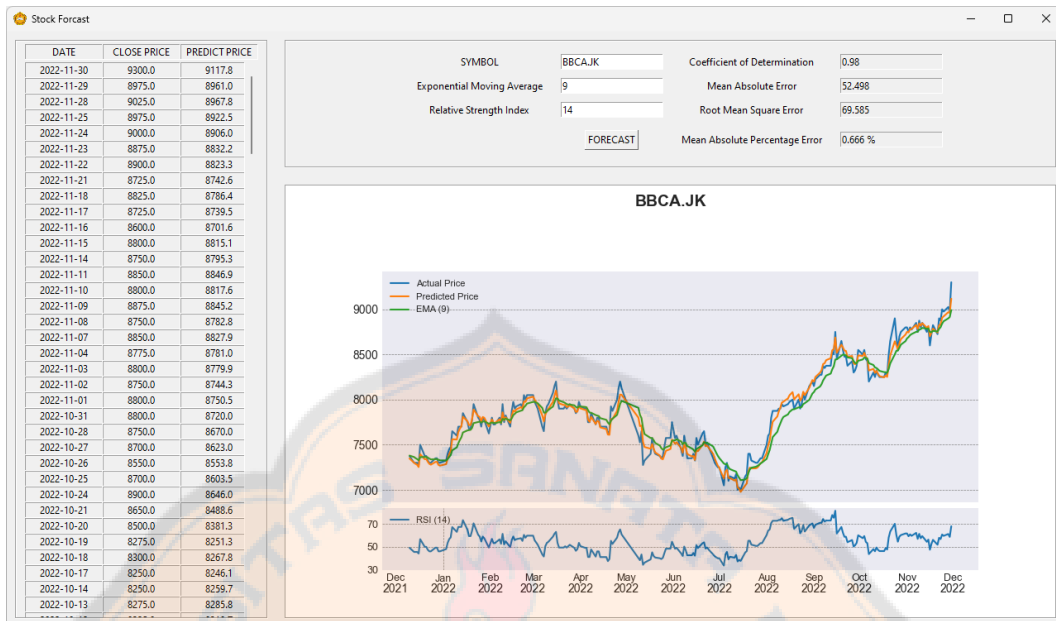
Lampiran 15 Hasil percobaan periode EMA 9 dan periode RSI 5



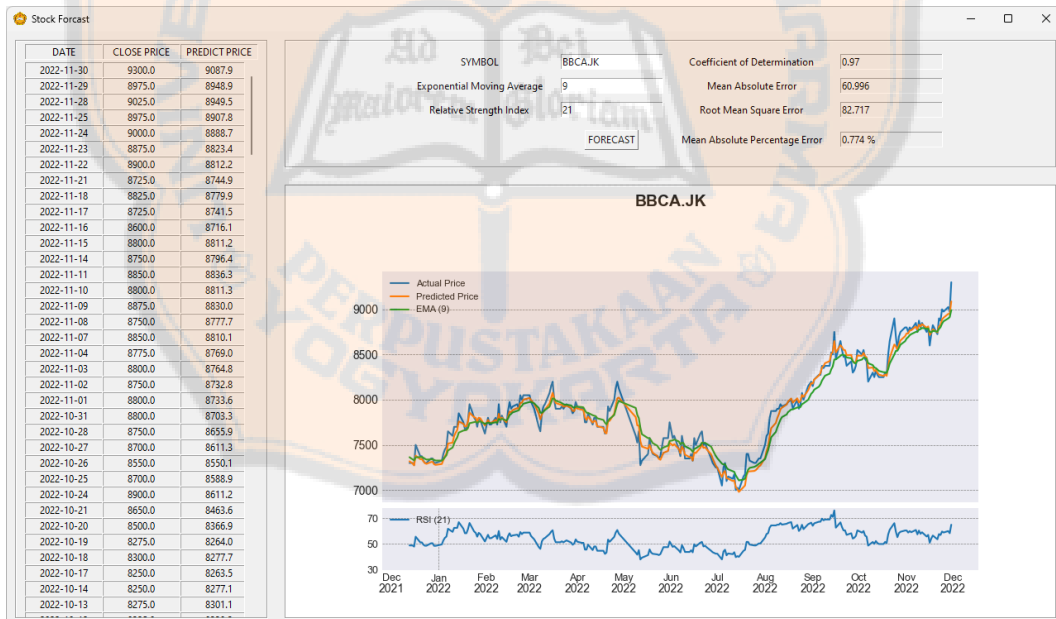
Lampiran 16 Hasil percobaan periode EMA 9 dan periode RSI 9



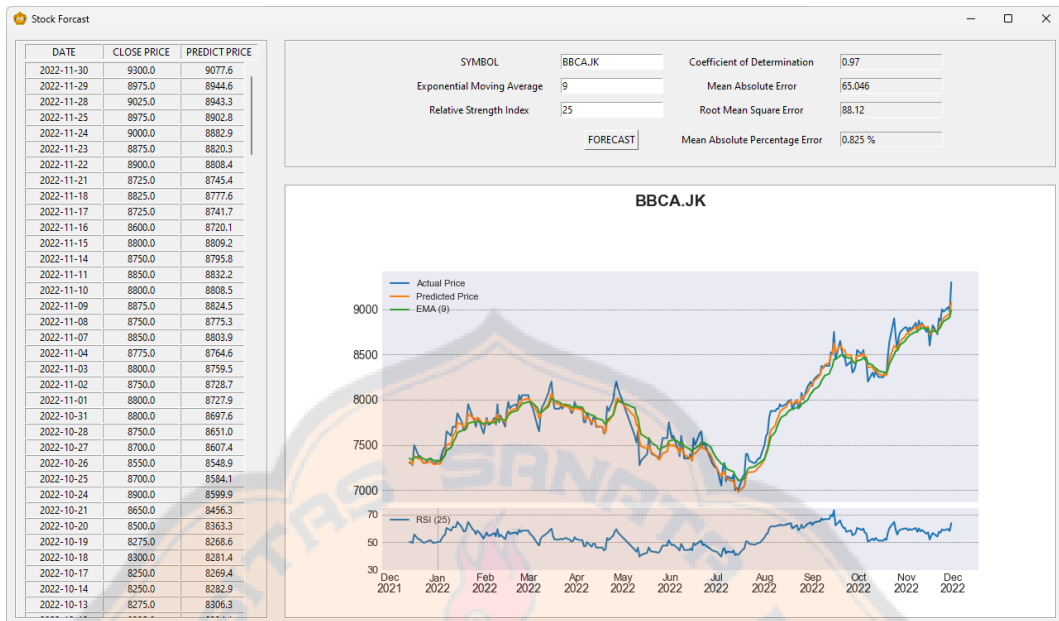
Lampiran 17 Hasil percobaan periode EMA 9 dan periode RSI 14



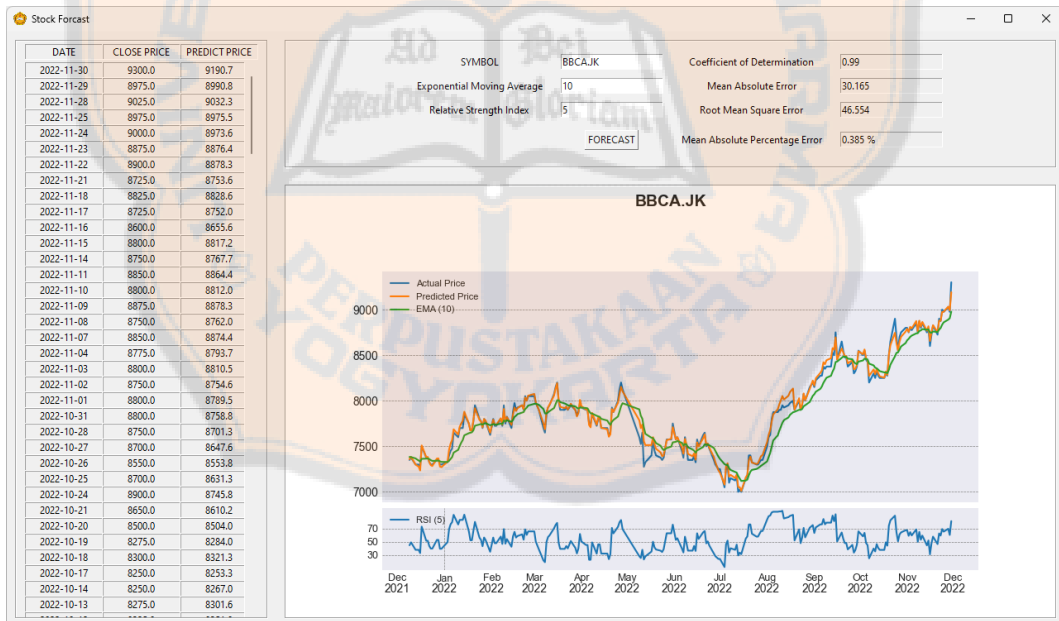
Lampiran 18 Hasil percobaan periode EMA 9 dan periode RSI 21



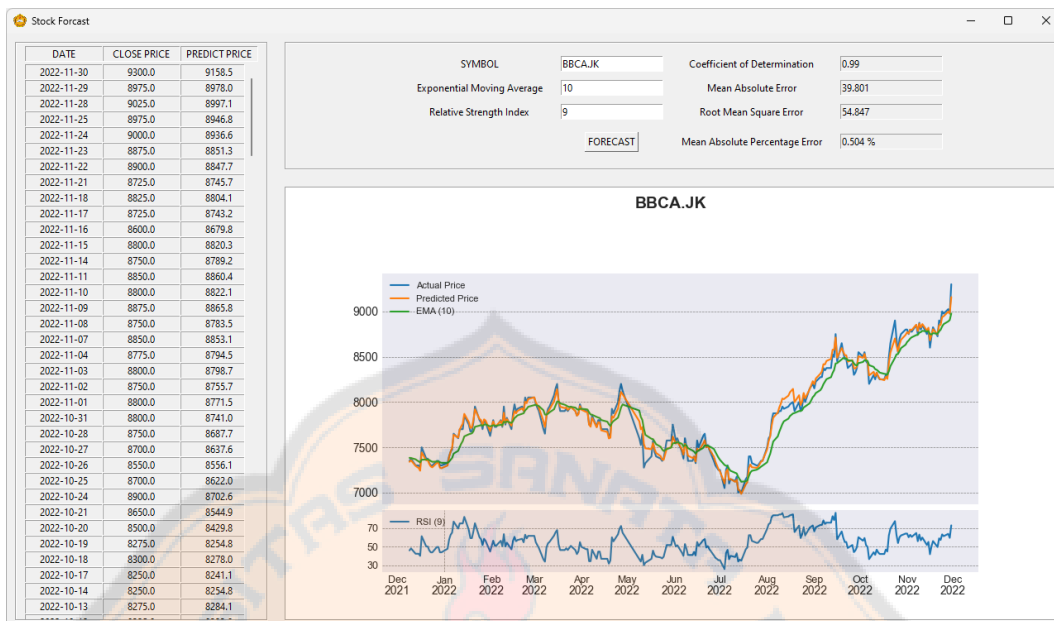
Lampiran 19 Hasil percobaan periode EMA 9 dan periode RSI 25



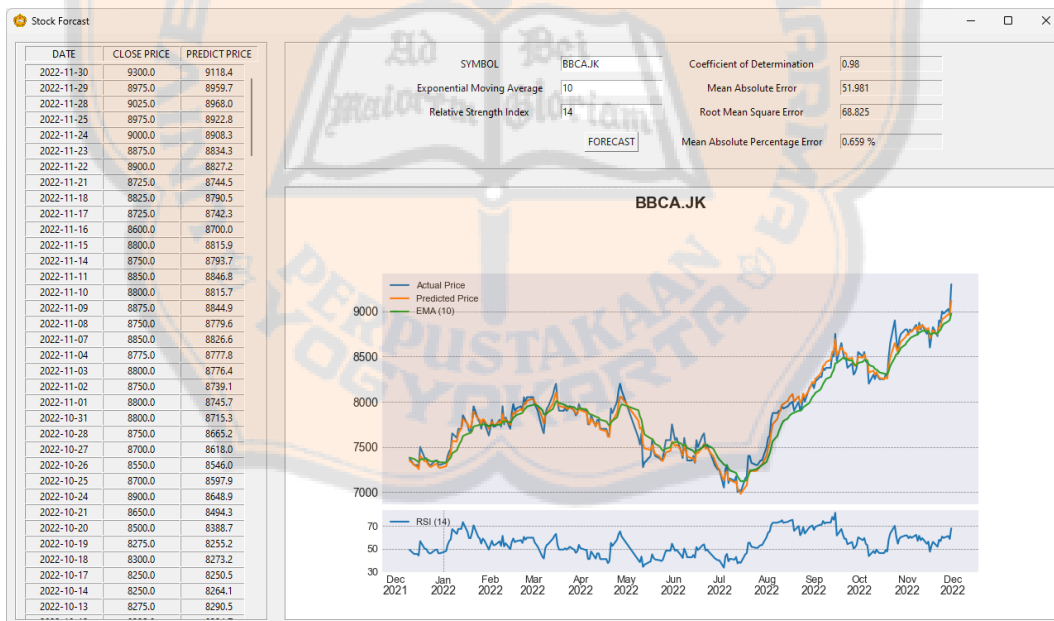
Lampiran 20 Hasil percobaan periode EMA 10 dan periode RSI 5



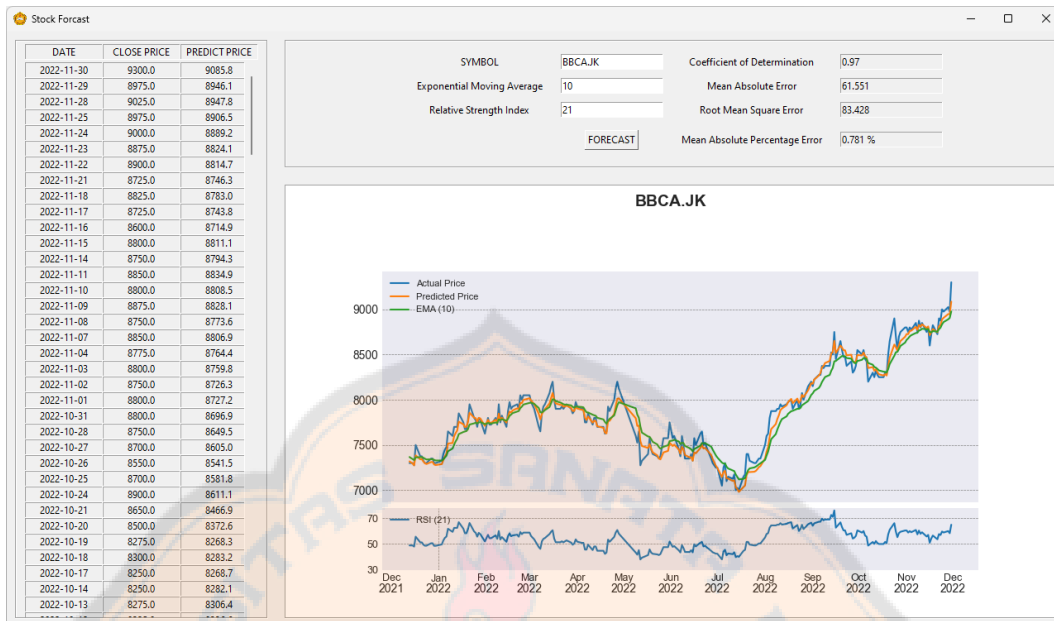
Lampiran 21 Hasil percobaan periode EMA 10 dan periode RSI 9



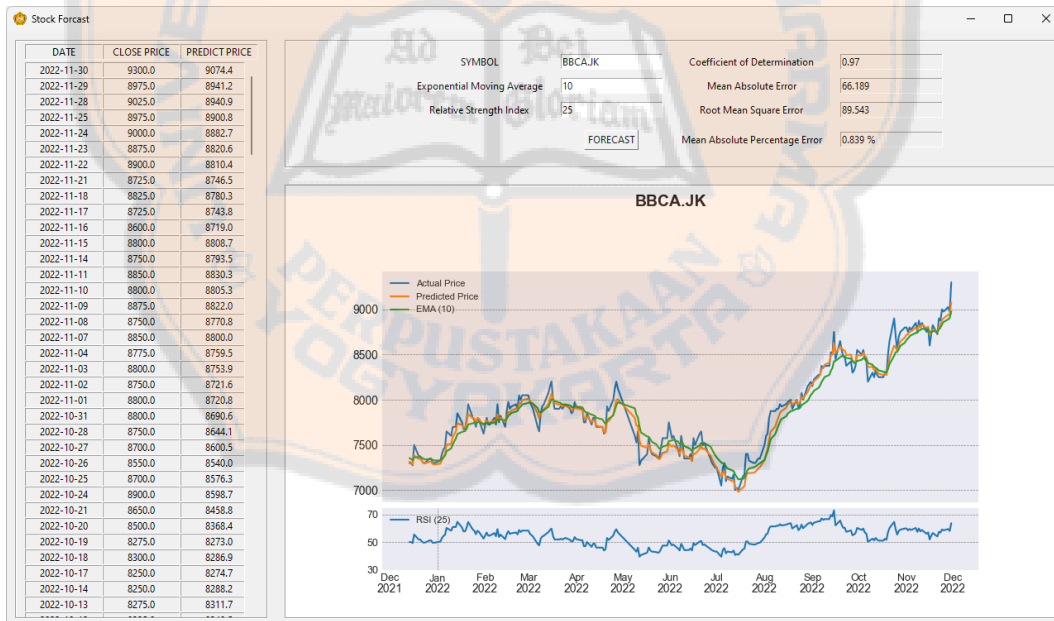
Lampiran 22 Hasil percobaan periode EMA 10 dan periode RSI 14



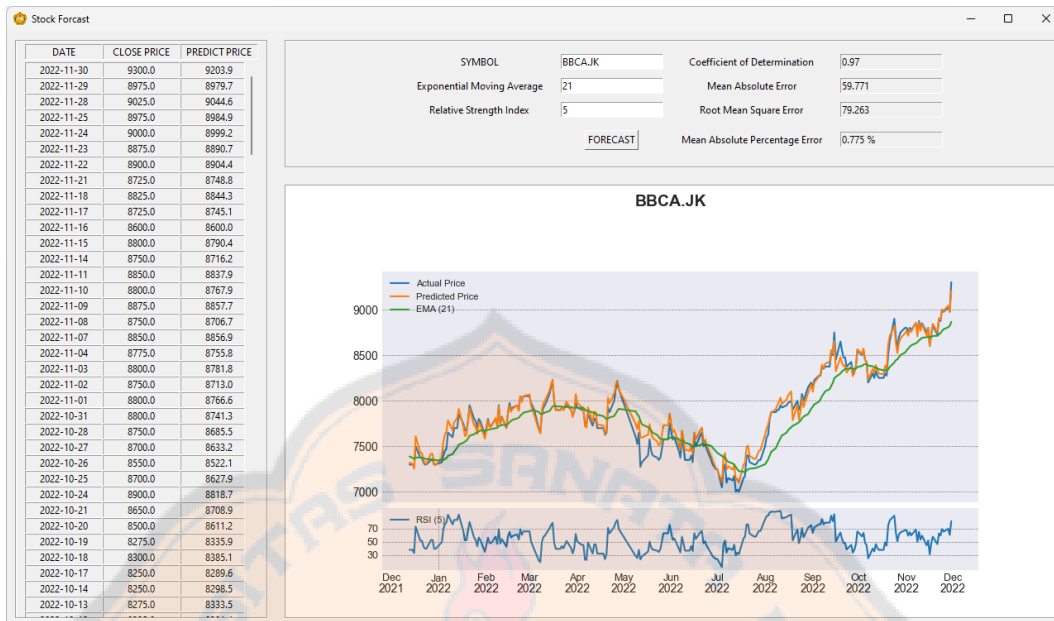
Lampiran 23 Hasil percobaan periode EMA 10 dan periode RSI 21



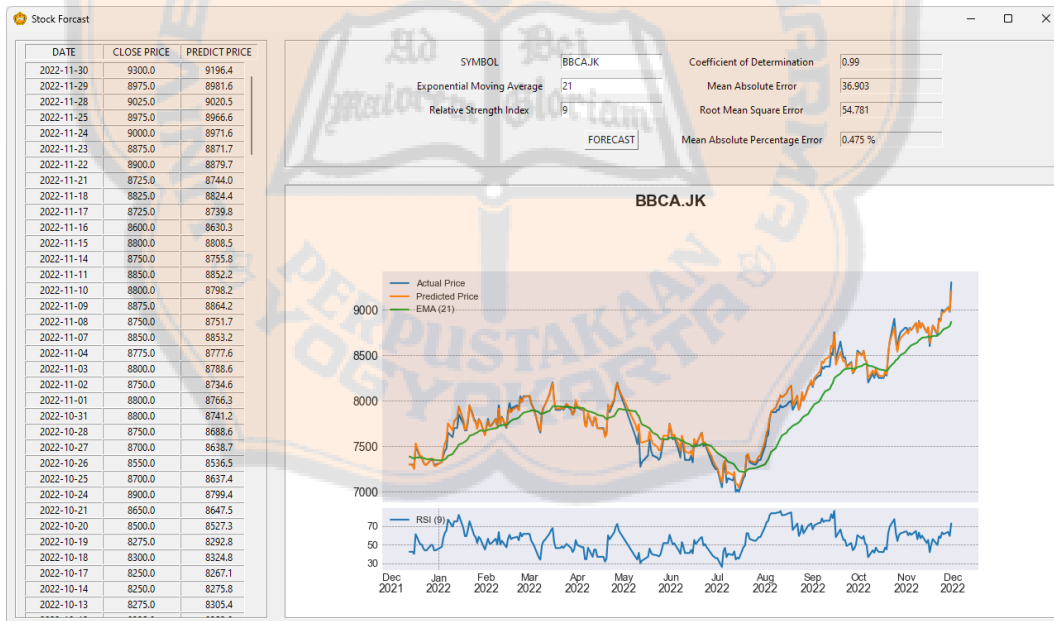
Lampiran 24 Hasil percobaan periode EMA 10 dan periode RSI 25



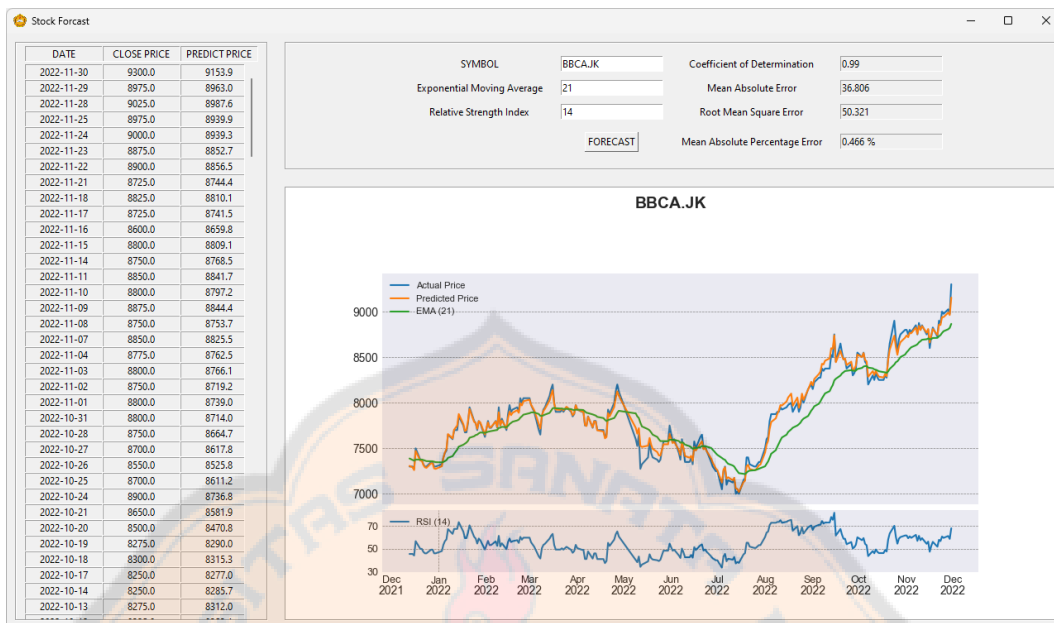
Lampiran 25 Hasil percobaan periode EMA 21 dan periode RSI 5



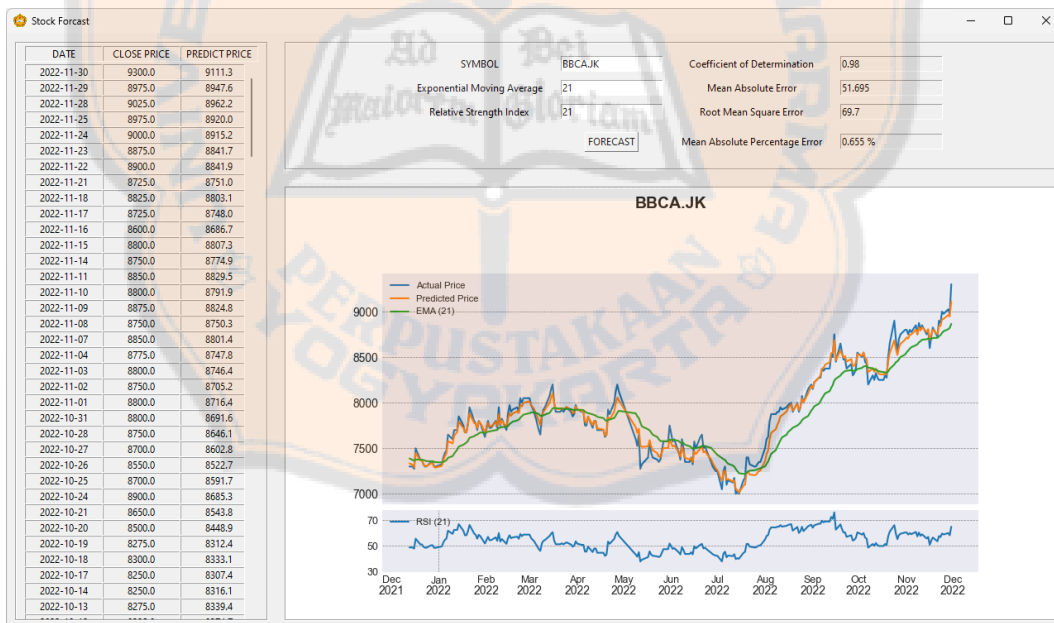
Lampiran 26 Hasil percobaan periode EMA 21 dan periode RSI 9



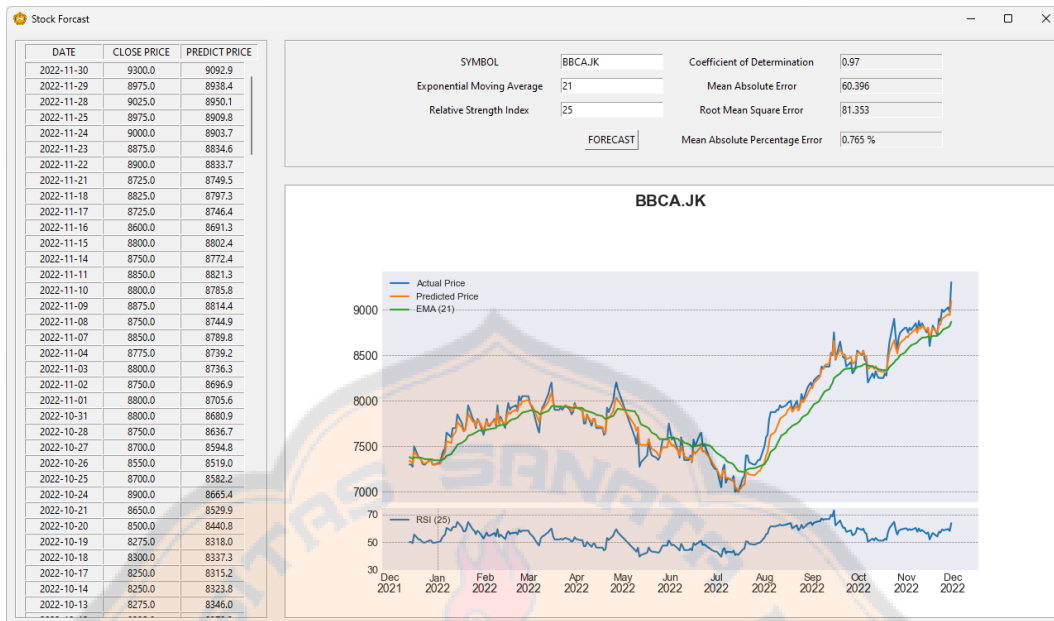
Lampiran 27 Hasil percobaan periode EMA 21 dan periode RSI 14



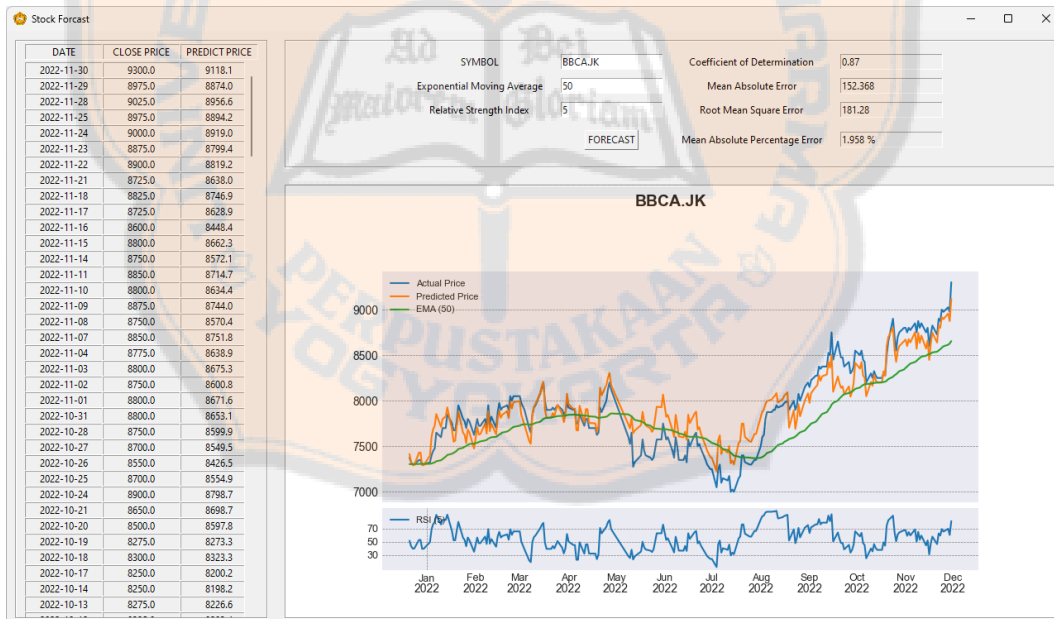
Lampiran 28 Hasil percobaan periode EMA 21 dan periode RSI 21



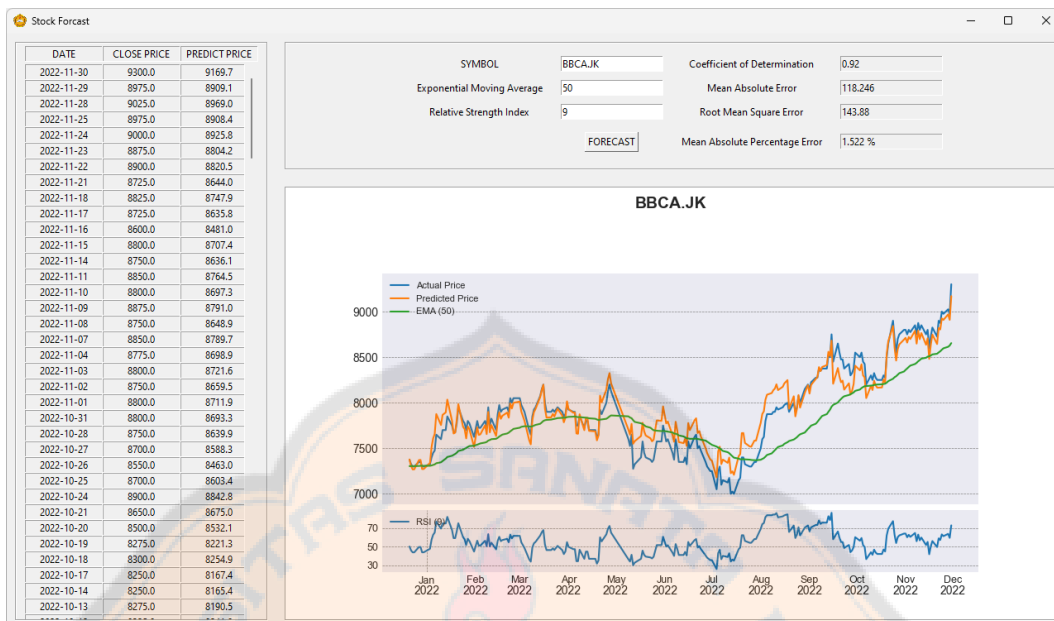
Lampiran 29 Hasil percobaan periode EMA 21 dan periode RSI 25



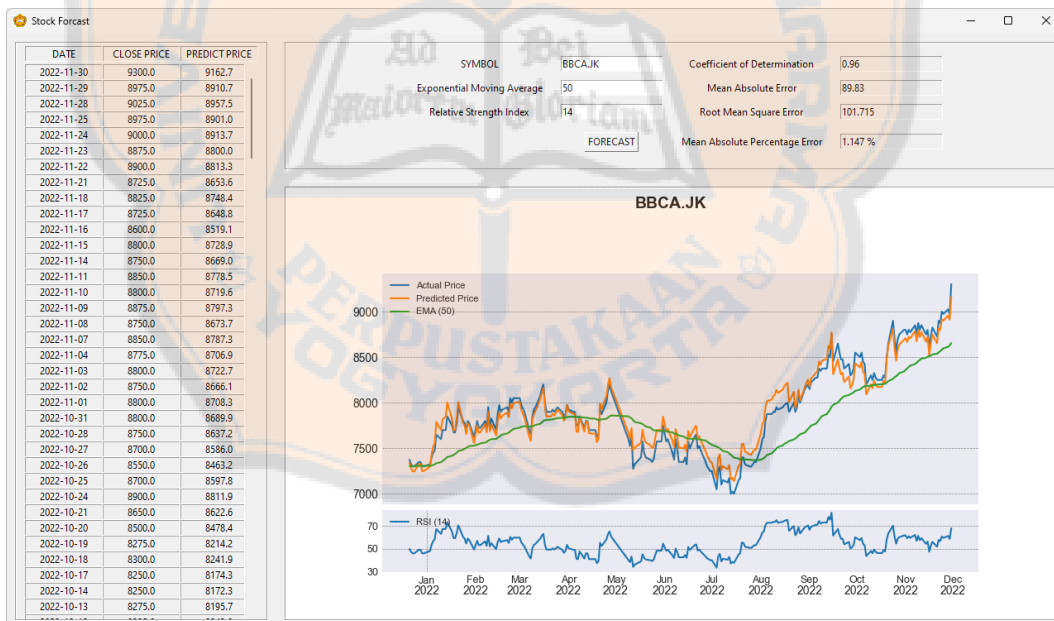
Lampiran 30 Hasil percobaan periode EMA 50 dan periode RSI 5



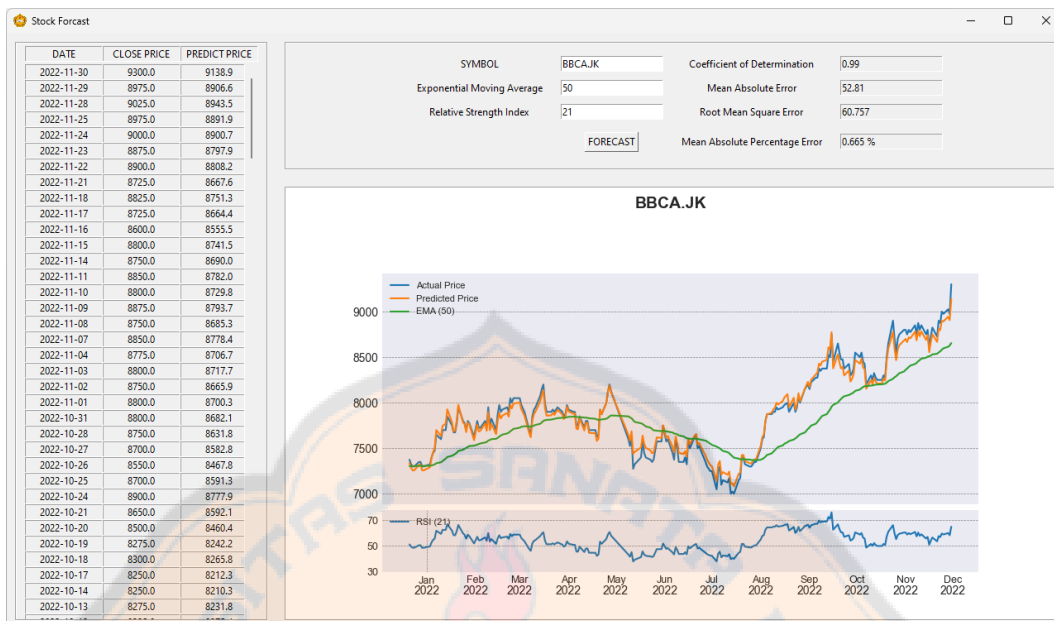
Lampiran 31 Hasil percobaan periode EMA 50 dan periode RSI 9



Lampiran 32 Hasil percobaan periode EMA 50 dan periode RSI 14



Lampiran 33 Hasil percobaan periode EMA 50 dan periode RSI 21



Lampiran 34 Hasil percobaan periode EMA 50 dan periode RSI 25

