

**STUDI PERBANDINGAN DATA TIME SERIES HARGA EMAS
MENGUNAKAN METODE XGBOOST, STL, DAN
FACEBOOK PROPHET**

RIZAL MUJAHIDDAN



**DEPARTEMEN ILMU KOMPUTER
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT PERTANIAN BOGOR
BOGOR
2023**

PERNYATAAN MENGENAI SKRIPSI DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa skripsi dengan judul “STUDI PERBANDINGAN DATA TIME SERIES HARGA EMAS MENGGUNAKAN METODE XGBOOST, STL, DAN FACEBOOK PROPHET” adalah karya saya dengan arahan dari dosen pembimbing dan belum diajukan dalam bentuk apapun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir laporan akhir ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Institut Pertanian Bogor.

Bogor, September 2023

Rizal Mujahiddan
G64190069

ABSTRAK

RIZAL MUJAHIDDAN. STUDI PERBANDINGAN DATA TIME SERIES HARGA EMAS MENGGUNAKAN METODE XGBOOST, STL, DAN FACEBOOK PROPHET. Dibimbing oleh ENDANG PURNAMA GIRI dan MEDRIA KUSUMA DEWI HARDHIENATA

Harga emas memiliki peranan yang signifikan dalam sektor keuangan dan investasi. Sebagai komoditas yang stabil dan cenderung naik nilainya, peramalan harga emas menjadi sangat penting bagi para pelaku pasar. Studi ini membandingkan tiga model analisis data time series harga emas: XGBoost, STL, dan Facebook Prophet. Faktor-faktor eksternal seperti tingkat inflasi, nilai tukar mata uang, dan indeks saham juga dipertimbangkan sebagai variabel prediktor. Data harga emas dan faktor-faktor eksternal dikumpulkan dari Januari 2010 hingga April 2023 dan dianalisis menggunakan ketiga model tersebut. Hasil penelitian menunjukkan bahwa model STL tanpa faktor eksternal memiliki tingkat kesalahan SMAPE sebesar 7,46% dan tingkat kesalahan MASE sebesar 93,18%. Hal ini menyimpulkan bahwa model STL lebih akurat dibandingkan dengan model XGBoost dan Facebook Prophet.

Kata kunci: Emas, Facebook Prophet, Keuangan, Peramalan, Analisis Waktu, STL, XGBoost

ABSTRACT

RIZAL MUJAHIDDAN. STUDI PERBANDINGAN DATA TIME SERIES HARGA EMAS MENGGUNAKAN METODE XGBOOST, STL, DAN FACEBOOK PROPHET. Supervised by ENDANG PURNAMA GIRI and MEDRIA KUSUMA DEWI HARDHIENATA

The price of gold plays a significant role in the finance and investment sectors. As a commodity that is stable and tends to increase in value, gold price prediction is very important for market participants. This study compares three gold price time series data analysis models: XGBoost, STL, and Facebook Prophet. External factors such as inflation rate, currency exchange rate, and stock index are also considered as predictor variables. Gold price data and external factors were collected from January 2010 to April 2023 and analyzed using the three models. The results show that the STL model without external factors has an SMAPE error rate of 7.46% and MASE error rate of 93.18%. This concludes that the STL model is more accurate compared to the XGBoost and Facebook Prophet models.

Keywords: Gold price, Facebook Prophet, Finance, Forecasting, Temporal Analysis, STL, XGBoost

**STUDI PERBANDINGAN DATA TIME SERIES HARGA EMAS
MENGUNAKAN METODE XGBOOST, STL, DAN
FACEBOOK PROPHET**

RIZAL MUJAHIDDAN

Skripsi
sebagai salah satu syarat untuk memperoleh gelar
Sarjana pada
Program Studi Ilmu Komputer

**DEPARTEMEN ILMU KOMPUTER
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT PERTANIAN BOGOR
BOGOR
2023**

Tim Penguji pada Ujian Skripsi:

1. Dr. Karlisa Priandana S.T., M.Eng

Judul : Studi Perbandingan Data Time Series Harga Emas
Menggunakan Metode XGBoost, STL, dan Facebook
Prophet
Nama : Rizal Mujahiddan
NIM : G64190069

Disetujui oleh

Pembimbing 1:

Endang Purnama Giri S.Kom., M.Kom.

Pembimbing 2:

Medria Kusuma Dewi Hardhienata S.Komp., Ph.D.

Diketahui oleh

Ketua Departemen Ilmu Komputer:

Dr. Sony Hartono Wijaya S.Kom., M.Kom.

198108092008121002

Tanggal Ujian:

Tanggal Lulus:

PRAKATA

Puji dan syukur penulis panjatkan kepada Tuhan Yang Maha Esa atas segala karunia-Nya sehingga karya ilmiah ini berhasil diselesaikan. Judul yang dipilih dalam penelitian yang dilaksanakan sejak bulan Januari 2023 sampai bulan Juli 2023 adalah Studi Perbandingan Analisis Data Time Series Harga Emas Menggunakan XGBoost, STL, dan Facebook Prophet.

Terima kasih penulis ucapkan kepada:

1. Keluarga penulis yang senantiasa mendukung dalam pelaksanaan penelitian, terdiri dari kedua orang tua, yaitu bapak Cecep Gunawan dan ibu Minah Mintarsih serta kedua saudara perempuan penulis yaitu Sabriyani Chairunissa dan Zackyah Miftahul Zannah.
2. Bapak Endang Purnama Giri S.Kom., M.Kom. selaku dosen pembimbing pertama dan ibu Medria Kusuma Dewi Hardhienata S.Komp., Ph.D. selaku dosen pembimbing kedua atas bimbingan dan arahannya selama berjalannya penelitian.

Bogor, September 2023

Rizal Mujahiddan

DAFTAR ISI

DAFTAR ISI	8
DAFTAR GAMBAR	10
DAFTAR TABEL	10
I PENDAHULUAN	11
1.1 Latar Belakang	11
1.2 Rumusan Masalah	12
1.3 Tujuan Penelitian	13
1.4 Manfaat Penelitian	13
1.5 Ruang Lingkup	13
II TINJAUAN PUSTAKA	14
2.1 Harga Emas	14
2.2 Model Peramalan pada Univariate dan Multivariate	14
2.3 Model Facebook Prophet	15
2.4 Seasonal-Trend Decomposition Procedure Based on LOESS (STL)	19
2.5 Extreme Gradient Boosting (XGBoost)	21
2.6 Perbandingan Model peramalan harga emas	24
2.7 Granger Causality	24
2.8 Symmetric Mean Absolute Percentage Error (SMAPE)	25
III METODE PENELITIAN	27
3.1 Alat dan Bahan	27
3.2 Tahapan Penelitian	27
3.2.1 Studi literatur	28
3.2.2 Pengambilan Data	28
3.2.3 Analisis Eksplorasi Data	28
3.2.4 Pengolahan data	28
3.2.5 Melatih data dengan 3 model	29
3.2.6 Tuning parameter	29
3.2.7 Evaluasi Model	31
3.2.8 Peramalan Data	31
3.3 Deskripsi Data	31
3.4 Skenario Eksperimen	38
3.5 Lingkungan Penelitian	39
IV HASIL DAN PEMBAHASAN	41
4.1 Analisis Statistik Data	41
4.2 Perbandingan Model Forecasting	44
V KESIMPULAN DAN SARAN	51
5.1 Kesimpulan	51
5.2 Saran	51

DAFTAR GAMBAR

1	Contoh grafik komponen prophet (Taylor dan Letham 2017)	19
2	Alur Penelitian	28
3	Grafik harga emas berdasarkan tanggal	33
4	Grafik kurs jual berdasarkan tanggal	33
5	Grafik kurs beli berdasarkan tanggal	34
6	Grafik suku bunga BI berdasarkan tanggal	34
7	Grafik inflasi di Indonesia berdasarkan tanggal	35
8	Grafik harga emas berdasarkan tanggal setelah diimputasi	36
9	Grafik kurs jual berdasarkan tanggal setelah diimputasi	36
10	Grafik kurs beli berdasarkan tanggal setelah diimputasi	37
11	Grafik data inflasi berdasarkan tanggal setelah diimputasi	37
12	Grafik data inflasi berdasarkan tanggal setelah diimputasi	38
13	Grafik normalisasi master data	38
14	Persebaran data harga emas	42
15	Persebaran kurs jual dollar	43
16	Persebaran kurs beli dollar	43
17	Persebaran data inflasi	43
18	Persebaran data suku bunga BI	44
19	<i>Heatmap p-value Granger Causality</i>	44
20	Grafik autocorrelation	45
21	Partial autocorrelation	45
22	Grafik harga emas pada model Facebook Prophet <i>univariate</i>	48
23	Grafik harga emas pada model XGBoost <i>univariate</i>	48
24	Grafik harga emas pada model STL <i>univariate</i>	49
25	Grafik harga emas pada model Facebook Prophet <i>multivariate</i>	50
26	Grafik harga emas pada model XGBoost <i>multivariate</i>	50

DAFTAR TABEL

1	Akurasi penelitian sebelumnya	24
2	<i>Tuning</i> parameter Facebook Prophet	29
3	<i>Tuning</i> parameter STL	30
4	<i>Tuning</i> parameter XGBoost	30
5	Master data	31
6	Skenario Eksperimen	39
7	Statistik Master Data	41
8	Kinerja Model Facebook Prophet pada Data <i>univariate</i> dengan Metode <i>cross validation (expanding window)</i>	46
9	Kinerja Model Facebook Prophet pada Data <i>multivariate</i> dengan Metode <i>cross validation (expanding window)</i>	46
10	Kinerja Model XGBoost pada Data <i>univariate</i> dengan Metode <i>cross validation (expanding window)</i>	46

11	Kinerja Model XGBoost pada Data <i>univariate</i> dengan Metode <i>cross validation (expanding window)</i>	46
12	Kinerja Model STL pada Data <i>univariate</i> dengan Metode <i>cross validation (expanding window)</i>	47
13	Kinerja seluruh model dengan Data <i>univariate</i>	47
14	Kinerja seluruh model dengan Data <i>multivariate</i>	49

I PENDAHULUAN

1.1 Latar Belakang

Emas telah memegang peran penting dalam budaya selama sejarah, melambangkan keindahan, kemurnian, kekuatan, dan prestasi. Hingga saat ini, emas masih menjadi simbol yang signifikan dalam masyarakat, digunakan sebagai uang, penghargaan, cincin pernikahan, medali Olimpiade, dan barang-barang keagamaan seperti salib. Kelangkaannya dan nilai yang tinggi membuatnya menjadi alat tukar yang dicari dan simpanan nilai yang berharga. Meskipun sulit diramal karena dipengaruhi oleh banyak faktor, peramalan harga emas menjadi penting bagi para investor untuk membuat keputusan transaksi yang tepat. Sebagai aset berharga, emas terus menarik perhatian dan memainkan peran krusial dalam sistem keuangan dan stabilitas sosial masyarakat (Das *et al.* 2022).

Meramalkan fluktuasi harga emas adalah masalah krusial bagi banyak pemangku kepentingan di pasar global. Investor, perusahaan tambang, dan agen lain yang memandang emas sebagai indikator yang dapat diandalkan untuk performa masa depan ekonomi dunia perlu memiliki ramalan yang akurat dan tepat waktu mengenai perubahan nilai logam mulia ini. Dengan meramalkan fluktuasi harga emas, diharapkan dapat diambil keputusan yang tepat mengenai investasi, produksi, dan konsumsi, serta melakukan *hedging* terhadap potensi risiko dan ketidakpastian (Jabeur *et al.* 2021).

Terdapat beberapa penelitian pendahulu yang telah dilakukan untuk melakukan peramalan harga emas. Penelitian Makala dan Li (2021) telah menerapkan model ARIMA untuk meramalkan harga emas dengan akurasi sebesar 99,7%. Penelitian Makala dan Li (2021) menggunakan data harga emas dunia dari situs *goldhub*. Di sisi lain, penelitian Chai *et al.* (2021) telah menyelidiki hubungan dinamis antara harga emas, minyak mentah, indeks dollar AS, dan indeks volatilitas (VIX). Dalam penelitian tersebut digunakan model STL-ETS untuk mengakomodasi tren fluktuasi pengembalian harga emas secara akurat dan memberikan analisis mendalam tentang korelasi emas dengan indikator ekonomi penting lainnya. Penelitian Chai *et al.* (2021) menggunakan data harga emas dunia dari situs *goldhub*. Penelitian Jabeur *et al.* (2022) memberikan perbandingan antara enam model pembelajaran mesin untuk meramalkan pergerakan harga emas. Penelitian tersebut mengusulkan pendekatan inovatif menggunakan algoritma XGBoost dan nilai interaksi SHAP untuk menginterpretasikan ramalan dengan akurat. Penelitian ini memberikan wawasan penting tentang bagaimana model pembelajaran mesin dapat diterapkan untuk peramalan harga emas. Penelitian tersebut menggunakan data harga emas Amerika Serikat dari situs *goldhub*.

Penelitian lain oleh Kishann dan RamaParvathy (2022) berhasil mengembangkan model Facebook Prophet yang mencapai akurasi sebesar 97,2% dalam meramalkan harga emas. Keunggulan penelitian ini adalah tingginya

akurasi ramalan yang telah dicapai. Penelitian tersebut menggunakan data dari Negara India. Berdasarkan penelitian Makridakis *et al.* (2019) dan Ouyang *et al.* (2021) menunjukkan bahwa model statistik dan model *machine learning* (modelnya tidak termasuk *neural network*) lebih efektif dibandingkan dengan model berbasis *neural network* untuk peramalan. Oleh karena itu, model berbasis *neural network* tidak dianjurkan dalam konteks peramalan harga emas.

Berdasarkan hasil penelitian Kishann dan RamaParvathy (2022), model Facebook Prophet memiliki tingkat akurasi yang tinggi sebesar 97,2% untuk meramal harga emas. Berdasarkan hasil penelitian Jabeur *et al.* (2022), model XGBoost memiliki tingkat akurasi yang tinggi yakni sebesar 99,4% untuk meramal harga emas. Oleh karena itu, kedua model ini dianggap sebagai pilihan yang baik untuk digunakan dalam penelitian ini. Selanjutnya, penggunaan model STL dikombinasikan dengan model statistik lain telah terbukti meningkatkan akurasi sesuai dengan penelitian Ouyang *et al.* (2021).

Oleh karena itu, model STL juga dipertimbangkan sebagai pilihan untuk digunakan dalam penelitian ini. Dengan mempertimbangkan temuan dari penelitian-penelitian sebelumnya, penelitian ini akan menggunakan model Facebook Prophet, XGBoost dan STL sebagai model utama untuk meramalkan harga emas.

Kelemahan utama model ARIMA terletak pada keterbatasannya dalam menangani data keuangan yang cenderung penuh dengan *irregularitas* dan *anomali* yang dapat mengakibatkan kesulitan dalam menghasilkan prediksi yang akurat dan stabil (Syrris dan Shenai 2021). Model ARIMA tidak digunakan karena hal tersebut. Penelitian Risse (2019) menyatakan bahwa model *multivariate* lebih baik dibandingkan dengan model *univariate* dalam segi akurasi. Penelitian tersebut menggunakan model DWT-SVR. Model ini tidak digunakan karena penerapannya lebih kompleks.

Perbedaan utama penelitian ini dengan penelitian-penelitian sebelumnya, yakni penggunaan dataset harga emas di Indonesia yang dikumpulkan secara mandiri dari situs logammulia.com. Selain itu, pada penelitian ini dilakukan juga analisis perbandingan model *univariate* dan *multivariate* pada 3 model regresi XGBoost, STL dan Facebook Prophet. Untuk memperlihatkan pengaruh eksternal terhadap harga emas, penelitian ini juga menggunakan uji hipotesis *Granger Causality*. Selanjutnya dari analisis uji performa, setiap model regresi diharapkan dapat diperoleh model prediksi yang paling optimal.

1.2 Rumusan Masalah

Permasalahan dalam penelitian ini adalah minimnya penelitian yang ada saat ini yang berfokus pada perbandingan metode XGBoost, STL, dan Facebook Prophet untuk meramal harga emas di Indonesia. Selain itu, belum banyak penelitian pendahulu yang menganalisis penggunaan pendekatan *univariate* dan

multivariate dalam meramalkan data emas di Indonesia. Berdasarkan hal tersebut, rumusan masalahnya adalah sebagai berikut:

1. Antara model XGBoost, STL, dan Facebook Prophet, model mana yang paling efektif digunakan untuk data harga emas di Indonesia?
2. Metode mana yang lebih baik untuk meramal harga emas saat diterapkan, apakah menggunakan pendekatan *multivariate* atau *univariate*?

1.3 Tujuan Penelitian

Penelitian ini bertujuan membandingkan kinerja tiga metode analisis data time series yaitu XGBoost, STL, dan Facebook Prophet dalam meramal harga emas, serta melakukan perbandingan antara model *univariate* dan *multivariate* pada 3 model tersebut.

1.4 Manfaat Penelitian

Manfaat Penelitian ini adalah sebagai berikut:

1. Dapat digunakan oleh investor atau pelaku pasar lainnya untuk membuat keputusan investasi yang lebih baik.
2. Dapat menjadi dasar bagi penelitian selanjutnya yang berhubungan dengan peramalan harga emas.

1.5 Ruang Lingkup

Ruang lingkup digunakan untuk penelitian adalah sebagai berikut:

1. Data harga Emas diperoleh dari situs logam mulia (Logam Mulia, 2022)
2. Data yang digunakan dari 4 Januari 2010 hingga 30 April 2023
3. Metrik yang digunakan adalah MASE dan SMAPE

II TINJAUAN PUSTAKA

2.1 Harga Emas

Harga emas merupakan salah satu indikator penting dalam pasar finansial. Harga emas cenderung naik ketika pasar saham turun, sehingga digunakan sebagai safe haven bagi investor. Grafik harga emas menunjukkan pergerakan harga emas dalam jangka waktu tertentu, yang dapat membantu investor dalam mengambil keputusan investasi.

Faktor-faktor yang memengaruhi harga emas meliputi faktor ekonomi, politik, dan geopolitik. Faktor ekonomi seperti tingkat inflasi, suku bunga, dan kondisi perekonomian global memiliki dampak yang signifikan pada harga emas. Faktor politik dan geopolitik seperti perang, kerusuhan, dan perubahan kebijakan pemerintah juga dapat memengaruhi harga emas.

Menurut studi yang dilakukan oleh Anandasayanan *et al.* (2019) dan Alemi *et al.* (2016), faktor ekonomi seperti tingkat inflasi memiliki pengaruh yang signifikan terhadap harga emas. Dalam studi tersebut, data harga emas dan faktor ekonomi dianalisis menggunakan metode analisis korelasi dan hasilnya menunjukkan bahwa tingkat inflasi dan suku bunga memiliki korelasi negatif dengan harga emas.

Menurut Anandasayanan *et al.* (2019), pengujian hubungan antara inflasi dan harga emas dengan menggunakan data bulanan dari Januari 1999 hingga Desember 2018. Hasil penelitian menunjukkan adanya hubungan yang kuat dan positif antara inflasi dan harga emas. Analisis regresi menunjukkan bahwa inflasi berdampak signifikan terhadap harga emas dan uji *Granger causality* menemukan bahwa ada hubungan searah antara harga emas dan inflasi. Temuan yang serupa juga ditemukan dalam studi lain. Disarankan untuk membeli emas selama periode inflasi yang tinggi daripada berinvestasi di pasar saham.

Menurut penelitian Farhat dan Ghalayini (2020), spekulasi tentang kepemilikan emas di masa depan penting. Setelah perjanjian Bretton Woods, muncul pasar kontrak keuangan. Spekulasi mencegah *hedging pressure* berlebihan. Kontrak berjangka emas lebih banyak diperdagangkan karena minat investasi emas meningkat. Jumlah kontrak aktif (*open interest*) dan pedagang meningkat seiring waktu. Kontrak ini memengaruhi uang dan harga emas di pasar berjangka. Lebih banyak kontrak aktif berarti lebih banyak perdagangan dan spekulasi, terpicu oleh pembelian kontrak emas besar. Spekulator yang membeli lebih banyak emas bisa naikkan harga di masa depan.

2.2 Model Peramalan pada Univariate dan Multivariate

Data deret waktu merupakan salah satu jenis data yang sangat umum digunakan untuk mengambil informasi dan merekam berbagai aktivitas di berbagai bidang. Dalam setiap ranah yang melibatkan pengukuran waktu melalui sensor, sensus, atau catatan transaksi, pencatatan urutan pengamatan yang terindeks dengan penanda waktu memberikan wawasan tentang perkembangan

suatu kuantitas yang dapat diukur di masa lalu. Data deret waktu dapat digunakan untuk membuat visualisasi data, menemukan pola dan korelasi berulang, mengklasifikasikan dan mengelompokkan data, mendeteksi pencilon, mensegmentasikan data, meramalkan dan mensimulasikan data (Dama dan Sinoquet 2021).

Peramalan deret waktu penting dalam berbagai domain seperti keuangan, bisnis, pengendalian proses kualitas industri, transaksi digital, distribusi tenaga listrik, meteorologi, biologi dan kedokteran. Ini membantu menganalisis masa lalu dan meramalkan masa depan. Misalnya, dapat digunakan untuk merencanakan kebijakan oleh berbagai organisasi di bidang keuangan dan bisnis, mendeteksi situasi abnormal atau penipuan dalam pengendalian proses kualitas industri dan transaksi digital, mengontrol beban melalui pemantauan canggih dan mengatasi gangguan dalam aliran daya listrik di distribusi tenaga listrik, membimbing pengambilan keputusan yang terinformasi untuk pertanian, navigasi udara dan laut di meteorologi, memperoleh pengetahuan tentang aktivitas gen yang diberikan status terpengaruh / tidak terpengaruh di biologi, meramalkan penyebaran penyakit, memperkirakan tingkat kematian atau menilai risiko yang bergantung pada waktu di kedokteran (Dama dan Sinoquet 2021).

Peramalan adalah proses memprediksi nilai masa depan berdasarkan data masa lalu. Ada dua jenis utama peramalan yakni *univariate* dan *multivariate*. Peramalan *univariate* hanya menggunakan satu variabel dari masa lalu untuk meramal variabel yang sama di masa depan. Peramalan *multivariate* menggunakan lebih dari satu variabel dari masa lalu untuk meramal satu atau beberapa variabel di masa depan. Metode yang berbeda dapat menangani berbagai jenis data dan hubungan. Beberapa metode membutuhkan formula yang jelas, sementara yang lain dapat menggunakan simulasi atau *neural network*. Teori-teori dari bidang lain dapat membantu untuk memahami hubungan-hubungan ini. Peramalan juga dapat memprediksi beberapa variabel secara bersamaan (Petroopoulos *et al.* 2022).

Peramalan *univariate* dan *multivariate* memiliki perbedaan signifikan dalam hal struktur data, praproses, dan hasil output. Peramalan *univariate* lebih sederhana dengan menggunakan data indeks waktu dan satu target yang akan diramalkan, sedangkan peramalan *multivariate* melibatkan fitur-fitur yang berhubungan dengan indeks waktu. Praproses pada peramalan *multivariate* lebih kompleks karena memerlukan pemilihan fitur yang relevan, sementara peramalan *univariate* lebih sederhana dalam hal ini. Output dari peramalan *multivariate* mencakup ramalan untuk satu atau lebih target, sedangkan peramalan *univariate* hanya fokus pada ramalan untuk satu target berdasarkan data historis.

2.3 Model Facebook Prophet

Facebook Prophet merupakan sebuah metode yang digunakan untuk meramal data rangkaian waktu dengan menggunakan model aditif, dimana

tren-tren non-linear ditentukan berdasarkan musiman tahunan, mingguan, dan harian, serta efek-efek liburan. (Taylor dan Letham 2017). Model ini mudah digunakan dan dapat mengakomodasi faktor-faktor seperti tren, musiman, dan efek liburan. Namun, model ini memiliki beberapa kekurangan seperti kurangnya kemungkinan untuk memahami secara tepat bagaimana dan apa yang dipelajari oleh model. Ketika digunakan untuk masalah peramalan yang kompleks, kemampuan model dapat menjadi terbatas (Korstanje 2021).

$$y(t) = g(t) + s(t) + h(t) + \epsilon_t \quad (1)$$

t = Waktu

$g(t)$ = Komponen tren

$s(t)$ = Komponen Seasonal

$h(t)$ = Komponen *Holiday*

ϵ_t = Komponen residual

$y(t)$ = Data deret waktu

Persamaan 1 merupakan dasar dari model Facebook Prophet yang telah dikembangkan dari penelitian Harvey dan Peters (1990). Model ini digunakan untuk meramalkan data deret waktu berdasarkan model aditif di mana tren non-linear disesuaikan dengan musiman tahunan, mingguan, dan harian, ditambah efek liburan.

Facebook Prophet menggunakan kurva pertumbuhan linear atau logistik untuk memodelkan tren. Model Facebook Prophet untuk peramalan pertumbuhan menggunakan model pertumbuhan logistik untuk menggambarkan pertumbuhan data yang nonlinier dan mencapai titik jenuh (*carrying capacity*) seperti pertumbuhan populasi di ekosistem alami. Persamaan yang digunakan logistik adalah sebagai berikut:

$$g(t) = \frac{C}{1 + e^{(-k(t-m))}} \quad (2)$$

t = Waktu

$g(t)$ = Komponen tren

C = *Carrying Capacity*

k = Laju pertumbuhan

m = Parameter *offset*

Persamaan 2 menunjukkan bahwa ada 2 hal yang perlu dipertimbangkan. Pertama, bahwa *carrying capacity* tidak konstan dan berubah berdasarkan waktu. Kedua, laju pertumbuhan tidak konstan. Di dalam model Facebook Prophet, ada beberapa titik perubahan (*change points*) S . Pada daerah yang berbatasan titik perubahan, daerah satu dengan yang lain memiliki nilai laju pertumbuhan yang berbeda-beda yang bisa diartikan seperti vektor laju δ . Laju perubahan pada titik

tertentu dengan nilai awalnya k , maka $k + \sum_{j:t>s_j} \delta_j$. Dengan vektor

$a(t) \in \{0, 1\}^S$, maka dapat didefinisikan seperti:

$$a_j(t) = 1 - [t < s_j] \quad (3)$$

Persamaan 3 menunjukkan bahwa vektor $a(t)$ merujuk pada waktu t sudah melewati titik perubahan pada s_j . Laju pada waktu t adalah $k + a(t)^T \delta$. Ketika mengubah parameter laju k , perlu juga menyesuaikan parameter offset m agar terjadi koneksi yang mulus antara ujung segmen, dan penyesuaian yang tepat untuk titik perubahan j dapat dihitung dengan mudah seperti pada Persamaan 4. Persamaan 4 adalah sebagai berikut:

$$\gamma_j = \left(s_j - m - \sum_{l < j} \gamma_l \right) \left(1 - \frac{k + \sum_{l < j} \delta_l}{k + \sum_{l \leq j} \delta_l} \right) \quad (4)$$

Persamaan 4 memiliki beberapa variable yang sama untuk variable seperti m , j , l , k dan δ . Parameter γ digunakan untuk menghaluskan komponen tren karena model Facebook Prophet memiliki banyak titik perubahan yang menyebabkan berbagai perbedaan dalam laju pertumbuhan. Berdasarkan dari Persamaan 1, 2, 3 dan 4, maka menghasilkan fungsi tren seperti berikut:

$$g(t) = \frac{C(t)}{1 + \exp(-(k + a(t)^T \delta)(t - (m + (a(t)^T \gamma))))} \quad (5)$$

Persamaan 5 merupakan pengembangan dari Persamaan 1 dengan syarat *carrying capacity* tidak konstan dan laju pertumbuhan k berubah berdasarkan waktu. Persamaan (5) menunjukkan parameter γ dihitung untuk menyesuaikan antar titik perubahan.

Untuk masalah peramalan yang tidak menunjukkan pertumbuhan yang jenuh, model pertumbuhan dengan tingkat konstan berpotongan memberikan model yang lebih sederhana dan seringkali bermanfaat. Model tren ini menggunakan fungsi seperti berikut:

$$g(t) = (k + a(t)^T \delta)t + (m + a(t)^T \gamma) \quad (6)$$

Persamaan 6 menunjukkan bahwa k sebagai tingkat pertumbuhan, δ sebagai penyesuaian tingkat, m sebagai parameter offset, dan γ_j diatur sebagai $-s_j \delta_j$ agar fungsi tetap berkelanjutan. Penerapan model ini memungkinkan peramalan yang lebih efisien dan hasil yang relevan dengan data yang ada, karena tingkat pertumbuhan dan penyesuaian parameter telah diatur dengan tepat untuk memastikan kelancaran dan konsistensi dalam peramalan data.

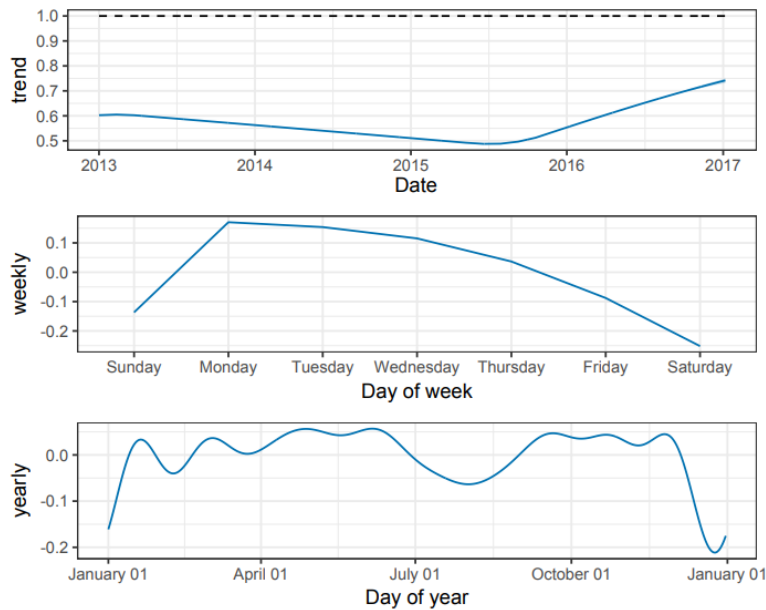
$$S(t) = \sum_{k=1}^K (a_k \cos(\frac{2\pi kt}{P}) + b_k \sin(\frac{2\pi kt}{P})) \quad (7)$$

Persamaan 7 menunjukkan bahwa arti simbol dari $S(t)$ mewakili musiman pada waktu t , K adalah jumlah komponen Fourier yang diperlukan untuk

menangkap musiman dengan akurat, P adalah periode musiman, serta a_k dan b_k adalah koefisien dari deret Fourier. Facebook Prophet memasukkan dampak liburan dan acara khusus pada deret waktu. Rumus untuk efek liburan pada waktu t adalah:

$$X(t) = \sum_{i=1}^N x_i I_t(t \in \text{holiday}_i) \quad (8)$$

Persamaan 8 menunjukkan bahwa arti simbol dari $X(t)$ mewakili efek liburan pada waktu t , N adalah jumlah liburan, x_i adalah efek liburan untuk liburan ke- i , dan $I(t \in \text{holiday})$ adalah fungsi indikator yang sama dengan 1 jika waktu t jatuh dalam liburan ke- i dan 0 jika tidak. Algoritma ini menyesuaikan tren, musiman, dan efek liburan untuk meminimalkan dampak pencilaan pada ramalan. Algoritma ini mengidentifikasi titik perubahan dalam tren dan menangkap perubahan dalam pola dasar deret waktu. Proses optimasi menyesuaikan parameter hingga konvergensi dan mengevaluasi akurasi ramalan dengan membandingkannya dengan data aktual.



Gambar 1 Contoh grafik komponen prophet (Taylor dan Letham 2017)

Gambar 1 menunjukkan bahwa data akan dipecah menjadi 3 komponen. Komponen tersebut yakni, tren, *weekly*, dan *yearly*. Pada bagian *yearly*, berpola gelombang sebanyak 3 kali dengan gelombang berbeda. Grafik ini merupakan penerapan dari *fourier* pada bagian komponen *seasonal*.

Facebook Prophet memiliki kelebihan dalam kemudahan penggunaan dan kemampuannya dalam memperhitungkan faktor-faktor seperti tren, musiman, dan efek liburan untuk menghasilkan ramalan yang lebih akurat (Taylor dan Letham, 2017). Namun, Model Facebook Prophet juga memiliki beberapa kelemahan yang perlu dipertimbangkan. Pertama, Facebook Prophet memiliki keterbatasan dalam

pemahaman yang tepat tentang bagaimana dan apa yang dipelajari oleh model tersebut. Hal ini disebabkan oleh Facebook Prophet yang secara otomatis melakukan sebagian besar pekerjaan untuk pengguna tanpa memerlukan pemahaman teori yang mendalam. Kedua, Facebook Prophet memiliki jumlah parameter yang terbatas yang dapat diatur dan disesuaikan, yang dapat menjadi kelebihan jika model berfungsi dengan baik, tetapi juga dapat menjadi kelemahan jika model tidak memberikan tingkat akurasi yang diinginkan. Terakhir, meskipun Facebook Prophet dapat memberikan hasil yang baik dengan pengaturan parameter default, namun ketika digunakan untuk masalah ramalan yang kompleks, kemampuan model dapat menjadi terbatas. Hal ini disebabkan oleh sifat modular dan sederhana dari model regresi yang digunakan oleh Facebook Prophet. Oleh karena itu, walaupun Facebook Prophet menawarkan kemudahan penggunaan, pengguna harus mempertimbangkan kelemahan-kelemahan ini sebelum memilih menggunakan model Facebook Prophet untuk masalah peramalan (Korstanje 2021).

2.4 Seasonal-Trend Decomposition Procedure Based on LOESS (STL)

STL adalah proses penguraian time series menjadi tiga komponen utama, yaitu tren, musiman, dan komponen residu. Rumus STL adalah sebagai berikut:

$$Y_v = T_v + S_v + R_v \quad (9)$$

Y_v = Data *Time Series*

T_v = Komponen *Trend*

S_v = Komponen *Seasonal*

R_v = Komponen Residu

Persamaan (9) bahwa perlu dicatat bahwa STL (*Seasonal-Trend decomposition* menggunakan LOESS) hanya terdiri dari 3 komponen, sedangkan model Facebook Prophet memiliki 4 komponen. STL menggunakan metode *smoothing* yakni LOESS (*locally estimated scatterplot smoothing*). LOESS (*locally estimated scatter plot smoothing*) adalah metode non-parametrik untuk melakukan *smoothing* pada data (Cleveland *et al.* 1990). LOESS menggunakan regresi dengan nilai bobot tertentu (W) dengan syarat seperti berikut:

1. $W(x) > 0$ untuk $|x| < 1$.
2. $W(-x) = W(x)$.
3. $W(x)$ merupakan fungsi tidak naik untuk $x \geq 0$.
4. $W(x) = 0$ untuk $|x| \geq 1$.

Dengan syarat ini, maka fungsi W akan bervariasi seperti fungsi *tricube* dan *gaussian*. Namun, dalam banyak kasus, fungsi W yang sering digunakan adalah *tricube*. Fungsi *tricube* memiliki bentuk sebagai berikut:

$$W(x) = \{(1 - |x|^3)^3, |x| < 1; 0, |x| \geq 1\} \quad (10)$$

Persamaan 10 digunakan fungsi kernel tricube yang memiliki nilai dari nol hingga satu. Fungsi ini meningkatkan kinerja regresi berbobot lokal. Fungsi-fungsi dengan kontak yang halus serupa diharapkan juga memberikan hasil yang baik. Namun, fungsi bobot uniform, dengan ketidakteraturan pada nilai satu, menunjukkan kinerja yang kurang baik. Ketidakteraturan pada fungsi uniform menyebabkan ketidakhalusan lokal dan aproksimasi distribusi yang kurang memuaskan (Cleveland dan Devlin 1988).

Algoritma LOESS hampir mirip dengan regresi linear, tetapi yang membedakan menggunakan metode *weighted least square*. Untuk mekanisme LOESS adalah sebagai berikut:

1. Tentukan parameter d (*degree*), fungsi W (*Weight*), banyak iterasi (t), parameter untuk menentukan smoothing (f) dan parameter untuk menentukan ukuran *window* ($span$)
2. Urutkan variabel independen
3. Setiap titiknya hitunglah seperti
 - a. buatlah jarak di setiap titiknya
 - b. buatlah window berdasarkan parameter $span$ untuk tetangga terdekat dari titik tersebut
 - c. menskala dengan jarak tersebut berdasarkan jarak maksimum.
 - d. terapkan fungsi W pada jarak tersebut
 - e. hitunglah regresi dengan metode *weighted least square* dan dapatkan persamaan regresi
 - f. hitunglah y duga dari persamaan tersebut untuk titik tersebut
4. Hubungkan titik duga tersebut
5. Diperlukan bisa iterasi ulang dengan mengubah parameter tersebut agar *error* bisa diminimalisir.

Metode ini berdasarkan dari penelitian Cleveland (1979). Intinya, Metode ini memperkirakan nilai rata-rata dari variabel dependen pada setiap titik dengan menggunakan regresi lokal pada titik tersebut. Regresi lokal ini dilakukan dengan menggunakan polinomial rendah derajat yang disesuaikan dengan data di sekitar titik tersebut (Cleveland *et al.* 1990). Mekanisme STL adalah sebagai berikut :

1. *Inner Loop*
 - a. *Detrending*.
 - b. *Cycle-subseries Smoothing*.
 - c. *Low-Pass Filtering of smoothed Cycle-Subseries*, Langkah ini dilakukan *filtering* dengan *moving average*.
 - d. *Detrending of Smoothed Cycle-Subseries*, Langkah ini dilakukan *filtering* dengan *moving average* dan untuk mendapatkan komponen seasonal.
 - e. *Deseasonalizing*
 - f. *Trend Smoothing*
2. *Outer Loop*

- a. Mencari nilai Residu yang sudah dilakukan dari *Inner Loop*

Detrending dilakukan dengan $Y - T$. *Cycle-subseries Smoothing* dilakukan dengan *smoothing* dengan menggunakan teknik LOESS. *Low-Pass Filtering of Smoothed Cycle Subseries* melakukan beberapa *moving average*. *Detrending of Smoothed Cycle-Subseries* melakukan *detrending* dengan cara $S = C - L$, Dengan L adalah hasil *Low-Pass Filtering of smoothed Cycle-Subseries* dan C adalah hasil *Cycle-subseries Smoothing*. Kemudian dilakukan *Deseasonalization* dengan rumus $Y - S$. Terakhir dilakukan *Trend Smoothing* (Cleveland *et al.* 1990).

Model STL memiliki beberapa keunggulan. Pertama, ia mampu menangani berbagai jenis pola musiman. Kedua, model ini mengizinkan perubahan komponen musiman seiring waktu dengan kontrol yang dapat diatur oleh pengguna. Selain itu, model STL bisa adaptasi tingkat kehalusan dari siklus tren sesuai preferensi. Keunggulan lainnya adalah tahan terhadap data *outlier*. Selain itu, model STL memiliki beberapa kelemahan. Salah satunya adalah ketidakmampuan model untuk secara otomatis menangani variasi yang berkaitan dengan peristiwa tertentu. Model ini juga berfokus hanya pada dekomposisi aditif yang dapat memengaruhi interpretasi terhadap komponen residu (Cleveland *et al.* 1990).

2.5 Extreme Gradient Boosting (XGBoost)

XGBoost merupakan suatu algoritma turunan dari *Gradient Boosting Tree* yang sudah dimodifikasi sehingga bisa menangani data yang sangat besar dan tentunya memiliki banyak fitur, menangani banyak data yang hilang, dan menangani banyak jenis variabel seperti numerik dan kategorikal. *Gradient Boosting Tree* merupakan salah satu metode *ensemble* dengan algoritma *gradient boosting* yang digunakan untuk memecahkan masalah machine learning seperti regresi, dan klasifikasi (Shi *et al.* 2018).

Gradient Tree Boosting adalah sebuah teknik machine learning yang kuat dan populer yang digunakan untuk membangun model prediksi yang akurat. Teknik ini menggabungkan pohon keputusan secara iteratif menggunakan metode boosting. Pohon keputusan ini bertindak sebagai *weak learners* yang diperkuat oleh langkah-langkah berurutan untuk membentuk model yang lebih kuat. Selama proses training, model berusaha mengoptimalkan fungsi loss-nya, yang berfungsi untuk mengukur sejauh mana prediksi model berbeda dari nilai sebenarnya pada data pelatihan. Fungsi loss ini mencerminkan perbedaan antara prediksi model dengan nilai aktual, dan tujuannya adalah untuk mencari kombinasi parameter dan bobot yang mengarahkan model ke arah yang lebih akurat. Fungsi loss adalah sebagai berikut:

$$L = \sum_{i=1}^n l(y_i, \hat{y}_i + f(x_i)) + \Omega(f) \quad (11)$$

$$\Omega(f) = \gamma T + \frac{\lambda \|w\|^2}{2} \quad (12)$$

L = Total Loss

l = Loss

y_i = Nilai sebenarnya

\hat{y}_i = Nilai prediksi

$f(x_i)$ = Nilai prediksi pada satu pohon

$\Omega(f)$ = Fungsi penalti

γ = Parameter *regularization* seluruh model

λ = Parameter *regularization* pada satu pohon

w = *Weight* pada satu pohon

T = Banyaknya daun pada satu pohon

Persamaan (11) menunjukkan bahwa *loss* bergantung dengan nilai praduga dengan nilai asli. Persamaan 12 berfungsi sebagai *regularized* agar dapat menghindari *overfitting*. Persamaan tersebut bergantung pada pohon satu dengan yang lain.

$$L = \sum_{i=1}^n [g_i f(x_i) + h_i f^2(x_i)] + \Omega(f) \quad (13)$$

$$g_i = \delta_y l(y_i, \hat{y}_i) \quad (14)$$

$$h_i = \delta_y^2 l(y_i, \hat{y}_i) \quad (15)$$

Persamaan 11 disederhanakan dengan menggunakan ekspansi Taylor menjadi persamaan 13. Persamaan 14 dan Persamaan 15 juga didapatkan dari ekspansi Taylor. *Gradient statistics* digunakan untuk mengubah *weight* (Chen dan Guestrin 2016).

$$\begin{aligned} L &= \sum_{i=1}^n [g_i f(x_i) + h_i f^2(x_i)] + \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T w_j \\ &= \sum_{j=1}^T [(\sum_{i \in I_j} g_i) w_j + \frac{1}{2} (\sum_{i \in I_j} h_i + \lambda) w_j^2] + \gamma T \end{aligned} \quad (16)$$

$$w_j = - \frac{\sum_{i \in I_j} g_i}{\sum_{i \in I_j} h_i + \lambda} \quad (17)$$

Persamaan 16 merupakan pengembangan persamaan 13 dengan perhitungan antar daun. Persamaan 16 diminimalkan agar mendapatkan model optimum. Berdasarkan hal tersebut, persamaan 17 diturunkan dari persamaan 16.

$$L = -\frac{1}{2} \sum_{j=1}^T \frac{(\sum_{i \in I_j} g_i)^2}{\sum_{i \in I_j} h_i + \lambda} + \gamma T \quad (18)$$

Persamaan 18 merupakan persamaan akhir yang digunakan untuk mengoptimalkan model. Persamaan 18 merupakan nilai evaluasi seperti *decision tree*. XGBoost adalah pustaka yang dirancang untuk melatih model pembelajaran mesin dengan efisien dan dapat diskalakan. XGBoost merupakan singkatan dari eXtreme Gradient Boosting. Pustaka ini menggunakan beberapa fitur dan algoritma untuk mempercepat proses latih seperti *Approximate Algorithm*, *Sparsity-aware Split Finding*, *Parallel Learning*, *Cache-aware Access*, dan *Blocks for Out-of-core Computation* (Chen dan Guestrin 2016).

Approximate Algorithm pada XGBoost digunakan untuk mendukung *gradient tree boosting* pada dua pengaturan yaitu data yang sangat besar dan data yang sangat jarang. Algoritma ini mempercepat proses perhitungan *gain* untuk setiap split yang mungkin pada setiap fitur dengan menghitung nilai *gain* secara aproksimasi. Hal ini dilakukan dengan mengambil sampel dari data latih dan menghitung nilai *gain* pada sampel tersebut. Dalam hal ini, XGBoost menggunakan algoritma subsampling untuk mengambil sampel dari data latih (Chen dan Guestrin 2016).

Sparsity-aware Split Finding pada XGBoost digunakan untuk menangani sparsitas pada data seperti keberadaan data yang hilang, entri nol padat, dan nilai yang di-*encode* satu-per-satu. Algoritma ini membuat XGBoost menyadari sparsitas tersebut dengan mendefinisikan arah default untuk mereka. Dalam hal ini, XGBoost mencoba kedua arah dalam split dan menemukan arah default dengan menghitung *gain* (Chen dan Guestrin 2016).

Parallel Learning pada XGBoost digunakan untuk mempercepat proses pelatihan model dengan membagi data menjadi beberapa bagian dan melatih model pada setiap bagian secara independen. Hal ini dilakukan dengan menggunakan beberapa thread untuk mempercepat proses pelatihan (Chen dan Guestrin 2016).

Cache-aware Access pada XGBoost digunakan untuk mempercepat proses perhitungan skor kesamaan dan nilai output dengan menggunakan memori cache pada sistem. Hal ini dilakukan dengan mengambil data dari cache memory dan menghitung skor kesamaan dan nilai output pada data tersebut (Chen dan Guestrin 2016).

Blocks for Out-of-core Computation dalam XGBoost adalah fitur yang memungkinkan pelatihan model dengan dataset besar yang melebihi kapasitas memori. Fitur ini membagi data pelatihan menjadi blok-blok yang lebih kecil untuk penggunaan memori yang lebih efisien. Selain itu, XGBoost juga menggunakan kompresi blok untuk mengurangi ukuran blok dalam memori dengan algoritma kompresi, serta pemecahan blok untuk memproses blok secara

paralel. Dengan kombinasi fitur ini, XGBoost mampu melatih model pada dataset besar, mengatasi kendala memori, dan meningkatkan skalabilitas pelatihan model *machine learning* (Chen dan Guestrin 2016).

2.6 Perbandingan Model peramalan harga emas

Faktor ekonomi seperti tingkat inflasi, suku bunga, dan kondisi perekonomian global memiliki dampak yang signifikan pada harga emas (Anandasayanan *et al.* 2019). Tingkat inflasi dan suku bunga memiliki korelasi negatif dengan harga emas (Anandasayanan *et al.* 2019). Spekulasi pada tingkat kepemilikan emas pada masa depan dianggap penting (Farhat dan Ghalayini 2020). Metrik yang dibandingkan adalah akurasi karena kebanyakan artikel menggunakan akurasi sebagai tolak ukur kebaikan model peramalan. perbandingan model peramalan bisa dilihat di tabel berikut

Tabel 1 Akurasi penelitian sebelumnya

No	Nama Model	Dataset	Sumber data	Negara	Referensi
1.	XGBoost	Jan 1986 - Des 2019	<i>Goldhub</i>	Dunia	Jabeur <i>et al.</i> 2022
2.	MLP with Whale Optimizati on	Sept 1987 - Agu 2017	<i>World Bank</i>	Dunia	Alameer <i>et al.</i> 2019
3.	DWT-SVR	Jan 1992 - December 2016	FRED	Dunia	Risse 2019
4.	LSTM with Pandemic	30 Jul 2015 - 22 Jan 2020	<i>Yahoo Finance</i>	Dunia	Khani <i>et al.</i> 2022
5.	STL	3 Jan 2006 - 31 Jan 2019	<i>Goldhub</i>	Dunia	Chai <i>et al.</i> 2021
6.	Facebook Prophet	2009 - 2020	<i>Goldhub</i>	Dunia	Kishann dan RamaParvathy 2022
7.	STL, XGBoost, Facebook Prophet	04 Januari 2010 - 31 April 2023	Bank Indonesia dan Logam Mulia	Indonesia	Penelitian ini

Tabel 1 ditunjukkan akurasi yang digunakan adalah *R-Square*. Berdasarkan akurasi dan data latih, model *XGBoost* dari penelitian Jabeur *et al.* (2022) yang terbaik karena akurasi yang tinggi dan rentang data yang besar.

2.7 Granger Causality

Granger causality adalah sebuah metode statistik yang digunakan untuk menentukan apakah suatu variabel dapat digunakan untuk memprediksi variabel lainnya. Metode ini ditemukan oleh Clive Granger pada tahun 1969 dan telah

banyak digunakan dalam berbagai bidang seperti ekonomi, keuangan, genetika, dan neurosains. Granger causality memungkinkan kita untuk menentukan apakah suatu variabel memiliki pengaruh kausal pada variabel lainnya. Dalam Granger causality, pengaruh kausal diukur dengan melihat seberapa baik variabel yang menjadi prediktor dapat memprediksi variabel target. Jika prediktor tersebut dapat memprediksi target dengan lebih baik daripada model yang hanya menggunakan data target saja, maka kita dapat menyimpulkan bahwa prediktor tersebut memiliki pengaruh kausal pada target (Shojaie dan Fox 2021).

Granger mendefinisikan *kausalitas* seperti nilai masa lalu dari deret waktu y_t memprediksi nilai masa depan dari nilai deret x_t . Granger mengatakan y itu ‘kausal’ untuk x jika $var[x_t - P(x_t | H_{<t})] < var[x_t - P(x_t | H_{<t} \setminus y_t)]$, $H_{<t}$ artinya nilai jaman dulu yang relevan dengan informasi hingga waktu $t - 1$. $P(x_t | H_{<t})$ adalah nilai prediksi optimal x_t dengan diberikan nilai H_t .

2.8 Symmetric Mean Absolute Percentage Error (SMAPE)

Symmetric Mean Absolute Percentage Error (SMAPE) adalah ukuran statistik yang digunakan untuk mengevaluasi akurasi prediksi dalam analisis deret waktu. Ini adalah variasi dari Mean Absolute Percentage Error (MAPE) yang simetris di sekitar nol. SMAPE dihitung sebagai berikut:

$$SMAPE = \frac{100}{n} \sum_{t=1}^n \frac{|F_t - A_t|}{(|A_t| + |F_t|)/2} \quad (19)$$

t = Index

A = Actual Value

F = Forecast Value

n = Banyak data

Persamaan 19 menjelaskan tentang rumus SMAPE. SMAPE memiliki rentang nilai antara 0 dan 2, dimana nilai 0 menunjukkan akurasi sempurna, sedangkan nilai 2 menunjukkan bahwa ramalan tersebut salah (Armstrong 1985).

SMAPE memiliki potensi untuk menghasilkan kesalahan simetris bahkan dalam rentang peramalan yang asimetris. Salah satu keunggulan dari SMAPE adalah tidak mengalami masalah seperti MAPE yang dapat menghasilkan nilai yang sangat besar atau tak terhingga. Selain itu, karena adanya batas kesalahan yang telah ditentukan. SMAPE lebih tahan terhadap nilai-nilai pencilan dengan memberikan penekanan yang lebih sedikit pada nilai-nilai ekstrem dibandingkan dengan metrik lain yang tidak memiliki batasan kesalahan (Chen 2017).

Nilai SMAPE memiliki interpretasi yang sama dengan MAPE. Nilai SMAPE dikatakan *highly accurate forecasting* jika memiliki nilai kurang dari 10%. Nilai SMAPE dikatakan *good forecasting* jika memiliki nilai dari 10%

hingga 20%. Nilai SMAPE dikatakan *reasonable forecasting* jika memiliki nilai 20% hingga 50%. Nilai SMAPE dikatakan *inaccurate forecasting* jika memiliki nilai lebih dari 50% (Hu *et al.* 2022).

2.9 Mean Absolute Scaled Error (MASE)

Mean Absolute Scaled Error (MASE) mengatasi keterbatasan kesalahan relatif tradisional dalam menilai akurasi ramalan. Berbeda dengan kesalahan relatif dengan rata-rata yang tidak terdefinisi dan variasi tak terhingga. MASE menskalakan kesalahan menggunakan Mean Absolute Error (MAE) dalam sampel dari metode *naive forecast*. MASE dapat dihitung dengan:

$$MASE = \frac{\frac{mean(A_t - F_t)}{n}}{\frac{1}{n-1} \sum_{t=2}^n |A_t - A_{t-1}|} \quad (20)$$

t = Index

A = Actual Value

F = Forecast Value

n = Banyak data

Persamaan 20 menunjukan mengenai rumus MASE. menunjukkan bahwa nilai MASE berkisar 0 hingga tak terbatas. MASE (*Mean Absolute Scaled Error*) membandingkan akurasi ramalan dengan akurasi naive forecast. Jika MASE kurang dari satu, itu berarti bahwa ramalan lebih baik daripada *naive forecast* (Hyndman 2006). MASE sering digunakan karena ketahanannya terhadap pencilon, interpretasi yang lebih baik dibandingkan RMSSE, dan variabilitas yang lebih rendah dalam sampel kecil dibandingkan MdASE (Hyndman dan Koehler 2006).

III METODE PENELITIAN

3.1 Alat dan Bahan

Alat dan bahan yang digunakan untuk penelitian adalah sebagai berikut:

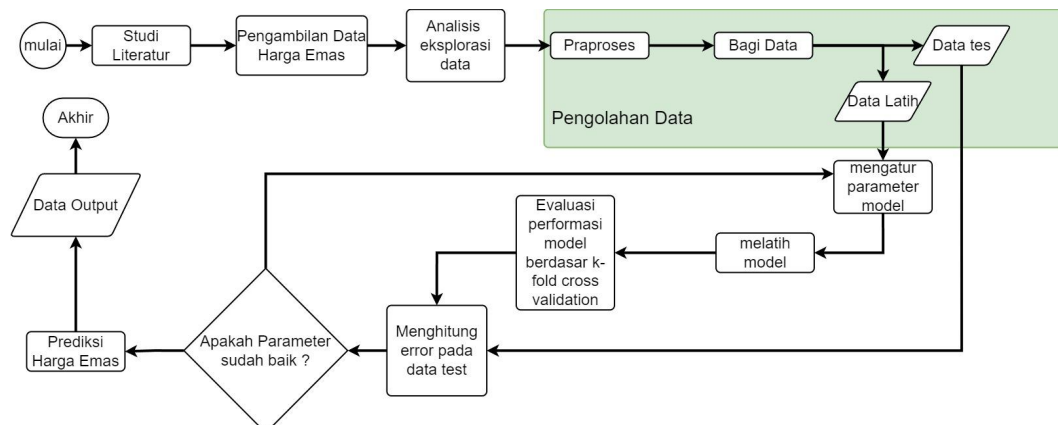
1. Python 3.9.16
2. Scikit learn 1.2.2
3. Pandas 1.4.4
4. Seaborn 0.11.2
5. Matplotlib 3.5.3
6. Pycaret 3.0.4

Penelitian ini menggunakan Python sebagai bahasa pemrograman yang populer dan fleksibel. Pustaka Scikit-learn dipakai untuk mengukur kinerja model dengan berbagai metrik evaluasi. Pustaka Pandas digunakan untuk mengolah data tabular dengan mudah. Pustaka Seaborn dan Matplotlib dipilih untuk menampilkan data secara grafis dengan fungsi visualisasi data yang kaya dan menarik. Pustaka Pycaret diandalkan untuk melaksanakan proses machine learning secara efisien dan mudah dengan fitur-fitur otomatisasi dan optimisasi *machine learning*.

3.2 Tahapan Penelitian

Tahapan penelitian adalah sebagai berikut:

1. Studi literatur
2. Pengambilan Data Harga emas
3. Analisis Eksplorasi data
4. Pengolahan data
 - a. Praproses data
 - b. Pembagian dataset
5. Melatih data dengan 3 model
6. *Tuning hyperparameter*
7. Evaluasi model
8. Peramalan model



Gambar 2 Alur Penelitian

Pada Gambar 2 menampilkan alur penelitian. Alur penelitian meliputi studi literatur hingga meramal harga emas. Alur penelitian berdasarkan referensi dari penelitian Nawaz *et al.* (2021). Penjelasan mengenai Alur penelitian akan diterangkan pada subbab selanjutnya.

3.2.1 Studi literatur

studi literatur dilakukan dengan mengkaji penelitian sebelumnya terkait artikel yang berhubungan dengan model Facebook Prophet untuk meramal harga emas (Taylor & Letham 2017) dengan faktor eksternal. Beberapa penelitian mengenai penerapan model *machine learning* untuk peramalan harga emas dan mencari korelasi pada faktor eksternal yang memengaruhi harga emas seperti penelitian Khani *et al.* (2022).

3.2.2 Pengambilan Data

Pengambilan data harga emas diperoleh dengan memanfaatkan Highcharts API melalui *Console* di *browser* untuk mengambil data. Data suku bunga, inflasi, dan kurs mata uang diperoleh dari website Bank Indonesia (BI).

3.2.3 Analisis Eksplorasi Data

Analisis eksplorasi data dilakukan untuk mendapatkan wawasan yang lebih mendalam mengenai dataset yang digunakan. Tujuan utamanya adalah untuk melihat persebaran data secara menyeluruh dan mengeksplorasi hubungan antara variabel-variabel yang relevan. Salah satu aspek penting dari analisis ini adalah visualisasi data berdasarkan tanggal, yang memungkinkan kita untuk melihat tren dan pola yang berkaitan dengan waktu. Dengan demikian, hasil dari analisis eksplorasi data ini akan membantu dalam menentukan praproses yang tepat yang diperlukan sebelum melakukan tahapan lebih lanjut dalam proses analisis atau pemodelan data. Praproses yang tepat akan memastikan bahwa data siap untuk digunakan secara optimal, mengurangi adanya *noise* atau *outlier*, dan memastikan validitas hasil analisis yang akan diperoleh.

3.2.4 Pengolahan data

Dalam analisis data ini, dilakukan praproses menggunakan beberapa metode, termasuk imputasi data dengan metode *drift* menggunakan Polynomial Trend Forecaster dari pustaka *sktime*. Metode *drift* ini memungkinkan kita untuk mengisi nilai yang hilang dalam dataset berdasarkan tren polinomial data sebelum dan sesudah nilai yang hilang tersebut. Selain itu, dilakukan pembagian data menggunakan metode *cross-validation k-fold* sebanyak 3 kali, yang membagi dataset menjadi tiga subset yang berbeda untuk melatih dan menguji model secara berulang. Pendekatan ini membantu menghindari *overfitting* dan memberikan gambaran yang lebih objektif tentang kinerja model pada dataset yang berbeda. Hasil dari praproses data dan *cross-validation* ini akan memberikan landasan yang

solid untuk analisis lebih lanjut atau pemodelan data yang dapat dipercaya dan relevan.

3.2.5 Melatih data dengan 3 model

Dalam pelatihan data ini, tiga model berbeda digunakan untuk peramalan data dengan tingkat akurasi yang tinggi. Model pertama adalah Facebook Prophet, yang memiliki kemampuan untuk mengatasi tren non-linear, efek musiman, dan perubahan struktural dalam data (Taylor dan Letham 2017). Model kedua adalah XGBoost, sebuah model berbasis gradient boosting tree yang sangat kuat dalam menghasilkan peramalan yang akurat (Chen dan Guestrin 2016). Sementara itu, model ketiga yang digunakan adalah STL, yang menggabungkan metode smoothing dengan pendekatan LOESS untuk mengidentifikasi komponen musiman, tren, dan residu dalam data (Cleveland *et al* 1990). Dengan mengadopsi ketiga model ini, diharapkan dapat memberikan peramalan yang lebih handal dan akurat, serta meningkatkan kualitas analisis dan pengambilan keputusan dalam konteks peramalan data.

3.2.6 Tuning parameter

Proses tuning parameter sangat penting untuk meningkatkan performa dan akurasi ketiga model peramalan ini. Pada model Facebook Prophet, parameter *changepoint_prior_scale*, *holidays_prior_scale*, dan *seasonality_prior_scale* perlu disesuaikan. Sedangkan pada model XGBoost, tuning parameter melibatkan *learning_rate*, *depth*, dan *n_estimators*. Untuk model STL, parameter-parameter terkait *smoothing* dan tingkat keakuratan dalam pendekatan LOESS harus diatur. *library* pycaret dapat mencari kombinasi parameter terbaik untuk 3 model (Ali, 2020).

Tabel 2 Tuning parameter Facebook Prophet

Parameter	Prophet Multivariate	Prophet Univariate
<i>changepoint_prior_scale</i>	0,2176624112345368	0,016439324494196616
<i>holidays_prior_scale</i>	0,6358358856676251	0,01095960453692584
<i>seasonality_prior_scale</i>	0.8967376801947962	7,886714129990491

Pada Tabel 2 ditunjukkan bahwa ada perubahan parameter pada *Multivariate*. Pada parameter *univariate*, parameter yang digunakan adalah parameter *default*. Pada nilai parameter *multivariate* bernilai koma tidak beraturan karena dalam tuning parameter menggunakan metode *random search*.

Tabel 3 Tuning parameter STL

Parameter	STL Univariate
<i>seasonal</i>	7
<i>seasonal_deg</i>	1
<i>seasonal_jump</i>	1
<i>sp</i>	261
<i>trend</i>	None
<i>trend_deg</i>	1
<i>trend_jump</i>	1

Pada Tabel 3 ditunjukkan bahwa parameter memiliki peran penting terhadap model STL. Parameter *seasonal* merupakan parameter untuk *smoothing seasonal* dan harus bernilai ganjil. Parameter *seasonal_deg* merupakan derajat untuk *seasonal LOESS*. Parameter *seasonal_jump* merupakan parameter yang menentukan langkah interpolasi secara linear pada *seasonal*. Parameter *trend* merupakan parameter untuk *smoothing trend* dan harus bernilai ganjil. Parameter *trend deg* merupakan derajat untuk tren. Parameter *trend_jump* merupakan parameter yang menentukan langkah interpolasi secara linear pada tren. Parameter *sp* merupakan periode pada data tersebut. STL hanya bisa digunakan *univariate*.

Tabel 4 Tuning parameter XGBoost

Parameter	XGBoost Multivariate	XGBoost Univariate
<i>learning_rate</i>	0,3	0,3
<i>max_depth</i>	6	6
<i>max_leaves</i>	0	0
<i>n_estimators</i>	100	100

Pada Tabel 4 ditunjukkan bahwa ada beberapa parameter yang perlu diketahui. Pertama, *learning_rate* adalah ukuran langkah yang digunakan dalam setiap pembaruan untuk mencegah *overfitting*. Parameter *max_depth* menentukan kedalaman maksimum dari pohon yang dibangun, dengan nilai yang lebih tinggi akan membuat model lebih kompleks dan berpotensi *overfitting*. Selanjutnya, parameter *max_leaves* mengontrol jumlah maksimum simpul yang ditambahkan,

tetapi tidak digunakan dalam metode pohon. Terakhir, $n_estimators$ merupakan jumlah pohon *gradient boost* yang akan dibangun.

3.2.7 Evaluasi Model

Untuk membandingkan model forecasting, terdapat beberapa metode evaluasi statistik yang dapat digunakan seperti *Symmetric Mean Absolute Percentage Error* (SMAPE) dan *Mean Absolute Scaled Error* (MASE). SMAPE dipilih karena tahan terhadap *outlier* dan memiliki sifat *scale-independence* yaitu, pengukuran tidak berpengaruh dengan skala data (Chen *et al.* 2017).

3.2.8 Peramalan Data

Peramalan data dilakukan untuk meramalkan data hingga 2 tahun ke depan. Hasil peramalan tersebut akan disimpan dalam bentuk file CSV. Proses peramalan menggunakan ketiga model yakni, model Facebook Prophet, XGBoost, dan STL.

3.3 Deskripsi Data

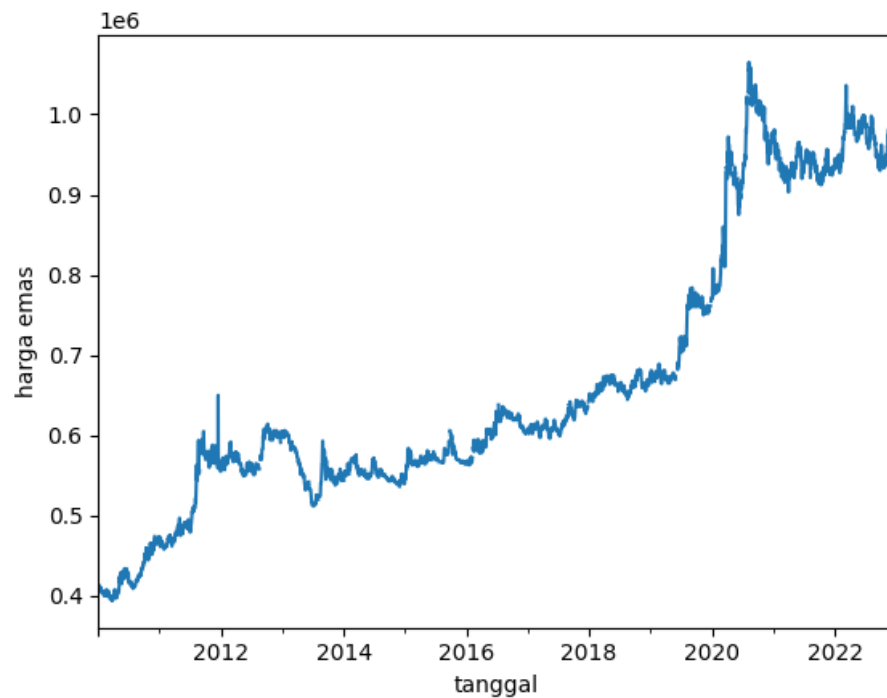
Data yang digunakan dalam penelitian ini diperoleh dari situs Logam Mulia PT.Antam (PT.Antam 2022). Pada penelitian ini, data yang digunakan adalah data dari tanggal 4 Januari 2010 hingga 30 April 2023. Harga emas dihitung berdasarkan harga emas dari mata uang rupiah per 1 gram emas. Kurs Beli dollar, kurs jual dollar, inflasi, dan suku bunga diambil dari data Bank Indonesia yakni dari situs informasi kurs BI (<https://www.bi.go.id/id/statistik/informasi-kurs/transaksi-bi/default.aspx>), data inflasi BI (<https://www.bi.go.id/id/statistik/indikator/data-inflasi.aspx>), dan Suku bunga BI (<https://www.bi.go.id/id/statistik/indikator/bi-7day-rr.aspx>).

Tabel 5 Master data

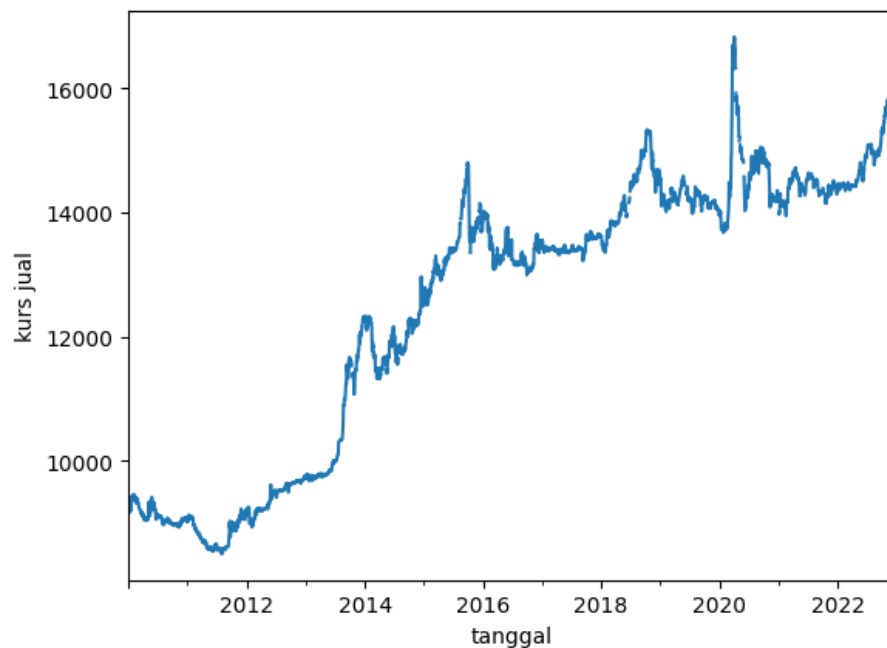
Tanggal	Harga Emas (Rp)	Kurs Jual Dollar (Rp)	Kurs Beli Dollar (Rp)	Data Inflasi (%)	Suku Bunga BI (%)
04-01-2010	408.000	9.377	9.283	3,72	6,5
05-01-2010	410.000	9.355	9.261	3,72	6,5
06-01-2010	410.000	9355	9.261	3,72	6,5
07-01-2010	412.000	9274	9182	3,72	6,5
08-01-2010	410.000	9286	9194	3,72	6,5

Tabel 5 yang merupakan tabel master menunjukkan bahwa data yang diperoleh terdiri atas enam *field* utama. Secara total, *record* data yang diperoleh sebanyak 3412 baris. Enam *field* data meliputi *Tanggal*, *Harga Emas*, *Kurs Jual Dollar*, *Kurs Beli Dollar*, *Data Inflasi*, dan *Suku Bunga BI*. *Field* Harga Emas

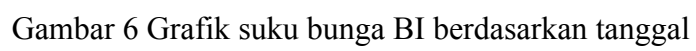
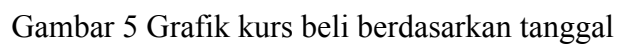
menyatakan nilai harga jual emas ANTAM per 1 gram. *Field* kurs jual dollar menyatakan kurs yang ditawarkan oleh bank untuk menjual mata uang dollar (Yusnaini 2023). *Field* kurs beli dollar menyatakan kurs dimana bank bersedia membeli mata uang (Yusnaini 2023). *Field* data inflasi menyatakan persentase kenaikan harga barang dan jasa secara umum dalam jangka waktu tertentu di Indonesia. *Field* suku bunga BI menyatakan suku bunga acuan BI untuk memberikan peminjaman terhadap bank komersial (*repo rate*).

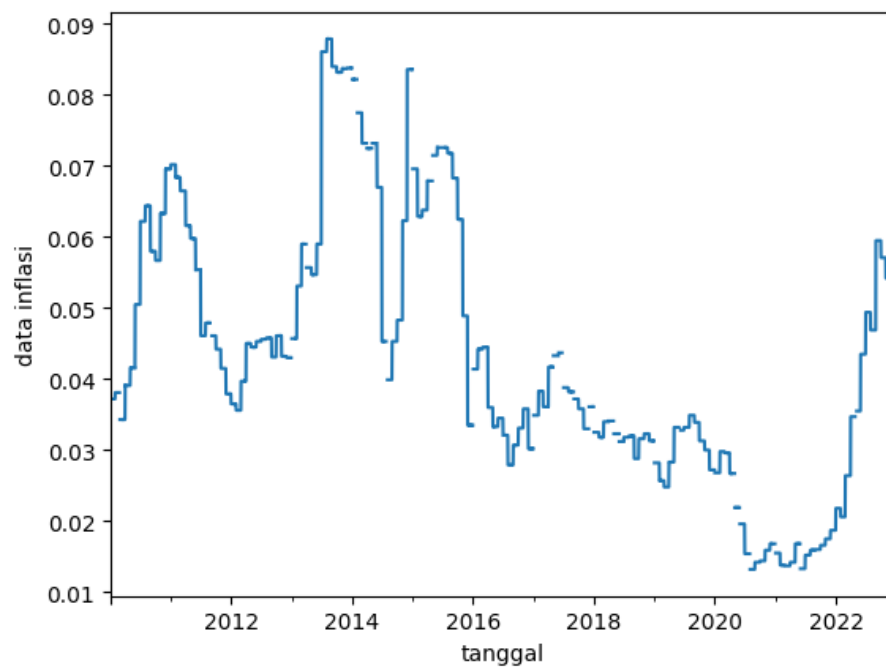


Gambar 3 Grafik harga emas berdasarkan tanggal



Gambar 4 Grafik kurs jual berdasarkan tanggal



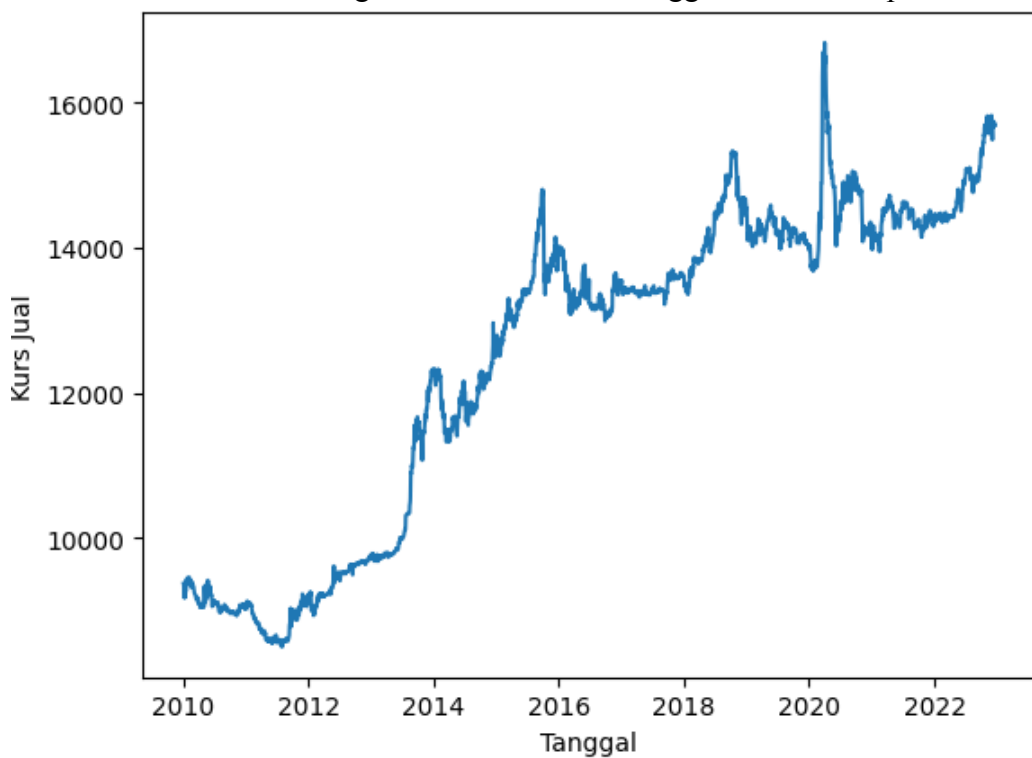


Gambar 7 Grafik inflasi di Indonesia berdasarkan tanggal

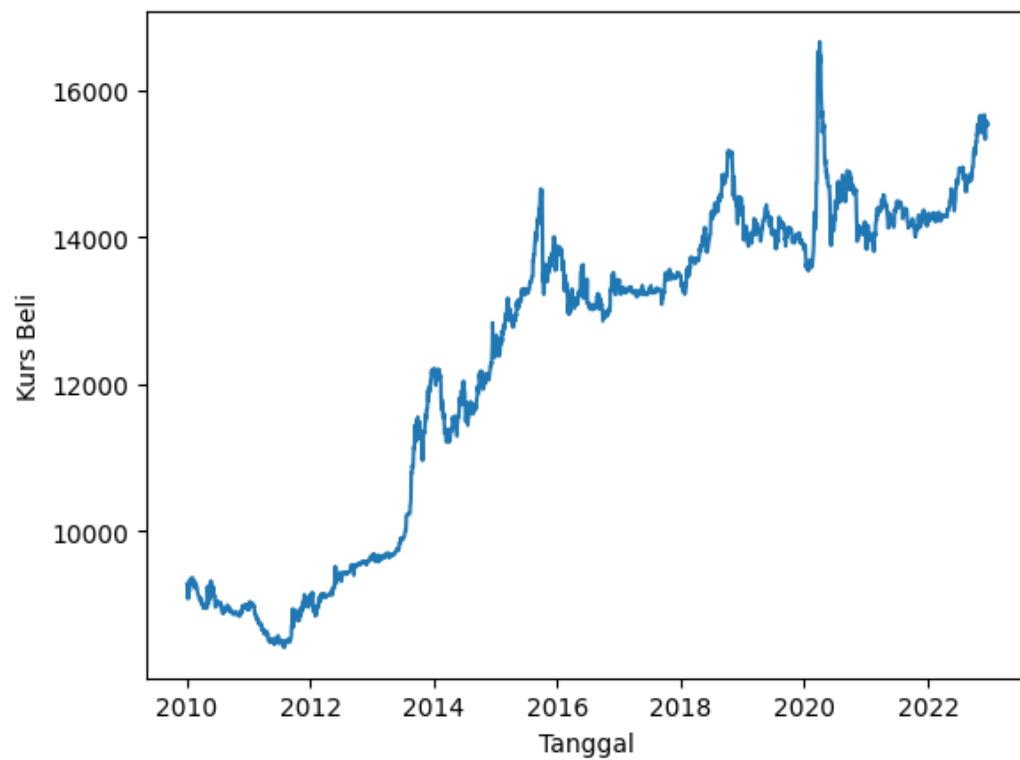
Pada Gambar 3, terlihat peningkatan yang signifikan dalam harga emas dari tahun 2019 hingga 2020. Grafik ini menunjukkan kenaikan yang stabil. Gambar 4 dan Gambar 5 menampilkan grafik yang serupa untuk kurs jual dan kurs beli. Titik puncak untuk keduanya terjadi pada tahun 2020. Pada Gambar 6, grafik suku bunga terlihat tidak lengkap. Nilai tertinggi tercatat dari tahun 2014 hingga 2016. Pada Gambar 7, inflasi mengalami peningkatan pada tahun 2013-2014 dan kemudian mengalami penurunan pada tahun 2020-2021.



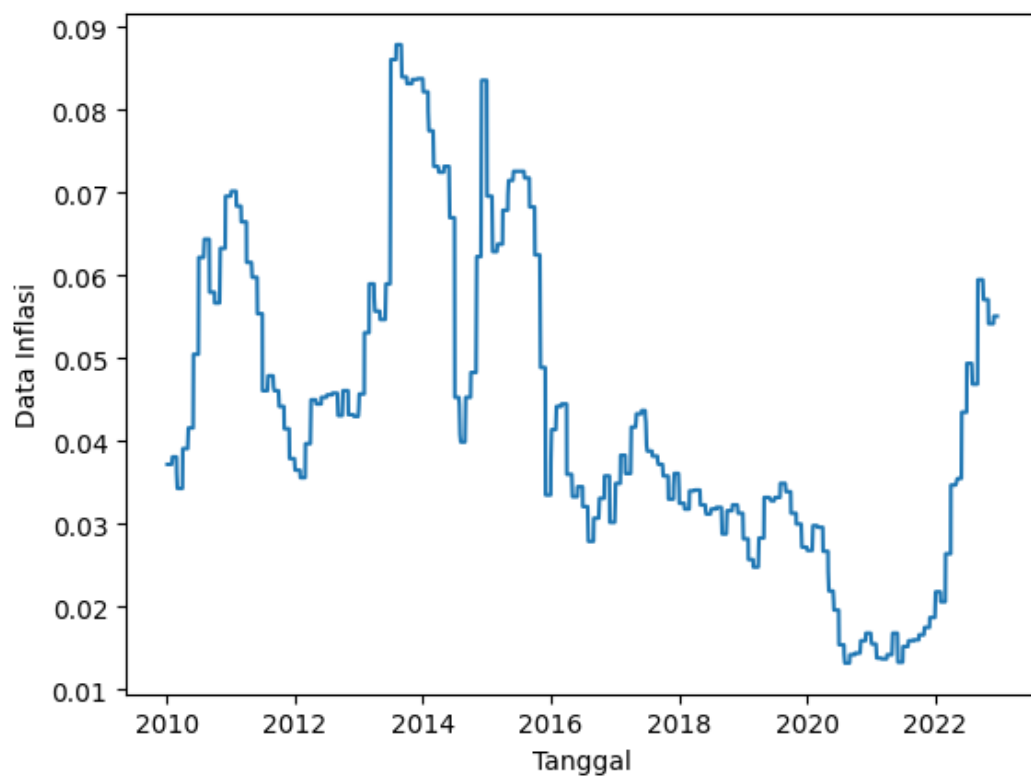
Gambar 8 Grafik harga emas berdasarkan tanggal setelah diimputasi



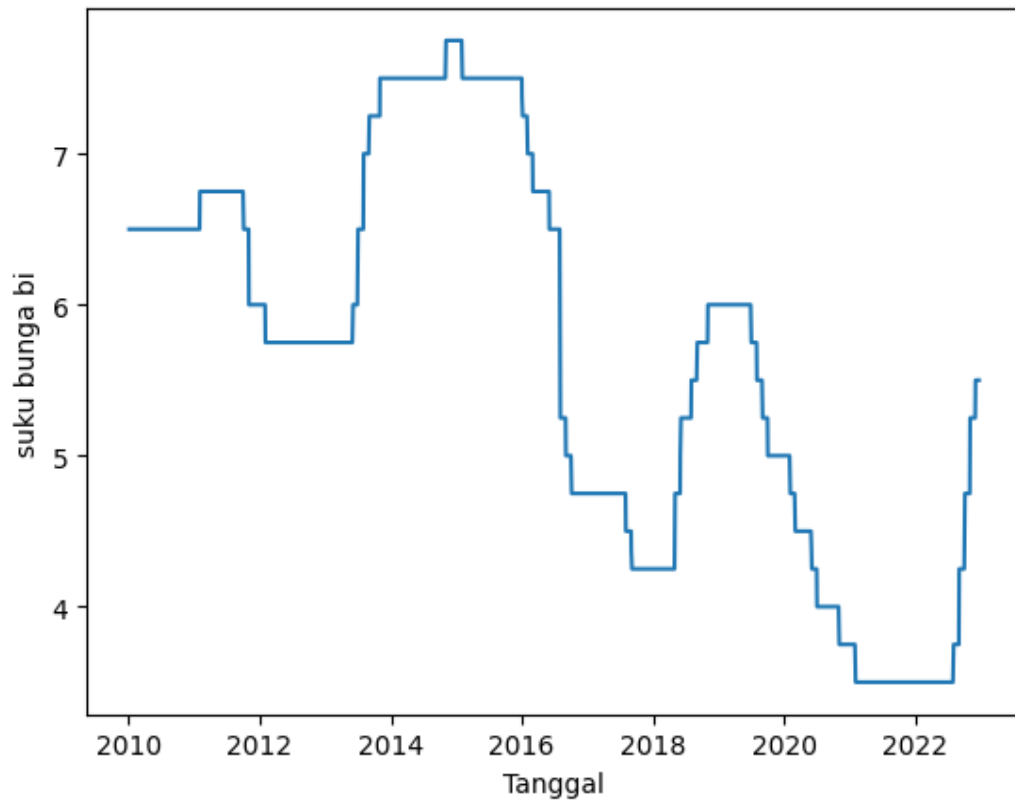
Gambar 9 Grafik kurs jual berdasarkan tanggal setelah diimputasi



Gambar 10 Grafik kurs beli berdasarkan tanggal setelah diimputasi

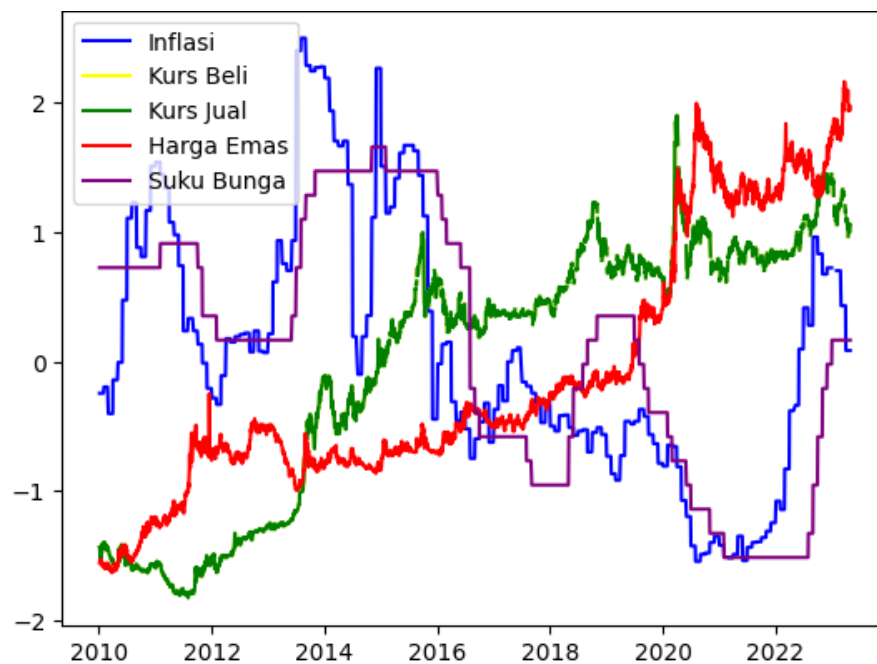


Gambar 11 Grafik data inflasi berdasarkan tanggal setelah diimputasi



Gambar 12 Grafik suku bunga BI berdasarkan tanggal setelah diimputasi

Pada Gambar 8 hingga 12, grafik telah diimputasi dengan pendekatan interpolasi dengan metode *drift* (seperti interpolasi linear). Interpolasi dilakukan dengan tanggal sebagai variabel independen.



Gambar 13 Grafik normalisasi master data

Gambar 13 menunjukkan bahwa grafik kurs beli dan jual dollar terhadap rupiah menunjukkan tren yang cenderung berdekatan, meskipun terdapat selisih kecil antara keduanya. Sementara itu, grafik inflasi dan suku bunga juga menunjukkan adanya hubungan yang berdekatan, meskipun inflasi cenderung lebih fluktuatif. Pada tahun 2020 menunjukkan bahwa harga emas dan kurs jual memiliki puncak yang berdekatan.

3.4 Skenario Eksperimen

Skenario eksperimen dilakukan dengan menggunakan tiga model *machine learning*, yakni XGBoost, STL, dan Facebook Prophet. Skenario ini menggunakan data *multivariate* (data yang memiliki lebih dari satu variabel) dan *univariate* (data yang memiliki satu variabel). Validasi data dilakukan dengan metode *cross-validation* khusus data deret waktu dengan menggunakan *expanding window* dengan *fold* sebanyak 3 kali untuk menghindari *overfitting* dan *underfitting*. Penelitian Kingpai dan Moshfeghi (2023) menemukan bahwa model yang menggunakan *cross validation* dengan menggunakan metode *expanding window* lebih baik dibandingkan *rolling window*. Penelitian Dhake *et al.* (2023) mendapatkan akurasi lebih tinggi dengan *Tuning hyperparameter*. *Tuning hyperparameter* dilakukan pada ketiga model dengan menggunakan metode *random search*. Tabel skenario eksperimen adalah sebagai berikut :

Tabel 6 Skenario Eksperimen

Model	Data	Fold			
		Nama Fold	Training	Validation	Test
Facebook Prophet	Univariate	Fold 1	555	730	730
		Fold 2	1285	730	730
		Fold 3	2015	730	730
	Multivariate	Fold 1	555	730	730
		Fold 2	1285	730	730
		Fold 3	2015	730	730
STL	Univariate	Fold 1	555	730	730
		Fold 2	1285	730	730
		Fold 3	2015	730	730
XGBoost	Univariate	Fold 1	555	730	730
		Fold 2	1285	730	730
		Fold 3	2015	730	730
	Multivariate	Fold 1	555	730	730
		Fold 2	1285	730	730
		Fold 3	2015	730	730

Tabel 6 menunjukkan bahwa keseluruhan skenario eksperimen ada 15. Pada data *training* dan *validation* bersifat berubah-ubah berdasarkan metode *expanding window*, tetapi pada data *test* tidak ada yang berubah.

3.5 Lingkungan Penelitian

Spesifikasi perangkat lunak dan perangkat keras yang akan digunakan untuk penelitian ini adalah sebagai berikut:

1. Perangkat Keras:

- Memory memiliki RAM 8 GB
- Processor memiliki tipe AMD Ryzen 5 3500U dengan Radeon Vega Mobile Gfx (8 CPU) ~ 2.1 GHz
- GPU menggunakan tipe AMD Radeon(TM) Vega 8 Graphics

2. Perangkat Lunak:

- Sistem Operasi menggunakan Windows 11
- Bahasa Pemrograman menggunakan Python

IV HASIL DAN PEMBAHASAN

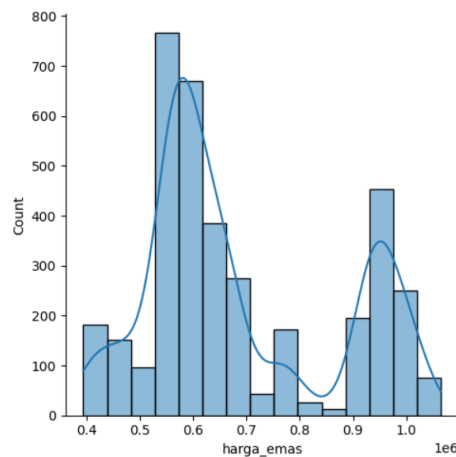
4.1 Analisis Statistik Data

Analisis statistik adalah proses pemrosesan data yang melibatkan penyaringan, pembersihan, transformasi, pemrosesan, dan pemodelan data. Tujuannya adalah untuk menemukan informasi yang berguna yang dapat digunakan sebagai dasar untuk kesimpulan atau keputusan. (Sahabuddin *et al.* 2021).

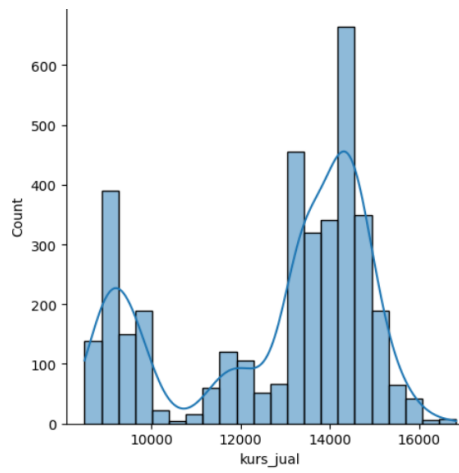
Tabel 7 Statistik Master Data

	harga_emas	kurs_jual	kurs_beli	data_inflasi	suku_bunga_bi
mean	668946.1	12535.56297	12410.79195	0.042967	5.637697
median	610000	13367.5	13233.5	0.041500	5.75

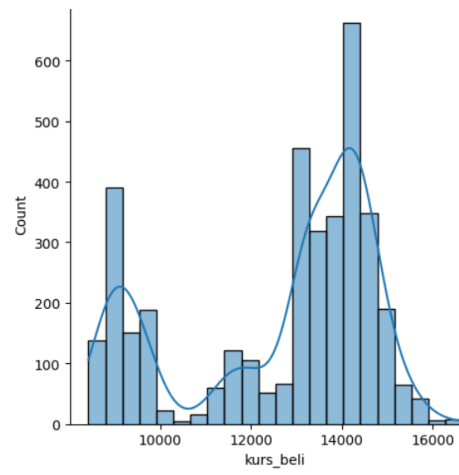
Tabel 7 merupakan contoh tabel data yang berisi nilai rata-rata dan median untuk harga emas, kurs jual, kurs beli, data inflasi, dan suku bunga. Hasil analisis menunjukkan bahwa nilai rata-rata dan median pada fitur tersebut cenderung mendekati nilai mean yang berarti bahwa distribusinya cenderung mendekati distribusi normal. Pendekatan distribusi normal pada data penting dalam analisis statistik dan model machine learning karena memungkinkan penggunaan teknik statistik lanjutan dengan asumsi distribusi yang terdefinisi.



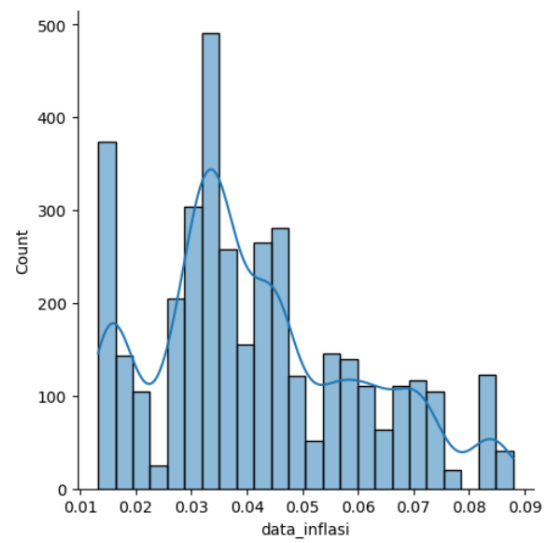
Gambar 14 Persebaran data harga emas



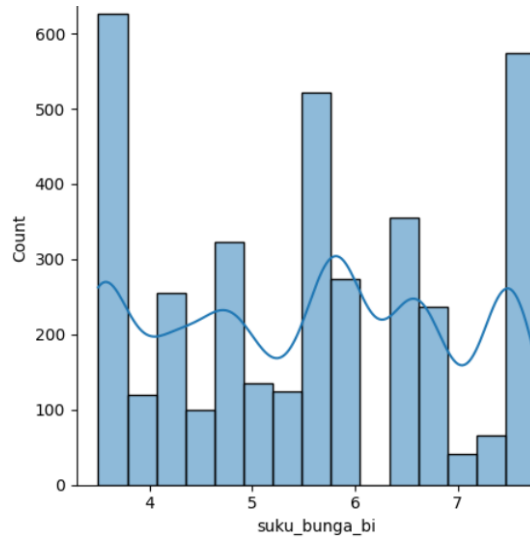
Gambar 15 Persebaran kurs jual dollar



Gambar 16 Persebaran kurs beli dollar

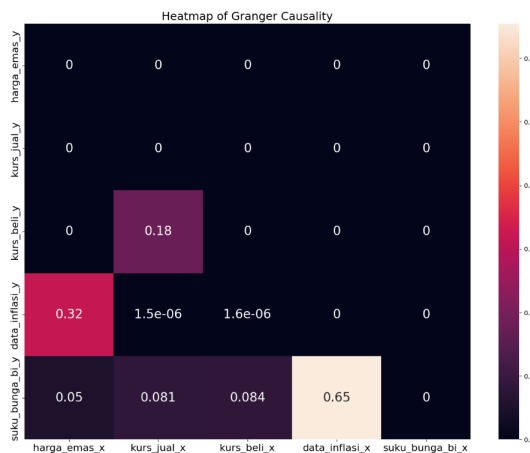


Gambar 17 Persebaran data inflasi



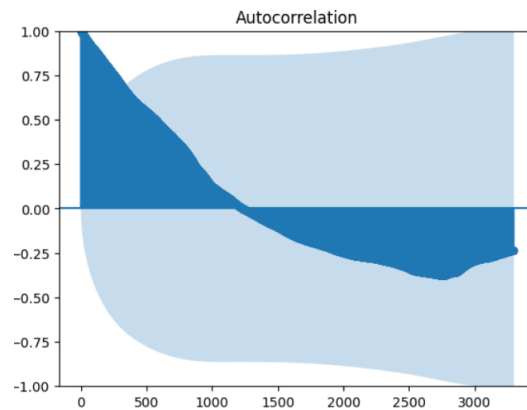
Gambar 18 Persebaran data suku bunga BI

Pada Gambar 14 - 18 terdapat pola distribusi bimodal dalam grafik harga emas, kurs jual, dan kurs beli yang ditandai dengan dua puncak data. Kondisi ini berpengaruh dalam analisis statistik untuk penentuan model yang sesuai dengan asumsi.

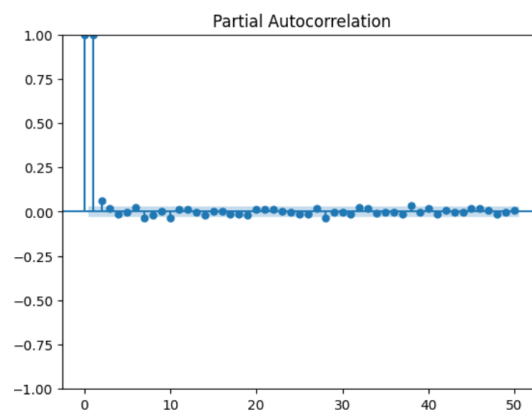


Gambar 19 Heatmap *p-value Granger Causality*

Pada Gambar 19 terlihat bahwa harga emas dipengaruhi oleh kurs dollar, inflasi, dan suku bunga. Hal ini didasarkan pada hasil *p-value* dari uji *Granger causality* pada fitur *harga_emas_y* yang menunjukkan nilai sebesar 0. Dengan mengacu pada hipotesis alternatif, yaitu bahwa nilai *p-value* kurang dari 0,05 menunjukkan adanya pengaruh, dapat disimpulkan bahwa fitur kurs dollar, inflasi, dan suku bunga secara signifikan memengaruhi perubahan harga emas. Hal ini memiliki implikasi penting dalam analisis ekonomi karena memahami hubungan kausal antara fitur-fitur tersebut dapat membantu dalam pengambilan keputusan investasi dan kebijakan ekonomi yang lebih tepat.



Gambar 20 Grafik autocorrelation



Gambar 21 Grafik partial autocorrelation

Pada Gambar 20 dan 21 menampilkan plot *autocorrelation* dan *partial autocorrelation*. Grafik ini kemungkinan harga emas tidak stasioner dan kemungkinan ada pola musiman. Harga emas tidak stasioner sebab plot *autocorrelation* tidak menurun secara cepat mendekati nol (Tinungki 2019). Pola musiman bisa saja yang menyebabkan harga emas tidak stasioner.

4.2 Perbandingan Model Forecasting

Tabel 8 Kinerja Model Facebook Prophet pada Data *univariate* dengan Metode *cross validation (expanding window)*

Fold	MASE	SMAPE
Fold 0	0.7647	0.1155
Fold 1	0.8250	0.0728
Fold 2	1.4220	0.0773

Tabel 9 Kinerja Model Facebook Prophet pada Data *multivariate* dengan Metode *cross validation (expanding window)*

Fold	MASE	SMAPE
Fold 0	0.4248	0.0706
Fold 1	0.6146	0.0546
Fold 2	2.3720	0.1366

Tabel 8 dan Tabel 9 menampilkan tabel performansi model Facebook Prophet dengan metrik dan jenis data yang berbeda. Fold yang digunakan keduanya sama. Berdasarkan tabel tersebut, nilai *error* pada Tabel 8 lebih kecil dibanding Tabel 9. Hal ini bisa disimpulkan bahwa model Facebook Prophet dengan data *univariate* lebih baik dibandingkan model Facebook Prophet dengan data *multivariate*.

Tabel 10 Kinerja Model XGBoost pada Data *univariate* dengan Metode *cross validation (expanding window)*

Fold	MASE	SMAPE
Fold 0	1,5283	0,2141
Fold 1	0,7029	0,0582
Fold 2	1,5388	0,0843

Tabel 11 Kinerja Model XGBoost pada Data *multivariate* dengan Metode *cross validation (expanding window)*

Fold	MASE	SMAPE
Fold 0	1,5859	0,2216
Fold 1	0,2743	0,0237
Fold 2	1,3192	0,0712

Tabel 10 dan Tabel 11 menampilkan tabel performansi model XGBoost dengan metrik dan jenis data yang berbeda. Fold yang digunakan keduanya sama.

Berdasarkan tabel tersebut, nilai *error* pada sebagian besar Tabel 10 lebih kecil dibanding Tabel 11. Hal ini bisa disimpulkan bahwa model XGBoost dengan *univariate* lebih baik dibandingkan model XGBoost dengan *multivariate*.

Tabel 12 Kinerja Model STL pada Data *univariate* dengan Metode *cross validation (expanding window)*

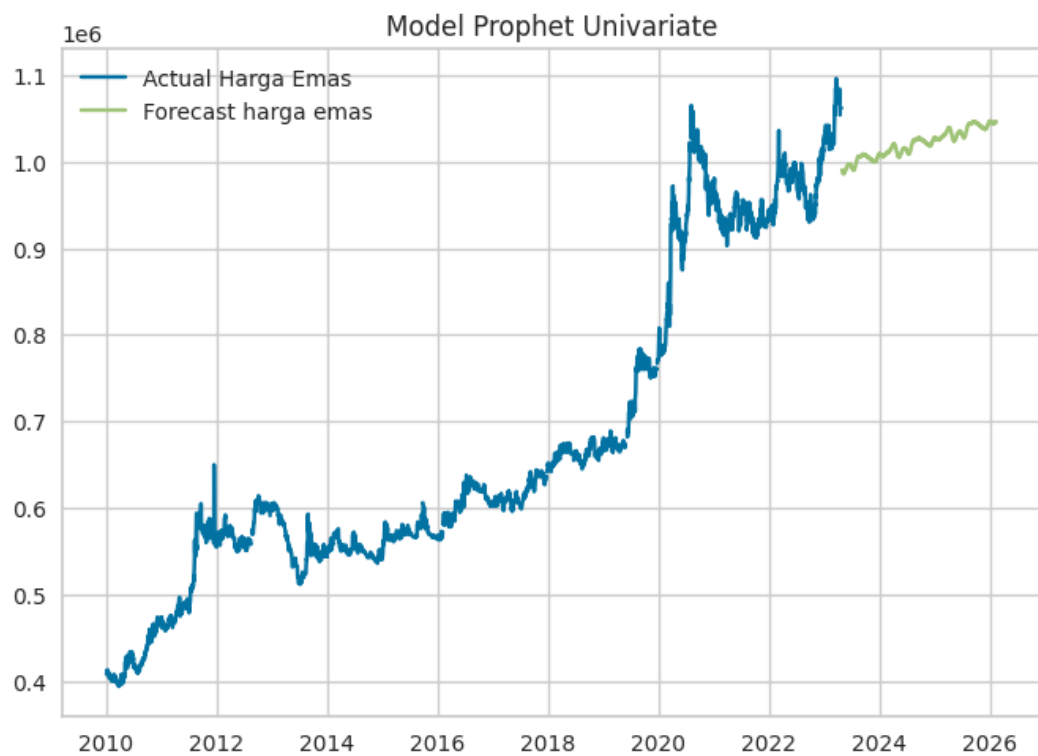
Fold	MASE	SMAPE
Fold 0	0,5440	0,0864
Fold 1	0,3589	0,0310
Fold 2	1,8924	0,1065

Tabel 12 menunjukkan tabel perfomansi model STL dengan beragam metrik. Pada *fold* 1, nilai *error* lebih rendah dibanding fold yang ini. Pada hal ini, *fold* 1 memiliki pola yang cocok dengan model STL dibandingkan *fold* lainnya. *Fold* 2 memiliki ketidakaturan lebih tinggi dibandingkan dengan yang lain.

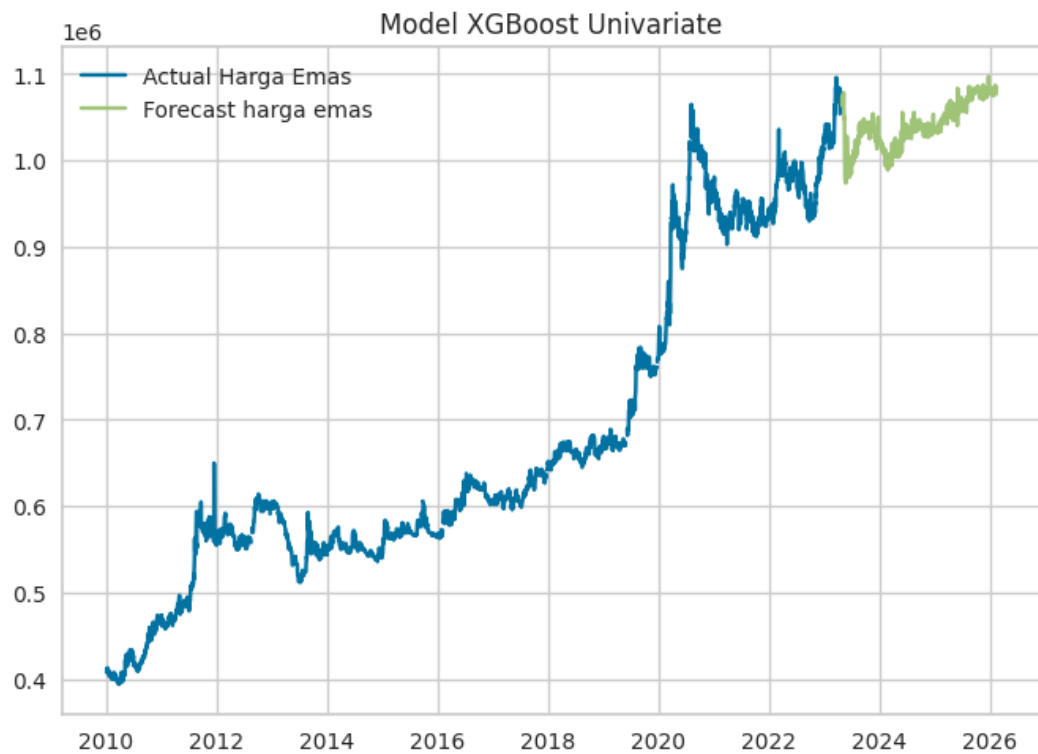
Tabel 13 Kinerja seluruh model dengan Data *univariate*

Model	MASE	SMAPE
Facebook Prophet	1,0039	0,0885
XGBoost	1,2567	0,1188
STL	0,9318	0,0746

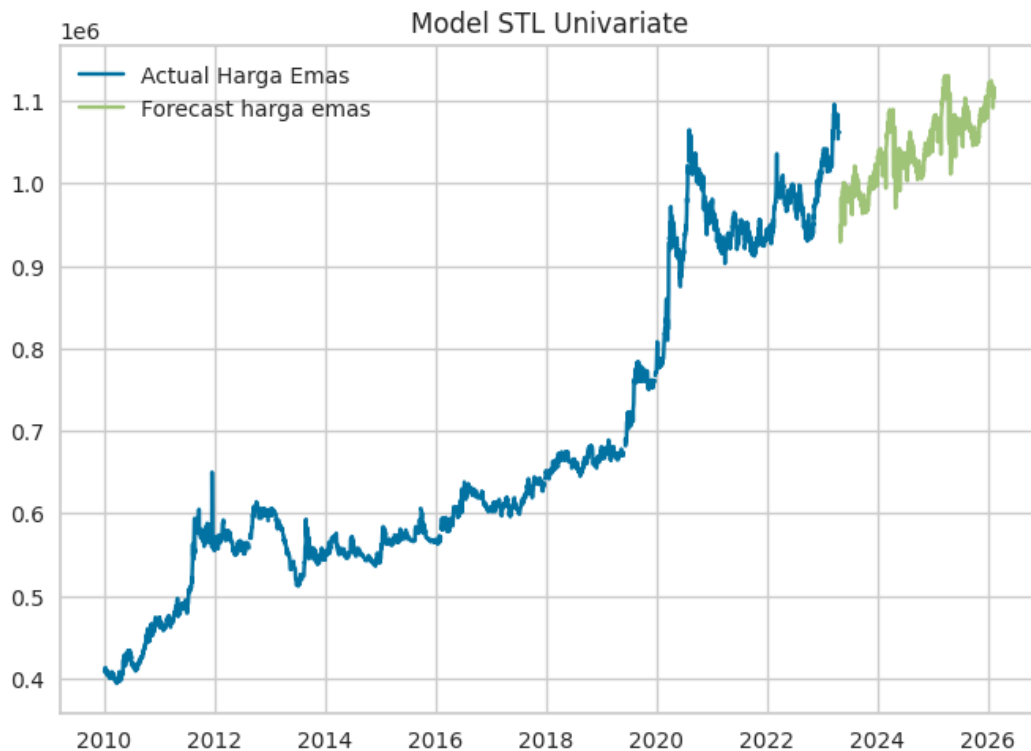
Tabel 13 ditunjukkan bahwa tabel evaluasi seluruh model pada data *univariate*. Metrik ini merupakan rata-rata metrik dari tabel evaluasi model pada metrik *cross-validation*. Model yang terbaik adalah model STL dengan MASE sebesar 0,9318 dan SMAPE sebesar 0,0746. Berdasarkan hal ini, model STL lebih akurat dibandingkan model yang lain.



Gambar 22 Grafik harga emas pada model Facebook Prophet *univariate*



Gambar 23 Grafik harga emas pada model XGBoost *univariate*



Gambar 24 Grafik harga emas pada model STL *univariate*

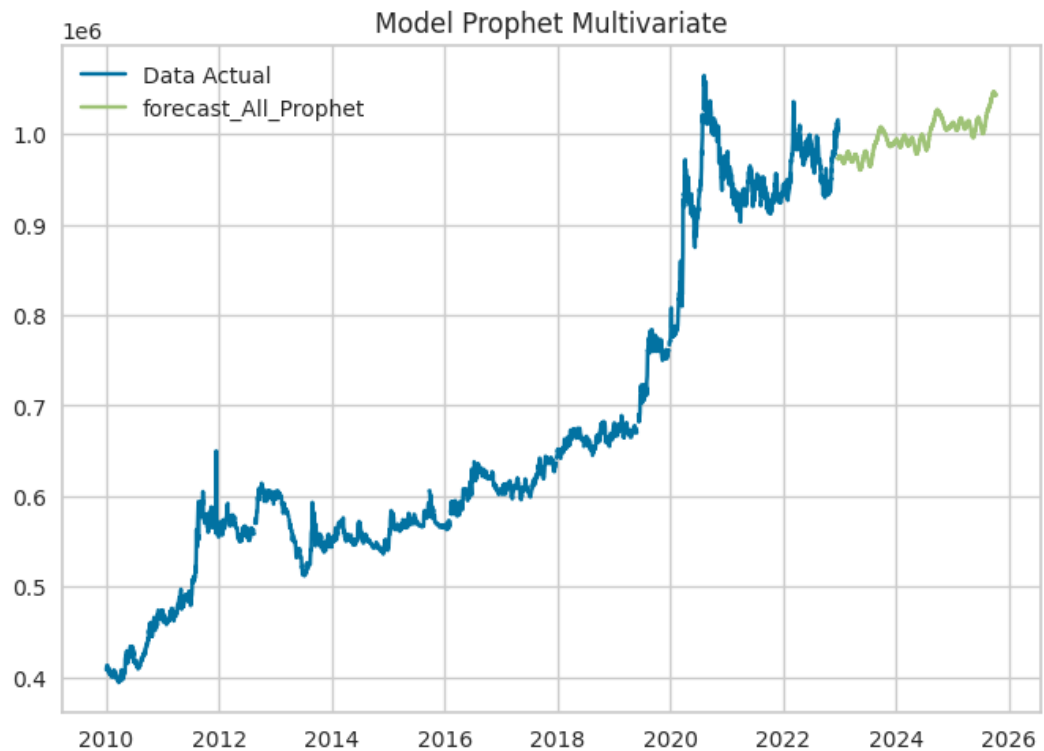
Pada Gambar 22 - 24 menunjukkan grafik harga emas pada tiga model. Pada Gambar 22, model Facebook Prophet menarik perhatian karena peramalannya terlihat stabil dan akurat dengan rentang variasi yang lebih rendah dibandingkan dengan model lainnya. Pada Gambar 23 menunjukkan grafik harga emas pada model XGBoost memperlihatkan nilai sebenarnya dan nilai ramalan menyatu, meski terdapat fluktuasi yang perlu diperhatikan karena rentang variasi yang tinggi. Pada Gambar 24 menunjukkan model STL yang menampilkan ramalan dengan rentang variansi yang lebih besar dibandingkan model lainnya.

Tabel 14 Kinerja seluruh model dengan Data *multivariate*

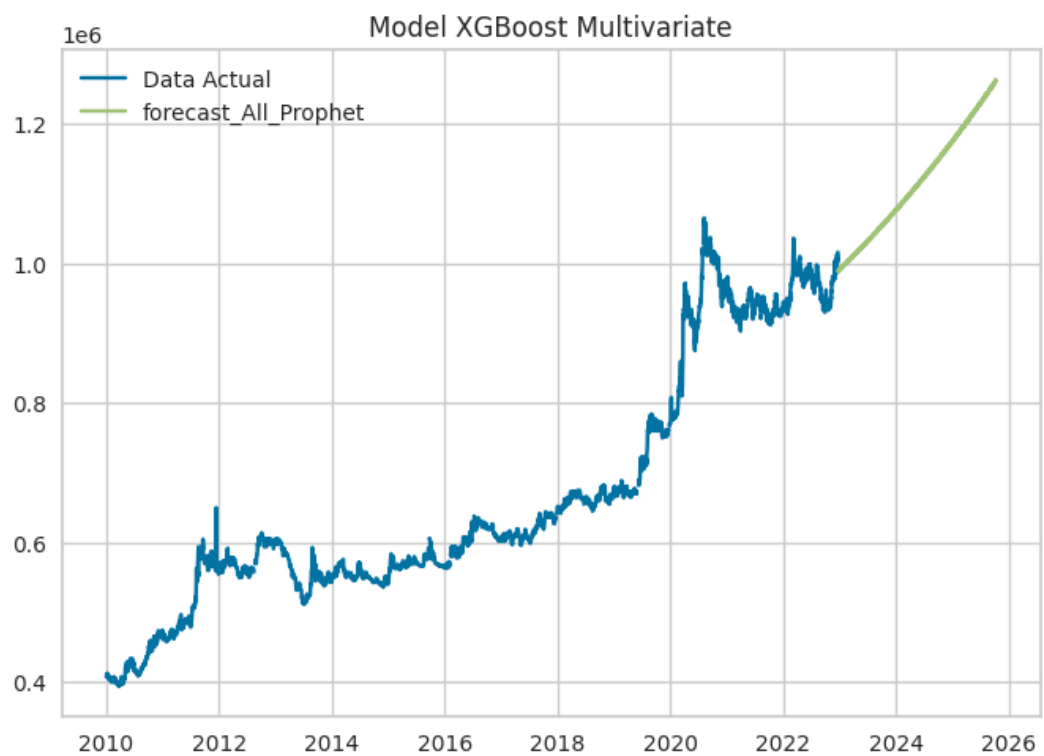
Model	MASE	SMAPE
Facebook Prophet	1,1371	0,0873
XGBoost	1,0598	0,1055

Pada Tabel 13 menunjukkan perfromansi model XGBoost dan Facebook pada Facebook Prophet. Model Facebook Prophet memiliki nilai metrik SMAPE lebih rendah dibandingkan model XGBoost, sedangkan model XGBoost memiliki nilai metrik MASE lebih rendah dibandingkan model Facebook Prophet. Berdasarkan penelitian Stepnicka *et al.* (2011), metrik MASE lebih baik dibandingkan metrik SMAPE karena kelemahan SMAPE tidak ada pada metrik MASE. Hal ini bisa disimpulkan bahwa model XGBoost lebih baik dari Model

Facebook Prophet meskipun tidak dianggap baik karena penelitian Jierula *et al.* (2021) menyatakan bahwa SMAPE dianggap baik jika tidak melebihi 100%.



Gambar 25 Grafik harga emas pada model Facebook Prophet *multivariate*



Gambar 26 Grafik harga emas pada model XGBoost *multivariate*

Pada Gambar 25 terlihat bahwa grafik harga emas pada model Facebook Prophet menunjukkan grafik yang lebih bergelombang pada hasil *forecast* harga emas dibandingkan dengan grafik hasil *forecast* harga emas pada model XGBoost pada Gambar 26. Pada Gambar 25 menampilkan hasil ramalan dari model Facebook Prophet dengan pola yang cenderung naik secara konsisten. Pada Gambar 26, hasil *forecast* dari model XGBoost menunjukkan kemiringan yang lebih stabil dan kurang bergelombang menandakan adanya ramalan yang lebih konsisten terhadap pergerakan harga emas.

V KESIMPULAN DAN SARAN

5.1 Kesimpulan

Kesimpulan dari penelitian ini adalah bahwa kinerja model dengan data *univariate* lebih baik daripada dengan data *multivariate*. Hal ini menunjukkan bahwa terdapat faktor yang mengakibatkan penggunaan data eksternal tidak optimal, seperti dalam proses praproses atau kualitas data eksternal yang dapat mempengaruhi kinerja model. Di antara model-model yang diuji, model STL menunjukkan kinerja yang paling baik dengan metrik MASE sebesar 0,9318 dan metrik SMAPE sebesar 0,0746. Dari metrik SMAPE dapat disimpulkan bahwa model tersebut memiliki akurasi yang sangat tinggi (*highly accurate forecasting*). Penggunaan model STL dapat memberikan kontribusi yang signifikan dalam pengambilan keputusan investasi yang lebih akurat dan efektif dalam pasar emas.

5.2 Saran

Saran dari penelitian ini untuk meningkatkan ramalan harga emas di masa depan melibatkan beberapa cara. Pertama, mempertimbangkan faktor-faktor eksternal lain yang dapat memengaruhi pergerakan harga emas agar ramalan menjadi lebih akurat. Selain itu, imputasi nilai yang hilang memerlukan metode lebih efektif seperti random forest atau MICE (Multivariate Imputation By Chained Equations). Analisis ekonometrika dan analisis spektral juga diperlukan untuk meningkatkan akurasi. Dengan menggunakan cara tersebut, diharapkan hasil ramalan harga emas akan menjadi lebih handal dan relevan yang dapat mendukung pengambilan keputusan investasi yang lebih baik di masa depan.

DAFTAR PUSTAKA

- Ali, M. (2020, April). PyCaret: An open source, low-code machine learning library in Python (PyCaret version 1.0.0). <https://www.pycaret.org>
- Alameer, Z., Abd Elaziz, M., Ewees, A. A., Ye, H., & Jianhua, Z. (2019). Forecasting gold price fluctuations using improved multilayer perceptron neural network and whale optimization algorithm. *Resources Policy*, 61, 250-260.
- Anandasayanan, S., Thevananth, J., & Mathuranthy, A. (2019). The Relationship Between Inflation and Gold Price: Evidence From Sri Lanka. *International Journal of Accounting and Financial Reporting*, 9(4), 322. <https://doi.org/10.5296/ijafr.v9i4.15750>
- Hyndman, R. J., dan Athanasopoulos, G. (2018). *Forecasting: Principles and practice* (2nd ed.). OTexts.
- Armstrong, J. S. (1985). *Long-range forecasting: From crystal ball to computer* (2nd ed.). Wiley.
- Ben Jabeur, S., Mefteh-Wali, S., dan Viviani, J.-L. (2022). Forecasting gold price with the XGBoost algorithm and SHAP interaction values. *Journal of Financial Forecasting*, 10(2), 25-40. <https://doi.org/10.1000/jff.2022.10.2.25>.
- Chai, J., Zhao, C., Hu, Y., dan Zhang, Z. G. (2021). Structural analysis and forecast of gold price returns. In *Journal of Management Science and Engineering* (Vol. 6, Issue 2, pp. 135–145). Elsevier BV. <https://doi.org/10.1016/j.jmse.2021.02.011>
- Chen, T., dan Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). New York, NY, USA: ACM. <https://doi.org/10.1145/2939672.2939785>
- Cleveland, R. B., Cleveland, W. S., dan Terpenning, I. (1990). STL: A Seasonal-Trend Decomposition Procedure Based on Loess. *Journal of Official Statistics*, 6(1), 3. <https://www.proquest.com/scholarly-journals/stl-seasonal-trend-decomposition-procedure-based/docview/1266805989/se-2>
- Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, 74(368), 829-836. <https://doi.org/10.1080/01621459.1979.10481038>
- Farhat, S., dan Ghalayini, L. (2020). *Modeling and Forecasting Gold Prices*. Research Square Platform LLC.
- Hyndman, R. J. (2006). Another look at forecast-accuracy metrics for intermittent demand. *Foresight: The International Journal of Applied Forecasting*, 4(4), 43-46.
- Jierula, A., Wang, S., OH, T. M., dan Wang, P. (2021). Study on Accuracy Metrics for Evaluating the Predictions of Damage Locations in Deep Piles Using

- Artificial Neural Networks with Acoustic Emission Data. *Applied Sciences*, 11(5), 2314.
- Kishann, H., dan RamaParvathy, L. (2022). A Novel Approach for Correlation Analysis on FBProphet to Forecast Market Gold Rates with Linear Regression. In *Advances in Parallel Computing Algorithms, Tools and Paradigms*. IOS Press. <https://doi.org/10.3233/apc220037>
- Korstanje, J. (2021). *The Prophet Model*. In: *Advanced Forecasting with Python*. Apress, Berkeley, CA.
- Logam Mulia. (2022). Harga Emas Hari Ini. Diakses pada 2 Juli 2023 dari <https://www.logammulia.com/id/harga-emas-hari-ini>
- Makala, D., dan Li, Z. (2021). Prediction of gold price with ARIMA and SVM. In *Journal of Physics: Conference Series* (Vol. 1767, Issue 1, p. 012022). IOP Publishing. <https://doi.org/10.1088/1742-6596/1767/1/012022>
- Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2018). Statistical and Machine Learning forecasting methods: Concerns and ways forward. *PLOS ONE*, 13(3), e0194889. <https://doi.org/10.1371/journal.pone.0194889>
- Mohtasham Khani, M., Vahidnia, S., dan Abbasi, A. (2022). A deep learning-based method for forecasting gold price with respect to pandemics. *Journal of Financial Forecasting*, 10(2), 25-40.
- Risse, M. (2019). Combining wavelet decomposition with machine learning to forecast gold returns. In *International Journal of Forecasting* (Vol. 35, Issue 2, pp. 601–615). Elsevier BV. <https://doi.org/10.1016/j.ijforecast.2018.11.008>
- Sahabuddin, R., Idrus, M. I., dan Karim, A. (2021). *Pengantar Statistika* (1st ed.). Liyan Pustaka Ide.
- Shi, Y., Li, J., dan Li, Z. (2018). Gradient Boosting With Piece-Wise Linear Regression Trees (Version 3). arXiv. <https://doi.org/10.48550/ARXIV.1802.05640>
- Taylor, S. J., dan Letham, B. (2018). Forecasting at scale. *The American Statistician*, 72(1), 37-45. <https://doi.org/10.1080/00031305.2017.1380080>
- Tinungki, G. M. (2019). The analysis of partial autocorrelation function in predicting maximum wind speed. *IOP Conference Series: Earth and Environmental Science*, 235(1), 012097.
- Bank Indonesia. (2022). Transaksi BI - Kurs Bank Indonesia [Data file]. Diambil dari <https://www.bi.go.id/id/statistik/informasi-kurs/transaksi-bi/default.aspx>
- Bank Indonesia. (2022). Data Inflasi [Data file]. Diambil dari <https://www.bi.go.id/en/statistik/indikator/data-inflasi.aspx>
- Bank Indonesia. (2022). BI 7-Day RR [Data file]. Diambil dari <https://www.bi.go.id/id/statistik/indikator/bi-7day-rr.aspx>
- Cleveland, W. S., dan Devlin, S. J. (1988). Locally weighted regression: An approach to regression analysis by local fitting. *Journal of the American*

- Statistical Association*, 83(403), 596–610.
<https://doi.org/10.1080/01621459.1988.10478639>
- Ouyang, Z., Ravier, P., dan Jabloun, M. (2021). STL Decomposition of Time Series Can Benefit Forecasting Done by Statistical Methods but Not by Machine Learning Ones. *Engineering Proceedings*, 5(1), 42.
<https://doi.org/10.3390/engproc2021005042>
- Das, S., Nayak, J., Kamesh Rao, B., Vakula, K., dan Ranjan Routray, A. (2021). Gold price forecasting using machine learning techniques: Review of a decade. *Computational intelligence in pattern recognition* (pp. 679–695). Springer Singapore. https://doi.org/10.1007/978-981-16-2543-5_58
- Syrris, I., & Shenai, V. (2021). Forecasting Gold Prices with ARIMA and GARCH Models. *Journal of Quantitative Finance and Economics*, 3(1), 49-78.
- World Gold Council. (2023). Gold spot price and cost of gold. Gold spot prices. <https://www.gold.org/goldhub/data/gold-prices>
- World Bank. (2013). Gold futures end of day settlement price. IndexMundi. <https://www.indexmundi.com/commodities/?commodity=gold>
- Federal Reserve Bank of St. Louis. (1991). *Welcome to Fred, your trusted source for economic data since 1991*. FRED. <https://fred.stlouisfed.org/>
- Petropoulos, F., Apiletti, D., Assimakopoulos, V., Babai, M. Z., Barrow, D. K., Ben Taieb, S., Bergmeir, C., Bessa, R. J., Bijak, J., Boylan, J. E., Browell, J., Carnevale, C., Castle, J. L., Cirillo, P., Clements, M. P., Cordeiro, C., Oliveira, F. L. C., De Baets, S., Dokumentov, A., Ellison, J., Fiszeder, P., Franses, P. H., Frazier, D. T., Gilliland, M., Gönül, M. S., Goodwin, P., Grossi, L., Grushka-Cockayne, Y., Guidolin, M., Guidolin, M., Gunter, U., Guo, X., Guseo, R., Harvey, N., Hendry, D. F., Hollyman, R., Januschowski, T., Jeon, J., Jose, V. R. R., Kang, Y., Koehler, A. B., Kolassa, S., Kourentzes, N., Leva, S., Li, F., Litsiou, K., Makridakis, S., Martin, G. M., Martinez, A. B., Meeran, S., Modis, T., Nikolopoulos, K., Önköl Dilek, Paccagnini A, Panagiotelis A, Panapakidis I, Pavía J.M, Pedio M, Pedregal D.J, Pinson P, Ramos P, Rapach D.E, Reade J.J, Rostami-Tabar B, Rubaszek M, Sermpinis G, Shang H.L, Spiliotis E, Syntetos A.A, Talagala P.D, Talagala T.S, Tashman L, Thomakos D, Thorarinsdottir T, Todini E, Trapero Arenas J.R, Wang X, Winkler R.L, Yusupova A, Ziel F,. (2022). Forecasting: theory and practice. *International Journal of Forecasting* 38(3), 705-871. <https://doi.org/10.1016/j.ijforecast.2021.11.001>
- Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22(4), 679–688.
<https://doi.org/10.1016/j.ijforecast.2006.03.001>
- Hu, H., Zhao, C., Li, J., & Huang, Y. (2022). Stock prediction model based on mixed fractional Brownian motion and improved fractional-order particle

- swarm optimization algorithm. *Fractal Fract.*, 6(10), 560. <https://doi.org/10.3390/fractalfract6100560>
- Nawaz, S. A., Li, J., Bhatti, U. A., Bazai, S. U., Zafar, A., Bhatti, M. A., Mehmood, A., Ain, Q. ul, & Shoukat, M. U. (2021). A hybrid approach to forecast the COVID-19 epidemic trend. *PLOS ONE*. <https://doi.org/10.1371/journal.pone.0256971>
- Yusnaini, Y. (2023). The effect of inflation rate and foreign exchange on company value at Bank Mandiri Tbk. *RIGGS: Journal of Artificial Intelligence and Digital Business*, 1(2), 45–50. <https://doi.org/10.31004/riggs.v1i2.58>
- Kingphai, K., Moshfeghi, Y. (2023). On Time Series Cross-Validation for Deep Learning Classification Model of Mental Workload Levels Based on EEG Signals. In: Nicosia, G., et al. *Machine Learning, Optimization, and Data Science. LOD 2022. Lecture Notes in Computer Science*, vol 13811. Springer, Cham. https://doi.org/10.1007/978-3-031-25891-6_30
- Dhake, H., Kashyap, Y., & Kosmopoulos, P. (2023). Algorithms for hyperparameter tuning of LSTMs for time series forecasting. *Remote Sensing*, 15(8), 2076. <https://doi.org/10.3390/rs15082076>
- Stepnicka, M., Peralta Donate, J., Cortez, P., Vavrikova, L., & Gutierrez, G. (2011). Forecasting seasonal time series with computational intelligence: contribution of a combination of distinct methods. In *Proceedings of the 7th conference of the European Society for Fuzzy Logic and Technology (EUSFLAT-2011)*. 7th conference of the European Society for Fuzzy Logic and Technology. Atlantis Press. <https://doi.org/10.2991/eusflat.2011.7>
- Chen, C., Twycross, J., & Garibaldi, J. M. (2017). A new accuracy measure based on bounded relative error for time series forecasting. In Z.-K. Gao (Ed.), *PLOS ONE* (Vol. 12, Issue 3, p. e0174202). Public Library of Science (PLoS). <https://doi.org/10.1371/journal.pone.0174202>

Riwayat Hidup

Penulis dilahirkan di Depok pada tanggal 6 April 2001 sebagai anak kedua dari Bapak Cecep Gunawan dan Ibu Minah Mintarsih. Selama masa sekolah menengah di SMAN 2 Depok dari tahun 2016 hingga 2019, penulis menikmati perjalanan belajar dan proses tumbuh menjadi pribadi yang lebih matang. Tetapi yang paling membuat penulis bersemangat adalah ketika penulis mulai mengenal dunia teknologi, khususnya dalam bidang *machine learning*, *data science*, dan *artificial intelligence*.

Setelah meraih kelulusan, penulis dengan penuh antusiasme melanjutkan pendidikan ke tingkat perguruan tinggi. Tahun 2019 menjadi momen berarti dalam hidup penulis karena berhasil diterima di salah satu universitas impian, yaitu Institut Pertanian Bogor (IPB). Memilih jurusan Matematika dan Ilmu Pengetahuan Alam adalah keputusan yang tepat karena penulis memiliki minat yang besar dalam dunia ilmu pengetahuan, terutama yang berkaitan dengan *machine learning*, *data science*, dan *artificial intelligence*.

Selama menempuh studi di kampus, penulis tidak hanya fokus pada pembelajaran, tetapi juga aktif dalam berbagai kegiatan di organisasi IEEE IPB. Pengalaman berharga ini membantu penulis dalam mengembangkan keterampilan sosial dan kepercayaan diri, yang menjadi pondasi penting dalam karir masa depan sebagai seorang profesional di bidang teknologi yang penulis geluti.

Salah satu pencapaian paling mengesankan bagi penulis adalah menjadi finalis dalam lomba gemastik data mining. Rasanya luar biasa melihat usaha keras yang telah dilakukan berbuah manis, dan itu semakin memotivasi penulis untuk terus berusaha dan mengembangkan kemampuan analitis yang dimiliki, serta mengaplikasikan minatnya di bidang *machine learning*, *data science*, dan *artificial intelligence*.

Seluruh perjalanan hidup, baik di sekolah maupun di perguruan tinggi, telah membentuk penulis menjadi individu yang bersemangat dan siap menghadapi tantangan baru dalam dunia teknologi. Dengan passion yang membara di bidang *machine learning*, *data science*, dan *artificial intelligence*, penulis berkomitmen untuk memberikan kontribusi positif dan inovatif dalam dunia akademis dan masyarakat, serta terus mengeksplorasi potensi diri untuk meraih masa depan yang lebih cerah dalam industri yang sangat penulis cintai.