

Prediksi Inflasi Bulanan Menggunakan LightGBM dan Optimasi *Hyperparameter* Berbasis Optuna

Monthly Inflation Prediction Using LightGBM and Optuna-Based Hyperparameter Optimization

Dian Maharani¹, Anggraini Puspita Sari ^{*2}, Muhammad Muharrom Al Haromainy³

Jl.Raya Rungkut Madya, Gunung Anyar, Surabaya, Indonesia

Email : dianmaharani1202@gmail.com¹, anggraini.puspita.if@upnjatim.ac.id^{*2},
muhammad.muharrom.if@upnjatim.ac.id³

**Corresponding author*

Received 7 June 2026; Revised 13 July 2026; Accepted 18 July 2026

Abstrak - Inflasi merupakan indikator makroekonomi penting yang mencerminkan perubahan tingkat harga barang dan jasa serta berpengaruh terhadap stabilitas ekonomi dan pengambilan kebijakan. Prediksi inflasi menjadi tantangan karena data inflasi memiliki karakteristik deret waktu yang dinamis, dipengaruhi oleh berbagai faktor ekonomi, serta mengandung hubungan nonlinier. Tujuan penelitian ini adalah menghasilkan model prediksi inflasi bulanan Indonesia dengan menerapkan metode LightGBM serta melakukan optimasi *hyperparameter* menggunakan Optuna guna meningkatkan kinerja model. Dataset penelitian mencakup data inflasi dan berbagai indikator makroekonomi Indonesia pada tahun 2014–2024 yang dikumpulkan dari Bank Indonesia dan Badan Pusat Statistik (BPS). Tahapan penelitian meliputi *preprocessing*, *feature engineering*, pembagian data, pemodelan, optimasi *hyperparameter*, dan evaluasi menggunakan MAE, RMSE, dan R^2 . Berdasarkan pengujian yang dilakukan, diperoleh hasil bahwa optimasi *hyperparameter* mampu meningkatkan performa model secara signifikan dengan menghasilkan MAE sebesar 0,0434, RMSE sebesar 0,0561, dan R^2 sebesar 0,9259. Hasil tersebut menunjukkan bahwa LightGBM yang dioptimasi menggunakan Optuna mampu memprediksi inflasi bulanan dengan tingkat akurasi yang tinggi dan efektif dalam menangkap pola nonlinier pada data inflasi.

Kata kunci - Inflasi, LightGBM, Optuna, Deret waktu, Prediksi

Abstract - Inflation is an important macroeconomic indicator that reflects changes in the general price level of goods and services and affects economic stability and policy-making. Inflation forecasting remains challenging because inflation data exhibit dynamic time-series characteristics, are influenced by various economic factors, and contain nonlinear relationships. This study aimed to develop a monthly inflation forecasting model for Indonesia using Light Gradient Boosting Machine (LightGBM) with Optuna-based hyperparameter optimization. The dataset consisted of inflation and macroeconomic variables from 2014 to 2024 obtained from Bank Indonesia and Statistics Indonesia. The research stages included preprocessing, feature engineering, data splitting, model development, hyperparameter optimization, and evaluation using MAE, RMSE, and the coefficient of determination (R^2). The results showed that hyperparameter optimization significantly improved model performance, achieving an MAE of 0,0434, RMSE of 0.0561, and R^2 of 0.9259. These findings indicate that Optuna-optimized LightGBM can forecast monthly inflation with high accuracy and effectively capture nonlinear patterns in inflation data.

Keyword - Inflation, LightGBM, Optuna, Time Series, Prediction

1. PENDAHULUAN

Inflasi merupakan salah satu indikator makroekonomi yang digunakan untuk menggambarkan perubahan tingkat harga barang dan jasa secara umum dalam suatu periode tertentu. Pergerakan inflasi menjadi perhatian penting karena memengaruhi daya beli masyarakat,

tingkat konsumsi, stabilitas ekonomi, serta kebijakan moneter yang ditetapkan oleh pemerintah dan bank sentral [1]. Ketidakstabilan inflasi dapat memberikan dampak terhadap aktivitas ekonomi, khususnya ketika terjadi kenaikan harga yang tidak terkendali [2]. Oleh karena itu, prediksi inflasi diperlukan untuk membantu proses pengambilan keputusan dan perumusan kebijakan ekonomi secara lebih tepat.

Prediksi inflasi menjadi penting karena kondisi ekonomi yang terus berubah menyebabkan pergerakan harga bersifat dinamis dan sulit diprediksi secara konsisten. Selain dipengaruhi oleh faktor internal ekonomi, inflasi juga dipengaruhi oleh *shock* eksternal, perubahan kebijakan moneter, harga komoditas, dan ekspektasi pasar yang dapat mengubah pola inflasi secara tiba-tiba [3]. Informasi prediksi inflasi dapat digunakan sebagai dasar dalam penyusunan kebijakan fiskal dan moneter, pengendalian harga, serta perencanaan ekonomi jangka pendek maupun jangka panjang [1]. Prediksi inflasi merupakan permasalahan yang kompleks karena data inflasi memiliki karakteristik deret waktu yang mengandung pola tren, musiman, dan nonlinear [4], [5]. Selain itu, berbagai variabel ekonomi seperti harga pangan, harga energi, nilai tukar, dan suku bunga saling berinteraksi sehingga meningkatkan kompleksitas proses prediksi. Data inflasi juga dapat mengalami fluktuasi yang tinggi pada periode tertentu akibat perubahan harga komoditas maupun kondisi ekonomi global sehingga menyebabkan pola inflasi sulit diprediksi secara konsisten [6].

Berbagai penelitian sebelumnya telah menggunakan pendekatan statistik maupun machine learning dalam prediksi inflasi. Machine learning dinilai mampu meningkatkan akurasi *forecasting* karena dapat mempelajari hubungan nonlinear yang kompleks pada data ekonomi [7]. Metode statistik seperti ARIMA dan ARIMAX banyak digunakan dalam prediksi inflasi Indonesia karena mampu memodelkan pola deret waktu berdasarkan hubungan historis data [8], [9]. Namun, pendekatan tersebut memiliki keterbatasan dalam menangani hubungan nonlinear dan perubahan pola data yang dinamis antar variabel ekonomi [7]. Penelitian yang dilakukan oleh Pebrianti et al. menunjukkan bahwa algoritma machine learning, seperti Random Forest dan XGBoost, memiliki kemampuan prediksi yang lebih unggul dibandingkan metode statistik konvensional karena mampu menangkap pola hubungan nonlinear di antara variabel-variabel ekonomi [10].

Seiring dengan perkembangan *machine learning*, beberapa penelitian mulai menggunakan algoritma berbasis *boosting* seperti XGBoost dan LightGBM untuk meningkatkan akurasi prediksi. Ivaşcu membandingkan beberapa algoritma *machine learning* seperti Random Forest, XGBoost, LSTM, dan LightGBM dalam prediksi inflasi di Romania [11]. Penelitian tersebut menemukan bahwa model machine learning efektif dalam merepresentasikan pola nonlinear pada inflasi, meskipun performanya dipengaruhi oleh karakteristik data dan jumlah observasi yang tersedia. Meskipun demikian, berbagai penelitian juga menunjukkan bahwa prediksi inflasi tetap menjadi permasalahan yang kompleks karena pola inflasi dapat berubah akibat *shock* ekonomi maupun perubahan kondisi pasar. Kondisi tersebut menyebabkan performa model tidak selalu stabil pada seluruh periode pengamatan, terutama ketika terjadi fluktuasi ekstrem dan perubahan pola data secara tiba-tiba [4].

Meskipun berbagai penelitian telah menerapkan metode statistik maupun *machine learning* untuk prediksi inflasi, sebagian besar penelitian masih berfokus pada perbandingan algoritma atau penggunaan model dengan konfigurasi parameter tertentu. Selain itu, penelitian yang secara khusus mengkaji penerapan LightGBM dengan optimasi *hyperparameter* berbasis Optuna untuk prediksi inflasi Indonesia masih relatif terbatas. Kondisi ini menunjukkan adanya peluang untuk mengembangkan model prediksi yang lebih optimal melalui kombinasi LightGBM dan Optuna.

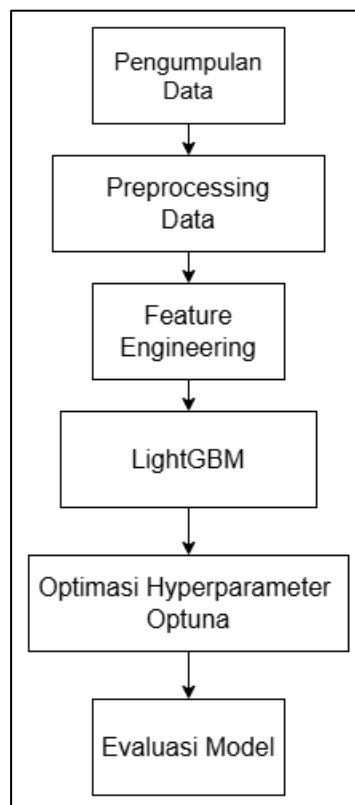
Selain pemilihan algoritma, penentuan *hyperparameter* yang tepat diperlukan untuk memperoleh hasil prediksi yang optimal, namun proses tersebut sering kali membutuhkan berbagai percobaan untuk menemukan konfigurasi terbaik [12]. Akiba et al. memperkenalkan Optuna sebagai *framework* optimasi *hyperparameter* otomatis yang mampu meningkatkan efisiensi proses pencarian parameter dibandingkan metode *tuning* tradisional [13]. Penelitian lain terkait kombinasi Optuna dan LightGBM menunjukkan bahwa optimasi *hyperparameter* dapat

meningkatkan performa prediksi serta membantu model menemukan konfigurasi yang lebih optimal pada data yang kompleks dan tidak seimbang [14]. Hasil tersebut konsisten dengan penelitian yang dilakukan oleh Sari et al., yang menunjukkan bahwa model yang dikembangkan memiliki tingkat akurasi prediksi yang lebih tinggi dibandingkan beberapa model pembandingan. Hal ini mengindikasikan bahwa pemilihan serta pengembangan model yang tepat merupakan faktor penting dalam meningkatkan kualitas hasil prediksi [15].

Berdasarkan permasalahan tersebut, penelitian ini menggunakan LightGBM dengan optimasi *hyperparameter* berbasis Optuna untuk prediksi inflasi bulanan Indonesia. Berbeda dengan sejumlah penelitian terdahulu yang menggunakan parameter bawaan atau melakukan penyesuaian *hyperparameter* secara manual, penelitian ini menerapkan Optuna sebagai metode optimasi *hyperparameter* otomatis sehingga diperoleh konfigurasi model yang lebih optimal. Pendekatan ini diharapkan mampu meningkatkan kemampuan LightGBM dalam mempelajari pola inflasi yang bersifat kompleks dan dinamis. Oleh sebab itu, penelitian ini bertujuan untuk mengembangkan model prediksi inflasi Indonesia menggunakan LightGBM yang dioptimasi dengan Optuna serta mengevaluasi performanya dalam menghasilkan prediksi yang akurat.

2. METODE PENELITIAN

Proses penelitian ini terdiri atas beberapa tahapan yang saling berkaitan dalam proses pembangunan model prediksi inflasi menggunakan algoritma LightGBM dengan optimasi *hyperparameter* berbasis Optuna. Tahapan penelitian dimulai dari pengumpulan data, *preprocessing*, pembentukan fitur, pembagian data, proses pemodelan, optimasi *hyperparameter*, hingga evaluasi model seperti yang ditunjukkan pada Gambar 1.



Gambar 1. Tahapan Penelitian

2.1. Pengumpulan Data

Data yang digunakan dalam penelitian ini merupakan data inflasi bulanan (MoM) beserta dengan beberapa variabel makroekonomi yang diduga memiliki pengaruh terhadap pergerakan inflasi. Data diperoleh dari sumber resmi dan lembaga terkait yang menyediakan publikasi data ekonomi secara berkala, yaitu Bank Indonesia dan Badan Pusat Statistik (BPS).

2.2. Preprocessing Data

a. Data Cleaning

Tahap awal *preprocessing* dilakukan dengan memeriksa konsistensi data, kesesuaian format, serta identifikasi data yang tidak valid maupun duplikasi data. Proses ini dilakukan agar seluruh variabel dapat digunakan dengan baik pada tahap analisis berikutnya.

b. Identifikasi Missing Value

Selanjutnya dilakukan pemeriksaan terhadap nilai hilang (*missing value*) pada dataset. Jika ditemukan data yang kosong, maka dilakukan penanganan sesuai karakteristik data agar informasi historis tetap dapat dipertahankan.

c. Transformasi Data

Transformasi data dilakukan untuk menyesuaikan struktur data sebelum memasuki tahap *feature engineering* dan pemodelan, yaitu transformasi data waktu dan transformasi persentase perubahan. Transformasi waktu digunakan untuk mengambil informasi temporal, seperti tahun dan bulan yang penting dalam proses prediksi deret waktu karena data time series mengandung ketergantungan temporal antar periode [16]. Selain itu, transformasi persentase perubahan dilakukan pada beberapa variabel ekonomi seperti nilai tukar, jumlah uang beredar, dan indeks keyakinan untuk mengubah data level menjadi data pertumbuhan bulanan. Transformasi tersebut umum digunakan pada data ekonomi untuk menangkap perubahan dinamis antar periode serta mengurangi pengaruh skala data yang berbeda antar variabel [17]. Tahap ini bertujuan agar data dapat diproses secara optimal pada model *machine learning* berbasis *time series*.

2.3. Feature Engineering

Tahap *feature engineering* dilakukan untuk menghasilkan fitur-fitur baru yang dapat meningkatkan kemampuan model dalam mengenali pola pergerakan inflasi. Pada penelitian ini, *feature engineering* meliputi pembentukan fitur temporal, *cyclical feature*, *trend feature*, *interaction feature*, *differencing feature*, *shock feature*, *lag feature*, dan *rolling statistics*. Selanjutnya, tahap *feature engineering* menghasilkan beberapa kategori fitur seperti *temporal*, *cyclical*, *lag*, *rolling*, *growth*, *dummy*, *interaction*, dan *acceleration feature* sebagaimana ditunjukkan pada Tabel 1.

Tabel 1. *Feature Engineering*

Kategori	Fitur yang Dihasilkan
<i>Temporal</i>	tahun, bulan, trend
<i>Cyclical</i>	sin_month, cos_month
<i>Lag Feature (lag1-3)</i>	inflasi, suku bunga, volatile food, administered price
<i>Rolling Feature</i>	rolling mean inflasi, volatile food, IKK
<i>Dummy Feature</i>	Shock bb, shock tarif listirk, panademic, idulfitri, awal/akhir tahun
<i>Interaction Feature</i>	food x adm, awal tahun x adm_diff
<i>Acceleration Feature</i>	Adm acceleration
<i>Growth Feature</i>	kurs_pct_change, M2_pct_change, ikk_pct_change
<i>Difference Feature</i>	food_diff1, adm_diff1, inflasi_diff1

Fitur-fitur tersebut digunakan untuk membantu model mempelajari pola historis, pola musiman, tren, serta perubahan dinamis pada data inflasi. Pembentukan fitur *lag*, *rolling feature*, dan *difference feature* menghasilkan beberapa nilai kosong (NaN) pada observasi awal karena memerlukan informasi dari periode sebelumnya. Nilai NaN tersebut kemudian dihapus (*drop*) sebelum proses pembagian data.

2.4. Pembagian Data

Setelah tahap feature engineering selesai, data kemudian dipisahkan menjadi data pelatihan (*training*) dan data pengujian (*testing*) berdasarkan urutan waktu pengamatan. Pembagian data dilakukan dengan perbandingan data *training* sebesar 80% dan data testing sebesar 20%. Data *training* digunakan dalam proses pelatihan model, sedangkan data testing digunakan untuk mengevaluasi performa akhir model terhadap data yang belum pernah digunakan sebelumnya.

2.5. Pemodelan LightGBM

Model prediksi pada penelitian ini dibangun menggunakan algoritma *Light Gradient Boosting Machine* (LightGBM). LightGBM merupakan algoritma machine learning berbasis *gradient boosting decision tree* yang dirancang untuk meningkatkan efisiensi proses pelatihan model serta mampu menangani hubungan nonlinear antar variabel. Secara umum, hasil prediksi model LightGBM dapat dinyatakan pada persamaan (1) [18].

$$\hat{y}_i = \sum_{k=1}^K f_k(x_i) \quad (1)$$

dengan $\hat{y}_i$ merupakan nilai prediksi pada data ke- i , sedangkan $f_k(x_i)$ merupakan fungsi *decision tree* ke- k . Fungsi objektif pada LightGBM digunakan untuk meminimalkan *loss function* dan regularisasi model yang dirumuskan pada persamaan (2) [18].

$$Obj = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k) \quad (2)$$

Pada persamaan (1) dan (2), $l(y_i, \hat{y}_i)$ menyatakan fungsi *loss* yang digunakan untuk mengukur kesalahan prediksi model, sedangkan $\Omega(f_k)$ merupakan fungsi regularisasi yang berperan dalam mengendalikan kompleksitas model. LightGBM menggunakan pendekatan *leaf-wise tree growth* sehingga proses pelatihan model menjadi lebih cepat dan efisien dibandingkan metode *boosting* konvensional.

2.6. Optimasi Hyperparameter Optuna

Optimasi *hyperparameter* dilakukan menggunakan *framework* Optuna untuk memperoleh kombinasi parameter model yang optimal. Pemilihan *hyperparameter* yang optimal memiliki peran penting dalam menentukan kemampuan model untuk mempelajari pola yang terdapat pada data. Optuna bekerja dengan melakukan pencarian parameter secara otomatis berdasarkan performa model pada setiap proses trial. Pada penelitian ini, beberapa *hyperparameter* yang dioptimasi ditunjukkan pada Tabel 2.

Tabel 2. *Hyperparameter* yang Diuji

<i>Hyperparameter</i>	Rentang Nilai
Learning rate	0,01 – 0.08
Num leaves	15 – 31
Max depth	4 – 7
Min data in leaf	5 - 15

Rentang nilai setiap *hyperparameter* ditetapkan dengan mempertimbangkan jumlah data penelitian yang relatif terbatas. Oleh karena itu, rentang parameter dipilih agar kompleksitas model tetap sesuai dengan karakteristik data dan mengurangi potensi *overfitting*. Parameter *learning_rate* digunakan untuk mengontrol kecepatan pembelajaran model agar proses *boosting* berlangsung lebih stabil. Sementara itu, *num_leaves* dan *max_depth* digunakan untuk mengatur kompleksitas pohon keputusan sehingga model tetap mampu mempelajari hubungan nonlinear tanpa menyebabkan *overfitting* pada data yang relatif kecil. Selain itu, *min_data_in_leaf* digunakan untuk membatasi jumlah minimum data pada setiap leaf agar model tidak terlalu sensitif terhadap *noise* dan fluktuasi ekstrem pada data inflasi.

2.7. Evaluasi Model

Tahap evaluasi dilakukan untuk mengukur kemampuan model dalam memprediksi pergerakan inflasi berdasarkan data aktual. Evaluasi model pada penelitian ini menggunakan beberapa metrik evaluasi, yaitu RMSE, MAE, dan koefisien determinasi (R^2).

a. *Root Mean Squared Error (RMSE)*

RMSE digunakan untuk mengukur rata-rata kesalahan prediksi dengan memberikan penalti yang lebih besar terhadap *error* bernilai tinggi [19]. RMSE dirumuskan pada persamaan (3).

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2} \quad (3)$$

b. *Mean Absolute Error (MAE)*

MAE digunakan untuk menghitung rata-rata selisih absolut antara nilai aktual dan nilai prediksi sehingga lebih mudah diinterpretasikan [19]. MAE dirumuskan pada persamaan (4).

$$MAE = \frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t| \quad (4)$$

c. *Koefisien Determinasi (R^2)*

Untuk mengevaluasi kemampuan model dalam merepresentasikan data aktual, digunakan koefisien determinasi (R^2). Nilainya berkisar antara 0 hingga 1, dengan nilai yang lebih tinggi menandakan bahwa model mampu menjelaskan variasi data dengan lebih baik. [20]. R^2 dirumuskan pada persamaan (5).

$$R^2 = 1 - \frac{\sum (y_t - \hat{y}_t)^2}{\sum (y_t - \bar{y})^2} \quad (5)$$

Ketiga metrik tersebut digunakan untuk memberikan gambaran performa model dalam memprediksi inflasi secara menyeluruh.

3. HASIL DAN PEMBAHASAN

Pada bagian ini dipaparkan hasil penelitian dan analisis yang diperoleh dari setiap tahapan pemodelan yang telah dilakukan. Pembahasan diawali dengan deskripsi data penelitian, dilanjutkan dengan proses *feature engineering* untuk membentuk variabel prediktor yang relevan. Selanjutnya, dilakukan pengembangan model LightGBM dengan optimasi *hyperparameter* menggunakan Optuna untuk memperoleh konfigurasi model yang optimal. Hasil pemodelan kemudian dievaluasi dan dianalisis guna menilai kemampuan model dalam memprediksi inflasi Indonesia serta mengidentifikasi konfigurasi model dengan kinerja terbaik.

3.1 . Deskripsi Data

Penelitian ini menggunakan data deret waktu (*time series*) bulanan inflasi Indonesia. Variabel dependen yang digunakan adalah inflasi *month-over-month* (MoM), sedangkan variabel independennya meliputi suku bunga, nilai tukar rupiah, jumlah uang beredar (M1 dan M2), ekspektasi inflasi, indeks keyakinan konsumen, *administered prices*, dan *volatile food*. Dataset terdiri dari data dengan frekuensi bulanan pada periode tahun 2014 hingga 2024.

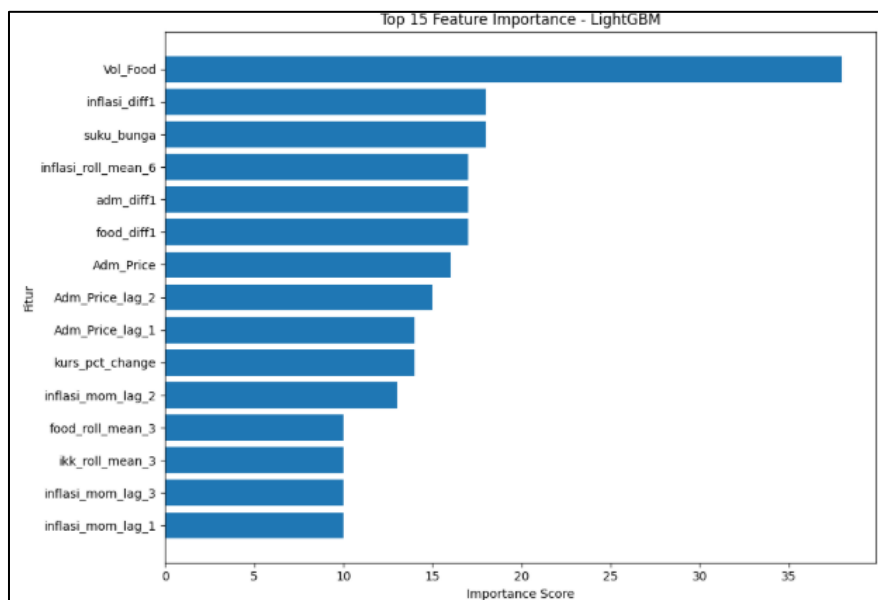
	inflasi_mom	M1 (miliar)	M2 (miliar)	suku_bunga	nilai_kurs	EH_3B1n	EH_6B1n	Adm_Price	Vol_Food	IKK
count	132.000000	1.320000e+02	1.320000e+02	132.000000	1.320000e+02	132.000000	132.000000	132.000000	132.000000	132.000000
mean	0.279015	1.662574e+06	6.232112e+06	5.458333	1.142698e+06	142.714394	144.409091	0.347121	0.360303	115.607576
std	0.371849	5.911751e+05	1.656109e+06	1.355708	5.397041e+05	13.515635	15.792033	1.044411	1.248935	12.544374
min	-0.450000	8.345320e+05	3.635060e+06	3.500000	1.316500e+04	112.700000	111.000000	-3.510000	-2.900000	77.300000
25%	0.070000	1.173173e+06	4.770366e+06	4.250000	1.189216e+06	133.250000	132.200000	-0.020000	-0.535000	112.750000
50%	0.210000	1.486151e+06	5.921536e+06	5.625000	1.377895e+06	141.100000	138.600000	0.160000	0.290000	120.150000
75%	0.440000	2.279922e+06	7.847711e+06	6.250000	1.460895e+06	152.650000	157.950000	0.480000	1.217500	123.725000
max	2.460000	2.839540e+06	9.246630e+06	7.750000	1.632939e+06	179.000000	177.800000	6.180000	3.530000	128.900000

Gambar 2. Deskripsi Data

Berdasarkan Gambar 2, dataset penelitian terdiri atas 132 observasi bulanan dengan variabel target berupa inflasi bulanan (*inflasi_mom*) dan sembilan variabel prediktor. Inflasi memiliki rata-rata 0,279% dengan rentang nilai antara -0,450% hingga 2,460%. Seluruh variabel prediktor menunjukkan variasi nilai yang cukup besar selama periode pengamatan, yang mencerminkan dinamika kondisi ekonomi Indonesia dan berpotensi memberikan informasi penting dalam proses pemodelan prediksi inflasi.

3.2 . Hasil Feature engineering

Tahap *feature engineering* dilakukan untuk menghasilkan fitur-fitur baru yang dapat menangkap pola historis inflasi dan hubungan antar variabel ekonomi. Fitur yang dibentuk meliputi fitur *lag*, *differencing*, *rolling statistics*, serta transformasi persentase perubahan pada beberapa variabel ekonomi. Fitur-fitur tersebut digunakan sebagai masukan model LightGBM untuk meningkatkan kemampuan dalam mempelajari pola inflasi bulanan.



Gambar 3. Feature Importance LightGBM

Gambar 3 menunjukkan 15 fitur dengan nilai *feature importance* tertinggi berdasarkan model LightGBM. Berdasarkan hasil tersebut, fitur Vol_Food memiliki kontribusi terbesar dalam prediksi inflasi diikuti oleh beberapa fitur historis inflasi seperti *inflasi_diff1*, *inflasi_roll_mean_6*, dan beberapa fitur *lag* inflasi. Hasil ini menunjukkan bahwa model LightGBM lebih banyak memanfaatkan informasi yang berasal dari perubahan harga kelompok *volatile food* serta pola historis inflasi dalam membentuk prediksi. Temuan ini sejalan dengan karakteristik inflasi di Indonesia. Menurut Farandy [21], inflasi Indonesia dalam satu dekade terakhir lebih banyak dipengaruhi oleh komponen *volatile food* dan *administered prices*, sedangkan inflasi inti (*core inflation*) relatif stabil. Hal ini menunjukkan bahwa perubahan harga pada kelompok *volatile food* berperan penting dalam membentuk fluktuasi inflasi. Temuan tersebut didukung oleh Theresia *et al.* [22] yang menjelaskan bahwa harga komoditas pangan strategis di Indonesia, seperti beras, cabai, bawang merah, telur, dan daging ayam, memiliki tingkat volatilitas yang tinggi dan cenderung persisten akibat faktor musim, pasokan, dan distribusi. Oleh karena itu, informasi dari kelompok *volatile food* menjadi relevan bagi model dalam memprediksi pergerakan inflasi.

Selain Vol_Food, beberapa fitur historis inflasi seperti *inflasi_diff1*, *inflasi_roll_mean_6*, dan *inflasi_mom_lag_1* juga termasuk dalam fitur dengan nilai *feature importance* yang relatif

besar. Hal ini menunjukkan bahwa informasi dari periode sebelumnya masih dimanfaatkan oleh model dalam memprediksi inflasi pada periode berikutnya. Selain itu, variabel makroekonomi seperti suku_bunga, Adm_Price, dan kurs_pct_change juga memberikan kontribusi terhadap proses prediksi, meskipun nilai *feature importance*-nya lebih rendah dibandingkan fitur-fitur utama. Secara keseluruhan, hasil ini menunjukkan bahwa model LightGBM memanfaatkan kombinasi informasi historis inflasi, indikator harga pangan, dan variabel makroekonomi dalam mempelajari pola inflasi bulanan.

3.3 . Hasil Optimasi Hyperparameter Optuna

Hasil optimasi *hyperparameter* menggunakan Optuna menghasilkan konfigurasi parameter yang lebih optimal dibandingkan parameter *default*. Konfigurasi tersebut membantu model menjaga keseimbangan antara kemampuan mempelajari hubungan nonlinear dan pengendalian kompleksitas model sehingga proses pelatihan menjadi lebih stabil pada data inflasi dengan jumlah observasi yang terbatas. Hasil parameter terbaik ditunjukkan pada Tabel 3.

Tabel 3. Hasil Parameter Terbaik

<i>Hyperparameter</i>	Nilai Terbaik
Learning rate	0,069
Num leaves	23
Max depth	5
Min data in leaf	7

Nilai parameter yang dihasilkan menunjukkan bahwa model menggunakan kompleksitas pohon yang relatif moderat sehingga mampu mempelajari pola data inflasi tanpa meningkatkan risiko *overfitting* secara berlebihan.

3.4 . Hasil dan Analisis Evaluasi Model

Hasil pengujian, model LightGBM yang telah dioptimasi menggunakan Optuna menunjukkan peningkatan performa yang signifikan dibandingkan model LightGBM dengan parameter. Perbandingan hasil evaluasi ditunjukkan pada Tabel 4.

Tabel 4. Perbandingan Evaluasi Model LightGBM dan LightGBM+Optuna

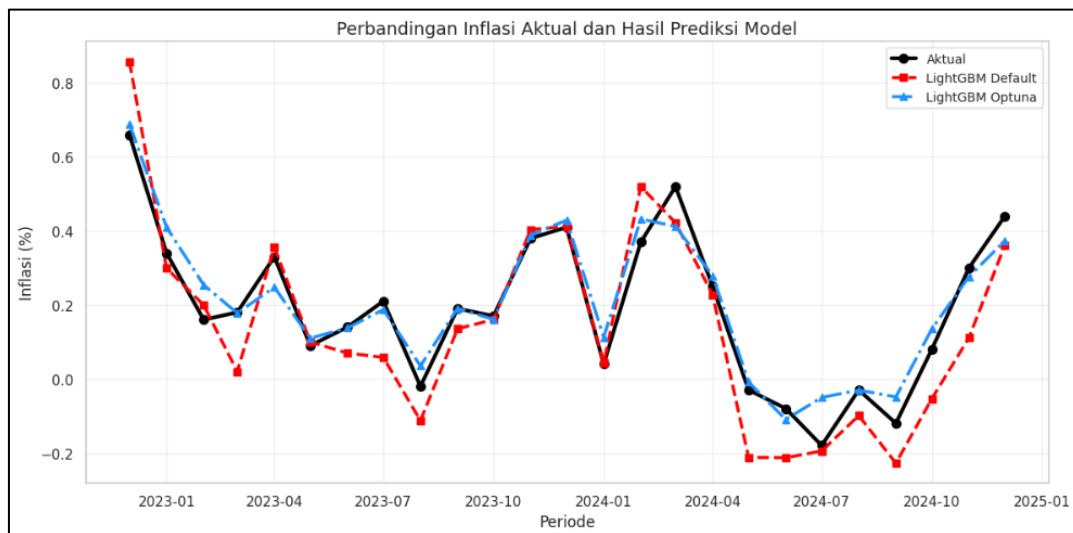
Model	MAE	RMSE	R ²
LightGBM	0,0825	0,1037	0,7470
LightGBM+Optuna	0,0434	0,0561	0,9259

Hasil tersebut mengindikasikan bahwa optimasi hyperparameter mampu meningkatkan kinerja model dalam memprediksi inflasi bulanan (*month-over-month inflation*). Model yang telah dioptimalkan menghasilkan nilai MAE sebesar 0,0434 dan RMSE sebesar 0,0561, yang lebih rendah dibandingkan model sebelum optimasi dengan MAE 0,0825 dan RMSE 0,1037. Penurunan nilai *error* tersebut mengindikasikan bahwa prediksi model menjadi lebih dekat terhadap nilai aktual. Selain itu, nilai koefisien determinasi (R²) meningkat dari 0,7470 pada model menjadi 0,9259 pada model hasil optimasi. Nilai tersebut menunjukkan bahwa sekitar 92,59% variasi inflasi bulanan mampu dijelaskan oleh model LightGBM yang telah dioptimasi. Peningkatan nilai R² menunjukkan bahwa model memiliki kemampuan yang lebih baik dalam menggeneralisasi pola hubungan nonlinear di antara variabel-variabel ekonomi.

Berdasarkan hasil evaluasi, optimasi menggunakan Optuna berhasil meningkatkan performa LightGBM secara signifikan dibandingkan penggunaan parameter *default*. Hal tersebut menunjukkan bahwa pemilihan *hyperparameter* yang tepat mampu meningkatkan kemampuan model dalam mempelajari pola data inflasi. Peningkatan ini mengindikasikan bahwa konfigurasi parameter yang optimal memungkinkan LightGBM membentuk struktur pohon keputusan yang lebih sesuai dengan karakteristik data sehingga menghasilkan prediksi yang lebih akurat.

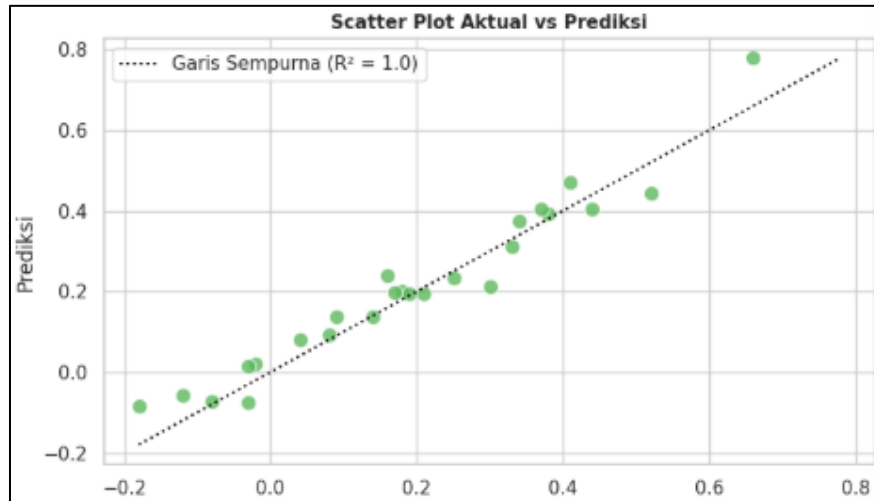
Hasil tersebut sejalan dengan karakteristik data inflasi pada penelitian ini yang dipengaruhi oleh berbagai indikator makroekonomi, seperti *Volatile Food*, *Administered Prices*,

suku bunga, nilai tukar, ekspektasi inflasi, dan Indeks Keyakinan Konsumen (IKK), yang memiliki hubungan kompleks dan tidak selalu bersifat linear terhadap inflasi. Sebagai algoritma berbasis *gradient boosting*, LightGBM mampu memodelkan hubungan nonlinier serta interaksi antarvariabel tanpa memerlukan asumsi linearitas. Ke et al. [18] menjelaskan bahwa LightGBM menggunakan strategi *leaf-wise tree growth* yang mampu meningkatkan akurasi prediksi dengan efisiensi komputasi yang baik. Selain itu, Goulet Coulombe et al. [23] menyatakan bahwa keunggulan utama *machine learning* dalam peramalan makroekonomi berasal dari kemampuannya menangkap hubungan nonlinier. Temuan tersebut juga didukung oleh Lindenmeyer dan Torrent [24], yang menunjukkan bahwa metode berbasis *boosting* efektif digunakan untuk memodelkan variabel makroekonomi karena mampu mengakomodasi hubungan linear maupun nonlinier dalam data sehingga meningkatkan kemampuan prediksi.



Gambar 4. Grafik Perbandingan Inflasi Aktual dan Prediksi

Gambar 4 menunjukkan perbandingan antara nilai inflasi aktual dengan hasil prediksi menggunakan model LightGBM dan LightGBM yang dioptimasi menggunakan Optuna pada data pengujian. Secara visual, kedua model mampu mengikuti pola umum pergerakan inflasi aktual, yang menunjukkan bahwa model berhasil mempelajari hubungan antara variabel input dan inflasi. Namun, model LightGBM dengan optimasi Optuna menunjukkan tingkat kesesuaian yang lebih baik terhadap data aktual dibandingkan model LightGBM tanpa optimasi. Hal ini ditunjukkan oleh garis prediksi LightGBM Optuna yang lebih sering berimpit atau berada dekat dengan garis aktual pada sebagian besar periode pengamatan. Sebaliknya, model LightGBM tanpa optimasi menunjukkan penyimpangan yang lebih besar, terutama saat terjadi perubahan inflasi yang relatif tajam. Pada awal periode pengujian, LightGBM menghasilkan prediksi yang cenderung lebih tinggi dibandingkan nilai aktual, sedangkan LightGBM Optuna memberikan prediksi yang lebih mendekati nilai sebenarnya. Pola serupa juga terlihat pada pertengahan hingga akhir periode pengujian ketika LightGBM menunjukkan kesalahan prediksi yang lebih besar, terutama saat inflasi berada pada level rendah maupun negatif. Secara keseluruhan, hasil ini menunjukkan bahwa optimasi hyperparameter mampu meningkatkan kemampuan LightGBM dalam mengikuti pola perubahan inflasi sehingga menghasilkan prediksi yang lebih akurat dibandingkan model tanpa optimasi.



Gambar 5. Scatter Plot Hasil Prediksi LightGBM+Optuna

Hasil *scatter plot* antara nilai aktual dan prediksi yang ditunjukkan pada Gambar 4, memperlihatkan bahwa sebagian besar titik berada di sekitar garis diagonal sempurna ($R^2 = 1$). Hal ini menunjukkan bahwa hasil prediksi memiliki tingkat kedekatan yang tinggi terhadap data aktual. Penyebaran titik yang relatif rapat mengindikasikan bahwa hasil prediksi model cenderung kecil dan stabil. Secara keseluruhan, hasil penelitian menunjukkan bahwa penggunaan optimasi *hyperparameter* berbasis Optuna mampu meningkatkan performa LightGBM secara signifikan dibandingkan penggunaan parameter. Kombinasi LightGBM dan Optuna menghasilkan model yang lebih *robust*, lebih stabil, dan lebih mampu menangkap pola inflasi bulanan secara akurat pada data pengujian. Oleh karena itu, pendekatan ini dinilai efektif untuk digunakan dalam prediksi inflasi berbasis *machine learning*.

4. KESIMPULAN

Berdasarkan penelitian yang dilakukan, prediksi inflasi bulanan menggunakan LightGBM dengan optimasi *hyperparameter* Optuna mampu menghasilkan performa yang baik. Hasil yang diperoleh menunjukkan bahwa proses optimasi *hyperparameter* berkontribusi dalam meningkatkan akurasi model, dengan menghasilkan MAE sebesar 0,0434, RMSE sebesar 0,0561, serta koefisien determinasi (R^2) sebesar 0,9259. Hasil penelitian juga menunjukkan bahwa kombinasi LightGBM dan Optuna efektif dalam mempelajari pola inflasi yang kompleks, dinamis, dan bersifat nonlinear. Selain itu, penerapan *feature engineering* membantu model menangkap informasi historis serta karakteristik data deret waktu yang berkontribusi terhadap peningkatan kualitas prediksi. Dengan demikian, tujuan penelitian untuk mengembangkan dan mengevaluasi model prediksi inflasi Indonesia menggunakan LightGBM yang dioptimasi dengan Optuna telah tercapai. Meskipun demikian, penelitian ini masih memiliki keterbatasan pada jumlah data yang digunakan serta kemungkinan adanya faktor eksternal yang belum sepenuhnya terwakili dalam model, seperti perubahan harga komoditas, kebijakan pemerintah, dan *shock* ekonomi. Oleh karena itu, penelitian selanjutnya dapat dilakukan dengan menggunakan periode data yang lebih panjang, menambahkan variabel ekonomi yang lebih beragam, serta membandingkan performa model dengan metode prediksi lainnya. Pengembangan tersebut diharapkan dapat meningkatkan akurasi prediksi dan memperluas pemanfaatan model sebagai alat pendukung dalam perencanaan serta pengambilan kebijakan ekonomi di masa mendatang.

REFERENSI

- [1] R. A. F. Salguero, “An Analysis of Monetary Policy Evidence and Theory through Meta-Analyses,” Sep. 2025, [Online]. Available: <http://arxiv.org/abs/2509.19591>
- [2] A. Alkhamov and B. Kriuk, “To What Extent Can Public Equity Indices Statistically Hedge Real Purchasing Power Loss in Compounded Structural Emerging-Market Crises? An Explainable ML-Based Assessment,” Jul. 2025, [Online]. Available: <http://arxiv.org/abs/2507.13055>
- [3] F. Ayivodji, E. Briand, K. Moran, and D. Stevanovic, “Monetary Policy in the Media Spotlight: Sentiments, Signals, and Economic Impact,” May 2026, [Online]. Available: <http://arxiv.org/abs/2605.15092>
- [4] M. C. Medeiros, G. F. R. Vasconcelos, Á. Veiga, and E. Zilberman, “Forecasting Inflation in a Data-Rich Environment: The Benefits of Machine Learning Methods,” *Journal of Business & Economic Statistics*, vol. 39, no. 1, pp. 98–119, Jan. 2021, doi: 10.1080/07350015.2019.1637745.
- [5] E. Nakamura, “Inflation forecasting using a neural network,” *Econ. Lett.*, vol. 86, no. 3, pp. 373–378, Mar. 2005, doi: 10.1016/j.econlet.2004.09.003.
- [6] J. H. Stock and M. W. Watson, “Why Has U.S. Inflation Become Harder to Forecast?,” *J. Money Credit Bank.*, vol. 39, no. s1, pp. 3–33, Feb. 2007, doi: 10.1111/j.1538-4616.2007.00014.x.
- [7] W. Badrawani, “An Interpretable Machine Learning Approach in Predicting Inflation Using Payments System Data: A Case Study of Indonesia,” 2025. doi: 10.48550/arXiv.2506.10369.
- [8] A. Yusnitasari, “Peramalan Inflasi Indonesia dengan Menggunakan Metode ARIMA Box-Jenkins,” *Jurnal Paradigma Multidisipliner (JPM)*, vol. 1, no. 2, 2020, doi: 10.1210/v1i2.25.
- [9] B. Noviana, “Forecasting Inflation in Indonesia Using ARIMAX : The Role of Money Supply and Exchange Rate,” vol. 13, no. 02, 2025, doi: 10.33019/equity.v13i2.370.
- [10] D. Pebrianti, H. Kurniawan, L. Bayuaji, and R. Rusdah, “XgBoost Hyper-Parameter Tuning Using Particle Swarm Optimization for Stock Price Forecasting,” *Jurnal Ilmiah Teknik Elektro Komputer dan Informatika*, vol. 9, no. 4, pp. 1179–1195, Jan. 2024, doi: 10.26555/jiteki.v9i4.27712.
- [11] C. Ivaşcu, “Can Machine Learning Models Predict Inflation?,” *Proceedings of the International Conference on Business Excellence*, vol. 17, no. 1, pp. 1748–1756, Jul. 2023, doi: 10.2478/picbe-2023-0155.
- [12] M. M. Al Haromainy, D. A. Prasetya, and A. P. Sari, “Improving Performance of RNN-Based Models With Genetic Algorithm Optimization For Time Series Data,” *TIERS Information Technology Journal*, vol. 4, no. 1, pp. 16–24, Jun. 2023, doi: 10.38043/tiers.v4i1.4326.
- [13] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, “Optuna: A Next-generation Hyperparameter Optimization Framework,” Jul. 2019, [Online]. Available: <http://arxiv.org/abs/1907.10902>
- [14] Y. Cai, J. Feng, Y. Wang, Y. Ding, Y. Hu, and H. Fang, “The Optuna-LightGBM-XGBoost Model: A Novel Approach for Estimating Carbon Emissions Based on the Electricity-Carbon Nexus,” *Applied Sciences*, vol. 14, no. 11, p. 4632, May 2024, doi: 10.3390/app14114632.
- [15] A. P. Sari, H. Suzuki, T. Kitajima, T. Yasuno, and D. A. Prasetya, “Prediction Model of Wind Speed and Direction Using Deep Neural Network,” *JEEMecs (Journal of Electrical Engineering, Mechatronic and Computer Science)*, vol. 3, no. 1, Feb. 2020, doi: 10.26905/jeemecs.v3i1.3946.
- [16] R. J. . Hyndman and George. Athanasopoulos, *Forecasting : principles and practice*. OTexts, 2021. Accessed: Jul. 10, 2026. [Online]. Available: <https://otexts.com/fpp3/>

- [17] H. Lütkepohl and F. Xu, “The role of the log transformation in forecasting economic variables,” *Empir. Econ.*, vol. 42, no. 3, pp. 619–638, Jun. 2012, doi: 10.1007/s00181-010-0440-1.
- [18] G. Ke *et al.*, “LightGBM: A Highly Efficient Gradient Boosting Decision Tree,” in *Advances in Neural Information Processing Systems*, I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, Eds., Curran Associates, Inc., 2017. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2017/file/6449f44a102fde848669bdd9eb6b76fa-Paper.pdf
- [19] Y. Si, S. Nadarajah, Z. Zhang, and C. Xu, “Modeling opening price spread of Shanghai Composite Index based on ARIMA-GRU/LSTM hybrid model,” *PLoS One*, vol. 19, no. 3 March, Mar. 2024, doi: 10.1371/journal.pone.0299164.
- [20] D. C. Montgomery, C. L. Jennings, and M. Kulahci, *Introduction to time series analysis and forecasting*, Second Edition. Hoboken, New Jersey: John Wiley & Sons, Inc., 2015. Accessed: Jul. 10, 2026. [Online]. Available: [https://cdn.oujdaibrary.com/books/78/78-introduction-to-time-series-analysis-and-forecasting-\(www.tawcer.com\).pdf](https://cdn.oujdaibrary.com/books/78/78-introduction-to-time-series-analysis-and-forecasting-(www.tawcer.com).pdf)
- [21] A. R. Farandy, “Analyzing Factors Affecting Indonesian Food Price Inflation,” *Jurnal Ekonomi dan Pembangunan*, vol. 28, no. 1, pp. 65–76, Dec. 2020, doi: 10.14203/JEP.28.1.2020.65-76.
- [22] A. Theresia, M. Ikhsan, F. N. Kacaribu, and S. Sumarto, “Spillover Effect of Food Producer Price Volatility in Indonesia,” *Economies*, vol. 13, no. 9, p. 256, Sep. 2025, doi: 10.3390/economies13090256.
- [23] P. Goulet Coulombe, M. Leroux, D. Stevanovic, and S. Surprenant, “How is machine learning useful for macroeconomic forecasting?,” *Journal of Applied Econometrics*, vol. 37, no. 5, pp. 920–964, Aug. 2022, doi: 10.1002/jae.2910.
- [24] G. Schultz Lindenmeyer and H. da Silva Torrent, “Boosting and Predictability of Macroeconomic Variables: Evidence from Brazil,” *Comput. Econ.*, vol. 64, no. 1, pp. 377–409, Jul. 2024, doi: 10.1007/s10614-023-10421-3.