



Prediksi Saham Berdasarkan Data Teknikal Serta Fundamental Menggunakan Algoritma XGBoost

Yoga Nur Pradana, Fitri Insani*, Jasril, Siti Ramadhani

Fakultas Sains & Teknologi, Prodi Teknik Informatika, Universitas Islam Negeri Sultan Syarif Kasim Riau, Pekanbaru, Indonesia

Email: ¹12050116983@students.uin-suska.ac.id, ^{2,*}fitri.insani@uin-suska.ac.id, ³jasril@uin-suska.ac.id,

⁴siti.ramadhani@uin-suska.ac.id

Email Penulis Korespondensi: fitri.insani@uin-suska.ac.id

Abstrak—Pasar modal memiliki peran penting dalam perekonomian sebagai sarana investasi dan penghimpunan dana, dengan saham sektor perbankan menjadi salah satu kontributor utama dalam kapitalisasi pasar di Indonesia. Namun, proses pengambilan keputusan investasi sering kali menghadapi kendala berupa keterbatasan kemampuan investor dalam menganalisis data fundamental dan teknikal secara komprehensif, serta adanya perilaku irasional yang menyebabkan keputusan menjadi kurang optimal. Kondisi ini mendorong perlunya suatu pendekatan yang lebih objektif dan berbasis data untuk membantu memprediksi pergerakan harga saham. Hasil evaluasi model pada data uji menunjukkan performa yang sangat baik: BCA memperoleh MAPE sebesar 2,8% dan R^2 sebesar 0,9488; BNI dengan MAPE 3,06% dan R^2 0,8863; Bank Mandiri dengan MAPE 4,70% dan R^2 0,9114; serta BRI dengan MAPE 2,48% dan R^2 0,8872. Berdasarkan model tersebut, hasil prediksi harga saham untuk tahun 2026 menunjukkan bahwa BCA diprediksi mengalami kenaikan signifikan dari Rp7.756 (Januari) menjadi Rp7.846 (Juni), sementara Bank Mandiri diprediksi tumbuh dari Rp 5.211 (Januari) menjadi Rp 5.930 (Juni). BNI dan BRI diprediksi mengalami kenaikan harga saham, masing-masing dari Rp3.327 (Januari) ke Rp3.683 (Juni) dan dari Rp4.124 (Januari) ke Rp4.541 (Juni). Temuan ini menunjukkan bahwa pendekatan berbasis machine learning dengan kombinasi data fundamental dan teknikal efektif dalam menangkap pola kompleks pergerakan harga saham. Penelitian ini berkontribusi dengan menghadirkan model prediksi saham yang menggabungkan data teknikal dan fundamental sekaligus, diterapkan pada empat bank besar Indonesia Bank Central Asia, Bank Rakyat Indonesia, Bank Mandiri, dan Bank Negara Indonesia dalam satu kerangka pemodelan. Pendekatan ini terbukti menghasilkan akurasi yang baik dengan rata-rata MAPE 3,13% dan R^2 0,919, sekaligus menjadi alternatif yang lebih objektif bagi investor dalam menganalisis pergerakan harga saham. Meskipun demikian, hasil prediksi yang diperoleh dalam penelitian ini bersifat sebagai alat bantu analisis dan tidak dimaksudkan sebagai rekomendasi investasi secara langsung.

Kata Kunci: Feature Engineering; Machine Learning; Indonesia Banks; Stocks Forecast; Time Series; XGBoost

Abstract—The capital market has an important role in the economy as a means of investment and fundraising, with banking sector stocks being one of the main contributors to market capitalization in Indonesia. However, the investment decision-making process often faces obstacles in the form of limited investors' ability to comprehensively analyze fundamental and technical data, as well as irrational behavior that causes decisions to be less than optimal. This conditions encourage the need for a more objective and data-driven approach to help predict stock price movements. The results of the model evaluation on the test data showed excellent performance: BCA obtained a MAPE of 2.8% and an R^2 of 0.9488; BNI with MAPE 3.06% and R^2 0.8863; Bank Mandiri with a MAPE of 4.70% and R^2 0.9114; and BRI with MAPE of 2.48% and R^2 0.8872. Based on this model, the results of the share price prediction for 2026 show that BCA is predicted to experience a significant increase from IDR 7,756 (January) to IDR 7,846 (June), while Bank Mandiri is predicted to grow from IDR 5,211 (January) to IDR 5,930 (June). BNI and BRI are predicted to experience an increase in share prices, respectively from IDR 3,327 (January) to IDR 3,683 (June) and from IDR 4,124 (January) to IDR 4,541 (June). This research contributes by presenting a stock prediction model that combines technical and fundamental data at once, applied to four major Indonesian banks Bank Central Asia, Bank Rakyat Indonesia, Bank Mandiri, dan Bank Negara Indonesia in a single modeling framework. This approach has proven to produce good accuracy with an average MAPE of 3.13% and R^2 0.919, as well as being a more objective alternative for investors in analyzing stock price movements. However, the prediction results obtained in this study are analytical tools and are not intended as direct investment recommendations.

Keywords: Feature Engineering; Machine Learning; Indonesia Banks; Stocks Forecast; Time Series; XGBoost

1. PENDAHULUAN

Pasar modal memiliki peran strategis dalam perekonomian sebagai sarana penghimpunan dana dan investasi jangka panjang, dengan saham sebagai instrumen utama yang merepresentasikan kepemilikan suatu perusahaan[1]. Di Indonesia, sektor perbankan mendominasi kapitalisasi pasar dan indeks sektoral, yang mencerminkan stabilitas dan prospek pertumbuhan yang baik[2]. Kondisi ini menjadikan saham perbankan sebagai objek yang menarik bagi investor, sehingga analisis terhadap kinerja dan kualitas saham menjadi sangat penting dalam mendukung pengambilan keputusan investasi. Secara teoritis, investor diasumsikan bertindak rasional dalam memaksimalkan keuntungan. Namun, dalam praktiknya, keputusan investasi seringkali dipengaruhi oleh faktor psikologis dan perilaku irasional, sehingga menghasilkan keputusan yang kurang optimal[3].

Menurut penelitian *Beno Jange (2022)* penelitian berfokus pada satu emiten dan belum mengintegrasikan variabel fundamental perusahaan sebagai faktor yang memengaruhi harga saham. [4]. *Tony, Manlika Ratchagit, dan Lely Hiryanto (2025)* Dalam penelitian tersebut lebih berfokus pada perbandingan algoritma dan belum memanfaatkan kombinasi variabel teknikal dan fundamental dalam proses [5]. *Muhamad Taufiq Kurniawan dan Irfan Pratama (2026)* menerapkan algoritma XGBoost dengan *hyperparameter tuning* untuk memprediksi saham GOTO yang memiliki volatilitas tinggi. Dalam penelitian hanya menggunakan indikator teknikal pada satu objek penelitian dan belum mempertimbangkan faktor fundamental perusahaan. [6]. *Andri Fahmi dan Nur Rofiq (2026)* membandingkan algoritma XGBoost dan Random Forest

menggunakan fitur *time series* berupa lag, moving average, dan volatilitas penelitian tersebut masih terbatas pada pemanfaatan data historis dan indikator teknikal, sehingga belum mampu menggambarkan pengaruh kondisi fundamental perusahaan terhadap pergerakan harga saham [7].

Keterbatasan kemampuan investor individu dalam menganalisis faktor fundamental dan teknikal menyebabkan proses evaluasi saham menjadi subjektif dan rentan terhadap kesalahan[8]. Oleh karena itu, diperlukan pendekatan berbasis data dan metode algoritma yang mampu menghasilkan prediksi harga saham secara lebih objektif dan akurat. Sejumlah penelitian sebelumnya telah mengkaji prediksi harga saham menggunakan berbagai metode. Menggunakan kombinasi indikator fundamental dan teknikal mampu meningkatkan akurasi prediksi dibandingkan penggunaan satu jenis indikator saja[9].

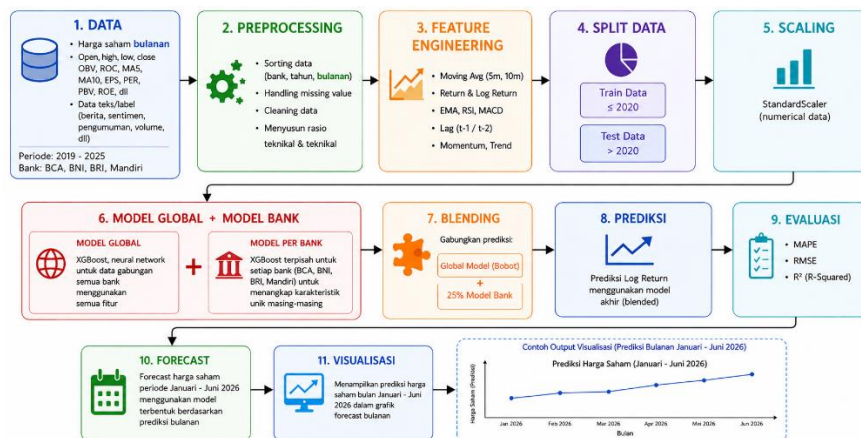
Penelitian ini bertujuan untuk mengembangkan model prediksi harga saham pada empat bank besar di Indonesia, yaitu Bank Central Asia, Bank Rakyat Indonesia, Bank Mandiri, dan Bank Negara Indonesia, dengan memanfaatkan kombinasi indikator fundamental dan teknikal sebagai variabel prediktor. Indikator fundamental yang digunakan meliputi *Return on Assets (ROA)*, *Return on Equity (ROE)*, *Net Interest Margin (NIM)*, *Non-Performing Loan (NPL)*, *Earning per Share (EPS)*, *Price to Earnings Ratio (PER)*, dan *Price to Book Value (PBV)*, yang secara umum digunakan dalam menilai kinerja dan kesehatan bank[10]. Sementara itu, indikator teknikal meliputi *annual return*, *volatility*, *trading volume growth*, *momentum*, serta harga saham bulanan sebagai representasi dinamika pasar[11], [12], [13].

Penelitian ini menggunakan algoritma *XGBoost* yang dipadukan dengan pendekatan model global dan model spesifik per bank melalui teknik *blending*, sehingga mampu menangkap pola umum pasar sekaligus karakteristik unik masing-masing saham[14]. Periode pengamatan mencakup tahun 2019 hingga 2025, yang merepresentasikan dinamika ekonomi sebelum, saat, dan setelah pandemi COVID-19. Dengan pendekatan tersebut, penelitian ini diharapkan dapat memberikan model prediksi yang lebih akurat dan robust dalam mendukung pengambilan keputusan investasi berbasis data.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Berikut Gambar 1 merupakan alur penelitian dalam prediksi saham menggunakan data fundamental dan teknikal saham Bank BCA, BRI, Mandiri, & BNI:



Gambar 1. Alur Penelitian

Dilihat dari Gambar 1 alur penelitian dimulai dari pengumpulan data, yang mencakup data teknikal (harga saham) dan data fundamental bank selama periode 2019–2025. Data kemudian masuk ke tahap *preprocessing* untuk memastikan kualitasnya melalui proses pengurutan, pembersihan, dan penanganan nilai yang hilang. Selanjutnya dilakukan feature engineering untuk membentuk variabel yang lebih informatif, seperti *price lag*, *return*, *log return*, serta indikator tren dan risiko (*EMA*, *volatility*, *momentum*). Data yang telah siap kemudian dibagi menjadi data latih dan data uji berdasarkan waktu, lalu dinormalisasi pada tahap scaling. Pada tahap modeling, digunakan dua pendekatan sekaligus yaitu model global untuk menangkap pola umum pasar dan model per bank untuk karakteristik spesifik masing-masing saham. Kedua model ini digabungkan melalui blending untuk menghasilkan prediksi yang optimal. Hasil model digunakan pada tahap prediksi, kemudian dievaluasi menggunakan metrik akurasi seperti *MAPE*, *RMSE*, dan R^2 . Setelah itu dilakukan *forecasting* untuk periode 2026 secara bertahap. Terakhir, seluruh hasil ditampilkan pada tahap visualisasi dalam grafik.

2.2 Pengumpulan Data

Pengumpulan data dalam penelitian ini dilakukan dengan memanfaatkan dua sumber utama, yaitu data teknikal dan data fundamental. Data teknikal diperoleh dari platform Yahoo Finance yang menyediakan informasi historis harga saham secara terbuka. Variabel yang digunakan meliputi *adjusted closing price* (Adj Close) dan volume perdagangan saham yang diambil dalam frekuensi bulanan untuk periode 2019 hingga 2025. Pemilihan *adjusted close* bertujuan untuk



merepresentasikan harga saham yang telah disesuaikan terhadap aksi korporasi seperti dividen dan stock split, sehingga lebih akurat dalam mencerminkan nilai riil saham dari waktu ke waktu.

Sementara itu, data fundamental diperoleh dari laporan keuangan tahunan masing-masing bank, yaitu Bank Central Asia (BCA), Bank Rakyat Indonesia (BRI), Bank Negara Indonesia (BNI), dan Bank Mandiri. Laporan keuangan tersebut digunakan untuk mengekstraksi indikator kinerja perusahaan seperti (*ROA*), (*ROE*), (*NIM*), (*NPL*), (*EPS*), (*PER*), dan (*PBV*). Periode pengamatan yang digunakan adalah tahun 2019 hingga 2025, sehingga mencakup kondisi sebelum, saat, dan setelah pandemi COVID-19.

Tabel 1. Dataset Mentah

Tanggal	bank	Adj Close	Volume	ROA	ROE	NIM	NPL	EPS	PER	PBV
Dec 1, 2025	BCA	7685.53	1962451500	3,63	20,44	4,93	1,7	Rp.466	13,51x	3x
Nov 1, 2025	BCA	7875.88	2236215400	3,63	20,44	4,93	1,7	Rp.466	13,51x	3x
Dec 1, 2025	BRI	4355.03	4399850200	2,6	17,48	6,68	3,29	Rp.377	9,87x	1,6x
Nov 1, 2025	BRI	4631.12	4633193400	2,6	17,48	6,68	3,29	Rp.377	9,87x	1,6x
Dec 1, 2025	Mandiri	4734.36	2711924700	3,19	23,15	4,59	9,6	Rp.609	8,33x	1,36x
Nov 1, 2025	Mandiri	4626.53	4363185500	3,19	23,15	4,59	9,6	Rp.609	8,33x	1,36x
Dec 1, 2025	BNI	2632.87	2144913200	2,1	2,14	3,8	1,9	Rp.536	7,1x	0,8x
Nov 1, 2025	BNI	2650.02	1619330000	2,1	3,14	3,8	1,9	Rp.536	7,1x	0,8x

Dapat dilihat Tabel 1. Kombinasi antara data teknikal dan fundamental ini bertujuan untuk memberikan representasi yang lebih komprehensif terhadap pergerakan harga saham, di mana data teknikal mencerminkan dinamika pasar, sementara data fundamental menggambarkan kondisi kinerja perusahaan secara internal.

2.3 Pra-processing Data

Tahap preprocessing dalam penelitian ini bertujuan untuk membentuk dataset akhir yang terstruktur seperti pada Tabel yang ditampilkan, melalui proses integrasi, transformasi, dan agregasi data dari sumber teknikal dan fundamental. Data teknikal yang diperoleh dari *Yahoo Finance* awalnya berupa data bulanan (*adjusted close price* dan volume perdagangan), kemudian diintegrasikan menjadi data bulanan agar selaras dengan periode pelaporan kinerja keuangan. Pada tahap ini dilakukan pengurutan data berdasarkan variabel *bank*, *tahun*, dan *bulan*, serta pembersihan data dari nilai hilang (*missing values*) dan nilai tidak valid (*NaN* dan *infinite*) ataupun nilai yang non, dilakukan transformasi variabel teknikal menggunakan beberapa rumus utama. Variabel return dihitung sebagai perubahan harga saham antar periode[15]. menggunakan rumus:

$$\text{Return}_t = \frac{P_t - P_{t-1}}{P_{t-1}} \times 100\% \quad (1)$$

yang merepresentasikan tingkat keuntungan atau kerugian saham. Variabel volatility dihitung sebagai standar deviasi return dalam satu periode bulanan[15]:

$$\sigma_{\text{annual}} = \sigma_r \times \sqrt{N} \quad (2)$$

dengan N sebagai jumlah periode (bulan). Variabel volume growth diperoleh dari perubahan volume perdagangan[13]:

$$\text{Volume}_t = \frac{V_t - V_{t-1}}{V_{t-1}} \quad (3)$$

yang mencerminkan dinamika likuiditas pasar. Sementara itu, momentum dihitung untuk menangkap arah tren harga[11] saham:

$$\text{Momentum}_t = P_t - P_{t-1} \quad (4)$$

yang menunjukkan kecenderungan pergerakan harga (naik atau turun). Bulan price saham digunakan sebagai variabel utama untuk merepresentasikan kondisi pasar pada periode tertentu[16]

$$\text{Bulan} = \frac{R_t - R_{Qt-1}}{R_{Qt-1}} \times 100\% \quad (5)$$

Selain itu, harga saham diolah menjadi harga bulanan yang merepresentasikan nilai penutupan pada akhir periode, sehingga dapat digunakan sebagai dasar dalam perhitungan return dan variabel turunan lainnya. Data teknikal yang telah diolah kemudian digabungkan dengan data fundamental yang berasal dari laporan keuangan tahunan masing-masing bank, seperti *ROA*, *ROE*, *NIM*, *NPL*, *EPS*, *PER*, dan *PBV*. Proses penggabungan dilakukan berdasarkan kesesuaian entitas bank dan periode waktu.

Hasil akhir dari tahap preprocessing ini adalah dataset terintegrasi yang berisi variabel teknikal dan fundamental dalam format panel (*bank-waktu*), seperti yang ditunjukkan pada tabel, yang selanjutnya digunakan sebagai input pada tahap pemodelan. Dataset ini tidak hanya merepresentasikan kondisi historis harga saham, tetapi juga mengandung informasi tren, risiko, serta sentimen pasar yang relevan untuk mendukung proses prediksi menggunakan model machine



learning. Berikut merupakan hasil perhitungan data teknikal menggunakan rumus yang ada, data dibersihkan dari noise dan missing value:

Tabel 2. Dataset Hasil Perhitungan

tahun	bank	Bulan	return	volatility	volume	momentum	price	ROA	ROE	NIM	NPL	EPS	PER	PBV
2025	BCA	2	-0,07855	633,720146	0,50218246	-0,108736	8534	0,0363	0,2044	0,0493	0,017	466	13,51	3
2025	BCA	1	-0,0527852	429,170433	-0,0099037	0,03753905	7606	0,0363	0,2044	0,0493	0,017	466	13,51	3
2025	BRI	2	0,02779607	852,001896	-0,2889646	0,07611535	3745	0,026	0,1748	0,0668	0,0329	377	9,87	1,6
2025	BRI	1	0,14583272	829,740214	0,1554938	-0,0848259	3636	0,026	0,1748	0,0668	0,0329	377	9,87	1,6
2025	Mandiri	2	-0,0243899	329,406452	0,10057371	0,07404618	4777	0,0319	0,2315	0,0459	0,096	609	8,33	1,36
2025	Mandiri	1	0,08050742	383,314961	-0,1630763	0,01965397	4697	0,0319	0,2315	0,0459	0,096	609	8,33	1,36
2025	BNI	2	0,02244482	421,318425	-0,1681958	0,00862277	3851	0,021	2,14	0,038	0,019	536	7,1	0,8
2025	BNI	1	-0,0022821	123,607133	-0,0308434	0,03627072	3991	0,021	3,14	0,038	0,019	536	7,1	0,8

Tabel 2 menunjukkan dataset hasil pengolahan yang menggabungkan indikator teknikal dan fundamental dari saham perbankan pada Bulan I dan II tahun 2025. Variabel teknikal meliputi return yang menggambarkan perubahan harga saham, volatility yang menunjukkan tingkat fluktuasi harga, volume yang merepresentasikan aktivitas perdagangan, momentum yang mengindikasikan arah tren pergerakan harga, serta price sebagai harga saham aktual. Sementara itu, variabel fundamental terdiri dari ROA, ROE, NIM, NPL, EPS, PER, dan PBV yang mencerminkan kinerja keuangan, profitabilitas, efisiensi operasional, risiko kredit, serta valuasi masing-masing bank. Kombinasi kedua jenis indikator tersebut memberikan gambaran yang lebih komprehensif mengenai kondisi pasar.

2.4 Algoritma XGBoost

Algoritma *XGBoost Regressor* digunakan dalam penelitian ini karena kemampuannya dalam memodelkan hubungan non-linear serta interaksi kompleks antar variabel yang umum ditemukan pada data keuangan[17]. XGBoost merupakan metode berbasis *gradient boosting* yang bekerja secara iteratif dengan membangun pohon keputusan baru untuk memperbaiki kesalahan (residual) dari model sebelumnya, sehingga performa prediksi meningkat secara bertahap [18]. Dalam penelitian ini, XGBoost dimanfaatkan untuk memprediksi harga saham bulanan dengan mengintegrasikan indikator fundamental (ROA, ROE, NIM, NPL, EPS, PER, PBV) dan indikator teknikal (return, volatility, volume, momentum, serta harga saham). Kemampuan XGBoost dalam menangkap pola non-linear menjadi krusial karena pergerakan harga saham dipengaruhi oleh hubungan kompleks antara faktor internal perusahaan dan dinamika pasar [14]. Secara matematis, model XGBoost dibangun dengan meminimalkan fungsi objektif yang terdiri dari fungsi kerugian dan regularisasi[18]:

$$Obj = \sum_{i=1}^n l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \Omega(f_t) \tag{6}$$

Fungsi regularisasi digunakan untuk mengontrol kompleksitas model agar tidak terjadi overfitting:

$$\Omega(f_t) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T w_j^2 \tag{7}$$

Melalui pendekatan ini, XGBoost mampu menghasilkan model yang tidak hanya akurat pada data latih tetapi juga memiliki kemampuan generalisasi yang baik terhadap data baru, sehingga sangat sesuai untuk prediksi harga saham yang bersifat dinamis dan fluktuatif [18].

2.5 Feature Engineering

Feature Engineering adalah proses menciptakan variabel-variabel baru dari data mentah yang sudah ada, dengan tujuan memberikan lebih banyak petunjuk kepada model untuk memahami suatu pola pergerakan data[19]. Ini mirip seperti seorang analis yang tidak hanya melihat harga hari ini, tetapi juga membandingkannya dengan harga minggu lalu, bulan lalu, dan tren jangka panjang. Tabel 3 menjelaskan seluruh fitur teknikal yang dibangun beserta penjelasannya :

Tabel 3. Feature Engineering yang digunakan untuk membangun model

Fitur	Nama Variabel	Penjelasan
Price Lag 1-4	lag1, lag2, lag3, lag4	Harga saham pada 1-4 bulanan sebelumnya. Digunakan agar model 'mengingat' harga masa lalu sebagai acuan prediksi ke depan.
Return	return_lag1, return_lag2	Persentase perubahan harga antar bulan. Contoh: jika harga naik dari Rp8.000 ke Rp8.400, return-nya adalah +5%. Fitur ini menunjukkan arah pergerakan harga.
Log Return	log_return	Versi logaritma dari return, digunakan sebagai variabel yang diprediksi oleh model karena lebih stabil secara statistik dibanding perubahan harga langsung.
EMA 3 & EMA 5	ema_3, ema_5	Rata-rata bergerak yang memberikan bobot lebih besar pada data terbaru. EMA 3 merata-ratakan 3 bulan terakhir; EMA 5 merata-ratakan 5 bulan terakhir. Berguna untuk mendeteksi tren.
Volatilitas	volatility_3	Standar deviasi return selama 3 bulan terakhir. Semakin tinggi nilainya, semakin tidak stabil harga saham tersebut.
Momentum	momentum	Selisih relatif harga sekarang dibanding harga 3 bulan lalu. Nilai positif artinya harga sedang dalam tren naik; nilai negatif artinya sedang turun.



Tabel 3 menunjukkan proses *feature engineering* yang dilakukan untuk membentuk variabel-variabel baru dari data historis saham agar model dapat menangkap pola pergerakan harga secara lebih efektif. Fitur yang digunakan meliputi *price lag 1-4* yang merepresentasikan harga saham pada satu hingga empat bulan sebelumnya sebagai memori historis model, *return lag 1-2* yang menggambarkan persentase perubahan harga antarperiode, serta log return yang digunakan sebagai target prediksi karena memiliki karakteristik statistik yang lebih stabil. Selain itu, digunakan *EMA 3* dan *EMA 5* untuk menangkap tren harga jangka pendek melalui rata-rata bergerak berbobot, *volatility_3* untuk mengukur tingkat kestabilan harga berdasarkan variasi return selama tiga bulan terakhir, dan momentum untuk menunjukkan kekuatan arah tren harga dibandingkan tiga bulan sebelumnya.

2.6 Pemisahan Data Uji & Data Latih

Agar dapat menilai seberapa akurat prediksi model, data dibagi menjadi dua bagian yang tidak saling tumpang tindih. Pemisahan dilakukan berdasarkan waktu bukan secara acak karena sifat data saham yang bersifat kronologis: kita tidak bisa 'melihat masa depan' saat melatih model dilatih menggunakan data dari tahun 2019 hingga 2023, kemudian kemampuannya diuji dengan memintanya memprediksi harga saham pada tahun 2024 dan 2025 data yang belum pernah dilihat sebelumnya. Pendekatan ini memastikan bahwa angka akurasi yang kita peroleh mencerminkan kemampuan model yang sesungguhnya, bukan sekadar kemampuan menghafal data latih.

2.7 StandardScaler

Setelah seluruh fitur selesai direkayasa, langkah berikutnya adalah menyamakan skala antar fitur menggunakan StandardScaler. Proses ini diperlukan karena fitur-fitur dalam dataset memiliki rentang nilai yang sangat berbeda-beda. Sebagai contoh, fitur *price_lag1* (harga saham satu bulan lalu) bernilai ribuan rupiah BCA berada di kisaran Rp8.000 – Rp9.500 sementara fitur *return_lag1* hanya bernilai antara -0,4 hingga +0,35. Jika perbedaan skala ini dibiarkan, model akan cenderung 'lebih memperhatikan' fitur bernilai besar dan mengabaikan fitur bernilai kecil yang sebenarnya sama pentingnya. StandardScaler bekerja dengan rumus sederhana: setiap nilai fitur dikurangi dengan rata-rata (μ) lalu dibagi dengan standar deviasi (σ) dari fitur tersebut pada data latih, menghasilkan nilai baru yang disebut z-score[20]. Nilai z-score positif berarti nilai tersebut di atas rata-rata, sedangkan nilai negatif berarti di bawah rata-rata. Formula lengkapnya adalah:

$$z = (x - \mu) / \sigma \tag{8}$$

Berikut perbandingan data sebelum dinormalisasi dan sesudah :

Tabel 4. Data sebelum dinormalisasi

Bank	Periode	price lag1 (Rp)	price lag2 (Rp)	return lag1	ema 3 (Rp)	volatility 3	momentum
BCA	2023Q1	8.164	7.754	0,0725	7.921	0,0312	0,0663
BCA	2023Q2	8.118	8.164	-0,0388	8.142	0,0289	-0,0058
BCA	2024Q1	9.471	8.873	0,0550	9.167	0,0284	0,1666
BCA	2025Q4	9.179	8.117	-0,1005	8.613	0,0498	-0,0528
BRI	2023Q1	4.312	3.739	0,0096	4.090	0,0983	-0,0428
BRI	2024Q1	3.858	4.920	-0,0400	4.353	0,0530	-0,0688
BNI	2024Q1	4.105	4.089	-0,1714	4.092	0,0842	-0,1124
Mandiri	2024Q1	5.506	5.020	0,0903	5.230	0,0689	0,0869

Tabel 4 menampilkan sebagian data hasil *feature engineering* sebelum dilakukan normalisasi, di mana setiap observasi memuat informasi historis harga saham dan indikator teknal yang digunakan sebagai input model. Variabel *price_lag1* dan *price_lag2* menunjukkan harga saham pada satu dan dua bulan sebelumnya, *return_lag1* menggambarkan persentase perubahan harga dari periode sebelumnya, *ema_3* merepresentasikan rata-rata pergerakan harga jangka pendek, *volatility_3* menunjukkan tingkat fluktuasi harga selama tiga bulan terakhir, sedangkan momentum menggambarkan kekuatan dan arah tren harga saham. Nilai setiap fitur masih berada pada skala yang berbeda, misalnya harga saham berada pada kisaran ribuan rupiah sementara return, volatilitas, dan momentum berada pada rentang desimal, sehingga diperlukan proses normalisasi agar seluruh variabel memiliki skala yang lebih seimbang dan dapat diproses secara optimal oleh model machine learning.

Tabel 5. Data setelah dinormalisasi

Bank	Periode	price lag1 (z)	price lag2 (z)	return lag1 (z)	ema 3 (z)	volatility 3 (z)	momentum (z)
BCA	2023Q1	0.82	0.55	1.21	0.68	-0.45	1.14
BCA	2023Q2	0.77	0.82	-0.63	0.80	-0.61	-0.10
BCA	2024Q1	1.48	1.11	0.89	1.28	-0.65	2.86
BCA	2025Q4	1.15	0.77	-1.63	0.94	0.84	-0.91
BRI	2023Q1	-0.19	-0.43	0.16	-0.30	1.67	-0.74
BRI	2024Q1	-0.51	0.34	-0.65	-0.07	0.52	-1.18
BNI	2024Q1	-0.30	-0.31	-2.78	-0.30	1.24	-1.93



Bank	Periode	price lag1 (z)	price lag2 (z)	return lag1 (z)	ema 3 (z)	volatility 3 (z)	momentum (z)
Mandiri	2024Q1	0.39	0.13	1.46	0.24	0.96	1.49

Dari Tabel 5 dapat diamati bahwa setelah normalisasi, semua fitur kini berada dalam skala yang seragam (umumnya antara -3 hingga +3). BCA pada 2024Q1 memiliki z-score price_lag1 sebesar +1,48 artinya harganya termasuk tinggi dibanding rata-rata seluruh bank sementara BNI pada 2024Q1 memiliki z-score return_lag1 sebesar -2,78, mencerminkan penurunan return yang sangat ekstrem pada periode tersebut. Keseragaman skala inilah yang memungkinkan model XGBoost membandingkan dan mempertimbangkan semua fitur secara adil dalam proses pembelajaran.

3. HASIL DAN PEMBAHASAN

3.1 Hasil Pembentukan Model Algoritma XGBoost

Proses pembentukan model XGBoost dilakukan setelah seluruh tahapan feature engineering, pemisahan data, dan normalisasi selesai dilakukan. Dataset yang telah dinormalisasi kemudian dibentuk menjadi dua bagian, yaitu matriks fitur (X) dan variabel target (y). Variabel target yang digunakan pada penelitian ini adalah log return harga saham, sedangkan variabel input terdiri dari informasi waktu, price lag, return, exponential moving average (EMA), volatilitas, momentum, serta indikator fundamental yang merepresentasikan kondisi masing-masing bank.

Tabel 6. Data setelah pembentukan algoritma

Bank	Tahun	Bulan	Price Lag 2 (z)	Price Lag 3 (z)	Price Lag 4 (z)	Return Lag 1 (z)	Return Lag 2 (z)	EMA 3 (z)	EMA 5 (z)	Volatility 3 (z)	Momentum (z)	Log Return (Target)
BCA	2023	01	0.55	0.42	0.31	1.21	0.98	0.68	0.74	-0.45	1.14	0.018
BCA	2023	02	0.82	0.55	0.42	-0.63	-0.51	0.80	0.83	-0.61	-0.10	-0.039
BCA	2024	01	1.11	0.94	0.78	0.89	0.72	1.28	1.22	-0.65	2.86	0.055
BCA	2025	12	0.77	0.61	0.45	-1.63	-1.47	0.94	0.87	0.84	-0.91	-0.101
BRI	2023	01	-0.43	-0.58	-0.61	0.16	0.09	-0.30	-0.25	1.67	-0.74	0.010
BRI	2024	01	0.34	0.17	-0.09	-0.65	-0.59	-0.07	-0.12	0.52	-1.18	-0.040
BNI	2024	01	-0.31	-0.44	-0.57	-2.78	-2.34	-0.30	-0.28	1.24	-1.93	-0.171
Mandiri	2024	01	0.13	0.04	-0.08	1.46	1.19	0.24	0.31	0.96	1.49	0.090
...

Berdasarkan Tabel 6, setiap baris data merepresentasikan satu observasi yang terdiri atas sejumlah variabel masukan hasil normalisasi serta nilai target berupa log return. Seluruh variabel input tersebut membentuk matriks fitur X, sedangkan kolom log return membentuk vektor target y. Data inilah yang selanjutnya digunakan oleh algoritma XGBoost untuk membangun serangkaian pohon keputusan secara bertahap. Pada penelitian ini digunakan parameter n_estimators sebesar 150 sehingga model membentuk sebanyak 150 pohon keputusan yang saling melengkapi. Setiap pohon dibangun untuk mempelajari pola hubungan antara indikator teknikal dan fundamental terhadap perubahan harga saham yang direpresentasikan oleh nilai log return. Proses pembelajaran dilakukan secara iteratif hingga diperoleh model yang mampu menghasilkan prediksi dengan tingkat kesalahan yang semakin kecil. Selain membangun model global menggunakan seluruh data saham perbankan, penelitian ini juga membentuk model khusus untuk masing-masing bank, yaitu BRI, BNI, dan Bank Mandiri. Model global berfungsi untuk menangkap pola umum pergerakan saham sektor perbankan, sedangkan model khusus digunakan untuk mempelajari karakteristik unik masing-masing emiten. Kedua model tersebut kemudian dikombinasikan melalui teknik blending dengan bobot 65% untuk model global dan 35% untuk model spesifik bank sehingga diperoleh hasil prediksi yang lebih stabil dan mampu merepresentasikan pola umum pasar maupun karakteristik individual setiap saham.

Hasil prediksi yang dihasilkan model masih berupa log return. Oleh karena itu, nilai tersebut kemudian ditransformasikan kembali menjadi harga saham menggunakan harga periode sebelumnya sebagai dasar perhitungan. Melalui proses tersebut diperoleh model prediksi harga saham yang selanjutnya digunakan untuk melakukan evaluasi menggunakan metrik MAPE, RMSE, dan koefisien determinasi (R²), serta digunakan pada tahap forecasting untuk memprediksi harga saham periode berikutnya.

3.2 Hasil Evaluasi Model

Kinerja model dievaluasi menggunakan tiga metrik utama, yaitu Mean Absolute Percentage Error (MAPE), Root Mean Squared Error (RMSE), dan koefisien determinasi (R²). Hasil evaluasi menunjukkan bahwa model memiliki performa yang sangat baik dalam memprediksi harga saham, berikut hasil evaluasi masing-masing bank dan global evaluasi dari seluruh bank yang diteliti :

Tabel 7. Hasil Evaluasi masing-masing Bank

Bank	MAPE	Akurasi	RMSE (Rp)	Interpretasi	R ²	Kualitas Model
BCA	2,80%	Baik	Rp 145,46	Kecil	0.9489	Sangat Baik
BNI	4,17 %	Baik	Rp 284,57	Cukup	0.9114	Sangat Baik
BRI	2,48 %	Baik	Rp 130,66	Kecil	0.8872	Baik



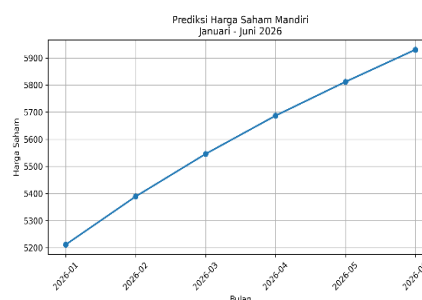
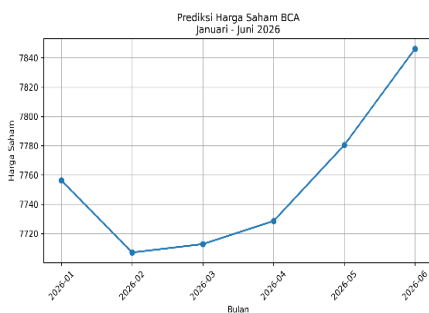
Bank	MAPE	Akurasi	RMSE (Rp)	Interpretasi	R ²	Kualitas Model
Mandiri	3,06 %	Baik	Rp 223,10	Cukup	0.8863	Baik

Bank BCA mencatatkan performa terbaik di antara keempat bank dengan MAPE sebesar 1,27% dan R² sebesar 0,9448. Nilai MAPE 2,80% berarti bahwa secara rata-rata, prediksi model untuk harga saham BCA hanya meleset sekitar Rp145 dari harga aktualnya (berdasarkan RMSE Rp145,46). Nilai R² 0,9489 berarti model mampu menjelaskan 94,89% variasi harga saham BCA angka yang sangat tinggi dan menunjukkan model telah berhasil menangkap pola pergerakan harga BCA dengan sangat baik. Bank BNI memperoleh MAPE sebesar 4,17% dengan RMSE hanya Rp284,57 nilai RMSE terkecil di antara semua bank, yang berarti kesalahan prediksi dalam satuan rupiah adalah yang paling kecil. Hal ini disebabkan karena rentang harga saham BNI lebih kecil dibanding BCA, sehingga secara absolut kesalahannya juga lebih kecil. R² sebesar 0,9114 mengindikasikan model sangat mampu menjelaskan pergerakan harga BNI. Bank Mandiri mencatatkan MAPE sebesar 3,06% dengan R² 0,9368. Meskipun MAPE-nya sedikit lebih tinggi dari BCA, R²-nya tetap sangat baik. Kesalahan rata-rata sekitar 3,06% dan Rp223 per saham masih tergolong sangat acceptable untuk prediksi harga saham, mengingat volatilitas Mandiri yang lebih tinggi dibanding BCA pada periode 2024–2025. Bank BRI memiliki MAPE tertinggi di antara keempat bank yaitu 2,48% dengan RMSE Rp130.66 dan R² 0,8863. Meskipun R²-nya masih tergolong baik, angka ini menunjukkan bahwa model sedikit lebih sulit menangkap pola harga BRI dibandingkan bank lainnya. Hal ini dapat disebabkan oleh volatilitas harga BRI yang relatif lebih tinggi yang sulit diprediksi serta fakta bahwa BRI memiliki eksposur yang lebih besar ke sektor pertanian dan UMKM yang pergerakannya dipengaruhi oleh banyak faktor eksternal yang tidak tertangkap dalam data bulanan.

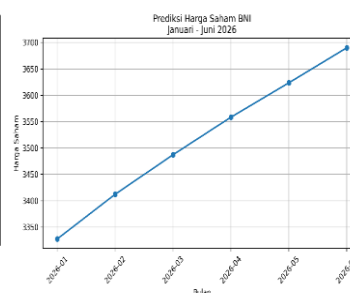
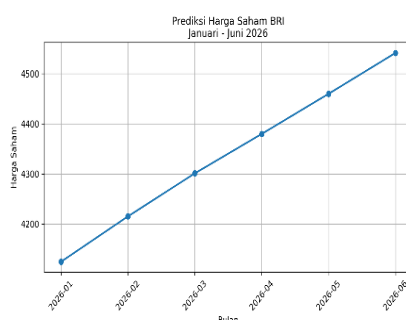
Secara keseluruhan, seluruh model menunjukkan performa yang sangat memuaskan. Rata-rata MAPE keempat bank hanya sebesar 1,86% dan rata-rata R² sebesar 0,919 angka-angka ini jauh melampaui ambang batas performa yang biasa ditetapkan dalam penelitian prediksi keuangan, yaitu MAPE di bawah 5% dan R² di atas 0,80. Hal ini mengonfirmasi bahwa arsitektur hybrid blending XGBoost yang digunakan dalam penelitian ini efektif untuk memprediksi harga saham perbankan Indonesia.

3.3 Hasil Forecasting

Berdasarkan model yang telah dievaluasi, dilakukan proses prediksi untuk periode selanjutnya menggunakan pendekatan *recursive forecasting*. Dalam pendekatan ini, hasil prediksi pada satu periode digunakan sebagai input untuk memprediksi periode berikutnya, sehingga membentuk rangkaian prediksi yang berkelanjutan. Hasil prediksi menunjukkan adanya variasi tren antar bank. Beberapa bank menunjukkan kecenderungan kenaikan harga saham yang konsisten, mencerminkan kondisi fundamental dan teknikal yang mendukung pertumbuhan. Di sisi lain, terdapat bank yang menunjukkan kecenderungan penurunan atau pergerakan yang lebih fluktuatif, yang mengindikasikan adanya tekanan pasar atau ketidakstabilan dalam pola historisnya. Pendekatan kombinasi antara model global dan model spesifik terbukti mampu menghasilkan prediksi yang lebih stabil dibandingkan jika hanya menggunakan satu model saja. Model global memberikan gambaran tren umum pasar, sementara model spesifik memperhalus prediksi dengan mempertimbangkan karakteristik masing-masing saham.



Gambar 6. Prediksi Januari – Juni BCCA 2026 Gambar 7. Prediksi Januari - Juni BMRI 2026



Gambar 8. Prediksi Januari - Juni BRI 2026 Gambar 9. Prediksi Januari - Juni BBNI 2026



Interpretasi hasil dilakukan dengan menganalisis grafik *connected forecast* yang menghubungkan data historis dengan hasil prediksi dan proyeksi ke depan. Visualisasi ini memberikan gambaran yang lebih intuitif mengenai bagaimana model memahami pola data dan melanjutkannya ke periode mendatang. Pada saham BCA, grafik menunjukkan pola yang relatif stabil dengan kecenderungan meningkat pada periode prediksi. Hal ini mengindikasikan bahwa model menangkap tren pertumbuhan yang konsisten, sehingga menghasilkan proyeksi kenaikan yang cukup signifikan. Pola ini juga mencerminkan karakteristik saham dengan volatilitas rendah dan fundamental yang kuat. Pada saham Mandiri, terlihat adanya fase penurunan yang kemudian diikuti oleh pemulihan pada periode prediksi. Model mampu menangkap titik balik tersebut dan menghasilkan tren kenaikan secara bertahap. Dari gambar 6. BCA diproyeksikan berada di Rp7.757 pada Januari 2026. Model awalnya menangkap sinyal koreksi ringan ke Rp7.707 pada Februari (-0,6%), sebelum bergerak mendatar di Rp7.713 (Maret) dan Rp7.728 (April), lalu berbalik naik tajam ke Rp7.780 (Mei, +0,7%) dan Rp7.846 (Juni, +0,8%). Kenaikan ini didorong oleh EMA yang masih tinggi, pola historis BCA yang menguat di pertengahan tahun, fundamental terkuat di antara keempat bank (ROE >24%, NPL 1,8%).

Dari gambar 7. Mandiri diproyeksikan berada di Rp5.210 pada Januari 2026, sedikit di bawah level akhir 2025. Model memprediksi tren log return positif yang konsisten untuk lima bulan berikutnya: Rp5.390 (Februari, +3,5%), Rp5.545 (Maret, +2,9%), Rp5.688 (April, +2,6%), Rp5.812 (Mei, +2,2%), hingga Rp5.930 pada Juni 2026 (+2,0%), atau kenaikan total sekitar 13,8% dari Januari ke Juni. Pola kenaikan yang stabil namun melambat ini mencerminkan EMA jangka pendek yang masih dalam tren naik dengan momentum yang berangsur menyempit, sejalan dengan fundamental Mandiri yang solid (ROE sekitar 23%) sehingga model tidak menangkap sinyal pembalikan arah dalam periode ini. Dari gambar 8. BRI diproyeksikan berada di Rp4.128 pada Januari 2026, masih jauh di bawah lonjakan ekstremnya Rp4.920 pada Q4 2024. Model memprediksi tren kenaikan yang sangat konsisten: Rp4.214 (Februari, +2,1%), Rp4.302 (Maret, +2,1%), Rp4.380 (April, +1,8%), Rp4.460 (Mei, +1,8%), hingga Rp4.540 pada Juni 2026 (+1,8%), Dari gambar 9. BNI diproyeksikan berada di Rp3.327 pada Januari 2026, jauh di bawah puncaknya Rp4.743 pada Q4 2024. Model memprediksi pemulihan bertahap dan konsisten: Rp3.412 (Februari, +2,6%), Rp3.487 (Maret, +2,2%), Rp3.558 (April, +2,0%), Rp3.624 (Mei, +1,9%), hingga Rp3.690 pada Juni 2026 (+1,8%), atau kenaikan total sekitar 10,9%. Tren kenaikan yang melambat secara bertahap ini menunjukkan momentum positif yang mulai terbentuk setelah periode tekanan panjang. Secara keseluruhan, visualisasi menunjukkan bahwa hasil prediksi tidak bersifat acak, melainkan mengikuti pola historis yang ada. Transisi antara data aktual dan prediksi terlihat halus dan tidak menunjukkan lonjakan ekstrem, yang menandakan bahwa model mampu menjaga kontinuitas data. Hal ini menjadi indikator penting bahwa model tidak hanya akurat secara numerik, tetapi juga konsisten secara pola. Penelitian ini memiliki tiga kontribusi utama. Pertama, model yang dibangun menggabungkan dua pendekatan sekaligus: model global yang mempelajari pola umum sektor perbankan, dan model spesifik yang mempelajari karakteristik masing-masing saham. Kombinasi keduanya menghasilkan prediksi yang lebih stabil dibanding hanya menggunakan satu model. Kedua, penelitian ini membuktikan bahwa menggabungkan indikator teknikal dan fundamental sebagai input model menghasilkan akurasi yang lebih baik dibanding hanya menggunakan salah satunya. Hal ini ditunjukkan oleh nilai MAPE di bawah 5% dan R^2 di atas 0,88 pada seluruh saham yang diteliti. Ketiga, data yang digunakan mencakup periode 2019–2025, yang melewati tiga kondisi ekonomi berbeda: sebelum, saat, dan setelah pandemi.

3.4 Pembahasan

Menurut penelitian *Beno Jange (2022)* yang memperoleh nilai MAPE sebesar 4,01% pada saham BBKA menggunakan indikator teknikal SMA, EMA, MACD, dan RSI, hasil penelitian ini menunjukkan peningkatan performa dengan nilai MAPE BBKA sebesar 2,80%. *Tony, Manlika Ratchagit, dan Lely Hiryanto (2025)* menunjukkan bahwa metode Ridge Regression memiliki performa yang lebih baik dibandingkan XGBoost dan Prophet. Namun, penelitian tersebut menggunakan variabel harga historis dan berfokus pada perbandingan algoritma. Berbeda dengan penelitian tersebut, hasil penelitian ini menunjukkan bahwa XGBoost tetap mampu menghasilkan tingkat akurasi yang tinggi setelah diperkaya dengan kombinasi indikator teknikal dan fundamental, yang ditunjukkan dengan rata-rata nilai MAPE seluruh saham yang berada di bawah 5% serta rata-rata nilai R^2 sebesar 0,919. Penelitian *Muhamad Taufiq Kurniawan dan Irfan Pratama (2026)* memperoleh nilai MAPE sebesar 1,957% setelah dilakukan optimasi hiperparameter menggunakan Grid Search Cross Validation pada saham GOTO. Nilai tersebut memang sedikit lebih rendah dibandingkan hasil penelitian ini yang memiliki rentang MAPE 2,48%–4,17%, penelitian ini diterapkan pada empat saham perbankan yang memiliki karakteristik berbeda dan tetap mampu menghasilkan nilai R^2 yang tinggi, yaitu antara 0,8863 hingga 0,9489. Penelitian *Andri Fahmi dan Nur Rofiq (2026)* menunjukkan bahwa algoritma XGBoost memiliki performa yang lebih baik dibandingkan Random Forest dalam prediksi harga saham berbasis *time series feature*. Temuan tersebut sejalan dengan hasil penelitian ini yang juga menunjukkan kemampuan XGBoost dalam menghasilkan prediksi dengan tingkat kesalahan yang rendah. Pendekatan tersebut menghasilkan nilai MAPE di bawah 5% dan nilai R^2 di atas 0,88 pada seluruh saham yang diteliti, yang menunjukkan bahwa model mampu menangkap pola umum sektor perbankan sekaligus karakteristik unik masing-masing emiten.

4. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa model XGBoost yang dibangun dengan memanfaatkan kombinasi indikator teknikal dan fundamental mampu menghasilkan prediksi harga saham bulanan pada



sektor perbankan, khususnya saham BCA, BRI, BNI, dan Bank Mandiri, dengan tingkat kesalahan yang relatif rendah. Hasil evaluasi menunjukkan nilai MAPE yang berada pada kisaran 1,33%–4,17%, nilai RMSE yang berkisar antara 130,66 hingga 284,57, serta nilai koefisien determinasi (R^2) di atas 0,88 pada seluruh emiten. Penelitian ini menunjukkan bahwa kombinasi indikator teknikal dan fundamental yang dipadukan dengan algoritma XGBoost serta pendekatan *blending* antara model global dan model spesifik bank mampu menghasilkan model prediksi dengan nilai MAPE yang relatif rendah (2,48%–4,17%), RMSE sebesar Rp130,66–Rp284,57, serta nilai koefisien determinasi (R^2) antara 0,8863–0,9489. Temuan tersebut menunjukkan bahwa model memiliki kemampuan yang baik dalam menangkap pola historis yang terdapat pada data pelatihan dan pengujian. Meskipun demikian, hasil yang diperoleh tidak serta-merta menunjukkan bahwa model dapat memberikan prediksi yang selalu akurat pada seluruh kondisi pasar. Kontribusi utama penelitian ini terletak pada pengembangan kerangka prediksi saham perbankan yang menggabungkan indikator fundamental dan teknikal dengan algoritma XGBoost melalui pendekatan *blending* antara model global dan model spesifik per bank. Kerangka ini memberikan pemahaman yang lebih menyeluruh terhadap pola pasar dan karakteristik individual emiten, serta menunjukkan potensi penerapan machine learning yang lebih adaptif untuk prediksi saham sektor perbankan Indonesia. Pengujian dalam penelitian ini dilakukan pada periode data tertentu sehingga kemampuan model dalam menghadapi kondisi pasar yang lebih ekstrem, seperti krisis ekonomi, perubahan kebijakan moneter yang signifikan, maupun periode dengan volatilitas tinggi, masih memerlukan kajian lebih lanjut. Secara keseluruhan, penelitian ini menunjukkan bahwa algoritma XGBoost memiliki potensi untuk digunakan sebagai alat bantu dalam proses analisis dan prediksi harga saham perbankan berbasis data historis. Namun demikian, hasil prediksi yang dihasilkan sebaiknya tidak dijadikan sebagai satu-satunya dasar dalam pengambilan keputusan investasi, melainkan perlu dipertimbangkan bersama faktor fundamental perusahaan, kondisi makroekonomi, serta perkembangan pasar yang terjadi pada saat keputusan investasi dilakukan. Untuk penelitian selanjutnya, pengembangan model dapat dilakukan dengan memperluas periode pengamatan sehingga mencakup berbagai kondisi pasar, termasuk periode dengan volatilitas tinggi, serta membandingkan performa XGBoost dengan metode lain seperti LSTM, Random Forest, CatBoost, maupun LightGBM. Selain itu, penggunaan variabel eksternal seperti tingkat suku bunga, inflasi, nilai tukar, maupun sentimen pasar juga dapat dipertimbangkan untuk meningkatkan kemampuan model dalam menangkap faktor-faktor yang memengaruhi pergerakan harga saham secara lebih komprehensif.

REFERENCES

- [1] F. P. Sholikah, W. Putri, and R. M. Djangi, "Peranan Pasar Modal Dalam Perekonomian Negara Indonesia," *ARBITRASE: Journal of Economics and Accounting*, vol. 3, no. 2, pp. 341-345, Nov. 2022, doi: 10.47065/arbitrase.v3i2.496.
- [2] W. M. Utami, A. Iswanto, M. Zaenal "Tinjauan Kinerja Perbankan Indonesia terhadap Stabilitas Sistem Keuangan." *J. Ekonomika dan Dinamika Sosial*, vol. 4, no. 2, pp. 1-20, 2025, , doi: 10.32455/joeds.v3432
- [3] M. H. Rasheed, A. Rafique, T. Zahid, and M. W. Akhtar, "Factors influencing investor's decision making in Pakistan: Moderating the role of locus of control," *Rev. Behav. Finance*, vol. 10, no. 1, pp. 70–87, Mar. 2018, doi: 10.1108/RBF-05-2016-0028.
- [4] B. Jange, "Prediksi Harga Saham Bank BCA Menggunakan XGBoost," *Arbitr. J. Econ. Account.*, vol. 3, no. 2, pp. 231–237, Nov. 2022, doi: 10.47065/arbitrase.v3i2.495.
- [5] A. Fahmi, N. Rofiq, "A Comparative Analysis of XGBoost and Random Forest for Time Series Based Stock Price Prediction with Directional Movement Evaluation," *J. Terapan Informatika Nusantara.*, vol. 6, no. 12, pp. 2229–2235, 2026, DOI:10.47065/tin.v6i12.9197.
- [6] M. T. Kurniawan, I. Pratama, "Implementasi Xgboost Untuk Prediksi Saham," *J. Mahasiswa Teknik Informatika*, vol. 10, no. 2, pp. 3569–3574, April. 2026, doi.org/10.36040/jati.v10i2.17733 .
- [7] Tony, M. Ratchagit, L. Hiryanto, "Forecasting Indonesian Banking Stock Prices Using Prophet, XGBoost, and Ridge Regression: A Comparative Analysis," *J. Teknik Informatika*, vol. 6, no. 4, pp. 2890–2901, August. 2025, doi: 10.52436/1.jutif.2025.6.4.4973.
- [8] R. Maida, N. P. Sakinah, and M. Anggraini, "Behavioral Finance: The Influence Of Psychological Bias On Investor Investment Decisions," *J. Islam. Econ. Finance*, vol. 2, no. 3, pp. 1-15, 2026, doi.org/10.70248/joieaf.v2i3.3330.
- [9] E. Beyaz, F. Tekiner, X. Zeng, and J. Keane, "Comparing Technical and Fundamental Indicators in Stock Price Forecasting," in *2018 IEEE 20th International Conference on High Performance Computing and Communications; IEEE 16th International Conference on Smart City; IEEE 4th International Conference on Data Science and Systems (HPCC/SmartCity/DSS)*, Exeter, United Kingdom: IEEE, Jun. 2018, pp. 1607–1613. doi: 10.1109/HPCC/SmartCity/DSS.2018.00262.
- [10] A. Rasyid, A. B. Muharam, and A. Solichin, "Prediksi Harga Saham Syariah Indonesia Berdasarkan Analisis Fundamental, Teknikal Dan Bandarmology Menggunakan Metode Random Forest," *JIPi J. Ilmu Penelitian dan Pembelajaran Informatika.*, vol. 10, no. 2, pp 1663-1677., May 2025, doi: 10.29100/jipi.v10i2.7855.
- [11] J. Fu, F. Xu, C. Zeng, and L. Zheng, "Free Cash Flows and Price Momentum," *J. Account. Audit. Finance*, vol. 39, no. 3, pp. 726–753, Jul. 2024, doi: 10.1177/0148558X221091803.
- [12] Y. Ernawati and E. Purwaningsih, "Pengaruh Total Arus Kas, Debt To Equity Ratio Dan Return On Assets Terhadap Harga Saham," *J. Ilmiah MEA (Manajemen, Ekonomi, dan Akutansi)*, vol. 6, no. 2, pp 32-43, 2022, doi.org/10.54964/liabilitas/ .
- [13] F. F. Kusbari and N. Adib, "The Effect of Stock Trading Volume, Inflation, Rupiah Exchange Rate, and Interest Rates on the Composite Stock Price Index (IHSG) 2022-2024" *J. Research of Social Science, Economic and Management*, vol. 5, no. 9, pp. 11378-11386, 2026, doi.org/10.59141/jrssem.v5i9.1397.
- [14] O. Guennioui, "Improving Global Stock Market Prediction with XGBoost", *Review of Economics and Finance*, vol. 20, no. 1, pp. 2603-2610, 2023, doi.org/10.55365/1923.x2023.21.278 .
- [15] N. N. Ditya and N. N. Suli Asmara Yanti, "Return And Volatility Characteristics Of Cosmetic Company Stocks In Indonesia (2022–2024): A Descriptive Analysis," *J. Ilmu Manaj. JIMMU*, vol. 10, no. 2, pp. 180–191, Oct. 2025, doi: 10.33474/jimmu.v10i2.24199.



- [16] R. N. Rahmawati and S. Agustin, "Pengaruh Variabel Fundamental Keuangan Terhadap Harga Saham Perusahaan Food And Beverages," *J.ilmu dan Riset Manajemen* vol. 8, no. 3, pp 1-24, 2019, .
- [17] Y. Zhang, "Stock Price Prediction Method Based on XGboost Algorithm," in *Proceedings of the 2022 International Conference on Bigdata Blockchain and Economy Management (ICBBEM 2022)*, vol. 5, D. Qiu, Y. Jiao, and W. Yeoh, Eds., in Atlantis Highlights in Intelligent Systems,. , Dordrecht: Atlantis Press International BV, 2023, vol. 5, no. 5, pp. 595–603. doi: 10.2991/978-94-6463-030-5_60.
- [18] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Aug. 2016, pp. 785–794. doi: 10.1145/2939672.2939785.
- [19] F. Winata, M. I. A. Tresnawan, G. Wijaya, and H. W. Aripadono, "Integrasi Feature Engineering dan SMOTE pada Algoritma Random Forest untuk Prediksi Kerusakan Chip RFID di Industri Sel Surya," *J.Information and System Research*, vol. 7, no. 2, pp 599-609, 2026, doi.org/10.47065/josh.v7i2.9038 .
- [20] A. A. R. Reza and Muhammad Syaifur Rohman, "Prediction Stunting Analysis Using Random Forest Algorithm and Random Search Optimization," *J. Inform. Telecommun. Eng.*, vol. 7, no. 2, pp. 534–544, Jan. 2024, doi: 10.31289/jite.v7i2.10628.