

ANALISIS PERBANDINGAN KINERJA MODEL LSTM, XGBOOST, DAN HYBRID LSTM-XGBOOST UNTUK PREDIKSI HARGA SAHAM

Annita Fadhilah Aprilia^{1*}, Wahyu S. J. Saputra², Andri Fauzan Adziima³

^{1,2,3}Program Studi Sains Data, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional “Veteran”
Jawa Timur, Surabaya, Indonesia

* Penulis korespondensi: 22083010033@student.upnjatim.ac.id

ABSTRAK

Pergerakan harga saham PT Bank Rakyat Indonesia (Persero) Tbk yang fluktuatif dan dipengaruhi oleh berbagai faktor menjadikan prediksi harga saham sebagai tantangan yang kompleks dalam analisis pasar keuangan. Penelitian ini bertujuan untuk membandingkan kinerja model *Long Short-Term Memory* (LSTM), *Extreme Gradient Boosting* (XGBoost), dan model *hybrid* LSTM-XGBoost dalam memprediksi harga penutupan saham harian berdasarkan data historis selama sepuluh tahun (2015–2025). Model LSTM digunakan untuk menangkap pola temporal dan ketergantungan jangka panjang pada data deret waktu, sedangkan XGBoost diterapkan untuk memodelkan hubungan non-linear serta memperbaiki kesalahan prediksi yang tidak tertangkap LSTM. Pada pendekatan *hybrid*, prediksi awal LSTM dikoreksi oleh XGBoost melalui pemodelan residual sehingga menghasilkan prediksi yang lebih akurat dan stabil. Evaluasi kinerja dilakukan menggunakan metrik *Root Mean Squared Error* (RMSE), *Mean Absolute Error* (MAE), dan koefisien determinasi (R^2). Hasil pengujian menunjukkan bahwa model *hybrid* LSTM-XGBoost memberikan performa terbaik dengan nilai RMSE sebesar 193,98, MAE sebesar 159,56, dan R^2 sebesar 0,93, lebih unggul dibandingkan model LSTM (RMSE=200,64; MAE=164,92; R^2 =0,928) maupun XGBoost (RMSE=586,18; MAE=395,99; R^2 =0,365). Prediksi 7 hari ke depan menunjukkan tren kenaikan harga secara bertahap dari Rp 3.831,83 hingga Rp 3.934,18 dengan fluktuasi yang wajar, yang dapat menjadi acuan bagi investor dalam pengambilan keputusan jangka pendek. Penelitian ini menunjukkan bahwa pendekatan *hybrid* mampu memanfaatkan keunggulan masing-masing model untuk meningkatkan akurasi prediksi harga saham pada data deret waktu yang dinamis dan kompleks.

Kata kunci: *Long Short Term-Memory* (LSTM), *Extreme Gradient Boosting* (XGBoost), *Hybrid* Model, Prediksi Harga Saham, PT Bank Rakyat Indonesia (Persero) Tbk, *Time Series*.

1 PENDAHULUAN

Investasi saham merupakan salah satu instrumen keuangan yang banyak diminati karena menawarkan potensi keuntungan yang relatif tinggi dibandingkan instrumen investasi lainnya. Karena menawarkan potensi imbal hasil yang tinggi dibandingkan dengan opsi investasi lainnya, investasi saham menjadi salah satu instrumen keuangan paling populer. Akan tetapi harga saham dipengaruhi oleh banyak hal, seperti makroekonomi, kebijakan pemerintah, kinerja perusahaan, dan perasaan pasar. Karena ketidakpastian ini, sulit untuk memprediksi harga saham secara akurat dengan metode konvensional. Dengan volume transaksi yang tinggi, saham PT Bank Rakyat Indonesia (Persero) Tbk yang diperdagangkan di Bursa Efek Indonesia juga mengalami kondisi ini. Pendekatan prediksi yang dapat menangkap pola historis sambil beradaptasi dengan perubahan pasar diperlukan karena fluktuasi harga saham BRI mencerminkan karakteristik data seri waktu yang kompleks, non-linear, dan berfluktuasi tinggi.

Kemampuan untuk memprediksi pergerakan harga saham sangat penting bagi investor dan analis karena membantu mereka membuat keputusan investasi yang lebih rasional dan berbasis

data. Apabila salah memprediksi pergerakan harga saham, risikonya akan kehilangan banyak uang, terutama jika memiliki saham dengan banyak likuiditas seperti saham perbankan. Oleh karena itu, sangat penting untuk membuat model prediksi harga saham yang akurat dan tahan lama yang dapat beradaptasi dengan perubahan pasar. Dengan kemajuan dalam teknologi komputasi, metode *machine learning* dan *deep learning* telah menjadi alternatif yang lebih fleksibel daripada metode statistik tradisional, yang biasanya memiliki ketergantungan waktu dan keterbatasan dalam menangkap pola nonlinier.

Salah satu model *deep learning* yang banyak digunakan dalam pemodelan data deret waktu keuangan adalah *Long Short-Term Memory* (LSTM) yang diperkenalkan oleh Hochreiter dan Schmidhuber untuk mengatasi masalah ketergantungan jangka panjang dalam Jaringan Saraf Rekuren. LSTM memungkinkan model menyimpan informasi penting selama periode waktu tertentu, yang menjadikannya model yang sangat efektif dalam mempelajari pola temporal jangka pendek dan jangka panjang dalam pemodelan data *time series* keuangan. Penelitian oleh (Fiqih Nur Iman, 2023) menunjukkan bahwa LSTM mampu memberikan performa prediksi yang baik pada saham perbankan Indonesia. Hasil serupa juga dilaporkan oleh (Rosyd et al., 2024) pada saham Bank BRI dan Bank BCA. Meskipun demikian, penelitian-penelitian tersebut mencatat bahwa performa LSTM dapat menurun ketika data mengandung noise tinggi dan volatilitas ekstrem, serta memerlukan proses pelatihan yang relatif kompleks.

Selain LSTM, algoritma *Extreme Gradient Boosting* (XGBoost) juga banyak digunakan dalam prediksi harga saham. XGBoost merupakan algoritma *machine learning* berbasis *ensemble* pohon keputusan yang unggul dalam menangkap hubungan nonlinier antarvariabel dan relatif stabil terhadap data yang kompleks. (Pratama & Banowosari, 2024) menunjukkan bahwa XGBoost mampu menghasilkan akurasi prediksi yang baik pada saham sektor perbankan Indonesia dengan tingkat kesalahan yang relatif kecil. (Nurmaya et al., 2023) juga melaporkan bahwa XGBoost memiliki efisiensi komputasi yang baik serta mekanisme regularisasi yang efektif dalam mengurangi *overfitting*. Namun, XGBoost tidak secara eksplisit memodelkan ketergantungan temporal, sehingga kurang optimal dalam menangkap pola urutan waktu jangka panjang yang menjadi karakter utama pergerakan harga saham.

Perbedaan karakteristik antara LSTM dan XGBoost menunjukkan bahwa masing-masing model memiliki keunggulan dan keterbatasan yang saling melengkapi. LSTM unggul dalam menangkap pola temporal jangka panjang, tetapi sensitif terhadap *noise*, sedangkan XGBoost kuat dalam memodelkan hubungan nonlinier dan koreksi kesalahan, namun kurang optimal dalam mempelajari dinamika waktu. Oleh karena itu, sejumlah penelitian mengusulkan pendekatan *hybrid* dengan mengombinasikan LSTM dan XGBoost untuk memanfaatkan keunggulan masing-masing model. (Shi et al., 2021) menunjukkan bahwa model *hybrid* mampu meningkatkan akurasi prediksi dibandingkan model tunggal. (Ajder et al., 2025) serta (Dolon, 2025) juga melaporkan bahwa *hybrid* LSTM–XGBoost lebih stabil dalam menghadapi data yang volatil dan mengandung pola tidak terduga.

Meskipun model *hybrid* LSTM–XGBoost menunjukkan hasil yang signifikan, sebagian besar penelitian sebelumnya berfokus pada penerapan model tertentu tanpa melakukan perbandingan kinerja yang komprehensif antara LSTM, XGBoost, dan model *hybrid* pada objek penelitian yang sama. Akibatnya, belum terdapat kesimpulan yang jelas mengenai sejauh mana penggabungan LSTM dan XGBoost benar-benar memberikan peningkatan performa dibandingkan penggunaan masing-masing model secara terpisah, khususnya pada saham sektor perbankan Indonesia.

Berdasarkan hal tersebut, penelitian ini bertujuan untuk melakukan analisis perbandingan kinerja model LSTM, XGBoost, dan *hybrid* LSTM–XGBoost dalam memprediksi harga saham PT Bank Rakyat Indonesia (Persero) Tbk. Melalui pendekatan *comparative study*, penelitian ini diharapkan dapat memberikan gambaran yang jelas mengenai keunggulan dan keterbatasan masing-masing model, serta membuktikan apakah penggabungan LSTM dan XGBoost mampu menghasilkan performa prediksi yang lebih baik dibandingkan model tunggal.

2 METODE

2.1 Jenis Penelitian dan Sumber Data

Jenis penelitian yang digunakan dalam penelitian ini adalah penelitian kuantitatif dengan pendekatan deskriptif. Pendekatan ini dipilih karena penelitian berfokus pada pengolahan dan analisis data numerik berupa harga saham untuk menggambarkan serta membandingkan kinerja beberapa model prediksi. Data yang digunakan merupakan data sekunder yang diperoleh dari situs Investing.com, yang menyediakan data historis harga saham secara lengkap dan terstruktur (Investing.com, 2025). Objek penelitian dalam studi ini adalah saham PT Bank Rakyat Indonesia (Persero) Tbk yang tercatat di Bursa Efek Indonesia. Data yang digunakan berupa data harga penutupan saham dengan periode waktu mulai tahun 2015 hingga tahun 2025, sehingga mencakup rentang waktu selama sepuluh tahun. Penggunaan data jangka panjang bertujuan untuk menangkap pola pergerakan harga saham secara lebih komprehensif, termasuk tren jangka panjang, fluktuasi musiman, serta dinamika pasar pada berbagai kondisi ekonomi. Data harga penutupan dipilih karena mencerminkan kesepakatan akhir pasar pada setiap periode perdagangan dan sering digunakan sebagai acuan dalam analisis prediksi harga saham.

Data yang digunakan selanjutnya diolah melalui tahap praproses, meliputi pemeriksaan kelengkapan data, penanganan nilai hilang, serta penyesuaian format sehingga sesuai dengan kebutuhan model LSTM, XGBoost, dan *Hybrid* LSTM–XGBoost. Teknik pengambilan sampel yang digunakan adalah *purposive sampling*, dengan pemilihan saham PT Bank Rakyat Indonesia (Persero) Tbk berdasarkan pertimbangan likuiditas yang tinggi, kapitalisasi pasar yang besar, serta perannya sebagai salah satu saham perbankan utama di Bursa Efek Indonesia (Putu & Kusuma, 2024). Karakteristik tersebut menjadikan saham BRI representatif untuk menguji dan membandingkan kinerja prediksi harga saham.

Adapun kriteria data yang digunakan dalam penelitian ini adalah sebagai berikut:

- Saham yang diteliti merupakan saham PT Bank Rakyat Indonesia (Persero) Tbk yang tercatat di Bursa Efek Indonesia selama periode penelitian tahun 2015–2025.
- Saham memiliki data harga penutupan yang tersedia secara lengkap pada periode penelitian.
- Data tidak mengalami perhentian perdagangan yang signifikan dalam periode pengamatan.
- Data harga saham dapat diakses secara konsisten melalui situs Investing.com.

2.2 Metode Analisis Data

Metode analisis data dalam penelitian ini dilakukan dengan pendekatan komputasional berbasis *machine learning* dan *deep learning* untuk memprediksi harga saham PT Bank Rakyat Indonesia (Persero) Tbk. Analisis difokuskan pada perbandingan kinerja tiga model, yaitu *Long Short-Term Memory* (LSTM), *Extreme Gradient Boosting* (XGBoost), dan *hybrid* LSTM–XGBoost. Seluruh tahapan analisis disusun secara berurutan sesuai dengan proses pemodelan yang diterapkan pada kode program.

- Normalisasi Data (*Min-Max Scaling*)

Sebelum dilakukan pemodelan, data harga saham dinormalisasi menggunakan metode *Min–Max Scaling* agar seluruh fitur berada pada rentang nilai yang sama. Proses ini bertujuan untuk mempercepat konvergensi model dan mencegah dominasi fitur tertentu dalam proses pelatihan, khususnya pada model LSTM.

Normalisasi Min–Max dirumuskan sebagai berikut (Zuzzaifa & Sancoko, 2025):

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (1)$$

Keterangan:

X' : nilai hasil normalisasi

X : nilai asli data

X_{min} : nilai minimum data

X_{max} : nilai maksimum data

b. Sequence Windowing

Data saham disusun dalam bentuk deret waktu menggunakan pendekatan *sliding window* dengan panjang urutan (*sequence length*) sebanyak 30 periode. Setiap urutan data digunakan untuk memprediksi harga penutupan pada satu periode berikutnya ($t + 1$). Secara matematis, proses ini dapat dinyatakan sebagai (Iskandar et al., 2025):

$$\begin{aligned} X_t &= \{x_{t-n}, x_{t-n+1}, \dots, x_{t-1}\} \\ y_t &= x_t \end{aligned} \quad (2)$$

Keterangan:

n : panjang *sequence*

X_t : data input deret waktu

y_t : target prediksi harga saham

c. Long Short-Term Memory (LSTM)

Model LSTM digunakan untuk menangkap pola temporal jangka pendek dan jangka panjang pada data harga saham. LSTM memiliki tiga mekanisme *gate* utama, yaitu *forget gate*, *input gate*, dan *output gate*. Persamaan utama LSTM adalah sebagai berikut (Fatimah et al., 2025):

Forget Gate

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (3)$$

Input Gate

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (4)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (5)$$

Cell State Update

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \quad (6)$$

Output Gate

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (7)$$

$$h_t = o_t \cdot \tanh(C_t) \quad (8)$$

d. *Extreme Gradient Boosting* (XGBoost)

Model XGBoost digunakan sebagai pembanding kedua dengan pendekatan *machine learning* berbasis *ensemble* pohon keputusan. XGBoost meminimalkan fungsi objektif yang terdiri dari loss function dan regularisasi.

Fungsi objektif XGBoost dirumuskan sebagai (Chen & Guestrin, 2016):

Objective Function XGBoost

$$obj(\theta) = \sum_{i=1}^n L(y_i, \hat{y}_i^t) + \sum_{k=1}^t R(f_k) \quad (9)$$

Keterangan:

$L(y_i, \hat{y}_i^t)$: fungsi loss I untuk menghitung selisih antara nilai aktual (y_i) dan prediksi ($\hat{y}_i^t$) pada iterasi ke- t .

$R(f_k)$: fungsi regularisasi yang mengontrol kompleksitas model untuk menghindari *overfitting*.

n : jumlah data

Fungsi regularisasi XGBoost

$$R(f_k) = \gamma^T + \frac{1}{2}\lambda \sum_{j=1}^t \omega_j^2 \quad (10)$$

Keterangan:

λ : Regularisasi pada bobot model

ω_j : Bobot dari fitur ke- j

e. Model *Hybrid* LSTM-XGBoost

Model *hybrid* dibangun dengan mengombinasikan keunggulan LSTM dan XGBoost. LSTM digunakan untuk memprediksi harga saham dan mengekstraksi fitur temporal, sedangkan XGBoost digunakan untuk memodelkan residual error dari LSTM.

Residual dihitung menggunakan persamaan (Diaz et al., 2025):

$$e_t = y_t - \hat{y}_t^{LSTM} \quad (11)$$

Residual tersebut kemudian dipelajari oleh XGBoost, dan hasil prediksi akhir model *hybrid* dirumuskan sebagai:

$$\hat{y}_t^{Hybrid} = \hat{y}_t^{LSTM} + \alpha \cdot \hat{e}_t^{XGBoost} \quad (12)$$

Keterangan:

α : faktor penyesuaian residual

$\hat{e}_t^{XGBoost}$: prediksi residual oleh XGBoost

f. Evaluasi Kinerja Model

Kinerja model LSTM, XGBoost, dan *hybrid* LSTM–XGBoost dievaluasi menggunakan tiga metrik evaluasi regresi yang umum digunakan dalam prediksi harga saham (Alkayes & Sugihartono, 2025).

Root Mean Squared Error (RMSE)

$$MSE = \sqrt{\sum \frac{(y_t - \hat{y}_t)^2}{n}} \quad (13)$$

Keterangan :

y_t : Nilai aktual

$\hat{y}_t$: Nilai prediksi

n : Jumlah data

Mean Absolute Error (MAE)

$$MAE = \sum \frac{|y_t - \hat{y}_t|}{n} \quad (14)$$

Keterangan:

y_t : Nilai aktual.

$\hat{y}_t$: Nilai prediksi.

n : Jumlah data.

Koefisien Determinasi (R^2)

$$R^2 = 1 - \sum \frac{(y_t - \hat{y}_t)^2}{(y_t - \bar{y}_t)^2} \quad (15)$$

2.3 Teknik Analisis Data

2.3.1 Model LSTM

Prosedur analisis data menggunakan model *Long Short-Term Memory* (LSTM) adalah sebagai berikut :

- Melakukan praproses data harga saham dan menyusunnya berdasarkan urutan waktu.
- Membentuk fitur tambahan berupa *moving average* dan volatilitas harga.
- Menyusun data ke dalam bentuk deret waktu menggunakan metode *sliding window*.
- Membagi data secara berurutan menjadi data pelatihan, validasi, dan pengujian.
- Melakukan normalisasi data menggunakan metode *Min-Max Scaling*.
- Melatih model LSTM dengan fungsi loss Huber dan *optimizer* Adam.
- Menghasilkan prediksi harga saham pada data pengujian.

2.3.2 Model XGBoost

Model XGBoost digunakan sebagai model pembanding kedua dengan pendekatan *machine learning* berbasis *ensemble* pohon keputusan. Prosedur analisis data menggunakan model XGBoost dilakukan sebagai berikut:

- Menggunakan data fitur numerik hasil praproses tanpa pembentukan *sequence*.
- Melakukan normalisasi dan pembagian data yang sama dengan model LSTM.
- Melatih model XGBoost menggunakan fungsi objektif regresi dengan regularisasi.
- Menghasilkan prediksi harga saham pada data pengujian sebagai model pembanding.

2.3.3 Model Hybrid LSTM-XGBoost

Model *hybrid* LSTM–XGBoost dikembangkan untuk menggabungkan keunggulan LSTM dalam menangkap pola temporal dan kemampuan XGBoost dalam memodelkan hubungan nonlinier serta memperbaiki kesalahan prediksi. Tahapan analisis model hybrid adalah sebagai berikut:

- Mengekstraksi fitur laten dari lapisan tersembunyi model LSTM.
- Menghitung residual dari selisih nilai aktual dan prediksi LSTM.
- Melatih model XGBoost untuk mempelajari pola kesalahan prediksi (residual).
- Menggabungkan prediksi LSTM dan koreksi residual XGBoost untuk menghasilkan prediksi akhir.

2.3.4 Evaluasi Model

Evaluasi kinerja model dilakukan untuk menilai dan membandingkan performa prediksi dari model LSTM, XGBoost, dan *Hybrid LSTM–XGBoost*. Prosedur evaluasi meliputi:

- Menghitung *Root Mean Squared Error* (RMSE) untuk mengukur besar kesalahan prediksi rata-rata dengan memberikan penalti lebih besar pada kesalahan yang ekstrem.
- Menghitung *Mean Absolute Error* (MAE) untuk mengukur rata-rata selisih absolut antara nilai aktual dan nilai prediksi.
- Menghitung koefisien determinasi (R^2) untuk mengukur kemampuan model dalam menjelaskan variasi data harga saham.

2.3.5 Prediksi harga saham Masa Depan

Setelah model *hybrid LSTM–XGBoost* menunjukkan performa terbaik berdasarkan hasil evaluasi, model tersebut digunakan untuk melakukan prediksi harga saham pada periode mendatang. Prosedur prediksi harga saham masa depan dilakukan sebagai berikut:

- Penggunaan *sequence* data terakhir dari data historis digunakan sebagai input awal untuk menghasilkan prediksi ke depan.
- Prediksi bertahap (*rolling forecast*) pada prediksi saham untuk beberapa hari ke depan, di mana hasil prediksi sebelumnya digunakan sebagai bagian dari input pada langkah selanjutnya.
- Visualisasi hasil prediksi harga saham masa depan yang divisualisasikan dalam bentuk grafik untuk memudahkan interpretasi tren pergerakan harga saham.

3 HASIL DAN PEMBAHASAN

3.1 Load Data

Data saham BBRI dari tahun 2015 hingga 2025 dimuat dari file CSV yang berisi tanggal (Date), harga pembukaan (Open), harga tertinggi (High), harga terendah (Low), harga penutupan (Close), dan volume perdagangan (Volume). Kolom Volume dibersihkan dari format teks seperti “K”, “M”, atau “B” menjadi angka numerik agar bisa diproses model. Selain itu, dibuat variabel tambahan target yang merupakan harga penutupan hari berikutnya, sehingga setiap baris data memiliki informasi harga saat ini dan harga yang akan datang untuk keperluan prediksi.

Dataset kemudian dibagi menjadi tiga bagian yaitu train, validation, dan test. Sekitar 10% data dari bagian pelatihan dijadikan data validasi untuk memantau performa model selama pelatihan, sedangkan 20% data terakhir (tahun 2023–2025) digunakan sebagai data uji agar model diuji pada kondisi pasar terbaru. Dengan pembagian ini, model dapat belajar dari data historis, memvalidasi performanya, dan diuji pada data nyata yang relevan, menggunakan variabel asli (Open, High, Low, Close, Volume) sebagai input dan variabel target sebagai fokus prediksi.

Tabel 1. Dataset BBRI

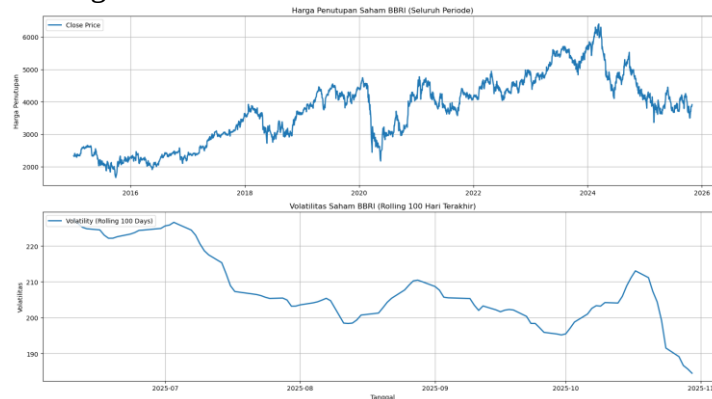
Date	Open	High	Low	Close	Volume	Target
2015-01-02	2305	2345	2305	2330	45160000.0	2320.0
2015-01-05	2320	2340	2315	2320	49570000.0	2305.0
2015-01-06	2310	2325	2300	2305	65090000.0	2355.0
2015-01-07	2320	2355	2315	2355	146160000.0	2395.0
2015-01-08	2360	2395	2360	2395	146660000.0	2405.0
...	...	...	...	...	...	...
2025-10-24	3820	3910	3820	3850	313900000.0	3860.0

Date	Open	High	Low	Close	Volume	Target
2025-10-27	3880	3930	3750	3860	332540000.0	3850.0
2025-10-28	3850	3860	3780	3850	239030000.0	3890.0
2025-10-29	3830	3890	3790	3890	215010000.0	3910.0
2025-10-30	3900	3990	3870	3810	287910000.0	3980.0

Sumber: (Investing.com, 2025)

3.2 Eksplorasi Data

Tahap eksplorasi data dilakukan untuk memahami karakteristik, pola, serta dinamika pergerakan harga saham BBRI sebelum proses pemodelan dilakukan. Visualisasi data digunakan sebagai alat utama untuk mengidentifikasi tren, fluktuasi, dan tingkat risiko yang terkandung dalam data deret waktu harga saham.



Gambar 1. Grafik Harga Penutupan dan Volatilitas Saham BBRI

Berdasarkan grafik harga penutupan, secara umum pergerakan harga saham BBRI menunjukkan pola fluktuatif dengan kecenderungan tren meningkat dari tahun 2020 hingga awal tahun 2024. Penurunan harga yang cukup signifikan pada awal tahun 2020 diduga kuat dipengaruhi oleh ketidakstabilan ekonomi global akibat pandemi COVID-19 yang berdampak pada pasar keuangan nasional maupun internasional. Selanjutnya, harga saham BBRI mencapai titik tertinggi pada tahun 2024 sebelum mengalami tren penurunan yang cukup tajam hingga tahun 2025. Penurunan ini diperkirakan dipengaruhi oleh berbagai faktor eksternal, seperti kondisi politik nasional, implementasi kebijakan baru pemerintah, perubahan kebijakan perdagangan Amerika Serikat, serta faktor makroekonomi lainnya yang berkontribusi terhadap meningkatnya ketidakpastian pasar. Grafik volatilitas saham menggambarkan tingkat fluktuasi harga saham dalam periode 100 hari terakhir. Peningkatan nilai volatilitas pada awal tahun 2025 menunjukkan bahwa pergerakan harga saham menjadi semakin tidak stabil, yang mengindikasikan meningkatnya risiko investasi pada periode tersebut. Sebaliknya, periode dengan volatilitas yang lebih rendah mencerminkan kondisi pasar yang relatif lebih stabil.

3.3 Evaluasi Model Prediksi

3.3.1 Model LSTM

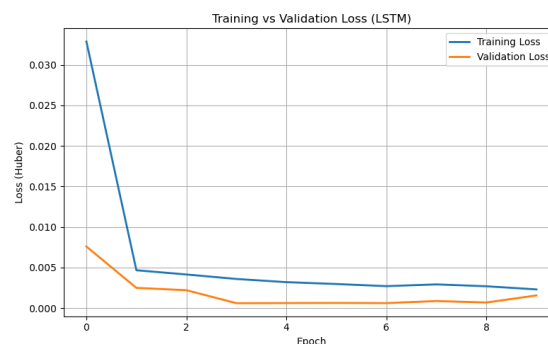
Model *Long Short-Term Memory* (LSTM) pada penelitian ini digunakan untuk mempelajari pola temporal dan ketergantungan jangka pendek hingga menengah pada data harga saham BBRI. Model ini berperan sebagai *feature extractor* utama yang bertugas menangkap dinamika pergerakan harga saham sebelum digunakan pada tahap pemodelan selanjutnya. Arsitektur dan

parameter model LSTM yang digunakan dalam penelitian ini ditampilkan pada Tabel 2 sebagai berikut.

Tabel 2. Parameter Model LSTM

Parameter	Nilai
<i>Batch Size</i>	32
<i>Epochs (maksimum)</i>	40
<i>Neurons (LSTM)</i>	48
<i>Dropout</i>	0,2
<i>Dense Layer</i>	16 Neuron (ReLU)
<i>Output Layer</i>	1 Neuron
<i>Optimizer</i>	Adam
<i>Learning Rate Awal</i>	0,001
<i>Loss Function</i>	Huber Loss
<i>Early Stopping</i>	Patience = 6
<i>ReduceLROnPlateau</i>	Factor = 0,5; Patience = 3

Berdasarkan Tabel 2, arsitektur model terdiri dari satu lapisan LSTM dengan 48 neuron, yang diikuti oleh Dropout sebesar 0,2 untuk mengurangi risiko *overfitting*. Keluaran dari lapisan LSTM selanjutnya diproses oleh Dense layer dengan 16 neuron dan fungsi aktivasi ReLU sebagai representasi fitur laten, serta satu neuron pada lapisan keluaran yang menghasilkan nilai prediksi harga saham. Proses pelatihan model dilakukan menggunakan *optimizer* Adam dengan *learning rate* awal sebesar 0,001 dan fungsi kerugian Huber loss. Pemilihan Huber loss bertujuan untuk meningkatkan ketahanan model terhadap keberadaan *outlier* pada data harga saham. Model dilatih menggunakan batch size 32 dengan jumlah maksimum 40 epoch, serta dilengkapi dengan mekanisme *Early Stopping* dengan *patience* 6 epoch untuk menghentikan pelatihan secara otomatis ketika performa pada data validasi tidak lagi mengalami perbaikan. Selain itu, digunakan pula ReduceLROnPlateau untuk menurunkan *learning rate* secara adaptif ketika nilai *validation loss* mengalami stagnasi.



Gambar 2. Grafik Training vs Validation Loss

Berdasarkan grafik *training loss* dan *validation loss* pada Gambar 2, terlihat bahwa nilai loss mengalami penurunan yang signifikan pada epoch awal, yang menunjukkan proses pembelajaran model berlangsung dengan cepat. Pada epoch pertama, *training loss* tercatat sebesar 0,0329 dan *validation loss* sebesar 0,0076. Selanjutnya, kedua nilai loss menurun secara bertahap hingga mencapai kondisi stabil dengan *validation loss* minimum pada kisaran 6×10^{-4} . Selama proses pelatihan, selisih antara *training loss* dan *validation loss* relatif kecil dan tidak menunjukkan pola divergensi yang signifikan. Hal ini mengindikasikan bahwa model LSTM tidak mengalami

overfitting dan memiliki kemampuan generalisasi yang baik terhadap data yang tidak digunakan dalam pelatihan. Penyesuaian *learning rate* yang terjadi secara otomatis pada beberapa epoch turut berkontribusi dalam menjaga stabilitas dan konvergensi model.



Gambar 3. Grafik Prediksi dengan LSTM

Hasil grafik prediksi harga saham BBRI.JK menggunakan model *Long Short-Term Memory* (LSTM) pada Gambar 3 menunjukkan bahwa model mampu mengikuti pola dan tren pergerakan harga saham secara umum. Kurva hasil prediksi terlihat bergerak searah dengan harga aktual, baik pada fase kenaikan maupun penurunan, meskipun pada beberapa periode tertentu masih terdapat deviasi antara nilai prediksi dan nilai aktual. Deviasi tersebut umumnya muncul pada saat terjadi perubahan harga yang relatif tajam, yang mencerminkan dinamika pasar saham yang bersifat fluktuatif. Evaluasi kinerja model berdasarkan metrik kesalahan menunjukkan bahwa model LSTM menghasilkan nilai RMSE sebesar 200,64, Mean Absolute Error (MAE) sebesar 164,92, dan koefisien determinasi (R^2) sebesar 0,92. Nilai RMSE dan MAE yang relatif rendah menunjukkan bahwa selisih antara harga prediksi dan harga aktual berada pada kisaran yang masih dapat diterima. Nilai R^2 sebesar 0,928 mengindikasikan bahwa variasi data harga saham aktual dapat dijelaskan oleh model LSTM, sehingga menunjukkan kemampuan prediksi yang sangat baik. Dengan nilai R^2 yang mendekati 1 serta kesalahan prediksi yang relatif kecil, model LSTM memiliki tingkat akurasi yang tinggi dalam memodelkan pola pergerakan harga saham.

3.3.2 Model XGBoost

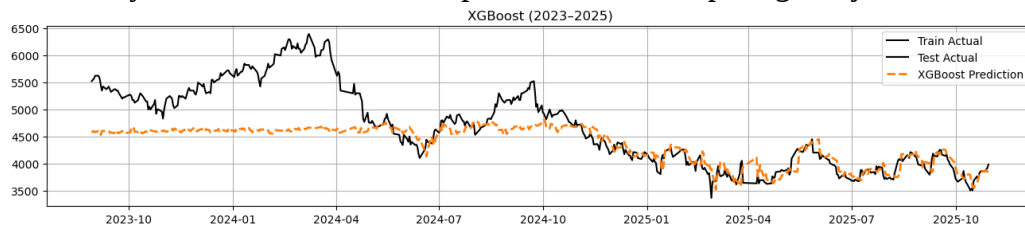
Model *Extreme Gradient Boosting* (XGBoost) pada penelitian ini digunakan sebagai baseline model untuk membandingkan performa prediksi dengan model berbasis sekuensial. Konfigurasi parameter model disajikan secara ringkas pada Tabel berikut.

Tabel 3. Parameter Model XGBoost

Parameter	Nilai
Model	XGBoost Regressor
Jumlah Estimator	200
Learning Rate	0,1
Max Depth	2
Subsample	0,7
Colsample by Tree	0,7
Objective	Regression
Input Data	Flattened time series features
Sample Weight	Digunakan
Random State	SEED

Berdasarkan implementasi pada kode, model XGBoost dikonfigurasi dengan 200 estimator, learning rate sebesar 0,1, dan max depth sebesar 2, serta parameter subsample dan colsample_bytree masing-masing sebesar 0,7. Pengaturan ini bertujuan untuk mengurangi risiko *overfitting* dengan membatasi kompleksitas pohon keputusan serta meningkatkan kemampuan

generalisasi model. Selain itu, proses pelatihan dilakukan dengan mempertimbangkan sample weight untuk menyesuaikan kontribusi setiap observasi terhadap fungsi objektif.



Gambar 4. Grafik Prediksi dengan XGBoost

Hasil evaluasi prediksi Gambar 4, model XGBoost menunjukkan bahwa kurva prediksi XGBoost cenderung lebih halus dan kurang mampu mengikuti fluktuasi harga saham yang bersifat dinamis. Model ini mengalami kesulitan dalam menangkap perubahan tren jangka pendek, terutama pada periode dengan volatilitas tinggi. Hal ini mengindikasikan keterbatasan XGBoost dalam memodelkan pola sekuensial ketika data deret waktu direpresentasikan dalam bentuk fitur statis. Evaluasi kinerja model XGBoost berdasarkan metrik kesalahan menunjukkan nilai RMSE sebesar 586,18, MAE sebesar 395,99, dan koefisien determinasi (R^2) sebesar 0,36.5. Nilai RMSE dan MAE yang relatif besar mengindikasikan adanya deviasi yang cukup signifikan antara harga prediksi dan harga aktual. Sementara itu, nilai R^2 yang rendah menunjukkan bahwa model XGBoost hanya mampu menjelaskan variasi data harga saham, sehingga performanya masih terbatas dibandingkan dengan model berbasis sekuensial seperti LSTM.

3.3.3 Hybrid Model LSTM-XGBoost

Pada penelitian ini, dikembangkan model *hybrid* LSTM–XGBoost untuk meningkatkan akurasi prediksi harga saham BBRI dengan mengombinasikan keunggulan model sekuensial dan model ensemble berbasis pohon. Model LSTM berperan sebagai *feature extractor* dan prediktor awal yang menangkap pola temporal data deret waktu, sedangkan XGBoost digunakan untuk memodelkan *residual error* yang tidak dapat ditangkap oleh LSTM. Pendekatan ini memungkinkan XGBoost fokus pada pola nonlinier dan kompleks yang tersisa setelah prediksi LSTM, sehingga menghasilkan prediksi akhir yang lebih akurat dan stabil.



Gambar 5. Grafik Prediksi dengan LSTM-XGBoost

Berdasarkan visualisasi prediksi pada Gambar 5, model *hybrid* LSTM–XGBoost menunjukkan performa yang paling stabil dan konsisten dalam mengikuti tren pergerakan harga saham BBRI dibandingkan model LSTM dan XGBoost tunggal. Kurva prediksi *hybrid* mampu mengikuti arah tren utama dengan deviasi yang lebih kecil, menandakan kemampuannya dalam mengurangi dampak fluktuasi ekstrem tanpa kehilangan sensitivitas terhadap perubahan harga yang signifikan. Model *hybrid* ini mengombinasikan kemampuan LSTM dalam menangkap pola temporal dan dependensi sekuensial dengan XGBoost yang berperan dalam memodelkan residual

kesalahan untuk menangkap hubungan nonlinier yang belum terakomodasi. Pendekatan ini membantu model memperbaiki kesalahan prediksi sehingga hasil prediksi menjadi lebih akurat. Hasil evaluasi menunjukkan bahwa model *hybrid* memperoleh nilai RMSE sebesar 193,98, MAE sebesar 159,56, dan koefisien determinasi (R^2) sebesar 0,93. Nilai kesalahan yang lebih rendah serta R^2 yang tinggi mengindikasikan bahwa model *hybrid* mampu menjelaskan variasi harga saham BBRI dengan baik dan memiliki kemampuan generalisasi yang tinggi. Secara keseluruhan, hasil ini menegaskan bahwa model *hybrid* LSTM–XGBoost berhasil menggabungkan keunggulan LSTM dalam memodelkan pola temporal dengan kemampuan XGBoost dalam menangkap hubungan non-linear pada residual kesalahan. Kombinasi tersebut menghasilkan prediksi yang lebih stabil dan akurat dibandingkan model tunggal, baik pada data historis (*training*) maupun data uji (*testing*). Oleh karena itu, model *hybrid* ini dinilai paling optimal dan layak dijadikan pendekatan utama dalam prediksi harga saham BBRI pada penelitian ini.

3.3.4 Perbandingan Kinerja Model

Tabel berikut menyajikan perbandingan kinerja tiga model yang digunakan dalam prediksi harga saham BBRI, yaitu LSTM, XGBoost, dan *Hybrid* LSTM–XGBoost, berdasarkan tiga metrik evaluasi yaitu *Root Mean Squared Error* (RMSE), *Mean Absolute Error* (MAE), dan koefisien determinasi (R^2).

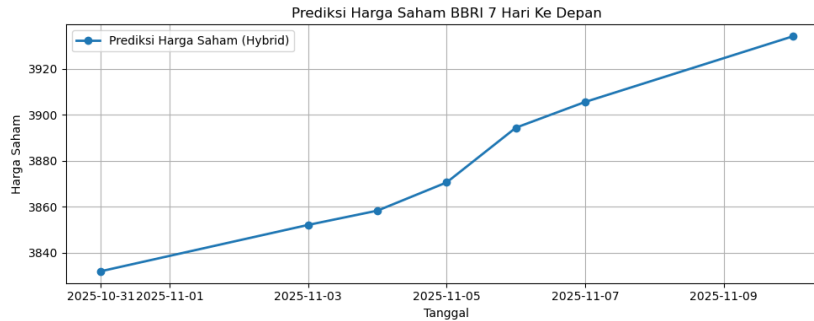
Tabel 4. Perbandingan Metrik Evaluasi Ketiga Model

<i>Model</i>	<i>RMSE</i>	<i>MAE</i>	<i>R²</i>
<i>LSTM</i>	200.64	164.92	0.928
<i>XGBoost</i>	586.18	395.99	0.365
<i>Hybrid LSTM–XGBoost</i>	193.98	159.56	0.930

Hasil evaluasi pada Tabel 4 menunjukkan bahwa model *Hybrid* LSTM–XGBoost memiliki kinerja terbaik dibandingkan model LSTM dan XGBoost tunggal. Model *hybrid* menghasilkan nilai RMSE sebesar 193,98, MAE sebesar 159,56, dan R^2 sebesar 0,93, yang menunjukkan kesalahan prediksi lebih rendah dan kemampuan penjelasan variasi data yang lebih baik. Penurunan RMSE dan MAE dibandingkan model LSTM menandakan bahwa koreksi residual menggunakan XGBoost mampu memperbaiki kesalahan prediksi LSTM. Sebaliknya, model XGBoost tunggal menghasilkan kesalahan paling tinggi dan nilai R^2 yang rendah, sehingga kurang optimal tanpa pemodelan temporal. Secara keseluruhan, penggabungan LSTM dan XGBoost menghasilkan model yang lebih akurat, stabil, dan representatif dalam memprediksi harga saham BBRI.

3.4 Prediksi 7 Hari dengan Model *Hybrid*

Berdasarkan performa terbaik yang ditunjukkan oleh model *hybrid* LSTM–XGBoost, dilakukan proyeksi harga saham BBRI untuk 7 hari ke depan menggunakan data historis terbaru. Prediksi ini divisualisasikan dalam bentuk grafik untuk menunjukkan tren dan potensi fluktuasi harga. Proses prediksi dilakukan dengan memanfaatkan LSTM untuk menangkap pola temporal, kemudian disempurnakan oleh XGBoost melalui koreksi residual dan pemodelan hubungan non-linear, sehingga menghasilkan prediksi yang lebih akurat dan stabil.



Gambar 6. Grafik Prediksi Harga Saham (*Hybrid*)

Kurva pada grafik menunjukkan tren kenaikan harga saham secara bertahap dari hari ke hari. Pada awal periode prediksi, harga diproyeksikan berada di sekitar Rp 3.831,83 pada 31 Oktober 2025, kemudian meningkat secara bertahap hingga mencapai Rp 3.934,18 pada 10 November 2025. Beberapa hari menunjukkan kenaikan yang relatif stabil, sementara hari-hari lainnya memperlihatkan fluktuasi kecil, mencerminkan ketidakpastian pasar. Tren yang meningkat ini terlihat konsisten dengan nilai prediksi harian yang tercantum pada Tabel 5.

Tabel 5. Prediksi Harga Saham BBRI 7 Hari ke Depan

Tanggal	Prediksi Harga (Rp)
2025-10-31	3.831,83
2025-11-03	3.852,04
2025-11-04	3.858,25
2025-11-05	3.870,55
2025-11-06	3.894,39
2025-11-07	3.905,60
2025-11-10	3.934,18

Setiap prediksi harian dihitung dengan memperbarui urutan data (*sequence*) secara bertahap, menggunakan harga terakhir sebagai acuan untuk hari berikutnya. Model juga menghitung rata-rata pergerakan (*moving average*) dan volatilitas harga, sehingga dapat memperkirakan ketidakpastian pasar. Rentang ketidakpastian ini tercermin dari fluktuasi harga yang mungkin terjadi, menandakan periode stabil atau periode dengan risiko perubahan harga yang lebih tinggi. Hasil prediksi menunjukkan tren kenaikan yang relatif konsisten, sehingga memberikan perkiraan harga harian yang dapat dijadikan referensi bagi investor untuk membuat keputusan jangka pendek.

4 KESIMPULAN

Saham PT Bank Rakyat Indonesia (Persero) Tbk menunjukkan tren kenaikan yang relatif konsisten, meskipun pada tahun 2025 volatilitas harga meningkat sehingga pergerakan saham menjadi lebih fluktuatif dibandingkan tahun-tahun sebelumnya. Untuk membantu investor mengambil keputusan yang lebih tepat, prediksi harga saham masa depan menjadi sangat penting. Penelitian ini membuktikan bahwa model *hybrid* LSTM–XGBoost mampu memprediksi harga saham dengan lebih akurat dibandingkan model tunggal. Penggabungan kedua model dilakukan karena LSTM unggul dalam menangkap pola temporal dan ketergantungan jangka panjang pada data deret waktu, sedangkan XGBoost kuat dalam memodelkan hubungan non-linear serta

memperbaiki kesalahan prediksi (residual) yang tidak tertangkap LSTM. Kombinasi ini menghasilkan prediksi yang lebih stabil dan akurat. Hasil evaluasi menunjukkan bahwa model *hybrid* mencapai RMSE sebesar 193,98, MAE sebesar 159,56, dan koefisien determinasi (R^2) sebesar 0,93, yang lebih baik dibandingkan model LSTM (RMSE 200,64; MAE 164,92; R^2 0,928) maupun XGBoost (RMSE 586,18; MAE 395,99; R^2 0,365). Prediksi 7 hari ke depan menunjukkan kenaikan harga secara bertahap dari Rp 3.831,83 hingga Rp 3.934,18, dengan fluktuasi yang wajar, sehingga hasilnya dapat dijadikan acuan bagi investor untuk pengambilan keputusan jangka pendek. Untuk pengembangan lebih lanjut, model *hybrid* ini dapat disempurnakan dengan menambahkan fitur tambahan seperti berita terkini atau sentimen media sosial. Dengan demikian, model tidak hanya mengandalkan pola historis, tetapi juga dapat menyesuaikan prediksi berdasarkan informasi eksternal, sehingga meningkatkan kemampuan responsif terhadap perubahan pasar secara *real-time*.

UCAPAN TERIMA KASIH

Penulis mengucapkan apresiasi dan ucapan terima kasih kepada Universitas Terbuka sebagai penyelenggara Seminar Nasional yang telah memberikan wadah ilmiah bagi diseminasi hasil penelitian ini. Ucapan terima kasih juga disampaikan kepada dosen pembimbing, Wahyu Syaifullah J. S. dan Andri Fauzan Adziima, atas bimbingan, arahan, serta masukan yang diberikan selama proses penyusunan penelitian. Selain itu, penulis mengucapkan terima kasih kepada pihak Investing.com atas ketersediaan data historis saham PT Bank Rakyat Indonesia (Persero) Tbk yang digunakan dalam penelitian ini, serta kepada seluruh pihak yang telah memberikan dukungan dan kontribusi sehingga penelitian ini dapat diselesaikan dengan baik.

DAFTAR PUSTAKA

- Ajder, A., Hamza, H. A. A., & Ayaz, R. (2025). Wavelet-Enhanced Hybrid LSTM-XGBoost Model for Predicting Time Series Containing Unpredictable Events. *IEEE Access*, 13(April), 58671–58679. <https://doi.org/10.1109/ACCESS.2025.3556540>
- Alkayes, A. F., & Sugihartono, T. (2025). *Perbandingan Algoritma XGBoost dan LSTM dalam Prediksi Harga Saham Tesla Menggunakan Data Tahun 2025* Abstrak *Comparison of XGBoost and LSTM Algorithms in Tesla Stock Price Prediction Using 2025 Data Abstract*. 5(6), 1563–1573.
- Chen, T., & Guestrin, C. (2016). *XGBoost: A Scalable Tree Boosting System*. 785–794. <https://doi.org/10.1145/2939672.2939785>
- Diaz, S., Fernanda, M., Jc, C., Henry, C., Sánchez, M., Duque, J. S., Cecilia, E., Dussan, E., & Alejandra, O. F. R. (2025). *A hybrid LSTM-XGBoost model with residual correction for air quality prediction using SSA*.
- Dolon, M. S. A. (2025). *Review of Applied Science and Technology HYBRID MACHINE LEARNING – DRIVEN FINANCIAL FORECASTING MODELS: INTEGRATING LSTM, PROPHET* ., 04(01), 1–34. <https://doi.org/10.63125/nr1j8527>
- Fatimah, A., Farisa, S., & Haviana, C. (2025). *Implementasi Bidirectional Long Short-Term Memory untuk Identifikasi Entitas Saham Implementation of Bidirectional Long Short-Term Memory for Stock Entity Identification*. 11(April), 78–90.
- Fiqih Nur Iman, D. W. (2023). *Prediksi harga saham menggunakan metode long short term memory*. 1(3), 601–616.
- Investing.com. (2025). *Historical stock prices of PT Bank Rakyat Indonesia (Persero) Tbk (BBRI)*. Investing.Com. <https://www.investing.com/>

- Iskandar, M. I., Mudzakir, T. Al, Cahyana, Y., & Pratama, A. R. (2025). *BULLETIN OF COMPUTER SCIENCE RESEARCH Prediksi Pola Pergerakan Saham Adro . Jk Melalui Model LSTM Berbasis Data Historis*. 5(4). <https://doi.org/10.47065/bulletincsr.v5i4.554>
- Nurmaya, S. T., Soesilo, B., & Fatah, D. A. (2023). *PREDIKSI HARGA SAHAM PADA BANK BCA TBK DENGAN TEKNIK XGBOOST (EXTREME GRADIENT BOOSTING)*.
- Pratama, B., & Banowosari, L. Y. (2024). *Comparison of extreme gradient boosting (xgboost) method with long short-term memory (lstm) for stock prediction pt. bank mandiri tbk. (bmri)*. 7, 5631–5636.
- Putu, N., & Kusuma, N. (2024). *PREDIKSI HARGA SAHAM BLUE CHIP PADA INDEKS IDX30 MENGGUNAKAN ALGORITMA RECURRENT NEURAL NETWORK (RNN)*. 23(1), 90–97.
- Rosyd, A., Purnamasari, A. I., & Ali, I. (2024). *PENERAPAN METODE LONG SHORT TERM MEMORY (LSTM) DALAM MEMPREDIKSI HARGA SAHAM PT BANK CENTRAL ASIA*. 8(1), 501–506.
- Shi, Z., Hu, Y., Mo, G., & Wu, J. (2021). *Attention-based CNN-LSTM and XGBoost hybrid model for stock prediction*. 14(8), 1–7.
- Zuzzaifa, N., & Sancoko, S. D. (2025). *Implementasi Algoritme Long Short-Term Memory untuk Prediksi Harga Saham BBKA dan BBRI*. 19(2), 91–97.