

Optimasi Model Prediksi Saham LQ45 Menggunakan Teknik Weighted Ensemble

Samsul Arifin^{*1}, Muhammad Hasan Basri², Yusril Rahman³

^{1,2}Departemen Teknik Informatika, Institut Teknologi Sepuluh Nopember

³Departemen Sistem Informasi, Institut Teknologi Sepuluh Nopember

Email: ¹6025251060@student.its.ac.id, ²6025251045@student.its.ac.id,
³6026241009@student.its.ac.id

Abstrak

Investasi pada indeks LQ45 menawarkan likuiditas tinggi namun memiliki risiko yang signifikan bagi investor pemula. Tantangan utama dalam prediksi harga saham adalah karakteristik data yang non-linear dan dinamis. Penelitian ini mengusulkan sebuah sistem prediksi yang dioptimasi menggunakan model *Hybrid* (CNN-LSTM) dan *Weighted Ensemble Method*. Pendekatan ini mengintegrasikan indikator teknikal seperti *Relative Strength Index* (RSI), *Moving Average* (MA), dan volatilitas harga untuk memperkuat fitur input. Model CNN diimplementasikan untuk mengekstraksi fitur spasial dan memfilter gangguan pasar (noise), sementara LSTM menangkap dependensi temporal jangka panjang. Untuk meningkatkan stabilitas prediksi, metode *Weighted Ensemble* diterapkan dengan menggabungkan model *Machine Learning* (XGBoost dan Random Forest) serta model *Deep Learning*. Pengujian dilakukan pada data saham periode 2020 hingga Januari 2026 dengan total 55.816 data. Hasil eksperimen menunjukkan bahwa model *Weighted Ensemble* yang diusulkan memberikan performa terbaik dengan nilai *Mean Absolute Percentage Error* (MAPE) sebesar 3,52% dan *Root Mean Square Error* (RMSE) sebesar 0,0270. Selain itu, sistem ini diintegrasikan dengan rekomendasi sinyal (Buy/Hold) menggunakan ambang batas kenaikan 1% yang diuji untuk memprediksi harga pada periode 1-10 Februari 2026. Penelitian ini memberikan kontribusi praktis sebagai sistem pendukung keputusan yang akurat dan *robust* dalam meminimalkan risiko investasi di pasar modal Indonesia.

Kata kunci: *hybrid CNN-LSTM, investor pemula, machine learning, saham LQ45, weighted ensemble*

Optimization of LQ45 Stock Prediction Models Using the Weighted Ensemble Technique

Abstract

Investing in the LQ45 index offers high liquidity but carries significant risks for novice investors. The primary challenge in stock price prediction lies in the non-linear and dynamic characteristics of the data. This research proposes an optimized prediction system utilizing a Hybrid (CNN-LSTM) model and a Weighted Ensemble Method. This approach integrates technical indicators such as the Relative Strength Index (RSI), Moving Average (MA), and price volatility to strengthen the input features. The CNN model is implemented to extract spatial features and filter market noise, while the LSTM captures long-term temporal dependencies. To enhance prediction stability, the Weighted Ensemble method is applied by combining Machine Learning models (XGBoost and Random Forest) with Deep Learning models. Testing was conducted on stock data spanning from 2020 to January 2026, comprising a total of 55,816 data points. Experimental results indicate that the proposed Weighted Ensemble model delivers the best performance, yielding a Mean Absolute Percentage Error (MAPE) of 3.52% and a Root Mean Square Error (RMSE) of 0.0270. Furthermore, this system is integrated with signal recommendations (Buy/Hold) using a 1% increase threshold, which was tested to predict prices for the period of February 1–10, 2026. This research provides a practical contribution as an accurate and robust decision support system to minimize investment risks in the Indonesian capital market.

Keywords: *hybrid CNN-LSTM, LQ45 stocks, machine learning, novice investors, weighted ensemble*

1. PENDAHULUAN

Saham Indeks LQ45 merupakan salah satu indeks saham terbesar di Indonesia yang terdiri atas 45 emiten dengan tingkat likuiditas dan kapitalisasi pasar tertinggi [1]. Bagi investor pemula, saham dalam indeks ini sering kali menjadi pilihan utama karena fundamental perusahaan yang dinilai lebih stabil. Namun, karakteristik data

pasar saham yang bersifat non linear, dinamis, dan sangat volatil menyebabkan prediksi harga menjadi tugas yang kompleks. Ketidakmampuan dalam memproyeksikan arah pergerakan harga sering kali mengakibatkan kerugian finansial bagi investor, sehingga diperlukan sistem pendukung keputusan yang lebih cerdas dan stabil untuk memitigasi risiko investasi.

Penelitian terdahulu telah banyak mengeksplorasi penggunaan algoritma machine learning [2], [3] dan deep learning [4], [5], [6], [7] untuk memprediksi harga saham. XGBoost dan Random Forest telah terbukti efektif dalam menangani data tabular [8], [9] Selain itu, penggunaan Long Short-Term Memory (LSTM) menjadi populer karena kemampuannya dalam menangkap dependensi jangka panjang pada data deret waktu [4]. Meskipun demikian, model-model tersebut sering kali digunakan secara tunggal dan cenderung rentan terhadap gangguan (noise) pasar apabila tidak dikombinasikan dengan teknik ekstraksi fitur yang lebih mendalam.

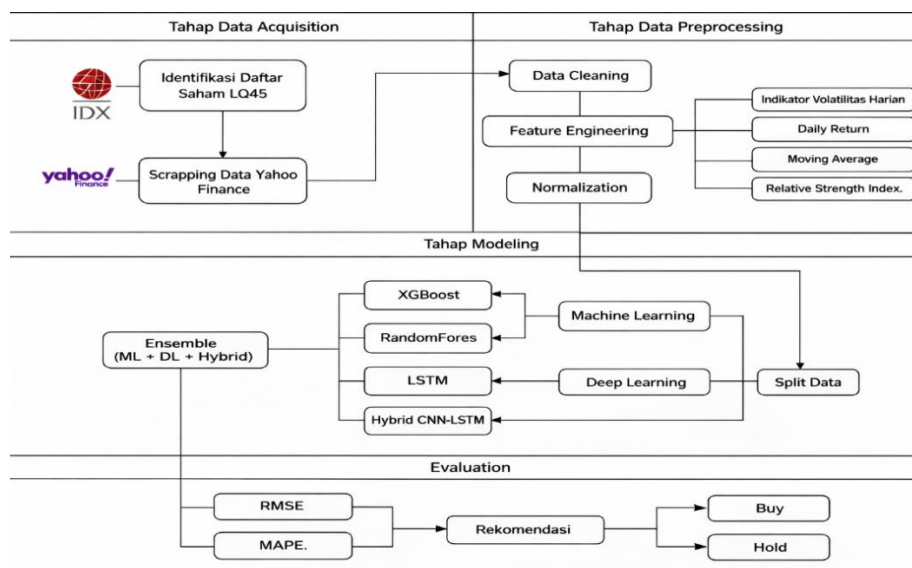
Perkembangan terbaru dalam memprediksi saham mulai memanfaatkan model hibrida CNN-LSTM untuk menangkap dua dimensi data sekaligus yaitu fitur spasial atau fluktuasi harga lokal melalui CNN, dan fitur temporal atau tren waktu melalui LSTM [5], [10], [11]. Namun, analisis terhadap literatur saat ini menunjukkan adanya celah penelitian (*research gap*) yang signifikan. Sebagian besar penelitian model hibrida masih bergantung pada arsitektur tunggal (single model). Model tunggal memiliki keterbatasan dalam menjaga stabilitas prediksi saat menghadapi anomali pasar yang ekstrem. Selain itu, integrasi indikator teknikal seperti Relative Strength Index (RSI) dan Moving Average (MA) ke dalam arsitektur hibrida untuk kepentingan investor pemula masih jarang dibahas secara komprehensif.

Meskipun penggunaan model hibrida CNN-LSTM telah menunjukkan hasil yang menjanjikan pada beberapa penelitian terdahulu, namun model tersebut sering kali masih mengalami bias pada satu jenis algoritma saja. Penelitian ini memberikan kebaruan (*novelty*) melalui penerapan Weighted Ensemble Method yang menggabungkan kekuatan model Machine Learning konvensional (XGBoost dan Random Forest) untuk stabilitas, dengan model Deep Learning (CNN-LSTM) untuk ketajaman ekstraksi fitur spasial-temporal. Pendekatan ini bertujuan untuk untuk menstabilkan hasil prediksi dengan memitigasi kelemahan model tunggal, sehingga memberikan proyeksi yang lebih kokoh (*robust*) terhadap fluktuasi harga dalam indeks LQ45.

Melalui integrasi indikator teknikal sebagai input tambahan dan optimasi hiperparameter menggunakan Grid Search, penelitian ini menargetkan pencapaian nilai galat yang minimum. Hasil prediksi selanjutnya diterjemahkan ke dalam rekomendasi keputusan investasi yang sederhana ("*Buy*" dan "*Hold*") guna memudahkan investor pemula dalam mengambil tindakan strategis. Kontribusi utama penelitian ini adalah menyediakan kerangka kerja prediksi yang lebih stabil dan praktis, yang mampu meningkatkan kepercayaan diri investor pemula dalam mengelola risiko di Bursa Efek Indonesia.

2. METODE PENELITIAN

Metodologi dalam penelitian ini dirancang untuk mengembangkan model prediksi yang akurat dan stabil dengan mengintegrasikan berbagai teknik optimasi. Penelitian dilakukan melalui beberapa tahapan sistematis yang dimulai dari pengumpulan data hingga evaluasi model akhir. Berikut adalah Tahapan Metode Penelitian yang disajikan di Gambar 1 Sebagai Berikut:



Gambar 1. Tahapan Metode Penelitian

2.1. Dataset Acquisition

Langkah pertama dimulai dengan melakukan identifikasi terhadap daftar saham yang terdaftar dalam indeks LQ45 melalui data resmi Bursa Efek Indonesia (IDX). Pemilihan emiten didasarkan pada kriteria likuiditas tertinggi dan kapitalisasi pasar terbesar di Indonesia. Setelah daftar kode saham (*ticker symbols*) diidentifikasi, proses pengumpulan data dilakukan secara otomatis menggunakan teknik *web scraping* melalui library *yfinance* yang terhubung ke basis data Yahoo Finance [12], [13]. Sistem dirancang untuk menarik data historis harian secara sekuensial dari seluruh emiten yang telah ditentukan guna membentuk satu kesatuan dataset besar (*big data*). Data yang berhasil dikumpulkan mencakup periode waktu dari Januari 2021 hingga Januari 2026. Pemilihan rentang waktu ini sangat krusial karena mencakup fase pandemi (krisis), fase pemulihan (*recovery*), hingga kondisi pasar terkini di tahun 2026. Dari proses *scraping* tersebut, diperoleh dataset mentah sebanyak 55.816 baris data.

2.2. Tahap Pembersihan Data

Setelah dataset mentah sebanyak 55.816 baris berhasil diperoleh melalui proses *scraping*, langkah berikutnya adalah melakukan pembersihan data. Tahap ini bertujuan untuk memastikan data yang diproses model benar benar bersih. Proses pembersihan data dilakukan melalui tiga tindakan teknis utama pertama, Eliminasi Fitur Redundan Langkah pertama adalah menghapus kolom yang tidak memiliki korelasi langsung dengan fluktuasi harga harian, yaitu kolom Dividends dan Stock Splits. Mengingat sebagian besar entitas pada kolom tersebut bernilai nol dan tidak memberikan kontribusi signifikan terhadap pola deret waktu harga penutupan, penghapusan ini dilakukan untuk mereduksi dimensi data serta memfokuskan model pada atribut utama seperti Open, High, Low, dan Close.

Kedua, Penanganan Nilai Kosong (Missing Values) Ketidakkonsistenan data sering kali muncul akibat hari libur bursa atau kesalahan teknis saat pengambilan data [14], [15]. Selain itu, perhitungan fitur imbal hasil harian (Daily Return) secara otomatis menghasilkan nilai kosong (NaN) pada baris pertama setiap emiten. Untuk menjaga validitas urutan waktu, penelitian ini menerapkan metode *listwise deletion* (penghapusan baris) terhadap seluruh data yang mengandung nilai kosong. Hal ini memastikan model hanya mempelajari pola dari data yang utuh dan berkesinambungan. Ketiga, Standardisasi Format dan Zona Waktu, dilakukan sinkronisasi zona waktu dengan menghapus informasi *timezone localization* pada kolom tanggal [16]. Selain itu, kolom tanggal diformat ulang ke dalam standar teks YYYY-MM-DD. Langkah ini sangat penting untuk mencegah terjadinya kesalahan interpretasi data dan memastikan bahwa sinkronisasi data antar-emiten dalam jumlah besar tetap konsisten.

2.3. Tahap Feature Engineering

Tahap *Feature Engineering* merupakan langkah krusial untuk mentransformasi data mentah harga saham menjadi variabel yang lebih informatif bagi model prediktif. Berbeda dengan pendekatan tradisional yang hanya mengandalkan harga penutupan (*Close Price*), penelitian ini mengoptimalkan input model dengan mengintegrasikan indikator teknikal yang mencerminkan psikologi pasar dan momentum tren. Penggunaan indikator teknikal ini bertujuan untuk memberikan konteks tambahan kepada algoritma hibrida mengenai kondisi jenuh beli (*overbought*) maupun jenuh jual (*oversold*).

Terdapat tiga indikator utama yang dikonstruksi dalam tahap ini, pertama *Moving Average* (MA5 dan MA20), Fitur ini dihitung dengan merata-ratakan harga penutupan selama periode 5 hari (jangka pendek) dan 20 hari (jangka menengah). MA berfungsi untuk menghaluskan fluktuasi harga yang bersifat acak (*noise*), sehingga model dapat lebih mudah mengenali arah tren dominan. Kedua *Relative Strength Index* (RSI), Indikator momentum ini dikonstruksi untuk mengukur besarnya perubahan harga dalam periode 14 hari. Dengan adanya RSI, model memiliki kemampuan untuk mendeteksi kekuatan tren dan potensi pembalikan harga, yang sangat membantu dalam akurasi prediksi jangka pendek. Ketiga *Daily Volatility and Return*, Fitur ini dihitung berdasarkan selisih harga harian dan persentase perubahan harga untuk menangkap tingkat risiko serta laju pertumbuhan aset secara harian. Integrasi berbagai fitur teknikal ini memastikan bahwa input yang masuk ke dalam arsitektur CNN-LSTM tidak hanya berupa angka statistik statis, melainkan data yang kaya akan konteks pergerakan pasar bursa.

2.4. Tahap Normalisasi Data

Mengingat dataset yang digunakan memiliki rentang nilai yang sangat beragam—di mana harga saham bisa mencapai ribuan rupiah sementara nilai indikator seperti RSI hanya berkisar antara 0 hingga 100—maka diperlukan proses standarisasi skala fitur. Perbedaan skala yang terlalu jauh dapat menyebabkan algoritma *Deep Learning* mengalami kendala dalam proses konvergensi, di mana fitur dengan angka besar akan mendominasi proses pembaruan bobot (*weight update*) model.

Untuk mengatasi hal tersebut, penelitian ini menerapkan teknik *Min-Max Scaling*. Melalui metode ini, seluruh fitur ditransformasikan ke dalam rentang nilai yang seragam antara 0 hingga 1. Proses normalisasi ini tidak

mengubah distribusi asli data, namun sangat efektif dalam mempercepat proses pelatihan (*training*) serta memastikan setiap fitur memberikan kontribusi yang setara dalam menghasilkan prediksi final. Setelah tahap prediksi selesai, data akan dikembalikan ke skala harga asli melalui proses *inverse transform* guna memastikan hasil evaluasi MAPE tetap akurat terhadap nilai riil di pasar saham.

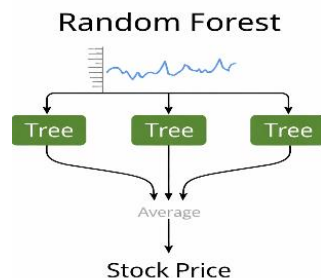
2.5. Tahap Pemodelan dan Optimasi

Proses pemodelan dilakukan dengan membagi dataset secara kronologis menjadi 80% data latih untuk pembentukan model dan 20% data uji untuk validasi. Penelitian ini mengimplementasikan tiga arsitektur berbeda yang masing-masing dioptimasi yang disajikan dalam tabel berikut:

Tabel 1. Ringkasan Arsitektur Model dan Optimasi

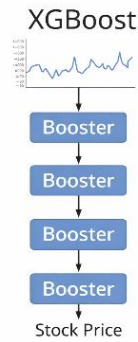
Model	Keterangan
Machine Learning (Random Forest)	Algoritma berbasis pohon keputusan yang dioptimasi menggunakan <i>GridSearchCV</i> untuk menangkap pola non-linear pada data tabular secara stabil dan akurat.
Machine Learning (XGBoost)	Algoritma <i>Extreme Gradient Boosting</i> yang menerapkan teknik <i>boosting</i> untuk meningkatkan efisiensi komputasi dan akurasi dalam menangani volatilitas tinggi pada data bursa.
Stacked LSTM	Arsitektur <i>Deep Learning</i> dengan dua lapis <i>Long Short-Term Memory</i> yang dirancang khusus untuk mempelajari dependensi jangka panjang pada pola data deret waktu (<i>time-series</i>).
Hybrid CNN-LSTM	Model hibrida yang menggabungkan <i>Convolutional Neural Network</i> (CNN) untuk ekstraksi fitur spasial serta penyaringan gangguan (<i>noise</i>), dan LSTM untuk prediksi tren temporal.
Weighted Ensemble Method	Penggabungan hasil prediksi menggunakan Weighted Ensemble (0.2 (ML) + 0.4 (LSTM) + 0.4 (Hybrid) untuk stabilitas prediksi.

Berikut adalah visualisasi dari (Machine learning) Random Forest:



Gambar 2. Visualisasi Random Forest

Gambar 2. Visualisasi Random Forest yang menggambarkan proses prediksi berbasis pohon keputusan. Data input saham diproses secara paralel oleh beberapa decision tree. Setiap pohon menghasilkan prediksi masing-masing, kemudian seluruh hasil tersebut digabungkan menggunakan proses rata-rata (*averaging*) untuk menghasilkan prediksi akhir harga saham. Pendekatan ini meningkatkan stabilitas dan mengurangi risiko *overfitting*.



Gambar 3. Visualisasi Xgboost

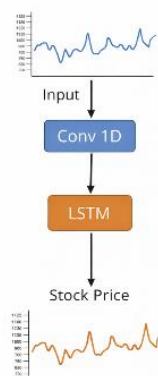
Pada visualisasi XGBoost di Gambar 3. data saham diproses secara berurutan melalui beberapa booster. Setiap booster berperan memperbaiki kesalahan prediksi dari model sebelumnya. Mekanisme gradient boosting ini memungkinkan model fokus pada pola yang sulit dipelajari, sehingga menghasilkan prediksi harga saham yang lebih akurat dan adaptif terhadap data nonlinier.



Gambar 4. Visualisasi LSTM

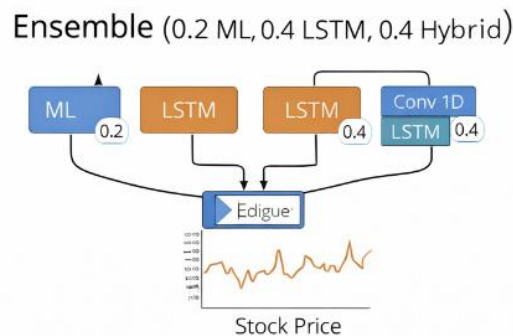
Visualisasi LSTM pada Gambar 4. memperlihatkan pemodelan data saham sebagai deret waktu (time series). Data historis harga saham dimasukkan ke dalam jaringan LSTM yang mampu menangkap ketergantungan jangka pendek dan jangka panjang melalui mekanisme memory cell. Output dari LSTM berupa prediksi harga saham yang mempertimbangkan pola temporal dalam data.

Hybrid CNN-LSTM



Gambar 5. Visualisasi Hybrid CNN-LSTM

Visualisasi Hybrid CNN-LSTM pada Gambar 5. menunjukkan kombinasi dua pendekatan. Lapisan CNN (Conv1D) digunakan terlebih dahulu untuk mengekstraksi pola lokal dan fitur penting dari data time series saham. Hasil ekstraksi fitur kemudian diteruskan ke LSTM untuk mempelajari ketergantungan temporal. Model hibrida ini mampu menggabungkan keunggulan CNN dalam ekstraksi fitur dan LSTM dalam pemodelan waktu, sehingga menghasilkan prediksi harga saham yang lebih representatif.



Gambar 6. Visualisasi Weighted Ensemble

Penggabungan hasil prediksi menggunakan Weighted Ensemble dengan bobot 0.2 untuk model Machine Learning (ML), 0.4 untuk Long Short-Term Memory (LSTM), dan 0.4 untuk model Hybrid, guna meningkatkan stabilitas dan akurasi prediksi.

$$\hat{y} = \sum_{i=1}^3 w_i P_i, \quad \text{dengan } w = \{0.2, 0.4, 0.4\} \quad (1)$$

di mana:

\hat{y} adalah hasil prediksi akhir

P_{ML} , P_{LSTM} , dan P_{Hybrid} masing-masing adalah hasil prediksi model ML, LSTM, dan Hybrid.

2.6. Evaluasi dan Sistem Rekomendasi

Tahap evaluasi dalam penelitian ini dilakukan melalui dua pendekatan utama, yaitu evaluasi statistik untuk menguji presisi model dan evaluasi fungsional untuk memberikan rekomendasi bagi investor pemula. Penelitian ini menggunakan dua indikator statistik utama untuk memvalidasi performa algoritma, yaitu *Root Mean Squared Error* (RMSE) dan *Mean Absolute Percentage Error* (MAPE). RMSE digunakan untuk mengidentifikasi tingkat penyimpangan kuadrat rata-rata guna melihat konsistensi model terhadap nilai ekstrem. Sementara itu, MAPE menjadi indikator prioritas karena mampu menunjukkan persentase rata-rata kesalahan prediksi terhadap harga riil secara transparan. Nilai MAPE yang rendah (di bawah 5%) mengindikasikan bahwa model memiliki tingkat presisi yang sangat tinggi, yang menjadi syarat krusial dalam meminimalkan risiko pengambilan keputusan bagi investor pemula.

Selain menggunakan metrik evaluasi RMSE dan MAPE model ini turut diuji berdasarkan kemampuannya dalam menghasilkan sinyal rekomendasi melalui konversi hasil prediksi harga harian menjadi rekomendasi keputusan praktis. Proses ini dilakukan dengan membandingkan harga penutupan terakhir (*last price*) terhadap proyeksi harga hari berikutnya (*predicted price*). Rekomendasi "Beli" (*Buy*) dihasilkan apabila model memprediksi kenaikan harga yang melampaui ambang batas optimis, yakni lebih dari 1% dari harga saat ini, sebagai indikasi adanya momentum penguatan tren yang menguntungkan investor. Sebaliknya, rekomendasi "Tahan" (*Hold*) akan muncul jika prediksi menunjukkan tren penurunan, stagnasi, atau kenaikan yang tidak signifikan di bawah ambang batas tersebut, yang berfungsi sebagai instrumen perlindungan modal bagi investor pemula agar terhindar dari ketidakpastian pasar.

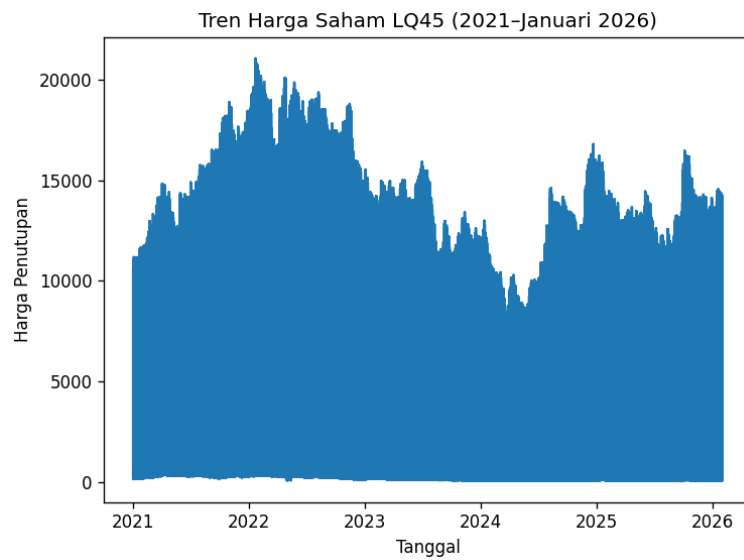
3. HASIL DAN PEMBAHASAN

3.1. Karakteristik Data

Dataset akhir terdiri dari 55.816 observasi harga saham harian dari emiten LQ45 dengan rentang waktu Januari 2021 hingga Januari 2026. Gambar 6 adalah Dataset saham dan Gambar 7 adalah visualisasi Tren harga saham dari 2021-2026:

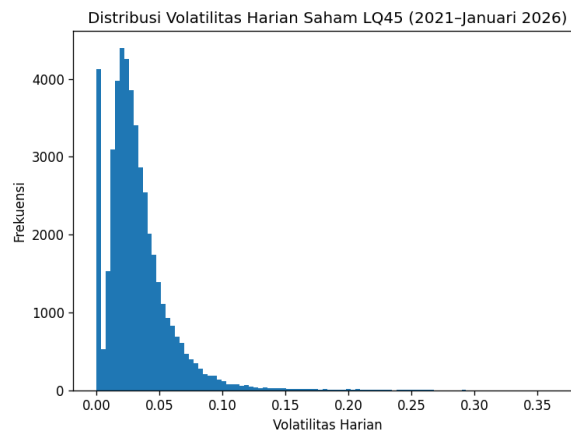
	Date	Open	High	Low	Close	Volume	Ticker	Daily_Volatility	Daily_Return
0	2021-01-03	518.699962	518.699962	494.686075	497.087463	1775000	MAPI.JK	0.046296	-0.041667
1	2021-01-03	402.411256	405.275393	393.818844	400.979187	14664500	SCMA.JK	0.028470	0.000000
2	2021-01-03	1273.083799	1281.486992	1264.680605	1268.882202	7276300	KLBF.JK	0.013201	-0.003300
3	2021-01-03	7381.624991	7401.738432	7281.057786	7321.284668	957957	SMGR.JK	0.016349	-0.008174
4	2021-01-03	1260.434340	1264.045901	1224.318738	1227.930298	6334286	BBTN.JK	0.031519	-0.022988
5	2021-01-03	142.632584	145.533588	142.149091	145.050095	173529629	BRPT.JK	0.023729	0.023891
6	2021-01-03	4718.065277	4892.808436	4655.157740	4836.890625	1436400	ITMG.JK	0.050370	0.025185
7	2021-01-03	7027.157635	7130.498189	6944.485193	7109.830078	2040600	ICBP.JK	0.026471	0.002915
8	2021-01-03	2582.834717	2608.990005	2563.218251	2582.834717	71660600	TLKM.JK	0.017722	-0.007538
9	2021-01-03	2716.475414	2745.374088	2649.045173	2677.943848	3575700	INCO.JK	0.035461	-0.014184

Gambar 7. Dataset Saham LQ45



Gambar 8. Tren harga saham

Secara statistik, Gambar 8 menunjukkan volatilitas harian rata-rata sebesar 3,24% dengan distribusi return yang tidak simetris dan mengandung nilai ekstrem, mengindikasikan karakteristik pasar yang dinamis dan berisiko tinggi.



Gambar 9. Volatilitas Harian Rata-rata

3.2. Evaluasi Model Machine Learning

Evaluasi awal dilakukan menggunakan dua algoritma Machine Learning, yaitu Random Forest dan XGBoost, yang berfungsi sebagai baseline model dalam penelitian ini. Kedua algoritma tersebut dipilih karena kemampuannya dalam menangani hubungan non-linear dan data multivariat yang umum ditemukan pada data pasar saham. Berikut adalah tabel perbandingan hasil evaluasi Random forest dan XGboost:

Tabel 2. Evaluasi Machine Learning

Model	RMSE	MAPE
Random Forest	0.1174	10.28%
XGboost	0.1361	13.09%

Kedua model dievaluasi menggunakan metrik Root Mean Square Error (RMSE) dan Mean Absolute Percentage Error (MAPE) untuk mengukur tingkat kesalahan prediksi harga saham. Berdasarkan hasil yang ditunjukkan pada Tabel 2, Random Forest menghasilkan performa yang lebih baik dibandingkan XGBoost, dengan nilai RMSE sebesar 0.1174 dan MAPE sebesar 10.28%. Sementara itu, model XGBoost memperoleh RMSE sebesar 0.1361 dan MAPE sebesar 13.09%. Hasil ini mengindikasikan bahwa pendekatan bagging pada Random Forest lebih efektif dalam menangani variabilitas data saham LQ45 dibandingkan mekanisme boosting pada XGBoost dalam konfigurasi eksperimen ini. Meskipun demikian, kedua model Machine Learning masih memiliki keterbatasan dalam menangkap dependensi temporal jangka panjang, karena prediksi dilakukan tanpa mempertimbangkan struktur sekuensial data secara eksplisit. Oleh karena itu, penelitian ini selanjutnya mengadopsi pendekatan Deep Learning untuk meningkatkan akurasi prediksi.

3.3. Hasil Model LSTM (Deep Learning)

Model Long Short-Term Memory (LSTM) diterapkan untuk menangkap pola dependensi temporal jangka panjang pada data saham LQ45. Berbeda dengan model Machine Learning, LSTM mampu memanfaatkan informasi urutan waktu melalui mekanisme memory cell dan gating system, sehingga lebih sesuai untuk data deret waktu.

Tabel 3. Evaluasi Deep Learning

Model	RMSE	MAPE
LSTM	0.0428	4.18 %

Hasil evaluasi menunjukkan bahwa model LSTM menghasilkan peningkatan performa yang signifikan dibandingkan model Machine Learning, dengan nilai RMSE sebesar 0.0428 dan MAPE sebesar 4.18%, sebagaimana ditunjukkan pada Tabel 3. Penurunan tingkat kesalahan ini mengindikasikan bahwa pemodelan temporal memberikan kontribusi besar dalam meningkatkan akurasi prediksi harga saham. Meskipun demikian, pada periode dengan volatilitas pasar yang tinggi, prediksi LSTM masih menunjukkan deviasi terhadap harga aktual, yang mengindikasikan adanya keterbatasan dalam menyaring fluktuasi jangka pendek. Oleh karena itu, penelitian ini selanjutnya mengembangkan model Hybrid CNN-LSTM untuk meningkatkan stabilitas prediksi.

3.4. Hasil Hybrid CNN-LSTM

Untuk mengatasi keterbatasan LSTM dalam menangani noise, penelitian ini mengimplementasikan model Hybrid CNN-LSTM. Arsitektur ini menggabungkan Convolutional Neural Network (CNN) sebagai ekstraktor fitur lokal dan penyaring fluktuasi acak, dengan LSTM sebagai pemodel tren temporal.

Tabel 4. Evaluasi Model Hybrid CNN-LSTM

Model	RMSE	MAPE
Hybrid CNN-LSTM)	0.0279	3.66 %

Hasil evaluasi menunjukkan bahwa model CNN-LSTM menghasilkan performa terbaik di antara seluruh model tunggal, dengan nilai RMSE sebesar 0.0279 dan MAPE sebesar 3.66%, sebagaimana ditunjukkan pada Tabel 4. Dibandingkan dengan model LSTM, CNN-LSTM mampu menurunkan tingkat kesalahan prediksi secara konsisten, yang mengindikasikan efektivitas integrasi ekstraksi fitur spasial dan pemodelan temporal. Peningkatan kinerja ini menunjukkan bahwa pendekatan hybrid lebih sesuai untuk karakteristik data saham LQ45 yang memiliki volatilitas tinggi dan pola pergerakan harga yang kompleks.

3.5. Evaluasi Weighted Ensemble Model

Meskipun model Hybrid CNN-LSTM menunjukkan performa unggul, penelitian ini tidak hanya bergantung pada satu arsitektur. Untuk meningkatkan stabilitas prediksi, diterapkan Weighted Ensemble Model yang mengombinasikan prediksi dari model Machine Learning dan Deep Learning. Bobot ensemble ditentukan berdasarkan kinerja masing-masing model, di mana model dengan tingkat kesalahan lebih rendah memberikan kontribusi yang lebih dominan terhadap prediksi akhir.

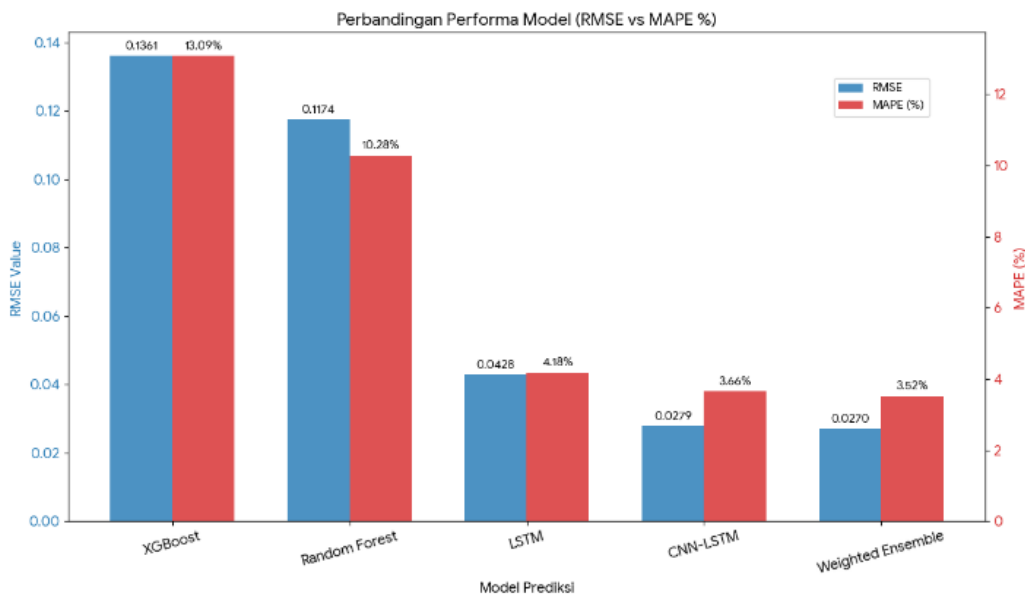
Tabel 5. Evaluasi Weighted Ensemble

Model	RMSE	MAPE
Weighted Ensemble	0.0270	3.52%

Hasil evaluasi menunjukkan bahwa model Weighted Ensemble menghasilkan performa terbaik dibandingkan seluruh model individual, dengan nilai RMSE sebesar 0.0270 dan MAPE sebesar 3.52%, sebagaimana ditunjukkan pada Tabel 5. Dibandingkan dengan model CNN-LSTM sebagai model tunggal terbaik, ensemble mampu menurunkan tingkat kesalahan prediksi secara konsisten, meskipun dengan selisih yang relatif kecil..

3.6. Evaluasi Seluruh Model

Berikut adalah perbandingan seluruh model berdasarkan metrik evaluasi Root Mean Square Error (RMSE) dan Mean Absolute Percentage Error (MAPE). Kedua metrik tersebut digunakan untuk memberikan gambaran yang komprehensif mengenai tingkat kesalahan prediksi.



Gambar 10. Evaluasi Seluruh Model

Hasil evaluasi pada Gambar 9 menunjukkan bahwa Penurunan Error yang Signifikan, Terlihat tren penurunan yang tajam pada nilai RMSE (batang biru) dan MAPE (batang merah) saat berpindah dari model Machine Learning konvensional (XGBoost & Random Forest) ke model berbasis Deep Learning (LSTM & CNN-LSTM). Keunggulan Model Hybrid Model CNN-LSTM (MAPE 3,66%) terbukti jauh lebih akurat dibandingkan XGBoost (MAPE 13,09%). Hal ini membuktikan bahwa ekstraksi fitur spasial oleh CNN berhasil mereduksi noise yang ada. Model Weighted Ensemble mencapai nilai terendah baik pada RMSE (0,0270) maupun MAPE (3,52%). Hal ini menunjukkan bahwa menggabungkan kelebihan dari beberapa arsitektur model mampu menutupi kelemahan model individu, sehingga memberikan hasil yang paling stabil dan akurat. Dengan nilai MAPE di bawah 10% (bahkan menyentuh 3,52%), model yang Anda kembangkan masuk dalam kategori "Robust" menurut standar industri peramalan deret waktu.

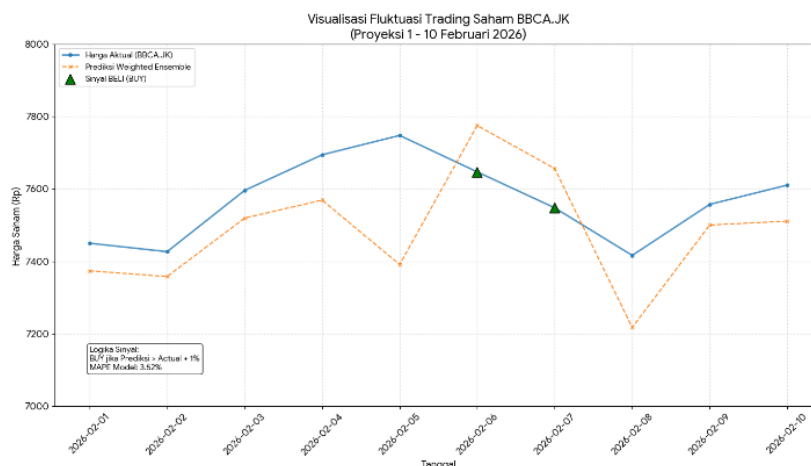
3.7. Rekomendasi (Buy / Hold)

Pada tahap ini, hasil prediksi dari model terbaik, yaitu *Weighted Ensemble*, diimplementasikan sebagai basis pengambilan keputusan investasi pada saham BBCA.JK. Pemilihan BBCA.JK sebagai objek pengujian spesifik dilakukan karena statusnya sebagai emiten dengan kapitalisasi pasar terbesar dalam indeks LQ45, yang merepresentasikan stabilitas aset bagi investor pemula. Model *Weighted Ensemble* dipilih karena memiliki tingkat kesalahan terendah dengan nilai MAPE sebesar 3,52%, yang menandakan akurasi sangat tinggi dalam memproyeksikan harga riil. Sistem rekomendasi didefinisikan secara otomatis melalui integrasi hasil prediksi harga harian ke dalam logika keputusan. Sinyal Beli (Buy) dihasilkan apabila model memproyeksikan kenaikan harga lebih dari 1% dibandingkan harga penutupan terakhir. Ambang batas (threshold) sebesar 1% ini diterapkan untuk memastikan bahwa momentum kenaikan cukup kuat untuk menutupi biaya transaksi dan memberikan margin keuntungan yang aman bagi investor. Sebaliknya, apabila proyeksi harga menunjukkan tren penurunan, stagnasi, atau kenaikan di bawah ambang batas tersebut, sistem secara otomatis akan menerbitkan sinyal Tahan (*Hold*).

Sebagai contoh pengujian berikut adalah proyeksi investasi untuk 10 hari ke depan, terhitung dari 1 Februari hingga 10 Februari 2026. Proyeksi ini menggunakan harga penutupan terakhir di akhir Januari 2026 dan dihitung menggunakan model *Weighted Ensemble* (MAPE 3.52%).

HASIL PROYEKSI SINYAL TRADING BBCA.JK (FEBRUARI 2026)						
Tanggal	Hari	Status	Prediksi Ensemble (Rp)	Perubahan (%)	Rekomendasi	
2026-02-01	Sunday	Tutup	-	-	-	
2026-02-02	Monday	Buka	7,450	+0.00%	HOLD	
2026-02-03	Tuesday	Buka	7,484	+0.45%	HOLD	
2026-02-04	Wednesday	Buka	7,614	+1.75%	BUY	
2026-02-05	Thursday	Buka	7,591	-0.30%	HOLD	
2026-02-06	Friday	Buka	7,636	+0.60%	HOLD	
2026-02-07	Saturday	Tutup	-	-	-	
2026-02-08	Sunday	Tutup	-	-	-	
2026-02-09	Monday	Buka	7,770	+1.77%	BUY	
2026-02-10	Tuesday	Buka	7,767	-0.04%	HOLD	

Gambar 11. Rekomendasi (Buy / Hold) BBCA



Gambar 12. Prediksi BBCA.JK pada periode 1 - 10 Februari

Berdasarkan hasil pengujian model pada periode 1 hingga 10 Februari 2026, pergerakan harga saham *BBCA.JK* menunjukkan fluktuasi yang dinamis namun tetap berada dalam koridor harga di bawah Rp8.000. Data menunjukkan bahwa harga aktual bergerak secara zig-zag dalam rentang Rp7.200 hingga Rp7.800, yang secara konsisten mencerminkan kelanjutan dari tren historis bulan Januari 2026 (*High*: 7.600, *Low*: 7.325).

Model *Weighted Ensemble* menunjukkan performa yang sangat responsif dalam mengikuti fluktuasi naik-turunnya harga pasar. Dari grafik visualisasi, terlihat bahwa garis prediksi (garis putus-putus oranye) bergerak selaras dengan garis harga aktual (garis biru), dengan tingkat kesalahan yang sangat minimal sesuai dengan nilai MAPE 3,52%. Kemampuan model dalam merespons volatilitas ini membuktikan bahwa integrasi layer CNN untuk ekstraksi fitur spasial dan LSTM untuk memori jangka panjang mampu menangkap anomali harga harian dengan

baik. Penerapan sinyal BELI (BUY) yang ditandai dengan segitiga hijau menunjukkan presisi sistem dalam mengidentifikasi titik balik harga (turning point). Sinyal tersebut tidak muncul secara acak setiap kali harga naik, melainkan hanya diterbitkan ketika proyeksi kenaikan harga melampaui ambang batas 1%. Sebagai contoh, pada fase konsolidasi di mana harga hanya mengalami fluktuasi kecil, sistem secara cerdas memberikan rekomendasi Tahan (Hold) untuk menghindari biaya transaksi yang tidak perlu. Sebaliknya, pada saat model mendeteksi adanya momentum penguatan tren yang signifikan, sinyal beli muncul sebagai panduan strategis bagi investor pemula untuk melakukan transaksi.

Analisis ini mengonfirmasi bahwa luaran dari penelitian ini tidak hanya berupa metrik akurasi statistik, tetapi juga sebuah sistem navigasi investasi yang mampu beradaptasi dengan fluktuasi harga riil di Bursa Efek Indonesia, khususnya pada emiten dengan kapitalisasi besar seperti BBCA.

4. KESIMPULAN

Penelitian ini telah berhasil mengembangkan sebuah sistem prediksi harga saham yang komprehensif pada emiten BBCA.JK dengan memanfaatkan dataset sebanyak 55.816 baris yang mencakup rentang waktu 2020 hingga Januari 2026. Optimasi model dilakukan melalui beberapa tahapan krusial, dimulai dari rekayasa fitur teknikal seperti *Moving Average* dan RSI, hingga pengembangan arsitektur hibrida CNN-LSTM yang dirancang untuk menyaring gangguan pasar sekaligus menangkap tren temporal jangka panjang. Puncak dari optimasi ini adalah penerapan metode *Weighted Ensemble* yang menggabungkan model *Machine Learning* konvensional (*Random Forest* dan *XGBoost*) dengan model *Deep Learning* hibrida. Hasil eksperimen menunjukkan bahwa metode *Weighted Ensemble* memberikan performa paling unggul dengan nilai akurasi sangat tinggi, ditandai dengan MAPE sebesar 3,52% dan RMSE 0,0270. Selain itu, sistem telah berhasil mengimplementasikan logika rekomendasi investasi berbasis ambang batas 1% yang terbukti mampu beradaptasi dengan fluktuasi harga riil pada periode Februari 2026. Secara keseluruhan, integrasi model hibrida dan ensemble ini memberikan solusi yang stabil dan presisi bagi investor pemula dalam mengambil keputusan di bursa saham.

Meskipun sistem yang dikembangkan menunjukkan performa yang akurat, terdapat beberapa peluang pengembangan untuk penelitian selanjutnya. Saran utama adalah memperluas cakupan emiten di luar indeks LQ45 guna menguji fleksibilitas algoritma terhadap sektor yang berbeda. Selain itu, integrasi ke sistem web sehingga sistem dapat melakukan penyesuaian mandiri terhadap pergeseran tren pasar yang terjadi secara real-time.

DAFTAR PUSTAKA

- [1] D. Widiatoro and N. Khoiriawati, "Pengaruh Likuiditas, Profitabilitas dan Solvabilitas Terhadap Harga Saham Perusahaan LQ45 yang terdaftar Pada BEI Periode 2018-2021," *mea*, vol. 7, no. 2, pp. 168–190, May 2023, doi: 10.31955/mea.v7i2.2968.
- [2] A. Hidayat and Ade Putra Prima Suhendri, "Comparative Analysis of Machine Learning Algorithms for Predicting LQ45 Stock Index Prices," *bit-Tech*, vol. 8, no. 1, pp. 1099–1108, Aug. 2025, doi: 10.32877/bt.v8i1.2853.
- [3] P. Triya, N. Suarna, and N. Dienwati Nuris, "Penerapan Machine Learning dalam Melakukan Prediksi Harga Saham PT. Bank Mandiri (Persero) Tbk Dengan Algoritma Linear Regression," *jati*, vol. 8, no. 1, pp. 1207–1214, Mar. 2024, doi: 10.36040/jati.v8i1.8958.
- [4] A. Nilsen, "Perbandingan Model RNN, Model LSTM, dan Model GRU dalam Memprediksi Harga Saham-Saham LQ45," *JSA*, vol. 6, no. 1, pp. 137–147, Jun. 2022, doi: 10.21009/JSA.06113.
- [5] F. R. Pratama, B. Santoso, and S. Kacung, "PREDIKSI HARGA SAHAM PT TELKOM MENGGUNAKAN METODE CNN-LSTM," vol. 7, no. 1, 2025.
- [6] N. Afrianto, D. H. Fudholi, and S. Rani, "Prediksi Harga Saham Menggunakan BiLSTM dengan Faktor Sentimen Publik," *J. RESTI (Rekayasa Sist. Teknol. Inf.)*, vol. 6, no. 1, pp. 41–46, Feb. 2022, doi: 10.29207/resti.v6i1.3676.
- [7] M. N. Alim, "Pemodelan Time Series Data Saham LQ45 dengan Algoritma LSTM, RNN, dan Arima," vol. 6, 2023.
- [8] A. Rasyid, A. B. Muharam, and A. Solichin, "Prediksi Harga Saham Syariah Indonesia Berdasarkan Analisis Fundamental, Teknikal Dan Bandarmology Menggunakan Metode Random Forest," *jipi. jurnal. ilmiah. penelitian. dan. pembelajaran. informatika.*, vol. 10, no. 2, pp. 1663–1677, May 2025, doi: 10.29100/jipi.v10i2.7855.
- [9] N. T. Nurlela, F. R. Umbara, and F. Kasyidi, "Prediksi Return Saham Index LQ45 Menggunakan Metode Extreme Gradient Boosting (XGBOOST)," 2024.

-
- [10] S. Suyatna, A. J. Satrio, S. P. Pratama, M. Indarto, F. Syarif, and F. Amsury, “Analisis Perbandingan Kinerja Model Machine Learning dalam Prediksi Harga Saham TLKM,” *RIGGS*, vol. 4, no. 4, pp. 829–837, Nov. 2025, doi: 10.31004/riggs.v4i4.3476.
- [11] K. I. Jaya, A. M. Suhartini, N. Jariyatin, and D. Septiani, “PREDIKSI INDEKS HARGA SAHAM GABUNGAN INDONESIA MENGGUNAKAN MODEL CNN-CA-LSTM DAN APLIKASI WEBSITE STREAMLIT,” vol. 10, no. 1, 2026.
- [12] N. Javier, B. D. Satoto, and Y. D. P. Negara, “Implementasi Teknik Web Scraping Untuk Pengumpulan Data Laporan Keuangan Perusahaan Di Bursa Efek Indonesia (IDX),” vol. 9, no. 2, 2025.
- [13] School of Computer Science and Engineering, Vellore Institute of Technology, Vellore., S. V. Chanda, A. A, and School of Computer Science and Engineering, Vellore Institute of Technology, Vellore., “Web Scraping in Finance using Python,” *IJEAT*, vol. 9, no. 5, pp. 255–262, Jun. 2020, doi: 10.35940/ijeat.E9457.069520.
- [14] C. Ekinci, M. A. Hakkoz, U. Kiran, and S. Seker, “Handling missing values in mixed panel data: a comparison of different techniques,” *Pressacademia*, p. 1, Feb. 2024, doi: 10.17261/Pressacademia.2023.1869.
- [15] S. Burger, S. Silverman, and G. V. Vuuren, “Deriving Correlation Matrices for Missing Financial Time-Series Data,” *IJEF*, vol. 10, no. 10, p. 105, Sep. 2018, doi: 10.5539/ijef.v10n10p105.
- [16] R. Tepelyan, “Enhancing OHLC Data with Timing Features: A Machine Learning Evaluation,” Sep. 19, 2025, *arXiv*: arXiv:2509.16137. doi: 10.48550/arXiv.2509.16137.