

Analisis Komparatif Model Machine Learning dan Deep Learning pada Peramalan Harga Saham Time Series

Edi Ismanto^{1,*}, Ahmad Gunawan Dalimunthe², Muhammad Iqbal³, Fauza Addinunnisa^{4*}

^{1,2,3,4} Program Studi Teknik Informatika, Universitas Muhammadiyah Riau, Pekanbaru, Indonesia
Email: ^{1*} edi.ismanto@umri.ac.id, ²230401024@student.umri.ac.id, ³230401010@student.umri.ac.id,
⁴210401130@student.umri.ac.id,

(* Email Corresponding Author: edi.ismanto@umri.ac.id)

Received: January 24, 2026 | Revision: January 29, 2026 | Accepted: January 29, 2026

Abstrak

Peramalan harga saham harian merupakan permasalahan yang menantang dalam bidang keuangan dan data science akibat tingginya volatilitas pasar serta pengaruh berbagai faktor eksternal. Penelitian ini menyajikan analisis komparatif antara beberapa model Machine Learning (ML) dan Deep Learning (DL) untuk peramalan harga saham berbasis indikator teknikal. Model ML yang dievaluasi meliputi Random Forest, Support Vector Regressor (SVR), dan XGBoost, sedangkan pendekatan DL mencakup Long Short-Term Memory (LSTM) dan Dense Neural Network (DNN). Data yang digunakan berupa data historis harga saham yang diperkaya dengan indikator teknikal seperti Moving Average (MA), Relative Strength Index (RSI), dan Bollinger Bands (BB). Kinerja model dievaluasi menggunakan Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), serta koefisien determinasi (R^2). Hasil eksperimen menunjukkan bahwa Support Vector Regressor menghasilkan kinerja prediksi terbaik, diikuti oleh Random Forest dan XGBoost. Model Deep Learning menunjukkan performa yang relatif lebih rendah, yang berkaitan dengan keterbatasan jumlah data serta kompleksitas proses tuning hiperparameter. Temuan ini mengindikasikan bahwa model Machine Learning, khususnya SVR, lebih efektif untuk peramalan harga saham dalam konteks dan kondisi eksperimental penelitian ini.

Kata Kunci: Machine Learning, Deep Learning, LSTM, XGBoost, Support Vector Regressor

Abstract

Daily stock price forecasting remains a challenging problem in finance and data science due to high market volatility and the influence of various external factors. This study presents a comparative analysis of several Machine Learning (ML) and Deep Learning (DL) models for stock price forecasting based on technical indicators. The evaluated ML models include Random Forest, Support Vector Regressor (SVR), and XGBoost, while the DL approaches comprise Long Short-Term Memory (LSTM) and Dense Neural Network (DNN). The dataset consists of historical stock price data enriched with technical indicators such as Moving Average (MA), Relative Strength Index (RSI), and Bollinger Bands (BB). Model performance is evaluated using Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and the coefficient of determination (R^2). Experimental results indicate that the Support Vector Regressor achieves the best predictive performance, followed by Random Forest and XGBoost. The Deep Learning models exhibit relatively lower performance, which is associated with limited data availability and the complexity of hyperparameter tuning. These findings suggest that Machine Learning models, particularly SVR, are more effective for stock price forecasting under the experimental setting of this study.

Keywords: Machine Learning, Deep Learning, LSTM, XGBoost, Support Vector Regressor

1. PENDAHULUAN

Perkembangan pesat sektor ekonomi global yang ditandai dengan meningkatnya aktivitas pasar keuangan telah mendorong kebutuhan akan strategi pengambilan keputusan investasi yang semakin akurat dan berbasis data. Investasi pada hakikatnya merupakan penempatan sejumlah dana pada masa kini dengan harapan memperoleh keuntungan di masa mendatang, yang tingkat keberhasilannya sangat bergantung pada kemampuan investor dalam memahami dan memprediksi dinamika pasar [1]. Salah satu instrumen investasi yang paling diminati adalah saham, karena potensi keuntungannya yang relatif tinggi dibandingkan instrumen keuangan lainnya. Namun demikian, investasi saham juga memiliki tingkat risiko yang tinggi akibat fluktuasi harga yang dipengaruhi oleh berbagai faktor, baik internal perusahaan maupun kondisi eksternal ekonomi makro [2], [3].

Di Indonesia, minat masyarakat terhadap investasi saham mengalami peningkatan yang sangat signifikan dalam beberapa tahun terakhir. Data Bursa Efek Indonesia (BEI) menunjukkan bahwa jumlah investor ritel meningkat dari sekitar 4 juta pada tahun 2020 menjadi lebih dari 11 juta pada awal tahun 2024. Peningkatan ini didorong oleh kemudahan akses teknologi digital, pertumbuhan platform perdagangan daring, serta masifnya edukasi literasi keuangan melalui media digital [4]. Kondisi tersebut menuntut

tersedianya metode analisis dan peramalan harga saham yang tidak hanya akurat, tetapi juga adaptif terhadap karakteristik pasar modal Indonesia yang cenderung volatil dan dipengaruhi oleh sentimen jangka pendek. Oleh karena itu, prediksi harga saham masih menjadi tantangan yang relevan dan menarik bagi para peneliti, tercermin dari banyaknya penelitian yang terus dilakukan dalam domain ini [5].

Secara tradisional, analisis harga saham dilakukan menggunakan pendekatan fundamental dan teknikal. Analisis fundamental menitikberatkan pada kinerja keuangan perusahaan, kondisi industri, dan faktor ekonomi makro, sedangkan analisis teknikal berfokus pada pola historis pergerakan harga dan volume perdagangan. Meskipun kedua pendekatan tersebut masih banyak digunakan, keterbatasannya terletak pada ketidakmampuan dalam menangkap pola nonlinier yang kompleks serta interaksi antarvariabel dalam skala data yang besar. Seiring dengan berkembangnya era big data, pendekatan konvensional ini mulai bergeser ke arah pemanfaatan teknologi kecerdasan buatan, khususnya *Machine Learning* (ML) dan *Deep Learning* (DL).

Metode *machine learning* dan *deep learning* memiliki kemampuan untuk memproses data historis dalam jumlah besar, mengekstraksi pola tersembunyi, serta mempelajari hubungan nonlinier yang sulit ditangkap oleh model statistik klasik. Berbagai penelitian menunjukkan bahwa algoritma ML seperti *Linear Regression*, *Random Forest*, *Support Vector Machine*, dan metode *boosting*, serta model DL seperti *Long Short-Term Memory* (LSTM), mampu menghasilkan performa yang lebih baik dalam peramalan harga saham dibandingkan pendekatan tradisional [6], [7]. Model *deep learning*, khususnya LSTM, dirancang untuk menangani data deret waktu (time series) dengan memori jangka panjang, sehingga sangat sesuai untuk memodelkan dependensi temporal pada data harga saham.

Namun demikian, meskipun penggunaan machine learning dan deep learning dalam prediksi harga saham telah berkembang pesat, masih terdapat sejumlah permasalahan yang perlu dikaji lebih lanjut. Salah satu tantangan utama adalah pemilihan algoritma yang paling sesuai dengan karakteristik data yang digunakan. Jenis data, skala waktu, panjang deret waktu, serta tingkat volatilitas pasar dapat memengaruhi kualitas hasil peramalan secara signifikan, karena setiap algoritma memiliki asumsi dan mekanisme pembelajaran yang berbeda [8], [9]. Selain itu, hasil penelitian terdahulu menunjukkan bahwa tidak ada satu model yang secara konsisten unggul dalam semua kondisi pasar, sehingga diperlukan studi perbandingan yang sistematis untuk memperoleh gambaran komprehensif terkait kelebihan dan keterbatasan masing-masing metode.

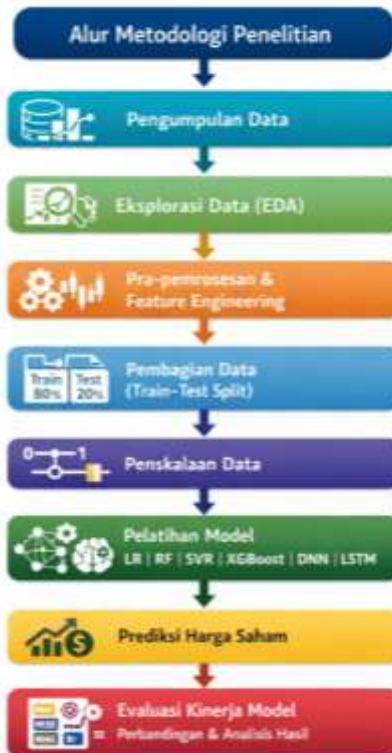
Sebagian penelitian sebelumnya cenderung hanya memfokuskan pada satu atau dua algoritma, atau membandingkan metode dengan konfigurasi dan dataset yang berbeda, sehingga menyulitkan penarikan kesimpulan yang objektif. Oleh karena itu, eksperimen komparatif menggunakan dataset yang sama, metrik evaluasi yang seragam, serta konteks pasar yang spesifik menjadi sangat penting untuk menilai performa relatif antara model *machine learning* dan *deep learning* dalam peramalan harga saham. Pendekatan perbandingan ini diharapkan dapat memberikan kontribusi ilmiah berupa rekomendasi model yang paling efektif dan akurat, khususnya untuk data time series harga saham.

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk menganalisis dan membandingkan kinerja beberapa metode machine learning dan deep learning dalam melakukan peramalan harga saham berbasis data historis deret waktu. Metode yang digunakan meliputi *Linear Regression* sebagai model baseline, *Random Forest* dan *XGBoost* sebagai representasi algoritma *ensemble learning*, *Support Vector Regressor* (SVR) sebagai model berbasis margin, serta *Long Short-Term Memory* (LSTM) sebagai model deep learning yang dirancang khusus untuk data time series. Evaluasi performa model dilakukan menggunakan metrik *Mean Absolute Error* (MAE) dan koefisien determinasi (R^2), yang umum digunakan untuk mengukur tingkat akurasi dan kemampuan model dalam menjelaskan variasi data aktual.

Dengan adanya penelitian ini, diharapkan dapat diperoleh pemahaman yang lebih mendalam mengenai efektivitas relatif antara pendekatan *machine learning* dan *deep learning* dalam peramalan harga saham, serta memberikan kontribusi praktis bagi investor dan akademisi dalam memilih model yang sesuai untuk mendukung pengambilan keputusan investasi yang lebih informasional dan berbasis data.

2. METODOLOGI PENELITIAN

Metodologi penelitian yang sistematis dan terstruktur sangat penting untuk memastikan hasil penelitian yang valid, reliabel, dan dapat direplikasi. Penelitian ini menggunakan pendekatan kuantitatif eksperimental dengan membandingkan performa beberapa model *Machine Learning* (ML) dan *Deep Learning* (DL), yaitu *Linear Regression*, *Random Forest*, *XGBoost*, *Support Vector Regressor* (SVR), *Long Short-Term Memory* (LSTM), dan *Deep Neural Network* (DNN), dalam melakukan peramalan harga saham berbasis data time series historis. Tahapan metodologi penelitian di tampilkan pada Gambar 1 yang meliputi pengumpulan dan eksplorasi data, pra-pemrosesan, pembagian dan penskalaan data, pemodelan, serta evaluasi kinerja model.



Gambar 1. Struktur/Alur Metodologi Penelitian

2.1 Pengumpulan dan Eksplorasi Data

Data yang digunakan dalam penelitian ini merupakan data saham berbentuk deret waktu (*time series*), sehingga peramalan harga saham dilakukan dengan memanfaatkan informasi harga historis pada periode sebelumnya untuk memprediksi harga pada periode berikutnya. Dataset terdiri dari 3.126 *record* data harian dengan atribut utama meliputi *Date*, *Open*, *High*, *Low*, *Close*, *Adj Close*, dan *Volume*.

Tahap eksplorasi data awal (*Exploratory Data Analysis* / EDA) dilakukan untuk memahami karakteristik dan struktur data. Proses ini mencakup pemeriksaan nilai hilang (missing values), tipe data, serta perhitungan statistik deskriptif dasar seperti nilai minimum, maksimum, rata-rata, dan standar deviasi. Selain itu, visualisasi data menggunakan grafik harga penutupan (*Close price*) ditampilkan untuk mengidentifikasi tren umum, pola musiman, serta tingkat volatilitas harga saham selama periode pengamatan. Proses EDA dilakukan menggunakan pustaka Pandas dan Matplotlib pada bahasa pemrograman Python.

2.2 Pre-processing

Tahap pra-pemrosesan dimulai dengan mengonversi kolom tanggal (*Date*) ke dalam format datetime dan menetapkannya sebagai indeks data untuk memfasilitasi analisis berbasis time series. Selanjutnya, dilakukan rekayasa fitur (feature engineering) dengan menambahkan sejumlah indikator teknikal yang umum digunakan dalam analisis pasar saham guna meningkatkan kemampuan prediktif model.

Indikator teknikal yang digunakan dalam penelitian ini meliputi:

- a. Moving Average (MA), yaitu rata-rata harga penutupan dengan periode 7 dan 21 hari.Expon
- b. ential Moving Average (EMA), yang memberikan bobot lebih besar pada harga terbaru.
- c. Moving Average Convergence Divergence (MACD), sebagai indikator momentum yang menunjukkan hubungan antara dua EMA.
- d. Relative Strength Index (RSI), yang mengukur kecepatan dan perubahan pergerakan harga.
- e. Bollinger Bands, sebagai indikator volatilitas yang terdiri dari pita atas, tengah, dan bawah.

Seluruh perhitungan indikator teknikal dilakukan menggunakan fungsi-fungsi pada pustaka Pandas dan NumPy. NumPy merupakan pustaka fundamental untuk komputasi numerik pada Python yang pertama kali diperkenalkan pada tahun 2005 [10].

2.3 Pembagian dan Penskalaan Data

Variabel target (y) didefinisikan sebagai harga penutupan hari berikutnya, yang dibuat dengan menggeser kolom *Close* satu langkah ke belakang. Dataset kemudian dibagi menjadi data pelatihan (80%) dan data pengujian (20%) dengan *shuffle=False* untuk menjaga urutan temporal data. Semua fitur prediktor (X) kemudian dinormalisasi menggunakan *MinMaxScaler* dari *Scikit-learn* untuk mengubah skala data ke rentang [0, 1], sebuah langkah penting untuk model yang sensitif terhadap skala seperti SVR dan jaringan saraf.

Variabel target (y) didefinisikan sebagai harga penutupan (*Close*) pada hari berikutnya, yang diperoleh dengan menggeser kolom *Close* satu langkah ke belakang (*one-step ahead forecasting*). Dataset kemudian dibagi menjadi data pelatihan (*training set*) sebesar 80% dan data pengujian (*testing set*) sebesar 20%. Pembagian data dilakukan tanpa pengacakan (*shuffle = False*) untuk menjaga urutan temporal data dan menghindari kebocoran informasi (*data leakage*).

Seluruh fitur prediktor (X) dinormalisasi menggunakan metode *Min-Max Scaling* dengan rentang [0, 1] menggunakan pustaka *Scikit-learn*. Tahap penskalaan ini sangat penting, khususnya untuk model yang sensitif terhadap perbedaan skala data seperti SVR dan jaringan saraf.

2.4 Model Machine Learning dan Deep Learning

Penelitian ini mengimplementasikan lima model prediksi yang terdiri dari metode *Machine Learning* dan *Deep Learning* sebagai berikut:

a. *Random Forest* (RF)

Random Forest merupakan metode *ensemble learning* berbasis kumpulan pohon keputusan (*decision trees*) yang dibangun dari subset data secara acak. Prediksi akhir diperoleh dengan menggabungkan hasil prediksi masing-masing pohon, sehingga mampu mengurangi *overfitting* dan meningkatkan akurasi model [11].

b. *Support Vector Regressor* (SVR)

SVR bekerja dengan membentuk fungsi regresi yang memaksimalkan margin dengan batas deviasi tertentu (ϵ -*insensitive loss*), sehingga mampu menghasilkan prediksi yang mendekati nilai aktual, khususnya pada data nonlinier [12].

c. *Long Short-Term Memory* (LSTM)

LSTM merupakan varian dari *Recurrent Neural Network* (RNN) yang dirancang untuk mempelajari ketergantungan jangka panjang pada data deret waktu. Model ini menggunakan mekanisme *input gate*, *forget gate*, dan *output gate* untuk mengontrol aliran informasi di dalam sel memori [13].

d. *Deep Neural Network* (DNN)

DNN adalah jaringan saraf dengan banyak lapisan tersembunyi (*multi-layer*) yang mampu memodelkan hubungan nonlinier kompleks. Model ini meniru cara kerja neuron pada otak manusia dan banyak digunakan dalam pengambilan keputusan berbasis data [14].

e. XGBoost

XGBoost merupakan algoritma *boosting* berbasis pohon keputusan yang dioptimalkan untuk kecepatan dan kinerja tinggi. Metode ini banyak digunakan dalam berbagai bidang karena kemampuannya menangani data besar dan kompleks secara efisien [15].

2.5 Evaluasi Kinerja Model

Evaluasi kinerja setiap model dilakukan menggunakan data pengujian dengan beberapa metrik evaluasi standar, yaitu Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), dan koefisien determinasi (R^2). Selain itu, digunakan metrik akurasi kustom berbasis rasio RMSE terhadap nilai rata-rata harga aktual untuk memberikan interpretasi kesalahan dalam bentuk persentase. Evaluasi ini bertujuan untuk menentukan model yang paling akurat dan stabil dalam memprediksi harga saham. Evaluasi dilakukan menggunakan persamaan (1), (2), (3), dan (4) berikut:

a. Rumus Mean Squared Error (MSE)

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (1)$$

b. Rumus Root Mean Squared Error (RMSE)

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (P_i - Q_i)^2} \quad (2)$$

c. Rumus Mean Absolute Error (MAE)

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (3)$$

d. Rumus *R-squared* (R^2)

$$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2} \quad (4)$$

3. HASIL DAN PEMBAHASAN

Pada bagian ini dilakukan evaluasi dan perbandingan kinerja lima model prediksi harga saham yang digunakan dalam penelitian, yaitu *Support Vector Regressor* (SVR), *Random Forest* (RF), XGBoost, *Long Short-Term Memory* (LSTM), dan *Deep Neural Network* (DNN). Evaluasi dilakukan menggunakan metrik akurasi prediksi, *Root Mean Squared Error* (RMSE), dan koefisien determinasi (R^2) untuk menilai tingkat ketepatan dan kemampuan model dalam merepresentasikan variasi data harga saham aktual.

3.1 Hasil Evaluasi Kinerja Model

Rekapitulasi hasil evaluasi kinerja masing-masing model disajikan pada Tabel 1. Berdasarkan hasil tersebut, model *Support Vector Regressor* (SVR) menunjukkan performa terbaik dibandingkan model lainnya, diikuti oleh *Random Forest* dan XGBoost. Sementara itu, model *deep learning* (LSTM dan DNN) menghasilkan performa yang relatif lebih rendah pada dataset yang digunakan.

Tabel 1. Perbandingan Kinerja Model Prediksi Harga Saham

Model	Akurasi (%)	RMSE	R^2
<i>Support Vector Regressor</i> (SVR)	98.50	10.0973	0.9904
<i>Random Forest</i> (RF)	98.38	10.8734	0.9889
XGBoost	98.18	12.1932	0.9860
<i>Long Short-Term Memory</i> (LSTM)	97.13	14.20 – 15.00	0.972 – 0.976
<i>Deep Neural Network</i> (DNN)	96.51	16.50 – 17.80	0.958 – 0.965

Berdasarkan Tabel 1, SVR memperoleh akurasi tertinggi sebesar 98.50% dengan nilai RMSE terendah, yaitu 10.0973, serta koefisien determinasi (R^2) sebesar 0.9904. Nilai R^2 yang mendekati 1 menunjukkan bahwa model SVR mampu menjelaskan hampir seluruh variasi data harga saham aktual. Random Forest dan XGBoost juga menunjukkan performa yang sangat kompetitif, namun masih berada sedikit di bawah SVR.

Nilai RMSE dan R^2 pada model LSTM dan DNN menunjukkan performa yang lebih rendah dibandingkan model *machine learning* klasik, yang mengindikasikan bahwa kompleksitas model *deep learning* belum sepenuhnya terkompensasi oleh ukuran dan karakteristik dataset yang digunakan. Hal ini dapat disebabkan oleh keterbatasan jumlah data, karakteristik volatilitas harga saham, serta kompleksitas model *deep learning* yang memerlukan data dalam skala lebih besar agar dapat bekerja secara optimal.

3.2 Pembahasan dan Analisis Hasil

Keunggulan model SVR dalam penelitian ini dapat diatribusikan pada kemampuannya dalam menemukan hyperplane optimal di ruang fitur berdimensi tinggi yang dibentuk oleh indikator-indikator teknis. Penggunaan kernel Radial Basis Function (RBF) memungkinkan SVR untuk memetakan hubungan nonlinier antara variabel input dan harga saham secara efektif, sehingga menghasilkan prediksi yang lebih akurat dan stabil dibandingkan model lainnya.

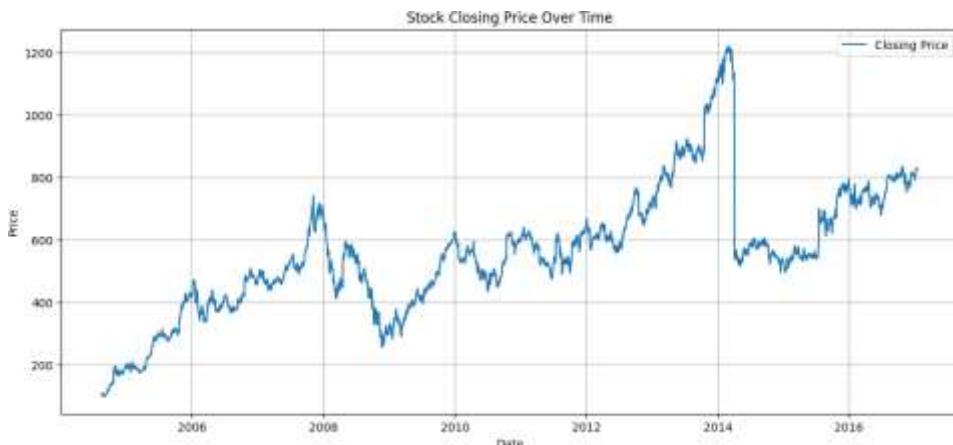
Untuk memperkuat hasil kuantitatif tersebut, dilakukan visualisasi perbandingan antara harga saham aktual dan harga hasil prediksi menggunakan model SVR. Visualisasi ini ditampilkan pada Gambar 1.



Gambar 2. Perbandingan Harga Aktual vs. Harga Prediksi oleh Model SVR.

Berdasarkan Gambar 2, terlihat bahwa kurva harga prediksi yang dihasilkan oleh model SVR mengikuti pola pergerakan harga aktual dengan sangat baik. Selisih antara nilai prediksi dan nilai aktual relatif kecil pada sebagian besar periode pengujian, yang mengonfirmasi keunggulan performa SVR tidak hanya secara numerik, tetapi juga secara visual. Hasil ini menunjukkan bahwa SVR memiliki kemampuan generalisasi yang baik dalam menangkap pola temporal dan dinamika pergerakan harga saham.

Secara keseluruhan, hasil penelitian ini menunjukkan bahwa model machine learning klasik, khususnya *Support Vector Regressor*, masih sangat kompetitif dan bahkan unggul dibandingkan model deep learning pada konteks peramalan harga saham dengan dataset time series berukuran menengah dan berbasis indikator teknikal. Gambar 2 menyajikan grafik harga penutupan saham dari tahun 2005 hingga 2016 yang menunjukkan dinamika pasar yang cukup fluktuatif.



Gambar 3. Grafik Harga Penutupan Saham dari Tahun 2005 hingga 2016

Berdasarkan Gambar 3, pada periode awal tahun 2005 hingga 2007 terlihat adanya tren kenaikan harga yang relatif stabil. Kondisi ini mengindikasikan fase pertumbuhan pasar dengan volatilitas yang masih terkendali. Namun, pada sekitar tahun 2008 terjadi penurunan harga yang sangat tajam, yang sejalan dengan terjadinya krisis keuangan global (global financial crisis) yang secara luas memengaruhi pasar saham di berbagai negara. Penurunan drastis ini mencerminkan meningkatnya ketidakpastian pasar dan tekanan jual yang tinggi pada periode tersebut.

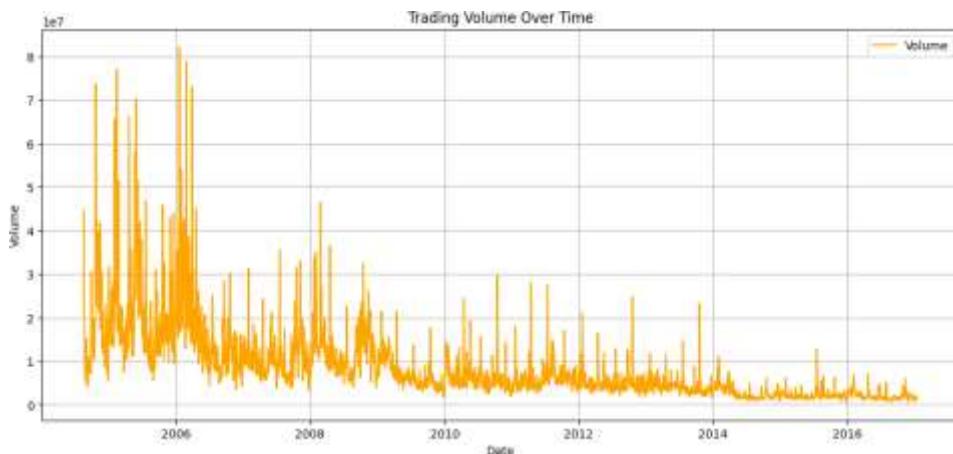
Memasuki tahun 2009 hingga 2013, grafik harga saham menunjukkan tren kenaikan secara bertahap, yang dapat diinterpretasikan sebagai fase pemulihan pasar pascakrisis. Pada periode ini, pergerakan harga relatif lebih stabil dibandingkan masa krisis, meskipun fluktuasi jangka pendek tetap terjadi. Puncak harga saham terlihat pada awal tahun 2014, dengan nilai harga penutupan mencapai lebih dari 1.200 satuan. Lonjakan ini kemudian diikuti oleh penurunan harga yang signifikan dalam waktu singkat.

Penurunan tajam setelah puncak harga tersebut dapat diasosiasikan dengan berbagai faktor korporasi maupun pasar, seperti pembagian dividen dalam jumlah besar, aksi korporasi berupa stock split, atau peristiwa keuangan tertentu yang berdampak langsung terhadap struktur harga saham. Setelah fase koreksi

tersebut, pergerakan harga kembali menunjukkan pola yang lebih stabil, dengan tren kenaikan moderat hingga akhir periode pengamatan pada tahun 2016.

Fluktuasi harga yang terlihat sepanjang periode pengamatan ini menegaskan bahwa data harga saham memiliki karakteristik nonlinier, dinamis, dan dipengaruhi oleh kejadian eksternal yang sulit diprediksi. Oleh karena itu, model peramalan harga saham dituntut untuk mampu menangkap pola volatilitas jangka pendek sekaligus tren jangka panjang yang terkandung dalam data historis. Hasil yang diperoleh dalam penelitian ini menunjukkan bahwa pemanfaatan indikator teknikal yang dikombinasikan dengan model machine learning, khususnya Support Vector Regressor, mampu merepresentasikan dinamika tersebut dengan baik dan menghasilkan estimasi harga yang akurat.

Analisis ini juga memperkuat temuan sebelumnya bahwa pemodelan berbasis data historis dan indikator teknikal memiliki peran penting dalam meningkatkan kinerja prediksi harga saham, terutama pada dataset time series dengan periode pengamatan yang panjang dan tingkat volatilitas yang bervariasi. Selanjutnya, untuk melengkapi analisis dinamika pasar saham, dilakukan pengamatan terhadap pola volume perdagangan selama periode penelitian. Gambar 3 menampilkan grafik volume perdagangan saham dari tahun 2005 hingga 2016 yang memberikan informasi penting mengenai tingkat aktivitas pasar dan likuiditas saham sepanjang waktu.



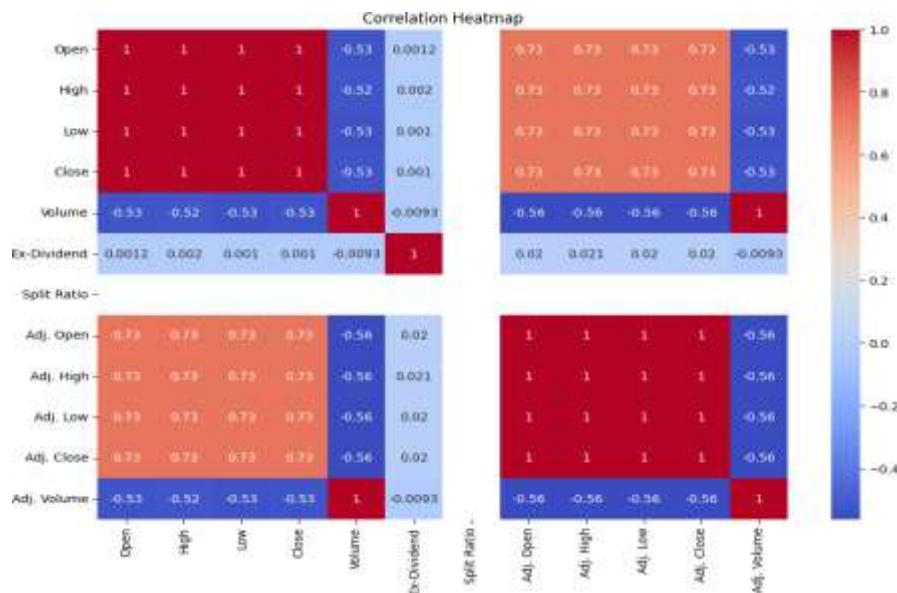
Gambar 4. Grafik Volume Perdagangan Saham dari Tahun 2005 hingga 2016

Berdasarkan Gambar 4, terlihat bahwa volume perdagangan saham mengalami fluktuasi yang relatif tinggi pada awal periode pengamatan, khususnya pada tahun 2005 hingga 2006. Pada periode ini, terjadi lonjakan volume perdagangan yang mencapai lebih dari 80 juta saham per hari, yang mencerminkan tingginya intensitas transaksi dan partisipasi pelaku pasar. Kondisi tersebut dapat dikaitkan dengan berbagai faktor eksternal, seperti sentimen pasar yang positif, rilis informasi ekonomi penting, maupun perubahan kebijakan moneter yang memengaruhi perilaku investor.

Namun demikian, seiring berjalannya waktu, volume perdagangan menunjukkan tren penurunan yang cukup konsisten. Setelah tahun 2007, grafik volume mencerminkan penurunan aktivitas perdagangan secara bertahap, hingga volume rata-rata berada di bawah 10 juta saham per hari setelah tahun 2012. Penurunan ini dapat mengindikasikan berkurangnya minat investor terhadap saham yang bersangkutan, perubahan kondisi makroekonomi, atau pergeseran strategi perdagangan oleh investor institusional yang cenderung lebih selektif.

Fluktuasi volume perdagangan ini merupakan salah satu indikator penting dalam analisis dan peramalan harga saham. Perubahan volume sering kali menjadi sinyal awal terjadinya pergeseran tren harga atau perubahan momentum pasar. Oleh karena itu, variabel volume perdagangan dimasukkan sebagai salah satu fitur utama dalam model prediksi yang dikembangkan dalam penelitian ini, dengan tujuan untuk menangkap hubungan antara aktivitas transaksi dan pergerakan harga saham.

Selain analisis univariat terhadap harga dan volume, dilakukan pula analisis multivariat untuk memahami hubungan antarvariabel dalam dataset. Gambar 4 menyajikan matriks korelasi antar variabel yang digunakan dalam penelitian ini.



Gambar 5. Matriks Korelasi Antar Variabel pada Dataset Saham

Berdasarkan Gambar 5, terlihat bahwa variabel harga *Open*, *High*, *Low*, dan *Close* memiliki korelasi yang sangat tinggi satu sama lain, dengan nilai korelasi mendekati 1. Kondisi ini menunjukkan bahwa pergerakan harga saham cenderung seragam dalam satu hari perdagangan, di mana perubahan pada satu jenis harga sangat berkaitan dengan perubahan pada harga lainnya. Korelasi tinggi juga ditemukan pada variabel harga penyesuaian (*Adjusted Open*, *Adjusted High*, *Adjusted Low*, dan *Adjusted Close*), yang menunjukkan konsistensi pergerakan harga setelah memperhitungkan aksi korporasi.

Sementara itu, variabel volume perdagangan (*Volume* dan *Adjusted Volume*) menunjukkan korelasi negatif moderat terhadap variabel harga saham, dengan nilai korelasi berkisar antara -0.53 hingga -0.56. Korelasi negatif ini mengindikasikan bahwa peningkatan harga saham cenderung diikuti oleh penurunan volume perdagangan, dan sebaliknya. Fenomena ini dapat mencerminkan kondisi pasar yang relatif kurang likuid pada saat harga berada pada level tinggi, atau adanya aksi distribusi oleh pelaku pasar ketika harga mendekati titik puncak. Variabel lain seperti *Ex-Dividend* dan *Split Ratio* menunjukkan korelasi yang sangat rendah terhadap variabel harga maupun volume. Hal ini mengindikasikan bahwa variabel-variabel tersebut bersifat relatif independen atau hanya berdampak pada periode tertentu, sehingga pengaruhnya terhadap pergerakan harga harian tidak bersifat langsung. Analisis korelasi ini menjadi dasar penting dalam proses pemilihan fitur (*feature selection*). Fitur-fitur yang memiliki korelasi sangat tinggi berpotensi menimbulkan redundansi informasi, khususnya pada model tertentu, sementara fitur dengan korelasi yang lebih rendah namun relevan secara teoritis tetap dipertahankan karena dapat memberikan sinyal tambahan yang berguna dalam proses peramalan harga saham. Dengan pendekatan ini, model yang dibangun diharapkan mampu menangkap informasi yang relevan secara optimal tanpa meningkatkan kompleksitas yang tidak diperlukan.

Perbedaan performa antara Random Forest dan Dense Neural Network dalam penelitian ini juga dapat dijelaskan dari karakteristik masing-masing algoritma. Random Forest sebagai model ensemble berbasis pohon keputusan cenderung lebih stabil, tidak sensitif terhadap skala data, serta mampu bekerja optimal pada dataset berukuran menengah. Sebaliknya, Dense Neural Network memiliki kompleksitas pelatihan yang lebih tinggi dan membutuhkan jumlah data yang besar serta tuning hiperparameter yang intensif agar dapat menghasilkan performa optimal. Kondisi ini menjelaskan mengapa pada studi ini model Random Forest mampu menghasilkan kinerja yang lebih baik dibandingkan DNN. Secara keseluruhan, hasil ini menegaskan bahwa pemilihan model prediksi harga saham harus mempertimbangkan kesesuaian antara karakteristik algoritma dan sifat data *time series* yang digunakan.

4. KESIMPULAN

Penelitian ini berhasil melakukan analisis komparatif terhadap beberapa model machine learning dan deep learning untuk peramalan harga saham. Hasil eksperimen menunjukkan bahwa model *Support Vector Regressor* (SVR) menghasilkan kinerja prediksi terbaik dibandingkan *Random Forest*, XGBoost, LSTM, dan Dense Neural Network berdasarkan metrik evaluasi yang digunakan pada dataset penelitian. Penambahan fitur berbasis indikator teknis terbukti menjadi faktor penting yang secara konsisten meningkatkan kemampuan prediktif seluruh model yang diuji. Temuan ini mengindikasikan bahwa pada

permasalahan peramalan time series saham dengan ukuran data dan rekayasa fitur yang tepat, model machine learning klasik seperti SVR dapat memberikan performa yang kompetitif, bahkan melampaui arsitektur deep learning yang lebih kompleks. Sebagai arah penelitian selanjutnya, studi ini merekomendasikan eksplorasi optimisasi hyperparameter yang lebih komprehensif, pemanfaatan data alternatif seperti sentimen berita dan media sosial, serta pengembangan arsitektur model hibrida untuk meningkatkan akurasi prediksi.

REFERENCES

- [1] A. R. S. Arya Sapta Dinata, "Pengaruh Tingkat Investasi Terhadap Pertumbuhan Ekonomi Dalam Perspektif Islam," *J. Bina Bangsa Ekon.*, vol. 18, no. 1, p. 215, 2024, [Online]. Available: <https://jbbe.lppmbinabangsa.id/index.php/jbbe/article/view/591/428>
- [2] H. Wang, J. Wang, L. Cao, Y. Li, Q. Sun, and J. Wang, "A Stock Closing Price Prediction Model Based on CNN-BiSLSTM," *Complexity*, vol. 2021, 2021, doi: 10.1155/2021/5360828.
- [3] I. K. Nti, A. F. Adekoya, and B. A. Weyori, "A systematic review of fundamental and technical analysis of stock market predictions," *Artif. Intell. Rev.*, vol. 53, no. 4, pp. 3007–3057, 2020, doi: 10.1007/s10462-019-09754-z.
- [4] A. Hardian *et al.*, "Indonesian Research Journal on Education," vol. 5, pp. 1079–1085, 2025.
- [5] A. Arfan and L. ETP, "Perbandingan Algoritma Long Short-Term Memory dengan SVR pada Prediksi Harga Saham di Indonesia," *Petir*, vol. 13, no. 1, 2020, doi: 10.33322/petir.v13i1.858.
- [6] Z. Hu, Y. Zhao, and M. Khushi, "A Survey of Forex and Stock Price Prediction Using Deep Learning," *Applied System Innovation*, vol. 4, no. 1. 2021. doi: 10.3390/asi4010009.
- [7] S. Sudriyanto, F. Syahro, and N. Fitriani, "Perbandingan Performa Model Machine Learning Support Vector Machine, Neural Network, Dan K-Nearest Neighbors Dalam Prediksi Harga Saham," *J. Adv. Res. Inform.*, vol. 2, no. 1, pp. 13–21, 2023, doi: 10.24929/jars.v2i1.2983.
- [8] G. Budiprasetyo, M. Hani'ah, and D. Z. Aflah, "Prediksi Harga Saham Syariah Menggunakan Algoritma Long Short- Term Memory (LSTM)," *J. Nas. Teknol. dan Sist. Inf.*, vol. 8, no. 3, pp. 164–172, 2023, doi: 10.25077/teknosi.v8i3.2022.164-172.
- [9] I. P. Putri, T. Tertiaavini, and N. Arminarahmah, "Analisis Perbandingan Algoritma Machine Learning untuk Prediksi Stunting pada Anak," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 4, no. 1, pp. 257–265, 2024, doi: 10.57152/malcom.v4i1.1078.
- [10] A. Pannadhitthana Candra, "Analisis Data Menggunakan Python: Memperkenalkan Pandas dan NumPy," *J. Inf. Syst. Educ. Dev.*, vol. 3, no. 1, pp. 11–16, 2025, doi: 10.62386/jised.v3i1.118.
- [11] A. Hadi, M. Qamal, and Y. Afrillia, "Comparison of Random Forest Algorithm Classifier and NaÃ¯ve Bayes Algorithm in Whatsapp Message Type Classification," *J. Renew. Energy, Electr. Comput. Eng.*, vol. 5, no. 1, pp. 9–17, 2025, doi: 10.29103/jreece.v5i1.21227.
- [12] E. N. Widyaningrum, R. A. Putri, M. A. Fathan, and N. R. Safitriani, "Stock Price Forecasting Using Autoregressive With Exogenous Variable Support Vector Regression (ARX – SVR)," *J. Mat. Stat. dan Komputasi*, vol. 21, no. 3, pp. 847–854, 2025, doi: 10.20956/j.v21i3.43613.
- [13] A. Hendrawan and E. I. Sela, "Analisis Sentimen Komentar Youtube Tentang Resesi Global 2023 Menggunakan LSTM," *J. Indones. Manaj. Inform. dan Komun.*, vol. 5, no. 1, pp. 587–593, 2024, doi: 10.35870/jimik.v5i1.526.
- [14] O. Oscar, N. Maulidah, A. Purnamawati, D. Putri, and H. F. Pardede, "Prediksi Tingkat Kesuksesan Promosi Bank Dengan Algoritma Dnn," *J. Inform.*, vol. 21, no. 1, pp. 23–33, 2021, doi: 10.30873/ji.v21i1.2866.
- [15] N. N. Pandika Pinata, I. M. Sukarsa, and N. K. Dwi Rusjanyathi, "Prediksi Kecelakaan Lalu Lintas di Bali dengan XGBoost pada Python," *J. Ilm. Merpati (Menara Penelit. Akad. Teknol. Informasi)*, vol. 8, no. 3, p. 188, 2020, doi: 10.24843/jim.2020.v08.i03.p04.