Indonesian Journal of Digital Business

Journal homepage: https://ejournal.upi.edu/index.php/IJDB

Peningkatan Prediksi Harga Saham Menggunakan Correlation-Driven PCA dan Temporal Attention-Based Deep Learning

Fahmi Iqbal Firmananda[1], Hidayati Rusnedy[2], Fazila Amalia[3]*

[1,3]Program Studi Digital Business, Universitas Pahlawan Tuanku Tambusai, Indonesia
[2]Program Studi Teknik Informatik, Universitas Pahlawan Tuanku Tambusai, Indonesia
Correspondence: E-mail: fahmiiqbalfirmananda@universitaspahlawan.ac.id

ABSTRACT

Fluktuasi harga saham yang tinggi menjadikan prediksi pergerakan harga saham sebagai tantangan yang kompleks. Berbagai pendekatan berbasis indikator teknikal telah digunakan, namun sering kali kurang efektif dalam menangkap pola nonlinier dan korelasi antar variabel. Penelitian ini mengusulkan integrasi Correlation-Driven Principal Component Analysis (CD-PCA) dengan Temporal Attention-Based Deep Learning sebagai metode untuk meningkatkan kualitas prediksi harga saham. CD-PCA digunakan untuk mentransformasi 1.195 indikator teknikal menjadi komponen utama yang mempertimbangkan korelasi terhadap variabel target, sehingga mampu mengurangi redundansi antar fitur tanpa kehilangan informasi penting. Data yang digunakan adalah saham perbankan besar Indonesia, yaitu BBCA, BBNI, BBRI, dan BMRI, dengan periode perdagangan dari 1 Januari 2015 hingga 30 Desember 2024. Eksperimen dilakukan dengan tiga skenario pembagian data, yaitu 70:30, 80:20, dan 90:10. Hasil penelitian menunjukkan bahwa model yang diusulkan mampu mengurangi hasil error prediksi secara konsisten. Pada skenario 70:30, saham BBCA mencatatkan MAPE terendah sebesar 2,17% dengan R^2 mencapai 0,85, menunjukkan performa prediksi yang baik ketika data latih cukup besar. Sementara itu, BBNI dan BBRI justru memberikan performa terbaik pada skenario 90:10 dengan MAPE masing-masing 3,56% dan 3,41%, serta R^2 sebesar

ARTICLE INFO

Article History:
Submitted/Received 29 September 2025
First Revised 29 Oktober 2025
Accepted 1 November 2025
First Available online 7 November 2025
Publication Date 7 November 2025

Keyword:
Correlation-Driven Principal Component Analysis,
Temporal Attention-Based Deep Learning,
Seleksi Fitur,
Prediksi Harga Saham,
Teknikal Indikator

0,85 dan 0,87. Sebaliknya, BMRI menunjukkan tingkat error lebih tinggi dan R^2 negatif pada skenario 70:30, menandakan kompleksitas pola harga saham yang sulit ditangkap. Secara keseluruhan, penelitian ini menegaskan bahwa integrasi CD-PCA dan Temporal Attention-Based Deep Learning berkontribusi terhadap pengurangan error prediksi harga saham, khususnya pada saham BBCA, BBNI, dan BBRI, sehingga berpotensi menjadi kerangka kerja andal dalam peramalan pasar saham di Indonesia.

© 2025 Kantor Jurnal dan Publikasi UPI

1. INTRODUCTION

Fluktuasi harga saham yang sangat cepat dan tidak terduga menjadikannya tantangan yang kompleks untuk diprediksi dengan akurat (Botunac et al., 2024). Dalam beberapa dekade terakhir, perkembangan teknologi machine learning telah membuka peluang baru untuk memprediksi pergerakan harga saham (Al-Khasawneh et al., 2024). Pendekatan berbasis indikator teknikal, seperti Simple Moving Average (SMA) (Khanpara et al., 2023; Priyatno, Ningsih, et al., 2024), Relative Strength Index (RSI) (Ming & Chen, 2024; Uma & Srinath Naidu, 2021), dan Bollinger Bands (Chen et al., 2022; S & Arya, 2024), telah diterapkan untuk menangkap tren historis data saham. Namun, metode konvensional ini sering kali tidak mampu menangkap hubungan nonlinier dan pola kompleks antar variabel dalam data, sehingga mengurangi akurasi prediksi.

Berbagai penelitian telah dilakukan untuk mengatasi keterbatasan ini, dengan menggunakan Principal Component Analysis (PCA) sebagai teknik pengurangan dimensi yang efektif. Penelitian (Mi et al., 2023; Sheng et al., 2022) menggabungkan PCA dengan Long Short-Term Memory (LSTM) untuk memprediksi indeks saham Shanghai, dan hasilnya menunjukkan peningkatan akurasi prediksi. Namun, penggunaan PCA tradisional masih menghadapi tantangan dalam menangani data saham yang penuh dengan hubungan nonlinier antar fitur.

Penelitian (Mansoor & Al-Sultan, 2025) menggunakan PCA sebelum menerapkan LSTM, yang menunjukkan penurunan kesalahan prediksi. Namun, PCA tradisional dalam pendekatan ini tidak cukup efektif dalam menangkap hubungan korelatif antar fitur . Di sisi lain, penelitian (Li et al., 2023) menggunakan PCA bersama regresi linier untuk meningkatkan akurasi prediksi, tetapi metode ini tidak optimal untuk data nonlinier. Penelitian (Ji, 2023) menunjukkan bahwa pengurangan dimensi dengan PCA dapat mengurangi overfitting dan mempercepat proses pelatihan pada algoritme seperti Stochastic Gradient Descent (SGD). Namun, hasil penelitian ini juga mengindikasikan bahwa PCA tidak selalu efektif dalam menjaga informasi penting dari variabel yang memiliki korelasi tinggi.

Penelitian (Zhu et al., 2025) menambahkan proses whitening pada PCA untuk meningkatkan prediksi harga saham, tetapi metode ini meningkatkan kompleksitas komputasi tanpa peningkatan yang signifikan dalam akurasi prediksi . Penelitian (Wang et al., 2023) mengintegrasikan PCA dengan IGRU (Improved Gated Recurrent Unit) untuk memprediksi harga saham. Hasilnya menunjukkan performa yang baik, tetapi tetap menghadapi kendala dalam menangani dimensi tinggi data saham. Penelitian (Uçkan, 2024) menunjukkan bahwa integrasi PCA dengan model deep learning berhasil meningkatkan akurasi prediksi harga saham pada pasar saham Turki. Namun, penelitian ini juga mencatat bahwa PCA tradisional sering mengabaikan hubungan korelatif antar fitur, sehingga informasi penting yang relevan untuk prediksi tidak sepenuhnya dimanfaatkan. Selain itu, PCA tradisional belum dioptimalkan untuk menangani pola temporal yang kompleks dalam data saham, yang dapat memengaruhi performa prediksi ketika diintegrasikan dengan model deep learning.

Penelitian ini mengusulkan pengembangan Correlation-Driven Principal Component Analysis (CD-PCA) sebagai modifikasi dari PCA tradisional untuk mempertimbangkan korelasi antar fitur dalam proses pengurangan dimensi. Selain itu, hasil CD-PCA diintegrasikan

dengan Temporal Attention-Based Deep Learning untuk memaksimalkan analisis pola temporal data saham. Pendekatan ini bertujuan untuk mengatasi tantangan tingginya dimensi data, hubungan nonlinier antar fitur, dan pola volatilitas dalam data saham.

2. METHODS

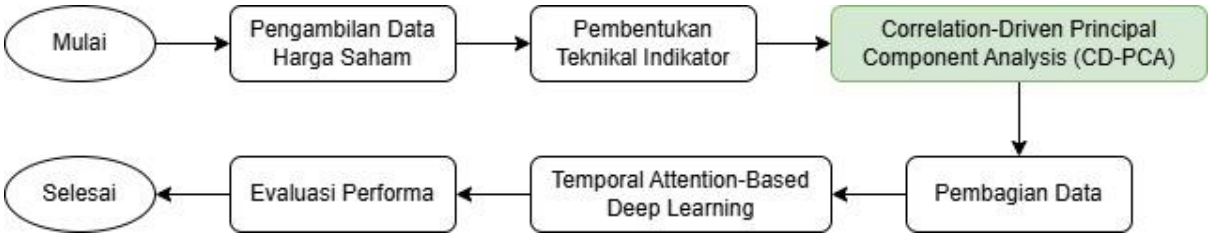


Figure 1 Framework Penelitian

Proses Penelitian dari pengambilan data hingga evaluasi prediksi harga saham. Figure 1 menampilkan secara keseluruhan tahap penelitian. Tahapan penelitian tersebut yaitu pengambilan data saham, pembentukan teknikal indikator, seleksi fitur menggunakan correlation-driven principal component analysis (CD-PCA), pembagian data, pemodelan menggunakan temporal attention-based deep learning, dan evaluasi performa prediksi.

2.1. Pengambilan Data

Data yang digunakan dalam penelitian ini berasal dari situs Yahoo Finance (finance.yahoo.com), yang merupakan salah satu sumber data finansial global terpercaya dan banyak digunakan dalam penelitian akademik maupun praktis di bidang pasar modal. Empat emiten perbankan besar Indonesia dipilih sebagai objek penelitian, yaitu PT Bank Central Asia Tbk (BBCA), PT Bank Rakyat Indonesia Tbk (BBRI), PT Bank Negara Indonesia Tbk (BBNI), dan PT Bank Mandiri Tbk (BMRI). Keempat emiten ini dipilih karena memiliki kapitalisasi pasar yang besar, tingkat likuiditas perdagangan yang tinggi, serta peran dominan dalam sektor perbankan nasional, sehingga data yang diperoleh dianggap representatif dalam mencerminkan pergerakan harga saham perbankan di Bursa Efek Indonesia (BEI). Periode data yang digunakan mencakup sepuluh tahun terakhir, mulai dari 1 Januari 2015 hingga 30 Desember 2024, sehingga mampu merekam dinamika jangka panjang pergerakan harga saham, termasuk fase pertumbuhan, koreksi, maupun gejolak pasar seperti pandemi COVID-19 pada tahun 2020. Rentang waktu yang panjang ini dipilih agar model prediksi tidak hanya didasarkan pada tren jangka pendek, tetapi juga dapat belajar dari variasi pola historis yang lebih luas. Variabel utama yang diunduh meliputi Date, Open, High, Low, Close, Adjusted Close, dan Volume, yang merupakan data historis standar dalam penelitian pasar saham.

2.2 Fitur Teknikal Indikator

Dalam penelitian ini digunakan berbagai indikator teknikal yang telah terbukti luas dalam literatur pasar modal sebagai alat untuk menganalisis pergerakan harga saham. Indikator teknikal dipilih karena mampu menangkap pola historis, tren, serta momentum pasar, sehingga dapat memberikan informasi penting bagi proses prediksi berbasis machine learning. Empat indikator utama yang digunakan adalah Simple Moving Average (SMA), Exponential Moving Average (EMA), Relative Strength Index (RSI), dan Bollinger Bands (BB). Indikator SMA dan EMA berfungsi sebagai penangkap tren harga. SMA menghitung rata-rata harga penutupan dalam periode tertentu, sedangkan EMA memberikan bobot lebih besar pada harga terbaru sehingga lebih responsif terhadap perubahan pasar. Indikator RSI digunakan untuk mengukur kekuatan relatif harga dan mengidentifikasi kondisi overbought

maupun oversold, yang sangat berguna dalam mendeteksi potensi pembalikan arah tren. Sementara itu, Bollinger Bands memanfaatkan deviasi standar untuk mengukur volatilitas harga, dengan pita atas dan bawah yang menunjukkan batas dinamis pergerakan harga saham. Rentang periode pembentukan indikator teknikal ditetapkan mulai dari 3 hari hingga 200 hari, sehingga cakupan indikator mencerminkan pergerakan jangka pendek, menengah, hingga jangka panjang. Dengan variasi periode ini, model dapat mempelajari pola dari fluktuasi harian hingga tren jangka panjang yang lebih stabil. Kombinasi indikator tren (SMA, EMA), momentum (RSI), dan volatilitas (Bollinger Bands) diharapkan mampu memberikan representasi yang komprehensif mengenai dinamika pasar saham perbankan Indonesia.

2.3 Seleksi Fitur Correlation-Driven Principal Component Analysis

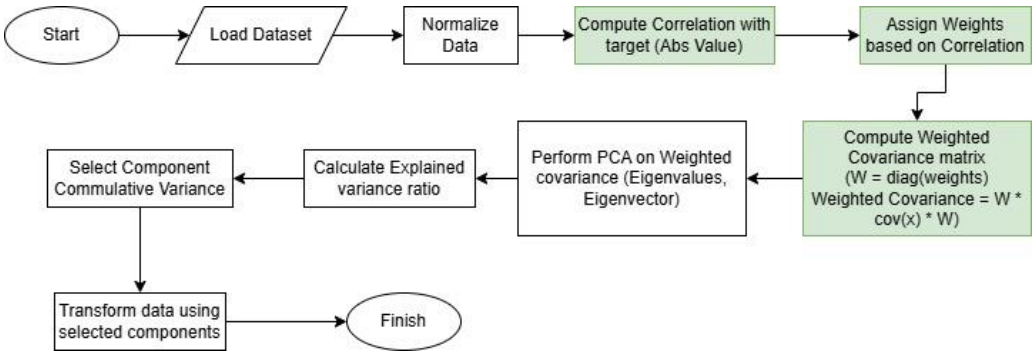


Figure 2 Proses Seleksi Fitur *Correlation-Driven Principal Component Analysis*

Proses seleksi fitur correlation-driven principal component analysis (CD-PCA) dapat dilihat pada Figure 2. Proses seleksi fitur dalam penelitian ini dilakukan menggunakan metode Correlation-Driven Principal Component Analysis (CD-PCA), yang merupakan pengembangan dari PCA tradisional dengan menambahkan komponen korelasi absolut terhadap variabel target. Tujuan utama metode ini adalah untuk menghasilkan representasi fitur baru yang tidak hanya mempertahankan informasi varian terbesar dari data, tetapi juga memperhatikan relevansi langsung terhadap target prediksi, dalam hal ini harga saham pada H+1. Tahapan proses CD-PCA ditunjukkan pada Figure 2. Pertama, data historis saham yang telah dilengkapi indikator teknikal dimuat dan dilakukan normalisasi menggunakan MinMaxScaler untuk menyeragamkan skala antar variabel. Selanjutnya, dihitung korelasi absolut antara setiap fitur dengan target. Korelasi ini menjadi dasar dalam pemberian bobot, di mana fitur yang memiliki korelasi lebih tinggi terhadap target mendapatkan bobot lebih besar. CA kemudian dilakukan pada matriks kovarians berbobot tersebut untuk memperoleh nilai eigen dan vektor eigen, yang digunakan dalam membentuk komponen utama. Selanjutnya dihitung explained variance ratio dari masing-masing komponen untuk menentukan kontribusi relatif terhadap total variansi data. Dari hasil ini, komponen utama dipilih berdasarkan variansi kumulatif yang dianggap cukup mewakili informasi data. Terakhir, data asli ditransformasi ke dalam ruang komponen utama yang telah dipilih, sehingga menghasilkan dataset baru dengan dimensi yang lebih efisien dan relevan untuk proses pemodelan prediksi. Langkah-langkah proses seleksi fitur ini dapat dilihat pada Algoritma 1.

Algoritma 1. Seleksi fitur *correlation-driven principle component analysis*

Input	:	Dataset saham dengan fitur teknikal

Output	:	Dataset dengan fitur baru berupa Principal Components
Step	:	1. Memuat dataset yang berisi fitur teknikal dan target harga saham.
		2. Normalisasi data menggunakan MinMaxScaler
		3. Menghitung korelasi absolut antara setiap fitur teknikal dan target: $r = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2} \cdot \sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}}$
		4. Menentukan bobot berdasarkan korelasi absolut $W = \text{diag}(r_{abs})$
		5. Menghitung matriks kovarians berbobot: $\text{Weighted Covariance} = W \cdot \text{Cov}(X) \cdot W$
		6. Melakukan PCA pada matriks kovarians berbobot untuk menghitung nilai eigen dan vektor eigen
		7. Menghitung rasio variansi setiap komponen utama
		8. Memilih komponen utama berdasarkan variansi kumulatif
		9. Mentransformasikan data asli menggunakan komponen utama terpilih

2.4. Pembagian Data

Tahap pembagian data merupakan langkah penting dalam proses pemodelan prediksi harga saham, karena menentukan bagaimana model dilatih dan dievaluasi. Dalam penelitian ini, data historis yang telah melalui tahap pembentukan indikator teknikal dan transformasi dengan Correlation-Driven PCA (CD-PCA) dibagi menjadi dua bagian utama, yaitu data latih (training set) dan data uji (testing set). Data latih digunakan untuk membangun dan menyesuaikan parameter model, sedangkan data uji digunakan untuk mengevaluasi performa model terhadap data yang tidak pernah dilihat sebelumnya, sehingga dapat mengukur kemampuan generalisasi model. Proses pembagian dilakukan dengan tiga skenario proporsi, yaitu 90:10, 80:20, dan 70:30. Pada skenario 90:10, sebagian besar data dialokasikan sebagai data latih (90%), sementara 10% sisanya digunakan sebagai data uji. Skema ini memungkinkan model belajar dari data historis yang sangat luas, namun hanya diuji pada sampel yang relatif kecil. Sebaliknya, pada skenario 70:30, porsi data uji menjadi lebih besar, yaitu 30% dari total data, sehingga memberikan evaluasi yang lebih ketat terhadap kinerja model, meskipun data latih menjadi lebih terbatas. Sementara itu, skenario 80:20 diposisikan sebagai kompromi antara keduanya. Pendekatan multi-skenario ini dipilih untuk mengevaluasi konsistensi kinerja model pada kondisi jumlah data latih yang berbeda. Dengan cara ini, hasil penelitian tidak hanya bergantung pada satu konfigurasi pembagian data, tetapi juga mencerminkan bagaimana model beradaptasi terhadap variasi ketersediaan data. Strategi ini penting dalam konteks prediksi harga saham, karena ketersediaan data di dunia nyata dapat bervariasi, baik karena keterbatasan historis maupun karena adanya kondisi pasar tertentu yang memengaruhi panjang data observasi.

2.5. Prediksi dengan Deep learning

Prediksi harga saham dilakukan menggunakan Long Short-Term Memory (LSTM) yang diperkaya dengan mekanisme temporal attention. LSTM dipilih karena mampu memproses

data runtun waktu dengan dependensi jangka panjang, sedangkan attention digunakan untuk memberikan bobot lebih besar pada langkah waktu yang paling relevan terhadap prediksi harga H+1. Arsitektur ini terdiri dari lapisan LSTM dengan 50 unit, lapisan perhatian temporal, serta kepala regresi berbasis Dense untuk menghasilkan nilai prediksi. Model dikompilasi menggunakan optimizer Adam dengan fungsi rugi MSE. Algoritma 2 merupakan temporal attention LSTM.

Algoritma 2. temporal attention LSTM

Langkah	**Deskripsi**
Input	Data runtun waktu (T, F) hasil transformasi CD-PCA, panjang jendela waktu (T), jumlah fitur per timestep (F).
1	Inisialisasi lapisan LSTM dengan 50 unit dan return_sequences=True untuk menghasilkan representasi tiap timestep h_t.
2	Hitung skor perhatian untuk setiap timestep menggunakan fungsi aktivasi tanh: $e_t = \tanh(W.h_t + b)$
3	Normalisasi skor perhatian menggunakan softmax sepanjang dimensi waktu: $\alpha_t = \frac{\exp(e_t)}{\sum_{k=1}^{T} \exp(r_k)}$
4	Hitung vektor konteks sebagai jumlah tertimbang representasi LSTM: $c = \sum_{t=1}^{T} \alpha_t . h_t$
5	Masukkan vektor konteks (c) ke lapisan Dense dengan 50 unit beraktivasi ReLU.
6	Gunakan lapisan Dense akhir dengan 1 neuron linear untuk menghasilkan prediksi harga $\hat{y}_{t+1}$
7	Kompilasi model menggunakan optimizer Adam dengan fungsi rugi MSE.
Output	Prediksi harga penutupan $\hat{y}_{t+1}$

2.6. Evaluasi

Penelitian ini dilakukan evaluasi dengan menggunakan *Mean squared error* (MSE) (Priyatno et al., 2023; Priyatno, Ramadhan Sudirman, et al., 2024), *Mean Absolute Error* (MAE) (Billah et al., 2024), *Mean absolute percentage error* (MAPE) (He & Wang, 2024), dan *R2 Score* (Haryono et al., 2024).

3. RESULTS AND DISCUSSION

Penelitian ini menggunakan data historis harga saham dari empat emiten perbankan besar di Indonesia, yaitu PT Bank Central Asia Tbk (BBCA), PT Bank Negara Indonesia Tbk (BBNI), PT Bank Rakyat Indonesia Tbk (BBRI), dan PT Bank Mandiri Tbk (BMRI). Data diperoleh dari Yahoo Finance dengan rentang waktu 01 januari 2015 hingga 30 desember 2024, sehingga mencakup periode sepuluh tahun perdagangan.

Setiap dataset terdiri atas 2.359 observasi harian yang merepresentasikan aktivitas perdagangan saham, dengan 1.195 variabel. Variabel-variabel tersebut meliputi data harga pasar (Open, High, Low, Close, Adjusted Close), volume transaksi, serta ratusan indikator teknikal yang dibentuk berdasarkan berbagai periode jangka waktu. Indikator teknikal yang digunakan mencakup Simple Moving Average (SMA), Exponential Moving Average (EMA), Relative Strength Index (RSI), dan Bollinger Bands, dengan rentang periode mulai dari 3 hari hingga 200 hari.

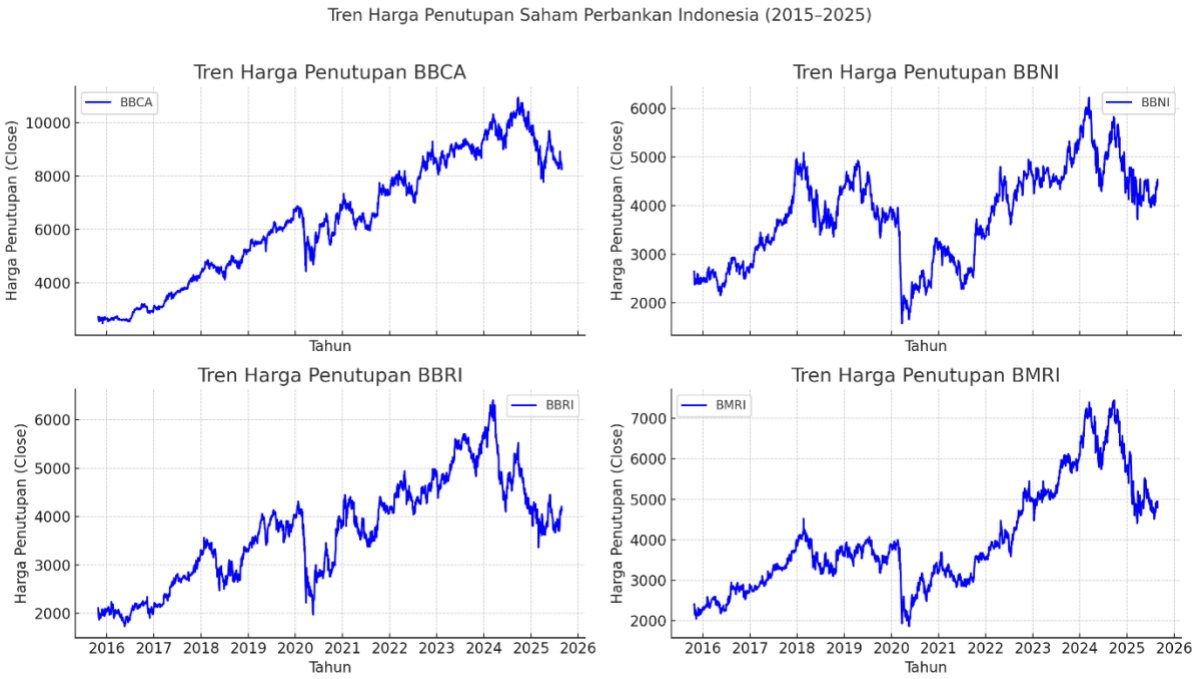


Figure 3 Grafik Pergerakan saham BBCA, BBNI, BBRI, dan BMRI

Figure 3 menampilkan grafik pergerakan harga penutupan saham BBCA, BBNI, BBRI, dan BMRI selama periode 2015 hingga 2025. Secara umum, keempat saham menunjukkan tren kenaikan jangka panjang, meskipun terdapat fluktuasi signifikan pada periode tertentu, khususnya saat terjadi gejolak pasar seperti pandemi COVID-19 pada tahun 2020.

Saham BBCA memperlihatkan pertumbuhan yang paling konsisten dan dominan, dengan harga penutupan yang terus meningkat hingga mencapai level di atas 10.000 pada tahun 2024 sebelum mengalami koreksi. Saham BBRI dan BMRI juga menunjukkan tren naik, dengan pola fluktuatif yang mencerminkan volatilitas sektor perbankan, namun tetap bergerak menuju level harga yang lebih tinggi dibandingkan awal periode pengamatan. Sementara itu, BBNI memperlihatkan pola yang relatif mirip, dengan tren kenaikan jangka panjang meskipun diiringi oleh beberapa periode koreksi tajam.

Pola ini menegaskan bahwa keempat saham perbankan besar Indonesia secara historis memiliki arah pertumbuhan positif, yang menunjukkan fundamental kuat sektor perbankan nasional. Tren kenaikan ini menjadi dasar penting bagi penerapan metode prediksi harga saham, karena meskipun ada volatilitas jangka pendek, kecenderungan jangka panjang tetap menunjukkan peningkatan nilai saham. Hal ini di pertegas oleh Table 1 tentang statistik dataset.

Table 1 Statistik Deskriptif harga penutupan saham BBCA, BBNI, BBRI, dan BMRI

Saham	Mean	Min	Max	Std Dev
BBCA	6403.56	2475.0	10950.0	2326.55
BBNI	3801.07	1580.0	6225.0	955.73
BBRI	3686.77	1731.8	6400.0	1062.32
BMRI	3988.19	1860.0	7450.0	1296.46

Tabel 1 menyajikan statistik deskriptif harga penutupan saham empat emiten perbankan besar di Indonesia. Saham BBCA memiliki rata-rata harga penutupan tertinggi, yaitu sekitar Rp6.403, dengan nilai maksimum mencapai Rp10.950. Hal ini menunjukkan posisi BBCA sebagai saham unggulan di sektor perbankan. Sementara itu, saham BBNI, BBRI, dan BMRI memiliki rata-rata harga penutupan yang relatif lebih rendah, yakni masing-masing Rp3.801, Rp3.687, dan Rp3.988. Dari sisi volatilitas, standar deviasi tertinggi juga dimiliki oleh BBCA (Rp2.326), menandakan adanya fluktuasi harga yang cukup besar, sejalan dengan tren kenaikan jangka panjang. Di sisi lain, BBNI, BBRI, dan BMRI menunjukkan standar deviasi yang lebih rendah, mencerminkan stabilitas relatif namun tetap mengalami fluktuasi signifikan pada periode tertentu. Data ini dimasukan sebagai inputan pada proses seleksi fitur menggunakan metode *correlation-driven proncipal component analysis (CD-PCA).*

Proses seleksi fitur dalam penelitian ini dilakukan dengan menggunakan metode Correlation-Driven Principal Component Analysis (CD-PCA). Berbeda dengan PCA tradisional yang hanya membangun matriks kovarians berdasarkan variansi murni, CD-PCA mengintegrasikan bobot korelasi absolut antara setiap variabel dengan target harga saham. Dengan demikian, komponen utama yang terbentuk tidak hanya merepresentasikan varian terbesar dalam data, tetapi juga menekankan relevansi terhadap variabel target. Transformasi CD-PCA menghasilkan 1.195 komponen utama, sesuai dengan jumlah variabel teknikal indikator yang tersedia dalam dataset awal. Setiap komponen utama merupakan kombinasi linear dari variabel asli dengan bobot yang ditentukan melalui korelasi, sehingga mampu mereduksi redundansi antar fitur tanpa kehilangan informasi penting. Data hasil CD-PCA dibagi menjadi 3 sekenario pembagian data, yaitu 70:30, 80:20, dan 90:10. Hal ini dapat dilihat pada Table 2.

Table 2 Pembagian data latih dan data uji saham BBCA

Rasio	Data Latih	Data Uji
70:30	1.649	708
80:20	1.885	472
90:10	2.121	236

Tabel 2 menunjukkan pembagian data latih dan data uji dengan tiga skenario rasio, yaitu 70:30, 80:20, dan 90:10. Walaupun tabel ini ditunjukkan pada saham BBCA, jumlah observasi pada saham BBNI, BBRI, dan BMRI sama sehingga hasil pembagian data pada dasarnya identik. Pada rasio 70:30, data latih terdiri dari 1.649 observasi dan data uji 708 observasi. Pada rasio 80:20, data latih meningkat menjadi 1.885 observasi dengan data uji 472 observasi. Sementara itu, pada rasio 90:10, data latih mencapai 2.121 observasi dengan data uji 236 observasi. Pendekatan ini digunakan untuk menguji konsistensi performa model pada berbagai skenario ketersediaan data.

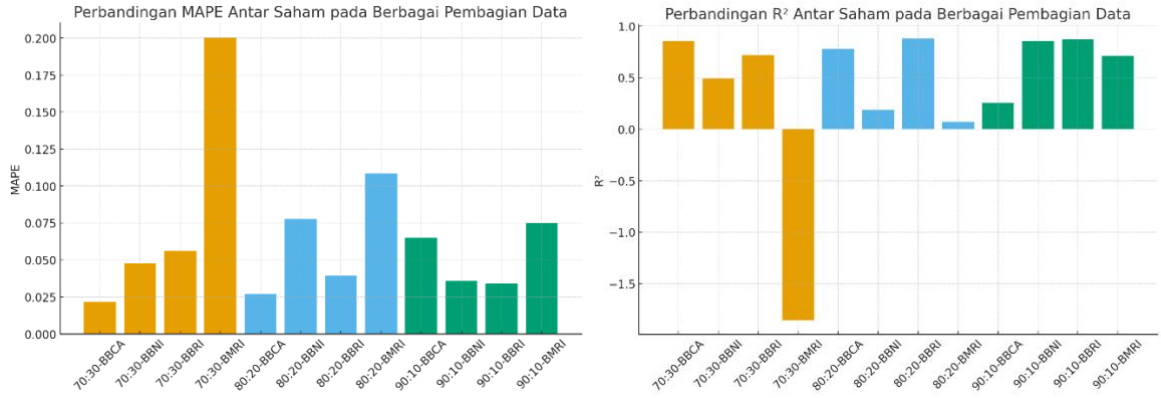
Table 3 Hasil Evaluasi Model Prediksi Harga Saham

Rasio	Saham	MSE	RMSE	MAE	MAPE	R^2
70:30	BBCA	0.000867	0.0294	0.0234	0.0217	0.8544
	BBNI	0.005627	0.0750	0.0525	0.0475	0.4874
	BBRI	0.005527	0.0743	0.0596	0.0562	0.7151
	BMRI	0.058985	0.2429	0.2143	0.2005	-1.8534
80:20	BBCA	0.001352	0.0368	0.0301	0.0268	0.7828

	BBNI	0.011666	0.1080	0.0865	0.0777	0.1893
	BBRI	0.002914	0.0540	0.0393	0.0395	0.8799
	BMRI	0.019783	0.1407	0.1219	0.1083	0.0724
90:10	BBCA	0.006791	0.0824	0.0678	0.0651	0.2528
	BBNI	0.001893	0.0435	0.0350	0.0356	0.8517
	BBRI	0.001381	0.0372	0.0303	0.0341	0.8708
	BMRI	0.007632	0.0874	0.0687	0.0750	0.7123

Hasil pada Tabel 3 menunjukkan variasi performa prediksi antar saham dan antar skenario pembagian data. Pada rasio 70:30, saham BBCA memberikan hasil terbaik dengan nilai MAPE yang sangat rendah (2,17%) dan R^2 sebesar 0,85, menandakan prediksi yang sangat mendekati data aktual. Sebaliknya, saham BMRI pada rasio ini mencatatkan nilai R^2 negatif (-,85), yang mengindikasikan performa model tidak optimal pada konfigurasi tersebut.

Pada skenario 80:20, saham BBRI menunjukkan performa paling stabil dengan R^2 sebesar 0,88 dan MAPE 3,95%. Sementara itu, performa BBNI dan BMRI relatif rendah dengan R^2 mendekati nol, menandakan bahwa model kesulitan menangkap pola harga pada skenario ini. Menariknya, pada rasio 90:10, BBNI dan BBRI mencapai hasil sangat baik, dengan nilai R^2 masing-masing 0,85 dan 0,87, serta MAPE di bawah 4%. Hal ini mengindikasikan bahwa kedua saham tersebut lebih responsif terhadap pembagian data uji yang lebih kecil. Di sisi lain, performa BBCA menurun (R^2 = 0,25), menunjukkan bahwa model memerlukan data latih yang lebih banyak untuk menjaga akurasi prediksi.



(a) MAPE (R^2)

Figure 4 Grafik perbandingan MAPE dan R2 antar saham

Figure 4 memperlihatkan perbandingan nilai Mean Absolute Percentage Error (MAPE) antar saham pada tiga skenario pembagian data. Terlihat bahwa saham BBCA konsisten memberikan tingkat kesalahan relatif yang paling rendah pada rasio 70:30 dan 80:20, masing-masing sebesar 2,17% dan 2,68%. Hal ini menunjukkan bahwa model mampu menangkap pola pergerakan harga saham BBCA dengan presisi tinggi ketika jumlah data latih cukup besar. Namun, pada skenario 90:10, nilai MAPE BBCA meningkat menjadi 6,51%, yang menandakan bahwa kinerja model menurun signifikan ketika data uji lebih sedikit, kemungkinan karena pola volatilitas jangka pendek lebih dominan pada periode tersebut.

Sebaliknya, saham BBNI dan BBRI memperlihatkan performa yang lebih baik pada skenario 90:10 dengan nilai MAPE masing-masing 3,56% dan 3,41%. Fenomena ini menunjukkan bahwa kedua saham tersebut memiliki pola harga yang lebih mudah

diprediksi ketika model diuji dengan sampel yang lebih kecil. Hal ini dapat diinterpretasikan sebagai bentuk kestabilan pola harga pada periode tertentu, yang membuat model mampu mengeneralisasi dengan baik meskipun data uji relatif terbatas. Sementara itu, BMRI menunjukkan kecenderungan berbeda, di mana nilai MAPE relatif tinggi pada semua skenario (10,83%–20,05%), menandakan adanya kompleksitas pola harga yang sulit ditangkap oleh model, atau kemungkinan adanya faktor eksternal yang memengaruhi volatilitas harga saham BMRI.

Figure 4 memperlihatkan distribusi nilai R^2 sebagai ukuran goodness-of-fit antar saham dan skenario pembagian data. Pada saham BBCA, nilai R^2 cukup tinggi pada skenario 70:30 (0,85) dan 80:20 (0,78), tetapi turun drastis pada 90:10 (0,25). Hal ini konsisten dengan kenaikan MAPE yang terlihat pada grafik pertama, sehingga dapat disimpulkan bahwa model prediksi untuk BBCA sangat bergantung pada ketersediaan data latih yang besar. Sebaliknya, saham BBNI dan BBRI justru mencatatkan nilai R^2 yang sangat baik pada rasio 90:10, masing-masing 0,85 dan 0,87, sehingga menunjukkan bahwa model lebih mampu menjelaskan variasi harga saham tersebut dalam kondisi data uji yang lebih kecil.

Performa BMRI kembali menjadi catatan penting, karena pada rasio 70:30 nilai R^2 justru negatif (-1,85), menandakan bahwa model prediksi lebih buruk dibandingkan baseline rata-rata sederhana. Walaupun nilai R^2 meningkat pada rasio 90:10 (0,71), tetap terlihat bahwa performa prediksi BMRI tidak sebaik tiga saham lainnya. Temuan ini dapat mengindikasikan bahwa pergerakan harga saham BMRI dipengaruhi oleh faktor volatilitas pasar atau dinamika fundamental perusahaan yang tidak sepenuhnya tercermin dalam indikator teknikal yang digunakan dalam penelitian ini.

Secara keseluruhan, interpretasi dari kedua grafik ini menegaskan bahwa integrasi Correlation-Driven PCA (CD-PCA) dan Temporal Attention-Based Deep Learning mampu menghasilkan performa prediksi yang kompetitif, terutama pada saham BBCA, BBNI, dan BBRI. Meskipun terdapat variasi kinerja antar skenario, kecenderungan utama yang dapat ditarik adalah bahwa model memiliki fleksibilitas dalam menangkap pola harga saham yang relatif stabil, sementara tantangan terbesar masih ditemukan pada saham dengan volatilitas tinggi seperti BMRI. Temuan ini memperlihatkan pentingnya pemilihan data latih dan uji yang tepat serta penyesuaian metode terhadap karakteristik saham yang dianalisis.

4. CONCLUSION

Penelitian ini menunjukkan efektivitas penggabungan Correlation-Driven Principal Component Analysis (CD-PCA) dengan Temporal Attention-Based Deep Learning dalam memprediksi harga saham perbankan besar di Indonesia (BBCA, BBNI, BBRI, dan BMRI). Metode CD-PCA berhasil mentransformasi 1.195 indikator teknikal menjadi komponen utama yang tidak hanya merepresentasikan variansi, tetapi juga mempertimbangkan korelasi dengan variabel target, sehingga mampu mereduksi redundansi tanpa kehilangan informasi yang relevan. Hasil eksperimen pada tiga skenario pembagian data (70:30, 80:20, dan 90:10) menunjukkan bahwa model yang diusulkan mampu mengurangi hasil error prediksi secara signifikan. Saham BBCA secara konsisten menunjukkan error yang rendah pada skenario dengan data latih lebih besar (70:30 dan 80:20), dengan nilai MAPE serendah 2,17% dan R^2 mencapai 0,85. Sebaliknya, BBNI dan BBRI justru memberikan hasil terbaik pada skenario 90:10, dengan nilai R^2 masing-masing 0,85 dan 0,87 serta MAPE di bawah 4%, yang mengindikasikan berkurangnya error meskipun data uji relatif kecil. Sementara itu,

BMRI masih menghadapi tantangan pada semua skenario, termasuk nilai R^2 negatif pada rasio 70:30, yang menunjukkan adanya volatilitas tinggi atau faktor eksternal lain yang belum sepenuhnya tertangkap oleh indikator teknikal. Secara keseluruhan, penelitian ini menegaskan bahwa integrasi CD-PCA dan Temporal Attention-Based Deep Learning berkontribusi pada pengurangan error prediksi harga saham yang konsisten, khususnya pada saham BBCA, BBNI, dan BBRI.

5. ACKNOWLEDGMENT

Terima kasih kepada Direktorat Jenderal Pendidikan Tinggi, Direktorat Jenderal Pendidikan Tinggi, Riset, dan Teknologi, Kementerian Pendidikan Tinggi, Ilmu Pengetahuan, dan Teknologi, atas dukungan dana melalui skema Penelitian Dosen Pemula (PDP) 2025 dengan nomor kontrak 0419/C3/DT.05.00/2025, 008/LL17/DT.05.00/PL/2025, dan 047/LPPM/KPD-DIKTI/VI/2025. Terima kasih kepada Lembaga Penelitian dan Pengabdian kepada Masyarakat Universitas Pahlawan Tuanku Tambusai yang telah memberikan bimbingan kepada dosen pemula dalam melaksanakan skema Penelitian Dosen Pemula..

6. AUTHORS' NOTE

The authors declare that there is no conflict of interest regarding the publication of this article. Authors confirmed that the paper was free of plagiarism.

7. REFERENCES

Al-Khasawneh, M. A., Raza, A., Khan, S. U. R., & Khan, Z. (2024). Stock Market Trend Prediction Using Deep Learning Approach. *Computational Economics*, *0123456789*. https://doi.org/10.1007/s10614-024-10714-1

Billah, M. M., Sultana, A., Bhuiyan, F., & Kaosar, M. G. (2024). Stock price prediction: comparison of different moving average techniques using deep learning model. *Neural Computing and Applications*, *36*(11), 5861–5871. https://doi.org/10.1007/s00521-023-09369-0

Botunac, I., Bosna, J., & Matetić, M. (2024). Optimization of Traditional Stock Market Strategies Using the LSTM Hybrid Approach. *Information*, *15*(3), 136. https://doi.org/10.3390/info15030136

Chen, C.-H., Chen, P.-Y., & Chun-Wei Lin, J. (2022). An Ensemble Classifier for Stock Trend Prediction Using Sentence-Level Chinese News Sentiment and Technical Indicators. *International Journal of Interactive Multimedia and Artificial Intelligence*, *7*(3), 53. https://doi.org/10.9781/ijimai.2022.02.004

Haryono, A. T., Sarno, R., & Sungkono, K. R. (2024). Stock price forecasting in Indonesia stock exchange using deep learning: a comparative study. *International Journal of Electrical and Computer Engineering (IJECE)*, *14*(1), 861. https://doi.org/10.11591/ijece.v14i1.pp861-869

He, X., & Wang, J. (2024). A Hybrid Forecasting System Based on Comprehensive Feature Selection and Intelligent Optimization for Stock Price Index Forecasting. *Mathematics*, *12*(23), 3778. https://doi.org/10.3390/math12233778

Ji, S. (2023). Application research of SGD algorithm based on PCA dimensionality reduction technique for stock price prediction. *2023 International Conference on Electronics and Devices, Computational Science (ICEDCS)*, 265–269. https://doi.org/10.1109/ICEDCS60513.2023.00054

Khanpara, P., Kadam, R., Lavingia, K., & Patel, S. (2023). Stock Trend Prediction: A Comparative Study using Different Approaches. *Proceedings - 5th International Conference on Smart Systems and Inventive Technology, ICSSIT 2023*, *Icssit*, 1697–1701. https://doi.org/10.1109/ICSSIT55814.2023.10060936

Li, Z., Lang, L., Sun, G., Cai, Z., & Luo, Z. (2023). Enhancing Multiple Linear Regression for Price Prediction: A PCA-Integrated Approach. *2023 4th International Conference on Computer, Big Data and Artificial Intelligence (ICCBD+AI)*, 337–341. https://doi.org/10.1109/ICCBD-AI62252.2023.00063

Mansoor, Z. K., & Al-Sultan, A. Y. (2025). Enhancing Stock Price Prediction Using Mutual Information, PCA, and LSTM: A Deep Learning Approach. *Fusion: Practice and Applications*, *17*(1), 196–208. https://doi.org/10.54216/FPA.170114

Mi, Y., Xu, D., & Gao, T. (2023). Application of PCA-LSTM algorithm for financial market stock return prediction and optimization model. *International Journal for Simulation and Multidisciplinary Design Optimization*, *14*(8), 1–6. https://doi.org/10.1051/smdo/2023009

Ming, L., & Chen, G. (2024). Stock Price Prediction Based on Relative Strength Index, Moving Average Convergence Divergence and XGBoost Model. *2024 IEEE 13th Data Driven Control and Learning Systems Conference (DDCLS)*, 1988–1993. https://doi.org/10.1109/DDCLS61622.2024.10606580

Priyatno, A. M., Ningsih, L., & Noor, M. (2024). Harnessing Machine Learning for Stock Price Prediction with Random Forest and Simple Moving Average Techniques. *Journal of Engineering and Science Application*, *1*(1), 1–8. https://doi.org/10.69693/jesa.v1i1.1

Priyatno, A. M., Ramadhan Sudirman, W. F., & Musridho, R. J. (2024). Feature selection using non-parametric correlations and important features on recursive feature elimination for stock price prediction. *International Journal of Electrical and Computer Engineering (IJECE)*, *14*(2), 1906. https://doi.org/10.11591/ijece.v14i2.pp1906-1915

Priyatno, A. M., Sudirman, W. F., Musridho, R. J., & Amalia, F. (2023). Impurity-Based Important Features for feature selection in Recursive Feature Elimination for Stock Price Forecasting. *Jurnal Teknik Industri Terintegrasi*, *6*(4), 1182–1194. https://doi.org/10.31004/jutin.v6i4.17726

S, J., & Arya, K. (2024). Stock Movement of Wipro Using Technical Analysis. *2024 10th International Conference on Advanced Computing and Communication Systems (ICACCS)*, *1*, 1033–1038. https://doi.org/10.1109/ICACCS60874.2024.10717324

Sheng, Y., Fu, K., & Wang, L. (2022). A PCA-LSTM Model for Stock Index Forecasting: A Case Study in Shanghai Composite Index. *2022 7th International Conference on Cloud Computing and Big Data Analytics (ICCCBDA)*, 412–417. https://doi.org/10.1109/ICCCBDA55098.2022.9778913

Uçkan, T. (2024). Integrating PCA with deep learning models for stock market Forecasting:

An analysis of Turkish stocks markets. *Journal of King Saud University - Computer and Information Sciences*, *36*(8), 102162. https://doi.org/10.1016/j.jksuci.2024.102162

Uma, K. S., & Srinath Naidu. (2021). Prediction of Intraday Trend Reversal in Stock Market Index Through Machine Learning Algorithms. In *Advances in Intelligent Systems and Computing: Vol. 1200 AISC* (pp. 331–341). https://doi.org/10.1007/978-3-030-51859-2_30

Wang, J., Liu, D., Jin, L., Sun, Q., & Xue, Z. (2023). A PCA-IGRU Model for Stock Price Prediction. *Journal of Internet Technology*, *24*(3), 621–629. https://doi.org/10.53106/160792642023052403007

Zhu, C., Lu, P., Feng, W., & Wang, Y. (2025). Bimodal Stock Price Prediction Based on Holt-Winters Exponential Smoothing and PCA Whitening Transformation. *IAENG International Journal of Computer Science*, *52*(1), 187–200.