

OPTIMIZATION OF STOCK TRADING STRATEGIES USING A HYBRID REINFORCEMENT LEARNING AND FORECASTING MODEL

OPTIMISASI STRATEGI PERDAGANGAN SAHAM MENGGUNAKAN MODEL HIBRIDA REINFORCEMENT LEARNING DAN FORECASTING

Rezha Ikhwan Hidayat¹, Anggit Dwi Hartanto²

Universitas Amikom, Yogyakarta, Indonesia

rezhaikwanh@students.amikom.ac.id¹, anggit@amikom.ac.id²

Abstract - Stock price prediction is an interesting challenge in machine learning due to the non-linear nature of the market. Although forecasting models can predict prices, they often do not provide optimal trading strategies. Reinforcement learning (RL) has the potential to optimize strategies, but it is highly dependent on the input states. This study integrates two methods—a CNN-LSTM forecasting model and RL (A3C)—to develop an algorithmic trading strategy. The model is evaluated using historical INDF stock data (2016–2024) with a data-split validation protocol of 80% training and 20% testing. Backtesting simulations on the period (Feb 2023–Dec 2024) show that the hybrid model achieves a cumulative total return of 121.44%. This result was obtained using an all-in trading strategy (one full position at a time) and includes transaction costs: a trading fee of 0.01% per transaction and a borrow interest rate of 0.0003% per day for short positions. This performance significantly outperforms traditional strategies: Buy and Hold (23.45%), MA Crossover (51.13%), RSI (9.09%), and MACD (–29.08%). The hybrid model also achieves a Sharpe Ratio of 2.381 (annualized, assuming a 0% risk-free rate).

Keywords - Stock Price Prediction, Algorithmic Trading, Hybrid Model, Reinforcement Learning, CNN-LSTM

Abstrak - Prediksi harga saham merupakan tantangan menarik dalam machine learning karena sifat pasar yang non-linear. Meskipun model peramalan dapat memprediksi harga, mereka seringkali tidak memberikan strategi perdagangan yang optimal. *Reinforcement learning* (RL) berpotensi mengoptimalkan strategi, namun sangat bergantung pada *state* yang dimasukkan. Penelitian ini mengintegrasikan dua metode—model peramalan CNN-LSTM dan RL (A3C)—untuk mengembangkan strategi perdagangan algoritmik. Model dievaluasi menggunakan data historis saham INDF (2016–2024) dengan protokol validasi pembagian data 80% latih dan 20% uji. Simulasi backtesting di uji pada (Feb 2023–Des 2024) menunjukkan model hibrida mencapai total return kumulatif 121,44%. Hasil ini dicapai dengan strategi perdagangan all-in (satu posisi penuh per waktu) dan telah memperhitungkan biaya transaksi biaya perdagangan trading fees sebesar 0.01% per transaksi dan biaya pinjaman borrow interest rate sebesar 0.0003% per hari untuk posisi short. Kinerja ini secara signifikan melampaui model A3C murni (60,79%) ini membuktikan kontribusi dari komponen *forecasting* serta *baseline* standar lainnya, termasuk Buy and Hold (23,45%), MA Crossover (51,13%), RSI (9,09%), dan MACD (-29,08%).

Kata Kunci - Prediksi Harga Saham, Perdagangan Algoritmik, Model Hibrida, *Reinforcement Learning*, CNN-LSTM

I. PENDAHULUAN

Saham merupakan instrumen pasar keuangan yang paling populer di kalangan masyarakat saat ini, dikarenakan saham mampu memberikan keuntungan yang menarik[1]. Meramalkan harga saham merupakan kegiatan yang menarik bagi para investor dan peneliti yang ingin mengalahkan pasar saham, Akan tetapi, meramalkan harga saham merupakan hal yang sulit[2]. Namun, Banyak penelitian telah dilakukan untuk memprediksi harga saham dengan tujuan memperoleh keuntungan dengan menggunakan berbagai teknik mulai dari tradisional sampai deep learning[3], [4], [5], [6]. Pentingnya menentukan waktu pasar yang tepat untuk strategi perdagangan [7] telah mendorong pengembangan metode peramalan akurat dan strategi adaptif, yang kini menjadi area penelitian krusial di bidang keuangan komputasional Pendekatan *forecasting* telah berevolusi dari model statistik tradisional seperti *linear regression*, *decision tree*[8], ARIMA[9] yang memiliki keterbatasan dalam menangani data non-linear [10]. beralih ke model *deep learning* seperti LSTM,CNN,dan GRU[11], yang lebih unggul dalam menangkap pola tersebut[12]. Namun, model-model ini hanya menghasilkan prediksi harga dan tidak memberikan rekomendasi aksi perdagangan atau strategi yang konkret.

Reinforcement Learning (RL) adalah metode optimasi strategi di mana agen secara otonom belajar mengambil keputusan untuk memaksimalkan imbalan [13]. Namun, kinerja agen RL sangat bergantung pada kualitas *input state representation*. Jika hanya mengandalkan data harga historis mentah, agen mungkin kesulitan mengekstraksi sinyal yang berguna, sehingga proses pembelajarannya menjadi tidak efisien[14]. Penelitian ini bertujuan merancang dan mengevaluasi model hibrida yang mengintegrasikan *forecasting* (CNN-LSTM) dengan *Reinforcement Learning* (A3C). Pendekatan ini menguji hipotesis bahwa dengan memperkaya *state* agen RL menggunakan wawasan prediktif, model hibrida akan secara signifikan mengungguli *baseline* standar *Buy and Hold*, MA, RSI, MACD dalam hal *total return* dan *Sharpe Ratio*. Studi kasus menggunakan saham PT. Indofood (INDF). Saham ini dipilih karena merupakan *blue chip* anggota LQ45 yang sangat likuid. Karakteristik harganya yang relatif stabil namun tetap volatil menjadikannya *dataset* ideal untuk menguji ketahanan model hibrida dalam membedakan sinyal valid dari *noise* pasar.

II. SIGNIFIKANSI STUDI

A. Studi Literature

Prediksi harga saham telah berevolusi dari metode tradisional seperti ARIMA[9] dan regresi linear[8] yang memiliki keterbatasan dalam menangani data *non-linear*[10], beralih ke model *deep learning* (DL) seperti LSTM dan GRU[11]. Meskipun model DL unggul dalam menangkap pola *non-linear*[12], fokus utamanya terbatas pada prediksi harga dan tidak memberikan rekomendasi aksi perdagangan. Sebagai solusi, *Reinforcement Learning* (RL)[13] digunakan untuk melatih agen mengambil keputusan optimal (beli, jual, atau tahan) guna memaksimalkan imbalan. Namun, kinerja RL sangat bergantung pada kualitas input atau representasi keadaan (*state representation*)[14]. Menggunakan data harga historis mentah saja terbukti kurang efisien. Oleh karena itu, penelitian modern berfokus pada pengayaan *state* agen RL dengan fitur-fitur yang lebih informatif, seperti indikator teknikal misalnya, MACD[4], [15] dan RSI[6], serta data eksternal seperti sentimen publik[3], [17], untuk menghasilkan strategi perdagangan yang lebih efektif dan menguntungkan[18].

B. Data

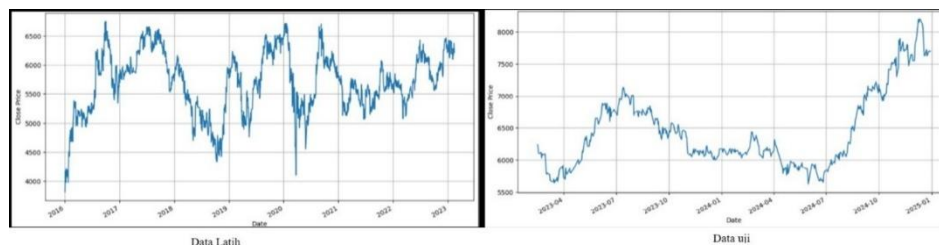
1. *Pengumpulan Data*, Data yang digunakan dalam penelitian ini bersumber dari Yahoo Finance. Data yang digunakan bernama INDF.JK (Indofood Sukses Makmur) dengan time frame harian dari rentang tahun 2016 sampai 2024. Data ini menggunakan Adjusted Close, yang berarti data harga telah disesuaikan secara otomatis untuk stock split dan dividen guna memastikan kontinuitas data. Struktur data mentah dari Yahoo Finance, beserta beberapa fitur teknikal awal yang diekstraksi, dirinci pada Tabel I.

TABEL I
NAMA KOLOM DATA

Nama kolom	Keterangan
Date	Tanggal pencatatan data perdagangan.
Open	Harga pembukaan saham pada hari perdagangan tersebut.
High	Harga tertinggi yang dicapai saham pada hari perdagangan tersebut
Low	Harga terendah yang dicapai saham pada hari perdagangan tersebut.
Close	Harga penutupan saham pada hari tersebut
Volume	Jumlah total lembar saham yang diperdagangkan pada hari tersebut, menunjukkan tingkat aktivitas dan likuiditas pasar.
Relative Strength Indeks (RSI)	Indikator momentum yang mengukur kecepatan dan perubahan pergerakan harga untuk mengidentifikasi kondisi jenuh beli (<i>overbought</i>) atau jenuh jual (<i>oversold</i>).
MA_20	Rata-rata harga penutupan saham selama 20 hari terakhir, digunakan untuk menghaluskan data harga dan mengidentifikasi tren jangka pendek.
MA_50	Rata-rata harga penutupan saham selama 50 hari terakhir, sering digunakan untuk mengidentifikasi tren jangka menengah hingga panjang.
MACD	Indikator momentum yang mengikuti tren dengan menunjukkan hubungan antara dua rata-rata pergerakan harga eksponensial (EMA)
Relative Strength Indeks (RSI)	Indikator momentum yang mengukur kecepatan dan perubahan pergerakan harga untuk mengidentifikasi kondisi jenuh beli (<i>overbought</i>) atau jenuh jual (<i>oversold</i>).

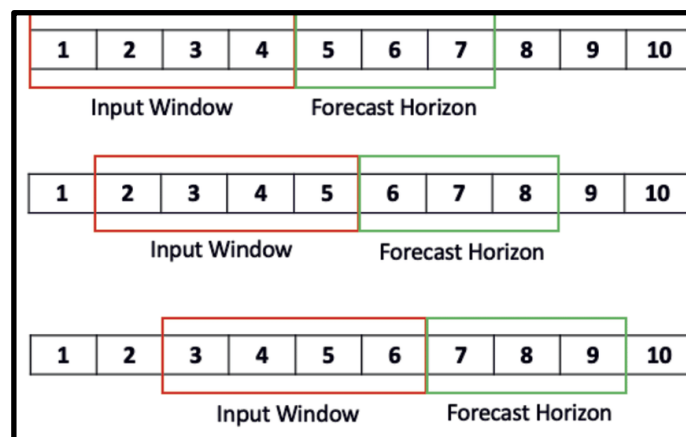
Selanjutnya, data ini melalui beberapa langkah pra-pemrosesan untuk membuat data lebih sesuai dengan model dan juga memperkaya fitur dari data yang akan digunakan. Berikut adalah langkah-langkah pra-pemrosesan data.

2. *Splitting Data* data di bagi ke dalam 2 tipe yaitu data latih dan juga data validasi dengan Ratio pembagian 80/20. 80% untuk data latih dan 20% untuk data validasi. Gambar 1 menampilkan plot data latih dan data validasi



Gambar 1. Plot data

3. *Windowing Data*: Dalam proses ini, data diambil sesuai dengan ukuran Window yang ditentukan. Tahapan ini bertujuan menyiapkan data agar memiliki bentuk yang sesuai dengan kebutuhan input model. Karena pada penelitian ini digunakan 60 data deret sebagai input dan 10 data berikutnya sebagai target prediksi, maka setiap sampel akan terdiri atas 70 data berurutan lalu dibagi 60 buat input dan 10 buat target.



Gambar 2. Simulasi windowing data

4. *Normalisasi dan fitur ekstraksi perwindow*, Setelah proses windowing, setiap sampel data terdiri dari window input (60 hari). Sebuah proses pra-pemrosesan dua Langkah ekstraksi fitur dan normalisasi diterapkan secara independen pada setiap window input. Normalisasi adalah proses menstandarkan skala data. Metode normalisasi yang digunakan dalam penelitian ini adalah Z-score, yaitu metode yang mengubah skala data berdasarkan nilai rata-rata dan standar deviasinya. Model neural network dapat belajar lebih baik jika data dinormalisasi[19].

$$Z = \frac{X - \mu}{\sigma} \quad (1)$$

Z: Skor Z (berapa jauh nilai dari rata-rata)

X: Nilai yang diamati

μ : Rata-rata

σ : standar deviasi

Fitur ekstraksi ini adalah langkah melakukan perhitungan ulang dari fitur yang sudah tersedia untuk menciptakan fitur baru. Ini adalah langkah yang penting untuk membuat model lebih mudah mengenali pola pada data. Data yang digunakan untuk fitur ekstraksi adalah *Close*, data ini di ubah menjadi fitur baru yaitu:

- Persen dari harga tertinggi: adalah persentase dari harga tertinggi pada data formula nya sebagai berikut:

$$P = \left(\frac{c}{m} \right) \times 100 \quad (2)$$

P : Percent From max

c : Close price

m : max price

- *Volatility* : tingkat fluktuasi atau perubahan harga suatu data dalam periode waktu tertentu formula nya sebagai berikut:

$$V = \frac{c - m}{r} \quad (3)$$

V : Volatility price

C : close price

m : min price

r : range price

- Ratio harga: adalah ratio harga penutupan saat ini berdasarkan harga penutupan pada awal data formula nya

$$R = \frac{c}{c_1} \quad (4)$$

R : Ratio harga

c : harga penutupan

c_1 : harga penutupan awal data

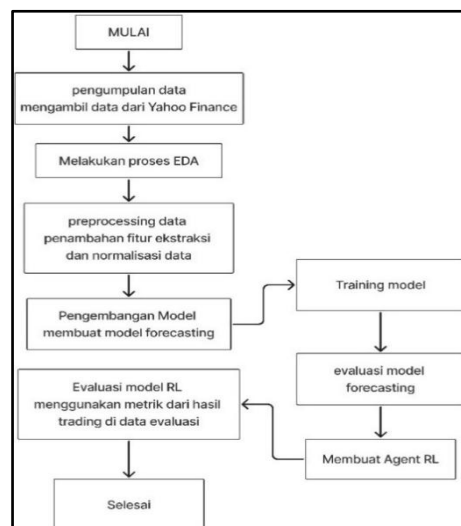
Pendekatan ini menghasilkan data yang bersifat relatif terhadap karakteristik data lokal. Sehingga *input* yang dihasilkan untuk model memiliki fitur-fitur yang di deskripsikan pada Tabel II.

TABEL II
DAFTAR FITUR MODEL PERAMALAN

fitur	Keterangan
Volume	Jumlah total lembar saham yang diperdagangkan pada hari tersebut, menunjukkan tingkat aktivitas dan likuiditas pasar.
Relative Strength Indeks (RSI)	Indikator momentum yang mengukur kecepatan dan perubahan pergerakan harga untuk mengidentifikasi kondisi jenuh beli (<i>overbought</i>) atau jenuh jual (<i>oversold</i>).
MA_20	Rata-rata harga penutupan saham selama 20 hari terakhir, digunakan untuk menghaluskan data harga dan mengidentifikasi tren jangka pendek.
MA_50	Rata-rata harga penutupan saham selama 50 hari terakhir, sering digunakan untuk mengidentifikasi tren jangka menengah hingga panjang.
MACD_line	garis utama dalam indikator MACD yang menunjukkan selisih antara rata-rata bergerak eksponensial
MACD_signal	rata-rata bergerak eksponensial (EMA) 9 hari dari MACD Line. Garis ini digunakan sebagai acuan sinyal
MACD_hist	menunjukkan selisih antara MACD Line dan MACD Signal
Persen form max	persentase dari harga max
Volatilitas	tingkat fluktuasi atau perubahan harga
Ratio	ratio adalah ratio harga penutupan saat ini berdasarkan harga penutupan awal

C. Metode Penelitian dan Evaluasi

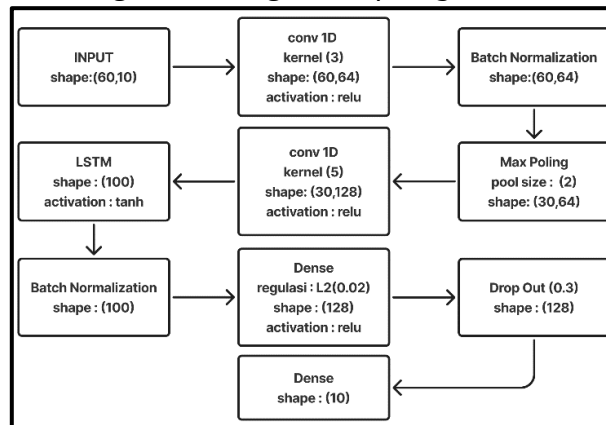
Penelitian ini dilaksanakan melalui beberapa tahapan utama, yang diawali dengan pengumpulan dan pra-pemrosesan data historis saham. Selanjutnya, dilakukan perancangan arsitektur untuk model forecasting CNN-LSTM dan model Reinforcement learning A3C. Kedua model tersebut kemudian diintegrasikan menjadi sebuah sistem hibrida. Tahap terakhir adalah pengujian dan evaluasi kinerja model hibrida dengan membandingkannya terhadap model pembanding menggunakan metrik evaluasi. Gambaran alur penelitian di tunjukan pada Gambar 3.



Gambar 3. Alur penelitian

1. Model Peramalan: CNN-LSTM

Data deret waktu finansial memiliki karakteristik unik yang mengandung pola spasial (fitur lokal dalam jendela waktu) dan dependensi temporal (tren jangka panjang) secara bersamaan. Untuk menangkap kedua karakteristik tersebut, model CNN-LSTM diadopsi dalam penelitian ini[2], [5], [11]. Penggunaan lapisan konvolusional (CNN) memungkinkan model untuk mengidentifikasi fitur-fitur penting dalam data mentah, sementara lapisan Long Short-Term Memory (LSTM) bertugas mempelajari hubungan sekuensial antar fitur tersebut. Sinergi ini memungkinkan model membangun pemahaman yang lebih komprehensif terhadap dinamika pasar dibandingkan dengan menggunakan arsitektur tunggal. Arsitektur model ini secara spesifik dirancang untuk memproses data masukan dengan bentuk (shape) (60, 10), yang merepresentasikan 60 hari data historis dengan 10 fitur. Setiap komponen arsitektur, mulai dari aliran data, perubahan shape tensor pada tiap lapisan, hingga detail *hyperparameter* seperti ukuran kernel (3, 5), pool size (2), jumlah unit LSTM (100), dropout rate (0.3), serta nilai regularisasi L2 (0.02) telah dijelaskan secara visual pada Gambar 4. Model ini dikompilasi menggunakan *optimizer* Adam (*learning rate* 0.001) dan fungsi *loss* kustom (dijelaskan selanjutnya), serta menggunakan *callback* ModelCheckpoint untuk menyimpan bobot terbaik berdasarkan *validation loss* guna mencegah *overfitting*.



Gambar 4. Arsitektur Model Peramalan

Secara umum fungsi kerugian yang digunakan adalah MSE akan tetapi ada sedikit perubahan, hasil keluaran model tidak langsung di hitung dengan target atau label dari data. Keluaran model akan di proses dengan fungsi *Exponentiated cumulative log return* dan akan di rekonstruksi ke harga. Hasil dari perhitungan tersebut yang akan di kalkulasi dengan fungsi MSE.

$$\hat{y} = y_0 \cdot e^{\sum_{i=1}^t r_i} \quad (5)$$

Keterangan:

y_0 : harga awal (start price)

r_t : *log return* yang diprediksi

$\hat{y}$: harga hasil rekonstruksi dari *log return* kumulatif

T : jumlah langkah waktu

$$MSE = \frac{1}{n} \sum_{i=1}^n (\hat{y} - y_i) \quad (6)$$

Keterangan:

y_i : harga target

$\hat{y}$: harga hasil rekonstruksi dari *log return* kumulatif

n : jumlah langkah waktu

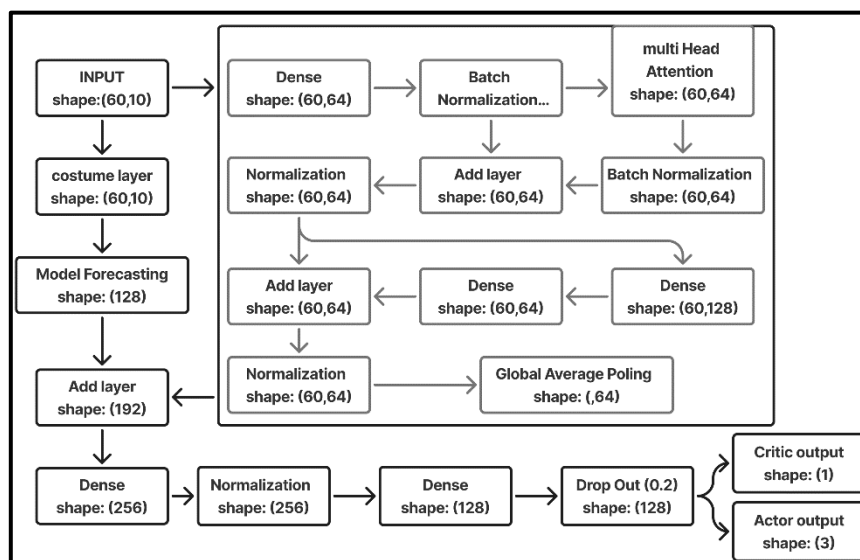
Tujuan model ini adalah untuk memprediksi harga saham selama 10 hari ke depan, akan tetapi model tidak memprediksi harga secara langsung, model memprediksi nilai *log return* 10 hari ke depan.

Hasil fitur ekstraksi dari model peramalan ini akan di jadikan data tambahan untuk model *reinforcement learning*.

2. Model Reinforcement learning: A3C

Model pengambilan keputusan menggunakan *Asynchronous Advantage Actor-Critic (A3C)*[21], sebuah algoritma *Reinforcement Learning (RL)* canggih. A3C menggabungkan fungsi *Advantage* untuk mengurangi varians dan menstabilkan pelatihan dengan skema pembelajaran asinkron agen paralel memperbarui model global secara efisien, sehingga sangat sesuai untuk optimasi strategi perdagangan saham.

Arsitektur model:



Gambar 5. Arsitektur Model RL

Arsitektur model A3C hybrid di gambar kan pada Gambar. 5 dirancang sebagai jaringan feature extractor dua-jalur yang memproses input state (60, 10) secara paralel. Jalur 1 (Prediktif) melewati input ke custom layer (normalisasi Z-score per-window, seperti dijelaskan di Bagian data normalisasi dan fitur ekstraksi) lalu ke Model Forecasting untuk menghasilkan vektor fitur (128). Jalur 2 memproses input yang sama melalui blok Dense/Multi-Head Attention dan Global Average Pooling untuk menghasilkan vektor kedua (64). Kedua vektor fitur ini kemudian digabungkan (concatenate) menjadi vektor hibrida (192), yang diproses melalui lapisan Dense (256, 128) dengan Normalisasi dan Dropout, sebelum akhirnya diumpankan ke dua head keluaran: Kritik (1) dan Aktor (3). adalah:

Environment:

Untuk melatih agen, penelitian ini memanfaatkan *environment* simulasi *gym-trading-env* (OpenAI Gym) yang dikonfigurasi pada *dataset* INDF. Lingkungan ini disetel secara realistis dengan **biaya perdagangan (trading fees)** sebesar 0.01% per aksi Beli/Jual dan **bunga pinjaman (borrow interest rate)** sebesar 0.0003% per hari untuk posisi *short*. **State Space** didefinisikan sebagai matriks (60, 10) yang mencakup 8 Fitur Pasar (Close, Volume, RSI, MA, MACD) dan 2 Fitur Dinamis Agen (posisi terakhir, portofolio). **Action Space** yang tersedia adalah [0, 1, 2] (Sell, Hold, Buy), dan **Reward Function** dihitung menggunakan *log return* portofolio (Formula 7) di bahas lebih lanjut di bagian selanjut nya.

$$r_t = \ln \left(\frac{p_t}{p_{t-1}} \right) \quad (7)$$

Keterangan:

r_t =reward pada langkah waktu
 p_t = portofolio valuasi pada langkah waktu

Hiperparameter:

Pemilihan *hiperparameter* merupakan langkah penting yang sangat memengaruhi kinerja model dan proses konvergensi dari model A3C. Nilai-nilai *hiperparameter* yang digunakan dalam eksperimen penelitian ini, yang dipilih berdasarkan serangkaian uji coba awal, disajikan pada Tabel 3.

TABEL III.
DAFTAR HIPERPARAMETER

Hiperparameter	Nilai
Learning rate	1e-5
Ukuran <i>Window Historis</i>	60
gama	0.85
Rollout (n-steps)	1
Entropy coefficient	0.06
Metode Gradient Clip	Global Norm(tf.clip_by_global_norm)
Gradient clip	40.0
Strategi Konvergensi	Berhenti di akhir
Total Timesteps	150.000
Frekuensi Checkpoint	1.000 step

3. Skenario Pengujian dan Metrik Evaluasi

Pengujian menggunakan dataset yang telah di siapkan menggunakan data validasi untuk menguji seberapa baik strategi yang diambil oleh model. Kinerja model dievaluasi menggunakan beberapa metrik standar industri yang perhitungannya dijabarkan di bawah ini.

1. **Total Return (Profit Multiplier):** Dihitung sebagai rasio nilai akhir portofolio terhadap nilai awalnya.
2. **Max Drawdown (MDD):** Mengukur kerugian maksimum dari puncak ke lembah (*peak-to-trough*) selama periode pengujian untuk mengukur risiko kerugian terbesar.
3. **Sharpe Ratio (Annualized):** Mengukur *return* yang disesuaikan dengan risiko. Rasio ini dihitung dari *return* harian dengan asumsi *Risk-Free Rate* (RFR) sebesar 0%, kemudian di analisis menggunakan faktor pengali $\sqrt{252}$.
4. **Sortino Ratio (Annualized):** Mirip dengan Sharpe, namun hanya menghukum *downside volatility*. *Downside deviation* didefinisikan sebagai standar deviasi dari *return* harian yang bernilai negatif (di bawah 0%). Rasio ini juga di analisis dengan faktor $\sqrt{252}$ dan RFR 0%.
5. **Calmar Ratio:** Mengukur *return* yang disesuaikan dengan risiko *drawdown* maksimum. Metrik ini dihitung dengan membagi *Return Tahunan (Annualized Return)* dengan nilai absolut dari *Max Drawdown* (MDD). *Return Tahunan* dihitung dengan menganalisis *Total Return* portofolio berdasarkan panjang hari perdagangan standar (252 hari).

Untuk mengukur keunggulan model hibrida yang diusulkan, kinerjanya akan dibandingkan dengan beberapa strategi yang umum dalam perdagangan saham seperti *buy and hold*, *MA crossover*, *RSI*, dan *MACD* [16], [22], sebagai model pembanding (*baseline*) berikut aturannya:

1. Strategi *buy and hold*: Strategi pasif di mana aset dibeli pada hari pertama periode pengujian dan ditahan hingga hari terakhir
2. *Moving Average (MA) Crossover*: Strategi ini menggunakan perpotongan antara dua indikator *Moving Average* untuk menghasilkan sinyal perdagangan, yaitu MA jangka pendek (misalnya, MA 20 hari) dan MA jangka panjang (misalnya, MA 50 hari).

Sinyal Beli (Golden Cross): Dihasilkan ketika garis MA jangka pendek memotong ke atas garis MA jangka panjang. Hal ini mengindikasikan potensi dimulainya tren naik (*uptrend*).

Sinyal Jual (Death Cross): Dihasilkan ketika garis MA jangka pendek memotong ke bawah garis MA jangka panjang, yang mengindikasikan potensi dimulainya tren turun (*downtrend*).

3. **RSI strategi:** Strategi ini memanfaatkan indikator momentum RSI untuk mengidentifikasi kondisi pasar yang jenuh beli (*overbought*) atau jenuh jual (*oversold*). Aturan dasarnya adalah:

Sinyal Beli: Dihasilkan ketika nilai RSI bergerak naik dari area *oversold* (misalnya, melewati batas 30 dari bawah ke atas). Ini menandakan bahwa tekanan jual mungkin telah berakhir.

Sinyal Jual: Dihasilkan ketika nilai RSI bergerak turun dari area *overbought* (misalnya, melewati batas 70 dari atas ke bawah), yang menandakan kemungkinan berakhirnya tren beli.

4. **Strategi MACD (Moving Average Convergence Divergence):** Strategi ini didasarkan pada hubungan antara garis MACD dan garis sinyalnya untuk mengidentifikasi perubahan momentum.

Sinyal Beli: Dihasilkan ketika garis MACD memotong ke atas garis sinyalnya. Perpotongan ini dianggap sebagai sinyal *bullish* yang menunjukkan momentum positif.

Sinyal Jual: Dihasilkan ketika garis MACD memotong ke bawah garis sinyalnya, yang dianggap sebagai sinyal *bearish* yang menunjukkan momentum negatif.

III. HASIL DAN PEMBAHASAN

Eksperimen menunjukkan bahwa model hibrida ini secara konsisten memiliki kinerja yang lebih unggul dibanding strategi pembanding. Model hibrida (A3C + CNN-LSTM) mencapai tingkat profitabilitas dan rasio risiko-imbal hasil yang jauh lebih tinggi. Untuk mengisolasi kontribusi setiap komponen, studi ablasi juga dilakukan dengan membandingkan model hibrida penuh terhadap model A3C Saja (tanpa jalur CNN-LSTM). Hasil pengujian lengkap dari semua model disajikan pada Tabel IV.

TABEL IV
HASIL PENGUJIAN

Model	Profit(%)	Max Drawdown	Sharpe Ratio	Sortino Ratio	Calmar Ratio
Hybrid	121.46	-0.133	2.381	3.919	4.258
A3C	60.79	-0.096	1.657	2.646	3.210
Buy and hold	23.44	-0.211	0.697	1.076	0.599
MA crossover	51.13	-0.081	1.314	2.184	3.249
RSI	9.09	-0.050	0.977	0.541	1.009
MACD	-29.08	-0.451	-0.880	-1.297	-0.391

Analisis terhadap Tabel V mengonfirmasi hipotesis utama penelitian ini. Poin terpenting datang dari studi ablasi, yang secara kuantitatif membuktikan kontribusi signifikan dari arsitektur hibrida. Model Hibrida (Profit 121.46%, Sharpe 2.381) secara substansial mengungguli model 'A3C Saja' (Profit 60.79%, Sharpe 1.657). Kesenjangan kinerja ini membuktikan bahwa sinergi dari modul *forecasting* CNN-LSTM (Jalur 1) berhasil memperkaya *state representation* agen, memungkinkannya mengambil keputusan yang lebih strategis, sejalan dengan temuan [17] dan [18]. Perbandingan dengan *baseline* konvensional juga menyoroti keunggulan model. Strategi *rule-based* seperti MACD gagal beradaptasi dan merugi (-29.08%), mengonfirmasi sifat *lagging*-nya [15]. Sementara itu, strategi RSI menunjukkan profil paling konservatif (MDD -0.050, Profit 9.09%). Model hibrida, sebaliknya, mampu menyeimbangkan risiko dan imbal hasil secara superior, dibuktikan dengan

pencapaian Sortino (3.919) dan Calmar Ratio (4.258) tertinggi, yang mengindikasikan manajemen *downside risk* yang cerdas untuk mencapai profitabilitas tinggi.

KESIMPULAN

Hasil penelitian menunjukkan bahwa model hibrida secara konsisten memiliki kinerja yang lebih unggul, mencapai Total Return 121.44% dan Sharpe Ratio (annualized) 2.381. Kinerja ini secara drastis melampaui strategi *baseline* konvensional seperti *Buy and Hold* Total Return 23.44% dan juga terbukti secara signifikan lebih baik daripada model A3C Profit 60.79% berdasarkan studi ablas. Temuan ini mengindikasikan bahwa integrasi wawasan prediktif dari model *forecasting* mampu memperkaya informasi pada model RL, sehingga menghasilkan strategi perdagangan yang lebih optimal, adaptif, dan menguntungkan. Namun, penelitian ini memiliki beberapa batasan utama yang harus ditegaskan. Pertama, generalisasi model masih belum teruji karena *dataset* terbatas pada satu saham (INDF.JK). Selain itu, *state* agen belum mencakup faktor eksternal seperti berita, sentimen pasar, atau *chart pattern*, dan evaluasi kinerja juga masih terbatas pada lingkungan simulasi. Secara metodologis, terdapat pula keterbatasan bahwa model belum divalidasi melalui uji signifikansi statistik seperti *paired bootstrap test* atau *Model Confidence Set* untuk memastikan keunggulan ini. Terakhir, strategi normalisasi Z-score per-window yang digunakan, meskipun valid secara kausal, berpotensi tidak optimal untuk menangkap tren jangka panjang.

Oleh karena itu, penelitian selanjutnya disarankan untuk fokus pada rencana eksperimen yang lebih spesifik dan akademis. Langkah prioritas mencakup menerapkan metode validasi yang lebih kuat seperti validasi *walk-forward*, yang lebih mencerminkan kondisi pasar yang terus berubah. Penting juga untuk menggunakan Model Confidence Set (MCS) atau *Deflated Sharpe Ratio (DSR)* untuk menguji signifikansi statistik dari strategi yang ditemukan. Setelah validasi ini selesai, *state* agen dapat diperkaya lebih lanjut dengan mengintegrasikan analisis sentimen dari berita dan pola chart sebagai *input state* tambahan.

REFERENSI

- [1] V. P. Ramadhan and F. Y. Pamuji, "Analisis Perbandingan Algoritma Forecasting dalam Prediksi Harga Saham LQ45 PT Bank Mandiri Sekuritas (BMRI)," *J. Teknol. dan Manaj. Inform.*, vol. 8, no. 1, pp. 39–45, 2022, doi: 10.26905/jtmi.v8i1.6092.
- [2] T. Kim and H. Y. Kim, "Forecasting stock prices with a feature fusion LSTM-CNN model using different representations of the same data," *PLoS One*, vol. 14, no. 2, pp. 1–23, 2019, doi: 10.1371/journal.pone.0212320.
- [3] N. Afrianto, D. H. Fudholi, and S. Rani, "Prediksi Harga Saham Menggunakan BiLSTM dengan Faktor Sentimen Publik," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 6, no. 1, pp. 41–46, 2022, doi: 10.29207/resti.v6i1.3676.
- [4] B. K. Kang, "Improving MACD Technical Analysis by Optimizing Parameters and Modifying Trading Rules: Evidence from the Japanese Nikkei 225 Futures Market," *J. Risk Financ. Manag.*, vol. 14, no. 1, 2021, doi: 10.3390/jrfm14010037.
- [5] J. Luo *et al.*, "Hybrid CNN-BiGRU-AM Model with Anomaly Detection for Nonlinear Stock Price Prediction," *Electron.*, vol. 14, no. 7, pp. 1–18, 2025, doi: 10.3390/electronics14071275.
- [6] A. K. Panigrahi, K. Vachhani, and S. K. Chaudhury, "Trend identification with the relative strength index (RSI) technical indicator –A conceptual study," *J. Manag. Res. Anal.*, vol. 8, no. 4, pp. 159–169, 2021, doi: 10.18231/j.jmra.2021.033.

- [7] M. Kadam, "The Technical Analysis for Gaining Profitability on Large Cap Stocks," *Int. J. Innov. Res. Sci. Eng. Technol.*, vol. 9, no. 2, pp. 63–69, 2022, doi: 10.17148/IARJSET.2022.9205.
- [8] R. Karim, M. K. Alam, and M. R. Hossain, "Stock market analysis using linear regression and decision tree regression," *2021 1st Int. Conf. Emerg. Smart Technol. Appl. eSmarTA 2021*, pp. 0–5, 2021, doi: 10.1109/eSmarTA52612.2021.9515762.
- [9] S. Khan and H. Alghulaiakh, "ARIMA model for accurate time series stocks forecasting," *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 7, pp. 524–528, 2020, doi: 10.14569/IJACSA.2020.0110765.
- [10] H. Mo, "Comparative Analysis of Linear Regression, Polynomial Regression, and ARIMA Model for Short-term Stock Price Forecasting," *Adv. Econ. Manag. Polit. Sci.*, vol. 49, no. 1, pp. 166–175, 2023, doi: 10.54254/2754-1169/49/20230509.
- [11] H. Song and H. Choi, "Forecasting Stock Market Indices Using the Recurrent Neural Network Based Hybrid Models: CNN-LSTM, GRU-CNN, and Ensemble Models," *Appl. Sci.*, vol. 13, no. 7, 2023, doi: 10.3390/app13074644.
- [12] S. C. Agrawal, "Deep learning based non-linear regression for Stock Prediction," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 1116, no. 1, p. 012189, 2021, doi: 10.1088/1757-899x/1116/1/012189.
- [13] M. Tran, D. Pham-Hi, and M. Bui, "Optimizing Automated Trading Systems with Deep Reinforcement Learning," *Algorithms*, vol. 16, no. 1, pp. 1–17, 2023, doi: 10.3390/a16010023.
- [14] M. Khushi and T. L. Meng, "Reinforcement learning in financial markets," *DatFa*, vol. 4, no. 3, pp. 1–17, 2019, doi: 10.3390/data4030110.
- [15] B. K. Kang, "Optimal and Non-Optimal MACD Parameter Values and Their Ranges for Stock-Index Futures: A Comparative Study of Nikkei, Dow Jones, and Nasdaq," *J. Risk Financ. Manag.*, vol. 16, no. 12, 2023, doi: 10.3390/jrfm16120508.
- [16] T. Chong, W.-K. Ng, and V. Liew, "Revisiting the Performance of MACD and RSI Oscillators," *J. Risk Financ. Manag.*, vol. 7, no. 1, pp. 1–12, 2014, doi: 10.3390/jrfm7010001.
- [17] F. Lima Paiva, L. Felizardo, R. Bianchi, and A. H. Reali Costa, *Intelligent Trading Systems: A Sentiment-Aware Reinforcement Learning Approach*. 2021. doi: 10.1145/3490354.3494445.
- [18] G. Papageorgiou, D. Gkaimanis, and C. Tjortjis, "Enhancing Stock Market Forecasts with Double Deep Q-Network in Volatile Stock Market Environments," *Electron.*, vol. 13, no. 9, 2024, doi: 10.3390/electronics13091629.
- [19] R. A. Saleh and H. T. Jawad, "The effect of some normalization methods on neural networks and robust methods with the presence of outliers," vol. 11, no. 3, pp. 3759–3767, 2025, doi: 10.22399/ijcesen.1716.
- [20] T. Tanaka, I. Nambu, Y. Maruyama, and Y. Wada, "Sliding-Window Normalization to Improve the Performance of Machine-Learning Models for Real-Time Motion Prediction Using Electromyography," *Sensors*, vol. 22, no. 13, 2022, doi: 10.3390/s22135005.
- [21] V. Mnih *et al.*, "Asynchronous methods for deep reinforcement learning," *33rd Int. Conf. Mach. Learn. ICML 2016*, vol. 4, no. September, pp. 2850–2869, 2016, doi: 10.48550/arXiv.1602.01783.
- [22] Y. Z. Chen, Y. Fang, X. Y. Li, and J. Wei, "A factor pricing model based on double moving average strategy," *Humanit. Soc. Sci. Commun.*, vol. 10, no. 1, pp. 1–13, 2023, doi: 10.1057/s41599-023-02362-x.