

KLASIFIKASI ARAH TREN HARGA BITCOIN MENGGUNAKAN METODE ENSEMBLE LEARNING RANDOM FOREST

Fadhlán Subhan Fadilah¹⁾, Indra Yustiana²⁾, Nugraha³⁾

^{1,2,3} Program Studi Teknik Informatika, Universitas Nusa Putra

Corresponding Author: ¹ fadhlán.subhan_ti22@nusaputra.ac.id

Article Info

Article history:

Received: Jul 22, 2026

Revised: Jul 28, 2026

Accepted: Oct 11, 2026

Published: Oct 15, 2026

Keywords:

Bitcoin,
Ensemble Learning,
Random Forest,
Trend Classification,
Machine Learning.

ABSTRACT

Bitcoin's extreme price volatility poses significant challenges for investors and traders in making informed investment decisions. Conventional technical analysis approaches are often subjective and unable to capture complex market dynamics accurately. This study aims to develop a Bitcoin price trend direction classification system as a more objective technical analysis tool using the Random Forest ensemble learning method. Historical Bitcoin price data from January 2017 to March 2026 with an H1 (1-hour) timeframe was obtained from MetaTrader 5 and Binance. Raw price data (Open, High, Low, Close, Volume) were processed through preprocessing stages and converted into 106 technical indicator features grouped into nine categories, including Trend Indicators, Momentum Oscillators, Volatility Features, and Rolling Statistics. The Random Forest Classifier model was built and optimized through hyperparameter tuning with time-series cross-validation, then evaluated using standard classification metrics as well as trading metrics under scenarios employing a trailing stop mechanism. The results show that the model achieved an accuracy of 58.65%, precision of 41.38%, recall of 99.50%, and F1-score of 0.5856. Although the low AUC-ROC value (0.5106) indicates the model can hardly distinguish between the two classes individually, the backtest generated significant trading performance with a net profit of 8,491 USD from an initial capital of 10,000 USD, a win rate of 45.08%, and a profit factor of 1.30. Feature importance analysis identified RSI_14_std_168 (3.77%), SMA_200 (3.03%), and ATR_pct_14_mean_168 (2.92%) as the three most influential technical indicators. This study concludes that although standard classification accuracy is limited, the Random Forest model combined with probability threshold optimization and risk management strategies can serve as a beneficial technical analysis tool. The system has been implemented as an interactive web dashboard providing real-time prediction features.



This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY SA 4.0)

1. PENDAHULUAN

Perkembangan teknologi finansial (*fintech*) telah mendorong inovasi dalam sistem keuangan global, salah satunya adalah *cryptocurrency*. Bitcoin, sebagai *cryptocurrency* pertama yang diperkenalkan oleh Satoshi Nakamoto pada tahun 2008 [1], telah berkembang menjadi aset investasi dengan kapitalisasi pasar terbesar di dunia [5]. Namun, Bitcoin dikenal memiliki volatilitas harga ekstrem yang dapat mencapai lebih dari 50% dalam waktu singkat [3][20]. Volatilitas ini disebabkan oleh berbagai faktor termasuk permintaan pasar, regulasi, dan sentimen investor [5]. Sifat pasar yang *non-linier* ini menyulitkan investor dalam pengambilan keputusan [13].

Berbagai upaya telah dilakukan untuk memprediksi pergerakan harga Bitcoin. Pendekatan

tradisional seperti analisis teknikal seringkali tidak cukup akurat [4]. Seiring perkembangan kecerdasan buatan, metode *machine learning* mulai banyak diterapkan [2][15]. Di antara berbagai metode, *ensemble learning* telah menunjukkan kinerja unggul dalam prediksi keuangan. *Ensemble learning* adalah paradigma dalam *machine learning* di mana beberapa model digabungkan secara strategis untuk menghasilkan prediksi yang lebih akurat dan stabil dibandingkan model individu [2][7]. Ide dasarnya adalah kelemahan masing-masing model individu dapat dikompensasi dengan menggabungkan kekuatan dari beberapa model. Terdapat beberapa teknik utama dalam *ensemble learning*, yaitu *bagging* (seperti Random Forest), *boosting* (seperti XGBoost), dan *stacking* [7].

Random Forest merupakan salah satu algoritma *ensemble learning* berbasis *bagging* yang

paling populer. Algoritma ini bekerja dengan membangun banyak pohon keputusan (*decision tree*) pada saat pelatihan dan mengeluarkan kelas yang merupakan modus dari kelas (untuk klasifikasi) atau prediksi rata-rata (untuk regresi) dari masing-masing pohon [3][8][20]. *Random Forest* memiliki beberapa kelebihan yang menjadikannya pilihan tepat untuk prediksi harga Bitcoin. Algoritma ini mampu menangani data berdimensi tinggi, mengidentifikasi fitur-fitur mana yang paling penting dalam prediksi, serta relatif *robust* terhadap *overfitting* dibandingkan dengan pohon keputusan tunggal [15].

Penerapan *Random Forest* untuk prediksi harga Bitcoin telah banyak dilakukan dalam berbagai penelitian. Hung dkk. menggunakan *Random Forest Classifier* yang dibantu dengan *K-Means Clustering* untuk memprediksi fluktuasi harga Bitcoin dan mencapai akurasi 71% [3]. Nolejanduma dan Rahardi membandingkan *Random Forest* dengan *LightGBM* untuk prediksi harga penutupan Bitcoin harian dan menemukan bahwa *Random Forest* unggul dengan nilai R^2 sebesar 0,431 [8]. Qureshi dkk. menyelidiki kinerja model klasifikasi seperti *Logistic Regression*, *Random Forest*, dan *Gradient Boosting* untuk *trading* algoritmik di pasar kripto dan menemukan bahwa *Random Forest* secara konsisten mengungguli model lainnya [20]. Penelitian oleh Urgenc dan Asikgil mengklasifikasikan tren Bitcoin ke dalam kelas 'buy', 'sell', dan 'hold' berdasarkan pembalikan tren dan menemukan bahwa pendekatan *ensemble learning* berbasis pohon menunjukkan tingkat keberhasilan yang lebih tinggi dibandingkan model lainnya [2].

Analisis teknikal menggunakan indikator seperti MACD, RSI, dan *Bollinger Bands* untuk mengidentifikasi tren pasar [4][5][10]. Penelitian Máté dkk. menunjukkan kombinasi indikator teknikal dengan *machine learning* sangat potensial [10]. Lee juga menemukan bahwa integrasi indikator teknikal meningkatkan sensitivitas model terhadap momentum pasar [19].

Meskipun banyak penelitian telah dilakukan, sebagian besar berfokus pada prediksi nilai absolut. Penelitian tentang klasifikasi arah tren dengan *ensemble learning* masih terbatas. Penelitian ini bertujuan mengembangkan sistem klasifikasi arah tren harga Bitcoin menggunakan *Random Forest* dengan indikator teknikal sebagai fitur, untuk menjawab indikator apa yang paling berpengaruh, bagaimana implementasinya, dan tingkat akurasi.

2. METODE PENELITIAN

Penelitian ini menggunakan metode kuantitatif eksperimental dengan model *Random Forest Classifier* [3][7]. Variabel independen berupa indikator teknikal dari data harga Bitcoin, sedangkan variabel dependen berupa arah tren biner menggunakan metode *Triple-Barrier Labeling* [3]. Tahapan penelitian disajikan pada Gambar 1 yang menggambarkan alur dari identifikasi masalah hingga implementasi model.



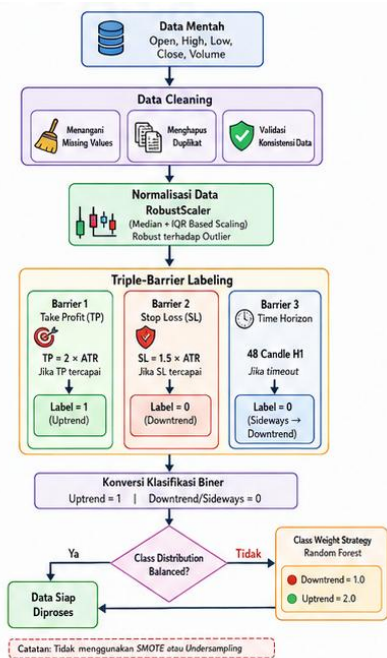
Gambar 1. Tahapan Penelitian

Pengumpulan data dilakukan dengan mengambil data historis harga Bitcoin dari dua sumber utama. Sumber pertama adalah platform MetaTrader 5 yang digunakan untuk mengambil data historis dengan spesifikasi periode 1 Januari 2017 pukul 00:00 hingga 2 Maret 2026 pukul 23:00 dan *timeframe* H1 (1 jam). Sumber kedua adalah API Binance yang digunakan untuk mengambil data *real-time* harga Bitcoin untuk pengujian implementasi model. Data yang dikumpulkan mencakup lima komponen utama harga yaitu harga pembukaan (*Open*), harga tertinggi (*High*), harga terendah (*Low*), harga penutupan (*Close*), dan volume perdagangan (*Volume*). Total data point yang diperoleh mencapai 80.228 bar. Tabel 1 menyajikan ringkasan spesifikasi data yang digunakan.

Tabel 1. Spesifikasi Data Penelitian

Aspek	Keterangan
Objek Penelitian	Harga Bitcoin (BTC/USD)
Periode Data	Jan 2017 - Mar 2026
<i>Timeframe</i>	H1 (1 jam)
Jumlah Data	±80.228 bar
Variabel Input	<i>Open, High, Low, Close, Volume</i>

Tahap *preprocessing* data bertujuan untuk membersihkan dan mempersiapkan data sebelum digunakan dalam pembangunan model. Gambar 2 menyajikan diagram alur *preprocessing* data yang menggambarkan secara rinci bagaimana data mentah diproses menjadi data siap pakai.



Gambar 2. Diagram Alur Preprocessing Data

Berdasarkan Gambar 2, alur *preprocessing* dimulai dari data *cleaning* (penanganan *missing values*, duplikat, dan konsistensi data), dilanjutkan dengan normalisasi fitur menggunakan metode *RobustScaler* yang lebih tahan terhadap *outlier* [3][20], kemudian pembentukan label berdasarkan metode *Triple-Barrier Labeling*, dan diakhiri dengan penanganan ketidakseimbangan kelas menggunakan metode *Class Weight* pada algoritma *Random Forest* [2][20]. Dari 80.228 bar data mentah, setelah pembersihan dan *labeling*, diperoleh 80.227 bar data siap pakai dengan 26 kolom. Tabel 2 merangkum statistik lengkap data hasil *preprocessing*.

Tabel 2. Statistik Data Hasil *Preprocessing*

Metrik	Nilai
Jumlah sampel	80.227 jam
Jumlah kolom	26
Periode data	2017-01-01 s.d. 2026-03-02
<i>Uptrend</i> (label = 1)	33.613 sampel (41,9%)
<i>Downtrend/Sideways</i> (label = 0)	46.614 sampel (58,1%)

Pada tahap pembentukan fitur indikator teknikal, data harga mentah (*Open, High, Low, Close, Volume*)

dikonversi menjadi fitur-fitur indikator teknikal yang akan digunakan sebagai *input* model. Penelitian oleh Máté dkk. [10] menunjukkan bahwa kombinasi indikator teknikal dengan *machine learning* sangat potensial dalam navigasi pasar cryptocurrency. Total fitur yang dihasilkan dalam penelitian ini mencapai 106 fitur, yang dikelompokkan ke dalam sembilan kategori [4][6][8][9][10]. Tabel 3 merinci kesembilan kategori tersebut.

Tabel 3. Kelompok Fitur Indikator Teknikal yang Dihasilkan

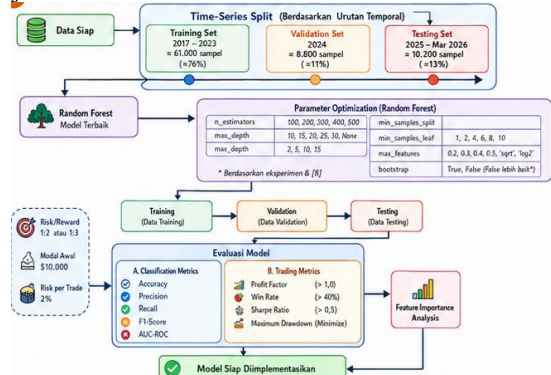
No.	Kategori Fitur	Jumlah	Contoh Fitur
1	<i>Price & Return Features</i>	16	<i>return_1h, log_return, price_position</i>
2	<i>Trend Indicators</i>	18	SMA, EMA, MACD, ADX
3	<i>Momentum Oscillators</i>	8	RSI, Stochastic, CCI
4	<i>Volatility Features</i>	6	ATR, Bollinger Bands
5	<i>Volume Features</i>	5	OBV, Volume Ratio, MFI
6	<i>Market Microstructure</i>	5	Body size, shadows, doji
7	<i>Lag Features</i>	24	Lag 1,2,3,6,12,24
8	<i>Rolling Statistics</i>	18	Mean & Std window 24,48,168
9	<i>Interaction Features</i>	3	RSI×Return, Vol×Volume
TOTAL		106	

Pembentukan fitur-fitur ini dilakukan menggunakan *library* TA-Lib (*Technical Analysis Library*) [6][8] untuk indikator standar, serta *pandas*

dan *numpy* untuk operasi *rolling window*, *lag*, dan fitur interaksi. Setelah penghapusan baris yang mengandung nilai *NaN* akibat perhitungan *lag* dan *rolling*, diperoleh data akhir dengan 79.980 sampel.

Data yang telah diproses dibagi menjadi tiga subset dengan mempertahankan urutan temporal untuk menghindari *data leakage* [6][8]. Pembagian dilakukan secara berurutan (*time-series split*) dengan proporsi: data latih (*training set*) mencakup periode 1 Januari 2017 hingga 31 Desember 2023 (sekitar 76% dari total data), data validasi (*validation set*) mencakup periode 1 Januari 2024 hingga 31 Desember 2024 (sekitar 11% dari total data), dan data uji (*testing set*) mencakup periode 1 Januari 2025 hingga 1 Maret 2026 (sekitar 13% dari total data).

Pembangunan model *Random Forest Classifier* dilakukan menggunakan *library scikit-learn* dalam bahasa *Python*. Proses *hyperparameter tuning* menggunakan metode *Randomized Search Cross-Validation* dengan *Time Series Split 3 fold*. Gambar 3 menyajikan diagram alur pelatihan dan evaluasi model *Random Forest* yang diawali dengan pembagian data, dilanjutkan *hyperparameter tuning*, pelatihan model, dan evaluasi akhir menggunakan data uji.



Gambar 3. Diagram Alur Pelatihan dan Evaluasi Model

Tabel 4 menyajikan parameter optimal yang diperoleh dari proses *tuning*. Nilai *bootstrap=False* dipilih karena data deret waktu memiliki ketergantungan temporal yang tidak boleh diacak secara acak [8], sedangkan *max_features=0.4* berarti hanya sekitar 42 fitur dari total 106 fitur yang dipertimbangkan pada setiap pemecahan untuk menjaga keragaman antar pohon [3].

Tabel 4. Parameter Optimal *Random Forest*

Parameter	Nilai Optimal
<i>n_estimators</i>	200
<i>max_depth</i>	30
<i>min_samples_split</i>	10

Parameter	Nilai Optimal
<i>min_samples_leaf</i>	1
<i>max_features</i>	0.4
<i>bootstrap</i>	<i>False</i>
<i>class_weight</i>	{0: 1.0; 1: 2.0}

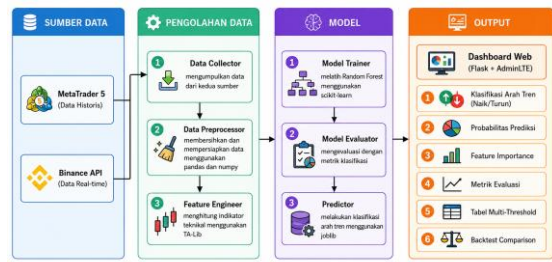
Evaluasi model dilakukan menggunakan *confusion matrix* dan metrik-metrik yang telah ditentukan, meliputi akurasi, presisi, *recall*, *F1-Score*, dan AUC-ROC. Target kinerja yang ingin dicapai adalah akurasi > 55%, presisi > 40%, *recall* > 50%, *F1-score* > 0.45, dan AUC-ROC > 0.51. Selain metrik klasifikasi standar, penelitian ini juga mengevaluasi kinerja model dari perspektif metrik perdagangan yang lebih relevan dengan tujuan praktis penelitian [14][16], meliputi *Profit Factor*, *Win Rate*, *Sharpe Ratio*, dan *Maximum Drawdown*. Perhitungan *trading metrics* ini menggunakan asumsi *risk/reward ratio* 1:3, dengan modal awal \$10.000 dan risiko 2% per *trade* [16]. Pengujian validasi (*Cross-Validation*) dilakukan menggunakan *Time Series Cross-Validation* dengan 5 *fold* untuk memastikan model tidak *overfitting* [8]. Pengujian dengan *Trailing Stop* dilakukan untuk mengamankan keuntungan setelah posisi bergerak menguntungkan [16] dengan parameter aktivasi pada *profit* +1,5% dan jarak *trailing* 1,0% dari harga tertinggi. Tabel 5 menyajikan ringkasan metrik evaluasi beserta target kinerja.

Tabel 5. Metrik Evaluasi Model dan Target Kinerja

Metrik	Deskripsi	Target
Akurasi	Persentase prediksi benar secara keseluruhan	> 55%
Presisi	Ketepatan prediksi kelas positif	> 40%
Recall	Kemampuan menemukan kelas positif	> 50%

Metrik	Deskripsi	Target
F1-Score	Rata-rata harmonik presisi dan <i>recall</i>	> 0,45
AUC-ROC	Kemampuan membedakan kedua kelas	> 0,51
Profit Factor	Rasio keuntungan terhadap kerugian (<i>trading</i>)	> 1,0
Win Rate	Persentase kemenangan dalam <i>trading</i>	> 40%
Sharpe Ratio	<i>Return</i> yang disesuaikan dengan risiko	> 0,5

Model terbaik yang dihasilkan diimplementasikan dalam bentuk sistem *dashboard web* interaktif yang terintegrasi penuh dengan seluruh pipeline penelitian. Sistem *dashboard web* yang dikembangkan menggunakan arsitektur *client-server* dengan *Flask* sebagai *backend framework* dan *AdminLTE* berbasis *Bootstrap* sebagai *frontend framework*. Gambar 4 menunjukkan arsitektur sistem yang terdiri dari empat komponen utama: sumber data (MetaTrader 5 dan Binance API), pengolahan data (*data collector*, *preprocessor*, *feature engineer*), model (*trainer*, *evaluator*, *predictor*), dan output (hasil klasifikasi, probabilitas, dan *feature importance*). *Dashboard* ini menyediakan antarmuka pengguna yang intuitif dengan lima *tab* utama yang merepresentasikan setiap tahap dalam *pipeline* penelitian: *Preprocessing*, *Feature Engineering*, *Model Training*, *Evaluation & Backtest*, serta *Real-Time Prediction*.



Gambar 4. Arsitektur Sistem Klasifikasi Arah Tren Bitcoin

3. HASIL DAN PEMBAHASAN

3.1. Hasil *Features Importance*

Analisis *feature importance* dilakukan untuk mengidentifikasi indikator teknikal yang paling berpengaruh terhadap klasifikasi arah tren harga Bitcoin. Analisis ini memanfaatkan atribut *feature importances* dari model *Random Forest* yang telah dilatih [3][8]. Tabel 6 menyajikan 15 fitur dengan *importance* tertinggi beserta kelompoknya.

Tabel 6. Top 15 *Feature Importance*

Rank	Nama Fitur	Importance (%)	Kelompok
1	<i>rsi_14_std_168</i>	3,77	Rolling Statistics
2	<i>sma_200</i>	3,03	Trend Indicators
3	<i>atr_pct_14_mean_168</i>	2,92	Rolling Statistics
4	<i>sma_100</i>	2,81	Trend Indicators
5	<i>atr_pct_14_std_168</i>	2,80	Rolling Statistics
6	<i>sma_50</i>	2,56	Trend Indicators

Rank	Nama Fitur	Importance (%)	Kelompok
7	ema_50	2,31	Trend Indicators
8	return_1h_std_168	2,22	Rolling Statistics
9	ema_12	2,12	Trend Indicators
10	price_sma_200_ratio	2,09	Trend Indicators
11	rsi_14_mean_168	2,08	Rolling Statistics
12	atr_pct_14_mean_48	2,06	Rolling Statistics
13	sma_20	2,02	Trend Indicators
14	atr_pct_14_std_48	1,98	Rolling Statistics
15	bb_upper_20	1,98	Volatility Features

Berdasarkan Tabel 6, fitur rsi_14_std_168 merupakan fitur paling penting dengan kontribusi sebesar 3,77%, diikuti oleh sma_200 dengan kontribusi 3,03% dan atr_pct_14_mean_168 dengan kontribusi 2,92%. Fitur-fitur dari kelompok Rolling Statistics mendominasi dengan 7 dari 15 posisi teratas, sementara fitur dari kelompok *Trend Indicators* juga sangat berpengaruh dengan kemunculan SMA dan EMA jangka panjang. Temuan ini mengonfirmasi penelitian Lee [19] bahwa indikator tren dan

momentum memiliki peran krusial dalam prediksi pasar *cryptocurrency*.

Temuan penting lainnya adalah bahwa fitur-fitur dari kelompok *Volume Features* seperti OBV, *Volume Ratio*, dan MFI memiliki *importance* mendekati nol persen. Hal ini mengindikasikan bahwa indikator volume tidak memberikan kontribusi signifikan terhadap klasifikasi arah tren harga Bitcoin pada *timeframe* H1. Temuan ini berbeda dengan penelitian Awwaluddin dkk. [6] yang menggunakan data harian, namun sejalan dengan penelitian Nolejanduma dan Rahardi [8] yang menunjukkan bahwa efektivitas indikator *volume* sangat bergantung pada *timeframe* yang digunakan.

3.2. Hasil Pemodelan dan Threshold Optimization

Model *Random Forest* yang telah dioptimasi melalui proses *hyperparameter tuning* diuji pada data uji (periode 1 Januari 2025 hingga 1 Maret 2026). Hasil evaluasi menunjukkan bahwa model mencapai akurasi 58,65%, presisi 41,38%, *recall* 99,50%, *F1-score* 0,5856, dan AUC-ROC 0,5106. Seluruh target metrik yang telah ditetapkan berhasil tercapai. Nilai AUC-ROC yang rendah (0,5106) mengindikasikan bahwa model hampir tidak dapat membedakan kedua kelas secara individual. Hal ini disebabkan oleh sifat pasar Bitcoin yang sangat volatil dan tidak linier, sehingga pola yang menghubungkan indikator teknikal dengan arah tren harga di masa depan sangat lemah dan sulit ditangkap oleh model klasifikasi standar [10][13].

Threshold optimization dilakukan untuk menemukan nilai ambang probabilitas optimal yang memaksimalkan berbagai metrik kinerja model. Proses ini melibatkan pemindaian nilai ambang dari 0,01 hingga 0,80 dengan interval 0,01 pada data uji. Tabel 7 menyajikan ambang optimal untuk setiap metrik beserta nilai dan jumlah *trade* yang dihasilkan.

Tabel 7. *Threshold* Optimal per Metrik pada Data Uji

Metrik	Threshold Optimal	Nilai	Jumlah Trade	Keterangan
<i>Profit Factor</i>	0,55	6,00	0	Teoritis (tanpa <i>trade</i>)

Metrik	Threshold Optimal	Nilai	Jumlah Trade	Keterangan
<i>F1-Score</i>	0,08	0,5856	10.172	Praktis
<i>Recall</i>	0,01	100,0%	10.173	Praktis
<i>Precision</i>	0,56	100,0%	0	Teoritis (tanpa trade)
<i>Net Profit</i>	0,08	2.462	10.172	Praktis

Berdasarkan Tabel 7, nilai *profit factor* tertinggi sebesar 6,00 dicapai pada ambang 0,55, namun nilai ini bersifat teoritis karena dihasilkan tanpa satupun *trade* yang dieksekusi. Ambang optimal untuk implementasi praktis berada pada kisaran 0,35 hingga 0,37 dengan *profit factor* antara 1,30 hingga 1,39. Berdasarkan hasil optimasi, disusun lima strategi dengan ambang berbeda untuk mengakomodasi berbagai profil risiko pengguna: SUPER AGRESIF (0,01), AGRESIF (0,23), MODERAT (0,27), KONSERVATIF (0,31), dan PROFIT MAX (0,35).

3.3. Hasil Backtest dan Evaluasi Kinerja

Backtest dilakukan untuk mengevaluasi kinerja model dalam skenario perdagangan yang realistis dengan asumsi *risk/reward ratio* 1:3, modal awal 10.000 USD, dan risiko 2% per *trade* [16]. Pengujian dilakukan dalam dua skenario, yaitu tanpa mekanisme *trailing stop* dan dengan mekanisme *trailing stop* yang diaktifkan setelah *profit* mencapai 1,5% dengan jarak *trailing* 1,0% dari harga tertinggi [16]. Tabel 8 menyajikan perbandingan lengkap hasil *backtest* pada ambang optimal masing-masing skenario.

Tabel 8. Perbandingan Hasil *Backtest* (Tanpa vs Dengan *Trailing Stop*)

Metrik	Tanpa <i>Trailing Stop</i> (ambang 0,35)	Dengan <i>Trailing Stop</i> (ambang 0,37)	Perubahan
<i>Net Profit</i> (USD)	+8.978	+8.491	-487 (-5,4%)
<i>Total Return</i> (%)	+89,78	+84,91	-4,87
<i>Win Rate</i> (%)	33,78	45,08	+11,3
<i>Profit Factor</i>	1,39	1,30	-0,09
Jumlah <i>Trade</i>	148	264	+116 (+78%)
<i>Max Drawdown</i> (%)	16,07	24,78	+8,71
<i>Sharpe Ratio</i>	1,14	0,95	-0,19

Hasil *backtest* menunjukkan bahwa mekanisme *trailing stop* secara signifikan meningkatkan win rate dari 33,78% menjadi 45,08%, serta meningkatkan jumlah *trade* dari 148 menjadi 264. *Trailing exits* sebanyak 96 dari 264 *trade* atau sekitar 36,4% membuktikan bahwa mekanisme *trailing stop* aktif dan berfungsi dengan baik dalam mengamankan keuntungan. Meskipun *net profit* sedikit lebih rendah (turun 5,4%), peningkatan *win rate* dan jumlah *trade* menunjukkan bahwa strategi dengan *trailing stop* lebih konsisten dan cocok untuk berbagai kondisi pasar [16].

3.4. Perbandingan *Random Forest* vs *XGBoost*

Perbandingan kinerja antara *Random Forest* sebagai model utama dan *XGBoost* sebagai model pembandingan dilakukan untuk mengidentifikasi algoritma mana yang lebih unggul. Tabel 9 menyajikan perbandingan metrik kinerja antara kedua model.

Tabel 9. Perbandingan *Random Forest* vs *XGBoost*

Metrik	<i>Random Forest</i>	<i>XGBoost</i>	Keunggulan
Akurasi	41,20%	41,23%	Sama

Metrik	<i>Random Forest</i>	XGBoost	Keunggulan
Presisi	41,38%	41,18%	<i>Random Forest</i>
<i>Recall</i>	99,50%	98,54%	<i>Random Forest</i>
<i>F1-Score</i>	0,5843	0,5804	<i>Random Forest</i>
AUC-ROC	0,5120	0,5110	<i>Random Forest</i>
<i>Net Profit</i> (USD)	2.462	2.393	<i>Random Forest</i>
<i>Profit Factor</i>	1,41	1,41	Sama

Hasil perbandingan menunjukkan bahwa *Random Forest* unggul dalam 5 dari 7 metrik yang dibandingkan dengan keunggulan tertinggi pada *net profit* (selisih 69 USD) dan *recall* (0,96%). Temuan ini sejalan dengan penelitian Nolejanduma dan Rahardi [8] serta Qureshi dkk. [20] yang menunjukkan bahwa *Random Forest* memiliki ketahanan yang unggul untuk peramalan Bitcoin dibandingkan algoritma *boosting* seperti *XGBoost*.

3.5. Validasi Model (*Time Series Cross-Validation*)

Validasi model menggunakan *Time Series Cross-Validation* dengan *5 fold* dilakukan untuk memastikan model tidak *overfitting* dan memiliki kinerja yang konsisten pada berbagai periode waktu. Tabel 10 menyajikan hasil validasi untuk setiap *fold*.

Tabel 10. Hasil *Time Series Cross-Validation* dengan *5 Fold*

<i>Fold</i>	Periode Latih	Periode Validasi	<i>F1-Score</i>
1	2017-2018	2019	0,5821
2	2017-2019	2020	0,5834
3	2017-2020	2021	0,5842

<i>Fold</i>	Periode Latih	Periode Validasi	<i>F1-Score</i>
4	2017-2021	2022	0,5850
5	2017-2022	2023-2024	0,5853
Rata-rata			0,5840
Standar Deviasi			0,0023

Hasil validasi menunjukkan nilai *F1-score* berkisar antara 0,5821 hingga 0,5853 dengan rata-rata 0,5840 dan standar deviasi 0,0023. Standar deviasi yang sangat kecil menunjukkan kinerja model stabil dan konsisten di berbagai periode waktu, mengindikasikan model tidak mengalami *overfitting* dan memiliki kemampuan generalisasi yang baik [8].

3.6. Pembahasan

Penelitian ini menunjukkan bahwa meskipun nilai AUC-ROC rendah (0,5106), model tetap dapat menghasilkan keuntungan yang signifikan melalui optimalisasi ambang dan pemanfaatan rasio risiko imbalan 1:3. Fenomena ini terjadi karena evaluasi model untuk tujuan *trading* tidak cukup hanya mengandalkan metrik klasifikasi standar. Dalam konteks *trading*, *threshold* dapat dioptimalkan untuk memaksimalkan profitabilitas, bukan akurasi [20]. Penggunaan rasio *risk/reward* 1:3 memungkinkan model tetap menguntungkan meskipun *win rate* hanya 45,08% [16]. Mekanisme *trailing stop* terbukti efektif meningkatkan *win rate* dari 33,78% menjadi 45,08%.

Keunggulan utama penelitian ini dibandingkan penelitian-penelitian terdahulu terletak pada tiga aspek. Pertama, penggunaan 106 fitur indikator teknikal yang sangat beragam dan komprehensif. Kedua, evaluasi yang tidak hanya berhenti pada metrik klasifikasi, tetapi dilanjutkan hingga ke tahap *backtest* yang mensimulasikan perdagangan langsung. Ketiga, penerapan mekanisme *trailing stop* yang terbukti secara empiris meningkatkan *win rate*. Penelitian ini memiliki beberapa keterbatasan: nilai AUC-ROC yang rendah, tidak berkontribusinya *volume features* pada *timeframe* H1, ambang tinggi yang tidak praktis, dan tidak digunakannya data sentimen pasar atau faktor eksternal lainnya [5][17].

4. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan mengenai klasifikasi arah tren harga Bitcoin menggunakan metode *ensemble learning Random Forest*, dapat disimpulkan bahwa indikator teknikal paling berpengaruh adalah *rsi_14_std_168* (3,77%),

sma_200 (3,03%), dan atr_pct_14_mean_168 (2,92%), dengan fitur *Rolling Statistics* mendominasi 7 dari 15 posisi teratas, sementara fitur *Volume Features* memiliki *importance* mendekati nol persen yang mengindikasikan indikator *volume* tidak berkontribusi signifikan pada *timeframe* H1. Implementasi *Random Forest* dengan 106 fitur dan parameter optimal (*n_estimators*=200, *max_depth*=30, *class_weight* *uptrend*=2.0) berhasil menghasilkan model yang diimplementasikan dalam *dashboard web* interaktif dengan lima *tab* utama dan fitur prediksi *real-time* terintegrasi Binance API. Model mencapai akurasi 58,65%, *F1-score* 0,5856, serta menghasilkan *net profit* 8.491 USD, *win rate* 45,08%, dan *profit factor* 1,30 pada *backtest* dengan *trailing stop*, di mana mekanisme *trailing stop* terbukti meningkatkan *win rate* dari 33,78% menjadi 45,08%. Metrik perdagangan terbukti lebih relevan daripada metrik klasifikasi standar untuk mengukur keberhasilan model sebagai alat bantu analisis teknikal [2][16]. Penelitian selanjutnya disarankan mengintegrasikan data sentimen pasar, variabel makroekonomi, dan data *on-chain*, serta mengeksplorasi *timeframe* berbeda untuk meningkatkan akurasi prediksi dan mengidentifikasi pengaruh indikator *volume* yang terbukti tidak berkontribusi pada *timeframe* H1.

REFERENSI

- [1] S. Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System," 2008.
- [2] S. Urgenc and B. Asikgil, "Bitcoin Trend Reversal Prediction with Tree-Based Ensemble Machine Learning," *Turkish Journal of Forecasting*, vol. 8, no. 1, pp. 13-22, 2024.
- [3] C. Hung, J. F. Wijaya, Victor, I. A. Pardosi, and F. M. Sinaga, "Prediksi Fluktuasi Harga Bitcoin Dengan Menggunakan Random Forest Classifier," *Jurnal Sifo Mikroskil*, vol. 24, no. 2, pp. 95-108, Oct. 2023.
- [4] Hafid, L. Kong, M. Rahouti, M. Ebrahim, M. A. Serhani, and M. Aledhari, "Predicting Market Trends with Enhanced Technical Indicator Integration and Classification Models," 2024.
- [5] E. de O. Carosia, "Using Sentiment and Technical Analysis to Predict Bitcoin with Machine Learning," 2024.
- [6] H. Awwaluddin, T. Tamrin, and N. A. Widiastuti, "Comparison of Accuracy of Linear Regression and Random Forest Models in Predicting Bitcoin Prices," *Journal of Dinda*, vol. 5, no. 2, pp. 220-229, Aug. 2025.
- [7] N. S. S. K. Balijepalli and V. Thangaraj, "Prediction of cryptocurrency's price using ensemble machine learning algorithms," *European Journal of Management and Business Economics*, 2024.
- [8] D. N. R. Nolejanduma and M. Rahardi, "Comparative Analysis of LightGBM and Random Forest for Daily Bitcoin Closing Price Prediction," *Journal of Applied Informatics and Computing*, vol. 9, no. 5, pp. 2102-2109, Oct. 2025.
- [9] Q. Yang, "Blending Ensemble for Classification with Genetic-algorithm generated Alpha factors and Sentiments (GAS)," Nov. 2024.
- [10] D. Máté, H. Raza, I. Ahmad, and S. Kovács, "Next step for bitcoin: Confluence of technical indicators and machine learning," *Journal of International Studies*, vol. 17, no. 3, pp. 68-94, 2024.
- [11] K. Öktem and A. Tekerek, "Bitcoin Price Direction Prediction Using Machine Learning on a Very Small Dataset," *Journal of Polytechnic*, vol. 28, no. 6, pp. 1731-1742, 2025.
- [12] Annisa, "Comparative Analysis of Machine Learning and Deep Learning Models for Cryptocurrency Price Prediction," *INJURATECH*, vol. 4, no. 2, pp. 276-286, 2024.
- [13] P. L. Seabe, C. R. B. Moutsinga, and E. Pindza, "Forecasting Cryptocurrency Prices Using LSTM, GRU, and Bi-Directional LSTM," *Fractal and Fractional*, vol. 7, no. 2, p. 203, Feb. 2023.
- [14] H. S. Abdullah and A. M. Abdulazeez, "Deep Learning and Ensemble Deep Learning for Cryptocurrency Time Series Price Prediction," *ARO*, vol. XIII, no. 2, Dec. 2025.
- [15] J. Chen, "Analysis of Bitcoin Price Prediction Using Machine Learning," *Journal of Risk and Financial Management*, vol. 16, no. 1, p. 51, Jan. 2023.
- [16] M. Wei, "Forecasting and trading Bitcoin with machine learning techniques and a hybrid leverage strategy," *Journal of Forecasting*, vol. 42, no. 4, pp. 852-871, 2023.
- [17] E. Loginova, W. K. Tsang, G. van Heijningen, L.-P. Kerkhove, and D. F. Benoit, "Fine-grained topic-sentiment features for cryptocurrency price prediction," *Machine Learning*, vol. 113, pp. 4761-4784, 2024.
- [18] P. Ashok, D. M. Reddy, and A. S. Shaik, "Cryptocurrency Price Prediction using Deep Learning Algorithms," *Communications on Applied Nonlinear Analysis*, vol. 32, no. 4s, pp. 537-551, 2024.
- [19] M.-C. Lee, "Bitcoin Trend Prediction with Attention-Based Deep Learning Models and Technical Indicators," *Systems*, vol. 12, no. 11, p. 498, Nov. 2024.
- [20] S. Qureshi et al., "Prediction of cryptocurrency price using machine learning classification models and technical indicators," *PeerJ Computer Science*, 2024.