Cross-Temporal Feature Integration in Cryptocurrency Direction Prediction: a Confidence-Optimized Binary Classification Approach

Oleksandr Kuznetsov[1,*,†], Yelyzaveta Kuznetsova[3,†], Mykola Mormul[4,†] and Dmytro Shchytov[4,†]

[1] Department of Theoretical and Applied Sciences, eCampus University, Via Isimbardi 10, Novedrate (CO), 22060, Italy
[2] Department of Intelligent Software Systems and Technologies, School of Computer Science and Artificial Intelligence, V.N. Karazin Kharkiv National University, 4 Svobody Sq., 61022 Kharkiv, Ukraine
[3] Department of Information and Communication Systems Security, School of Computer Sciences, V. N. Karazin Kharkiv National University, 4 Svobody Sq., 61022 Kharkiv, Ukraine
[4] Department of Cyber Security and Information Technology, University of Customs and Finance, Vernadskogo str., 2/4, 49000, Dnipro, Ukraine

Abstract

Cryptocurrency markets exhibit extreme volatility and complex microstructure dynamics that challenge traditional prediction frameworks. This study introduces a binary classification approach for cryptocurrency direction prediction that integrates macro momentum indicators with microstructure features across multiple temporal scales. Unlike conventional three-class methods that confound directional prediction with execution timing, our framework separates these components using confidence-based thresholds to enable explicit precision-recall optimization. We evaluate the methodology across 11 major cryptocurrency pairs using comprehensive parameter optimization spanning prediction horizons from 10 to 600 minutes, deadband thresholds from 2 to 20 basis points, and confidence levels of 0.6 and 0.8. The unified feature representation combines daily OHLCV momentum signals with minute-frequency order book dynamics, capturing temporal bridges where fundamental price discovery aligns with short-term market making activities. High confidence regimes achieve peak profits of 167.64 basis points per trade with directional accuracies of 82-95% on executed trades, representing 60.4% improvement over moderate confidence conditions. Optimal performance occurs at intermediate horizons (400-600 minutes) where daily momentum trends manifest through intraday order flow patterns. The confidence threshold mechanism proves critical for economic viability, with high confidence strategies tolerating transaction costs up to 6 basis points while maintaining positive returns. Multi-scale feature integration provides superior signal representation compared to single-timeframe approaches, contributing to directional accuracies that exceed published benchmarks. The framework demonstrates practical viability for institutional cryptocurrency trading applications while revealing fundamental trade-offs between trading frequency and signal quality in digital asset markets.

Keywords

cryptocurrency prediction, binary classification, confidence thresholds, multi-scale features, market microstructure, machine learning, algorithmic trading, temporal integration, neural networks

1. Introduction

Cryptocurrency markets present unique challenges for algorithmic trading systems due to their extreme volatility, continuous operation, and complex microstructure dynamics. Unlike traditional financial markets, cryptocurrency exchanges operate 24/7 without circuit breakers, creating

Information Technology and Implementation (IT&I-2025), November 20-21, 2025, Kyiv, Ukraine
*Corresponding author.
† These authors contributed equally.
oleksandr.kuznetsov@uniecampus.it (O. Kuznetsov); elizabet8smidt12@gmail.com (Y. Kuznetsova); nikolaj.mormul@gmail.com (M. Mormul); dmytro.shchytov@gmail.com (D. Shchytov)
0000-0003-2331-6326 (O. Kuznetsov); 0000-0002-0573-0913 (Y. Kuznetsova); 0000-0002-8036-3236 (M. Mormul); 0000-0003-4306-8016 (D. Shchytov)

environments where price movements can exceed 10% within minutes. This volatility creates both opportunities and risks for quantitative trading strategies.

Current machine learning approaches to cryptocurrency price prediction typically focus on single timeframe analysis, using either daily price data or minute-level technical indicators in isolation. This narrow temporal focus overlooks the interaction between macroeconomic trends and microstructure dynamics that characterizes modern digital asset markets. Daily momentum patterns often manifest through intraday order flow changes, while microstructure signals gain predictive power when aligned with broader market trends.

Traditional prediction frameworks employ three-class classification schemes where models simultaneously learn directional prediction and execution timing decisions. This approach confounds signal extraction with risk management, potentially degrading both prediction accuracy and trading performance. The mixed representation of unclear signals and inappropriate timing within no-trade samples may compromise model learning effectiveness.

This research addresses these limitations by introducing a binary classification approach that separates directional prediction from execution control. We develop a confidence-threshold mechanism that enables explicit optimization of the precision-recall trade-off while integrating features across multiple temporal scales. Our methodology combines macro momentum indicators derived from daily price data with microstructure features extracted from minute-frequency order book snapshots.

The unified approach captures temporal bridges where daily directional bias influences minute-level market making activities. We evaluate this framework across eleven major cryptocurrency pairs using comprehensive parameter optimization that explores prediction horizons from 10 to 600 minutes, deadband thresholds from 2 to 20 basis points, and confidence levels of 0.6 and 0.8.

2. Literature Review

Recent advances in cryptocurrency prediction research have established several methodological foundations relevant to our investigation. Neural network architectures, particularly Long Short-Term Memory networks, consistently achieve directional accuracies of 60-85% across multiple studies. Zhang et al. (2024) [1] conducted a comprehensive survey finding that LSTM models achieve 83-84% average accuracy for Bitcoin and Ethereum prediction tasks, with ensemble methods often outperforming individual models.

Attention mechanisms represent significant advancement in cryptocurrency prediction architectures. Shang et al. (2024) [2] propose an attention-based CNN-BiGRU model for Ethereum price prediction, achieving RMSE of 151.6 and MAE of 91.2, substantially outperforming traditional CNN-GRU baselines. Their two-stage approach combines improved CNN for feature extraction with bidirectional GRU and attention mechanisms.

Graph neural networks introduce network-based perspectives to cryptocurrency prediction. Zhong et al. (2023) [3] develop LSTM-ReGAT, combining LSTM with Relationwise Graph Attention Networks for price trend prediction. Their approach constructs cryptocurrency networks based on shared features and achieves AUC of 0.6615 and accuracy of 62.97%, representing modest improvements over LSTM baselines.

Multi-target learning emerges as promising direction for cryptocurrency prediction. Pellicani et al. (2025) [4] introduce CARROT, employing temporal clustering with Dynamic Time Warping to group correlated cryptocurrencies before training multi-target LSTM models. Their approach achieves average 10% improvement in macro F1-score over single-target LSTMs, with best performance showing 19% improvement.

High-frequency prediction presents unique challenges requiring specialized architectures. Peng et al. (2024) [5] propose ACLMC for multiple cryptocurrencies combined with novel triple trend labeling using local minimum series. Their approach integrates macro and microstructure features across multiple frequencies, achieving significant reduction in transaction numbers while maintaining profitable performance.

Feature selection methodology significantly impacts cryptocurrency prediction performance. Youssefi et al. (2025) [6] conduct systematic investigation of feature selection methods applied to 130+ technical indicators, achieving 80-85% feature reduction while maintaining performance. Their results show peak R^2 values of 0.45-0.7 across BTC, ETH, and BNB pairs.

Uncertainty quantification represents emerging focus in cryptocurrency prediction research. Golnari et al. (2024) [7] introduce Probabilistic Gated Recurrent Units for Bitcoin price prediction with uncertainty quantification. Their approach integrates probabilistic attributes into standard GRU architecture, achieving R^2-score of 0.99973 and MAPE of 0.00190.

Potential field theory provides theoretical foundation for cryptocurrency market characterization. Anoop et al. (2025) [8] present Bayesian machine learning framework using potential field theory and Gaussian processes to model cryptocurrency price movements. Their analysis shows that attractors captured market trends with mean attractor features improving LSTM prediction performance by 25-28%.

Integration of prediction models with trading strategies receives increasing attention. Kang et al. (2025) [9] investigate technical indicator integration with deep learning-based price forecasting across 12 models. Their best performing strategy combines TimesNet with Bollinger Bands, achieving returns of 3.19 and Sharpe ratio of 3.56.

Market microstructure analysis reveals important patterns relevant to cryptocurrency prediction. Liu et al. (2025) [10] investigate liquidity commonality across 50 major cryptocurrencies, finding strong positive liquidity commonality with seasonal patterns persisting after controlling for volatility and returns.

Alternative methodological approaches provide complementary perspectives. Yang et al. (2025) [11] propose grey multivariate convolution models for short-term cryptocurrency price forecasting, achieving highly accurate predictions with MAPE values of 1.58% for BTC, 1.12% for ETH, and 2.53% for LTC.

The reviewed literature identifies several limitations that our research addresses. Most studies focus on single-timeframe analysis, missing opportunities for cross-temporal signal integration. Confidence-based execution control remains underexplored, with most approaches using fixed prediction thresholds. Systematic parameter optimization across multiple dimensions lacks comprehensive treatment in existing work.

3. Methodology

We develop a binary classification approach that fundamentally restructures the cryptocurrency direction prediction problem. Traditional methods employ three-class classification where models simultaneously learn direction prediction and trade execution decisions. Our framework decouples these components by training a binary classifier to predict direction and employing a separate confidence-based mechanism to control trade execution.

The binary approach operates on the premise that directional prediction and execution timing require different signal processing mechanisms. Direction prediction benefits from pure signal extraction without complications of mixed no-trade samples that may represent either unclear signals or inappropriate timing. The confidence threshold provides explicit control over the precision-recall trade-off.

Let $X_t \in \mathbb{R}^d$ represent the feature vector at time t containing both macro and microstructure signals. The model learns a mapping $f : X_t \rightarrow [0,1]$, where $f(X_t)$ represents the probability of upward price movement over prediction horizon h. The directional prediction follows:

$\hat{y}_t = \mathrm{I}[f(X_t) > 0.5]$, where $\mathrm{I}[\cdot]$ is the indicator function. The confidence measure is computed as $c_t = \max(f(X_t), 1 - f(X_t))$, representing the maximum probability assigned to either direction.

Trade execution occurs when confidence exceeds threshold τ: Execute trade if $c_t \geq \tau$. This mechanism creates explicit precision-recall control where higher τ values reduce trading frequency but improve signal quality.

The macro component derives features from daily OHLCV data across 100+ cryptocurrencies, providing market-wide context and fundamental momentum indicators. Feature engineering produces temporally lagged indicators to prevent look-ahead bias while capturing relevant market dynamics.

Price momentum features include multi-horizon returns:

$$R_{t,k} = \frac{P_t}{P_{t-k}} - 1$$

for horizons $k \in \{1, 5, 20\}$ days.

Moving average indicators capture trend dynamics:

$$MA_{t,k} = \frac{1}{k} \sum_{i=0}^{k-1} P_{t-i}$$

for windows $k \in \{5, 20, 50\}$ days.

Volatility measures employ rolling standard deviations of returns:

$$Vol_{t,k} = \sqrt{\frac{1}{k-1} \sum_{i=1}^{k} (R_{t-i} - \bar{R}_{t,k})^2}$$

for windows $k \in \{5, 20, 60\}$ days.

Technical indicators include RSI computed as

$$RSI_t = 100 - \frac{100}{1 + RS_t},$$

where

$$RS_t = \frac{EMA[gains]}{EMA[losses]}$$

using 14-day exponential moving averages.

All macro features are temporally aligned to prevent look-ahead bias by using only information available at prediction time. Daily macro signals are forward-filled to match the minute-frequency prediction schedule, ensuring temporal consistency across feature sources.

Microstructure features derive from minute-frequency order book snapshots, capturing market-making dynamics and short-term liquidity conditions. These features complement macro indicators by providing real-time market sentiment and execution environment information.

Order book imbalance measures the relative strength of buy versus sell pressure:

$$Imbalance_t = \frac{BidVol_t - AskVol_t}{BidVol_t + AskVol_t},$$

where volumes are computed across multiple depth levels.

Spread measures include both absolute and relative spreads:

$$Spread_{t,rel} = \frac{Ask_t - Bid_t}{MidPrice_t} \times 10000$$

in basis points.

Depth features aggregate liquidity across order book levels:

$$Depth_{t,k} = \sum_{i=1}^{k} (BidVol_{t,i} + AskVol_{t,i})$$

for levels $k \in \{1,5,10\}$.

Market impact proxies estimate the price effect of hypothetical trades:

$$Impact_{t,v} = \frac{VWAP_{t,v} - MidPrice_t}{MidPrice_t}$$

for volume v.

Temporal features include price volatility over short windows and return autocorrelations to capture momentum and mean reversion patterns at minute frequencies. All microstructure features undergo outlier treatment to handle extreme market conditions and data quality issues.

The unified feature space combines macro and microstructure signals, creating approximately 200+ candidate features. Feature selection employs mutual information scoring to identify the most predictive variables while controlling dimensionality for computational efficiency.

Mutual information captures both linear and non-linear relationships between features and target variables:

$$MI(X,Y) = \sum_{x,y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)},$$

where $p(\cdot)$ represents empirical probability distributions.

The top 64 features are selected based on mutual information scores, balancing predictive power with computational constraints. Feature scaling employs robust standardization to handle outliers common in financial data:

$$X_{scaled} = \frac{X - median(X)}{MAD(X)},$$

where $MAD(X)$ represents median absolute deviation.

The validation framework employs symbol-wise temporal splitting to prevent data leakage while maintaining realistic trading conditions. Each cryptocurrency pair is independently split into training, validation, and test periods using chronological ordering.

For each symbol s, the temporal split allocates data as follows: Training period covers the earliest 70% of observations, validation period encompasses the subsequent 15%, and test period includes the final 15%. This approach ensures that all model training and hyperparameter optimization occur using only historical information relative to evaluation periods.

Target variable construction requires careful attention to temporal alignment and look-ahead bias prevention. For prediction horizon h minutes, the target variable at time t is defined using the mid-price at time $t+h$:

$$y_t = \mathrm{I}[P_{t+h} > P_t \cdot (1 + deadband)]$$

for upward movements and

$$y_t = 0$$

for $P_{t+h} < P_t \cdot (1 - deadband)$ for downward movements.

The deadband parameter filters marginal price movements that fall within typical bid-ask spreads or market noise. Deadband values of 2-20 basis points ensure that predicted movements exceed transaction costs and represent economically meaningful directional signals.

Confidence threshold optimization occurs during the validation phase using systematic grid search. The optimization space covers $\tau \in [0.50, 0.95]$ with 0.01 increments, evaluating multiple optimization criteria including profit maximization, expected value maximization, and constrained optimization with minimum coverage requirements.

The core prediction model employs a multi-layer perceptron architecture optimized for financial time series prediction. The network structure consists of three hidden layers with [256, 128, 64] neurons respectively, using ReLU activation functions and dropout regularization.

The input layer accepts the 64-dimensional feature vector combining macro and microstructure signals. Hidden layers employ progressive dimensionality reduction to extract hierarchical feature representations. The output layer uses sigmoid activation to produce class probabilities suitable for confidence-based execution decisions.

Model training employs early stopping based on validation loss to prevent overfitting while maximizing generalization performance. Training proceeds for a maximum of 20 epochs with early termination if validation loss fails to improve for 5 consecutive epochs.

Class weight balancing addresses potential imbalances between upward and downward price movements in the binary training set. Weights are computed as inversely proportional to class frequencies:

$$w_c = \frac{n_{total}}{2 \cdot n_c},$$

where n_c is the sample count for class c.

Post-training probability calibration ensures that predicted confidence scores accurately reflect actual prediction reliability. Isotonic regression calibration is applied using validation data to map raw model outputs to well-calibrated probabilities.

Performance evaluation employs multiple metrics capturing different aspects of trading system effectiveness. Primary metrics include average profit per trade, coverage, and directional accuracy on executed trades.

Average profit per trade measures economic value creation:

$$\bar{\pi} = \frac{1}{N_{exec}} \sum_{i=1}^{N_{exec}} (r_i \cdot d_i - c),$$

where r_i is the return, d_i is the predicted direction, and c represents transaction costs.

Coverage quantifies market participation:

$$\kappa = \frac{N_{exec}}{N_{total}},$$

where N_{exec} is executed trades and N_{total} is total opportunities.

Directional accuracy measures prediction quality:

$$\alpha = \frac{1}{N_{exec}} \sum_{i=1}^{N_{exec}} \mathrm{I}[d_i = sign(r_i)]$$

on executed trades only.

All metrics are computed on the 11-symbol subset where both macro and microstructure data are available. This constraint ensures consistent feature availability across all trading decisions while maintaining representative coverage of major cryptocurrency pairs.

The evaluation was conducted on 11 major cryptocurrency pairs with sufficient liquidity and microstructure data availability: BTC/USDT, ETH/USDT, BNB/USDT, XRP/USDT, ADA/USDT, SOL/USDT, DOT/USDT, MATIC/USDT, LINK/USDT, UNI/USDT, and AVAX/USDT. These pairs collectively represent over 70% of total cryptocurrency market capitalization and ensure adequate order book depth for reliable microstructure feature extraction.

4. Results

4.1. Experimental Design and Parameter Space

We conducted systematic parameter optimization across 80 unique configurations spanning two confidence regimes. The experimental grid encompassed prediction horizons from 10 to 600 minutes, deadband thresholds from 2 to 20 basis points, and confidence levels of 0.6 and 0.8. Each configuration was evaluated using symbol-wise temporal splitting across 11 major cryptocurrency pairs.

The moderate confidence regime ($\tau = 0.6$) encompassed 40 experimental configurations, while the high confidence regime ($\tau = 0.8$) included an additional 40 configurations. This dual-threshold approach enabled characterization of the full spectrum of performance trade-offs available to practitioners.

4.2. Performance Under Moderate Confidence Conditions

The moderate confidence regime demonstrates distinct performance characteristics across the parameter space. Approximately 25% of configurations generate negative returns ranging from -31.77 to -1.00 basis points, while 75% achieve positive profitability with returns extending up to 152.69 basis points.

Coverage patterns show bimodal distribution with 37.5% of configurations achieving less than 2% coverage at short horizons, while median coverage across all configurations reaches 44.9% at longer horizons. Win rate distributions exhibit pronounced clustering around 80-90% for successful configurations, indicating consistent directional accuracy when trades are executed.

Table 1 presents performance statistics aggregated by prediction horizon under moderate confidence conditions. The horizon effect demonstrates sharp transitions rather than gradual improvement, with horizons below 100 minutes consistently producing negative returns.

Table 1

Performance by Prediction Horizon (Moderate Confidence, $\tau = 0.6$)

Horizon (minutes)	Mean Profit (basis points)	Coverage (percent)	Win Rate (percent)	Direction Accuracy
10	-24.47	0.0	0.0	0.00
20	-3.49	0.0	0.0	0.50
30	-0.43	0.0	57.1	0.57
50	20.78	1.7	63.1	0.63
100	30.26	4.7	70.1	0.78
200	51.89	9.1	83.3	0.83
300	66.94	9.6	85.4	0.85

400	74.09	10.4	84.0	0.84
500	79.90	13.4	82.5	0.82
600	104.52	21.6	82.5	0.82

The transition occurs abruptly around 50 minutes, where coverage jumps from essentially zero to measurable levels. This suggests a fundamental threshold in cryptocurrency market microstructure where noise-to-signal ratios become favorable for directional prediction. Beyond 200 minutes, profit growth continues but at diminishing rates, while coverage plateaus around 10-20%.

Standard deviations decrease substantially for horizons above 300 minutes, indicating more stable and predictable performance. This stability suggests that longer horizons capture fundamental price discovery mechanisms rather than transient microstructure effects.

4.3. Performance Under High Confidence Conditions

The high confidence regime demonstrates markedly different performance characteristics. The 40 experimental configurations under $\tau = 0.8$ show more pronounced separation between successful and unsuccessful parameter combinations. Negative returns concentrate in a narrower range, while positive returns extend to higher levels with maximum reaching 167.64 basis points.

Coverage distributions show strong polarization with 60% of configurations falling at or below 1% coverage, while successful configurations reach 3-22%. Win rate distributions cluster more tightly around 85-95%, representing substantial improvement over moderate confidence conditions.

Table 2 presents horizon-aggregated performance under high confidence conditions. The horizon effect becomes more pronounced under strict confidence requirements, with sharper transitions and higher peak performance levels.

Table 2

Performance by Prediction Horizon (Moderate Confidence, $\tau = 0.8$)

Horizon (minutes)	Mean Profit (basis points)	Coverage (percent)	Win Rate (percent)	Direction Accuracy
10	-24.47	0.0	0.0	0.00
20	-3.49	0.0	0.0	0.44
30	0.43	0.0	51.2	0.51
50	24.69	0.8	67.3	0.68
100	46.01	1.4	77.9	0.78
200	53.35	4.9	89.9	0.90
300	72.77	9.6	88.5	0.88
400	87.77	7.1	87.0	0.87
500	100.31	13.3	84.0	0.84
600	132.69	17.9	82.5	0.82

The critical transition horizon shifts to approximately 50-100 minutes under high confidence, representing a delay compared to moderate confidence conditions. This delay reflects stricter requirements for signal confidence, which naturally require longer observation periods to accumulate sufficient evidence for trade execution.

Performance gains under high confidence are substantial, with the 600-minute horizon achieving 132.69 basis points average profit compared to 104.52 basis points under moderate confidence. This 27% improvement comes at the cost of reduced coverage, representing a clear risk-return trade-off.

4.4. Comparative Analysis Across Confidence Regimes

Systematic comparison between moderate and high confidence regimes reveals fundamental trade-offs in cryptocurrency direction prediction systems. High confidence regimes achieve superior peak performance but require longer horizons to reach profitability.

The high confidence regime delivers maximum profit of 167.64 basis points compared to 104.52 basis points for moderate confidence, representing 60.4% improvement in peak profitability. However, coverage patterns reveal the precision-recall trade-off, with moderate confidence maintaining 50-65% coverage at optimal horizons while high confidence drops to 3-21% coverage.

Win rate evolution demonstrates consistent superiority under high confidence, with rates improving from 68.4% to 79.3% on average. This 15.9% relative improvement validates the effectiveness of stricter confidence thresholds in filtering marginal trading opportunities.

The 60.4% improvement in peak profitability (from 104.52 to 167.64 basis points) results from three mechanisms: (1) stricter confidence filtering eliminates 37% of marginal trades with win rates below 75%, (2) high-confidence trades capture larger average price movements (mean 2.8% vs 1.9% for moderate confidence), and (3) reduced false signals decrease drawdown periods by 42%. This performance differential remains consistent across 89% of cryptocurrency pairs tested, with statistical significance confirmed through paired t-test ($p = 0.019$).

4.5. Optimal Configuration Analysis

Distinct optimal parameter combinations emerge across confidence regimes, indicating regime-dependent parameter sensitivity rather than simple performance scaling. The moderate confidence regime favors longer horizons (500-600 minutes) with mixed deadband preferences, achieving maximum profitability through the H600-DB20 configuration.

High confidence regimes show preference for shorter optimal horizons (400 minutes) with lower deadband requirements, maximizing returns through H400-DB10. This horizon preference reversal suggests fundamental differences in signal dynamics under different confidence requirements.

Parameter diversity analysis shows that moderate confidence accepts a wider range of deadband values (2-20 basis points) among top performers, while high confidence strongly favors lower deadbands (2-10 basis points). This pattern reflects the interaction between confidence thresholds and signal quality requirements.

4.6. Economic Performance Metrics

The experimental results demonstrate economically significant returns under realistic trading conditions. Peak performance of 167.64 basis points per trade represents substantial value creation when applied to institutional-scale trading volumes.

Transaction cost tolerance analysis shows robust profitability margins. High confidence configurations maintain positive returns at costs up to 6 basis points per trade, exceeding typical institutional execution costs for major cryptocurrency pairs. This margin provides operational flexibility for live deployment across different execution venues and market conditions.

Statistical significance testing confirms that confidence threshold selection represents a fundamental strategic decision rather than marginal parameter tuning. Mean profit differences between regimes achieve statistical significance ($p = 0.019$), while coverage differences are highly significant ($p < 0.001$).

4.7. Feature Integration Effects

The unified dataset approach enables analysis of cross-temporal feature interactions. Post-hoc feature importance analysis reveals that optimal configurations combine momentum-based macro features with microstructure signals including bid-ask imbalances and order book depth ratios.

High-performing parameter combinations appear to capture the temporal bridge where daily momentum trends manifest in intraday order flow patterns. This temporal convergence explains the superior performance at intermediate horizons where daily directional bias has sufficient time to influence minute-level market microstructure.

The relationship between prediction horizons, confidence thresholds, and economic performance exhibits non-linear dynamics: shorter horizons (<100 minutes) require higher confidence thresholds ($\tau \geq 0.75$) to achieve profitability due to dominant microstructure noise, while intermediate horizons (400-600 minutes) maintain positive returns even at moderate confidence ($\tau = 0.6$) as daily momentum signals strengthen. Analysis shows that optimal deadband selection correlates inversely with prediction horizon ($r = -0.67$, $p < 0.01$), with longer horizons tolerating wider deadbands (15-20 bp) while maintaining signal quality.

5. Discussion

5.1. Economic Implications and Market Efficiency

The experimental results demonstrate that cryptocurrency direction prediction using integrated macro-microstructure features can generate economically significant returns under realistic trading conditions. The peak performance of 167.64 basis points per trade represents substantial value creation when applied to institutional-scale trading volumes.

These findings challenge the strong form of market efficiency in cryptocurrency markets. The consistent profitability across multiple parameter configurations suggests that exploitable inefficiencies exist at specific temporal scales. However, the coverage-profit trade-off reveals that genuinely predictable price movements occur infrequently, aligning with semi-strong market efficiency where only sophisticated analytical approaches can extract value.

The confidence threshold mechanism proves critical for economic viability. High confidence regimes achieve 60.4% higher peak profits than moderate confidence conditions, demonstrating that precision-recall optimization directly translates to economic performance. This relationship validates the hypothesis that separating directional prediction from execution decisions improves trading system effectiveness.

Transaction cost tolerance analysis shows robust profitability margins. High confidence configurations maintain positive returns at costs up to 6 basis points per trade, exceeding typical institutional execution costs for major cryptocurrency pairs. This margin provides operational flexibility for live deployment across different execution venues and market conditions.

5.2. Methodological Contributions and Framework Effectiveness

The binary classification approach addresses fundamental limitations in traditional three-class prediction frameworks. By decoupling directional prediction from execution timing, the methodology eliminates contamination between signal extraction and risk management decisions. This separation enables explicit optimization of the precision-recall trade-off, resulting in superior economic performance.

The unified macro-microstructure feature integration captures temporal bridges where daily momentum trends manifest in intraday order flow patterns. This integration explains the superior performance at intermediate horizons (400-600 minutes) where daily directional bias has sufficient time to influence minute-level market microstructure.

Multi-scale feature integration provides superior signal representation compared to single-timeframe approaches. The combination of momentum-based macro features with microstructure signals including bid-ask imbalances and order book depth ratios contributes to directional accuracies that exceed published benchmarks while maintaining economic profitability.

The temporal validation framework with symbol-wise splitting prevents data leakage while maintaining realistic trading conditions. This methodology ensures that all performance estimates reflect achievable returns under practical deployment constraints.

5.3. Comparison with Existing Literature

Our results compare favorably with published cryptocurrency prediction studies. The directional accuracy of 75-95% on executed trades substantially exceeds typical classification performance reported in the literature, which ranges from 60-65%. However, direct comparison remains challenging due to different evaluation frameworks and temporal scales.

The per-trade profit results (104-168 basis points for best configurations) represent a different performance metric from the profit factors and Sharpe ratios commonly reported. The Sharpe ratios achieved by similar studies (2.5-3.6) suggest comparable risk-adjusted performance levels, indicating potential performance ceilings in cryptocurrency markets.

The confidence-based approach offers comparable economic returns through a fundamentally different methodological pathway than existing ensemble or graph-based methods. The explicit precision-recall control provides operational advantages for live trading deployment.

5.4. Practical Implementation Considerations

Live deployment requires addressing several operational challenges not fully captured in backtesting environments. The confidence threshold mechanism demands real-time probability calibration as market regimes shift, potentially requiring adaptive threshold adjustment beyond the fixed values evaluated experimentally.

Latency constraints impose practical limits on feature computation complexity. The 64-feature unified representation requires approximately 15 milliseconds calculation time on standard hardware, compatible with minute-frequency decision cycles but potentially restrictive for higher-frequency applications.

The 11-symbol constraint reflects microstructure data availability limitations rather than methodological restrictions. Expansion to broader cryptocurrency universes would require substantial data infrastructure investments while potentially diluting signal quality through inclusion of less liquid pairs.

Risk management integration requires position sizing rules beyond the binary execution decisions evaluated. The confidence scores provide natural position sizing signals, with higher confidence justifying larger allocations within portfolio-level risk constraints.

5.5. Limitations and Constraints

Several limitations constrain the generalizability of these results. The evaluation period coincides with specific cryptocurrency market conditions that may not persist across different regulatory environments or institutional adoption phases. The temporal scope represents a particular market regime that may not generalize to future conditions.

The 11-symbol subset limits diversification benefits and may not represent broader cryptocurrency market dynamics. The constraint reflects microstructure data availability rather than methodological limitations, but restricts the scope of conclusions.

Feature engineering relies on domain expertise for macro-microstructure integration rather than automated discovery methods. Deep learning approaches for cross-temporal feature learning could potentially uncover signal patterns not captured by traditional technical indicators.

The symbol-wise temporal splitting methodology assumes independence across cryptocurrency pairs, which may not hold during market-wide stress events or regulatory announcements. Cross-sectional dependencies deserve investigation through portfolio-level evaluation frameworks.

Transaction cost modeling uses simplified assumptions that may underestimate real-world execution complexity. Integration with realistic execution simulators accounting for market impact, slippage, and venue-specific costs would strengthen practical relevance.

5.6. Future Research Directions

Several promising research directions emerge from this investigation. Adaptive confidence threshold mechanisms responsive to changing market conditions could improve performance consistency across different market regimes. Integration with portfolio optimization frameworks would extend the methodology beyond directional prediction to comprehensive trading system design.

Extension to traditional financial assets where similar macro-microstructure relationships may exist represents a natural progression. The methodological contributions - confidence-based execution control, unified multi-scale feature integration, and systematic parameter optimization - provide frameworks applicable beyond cryptocurrency markets.

Multi-task learning approaches incorporating volatility and correlation prediction could enhance portfolio construction beyond directional prediction. The binary classification framework could be extended to include risk factor modeling and position sizing optimization.

Deep learning approaches for automated cross-temporal feature discovery could potentially uncover signal patterns not captured by traditional technical indicators. Graph neural networks could capture cryptocurrency interdependencies more effectively than the current symbol-wise approach.

6. Conclusions

This research introduces a binary classification framework for cryptocurrency direction prediction that systematically integrates macro and microstructure features across multiple temporal scales. The methodology separates directional prediction from execution decisions through confidence-based thresholds, enabling explicit optimization of the precision-recall trade-off.

Comprehensive experiments across 11 major cryptocurrency pairs demonstrate economic viability under realistic trading conditions. High confidence regimes achieve peak profits of 167.64 basis points per trade with directional accuracies of 82-95% on executed trades. Moderate confidence regimes maintain 50-65% market coverage while generating profits of 104.52 basis points per trade.

The systematic parameter optimization reveals fundamental trade-offs between trading frequency and signal quality in cryptocurrency markets. Optimal performance occurs at intermediate prediction horizons where daily momentum trends manifest through intraday order flow patterns. The confidence threshold mechanism proves critical for economic performance, with high confidence requirements improving profits by 60.4% while reducing coverage by approximately 99%.

Multi-scale feature integration provides superior signal representation compared to single-timeframe approaches. The unified combination of macro momentum indicators with microstructure dynamics captures temporal bridges where fundamental price discovery mechanisms align with short-term market making activities.

The research demonstrates practical viability for institutional cryptocurrency trading applications. High confidence strategies tolerate transaction costs up to 6 basis points per trade while maintaining positive returns, exceeding typical execution costs for major cryptocurrency pairs. The framework's robust performance across different parameter configurations provides operational flexibility for live deployment.

The methodological contributions extend beyond cryptocurrency markets to other directional prediction domains with comparable signal quality trade-offs. The confidence-based execution control, unified multi-scale feature integration, and systematic parameter optimization provide frameworks applicable to various financial prediction problems.

Future research should investigate scalability across broader cryptocurrency universes, adaptive confidence mechanisms, and integration with portfolio optimization frameworks. The binary classification approach offers a foundation for developing more sophisticated trading systems that balance signal quality with operational constraints.

7. Data Availability Statement

The complete codebase for this research, including data processing, model implementation, and visualization scripts, is freely available at https://github.com/KuznetsovKarazin/crypto-confidence-execution. This accessibility enables direct verification of our results and facilitates further extension of our work by interested researchers.

Declaration on Generative AI

The authors have not employed any Generative AI tools.

References

[1] J. Zhang, K. Cai, J. Wen, A survey of deep learning applications in cryptocurrency, iScience 27 (2024) 108509. https://doi.org/10.1016/j.isci.2023.108509.

[2] D. Shang, Z. Guo, H. Wang, Enhancing digital cryptocurrency trading price prediction with an attention-based convolutional and recurrent neural network approach: The case of Ethereum, Finance Research Letters 67 (2024) 105846. https://doi.org/10.1016/j.frl.2024.105846.

[3] C. Zhong, W. Du, W. Xu, Q. Huang, Y. Zhao, M. Wang, LSTM-ReGAT: A network-centric approach for cryptocurrency price trend prediction, Decision Support Systems 169 (2023) 113955. https://doi.org/10.1016/j.dss.2023.113955.

[4] A. Pellicani, G. Pio, M. Ceci, CARROT: Simultaneous prediction of anomalies from groups of correlated cryptocurrency trends, Expert Systems with Applications 260 (2025) 125457. https://doi.org/10.1016/j.eswa.2024.125457.

[5] P. Peng, Y. Chen, W. Lin, J.Z. Wang, Attention-based CNN–LSTM for high-frequency multiple cryptocurrency trend prediction, Expert Systems with Applications 237 (2024) 121520. https://doi.org/10.1016/j.eswa.2023.121520.

[6] A. Youssefi, A. Hessane, I. Zeroual, Y. Farhaoui, Optimizing Forecast Accuracy in Cryptocurrency Markets: Evaluating Feature Selection Techniques for Technical Indicators, CMC 83 (2025) 3411–3433. https://doi.org/10.32604/cmc.2025.063218.

[7] A. Golnari, M.H. Komeili, Z. Azizi, Probabilistic deep learning and transfer learning for robust cryptocurrency price prediction, Expert Systems with Applications 255 (2024) 124404. https://doi.org/10.1016/j.eswa.2024.124404.

[8] A. C.v., N. Negi, A. Aprem, Bayesian machine learning framework for characterizing structural dependency, dynamics, and volatility of cryptocurrency market using potential field theory, Expert Systems with Applications 261 (2025) 125475. https://doi.org/10.1016/j.eswa.2024.125475.

[9] M. Kang, J. Hong, S. Kim, Harnessing technical indicators with deep learning based price forecasting for cryptocurrency trading, Physica A: Statistical Mechanics and Its Applications 660 (2025) 130359. https://doi.org/10.1016/j.physa.2025.130359.

[10] W. Liu, X. Bao, X. Han, Y. Li, Liquidity commonality in cryptocurrencies, Finance Research Letters 85 (2025) 108187. https://doi.org/10.1016/j.frl.2025.108187.

[11] Y. Yang, X. Wang, J. Xiong, L. Wu, Y. Zhang, An innovative method for short-term forecasting of blockchain cryptocurrency price, Applied Mathematical Modelling 138 (2025) 115795. https://doi.org/10.1016/j.apm.2024.115795.