Predicting High-Frequency Cryptocurrency Direction using Hybrid CNN-LSTM Networks

Michel Saliba

March 15, 2026

Abstract

This preliminary study investigates the efficacy of a hybrid Convolutional Neural Network and Long Short-Term Memory (CNN-LSTM) architecture for high-frequency cryptocurrency forecasting. Utilizing Open, High, Low, Close, and Volume (OHLCV) data augmented with engineered technical features, our initial classification models demonstrate that a hierarchical architecture—where 1D Convolutions extract local microstructural features and LSTMs model temporal momentum—can achieve statistically significant out-of-sample directional hit rates. Preliminary multi-timeframe stability tests on XRP indicate that this predictive edge is heavily concentrated in lower timeframes (1-minute and 5-minute), rapidly degrading to a random walk by the 30-minute interval. Subsequent 5-minute cross-asset testing revealed that only two specific coins exhibited a measurable classification edge, suggesting that structural inefficiencies are more readily exploitable in niche, lower-liquidity assets. Building upon these directional baselines, we outline a roadmap for continuous price prediction via rigorous hyperparameter optimization and the implementation of a dynamic, rolling-window training architecture designed to systematically adapt to shifting market regimes.

1 Introduction

Financial time series, particularly in cryptocurrency markets, exhibit high noise-to-signal ratios and non-stationary volatility regimes. Standard Recurrent Neural Networks (RNNs) or LSTMs often overfit to this micro-level noise when fed raw OHLCV sequences.

The motivation behind the Hybrid CNN-LSTM approach is to force the network to learn a denoised, hierarchical representation of the market. Convolutional layers act as local feature extractors, learning spatial relationships across the multi-channel input (e.g., volume spikes corresponding with wicks). The LSTM subsequently models the temporal sequence of these condensed, high-level features. This architectural framework was initially inspired by Livieris et al. [LPP20], who demonstrated the efficacy of a similar hybrid CNN-LSTM methodology for gold price forecasting.

While this report outlines our baseline architecture and provides early evidence of an out-of-sample directional edge, it also serves as the foundation for our next developmental phase. To fully capture alpha in shifting market environments, we are extending this framework from binary classification to continuous price prediction. Furthermore, to combat the non-stationarity inherent in crypto assets, we propose transitioning to a dynamic, rolling-window training architecture. By continuously retraining on recent data and discarding obsolete sequences, this walk-forward approach ensures the model systematically adapts to the newest market trends and volatility regimes.

2 Methodology

Our baseline model utilizes historical market data to predict the probability of a positive forward 1-step return, framing the problem as a binary classification task.

2.1 Data Pre-processing and Feature Engineering

To address non-stationarity and define the classification target, prices are first transformed into log returns:

$$r_t = \ln\left(\frac{C_t}{C_{t-1}}\right) \tag{1}$$

where C_t represents the closing price at time step t. The target variable Y_t is then defined as a binary indicator for an upward forward movement:

$$Y_t = \begin{cases} 1 & \text{if } r_{t+1} > 0 \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

Feeding raw price data into neural networks often results in poor generalization. Therefore, alongside the core OHLCV data, the feature set is augmented with engineered technical indicators. Specifically, we incorporate the 14-period Relative Strength Index (RSI) to provide the network with bounded momentum context.

The sequence length is fixed at $N = 40$ steps. The input matrix at time t is $X_t \in \mathbb{R}^{C\times 40}$ (where C is the number of feature channels), which is standard-scaled using strictly in-sample data to prevent forward-looking bias.

2.2 Network Architecture and Training Limitations

The architecture is divided into two discrete processing blocks:

- **Spatial Feature Extraction (CNN):** The network applies two sequential 1D Convolutional layers (kernel size $k = 3$) separated by Max Pooling ($k = 2$) and ReLU activations. This compresses the 40-step input sequence into a condensed series of higher-order feature maps.
- **Temporal Sequence Modeling (LSTM):** The resulting tensor is fed into a 2-layer LSTM with a hidden dimension of 64 and a dropout rate of 0.2. The final hidden state h_N is passed to a fully connected layer optimized via Binary Cross-Entropy with Logits (BCEWithLogitsLoss) to output the unnormalized probability of a positive price movement.

Hardware and Convergence Constraints: It is critical to note that due to current computational resource limitations, the training loop for these preliminary tests was strictly capped at 25 epochs per run. Consequently, we cannot guarantee full convergence of the loss function. The results presented herein should be viewed as an unoptimized structural baseline rather than the architecture's theoretical maximum performance.

3 Preliminary Results and Backtesting

The model was evaluated strictly out-of-sample (OOS) across 10 independent runs to test structural stability. To translate the continuous network output into a discrete directional signal, we applied confidence thresholds. Signals are generated only when the model's pseudo-probability exceeds the noise band:

$$S_t = \begin{cases} 1 & \text{if } \hat{p} > 0.53 \\ -1 & \text{if } \hat{p} < 0.47 \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

where $\hat{p}$ is the sigmoid-transformed output.

3.1 Multi-Timeframe Stability

Timeframe stability tests on XRP reveal a strong inverse correlation between interval length and predictive edge.

As seen in Figure 1:

- **1-Minute Interval:** Achieved the highest average directional hit rate ($\sim$57.5%), exhibiting robust stability across runs.

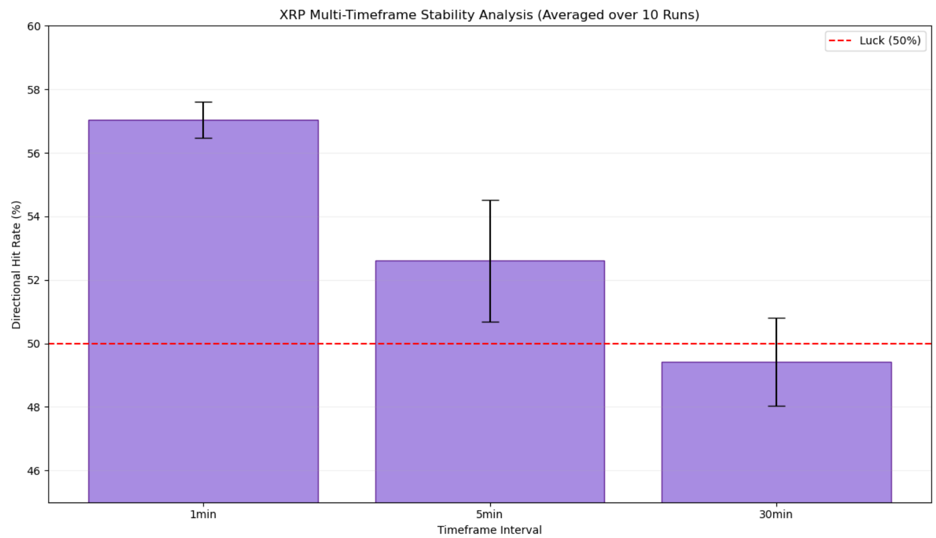

Figure 1: Timeframe stability tests indicating directional hit rates across different intervals for XRP.

- **5-Minute Interval:** Maintained a profitable edge at ∼52.5%, though variance between runs increased.
- **30-Minute:** Performance degraded significantly, showing a random walk.

3.2 Multi-Asset Agnostic Test

An agnostic test applying the same architecture across BTC, ETH, XRP, ADA, BCH and DOGE at the 5-minute interval confirmed varying degrees of microstructural edge.

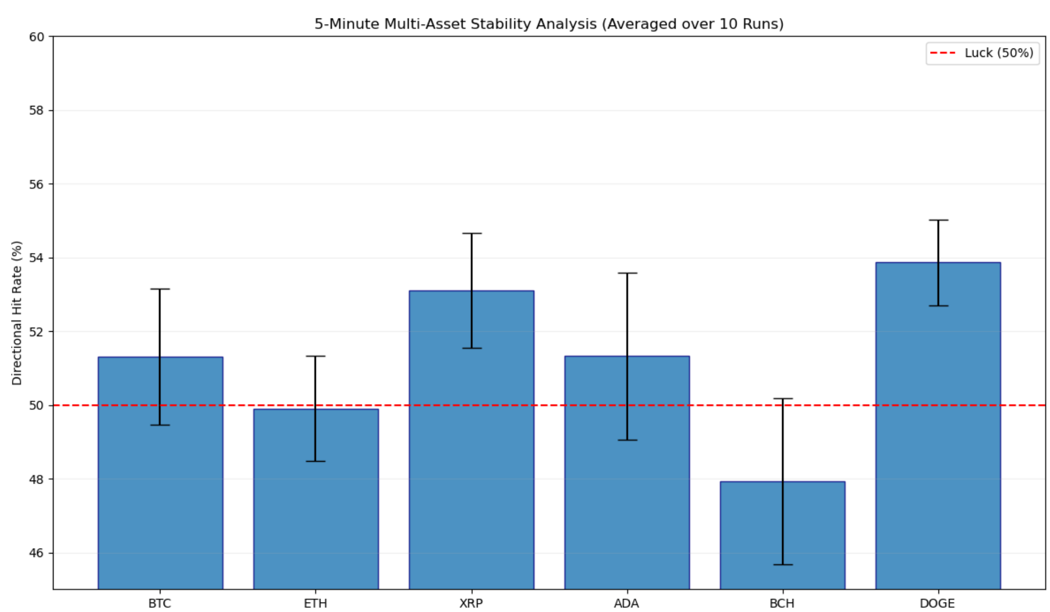

Figure 2: Multi-asset 5-minute stability test results showing hit rates across BTC, ETH, and XRP.

As shown in Figure 2, XRP and DOGE yielded the highest mean hit rates (> 52%), while BTC, ETH, ADA and BCH hovered around 50%, showing pure randomness. This variance suggests the current feature set and hyperparameters (e.g., 25 epochs) may be capturing asset-specific high-frequency order flow imbalances rather than a universally applicable momentum factor.

4 Conclusion and Future Directions

The baseline CNN-LSTM architecture demonstrates a measurable out-of-sample directional edge, specifically isolated to high-frequency intervals (1m, 5m).

To evolve this from a preliminary baseline to a production-ready trading model, our immediate next steps include:

1. **Extended Training & Hyperparameter Tuning:** Removing the 25-epoch cap and utilizing early stopping mechanisms to ensure model convergence, alongside rigorous tuning of learning rates and network dimensions.
2. **Multi-Timeframe (MTF) Fusion:** Bridging the performance gap by feeding 1m and 1h data into parallel network branches, fusing micro-structure edge with macro-trend context.
3. **Dynamic Walk-Forward Optimization:** Implementing a rolling-window training architecture to continuously adapt to shifting market regimes.
4. **Continuous Price Regression:** Transitioning from binary classification to direct price forecasting to estimate specific return magnitudes, enabling rigorous back-testing for high-frequency trading (HFT) strategies.

References

[LPP20] Ioannis E Livieris, Emmanuel Pintelas, and Panagiotis Pintelas. A cnn–lstm model for gold price time-series forecasting. *Neural computing and applications*, 32(23):17351–17360, 2020.