Report

Investment Strategies based on Machine Learning and Deep Learning for Bitcoin

Oussama Taghlaoui **Ismaili Abdelh** **Ossama Lboudali**

Abstract

Abstract — This report presents the design, implementation, and evaluation of quantitative investment strategies applied to Bitcoin (BTC/USDT), based on an ensemble of Machine Learning and Deep Learning models. By integrating massive heterogeneous databases (price history, advanced technical indicators, and macroeconomic filters), the study demonstrates the ability of these predictive algorithms to drastically outperform a passive "Buy & Hold" management strategy. The work highlights the hegemony of ensemble models (Random Forest) for maximizing gross return, as well as the structural advantage of deep sequential networks (GRU) for risk adjustment and drawdown minimization.

1 Introduction

1.1 General Project Context

Cryptocurrency markets, particularly Bitcoin (BTC), are characterized by extreme volatility, continuous data availability (24/7), and a strong influence from macroeconomic and global informational factors. In this context, purely linear prediction methods or those based solely on manual technical analysis reach their limits. The integration of Artificial Intelligence (AI) now makes it possible to effectively exploit this mass of temporal and textual data to extract viable trading signals.

1.2 Research Problem

How to design a robust predictive architecture combining price history, technical indicators, and financial news analysis to predict Bitcoin's price direction at an hourly level with profitability superior to classic market benchmarks?

1.3 Scientific and Technical Objectives

- **Algorithmic Development**: Implement and compare a wide range of classical machine learning and deep learning models (10 models in total).
- **Feature Engineering**: Build a rich and complex dataset (>90 features).
- **Production Deployment**: Design an analytical dashboard communicating with a high-performance API.

2 Methodology

2.1 Data Extraction and APIs

The architecture relies on several streams:

- **Binance API**: Hourly price history (OHLCV) with over 69,000 observations.
- **News RSS Feeds**: Real-time integration of articles from 9 specialized sources.
- **Macro Data**: Traditional indices (S&P500), DXY, physical gold, Hash Rate, and NVT Ratio.

2.2 Feature Description

The descriptive spectrum includes over 90 variables:

- **Technical Indicators**: RSI, MACD, Bollinger Bands, Moving Averages.
- **Lags**: Time delays to capture return autocorrelation.
- **Textual Variables**: Polarity scores derived from news via the extraction of Transformer-type embeddings.

2.3 Model Presentation

Three families of algorithms were developed concurrently:

1. **Classical ML**: Random Forest, XGBoost, SVM, Logistic Regression, and Naive Bayes.
2. **Chronological Deep Learning**: Multilayer Perceptron (MLP), Recurrent Neural Networks (LSTM, GRU).
3. **NLP/Hybrid Models**: Combination of textual features extracted with a CNN-BiLSTM network.

3 Experimental Results

3.1 Model Comparison

The evaluation is based on a strict temporal cross-validation set ("Test Set"). Although Accuracy is the primary metric, robustness requires considering precision, recall, and F1-score during optimization to limit false positives in trading.

Note: Accuracy in high-frequency trading (53%+) is considered statistically very significant to generate a positive mathematical expectancy of profit.

Table 1: Experimental Performance Metrics on the Test Set

Model / Algorithm	Accuracy	Complexity	Behavioral Nature
Random Forest	**53.27%**	Moderate	High noise tolerance, excellent F1-score
XGBoost	53.17%	Moderate	Highly competitive, sensitive to hyperparameters
Logistic Regression	52.66%	Low	Robust baseline for linear features
SVM	52.61%	High	High computational cost, consistent
MLP	52.39%	High	Detects tabular non-linearity well
Naive Bayes	52.27%	Low	Fast but assumes feature independence
LSTM	50.59%	Very High	Sensitive to noise on pure price history
GRU	51.69%	Very High	Better temporal smoothing than LSTM
LSTM-CNN	51.73%	Very High	Captures local patterns via CNN
CNN-BiLSTM (Hybrid)	48.79%	Very High	Impacted by the lack of massive labeled data

4 Backtesting

4.1 Backtesting Protocol

The ultimate validator of a financial model is not Accuracy, but investment simulation (Backtesting). Our environment simulates real broker conditions. Key metrics include: ROI (Return on Investment), Sharpe Ratio (Risk-adjusted return), and Maximum Drawdown.

4.2 Simulated Performances

The gap between Train and Test was meticulously monitored to analyze overfitting. GRU and LSTM networks required strict "Dropout" mechanisms as they tended to overfit on history. In contrast, the Random Forest demonstrates excellent generalization with very little overfitting.

Table 2: Financial Simulation Metrics (Backtesting)

Strategy / Model	Secured ROI	Sharpe Ratio	Max Drawdown
Benchmark (Buy & Hold)	*+39.4%*	*0.68*	*-34.8%*
Random Forest	**+425.3%**	**2.45**	-30.4%
XGBoost	+273.5%	2.00	-25.5%
Logistic Regression	+173.9%	1.58	-25.4%
MLP	+153.7%	1.48	-26.0%
SVM	+62.0%	0.88	-28.7%
GRU (Low Risk)	+59.4%	**2.70**	**-4.9%**
LSTM-CNN	+49.8%	2.29	-5.3%
Naive Bayes	+46.5%	0.75	-44.2%
Standard LSTM	+17.1%	0.03	-4.5%

5 Comparative Analysis

5.1 The Hegemony of Trees

Traditional algorithms based on decision trees (Random Forest, XGBoost) dominate the ranking by gross return (+425.3% ROI). Their intrinsic ability to evaluate high-dimensional variables without disproportionately capturing random market 'noise' makes them by far the most performant models.

5.2 Risk Adjustment via DL

Where Machine Learning is purely offensive, Deep Learning shows unique defensive capabilities. The **GRU** model displays the best Sharpe Ratio (2.70) and the lowest Drawdown (-4.9%). It acts as an extremely secure financial hedge, proving the superiority of recurrent networks to avoid catastrophic signals.

6 Dashboard

To actualize the AI architecture, the project includes a complete decision-support platform:

- **Backend (FastAPI)**: Dynamic interface loading pre-trained models ('.pkl', '.h5'), enabling fast live inference and scheduled scraping.
- **Frontend (React)**: Dashboard centralizing the metrics:
  - *AI Consensus Tracker*: Final signal (BUY, SELL, WAIT).
  - *Real-time Narrative Flow*: Display of sentiment labels (Bullish/Bearish).
  - *Leaderboard*: Interactive ranking of model accuracy.

7 Conclusion

7.1 Synthesis of the Optimal Model

The architecture scientifically confirms the viability of predictive Machine Learning in cryptocurrency. The **Random Forest** is indisputably the best model with a spectacular backtest (**+425.3% ROI**, **53.27% Accuracy**, and an excellent Sharpe Ratio of 2.45), overwhelmingly superior to classic investment strategies.

7.2 Limitations and Improvements

The NLP approach remains a bottleneck: the hybridization (CNN-BiLSTM) suffers from a lack of labeled data scaling. Future improvements focus on:

1. The massive expansion of historical press databases ("News Scaling").
2. The structural adjustment of deep DL algorithm hyperparameters (e.g., Focal Loss) to cross a new threshold of precision.