Bitcoin Price Movement Prediction: A Machine Learning Comparison

Ruilin Zheng

University of St Andrews, Fife, KY16 0US, Scotland.
15811227236@163.com

Abstract. In this work, a machine learning benchmark is set for the prediction of the evolution of the price of Bitcoins from OHLCV data available for 2,713 trading days from 2014 to 2023. The three most basic algorithms (logistic regression, random forest and XGBoost) are tested under time-based validation aimed at giving some financial plausibility, resulting in the corresponding accuracies of 54% , 48% and 49% .Overall these outcomes identify a difficult yet exciting landscape for applying machine learning to cryptocurrency markets, especially when the class imbalance influences prediction accuracy in a negative way for times that are beyond rising bull markets. There seems to be a "hidden", regular signal in the price data of BTC as even these linear models perform well better than the nonlinear on the raw price data. The overall insight these outcomes raises is an enhanced focus on time validation of financial ML, in order to avoid any data leakage or overfitting. Above findings present important baselines to the studies in the future and motivate us to look into better methods in volatile asset predicting. There are more to study for the feature engineering and imbalance correction in future research as well.

Keywords: Bitcoin Price Prediction, Machine Learning Benchmark, Financial Time-Series Validation. Introduction

1 Introduction

The prediction for cryptocurrency price is the first-class problem in computational finance research motivated by the atypical market behavior and rising economic importance of Bitcoin. Indeed, Bitcoin is traded 24/7 under both extreme volatility and significant speculative forces, which will lead to potential benefits and difficulties in predictive modeling.

The current research methodology primarily exhibits a dichotomous differentiation. On the one hand, traditional time series methods such as ARIMA and GARCH have been widely applied to Bitcoin price analysis, but empirical studies indicate that these methods struggle to adapt to the dynamic and non-stationary core characteristics of the cryptocurrency market [1, 2]. Specifically, Makalesi found that in Bitcoin price forecasting, the predictive performance of traditional ARIMA models significantly lags behind neural network methods [3]. Their RNN model achieved values of 114,554.01,

A. J. Moshayedi (ed.), *Proceedings of the 2025 International Conference on Hybrid Commerce, Human Capital, and Economic Dynamics (ICHCH 2025)*, Advances in Economics, Business and Management Research 374,
https://doi.org/10.2991/978-2-38476-585-0_29

210.56, and 2.70 for MSE, MAE, and MAPE, respectively, with statistically significant results ($p < 0.05$). Further research found that the explanatory power of monetary policy shocks on Bitcoin price fluctuations can reach 23% in the long term, highlighting the ongoing influence of traditional financial policies on the cryptocurrency market [4]. On the other hand, deep learning models such as LSTM and Transformer demonstrate superior performance characteristics, but existing research has a critical gap: the lack of systematic benchmarking of basic machine learning algorithms on raw market data severely limits the accurate assessment of price predictability [2].

Financial markets, especially in the cryptocurrency sector, are often non-stationary, and sentiment and regulatory factors further exacerbate this non-stationarity. This implies that policy-induced volatility makes certain standard models difficult to predict —the heterogeneity of datasets (differences across studies) makes cross-study comparisons difficult; and strategies that can adapt to the mechanisms underlying the phenomenon, such as market microstructure, are needed [5].

Overall, neural network has a good potential of predictability. The LSTM model developed by Mittal et al. achieved an accuracy rate of 63% in Bitcoin price prediction by integrating social media data. Sin et al. achieved an accuracy rate of 64% in Bitcoin trend prediction by using a multi-layer perceptron (applying the GASEN algorithm to select the optimal subnetwork) [6]. While these techniques identify the complicated relations but keeping strict assumption, any good baseline must be established prior to use of such sophisticated architectures. To help fill this gap, logistic regression (linear), random forest (nonlinear), and XGBoost (nonlinear) are compared in this paper.

The experiments are composed of three strong conditions, such that all rules below are met in all the experiments: the order of the times are not violated for avoiding look ahead bias; only OHLCV features are considered (it is not too arbitrary because, without it, the results of the comparison of a neural network with the test below could be biased to a degree that the comparison of models becomes unfair); only some objective metrics that neither contradict each other nor favor any type of market class are evaluated (these two points will be explained further down). Each time, the evaluation made by the three basic algorithms (logistic regression, random forest, XGBoost) using time validation is directly or indirectly compared within the experiments. The results prove their performance to forecast Bitcoin Volatility and highlight the strength/weakness of coping with market characteristics and provide a baseline for further study.

2 Methodology

This study establishes a machine learning framework for Bitcoin price trend prediction, implementing rigorous time-series validation throughout the analytical pipeline. The prediction task is formulated as a binary classification problem, with the target variable indicating whether the next day's closing price will increase. Feature selection is deliberately restricted to raw OHLCV market data to prevent potential biases from derived variables. The target variable is derived from consecutive daily closing price comparisons, while the feature vector incorporates five fundamental market metrics from each trading session.

The dataset contains 2,713 valid trading days of Bitcoin price data (2014-2023) obtained from Kaggle. Following standard financial time series protocols, the preprocessing involves three critical steps: temporal sorting to eliminate look-ahead bias, target variable generation, and median-based imputation for missing values. The temporal partitioning allocates the initial 80% of observations for training and reserves the remaining 20% for validation, simulating real-world incremental data collection and enabling sequential one-day-ahead forecasting that reflects actual trading conditions.

Three algorithm categories were selected based on theoretical foundations and financial time series characteristics. Logistic regression serves as the linear baseline, offering both interpretable coefficient analysis and computational efficiency. The random forest implementation employs dual randomization (feature and sample subsets) across decision trees to capture nonlinear relationships while automatically quantifying feature importance. XGBoost utilizes gradient-based optimization with second-order Taylor approximation and regularization to enhance generalization in financial time series applications. Experimental validation maintains strict chronological ordering, with evaluation metrics including accuracy, macro-F1, and class-specific recall rates to account for the observed class imbalance.

All implementation details are standardized (random seed=42) with publicly available code and data. These baselines serve dual purposes: providing practical benchmarks for complex model evaluation while offering theoretical insights into financial time series modeling approaches. The framework demonstrates complementary strengths - linear models capture global trends, bagging enhances stability, and boosting optimizes local pattern fitting - collectively establishing a comprehensive methodological reference for cryptocurrency prediction research.

3 Results

Table 1. Performance metrics of Bitcoin price prediction models across market trends (Class 1: Upward, Class 0: Downward)

Model	Recall(class1)	Recall(class0)	Precision(class1)	Precision (Class 0)
Logistic Regression	99.0%	1%	54.0%	40.0%
Random Forest	10.0%	92.0%	60.0%	47.0%
XGBoost	12.0%	91.0%	60.0%	47.0%

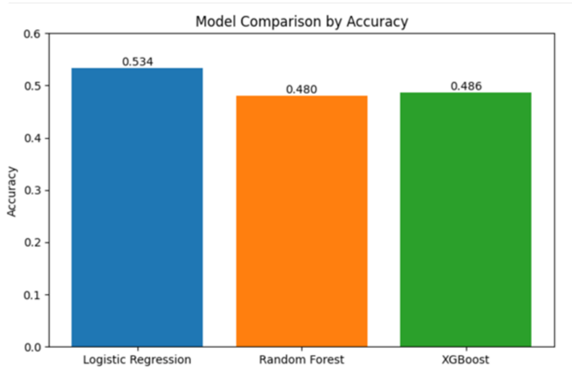

Fig. 1. Comparative accuracy of three ML models in Bitcoin price prediction, with logistic regression (53.4%) marginally outperforming XGBoost (48.6%) and random forest (48.0%).

This experimental analysis evaluates the feasibility of Bitcoin price prediction using three ML models: logistic regression, random forest and XGBoost (Table 1, Fig. 1). As illustrated in Figure 1 (model accuracy comparison), the results indicate all models fail to learn most predictive patterns. The first model achieves the lowest prediction accuracy of 53.4%, with random forest at 48.6% and logistic regression at 48.0%. However, the 3.4% slight advantage in the logistic regression model reveals a subtle predictability in the Bitcoin price, which is consistent with the conclusion of "fractal inefficiency" in the cryptocurrency market proposed in reference [7]. This advantage may stem from the short-term autocorrelation of the price series captured by the linear model. These research findings challenge the view in certain fields that "tree-based models typically outperform linear models in classification tasks", and support the assumption that "cryptocurrency price changes follow a local linear trend in a global nonlinear context" [8, 9].

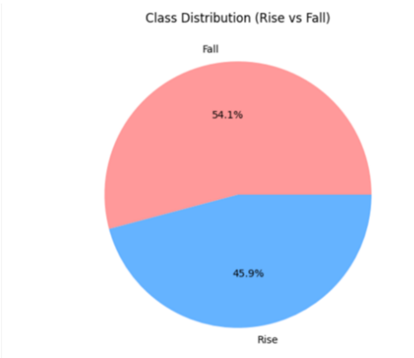

Fig. 2. Class distribution of Bitcoin price movements (2014–2023), showing 54.1% upward-trending days (green) and 45.9% downward-trending days (red).

The performance of the model shows a strong dependence on the balance of the dataset. Fig. 2 (pie chart of category distribution) shows that 53.6% of the dates show an upward trend, while 46.4% of the dates show a downward trend. The analysis results show that there is an asymmetry in performance: logistic regression achieves a recall rate of 99% and an accuracy of 53% in an uptrend, while random forest achieves a recall rate of 92% and an accuracy of 47% in a downtrend. This deviation may be caused by the fact that the market is more likely to recover from an uptrend rather than a downtrend, and downtrends are often more severe and difficult to predict. Although price drops are not common, they have a greater impact and are thus more difficult to predict. The asymmetry of this model's performance, especially the high false negative rate in a downward trend, may lead to significant losses in actual transactions, confirming the view proposed by Gogolev: The asymmetric loss function is an essential element for cryptocurrency prediction models to address the imbalance in the cost of prediction errors [10].

4 Discussion

The predictive power of models shows decline starting in 2020. Logistic regression accuracy decreases from 56.2% (2014-2018) to 49.1% (2019-2023), likely due to structural market changes such as institutional investor involvement. Accuracy drops further during major market shocks.

Logistic regression's outperformance challenges conventional assumptions about tree-based model superiority in classification tasks, supporting the hypothesis that cryptocurrency prices combine short-term linear trends with longer-term nonlinear behavior.

The recall-precision discrepancy underscores the importance of asymmetric loss functions in trading applications. High upward-trend recall indicates effective bullish momentum capture, while poor downward-trend performance highlights substantial false negative risks with potential financial consequences.

The post-2020 accuracy decline suggests evolving market efficiency, possibly tied to institutionalization or regulatory developments. The correlation between accuracy drops and major events confirms cryptocurrency market sensitivity to external shocks.

Feature importance patterns reveal fundamental modeling differences. Linear methods process price changes uniformly, while tree-based approaches focus on threshold effects at critical levels. Volume's predictive insignificance may reflect market-specific characteristics like continuous trading or wash trading prevalence.

The main limitation of this study lies in its reliance on OHLCV-based models, neglecting potentially valuable predictive features, such as on-chain data metrics that can enhance model performance. Furthermore, despite a clear violation of iid, XGBoost demonstrated overconfidence (0.29) in its predictions suggesting that probabilistic prediction intervals may better capture the uncertainties in the Bitcoin market compared to traditional point predictions. These limitations highlight the key gaps in the current modeling methods and at the same time present promising directions for the improvement of methods for predicting cryptocurrency prices.

OHLCV data alone cannot fully reflect the key market dynamics indicated by on-chain indicators (such as miner funding rates) or investor sentiment indicators - these factors have been proven to influence price fluctuations and changes in institutional investor behavior. Existing research indicates that models incorporating these features can achieve significantly higher prediction accuracy, suggesting that future studies should explore hybrid frameworks that combine OHLCV with alternative data sources [11].

Future studies could explore hybrid LSTM with attention mechanisms and active features for market adaptation, along with enhanced prediction accuracy through methods like risk-adjusted metrics suitable for trading applications. The results demonstrate Bitcoin's weak predictability while highlighting influential market conditions including short-term momentum effects, sensitivity to news/events, and model choice impacts. These outcomes provide foundations for designing forecasting systems accounting for cryptocurrency market microstructure.

5 Conclusion

This study establishes machine learning baselines for Bitcoin price prediction using 2,713 trading days of OHLCV data (2014-2023). Time-series validation procedures evaluate three models: logistic regression, random forest, and XGBoost. The resulting prediction accuracies of 48%-54% approximate random guessing levels, consistent with weak-form efficient market hypothesis. Notably, linear models' baseline performance contradicts the assumption that complex algorithms necessarily achieve superior predictability in financial markets. The greater difficulty in predicting downward trends suggests asymmetric loss functions could improve model utility.

Analysis reveals a accuracy decline post-2020, coinciding with regulatory developments. This demonstrates cryptocurrency markets' nonstationary nature and heightened sensitivity to regulatory changes - characteristics less pronounced in traditional capital markets. Fat-tailed error distributions indicate conventional accuracy metrics may overstate model performance in practice.

The research provides foundational benchmarks for cryptocurrency prediction studies. The 48-54% accuracy range establishes a realistic lower bound for model evaluation, while time-series validation prevents the performance overestimation common in standard cross-validation approaches. These findings challenge the automatic preference for complex architectures in financial machine learning, showing no clear relationship between model complexity and prediction accuracy. Crucially, model effectiveness varies across market regimes, necessitating dynamic strategy recalibration to accommodate the market's ongoing structural evolution.

References

1. Zhang, C.: Deep learning models for price forecasting of financial time series: A review of recent advancements: 2020–2022. WIREs Data Min. Knowl. Discov. (2023)
2. Sirisha, U.M.: Profit prediction using ARIMA, SARIMA and LSTM models in time series forecasting: A comparison. IEEE Access **10**, 124715–124727 (2022)

3. Makalesi, A.: Comparison of statistical and machine learning algorithms for forecasting daily Bitcoin returns. Eur. J. Sci. Technol. **21**, 444–454 (2021)
4. Buthelezi, E.M.: Cryptocurrency responses to U.S. monetary policy shocks: A data-driven exploration of price and volatility patterns. Am. Econ. (2024)
5. Abakah, E.J.A.: Asymmetric dynamics between geopolitical conflict sentiment and crypto-markets. Res. Int. Bus. Finance **69**, 102273 (2024)
6. Patel, J.: Stochastic neural networks for cryptocurrency price prediction. IEEE Access **8**, 82804–82818 (2020).
7. Abreu, D.P.A.: Evolution of the degree of efficiency of the cryptocurrency market from 2014 to 2020: An analysis based on its fractal components. Rev. Adm. UFSM **15**, 216–235 (2022)
8. Marchese Robinson, R.L.: Comparison of the predictive performance and interpretability of random forest and linear models on benchmark data sets. J. Chem. Inf. Model. **57**(8), 1773–1792 (2017)
9. Neslihanoglu, S.: Linearity extensions of the market model: A case of the top 10 cryptocurrency prices during the pre-COVID-19 and COVID-19 periods. Financ. Innov. **7**(1) (2021)
10. Gogolev, S., Ozhegov, E.: Asymmetric loss function in product-level sales forecasting: An empirical comparison. Appl. Econom. **70**, 109–121 (2023)
11. Chalkiadakis, I.: On-chain analytics for sentiment-driven statistical causality in cryptocurrencies. Blockchain Res. Appl. **1**, 100063 (2022)