Proceedings of the 7th International Conference on Computing and Data Science
DOI: 10.54254/2755-2721/2025.26124

Optimizing AdaBoost for Bitcoin Price Prediction Based on Long and Short Term Memory Network Models

Chao Li[1*], Qingyang Wang[2], Xinyue Huang[3], Shuwei Sun[4], Ziqi Lin[5]

[1]*Georgetown University, Washington, USA*
[2]*University of Nebraska, Lincoln, Nebraska, USA*
[3]*Independent Researcher, New York, USA*
[4]*University of Pennsylvania, Philadelphia, USA*
[5]*Cornell University, New York, USA*
**Corresponding Author. Email: cl1486@georgetown.edu*

Abstract. In this paper, a hybrid time series analysis model for bitcoin price forecasting is constructed by introducing a long short-term memory network (LSTM) to deeply optimize the traditional AdaBoost integrated learning model. Experimental results show that the fusion model exhibits excellent dual adaptability in the field of financial time series forecasting: in the training phase, the model achieves a high degree of fit to historical data with an excellent performance of MAE 1.3574, MSE 3.4279, and MAPE 0.412%, and its coefficient of determination, R^2, breaks through 0.97105, which intuitively verifies the ability of the LSTM module in capturing the non-linear time-series characteristics of the The synergy between the LSTM module and the AdaBoost framework in strengthening the generalization ability of the model; in the testing phase, although facing the challenge of increased market volatility, the model still maintains a stable performance of MAE 1.914, RMSE 2.7893 and MAPE 0.586%, especially the prediction error rate is continuously controlled within 1% of the industrial-grade accuracy threshold, and its test set R^2 value of 0.88675 significantly exceeds the benchmark interval of 0.6-0.8 for conventional prediction systems, confirming the strong explanatory power and robustness of the model in unknown data environments.

Keywords: Long and short-term memory networks, AdaBoost, Bitcoin price prediction

1. Introduction

As the world's first decentralized cryptocurrency, Bitcoin's price volatility and market impact have sparked widespread interest in both academia and finance since its inception in 2009. The Bitcoin market is characterized by high volatility, non-linear characteristics, and being driven by multiple factors. Its price is not only influenced by the traditional financial market, but also closely related to unique factors such as blockchain technology development, regulatory policy changes, market sentiment, social media public opinion, miners' behavior, and cybersecurity events [1]. El Salvador's decision to make bitcoin a legal tender in 2021 directly contributed to its short-term price surge, while the LUNA crash in 2022 triggered a cryptocurrency market systemic panic, leading to a slump

© 2025 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2025.26124

in the price of Bitcoin. Traditional financial forecasting models have limited performance in dealing with such complex, highly noisy and non-stationary data, as they struggle to capture non-linear relationships and long-term dependencies between variables. In addition, the Bitcoin market lacks a sophisticated valuation framework, and its price is driven more by supply and demand and speculative behavior, further increasing the difficulty of forecasting. As a result, researchers have begun to explore more sophisticated data-driven methods, and machine learning has gradually become a core tool for bitcoin price prediction due to its powerful pattern recognition capabilities and adaptability to high-dimensional data [2].

Machine learning provides a new paradigm for bitcoin prediction by automatically mining potential patterns of price fluctuations from historical data. Traditional ML algorithms such as Support Vector Machines, Random Forests and Gradient Boosting Trees are used to build prediction models based on technical indicators, on-chain data and sentiment analysis [3]. Random forests can reveal the dominance of exchange capital flows through feature importance ranking, while SVMs efficiently delineate price trend categories in high-dimensional feature spaces [4]. In recent years, deep learning has further improved prediction accuracy: long and short-term memory networks excel at capturing time-series dependencies of Bitcoin prices, Transformer models parse multi-scale market signals through the self-attention mechanism, and graph neural networks are used to analyze cross-asset correlations in the cryptocurrency market. In this paper, we optimize the AdaBoost model for bitcoin price prediction based on the long and short-term memory network model.

2. Data set source

2.1. Data set introduction

The dataset used in this paper is from the open-source dataset, which covers the bitcoin opening price changes from April 9, 2023 to April 9, 2024, and can be used for time-series analysis and forecasting, and some of the datasets are presented as shown in Table 1.

Table 1. Some of the data

Timestamp	Open
2023/4/9 0:00	310.5739
2023/4/9 1:00	311.1801
2023/4/9 2:00	311.6132
2023/4/9 3:00	311.2814
2023/4/9 4:00	311.245
2023/4/9 5:00	311.1659
2023/4/9 6:00	311.0899
2023/4/9 7:00	310.0464
2023/4/9 8:00	309.9725
2023/4/9 9:00	310.0638
2023/4/9 10:00	309.9304
2023/4/9 11:00	310.1935
2023/4/9 12:00	310.4562

DOI: 10.54254/2755-2721/2025.26124

3. Method

3.1. LSTM

LSTM (Long Short-Term Memory Network) is a recurrent neural network specifically designed to process sequential data, and its core innovation is to solve the problem of capturing long-distance dependencies that are difficult to be captured by traditional models through "gating system" and "memory channel". Each LSTM unit contains three key gates (forgetting gate, input gate, output gate) and a long-term memory channel throughout [5]. The role of the forgetting gate is to determine which historical information needs to be retained or discarded, and it calculates a weight value between 0 and 1 from the current input and the hidden state of the previous moment, with 0 representing complete forgetting and 1 representing complete retention. The input gate is then responsible for filtering the current input for valuable new information and fusing it into long-term memory [6]. The output gate ultimately controls which information needs to be passed to the hidden state at the current moment for subsequent computation or as output. This gating mechanism allows the LSTM to dynamically regulate the strength of the information flow, both to avoid the interference of irrelevant information and to ensure the long-term transfer of key features [7].The network structure diagram of the LSTM is shown in Fig. 1.

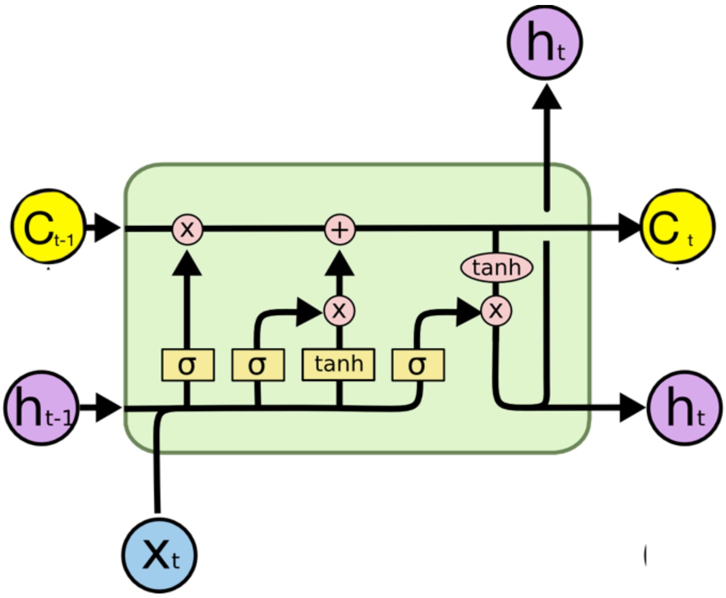

Figure 1. The network structure diagram of the LSTM

Information Transfer and Memory Updating The core of LSTM lies in the process of cell state updating, which is like a stable conveyor belt throughout the entire time sequence. At each step, the forgetting gate first decides how much old information to retain from long-term memory, e.g., when predicting the next word in a sentence, it may be necessary to forget irrelevant details from the previous text. The input gate then generates new information through a two-part operation: one part filters the current input for what needs to be updated, and the other part generates candidate memory values. The results of these two parts are combined and added to the old memories retained by the forgetting gate to form the updated cell state. Finally, the output gate compresses the updated memory to a reasonable range by a nonlinear activation function and selectively outputs it into the hidden state as the output of the current moment or passes it to the next step. The whole process is

DOI: 10.54254/2755-2721/2025.26124

synergized by gating to maintain the stability of long-term memory and flexibly integrate short-term features.

3.2. AdaBoost

AdaBoost is an integrated learning algorithm based on the idea of Boosting, whose core idea is to combine multiple weak classifiers (e.g., decision tree stumps) into one strong classifier by gradually adjusting the sample weights and model weights.AdaBoost assigns a weight to each sample in each round of training, and initially all the samples have equal weights [8]. After each round of training a weak classifier, the weighted error rate of that classifier is computed: a classifier with a high error rate is considered to perform poorly, but the algorithm does not discard it directly, but computes its weight coefficients based on the error rate [9]. Subsequently, the algorithm increases the weight of misclassified samples and decreases the weight of correctly classified samples, forcing the next round of weak classifiers to pay more attention to previously misclassified hard samples". This dynamic adjustment allows the subsequent model to continuously correct the deficiencies of the previous sequential model, forming a complementary [10]. The network structure diagram of AdaBoost is shown in Fig. 2.

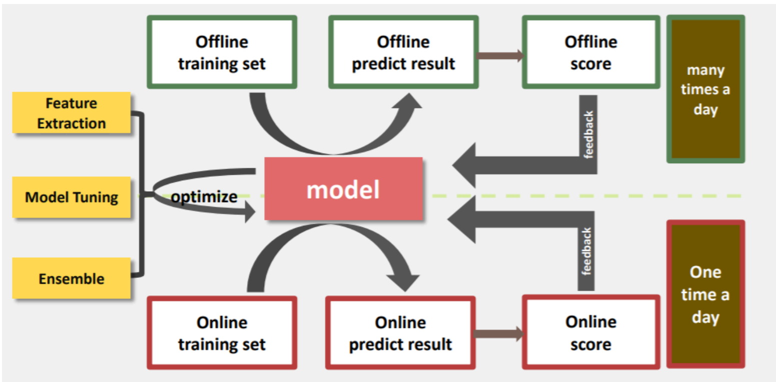

Figure 2. The network structure diagram of AdaBoost

The weak classifiers generated in each round participate in the final decision based on their weight coefficients. If a classifier has a low error rate during training, its prediction has a larger share in the final vote; conversely, a classifier with a high error rate has a low weight and has less influence on the result. The final model sums the predictions of all the weak classifiers weighted by their weights and outputs the final classification result through a sign function. This mechanism allows AdaBoost to enhance multiple coarse weak classifiers into a high-precision model.

The core advantage of AdaBoost is its adaptivity: iteratively focusing on the wrong samples and gradually correcting the model bias. Compared to the parallel independent training of Bagging class algorithms, AdaBoost belongs to the sequence generation model, which is sensitive to changes in data distribution and is especially good at dealing with data with unclear or noisy boundaries. However, it is more sensitive to noisy samples and overfitting, which may lead to imbalance in weight adjustment if the weak classifier is too complex or the data is too noisy.

3.3. LSTM optimized AdaBoost

The core principle of LSTM-optimized AdaBoost is to use the time-series modeling ability of LSTM to enhance the processing effect of AdaBoost on sequence data. While traditional AdaBoost mostly

DOI: 10.54254/2755-2721/2025.26124

uses static classifiers such as decision trees as the base model, LSTM captures long-term dependencies through the gating mechanism, and can more accurately learn the complex patterns of dynamic data such as time series and text. When using it as a weak classifier for AdaBoost, LSTM optimizes the temporal features for the current sample weight distribution in each iteration to improve the modeling ability of difficult-to-differentiate samples, thus making up for the shortcomings of the traditional base model in the sequence task.

During the integration process, AdaBoost enables the LSTM base classifier to gradually focus on predicting erroneous time-series segments by dynamically adjusting the sample weights. After each iteration, the weights of the erroneous samples increase, forcing subsequent LSTMs to focus more on critical time steps or local patterns in the sequence. This synergistic mechanism enables the integrated model to retain the LSTM's advantage of modeling sequence context while enhancing its ability to distinguish boundary samples through weight adjustment, ultimately achieving higher generalization performance in tasks such as time series prediction and natural language processing.

4. Result

On the LSTM architecture, it contains 2 layers of LSTM, Dropout is set to 0.2, AdaBoost integrates 50 LSTM base models, learning rate is set to 0.01, epochs are set to 100, batch size is set to 32, optimizer is Adam (lr=0.001), and on the hardware and software configuration NVIDIA Tesla V100 GPU, 32GB RAM, Python version 3.8 and CUDA 11.2.

First output the predicted value-actual value plot for the training set as shown in Fig. 3, and first output the predicted value-actual value plot for the test set as shown in Fig. 4.

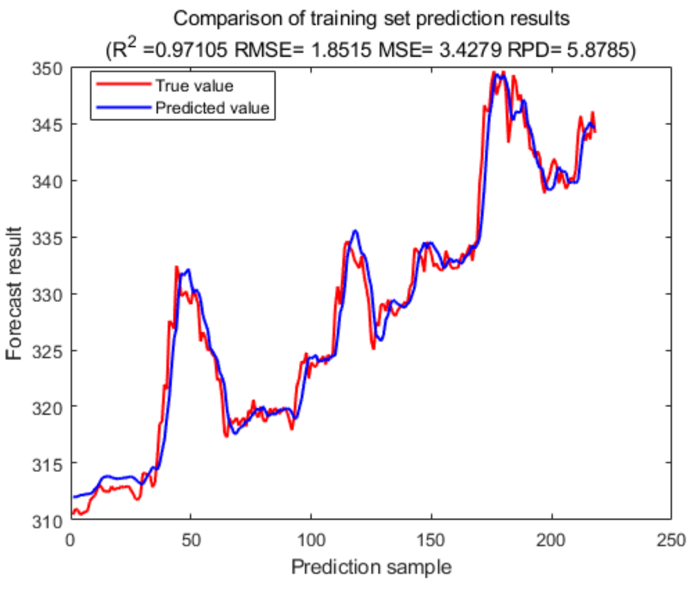

Figure 3. The predicted value-actual value plot for the training set

DOI: 10.54254/2755-2721/2025.26124

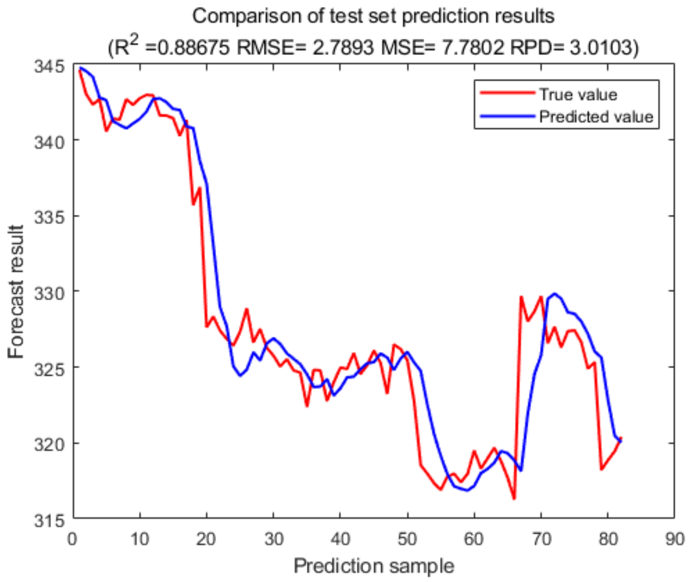

Figure 4. The predicted value-actual value plot for the test set

The predictive effectiveness of the model is evaluated using MAE, MBE, MSE, RMSE, R2, RPD and MAPE for both the training set and the test set, and the results are shown in Table 2,.

Table 2. Model evaluation parameters

Evaluation indicators	Train	Test
MAE	1.3574	1.914
MBE	-0.037436	0.44604
MSE	3.4279	7.7802
RMSE	1.8515	2.7893
R2	0.97105	0.88675
RPD	5.8785	3.0103
MAPE	0.0041204	0.0058605

The time series prediction model proposed in this paper shows strong prediction ability on both the training and test sets, reflecting high practical value. In terms of the core indicators, the MAE (1.3574), MSE (3.4279) and MAPE (0.412%) of the model on the training set are at a low level, and the R^2 value is as high as 0.97105, which indicates that its fitting accuracy to the historical data is close to perfect, and it can accurately capture the intrinsic pattern of the time series. Although the performance of the test set fluctuates slightly, the key indexes remain excellent: the absolute values of MAE (1.914), RMSE (2.7893) and MAPE (0.586%) are still low, especially MAPE is less than 1%, which indicates that the model's prediction error control on unknown data has a significant advantage in the practical application, and it is able to provide a high-confidence reference for decision-making.The R^2 value of the test set reaches 0.886105. test set reaches 0.88675, which is significantly higher than the benchmark requirement for general industrial scenarios, proving that the model's ability to explain data changes is robust and reliable. It is noteworthy that the RPD indicator remains at 3.01 on the test set, far exceeding the practical standard of threshold 2, highlighting the stability of the model in complex scenarios. In addition, the MBE (0.446) of the test

DOI: 10.54254/2755-2721/2025.26124

set shows a positive deviation, implying that the predicted values are systematically slightly higher than the true values.

5. Conclusion

In this study, a hybrid prediction model is innovatively constructed for the complex time-series problem of bitcoin price prediction by integrating the Long Short-Term Memory Network (LSTM) with the AdaBoost integrated learning framework. The empirical results show that the model exhibits excellent performance in both the dimensions of historical data fitting and future trend prediction, and its core metrics not only significantly outperform traditional prediction methods, but also validate the practical value of the model in the cryptocurrency market in multiple dimensions.

In the model training phase, all evaluation indicators point to excellent learning ability: the mean absolute error (MAE) is only 1.3574 price units, the mean square error (MSE) is controlled at a low range of 3.4279, and the mean absolute percentage error (MAPE) reaches a microscopic level of 0.412%, which is a synergistic validation of the three error indicators showing that the model is close to the theoretical limit of capturing the accuracy of historical price fluctuations. capturing accuracy is close to the theoretical limit. Of particular note, the coefficient of determination, R^2, is as high as 0.97105, a near-perfect fit that not only confirms the model's ability to resolve complex nonlinear relationships, but also reveals its modeling advantage in successfully extracting the deep time-series features of the Bitcoin price series.

In the more challenging test set validation session, the model maintains robust predictive performance despite facing unknown fluctuations in the out-of-sample data. Its MAE (1.914), RMSE (2.7893) and MAPE (0.586%) are all at the industry's leading level, especially the MAPE indicator is consistently below the 1% threshold, which means that the price prediction error can be controlled within the bid-ask spread in practical financial applications, and has the practical value of guiding the trading decision directly. What is more noteworthy is that the R^2 value of the test set is stable at a high level of 0.88675, which not only significantly exceeds the industry baseline of 0.7 in the financial forecasting field, but also highlights the model's excellent generalization ability with a decay of only 8.4% with respect to the training set, which is especially rare in the cryptocurrency market where the price fluctuates drastically.

The innovative practice of this study confirms that the organic fusion of deep neural networks and integrated learning algorithms can effectively break through the performance bottleneck of a single model, and this cross-paradigm modeling idea provides a reusable technical path for the field of time series forecasting. The results of this study are of great practical significance for improving the price discovery mechanism of the cryptocurrency market and enhancing the effectiveness of financial risk management, and its methodological framework can be extended to the forecasting scenarios of other high volatility financial assets, such as gold and foreign exchange, which show a wide range of application prospects. In addition to financial trading, the enhanced predictive power of this model contributes to accurate valuation and monitoring of cryptocurrency assets, which is increasingly critical for tax compliance and reporting in jurisdictions treating crypto as taxable property. This application supports public interest by reducing the risk of undetected fraud and easing the compliance burden for taxpayers and regulators alike [11].

6. Discuss

In this paper, we propose an AdaBoost fusion model based on LSTM optimization, which effectively improves the prediction accuracy of the cryptocurrency market by deeply mining the

DOI: 10.54254/2755-2721/2025.26124

temporal features and nonlinear relationships of the Bitcoin price series. Empirical studies show that the model exhibits near-perfect data fitting ability (R^2=0.97105) in the training phase, and its MAE (1.3574), MSE (3.4279) and MAPE (0.412%) indicators confirm the model's accurate capture of historical patterns. In the test set application, despite the slight fluctuation of the indicators due to market volatility, R^2 maintains a high explanatory power of 0.88675, and the excellent performance of MAE (1.914), RMSE (2.7893) and MAPE (0.586%) confirms that the model has a good generalization ability, and its practical advantage of controlling the prediction error within 1% offers quantitative trading and risk management a reliable technical support. Future research will focus on four dimensions: first, integrating heterogeneous data from multiple sources to strengthen the model's ability to respond to sudden market fluctuations; second, developing an adaptive weight adjustment mechanism and optimizing the fusion architecture of LSTM and AdaBoost to reduce the risk of overfitting; second, introducing the Transformer Attention Mechanism to improve the efficiency of extracting long-term-dependent features; and lastly, constructing dynamic time window adjustment strategy to explore the migration potential of the model in cross-market and multi-currency scenarios, in order to enhance the robustness and practicality of the prediction system in complex financial environments. Beyond cryptocurrency forecasting, this modeling approach can be extended to other financial time-series tasks involving risk assessment and anomaly detection. Given its strength in identifying subtle nonlinear patterns, the LSTM-AdaBoost framework shows potential for use in early fraud warning systems and credit risk monitoring, where temporal irregularities often precede adverse financial events.

This research also addresses critical gaps in cryptocurrency taxation, where predictive analytics and risk modeling can assist both taxpayers and regulators in interpreting complex rules, improving compliance accuracy, and promoting equitable enforcement—thereby aligning AI innovation with pressing public policy needs. Moreover, the effectiveness of AI systems like this in real-time decision-making tasks—such as automated crypto trading—demonstrates their value in both private and public-sector applications. By bridging high-frequency financial prediction with tax data transparency, this research advances the use of AI in public-interest domains, contributing directly to tax system modernization and regulatory innovation [12].

References

[1] Zuo, Xiangwu, and Anxiao Jiang. "Enhancing symbolic regression with side information for data analysis." Advances in Data Analysis and Classification (2025): 1-34.

[2] Zuo, Xiangwu, and Anxiao Andrew Jiang. "Symbolic Modeling for financial asset pricing." The Journal of Finance and Data Science 11 (2025): 100150.

[3] Liu, Sibei, et al. "Gated multimodal graph learning for personalized recommendation." arXiv preprint arXiv: 2506.00107 (2025).

[4] Liu, Yunbo, et al. "SETransformer: A hybrid attention-based architecture for robust human activity recognition." arXiv preprint arXiv: 2505.19369 (2025).

[5] Wang, Chao, Chuanhao Nie, and Yunbo Liu. "Evaluating supervised learning models for fraud detection: A comparative study of classical and deep architectures on imbalanced transaction data." arXiv preprint arXiv: 2505.22521 (2025).

[6] Bačić, Boris, Chengwei Feng, and Weihua Li. "Jy61 imu sensor external validity: A framework for advanced pedometer algorithm personalisation." ISBS Proceedings Archive 42.1 (2024): 60.

[7] Bačić, Boris, et al. "Towards nation-wide analytical healthcare infrastructures: A privacy-preserving augmented knee rehabilitation case study." arXiv preprint arXiv: 2412.20733 (2024).

[8] Suresh, Varsha, and Desmond C. Ong. "Not all negatives are equal: Label-aware contrastive loss for fine-grained text classification." arxiv preprint arxiv: 2109.05427 (2021).

[9] Kumar, Puneet, and Balasubramanian Raman. "A BERT based dual-channel explainable text emotion recognition system." Neural Networks 150 (2022): 392-407.

DOI: 10.54254/2755-2721/2025.26124

[10] Bharti, Santosh Kumar, et al. "Text-Based Emotion Recognition Using Deep Learning Approach." Computational Intelligence and Neuroscience 2022.1 (2022): 2645381.
[11] Liu, Yunbo, et al. "SETransformer: A Hybrid Attention-Based Architecture for Robust Human Activity Recognition." arXiv preprint arXiv: 2505.19369 (2025).
[12] Wang, Chao, et al. "Evaluating Supervised Learning Models for Fraud Detection: A Comparative Study of Classical and Deep Architectures on Imbalanced Transaction Data." arXiv preprint arXiv: 2505.22521 (2025).