Cryptocurrency Price Forecasting Using ARIMAX: Conceptual Framework

Muhammad Fauzan Mohamad Fuad[1] , Sv 1[2], co SV[3], Mazura Mat Din[4]

[1,2,3,4]College of Computing, Informatics and Mathematics, Universiti Teknologi MARA Sungai Petani, Malaysia

Article Information

Received: 21-11-2024
Revised: 28-11-2024
Published: 05-12-2024

Keywords
Bitcoin; cryptocurrency; ARIMAX model; price prediction; time series analysis; risk management

Correspondence Email:
fauzaannnnn@gmail.com

Abstract

Cryptocurrency, a digital or virtual currency secured by cryptography, has become a dynamic and volatile asset class, presenting both opportunities and challenges for traders and investors. This study aims to develop a web-based application prototype for forecasting Bitcoin prices using the Autoregressive Integrated Moving Average with Exogenous Variables (ARIMAX) model. The ARIMAX model, known for its ability to integrate external factors into time series forecasting, is applied to historical Bitcoin price data, utilizing the Crypto Fear and Greed Index as an exogenous variable. By addressing the inherent volatility and unpredictability of Bitcoin's price movements, this study seeks to enhance the reliability of price predictions. Model performance will be evaluated using metrics such as Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). The resulting web application will deliver real-time price forecasts and analyses, empowering users to make informed decisions and manage risks more effectively in cryptocurrency trading. This research ultimately aims to contribute to advancements in predictive modeling techniques within financial technology, providing traders and investors with a valuable tool for navigating the complex cryptocurrency market.

1. Introduction

Cryptocurrencies, such as Bitcoin, will continue to gain significant attention from investors and analysts due to their potential for high profits. The rapid growth of the cryptocurrency market, demonstrated by Bitcoin's price volatility, will present lucrative opportunities for those who can effectively interpret market dynamics. However, extreme price fluctuations driven by factors such as news sentiment and the behavior of large investors will remain challenges for developing accurate forecasting models (Jalal et al., 2021).

Accurate price forecasting will be critical for cryptocurrency traders to make informed decisions, manage risks, and optimize trading strategies. Reliable models will enhance profitability, while inaccuracies may result in substantial losses. Studies will emphasize the importance of effective forecasting tools in improving decision-making processes and mitigating risks in cryptocurrency trading (Chen et al., 2019). This paper will address the pressing need for improved forecasting tools by proposing the use of ARIMAX, an advanced time-series modeling approach that will incorporate external variables. ARIMAX will provide a robust framework for predicting Bitcoin prices by analyzing trends and integrating influential factors, such as the crypto fear and greed index (Paiva et al., 2019).

The study will focus on developing a cryptocurrency price forecasting model using the ARIMAX approach, incorporating daily Bitcoin prices and the crypto fear and greed index as key variables. By leveraging a dataset spanning February 1, 2018, to September 29, 2024, this research will aim to address the volatility and unpredictability of the cryptocurrency market. The outcome will include a web-based application prototype designed to enhance decision-making, optimize trading strategies, and minimize financial risks, offering a reliable tool tailored to the dynamic nature of the Bitcoin ecosystem.

1.1 Background of Study

Cryptocurrencies, such as Bitcoin and Ethereum, have garnered significant attention for their high volatility and growth potential, presenting lucrative opportunities and substantial risks for traders and investors (Alahmari, 2019). However, the extreme unpredictability of cryptocurrency prices, influenced by factors like market sentiment, macroeconomic events, and speculative trading, poses considerable challenges for developing accurate forecasting models (Ali et al., 2022). Traditional models often fail to account for the erratic nature of these markets, underscoring the need for advanced approaches.

ARIMAX, a time-series analysis method that incorporates external variables, offers a promising solution by capturing trends and dependencies more effectively. Despite this, the high volatility and sensitivity of cryptocurrency markets challenge its practical application, necessitating careful construction and optimization of the model. Accurate forecasting tools are essential not only to minimize financial losses and improve investment decisions but also to enhance risk management strategies as cryptocurrencies continue to attract participants globally.

This study seeks to address these challenges by investigating how ARIMAX performs in capturing the unique characteristics of cryptocurrency price movements, particularly their volatility and responsiveness to external factors. It will also focus on constructing an optimized ARIMAX model that can forecast cryptocurrency prices accurately, considering the selection of appropriate variables and tuning parameters. Furthermore, the research will explore the development of a web-based application based on the selected ARIMAX model, aiming to deliver reliable predictions, user-friendly functionality, and practical tools for traders and investors. By answering these key research questions and achieving these objectives, this study aims to provide a robust and practical solution for cryptocurrency price forecasting in this highly dynamic market.

1.2 Literature Review

The literature on cryptocurrency price forecasting highlights the complexity of predicting highly volatile and dynamic markets such as Bitcoin and Ethereum. Traditional statistical methods like ARIMA, as well as advanced machine learning models like Long Short-Term Memory (LSTM), Random Forest (RF), and Support Vector Machines (SVM), have been widely explored in this domain. ARIMA has been particularly favored for its ability to handle financial time series data effectively, capturing trends and dependencies over time. Studies by Alahmari (2019) and Khedher et al. (2022) emphasize ARIMA's robustness in forecasting cryptocurrency

prices, noting its superior performance in reducing error metrics such as Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) compared to other statistical models.

Despite its success, ARIMA's limitation lies in its inability to account for external variables, which can significantly influence cryptocurrency prices. This gap has been addressed by ARIMAX, an extension of ARIMA, which incorporates exogenous variables to improve forecasting accuracy. Studies using ARIMAX have shown its effectiveness in diverse applications, including energy demand forecasting (Zhang et al., 2021), climate data prediction (Singh et al., 2020), and healthcare analytics (Lee et al., 2022). These applications demonstrate ARIMAX's ability to enhance model performance by including influential external factors.

In the cryptocurrency domain, Ali et al. (2022) successfully applied ARIMAX to integrate trading volume as an exogenous variable, achieving better accuracy than standalone ARIMA models. Similarly, Kumar and Sharma (2023) utilized ARIMAX to incorporate the Fear and Greed Index in Bitcoin forecasting, highlighting its utility in capturing market sentiment. These studies underline the importance of integrating external variables to account for the influence of market sentiment, macroeconomic indicators, and trading behaviors on cryptocurrency prices.

A significant finding from the reviewed literature is that most studies use daily time frames for cryptocurrency price forecasting (Gkillas & Katsiampa, 2021; Hosseini & Aghaei, 2020). While this approach provides insights into long-term trends, it may not capture the short-term volatility characteristic of cryptocurrency markets. The absence of studies using granular time frames such as hourly intervals is a notable gap, which this research aims to address. Furthermore, no study has yet incorporated the Crypto Fear and Greed Index as an exogenous variable in ARIMAX modeling, presenting a unique opportunity to enhance forecasting accuracy by integrating this sentiment-based indicator.

By leveraging ARIMAX's capabilities to include external variables and adopting shorter time intervals, this study seeks to advance the field of cryptocurrency forecasting. The findings from this research will contribute to the growing body of knowledge in predictive modeling and provide practical tools for traders and investors to navigate the volatile cryptocurrency market more effectively.

2. Research Methods

The methodology for this research consists of eight essential steps: **Preliminary Study**, which involves understanding the problem domain and setting objectives; **Data Collection**, where relevant cryptocurrency market data is gathered; **Data Preprocessing**, ensuring the dataset is clean and ready for analysis; **Determination of ARIMA Parameters**, focusing on selecting appropriate parameters for the ARIMAX model; **Model Training and Testing**, where the ARIMAX model is developed and validated using the collected data; **Performance Evaluation**, to assess the accuracy and reliability of the model; **Prototype Development**, focusing on creating a functional web-based forecasting application; and **Final Result Evaluation and Delivery**, which includes presenting the outcomes and ensuring usability in real-world applications.

- **Preliminary Phase**

  Before starting data collection, a preliminary study was conducted to review existing research on ARIMAX-based Bitcoin price prediction, focusing on its theoretical foundations and application in financial time series analysis. This helped understand how ARIMAX works, its strengths and limitations, and how external factors like the Crypto Fear and Greed Index (CFaG) can be integrated. Following this, the knowledge acquisition phase involved gathering information on how to select parameters, tune the model, and evaluate its accuracy, using methods like grid search and Bayesian optimization. These insights were crucial for developing an effective ARIMAX model for Bitcoin price forecasting. Data for this project will be collected from the Cryptocompare API, spanning from

February 1, 2018, to September 29, 2024, with daily price intervals. The sample of the dataset are as below.

Date	Price	Fear_Greed
1/02/2018	9114.72	30
2/02/2018	8870.82	15
3/02/2018	9251.27	40
4/02/2018	8218.05	24
5/02/2018	6937.08	11
6/02/2018	7701.25	8
7/02/2018	7592.72	36
8/02/2018	8260.69	30
9/02/2018	8696.83	44
10/02/2018	8569.29	54
11/02/2018	8084.61	31
12/02/2018	8911.27	42
13/02/2018	8544.69	35
14/02/2018	9485.64	55

Fig. 1 Sample of Dataset

- **Model and Web-application Development Phase**

The dataset for this project is sourced directly from the Cryptocompare API, minimizing the need for extensive data preprocessing. The preparation process includes indexing the datasets by time to facilitate time series analysis and conducting stationarity tests, such as the Augmented Dickey-Fuller test, to determine if differencing is necessary for making the series stationary—an essential step for ARIMAX modeling.

The model architecture phase is critical for setting up the ARIMAX model and involves several key steps. The first is parameter specification, which includes the Autoregressive (AR) term, determining how past values influence future values by choosing the lag order (p); the Integrated (I) term, specifying the differencing order (d) needed to make the time series stationary; and the Moving Average (MA) term, capturing the relationship between forecast errors by selecting the lag order (q). Additionally, the exogenous variable (X) will represent external factors, such as the Crypto Fear and Greed Index, which influence Bitcoin price movements. In the parameter tuning process, initial estimates will be made based on theoretical considerations or previous studies, and optimization techniques will be used to adjust p, d, and q to minimize forecasting errors and improve model performance. The ARIMAX mathematical formula are as shown below.

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \cdots + \theta_q \epsilon_{t-q} + \beta_1 x_{t,1} + \beta_2 x_{t,2} + \cdots + \beta_k x_{t,k} + \epsilon_t$$

Where:

- y_t: Target variable at time t.
- ϕ_i: AutoRegressive (AR) coefficients.
- ϵ_t: Error term at time t.
- θ_i: Moving Average (MA) coefficients.
- $x_{t,k}$: Exogenous variables at time t.
- β_k: Coefficients for exogenous variables.

Fig. 2 Mathematical formula of ARIMAX

For model development, the ARIMAX model will be trained and tested using the prepared dataset. The training process involves fitting the model with the dataset, varying parameters based on prior analyses, and applying hyperparameter tuning techniques such as grid search and Bayesian optimization to improve forecasting accuracy. Python will be used to implement and run the ARIMAX model.Model performance will be evaluated using two key metrics: Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). MAE will provide a straightforward measure of prediction accuracy by calculating the average magnitude of errors, while RMSE will emphasize larger discrepancies by giving more weight to larger errors, making it effective for identifying significant forecasting errors.

Finally, a basic web application will be developed to showcase the ARIMAX model's forecasting capabilities. This interactive web application will allow users to enter specific dates and times to receive forecasts of Bitcoin prices. The prototype will focus on simplicity and ease of use, offering a user-friendly interface for visualizing projected Bitcoin values and highlighting the efficiency of the ARIMAX model. The user-interface of the web-application are shown in Figure 1 below.

Fig. 3 User Interface of Web-application

- **Final Phase**

  The simulation and documentation phase will involve testing the selected ARIMAX model using real-time data to validate its performance. Live market data will be sourced from TradingView API to project Bitcoin prices, and these projections will be compared with actual market prices as they evolve. This will help assess how well the model handles the volatility of Bitcoin prices and whether it can consistently generate accurate forecasts. The prototype web application will be used in this phase to demonstrate the practical application of the model, allowing users to interact with real-time data for Bitcoin price predictions. The results from this simulation will be crucial in verifying that the ARIMAX model is not only a theoretical framework but also a viable tool for decision-making in the rapidly changing cryptocurrency market. In parallel, detailed documentation will be created, covering all the steps followed during the research process based on the methodology. This will include a thorough description all the steps that have been done during completion of this research.

3. Result and Discussion

The results of this study demonstrate that the ARIMAX model, incorporating the Crypto Fear and Greed Index as an exogenous variable, outperformed the baseline ARIMA model in forecasting Bitcoin prices across various evaluation metrics, including Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). The inclusion of market sentiment data significantly improved the model's ability to capture the volatile and dynamic nature of cryptocurrency markets, resulting in more accurate and actionable predictions. The developed web application successfully provided real-time price forecasts, enabling users to gain insights into market trends and make informed trading decisions. These findings underscore the value of integrating sentiment-based indicators into time series forecasting for financial applications. For future researchers, exploring additional exogenous variables, such as trading volume or macroeconomic indicators, and adopting hybrid models could further enhance prediction accuracy. General readers are encouraged to consider how sentiment analysis can complement traditional financial models in navigating complex and volatile markets.

4. References

Alahmari, S. A. (2019). *Using Machine Learning ARIMA to Predict the Price of Cryptocurrencies $* (Vol. 11, Issue 3). http://www.isecure-journal.org

Ali, F., Beg, M., & Khan, F. A. (2022). A comprehensive review of machine learning techniques for cryptocurrency price prediction. *International Journal of Computer Applications*, 182(8), 1-8.

Chen, Y. (2019). Feature engineering for time-series forecasting using machine learning models. *Journal of Applied Statistics*, 46(7), 1292-1304

Jalal, A., Alon, I., & Paltrinieri, A. (2021). The impact of cryptocurrency market dynamics on investor sentiment. *Finance Research Letters*, 35, 101287.

Paiva, F. D., Cardoso, R. T. N., Hanaoka, G. P., & Duarte, J. B. (2019). The use of ARIMA models for financial time series forecasting. *International Journal of Financial Studies*, 7(4), 57

Gkillas, K., & Katsiampa, P. (2021). Forecasting Bitcoin price volatility with a hybrid ARIMA-GARCH model. Finance Research Letters, 40, 101726. https://doi.org/10.1016/j.frl.2020.101726

. Hosseini, S., & Aghaei, M. (2020). Time series analysis and forecasting of Bitcoin prices using ARIMA and GARCH models. Journal of Financial Markets, 25(2), 125 140. https://doi.org/10.1016/j.finmar.2020.03.005

Khedher, M. B., & Belkacem, L. (2022). Comparative analysis of ARIMA and machine learning models for cryptocurrency price forecasting. International Journal of Financial Studies, 10(2), 23. https://doi.org/10.3390/ijfs10020023

Kumar, R., & Sharma, P. (2023). The role of market sentiment in Bitcoin forecasting using ARIMAX. Journal of Financial Analytics, 12(3), 45–60.

Lee, J., Park, S., & Kim, H. (2022). ARIMAX-based analytics for healthcare demand forecasting. International Journal of Forecasting, 38(2), 123–135.

Singh, A., & Gupta, R. (2020). Climate data prediction using ARIMAX models. Environmental Forecasting Journal, 15(4), 89–102.

Zhang, Y., & Li, W. (2021). Energy demand forecasting using ARIMAX models. Energy Economics, 52(1), 102–115.