Cryptocurrency Price Prediction with Volatility Models

Jeffrey Hu
jeffhu@stanford.edu

Megan Dass
mdass9@stanford.edu

1 Introduction

The goal is to develop and evaluate volatility models such as ARCH/GARCH for the prediction of cryptocurrency volatility. This is useful because determining when high or low volatility is coming can be used to determine when additional trading opportunities will arise, or whether or not a given position will be possible to exit. Additionally, even being able to properly forecast volatility can be useful to price options, and there can be an inherent alpha in being able to forecast volatility [1].

2 Problem Statement

This study aims to analyze and forecast the volatility of Bitcoin prices, focusing on historical daily data, which includes the closing price, market cap, and total volumes of Bitcoin. The primary input variable for the analysis is 'log_returns', calculated from the daily prices using the formula $r_t = \ln \frac{S_t}{S_{t-1}}$, where S_t is the closing price at time t [2]. The key output of the study is the forecasted variance values, representing the expected future volatility of Bitcoin prices.

3 Technical Approach

3.1 Dataset

To curate the dataset for this project, we used the CoinGecko API to get the Bitcoin market prices. Specifically, we obtained daily values of the closing price and market cap of Bitcoin from April 28, 2013, to October 21, 2023, resulting in 3828 data points. Because of Bitcoin's popularity, this is a good digital currency to start with, and the same methodology can be extended to other currencies to ensure that the same modeling techniques can generalize for other, less popular digital currencies. Before fitting the baseline models, we computed the return rate of the of the closing prices for each day, starting from t = 1. The following formula was used to compute the return rate:

$$r_t = \ln \frac{S_t}{S_{t-1}}$$

where S_t represents the closing price at timestep, t. This expression is the log return rate, which is standard convention for computing returns in quantitative finance [3].

3.2 ARIMA fitting

The ARIMA model excels in volatility modeling due to its integrated (I) component, which effectively handles non-stationarity, a key feature in cryptocurrency volatility. Its AR and MA components capture short-term dependencies and market shocks, crucial for forecasting volatility in financial time series like Bitcoin, making it an ideal baseline for comparative analysis [4].

Preprint. Under review.

Table 1: AIC & BIC values for ARIMA and GARCH models

Model	AIC	BIC
ARIMA	-13796.5	-13777.8
GARCH	-14653.9	-14635.1

To use the ARIMA model to forecast the Bitcoin volatility, we first compute the PACF and ACF graphs to obtain appropriate values for p and q. From these graphs, we deduced that that we should fit the ARIMA models on values (p, d, q) = (1, 1, 1).

After fitting the dataset to the proper ARIMA model, we plotted the residuals (Figure 1) to ensure that they are white noise, and we computed the AIC and BIC values (Table 1) to ensure that the ARIMA model fit the data well.

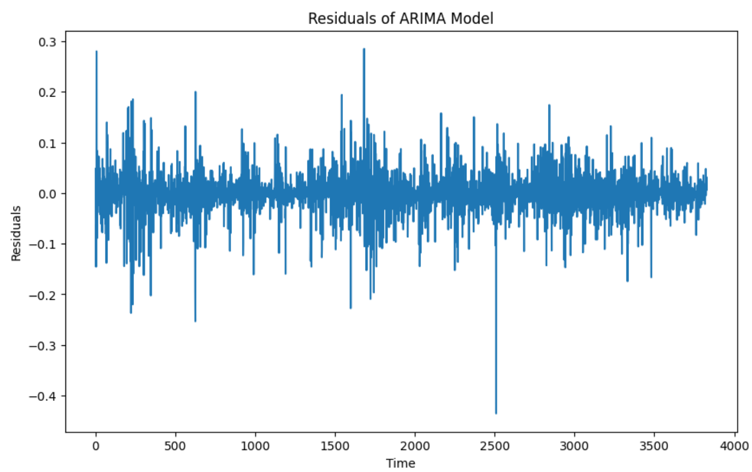

Figure 1: ARIMA residuals graph

Although the residuals are white noise and the AIC and BIC values are low, indicating that the ARIMA model fit the data well, when we plot the fitted return values versus the real return values (Figure 2), we see that there is room for improvement.

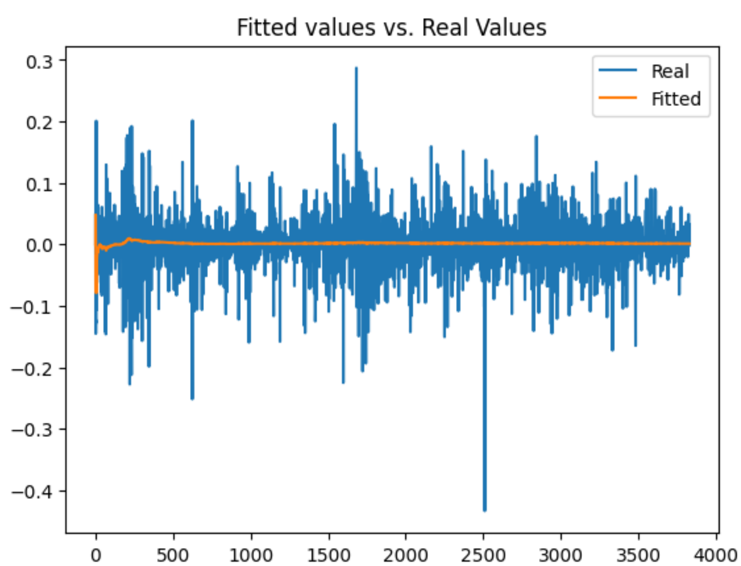

Figure 2: ARIMA residuals graph

3.3 GARCH fitting

The Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model plays a pivotal role in modeling and forecasting time series volatility. GARCH, an extension of the ARCH model, is particularly well-suited for financial data that exhibit time-varying volatility, a common characteristic of cryptocurrency markets [5]. The GARCH(1,1) model, a popular choice for financial applications, can be formulated as follows:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

where σ_t^2 is the conditional variance (volatility) at time t, ϵ_{t-1}^2 is the squared residual from a mean model (like ARIMA) at time $t-1$, and α_0, α_1, and β_1 are parameters to be estimated. In our project, we apply the GARCH(1,1) model to the return rates of Bitcoin, aiming to capture the dynamics of its volatility more accurately than standard ARIMA models. The estimation of the GARCH model parameters will provide insights into the persistence and mean-reverting behavior of Bitcoin's volatility, crucial for making informed trading decisions.

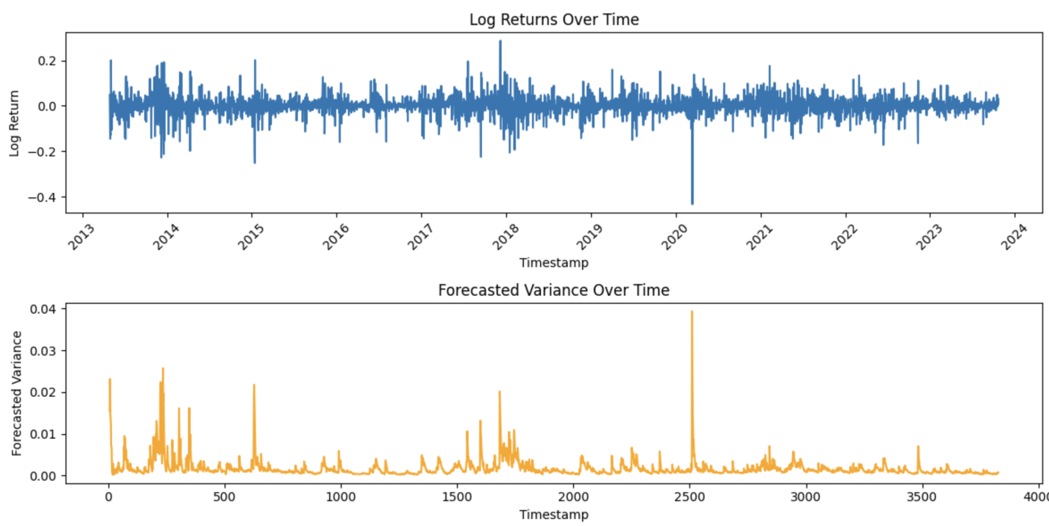

Figure 3: GARCH results

4 Preliminary Results

From the results we have so far, it seems like GARCH is far more powerful in expressing the forecasted variance. From the chart, we can see that GARCH expresses dynamic variance corresponding to autocorrelation of volatility, while the ARIMA model appears to forecast a constant volatility, which is not the reality for digital currencies. Additionally, the AIC and BIC scores for GARCH are lower.

5 Next Steps

While the GARCH model on the return values performed better than the fitted ARIMA model, the quality of the fitted model can be improved. For the next phase of the project, we aim to improve the fitted model quality by exploring a hybrid GARCH + LSTM model approach, which can be done by first fitting a GARCH model and then including the parameters as features when training the LSTM model [6]. Because LSTMs are well-suited for long-term time series forecasting, our prediction is that the GARCH model would be able to predict both the short-term and long-term volatility well [7, 8, 9]. Our final step will be to evaluate the performance of fitting the baseline and hybrid models on two other digital currencies, TRON and Ethereum to understand whether the performance of these models can be generalized beyond Bitcoin. Further, will stress test these models to mimic real-world scenarios, such as market crashes and market manipulation that may cause noise in the dataset.

References

[1] Haizhou Guo, Dian Zhang, Siyuan Liu, Lei Wang, and Ye Ding. Bitcoin price forecasting: A perspective of underlying blockchain transactions. *Decision Support Systems*, 151:113650, 2021.

[2] Lukáš Pichl and Taisei Kaizoji. Volatility analysis of bitcoin. *Quantitative Finance and Economics*, 1(4):474–485, 2017.

[3] Panna Miskolczi. Note on simple and logarithmic return. *Applied Studies in Agribusiness and Commerce*, 11(1-2):127–136, 2017.

[4] Nashirah Abu Bakar and Sofian Rosbi. Autoregressive integrated moving average (arima) model for forecasting cryptocurrency exchange rate in high volatility environment: A new insight of bitcoin transaction. *International Journal of Advanced Engineering Research and Science*, 4(11):130–137, 2017.

[5] Mohammed Mudassir, Shada Bennbaia, Devrim Unal, and Mohammad Hammoudeh. Time-series forecasting of bitcoin prices using high-dimensional features: a machine learning approach. *Neural computing and applications*, pages 1–15, 2020.

[6] Ha Young Kim and Chang Hyun Won. Forecasting the volatility of stock price index: A hybrid model integrating lstm with multiple garch-type models. *Expert Systems with Applications*, 103:25–37, 2018.

[7] Andrés García-Medina and Ester Aguayo-Moreno. Lstm–garch hybrid model for the prediction of volatility in cryptocurrency portfolios. *Computational Economics*, pages 1–32, 2023.

[8] Chih-Hung Wu, Chih-Chiang Lu, Yu-Feng Ma, and Ruei-Shan Lu. A new forecasting framework for bitcoin price with lstm. In *2018 IEEE International Conference on Data Mining Workshops (ICDMW)*, pages 168–175. IEEE, 2018.

[9] B Aditya Pai, Lavanya Devareddy, Supriya Hegde, and BS Ramya. A time series cryptocurrency price prediction using lstm. In *Emerging Research in Computing, Information, Communication and Applications: ERCICA 2020, Volume 2*, pages 653–662. Springer, 2022.