Forecasting of Price Returns of Meme Coins

Tejus Prabhu
Department of Statistics and Data Science
Christ (Deemed to be University), Bangalore, Karnataka, India

Dr. Nagaraja M.S.
Department of Statistics and Data Science
Christ (Deemed to be University), Bangalore, Karnataka, India

Abstract: *Due to people's increased access to the internet and social media, investing in meme coins has grown in popularity. Using a variety of time series models like ARMA, ARIMA, GARCH, ARMA-GARCH, etc., we address forecasting of a meme coin called Pepemon Pepeballs and understanding the behavioral patterns in order to choose the best model for the aforementioned. The McLeod-Li Test was employed to determine whether ARCH/GARCH effect was present in the mean model. According to the results, the ARMA (4,2)-GARCH (1,1) model was the most appropriate one for Pepemon Pepeballs.*

Keywords: *Meme Coins, Pepemon Pepeballs, Value at Risk*

Introduction

The prevalence of numerous online memes on social media platforms like Twitter, Reddit, Instagram, etc. has made a particular sort of cryptocurrency known as meme coins. In 2013, Dogecoin became the first meme coin to be released. Elon Musk's tweets in 2020 helped it gain popularity, and since then, several meme coins have been developed, including Shiba Inu, Mona Coin, Pepemon Pepeballs, and others.[16] Compared to other cryptocurrencies, meme coins exhibit more volatile behavior since they are perceived as riskier and any shift in attitude may greatly affect their rise or fall. The price of these cryptocurrencies is also influenced by government regulations, and significant actions taken by any big economy can have a significant impact on the meme coin's standing. Meme coins became increasingly prevalent over time, and the best ones were also featured on several trading websites. We therefore use time series analysis and models to represent the characteristics of the coins according to various timelines in a mathematical form, and we use it to predict the future prices of the coin to give people an idea of how to invest carefully and take more precautions to reduce their net loss. To choose the most appropriate time series model, we must use a variety of time series model, including AR, MA, ARMA, ARIMA, ARCH, GARCH, and others. The extreme volatility of meme coins means that the sole usage of the mean models would not be sufficient. Theincorporate the usage of variance models were implemented in 1982 by Robert Engle who used ARCH and the generalized version was implemented in 1986 by Tim Bollerslev.

Literature Review

Various articles showcased the time series analysis of multiple cryptocurrencies which are commonly used and are listed in multiple stock markets like Bitcoin, Ethereum, Lite Coin, Doge Coin, etc. But there is little research about lesser-known and unlisted cryptocurrencies and meme coins. Various studies have also been made on the prices of tradable commodities like oil, rubber, stocks etc.

The research has been made to add more cryptocurrencies into the limelight and create more awareness among people. Various time series and machine learning models have been created for other cryptocurrencies and have been found to be almost accurate with exceptions regarding external factors like the sentimental nature of the people.

The ARMA-GARCH has been increasing in popularity for conducting research on various domains as we would like to study the effects of the mean as well as the variance models, rather than just focusing on only the mean or variance components of the model. Gencay et al. (2001) used the ARMA-GARCH model in order to determine the volatility of the Canadian Stock Market and it was found that ARMA-GARCH was pretty effective in showcasing the amount of volatility. Alizadeh and Nomikos (2015) used the same model to calculate the amount of risk of crude oil pieces and it was able to effectively determine the volatility.

Research Method

Data collection

The information were gathered from Yahoo Finance between March 2021 and October 2022. The meme coin's prices, the date and other information are provided. The close prices from the provided dataset are used to generate the price returns, which are then translated into a format that is appropriate for the procedure. The data is referred to as primary because it was already available on a website and wasn't gathered through surveys and questionnaires.

The methodology has been carried out using software like Python and R. Time Series Analysis is a method that is used to analyze various sets of data points that are distributed and represented by periods. The periods can be in various forms like day, week, month, year, etc. The time series models used are:

AR(p)- AR stands for Auto-Regressive Model. It is not always considered to be a stationary process. AR says that the value of the output is dependent on the previous values in a linear stochastic manner.

yt=c+Õ1yt”1+Õ2yt”2+…+Õpyt”p+åt

Where åt is the white noise and Õ is the parameter of the AR Model.

MA(q)- MA stands for Moving Average. It shows that the response variable is cross-correlated with a non- dependent variable. In this, the output is cross-correlated with a non-identical form of a random variable. The Moving Average Model is always considered to be stationary in nature. It is considered to be a finite impulse response filter for a white noise process.

yt=c+åt+è1åt”1+è2åt”2+…+èqåt”q

where åt”q is the white noise and è is the parameter.

ARMA (p, q)- ARMA stands for the Auto Regressive Moving Average Model. It is a combination of the Auto Regressive and the Moving Average Models. p stands for the lag order of the AR model and q stands for the lag order of the q model. This model considers the impact of the previous lags with the obtained. residuals.

yt=c+åt+ Õ1yt”1 +Õ2yt”2+ …+ Õpyt”p +è1åt”1+ è2åt”2 +…+ èqåt”q

ARIMA (p, d, q) - ARIMA stands for the Auto Regressive Integrated Moving Average. It is considered to be the generalized version of the Auto Regressive Moving Average. Here, we have an additional component q, defined as the degree of differencing of the data. The differentiation has to be continued until we get stationary data.

ARCH:

ARCH stands for Auto-Regressive Conditional Heteroskedasticity. It was developed to check the effects of volatility in various fields like finance, gold, etc. It was initially designed to improve econometric models. It is used in situations in which there may be short periods of increased variation.

t"1

$ó_t{}^2 = ó_0 + á_1X^2 + ... + á_p\, X^2{}_{t"p}$

GARCH (p, q)

GARCH is the generalized version of ARCH and is used if there is an autocorrelation in the variance terms.

t"p t"1 t"q

$ót^2 = ù + á1X^2t"1 + ... + áp\, X^2 + â1ó^2 + ... + âqó^2$ where á, â and ù are the parameters.

ARMA-GARCH:

The ARMA-GARCH is the hybridized version of ARMA and the GARCH processes. It provides a better representation as the volatility can affect the overall returns of the meme coin.

ARIMA-GARCH:

The ARIMA-GARCH is the hybridized version of the ARIMA and the GARCH processes. From the initial stages of the analysis, we have to check the stationarity of the model using the Augmented Dicky-Fuller Test, which states that the data is stationary if the p-value is more significant than the significance level taken.

Akaike and Bayesian Information Criterion:

The Akaike Information Criterion is an estimator which is used to check the quality of the time series models. As if the value of the AIC is the lowest, regardless of the absolute value, it is considered to be the best model.

AIC = 2k – 2 ln (5Ø?Ü) BIC = k ln(n) – 2 ln (5Ø?Ü)

k is the number of estimated parameters and 5Ø?Ü is the maximized value of the likelihood function.

The McLeod Li test is used to check the presence of the ARCH effect on the time series model. It showcases thelag values graphically. If the majority of the lag values are lesser than the level of significance, we have to reject the null hypothesis. We check the stationarity of the time series data using various tests like the Augmented Dicky Fuller test.

The level of significance (has been set to 0.05 for all the calculations.

The Value at Risk is a statistical method that is used to calculate and derive the possibility of the potential financialloss of commodities like stock, cryptocurrencies, gold, etc. It is always reported as a positive value despite the value technically signifying a loss.

Results

From the implementation of the Augmented Dicky Fuller Test, the data is stationary from the beginning as the p- value was found to be lesser than 0.05. Hence, the implementation of models like ARIMA and ARIMA-GARCHis not feasible.

The AIC and BIC of the selected AR, MA and ARMA values are given in Table1.

Table 1: AIC Values for the mean models

Model	AIC Values
ARMA (4,2)	-956.93
ARMA (4,1)	-953.57
ARMA (4,3)	-954.93
ARMA (3,3)	-954.15
ARMA (3,2)	-954.21
ARMA (3,1)	-953.74
ARMA (2,1)	-955.06
ARMA (2.2)	-952.3
AR (3)	-955.32
AR (1)	-954.22
AR (2)	-956.86
MA (1)	-955.58
MA (2)	-955.78
MA (3)	-955.13
MA (4)	-953.78

The model ARMA (4,2) has been re-confirmed on the application of the auto.arima function in R. We have to check for the presence of conditional heteroskedasticity.

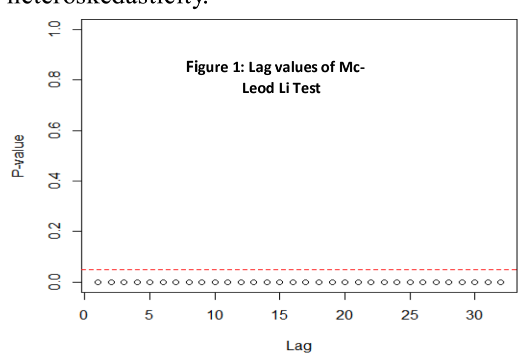

Figure 1: Lag values of Mc-Leod Li Test

On the application of the McLeod Li Test, the p values of the majority of the lags are less than 0.05. Hence, the null hypothesis is rejected and we can conclude that there is a significant presence of the ARCH/GARCH effect in the data. The GARCH value was calculated simultaneously with the ARMA (4,2) and the values of the AIC and BIC are showcased in Table2.

Model	AIC	BIC
ARMA(4,2)-GARCH(1,1)	**-1216.25**	**-1167.63**
ARMA(4,2)-GARCH(1,2)	-1215	-1161
ARMA(4,2)-GARCH(1,3)	-1213.3	-1155.84
ARMA(4,2)-GARCH(1,4)	-1212.9	-1151.02
ARMA(4,2)-GARCH(1,5)	-1209.05	-1142.75
ARMA(4,2)-GARCH(2,1)	-1214.25	-1161.21
ARMA(4,2)-GARCH(2,3)	-1211.3	-1149.42
ARMA(4,2)-GARCH(2,4)	-1209.3	-1143

To get a better interpretation of the forecasted values, we have to showcase the plots of the observed and the predicted values of the volatility.

In the plot below, the close price returns are depicted in dark blue and the fitted volatility is shown in red. We can observe that the volatility values almost fit in a similar pattern to that of the close price returns.

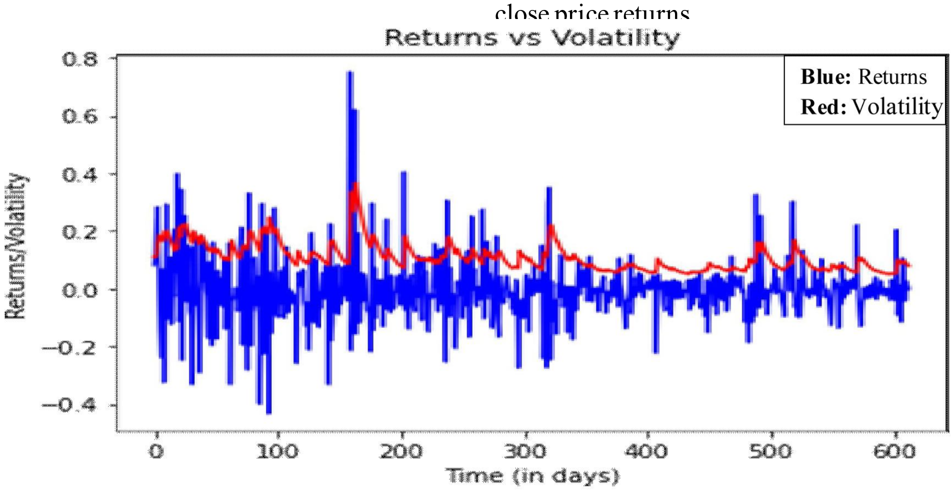

Figure 2: Returns vs Volatility Plot

In the graph below, we plot the predicted values of the ARMA-GARCH model as well as the future values of the Value at Risk for 90 days. We can see that there are positive as well as negative fluctuations that vary by each day. For some particular days, it could have been assumed that the coin would have collapsed but has made a quick recovery.

The plot in purple is the returns and the plot in red is the Value at Risk. The predictions are represented by black and green respectively.

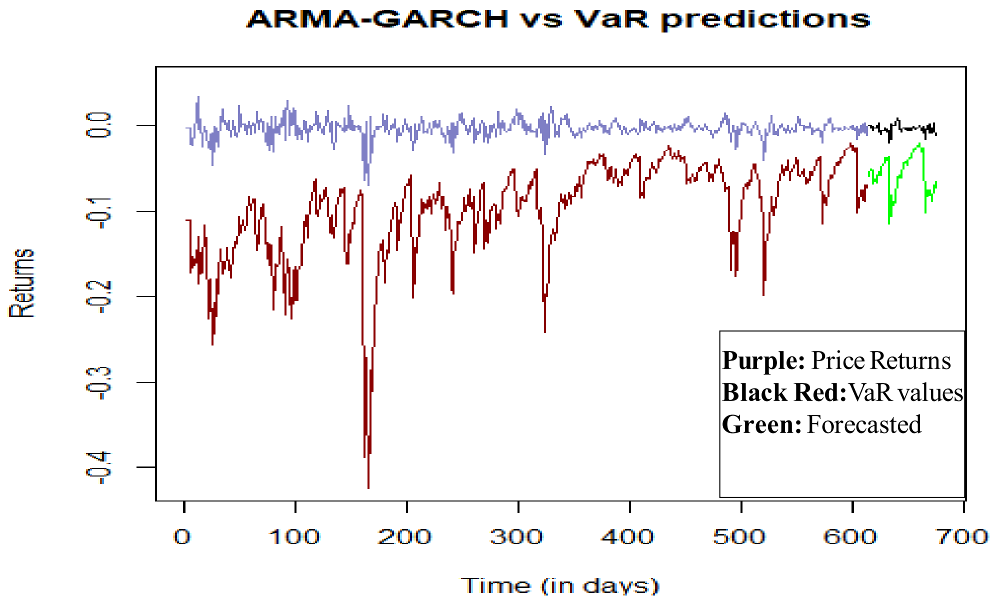

Figure 3: Plot of the ARMA-GARCH and the Value at Risk Observed and Predicted Values

Conclusion

We have fitted the best model for the price returns of Pepemon Pepeballs, which is given by the ARMA (4,2)- GARCH (1,1). The hybrid method performed better than the individual models as it contained both the mean as well as the variance models and incorporates the best of both worlds. For future studies, we can check for various unknown crypto currencies and more advanced versions of ARMA-GARCH, ARIMA-GARCH, etc.

Although the Value at Risk represents the risk of loss, it is always interpreted as a positive value. We can see that the coin did not suffer extremely high losses as compared to the ones shown by the VaR values.

From the given graphical representation of the returns, one must not rely on Pepemon Pepeballs in order to maximize earnings and should mainly be used to invest for further research purposes or for other non-serious purposes as the returns have a possibility of reducing further in the upcoming future.

References

Almansour B, Alshater M, Almansour A. Performance of ARCH and GARCH Models in Forecasting

Cryptocurrency Market Volatility. Industrial Engineering & Management Systems. 2021;20(2):130-139.

Chan S, Chu J, Nadarajah S, Osterrieder J. A Statistical Analysis of Cryptocurrencies. *Journal of Risk and Financial Management*. 2017;10(2):12.

Dash S, Ambastha K, Seraphim I, B S. Determining the Close Price of Bitcoin Using Regression Based on Blockchain Information. 2017;.

Conlon T, McGee R. Betting on Bitcoin: Does Gambling Volume on the Blockchain Explain Bitcoin Price Changes?. *SSRN Electronic Journal*. 2019;.

Ghani IM, Rahim HA. Modeling and forecasting of volatility using ARMA-GARCH: Case Study on malaysia natural rubber prices. IOP Conference Series: Materials Science and Engineering. 2019;548(1):012023.

Gao J. Research on stock price forecast based on Arima-GARCH model. E3S Web of Conferences. 2021;292:02030.

Juriæ V, Šimièeviæ V, Kajba D. Statistical Analysis of the Most Influential Cryptocurrencies. *SSRN Electronic Journal*. 2019;.

Kim J, Kwon D, Heo J, Kim C, Han Y. Time Series Classification of Cryptocurrency Price Trend Based on a Recurrent LSTM Neural Network. *Journal of Information Processing Systems*. 2019;15(3):694-706.

Koki C, Leonardos S, Piliouras G. Exploring the predictability of cryptocurrencies via Bayesian hidden Markov models. *Research in International Business and Finance*. 2022;59:101554.

Osterrieder J, Lorenz J. A Statistical Risk Assessment of Bitcoin and its Extreme Tail Behavior. *Annals of Financial Economics*. 2017;12(01):1750003.

Muhammad Fahmi A, Azah Samsudin N, Mustapha A, Razali N, Kamal Ahmad Khalid S. Regression based Analysis for Bitcoin Price Prediction. *International Journal of Engineering & Technology*. 2018;7(4.38):1070.

https://www.rdocumentation.org/packages/TSA/versions/1.3/topics/McLeod.Li.test

https://www.deccanherald.com/brandspot/pr-spot/the-evolution-of-meme-coins-dogecoin-shiba-inu-and- dogeliens-1138476.html

https://www.ndtv.com/business/meme-cryptocurrency-shiba-inu-meme-coin-elon-musks-twitter-buyout- gives-wings-to-dogecoins-price-3500399