Forecasting volatility of Bitcoin by applying GARCH and LSTM to improve momentum trading performance

Chan, Yun Kit Chan

ykchanbg@connect.ust.hk

Supervised by: Dr. David Rossiter

COMP 4971L – Independent Work

Department of Computer Science and Engineering

The Hong Kong University of Science and Technology

Fall 2023

Abstract

The use of volatility forecasts is crucial in estimating risk for financial institutions and extensive research has been done in this field. However, there is a lack of research on the application of volatility prediction in improving momentum trading performance. Therefore, this paper proposes a momentum trading strategy, CMI, which captures the fluctuation of Bitcoin and introduces a merge of GARCH and LSTM architecture. The paper also demonstrates that incorporating volatility prediction can improve trading performance by accurately identifying low-volatility periods and avoiding trades during these periods. This contributes to the existing literature on volatility in the financial market and highlights the potential benefits of incorporating volatility prediction in trading strategies.

Table of Contents

1. Introduction

1.1 Background

Bitcoin (BTC), the pioneering decentralized cryptocurrency, was introduced in 2009 by Nakamoto. Its emergence has sparked considerable attention from the media and academics. The decentralized nature of Bitcoin establishes a system wherein the verification of transactions is conducted without reliance on conventional intermediaries like banks *(The New York Times, 2022)*. Today, digital assets have evolved into a prominent financial instrument. There were more than 25,000 cryptocurrencies in the marketplace and worth more than $3,058 billion in November 2023 *(Coingecko, 2023)*.

The rapid growth of cryptocurrencies has presented significant investment opportunities, particularly in algorithmic trading. Researchers have found that the efficiency of the Bitcoin market is comparatively lower when compared to more established assets such as equities and foreign exchange *(Fung et al., 2021)*. Consequently, financial institutions have begun studying cryptocurrency investment to tap into its potential for greater profitability. Kosc et al. *(2019)* conducted a study on the momentum and contrarian effects of various cryptocurrency pairs, while Fil et al. *(2020)* employed the distance and cointegration method in pair trading to achieve consistent monthly returns. In addition to pair trading, several researchers have utilized neural networks to enhance investment decisions by generating trading signals. For instance, Nakano et al. *(2018)* explored Bitcoin technical trading using a neural network for return prediction, and Alkhodhairi et al. *(2018)* developed LSTM and GRU models for price prediction.

1.2 Volatility Prediction

In addition to return and price prediction, the forecasting of volatility also plays a crucial role in enhancing market risk models. Volatility, which indicates the degree of price fluctuations, is not directly observable and the fluctuations of volatility are often overlooked when implementing a trading strategy. For example, the Black-Scholes Model assumes constant volatility, but this parameter is dynamic in the real world. The fluctuations in volatility hold significant importance in estimating price behavior and managing risk effectively. Research conducted by Herremans et al. (2022) demonstrated that volatility predictions can help reduce risk when Bitcoin

trading.
Building upon these findings, this research aims to propose a momentum trading strategy that captures price fluctuations and further investigates the influence of volatility predictions on strategy performance.

With rapid Deep Learning development, researchers show increasing interest in employing neural networks especially Recurrent Neural Network (RNN) to this problem. One underlying reason is excellent performance in retaining information from previous steps in the time sequence data, making RNN effective in recognizing market fluctuations (Saxena, 2023). Long Short-Term Memory (LSTM) is an RNN architecture widely used for sequence prediction tasks. For instance, Sullivan et al. (Sullivan, 2018) employed LSTM to predict US equity market volatility; and Pratas et al. (2023) suggested that LSTM provides a promising tool for forecasting Bitcoin volatility, especially for short-term horizons.

Other than deep learning, researchers suggest that the generalized autoregressive conditional heteroskedasticity (GARCH) model outperforms feedforward artificial neural networks (ANN) and LSTM models in forecasting Bitcoin volatility, especially on shorter time horizons (one-day and three-day intervals). Gyamerah et al. *(2019)* have modeled Bitcoin volatility using various 3 GARCH variants - sGARCH, iGARCH, and tGARCH. Additionally, Bergsli et al. (2022) compared six different GARCH models and found that RGARCH and APARCH perform the best among them; and Klose (2022) used GARCH to forecast the volatility of cryptocurrency and gold.

However, the high volatility and unusual patterns and behavior of the cryptocurrency market make it challenging to apply GARCH (Langeland, 2015). To address this challenge, some scholars have proposed hybrid models with both classical and neural network techniques, resulting in significant improvements in prediction accuracy (Smyl et al., 2020). Zahid (2022) leveraged the strengths of both GARCH and RNN to enhance volatility forecasting and found that LSTM-GARCH outperform GARCH and GARCH-GRU. Similarly, Amirshahi et al. (2023) employed a GARCH-LSTM model to predict the volatility of 27 different cryptocurrencies. The existing literature suggests that integrating GARCH and neural network models may outperform using a single type of model.

1.3 Problem and hypotheses

In our study, we will begin by introducing a volume-based momentum trading strategy called the Candlestick Momentum Indicator (CMI). This strategy aims to quantify momentum, capture market fluctuations, and identify potential trading opportunities. We will then develop a GARCH-LSTM model to forecast realized Bitcoin volatility with the rolling window technique and incorporate the prediction into the CMI trading strategy. Our research contributes to the existing literature on hybrid models in several ways. Firstly, we will summarize and compare the predictive power of different GARCH variants found in various studies and select the combination that performs the best (Perez et al., 2021; Zahid et al., 2022; Madina et al., 2023). Secondly, we will enhance our LSTM model's prediction capabilities by incorporating parameters from the selected GARCH model, as well as Bitcoin log-returns and the CMI as inputs to improve the accuracy of our volatility forecasts. Lastly, we will integrate the predictions from our LSTM model into the CMI-based trading strategy. By identifying periods of expected low volatility, we aim to avoid low price fluctuation periods, enhance the performance of our trading strategy, and discuss their implication.

1.4 Paper Structure

This paper is structured as follows: Section 2 provides the theoretical background of GARCH, LSTM, and hybrid model, and Section 3 presents the data and benchmark model. Section 4 is an overview of the hybrid model and trading strategy. Section 5 describes the modeling result and backtesting result.

2. Technical Background

2.1 GARCH

ARCH and GARCH were developed by Engle and Bollerslev and GARCH has become one of the most important models for forecasting volatility. GARCH is a statistical modeling technique used to help predict the volatility of returns on financial assets. (Investopedia, 2021)

Since the original introduction, many variations of GARCH have emerged. These include Integrated GARCH (IGARCH), which restricts the volatility parameter, and Threshold GARCH (TGARCH), which models the volatility differently depending on whether the market is in a high or low volatility regime.

To model a time series using an ARCH process, return residuals ϵ_t is given by the following equation:

$$\epsilon = \sigma_t Z_t$$

, where Z_t is a strong white noise process and σ_t represents a time-dependent standard deviation. The representation of σ_2^2 is given by the following equation:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i \epsilon_{t-j}^2$$

, where $\alpha_0 > 0$, $\alpha_i \geq 0$ and $j > 0$. The equation is known as the $ARCH(p)$ model.

If an ARMA model is assumed for the error variance, the model is a GARCH model *(Bollerslev, 1986)*. Bollerslev extended the $ARCH(p)$ method to introduce the $GARCH(p, q)$ defined by the following equation:

$$\sigma_t^2 = w + \sum_{i=1}^{p} \alpha_i \, \epsilon_{t-i}^2 + \sum_{i=1}^{q} \beta_i \, \sigma_{t-i}^2$$

, where p is the order of the ARCH terms ϵ^2 and q is the order of the GARCH terms σ^2.

2.2 Long short-term memory

Long Short-Term Memory Networks (LSTM) is a type of deep learning model that belongs to the family of recurrent neural networks (RNN). Unlike traditional RNNs, LSTM is specifically designed to address the vanishing gradient problem commonly encountered in sequential data analysis and it is possible to forget past irrelevant information. LSTM was introduced by Hochreiter and Schmidhuber in 1997 as a solution to the limitations of traditional RNNs and other machine learning algorithms. It allows information to persist and be retained for longer periods, enabling the model to effectively capture long-term dependencies in sequential data.

An LSTM model consists of multiple memory cells that interact with each other through various gates. These gates include an input gate, an output gate, and a forget gate *(Siegelmann et al., 1991)*. At each time step, the LSTM network takes an input and updates its internal state based on the information from the input, the previous state, and the output from the previous time step.

The forget gate decides which information from the previous state should be discarded. The input gate determines how much of the new input should be stored in the memory cell, and each memory cell has three sigmoid layers and one tanh layer. The combination of the forget gate and the input gate controls the flow of information into and out of the memory cell. The LSTM network also has an output gate that regulates how much of the memory cell's content should be used to generate the output at the current time step. The output is presented by h_t, while c_t represents the value of the memory cell.

Figure 1 displays the structure of the LSTM cell, including three gates and the output of the LSTM unit.

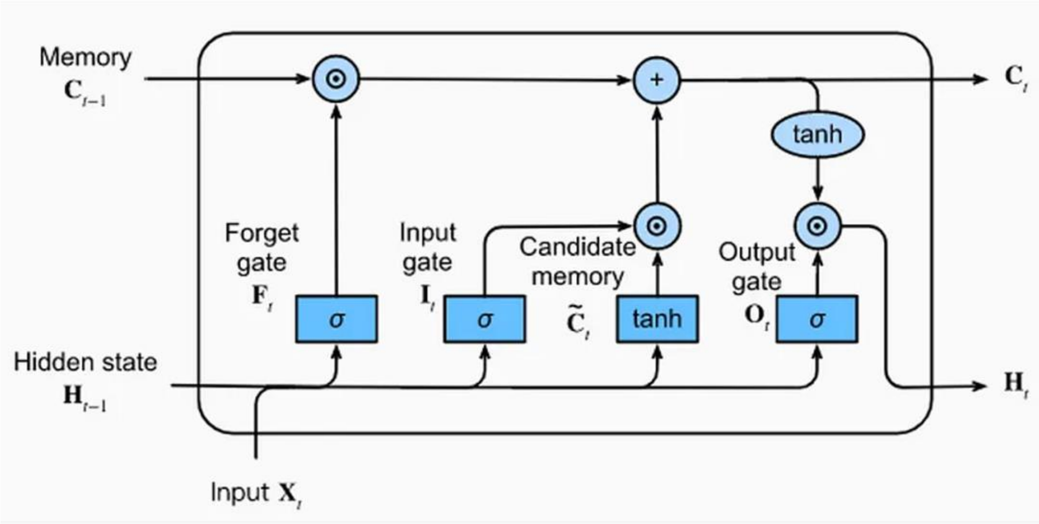


Figure 1. Structure of the LSTM cell
Image Source: Dive into Deep Learning

3. Technical Background

3.1 Data Description

We obtained the price and volume data of BTCUSDT Perpetual Contract (BTCUSDT.P) from Binance, the largest cryptocurrency exchange in terms of daily trading volume, through its official Application Programming Interface (API). The data covers a time interval of 4 hours, starting from 2019-09-09 00:00 (UTC+8) and ending on 2023-11-15 00:00 (UTC+8), with 9168 observations in total. We specifically used closing, high, and low prices denoted as P_t, P_t^h, and P_t^l, respectively.

We use BTCUSDT.P price data from a specific exchange instead of BTC spot price to closely simulate real-world trading execution. Since the BTC spot can only be used for long positions, we choose BTCUSDT.P for our long-short trading strategy. In this study, we will not be considering Bitcoin strategy ETFs in the equity market or BTC spot trading in decentralized exchanges (DEX) due to lower volatility and difficulties in data collection.

In this paper, we focus solely on studying Bitcoin. Bitcoin is chosen due to its status as the largest and oldest cryptocurrency, with high trading volume and market capitalization. Additionally, the availability of extensive historical data, especially price data of perpetual contracts, makes it a suitable choice for analysis. In addition, the strong correlation between cryptocurrencies suggests that the findings may also be applicable to other cryptocurrencies (Burnie et al., 2018). This will provide a more comprehensive evaluation of the strategy's effectiveness.

Additionally, Bitcoin is considered to have a lower susceptibility to market manipulation risks due to high market capitalization. Recent investigations have uncovered concerning instances of market manipulation in the cryptocurrency landscape. Deceptive actors are manipulating pool liquidity, causing small-sized altcoin prices to skyrocket by up to 22,000% *(Vanunu, 2023)*. This highlights the need for high liquidity in the market to protect against price manipulation. In low liquidity markets, a single large trade can significantly impact an asset's price, making it vulnerable to manipulation *(The Block, 2023)*. Therefore, maintaining price stability and a fair trading environment is crucial, and liquidity plays a crucial role in achieving this in the cryptocurrency market.

4. Methodology

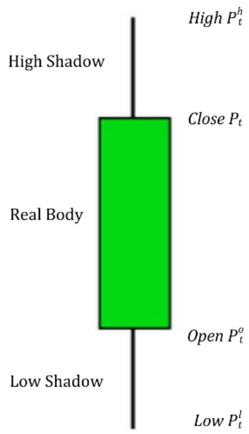
In the upcoming section, the proposed momentum trading strategy and volatility model methodology will be explained in detail. This will include the integration of the trading strategy with the volatility prediction results, as well as the criteria for opening and exiting trades. The associated trading costs will also be discussed.

This section consists of five subsections. Section 4.1 introduces the Candlestick Momentum Indicator (CMI), which will be applied to the momentum trading strategy. Section 4.2 discusses the design of GARCH and its variant models. Section 4.3 explores hybrid models for volatility forecasting and benchmarking models. Section 4.4 presents a trading strategy that integrates CMI and volatility forecasts. Finally, Section 4.5 describes the performance metrics used for evaluating volatility modeling and the trading strategy.

Training and testing period: In order to evaluate the performance of the proposed methodology, the data will be split into a training set and a testing set at a ratio of 1:2. We will split the data into the training set and testing set. For our model testing, we will be using data from 2021-01-28 16:00 to 2023-11-13 12:00, resulting in a total of 6113 trading periods over a duration of 1018 days. This will provide a comprehensive evaluation of the performance of the volatility models and trading strategy.

4.1 Momentum Indicator CMI

The Candlestick Momentum Indicator (CMI) is a unique momentum technical indicator that combines elements from candlestick charting with volume. It is designed to capture the momentum of price movements within a given time period. The formula for calculating CMI is as follows:



CMI = (Real Body + Low Shadow - High Shadow) / (High - Low) * Volume
CMI = (2 * Close - High - Low) / (High - Low) * Volume

The rationale behind the Candlestick Momentum Indicator (CMI) lies in its ability to capture and quantify the strength of buying and selling pressure in the market. A larger real body and low shadow suggest a stronger long momentum during the specified time period, while a larger high shadow suggests stronger selling pressure. To further refine the momentum measurement, the CMI formula normalizes the indicator by dividing it by the difference between the high and low prices. This normalization ensures that the CMI reflects the momentum relative to the price range. The CMI incorporates volume into the calculation by multiplying the normalized momentum by the trading volume. This amplification factor helps to emphasize the extent of momentum relative to the trading activity in the market.

The trading strategy is intended to capture significant reversion from high momentum positions. For example, when the CMI is high and overbought, it is expected that the momentum will drop, and therefore, the cryptocurrency will be shorted. The graph below illustrates the relationship between BTCUSDT.P price and CMI. CMI is used to summarize the momentum of past candlesticks, and the strategy will enter a trade in the opposite direction if the long/short momentum is excessively high.

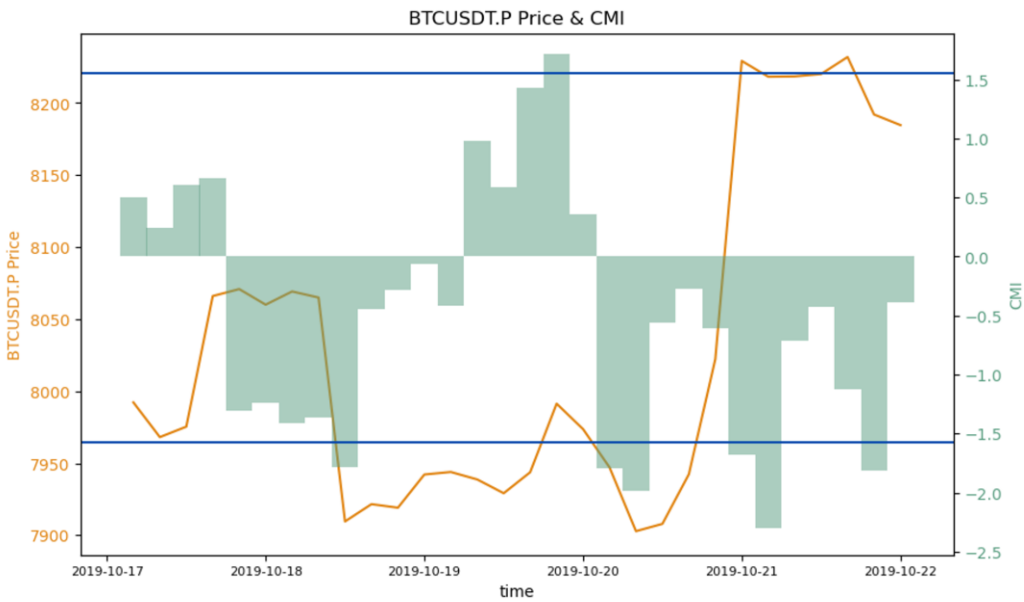


Figure 2: BTCUSDT.P Price & CMI

4.2 GARCH Design

Since the original introduction, many variations of GARCH have emerged. These include exponential GARCH (EGARCH), Glosten-Jagannathan-Runkle GARCH (GJR-GARCH), and Threshold GARCH (TGARCH). All the GARCH model variations seek to incorporate the direction, positive or negative, of returns in addition to the magnitude (addressed in the original model) (Investopedia, 2021).

Previous studies have utilized GARCH models for volatility prediction in the finance market *(Lim et al., 2013)*. Through analyzing and evaluating the performance of various GARCH models in different studies *(Perez et al., 2021; Zahid et al., 2022; Madina et al., 2023)*, we have selected 6 models to be used alongside LSTM-GARCH and as benchmark models based on the volatility prediction performance: GARCH, ADD-EGARCH, AVGARCH, EGARCH, GJRGARCH, TGARCH.

The log-returns will be the input for GARCH model. The log-returns r_t are calculated as:

$$\mathrm{r_t} = \log(\mathrm{p_t}) - \log(\mathrm{p_{t-1}})$$

, where p_t is the price data at time t.

4.3 LSTM-GARCH Hybrid Model

The LSTM-GARCH models incorporate 6 GARCH-based algorithms mentioned in section 4.3. The volatility estimates from GARCH and variant models are fed into an LSTM layer with 32 units, followed by 2 feed-forward layers with 8 and 1 neurons, respectively. Additionally, the input includes the standard deviation of volatility and CMI.

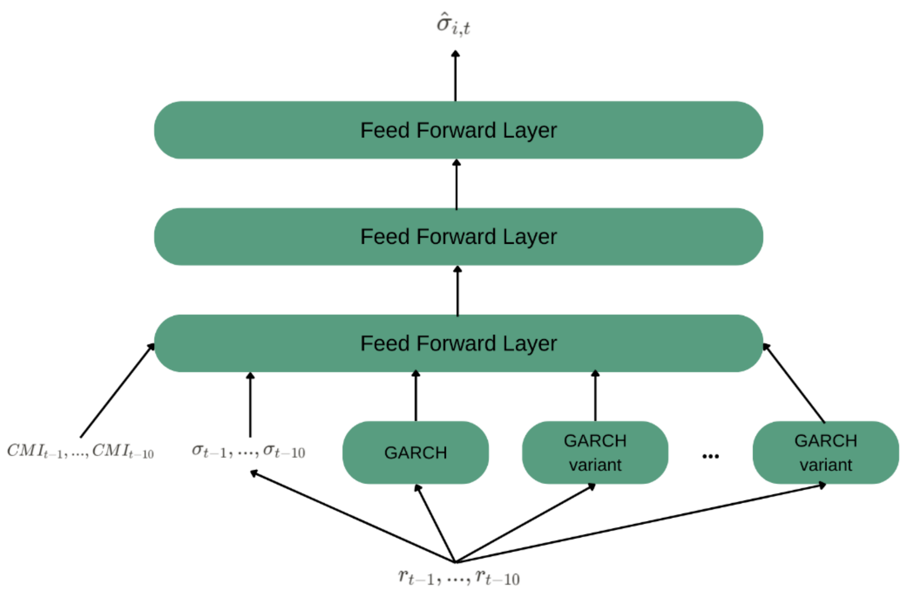


Figure 3: LSTM-GARCH architecture

CMI_{t-1} is defined as the CMI at time $t-1$. The true implied volatility, $\widehat{\sigma_{1,t}}$ is used as a response variable to train the models. This variable is the standard deviation of the future logarithmic returns:

$$\widehat{\sigma_{1,t}} = \sqrt{\frac{\sum_{n=0}^{i-1}(r_{t+n} - \overline{r_f})^2}{i-1}}$$

where r_{t+n} represents the daily logarithmic return after n unit of time periods, $\overline{r_f}$ is the mean logarithmic return from t_0 to t_{i-1}. In this paper, $i = 4$.

The proposed architectures and benchmark models will be fitted using the rolling window approach, with a fixed sample length for fitting the model and forecasting the next step. In this paper, the window size is set to 600 and the forecast horizon is 4 horizon (16 hours in total, 4 hours for each horizon). This process will be repeated until the entire period being analyzed has been forecasted.

All the coding work is built in Python. Some of the well-developed packages foster development and are very helpful for quantitative trading research. Firstly, the sklearn and Keras package helps in the data preparation process, including normalization and clearing. Regarding data scraping and data collection, Binance Public API Connector implements the process. Last but not least, we adopt ARCH and TensorFlow packages to construct the 6 GARCH models and neural network.

4.4 Benchmark Model

The 6 GARCH models used in the LSTM-GARCH framework will also serve as benchmark models for comparison. These models have been previously presented in the same subsection and will be evaluated based on their performance in predicting volatility and reversion of high-momentum positions.

4.5 Trading Logic

4.5.1 Without Volatility prediction

The trading strategy will implement a fixed volatility threshold, similar to the conventional approach of technical indicators such as RSI. For example, if the RSI is below 30, a short position will be taken, and if it is above 70, a long position will be taken. In this paper, low levels of CMI, below the 10th percentile, indicate an oversold or undervalued condition and generate buy signals, while high levels of CMI, above the 90th percentile, suggest an overbought or overvalued security and generate sell signal. The trade will be exited after 4 time periods or exit to get profit if CMI touches the opposite threshold. For this study, we will be using 4-hour price data, and trades will be closed after 16 hours.

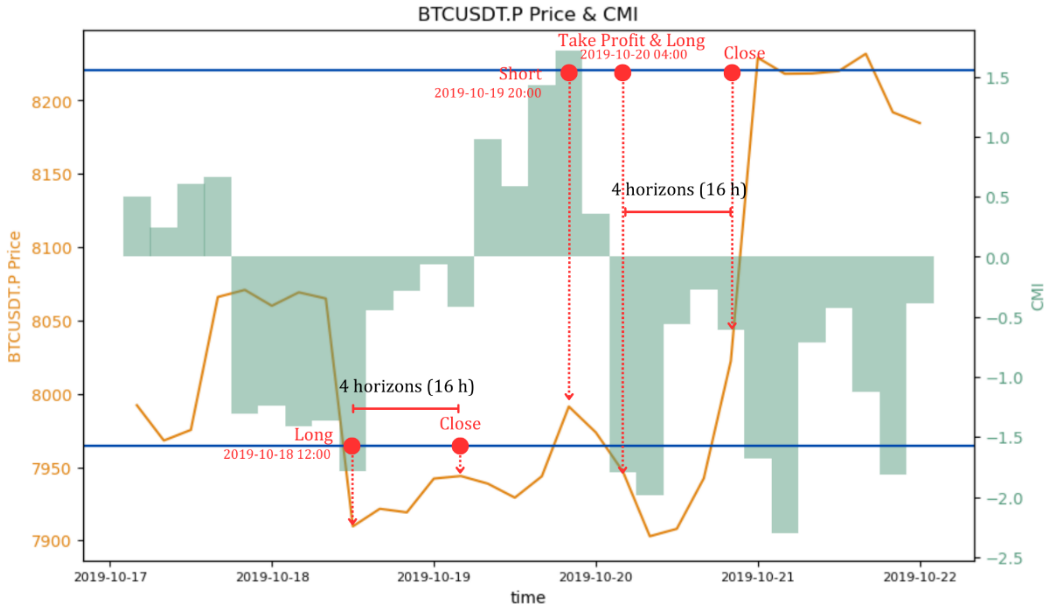


Figure 4: Demonstration of trading logic without Volatility prediction

Figure 4 demonstrates The trading logic without integrating volatility prediction is in. On 2019-10-18 12:00, the CMI is below the lower threshold (lower blue line), indicating a long trade entry. This trade will be closed after 4 horizons. On 2019-10-19 20:00, as the CMI rises above the upper threshold (upper blue line), a short trade is entered. As the CMI drops below the lower threshold at 2019-10-20 04:00, the trade is closed early and a long trade is entered simultaneously. This trade will also be closed after 4 horizons.

4.5.2 With Volatility prediction

By incorporating volatility prediction into the trading strategy, we aim to improve its profitability by avoiding trades during low volatility periods. In this study, we have set a threshold of the 20th percentile for volatility prediction, meaning that we will not enter a trade if the predicted volatility falls below this point.

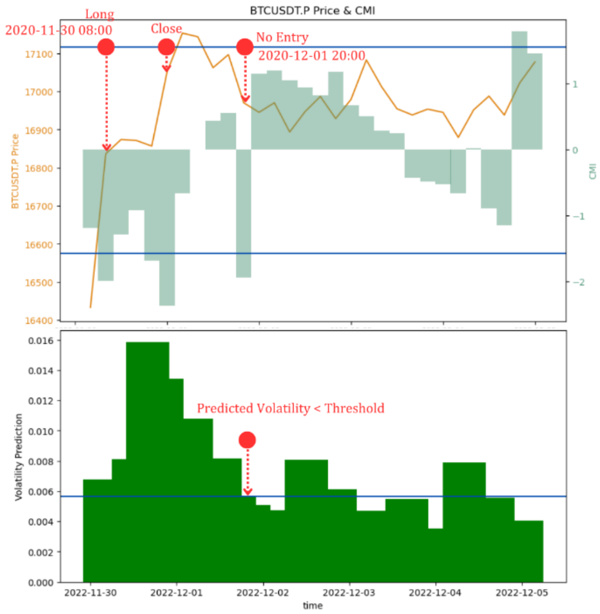


Figure 5: Demonstration of trading logic Volatility prediction

The optimized trading logic is shown in Figure 5. As seen in the green chart at the bottom, the predicted volatility on 2020-11-30 08:00 is greater than the predetermined threshold. Meanwhile, the CMI (shown in the shamrock color at the top) is also above the entry threshold, indicating a long trade entry. This trade will be closed after 4 horizons, which is equal to 16 hours. On 2022-12-01 20:00, although the CMI is below the short threshold, the predicted volatility is also below the predetermined threshold, resulting in no trade entry at this time.

4.5.3 Benchmark Strategy

As a benchmark model for our study, we will be using the buy-and-hold strategy for BTC. This strategy involves purchasing BTC and holding onto it for the entire duration of the study, without making any trades. The performance of this strategy will be compared to the performance of our proposed momentum trading strategy to assess its effectiveness.

4.6 Model Performance Metrics

These metrics will provide valuable insights into the effectiveness and profitability of the models, as well as the success of the trading strategy. The following section will discuss the different performance metrics used in this study. We will use Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and recall to evaluate the performance of volatility prediction.

4.6.1 MSE & RMSE

MSE and RMSE are two commonly used performance indicators for machine learning models. They measure the average difference between the model output and the actual values, providing an estimation of the model's accuracy in predicting the target value.

$$\mathrm{MSE} = \frac{\sum_{i=1}^{N}(x_i - \hat{x}_i)^2}{N}$$

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \hat{x}_i)^2}{N}}$$

, where N is the length of data, x_i is the actual observations in the time series and $\hat{x}_i$ is the estimated time series.

4.6.2 Recall & Precision

Precision and Recall are two commonly used metrics for evaluating the performance of a binary classification model. Precision measures the percentage of correct positive predictions out of all positive predictions, while Recall measures the percentage of correctly predicted positive values out of all actual positive values. In this study, these metrics will be used to evaluate the accuracy of the volatility prediction. In the trading strategy, a 20th percentile threshold will be used, meaning that only trades with a volatility prediction higher than the 20th percentile will be entered. Therefore, any volatility prediction below the 20th percentile will be classified as "Low volatility.”

However, prediction output is classified based on a fixed percentile within the data series, so the recall and precision values will be the same. This is because both measures are calculated using the same classification threshold. In this paper, the threshold is set to be the 20th percentile. We will only present recall in section 5.

4.7 Trading Strategy Performance Metrics

The performance of the strategy will be evaluated using Return & CAGR, Sharpe ratio, Sortino ratio, and Maximum Drawdown (MDD). These metrics will help determine the profitability, risk, and risk-adjusted returns of the strategy.

4.7.1 Return & CAGR

The return of a trading strategy is a key indicator of its profitability. It measures the profit earned over the testing period (1018 days in this paper). Compound Annual Growth Rate (CAGR) is another commonly used measure, which calculates the average annual return of the strategies over a year.

4.7.2 Sharpe ratio

The Sharpe ratio is a metric that compares the risk and return of an investment. It takes into account the volatility and risk of an investment over some time, rather than just the returns, to provide a more comprehensive measure of investment performance *(Investopedia, 2023)*. The Sharpe ratio is computed as follows:

$$\text{Sharpe Ratio} = \frac{\mathrm{R_p} - \mathrm{R_f}}{\sigma_\mathrm{p}}$$

, where R_p is the return of the portfolio, R_f is the risk-free rate and σ_p is the standard deviation of the portfolio's excess return.

4.7.3 Sortino ratio

The Sortino ratio is a variation of the Sharpe ratio that aims to differentiate between harmful and overall volatility by using the downside deviation of portfolio returns instead of the total standard deviation. This ratio takes into account the risk-free rate and is named after its creator, Frank A. Sortino *(Investopedia, 2020)*. The calculation of the Sortino ratio is described as follows:

$$\text{Sortino Ratio} = \frac{\mathrm{R_p} - \mathrm{R_f}}{\sigma_\mathrm{d}}$$

, where R_p is the return of the portfolio, R_f is the risk-free rate and σ_p is the standard deviation of the portfolio's excess return.

4.7.4 Maximum Drawdown

Maximum Drawdown (MDD) is a measure of the largest loss observed in a portfolio from its peak value to its lowest value before a new peak is achieved. It is used as an indicator of the potential downside risk over a specific time period *(Investopedia, 2022)*.

4.7.5 Trading Cost

While a higher number of trades may result in a higher overall profit, it also comes with additional trading and borrowing costs. To accurately assess the performance of the strategy, we will take into account these costs. For this study, we will be factoring in a trading cost of 0.05% for both entry and exit, based on the Binance USDT Futures trading fee. The study will not consider the funding rate as it is deemed negligible and will not significantly impact the results.

5. Result and Discussion

In this section, we will evaluate the proposed strategy and discuss the results obtained by applying the methodology discussed in section 4 to show the effectiveness of the proposed prediction model and trading strategy.

5.1 Volatility prediction evaluation

Before discussing the trading performance, we will start with the result of the volatility model. The following part compares the LSTM-GARCH model with the benchmark models (6 GARCH models, LSTM). Table 1 presents the MSE and RMSE of different models. We found that the GARCH and EGARCH models underperform tremendously with other benchmark models in terms of MSE and GARCH, while LSTM-GARCH. GARCH and AVGARCH have similar performance in terms of recall. LSTM-GARCH outperform all the benchmark by integrating GARCH volatility models and LSTM.

Model	**Performance (ranking)**		
	MSE	RMSE	Recall
GARCH	278140814	16677.6	0.78199
ADD-EGARCH	0.00173	0.04163	0.72690
AVGARCH	0.00124	0.03527	0.78370
EGARCH	1.0754e+50	1.0370e+25	0.70722
GJRGARCH	0.00321	0.05665	0.74699
TGARCH	0.00105	0.03245	0.74870
LSTM-GRACH	0.00017	0.01329	0.80792

Table 1: Performance of volatility prediction model

The results show that GARCH and EGARCH have extremely high MSE and RMSE values, which is due to their failure to capture the asymmetric effects of market shocks on volatility. The results also show that GARCH and the variant models have strong predicting power in volatility, with recall greater than 70%. On the other hand, LSTM-GARCH has the best performance, with the lowest MSE and RMSE values and the highest recall. This could be attributed to the fact that LSTM-GARCH combines the long-term memory capabilities of LSTM with the predictive power of GARCH and variant models, resulting in more accurate volatility predictions.

5.2 Strategy performance

In this section, trading results of integrating with different volatility models are backtested with the trading logic mentioned in section 4. Furthermore, we use a simple buy-and-hold strategy for BTCUSDT.P from Binance over the testing period. Table 2 describes the key results of the study.

Model/ Strategy	**Performance (ranking)**					
	Return (over testing period)	CAGR	Sharpe Ratio	Sortino ratio	MDD	No. of trade
GARCH	220.0%	53.5%	2.31610	3.55311	34.584%	479
ADD-EGARCH	190.5%	48.1%	2.12626	3.22801	35.247%	451
AVGARCH	176.5%	45.4%	2.01455	3.04632	37.352%	484
EGARCH	176.9%	45.5%	2.00611	3.04383	41.046%	526
GJRGARCH	188.3%	47.7%	2.12323	3.24552	40.257%	465
TGARCH	132.1%	36.4%	1.69428	2.54538	41.635%	477
LSTM-GRACH	227.7%	54.8%	2.33768	3.56882	33.969%	469
CMI (No prediction)	182.7%	46.6%	2.01879	3.08830	34.247%	580
BTC B&H	-1.93%	-0.7%	-0.00661	-0.01044	77.083%	-

Table 2: Backtest performance with different model/ strategy

Table 2 shows the main result obtained from this study. It consists of the return, Sharpe ratio, Sortino ratio, MDD, and the number of trades, after considering transaction costs. Based on the results, the buy-and-hold strategy for BTC has a negative return and the worst performance overall, indicating that it is not a viable strategy for cryptocurrency trading.

Even though Bitcoin has no profit in the backtesting period, the results also show that LSTM-GARCH has the highest return and the best risk-adjusted performance among the models evaluated. It has the highest Sharpe and Sortino ratios and the lowest MDD, indicating that it is the most effective model for predicting cryptocurrency market volatility and generating profitable trades. This suggests that incorporating volatility prediction can improve trading performance by identifying low-volatility periods and avoiding unprofitable trades.

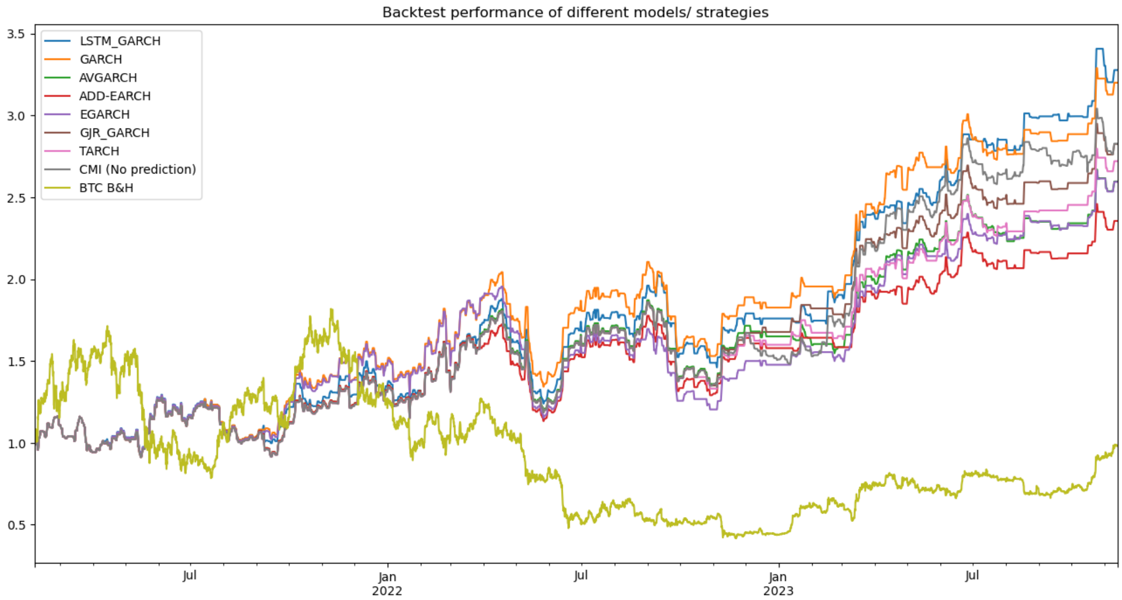


Figure 6: Backtest performance with different models/ strategies with time

In the bitcoin market, the CMI trading strategy demonstrated a high level of flexibility and profitability. While Bitcoin rose by 70% in the first four months of the testing period, the strategy only saw a 15% increase. However, the strategy maintained its profitability even during the bear market from November 2021 to January 2023. During this time, Bitcoin dropped by 75%, but the CMI trading strategy was able to capture stable profits by going long and short, resulting in a 42% increase. This further supports the effectiveness of the momentum trading strategy, as it was able to generate a 30% profit in the period of April 2023 to September 2023, during which Bitcoin remained within a 30% range. The results demonstrate the profitability of the trading strategy, even in times when Bitcoin is struggling to rise or fall.

5.3 Strategy performance

In conclusion, the results suggest that the momentum trading strategy proposed in this study may outperform the BTC market in terms of both absolute and risk-adjusted returns. The incorporation of volatility models further enhances the trading strategy by accurately identifying low-volatility periods and preventing trades during these periods.

6. Conclusion and Recommendations

6.1 Discussion of the result

In recent years, Bitcoin has gained significant attention from researchers due to its unique price patterns and speculative. These factors have sparked interest in various research topics related to this new asset class. While there is existing literature focusing on the momentum trading strategy in the cryptocurrency market, there is a lack of research on the impact of volatility on its performance. This highlights the potential for further exploration and contribution to the existing literature.

This study contributes to the development of a momentum strategy that captures the fluctuation of Bitcoin price movement. Our findings suggest that the merged use of various GARCH models and the CMI momentum indicator may be more effective in forecasting Bitcoin volatility. Additionally, we suggest the use of the LSTM-GARCH volatility model, which incorporates the long-term memory capabilities of LSTM and the predictive power of GARCH. This model has been shown to be more effective in predicting market volatility compared to traditional GARCH models. Our results are consistent with prior research and highlight the significance of these new methodologies for momentum trading.

6.2 Future Plan

Dealing with extreme observations can be challenging, particularly when analyzing cryptocurrency returns. Recent research *(Dakos et al., 2020)* has shown that the existence of extreme observations in cryptocurrency returns can result in misleading outcomes when assessing risk, especially for small-cap altcoins. This poses a potential challenge for non-robust volatility models such as GARCH. To validate this further, it is necessary to conduct additional research using other cryptocurrencies.

In addition, recent literature has explored the use of Transformer models for volatility prediction, which have demonstrated the potential to outperform traditional LSTM-GARCH models. Leveraging the power of the Transformer model may offer improved prediction performance. Therefore, incorporating a Transformer model into the analysis could be a promising approach to enhance volatility prediction.

7. References

1. Amirshahi, B., & Lahmiri, S. (2023). Hybrid deep learning and garch-family models for forecasting volatility of Cryptocurrencies. Machine Learning with Applications, 12, 100465. https://doi.org/10.1016/j.mlwa.2023.100465
2. Asante Gyamerah, S. (2019). Modelling the volatility of Bitcoin returns using GARCH models. Quantitative Finance and Economics, 3(4), 739–753. https://doi.org/10.3934/qfe.2019.4.739
3. Bergsli, L. Ø., Lind, A. F., Molnár, P., & Polasik, M. (2022). Forecasting volatility of bitcoin. Research in International Business and Finance, 59, 101540. https://doi.org/10.1016/j.ribaf.2021.101540
4. Cryptocurrency prices, charts, and crypto market cap. CoinGecko. (n.d.). https://www.coingecko.com/
5. Gupta, V., & Ahmad, M. (2019). Stock price trend prediction with long short-term memory neural networks. International Journal of Computational Intelligence Studies, 8(4), 289. https://doi.org/10.1504/ijcistudies.2019.103619
6. K. Alkhodhairi, R., R. Aljalhami, S., K. Rusayni, N., F. Alshobaili, J., A. Al-Shargabi, A., & Alabdulatif, A. (2021). Bitcoin candlestick prediction with deep neural networks based on Real Time Data. Computers, Materials & Continua, 68(3), 3215–3233. https://doi.org/10.32604/cmc.2021.016881
7. Pratas, T. E., Ramos, F. R., & Rubio, L. (2023). Forecasting bitcoin volatility: Exploring the potential of deep learning. Eurasian Economic Review, 13(2), 285–305. https://doi.org/10.1007/s40822-023-00232-0
8. Saxena, S. (2023, October 25). What is LSTM? introduction to long short-term memory. Analytics Vidhya. https://www.analyticsvidhya.com/blog/2021/03/introduction-to-long-short-term-memory-lstm/#:~:text=LSTM
9. V-lab: Bitcoin to US dollar garch volatility analysis. Real-time Financial Volatility, Correlation, And Risk Measurement, Modeling, And Forecasting. (n.d.). https://vlab.stern.nyu.edu/volatility/VOL.BTCUSD%3AFOREX-R.GARCH
10. Yaffe-bellany, D. (2022, September 15). Crypto’s long-awaited “merge” reaches the finish line. The New York Times. https://www.nytimes.com/2022/09/15/technology/ethereum-merge-crypto.html
11. 10.1. Long Short-Term Memory (LSTM) - Dive into Deep Learning 1.0.3 documentation. (n.d.). https://d2l.ai/chapter_recurrent-modern/lstm.html
12. Copeland, T. (2023, October 13). What is liquidity and why does it matter?. The Block. https://www.theblock.co/learn/251470/what-is-liquidity-and-why-does-it-

matter

13. Etal. (2023, December 5). Crypto deception unveiled: Check point research reports manipulation of pool liquidity skyrockets token price by 22,000%. Check Point Research. https://research.checkpoint.com/2023/crypto-deception-unveiled-check-point-research-reports-manipulation-of-pool-liquidity-skyrockets-token-price-by-22000/
14. García-Medina, A., & Aguayo-Moreno, E. (2023). LSTM–GARCH hybrid model for the prediction of volatility in cryptocurrency portfolios. Computational Economics. https://doi.org/10.1007/s10614-023-10373-8
15. Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735
16. Investopedia. (n.d.). Investopedia. https://www.investopedia.com/
17. Powell, F. (2023, December 4). 10 best crypto apps & exchanges of December 2023. Forbes. https://www.forbes.com/advisor/investing/cryptocurrency/best-crypto-exchanges/#:~:text=Global%20Crypto%20Exchanges,quality%20and%20safety%20vary%20widely
18. Ramos-Pérez, E., Alonso-González, P. J., & Núñez-Velázquez, J. J. (2021). Multi-transformer: A new neural network-based architecture for forecasting S&P volatility. Mathematics, 9(15), 1794. https://doi.org/10.3390/math9151794
19. SAP help portal. (n.d.). https://help.sap.com/docs/SAP_PREDICTIVE_ANALYTICS/41d1a6d4e7574e32b815f1cc87c00f42/5e5198fd4afe4ae5b48fefe0d3161810.html
20. Team, T. I. (n.d.). GARCH model: Definition and uses in statistics. Investopedia. https://www.investopedia.com/terms/g/garch.asp
21. Bornholt, G., Dou, P., & Malin, M. (2015). Trading volume and momentum: The International Evidence. Multinational Finance Journal, 19(4), 267–313. https://doi.org/10.17578/19-4-2
22. Catania, L., & Grassi, S. (2022). Forecasting cryptocurrency volatility. International Journal of Forecasting, 38(3), 878–894. https://doi.org/10.1016/j.ijforecast.2021.06.005
23. Fil, M., & Kristoufek, L. (2020). Pairs trading in cryptocurrency markets. IEEE Access, 8, 172644–172651. https://doi.org/10.1109/access.2020.3024619
24. Fung, K., Jeong, J., & Pereira, J. (2022). More to Cryptos Than Bitcoin: A garch modelling of heterogeneous cryptocurrencies. Finance Research Letters, 47, 102544. https://doi.org/10.1016/j.frl.2021.102544
25. Herremans, D., & Low, K. W. (2022). Forecasting bitcoin volatility spikes from

whale transactions and cryptoquant data using synthesizer transformer models. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.4247685

26. Kosc, K., Sakowski, P., & Ślepaczuk, R. (2019). Momentum and contrarian effects on the cryptocurrency market. Physica A: Statistical Mechanics and Its Applications, 523, 691–701. https://doi.org/10.1016/j.physa.2019.02.057
27. Nagel, S. (2002). Is Momentum Caused by Delayed Overreaction? .
28. Nakano, M., Takahashi, A., & Takahashi, S. (2018). Bitcoin technical trading with Artificial Neural Network. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3128726
29. Pratas, T. E., Ramos, F. R., & Rubio, L. (2023). Forecasting bitcoin volatility: Exploring the potential of deep learning. Eurasian Economic Review, 13(2), 285–305. https://doi.org/10.1007/s40822-023-00232-0
30. Ramos-Pérez, E., Alonso-González, P. J., & Núñez-Velázquez, J. J. (2021). Multi-transformer: A new neural network-based architecture for forecasting S&P volatility. Mathematics, 9(15), 1794. https://doi.org/10.3390/math9151794
31. Zahid, M., Iqbal, F., & Koutmos, D. (2022). Forecasting bitcoin volatility using hybrid GARCH models with machine learning. Risks, 10(12), 237. https://doi.org/10.3390/risks10120237
32. Azizi, S. P., Huang, C. Y., Chen, T. A., Chen, S. C., & Nafei, A. (2023). Bitcoin volatility forecasting: An artificial differential equation neural network. AIMS Mathematics, 8(6), 13907–13922. https://doi.org/10.3934/math.2023712