Forecasting Bitcoin Futures:
A Hybrid ARIMA, LSTM, and GARCH Approach
April 19, 2024

Jasmine Chen

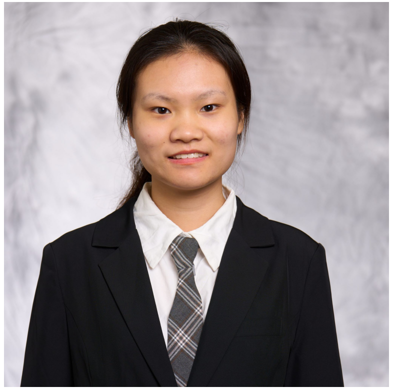
Ming-Hung Chen

Xinqian Li

Prithwish Maiti

Team Members






Jasmine Chen
pchen27@ncsu.edu

Ming-Hung Chen
mchen37@ncsu.edu

Xinqian Li
xli253@ncsu.edu

Prithwish Maiti
psmaiti@ncsu.edu

Outline

- Motivation
- Data Introduction
- Methodology
- Forecasting Models Introduction
- Conclusion
- Future Works

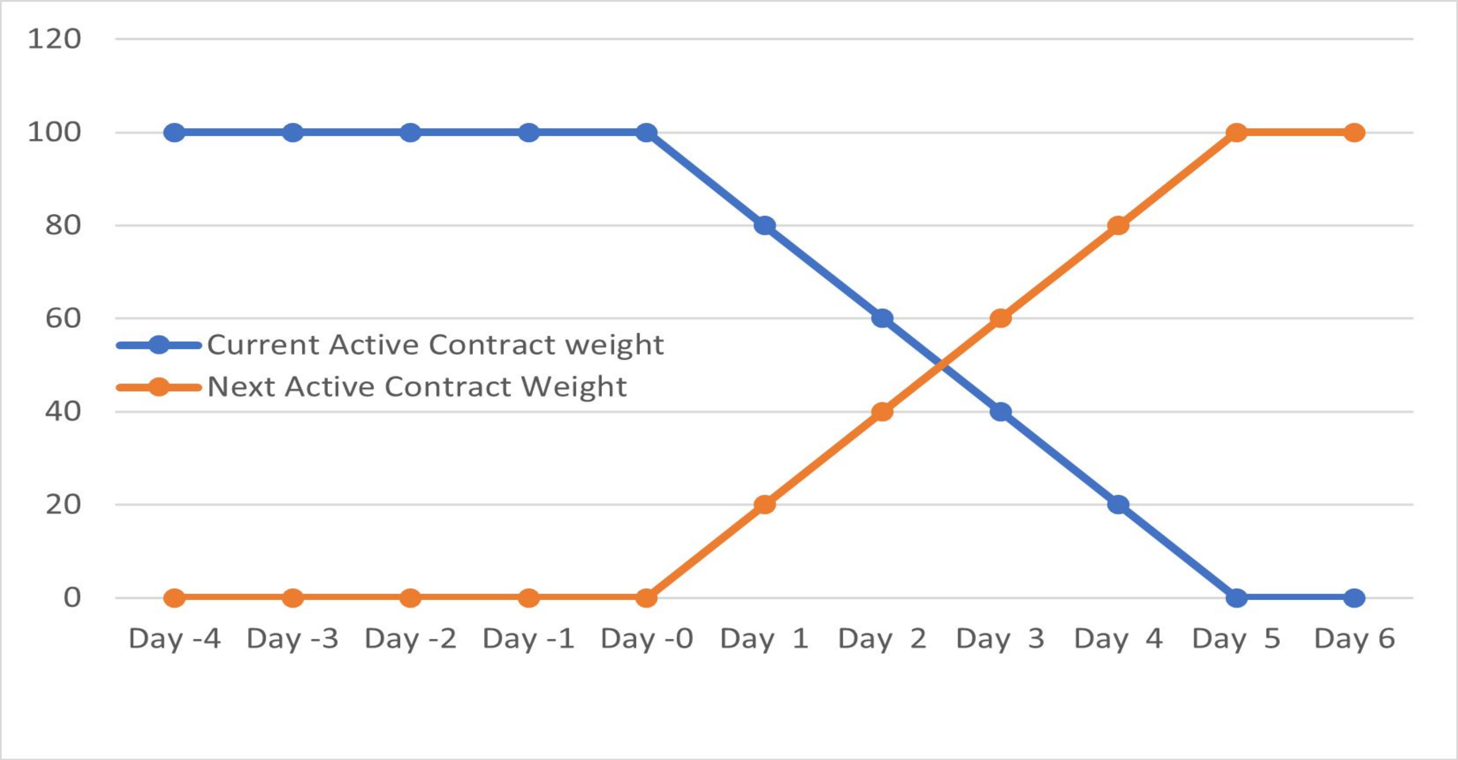
Motivation

- Futures provides price discovery and risk hedging capabilities in the cryptocurrency ecosystem.
- Crypto currencies are less correlated with or in fact negatively correlated with other commodities
- Example
  - Correlation with Treasury volatility: 0.1911
  - Correlation with Gold volatility: 0.2423

Data Introduction

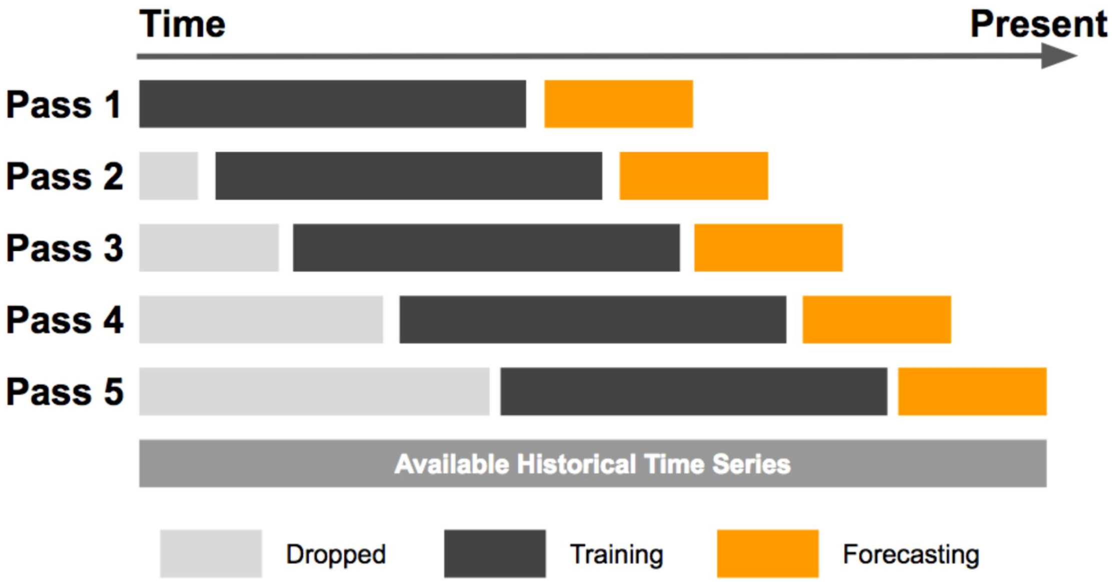
- Horizons Bitcoin Front Month Rolling Futures Index
- ER Bloomberg Ticker: HBITCNER
- The Index is a proprietary index owned and operated by Horizons ETFs Management Inc. ("Horizons ETFs").
- The index calculation agent is Solactive AG.

- 10- day Transition Phase of the Horizons Front Month Rolling Index



Methodology

Rolling Window Analysis & Forecasting



ARIMA

Autoregressive Integrated Moving Average

$$\underbrace{(1 - \phi_1 L_{t-1} - \phi_2 L_{t-2} - \cdots - \phi_p L_{t-p})}_{\mathrm{AR}(p)} \underbrace{(1 - L)^d}_{d \text{ differences}} Y_t = \underbrace{(1 + \theta_1 L_{t-1} + \theta_2 L_{t-2} + \cdots + \theta_q L_{t-q})}_{\mathrm{MA}(q)} \varepsilon_t$$

L = lag operator i.e. $L_{t-1} Y_t = Y_{t-1}$

p = order of the autoregressive part;

d = degree of first differencing involved;

q = order of the moving average part.

ARIMA(1,0,1)

$$Y_t = \phi_1 Y_{t-1} + \theta_1 \varepsilon_{t-1} + \varepsilon_t$$

p = 1 = order of the autoregressive part;

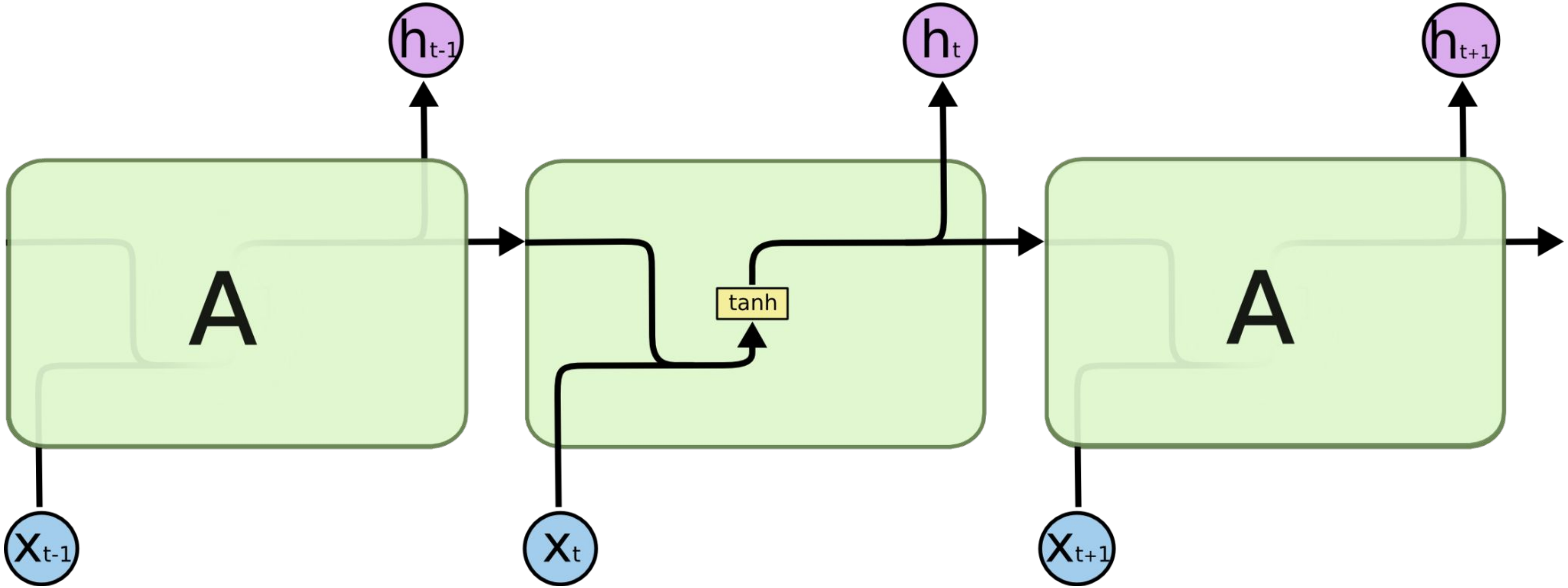
d = 0 = degree of first differencing involved;

q = 1 = order of the moving average part.

LSTM

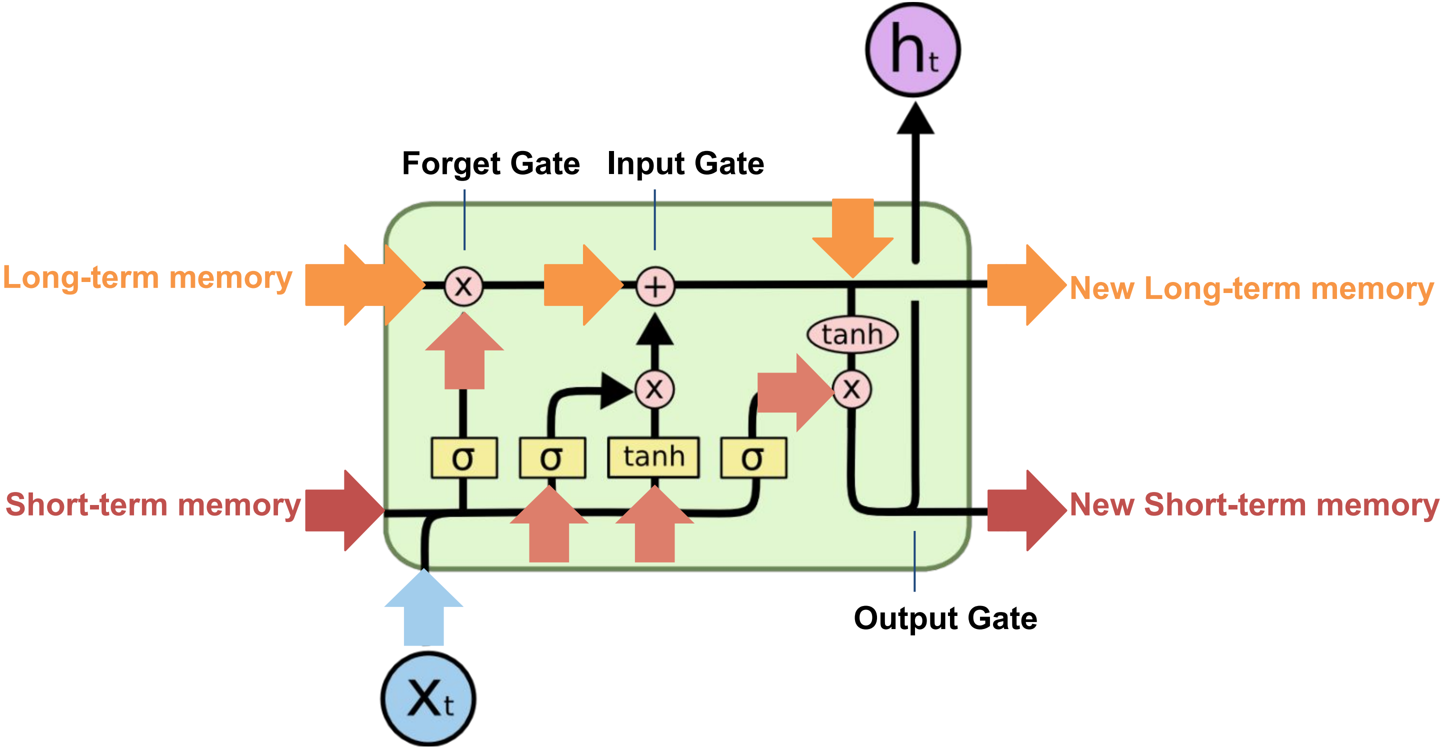
Long-Short Term Memory

- Standard Recurrent Neural Network

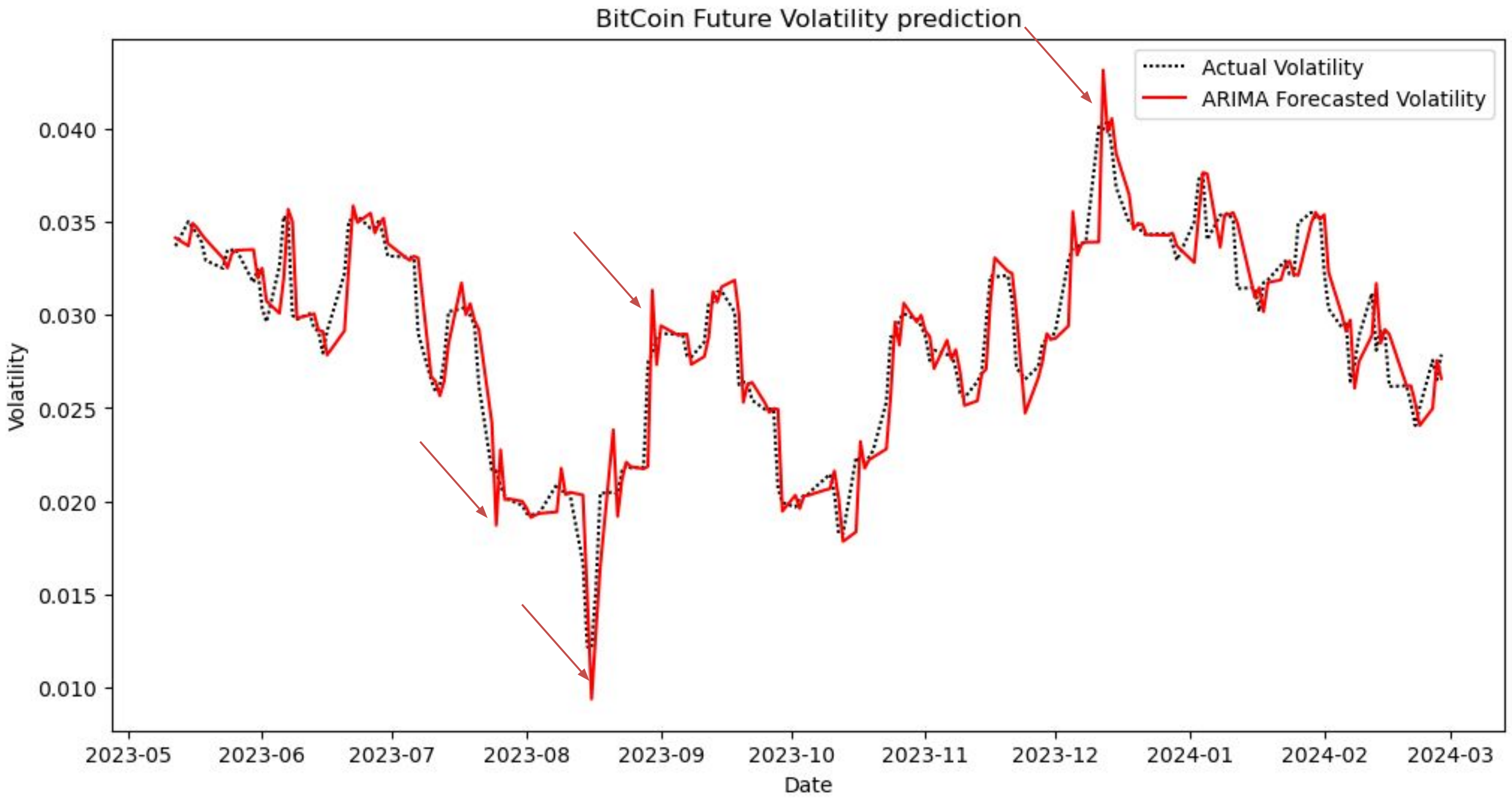


LSTM

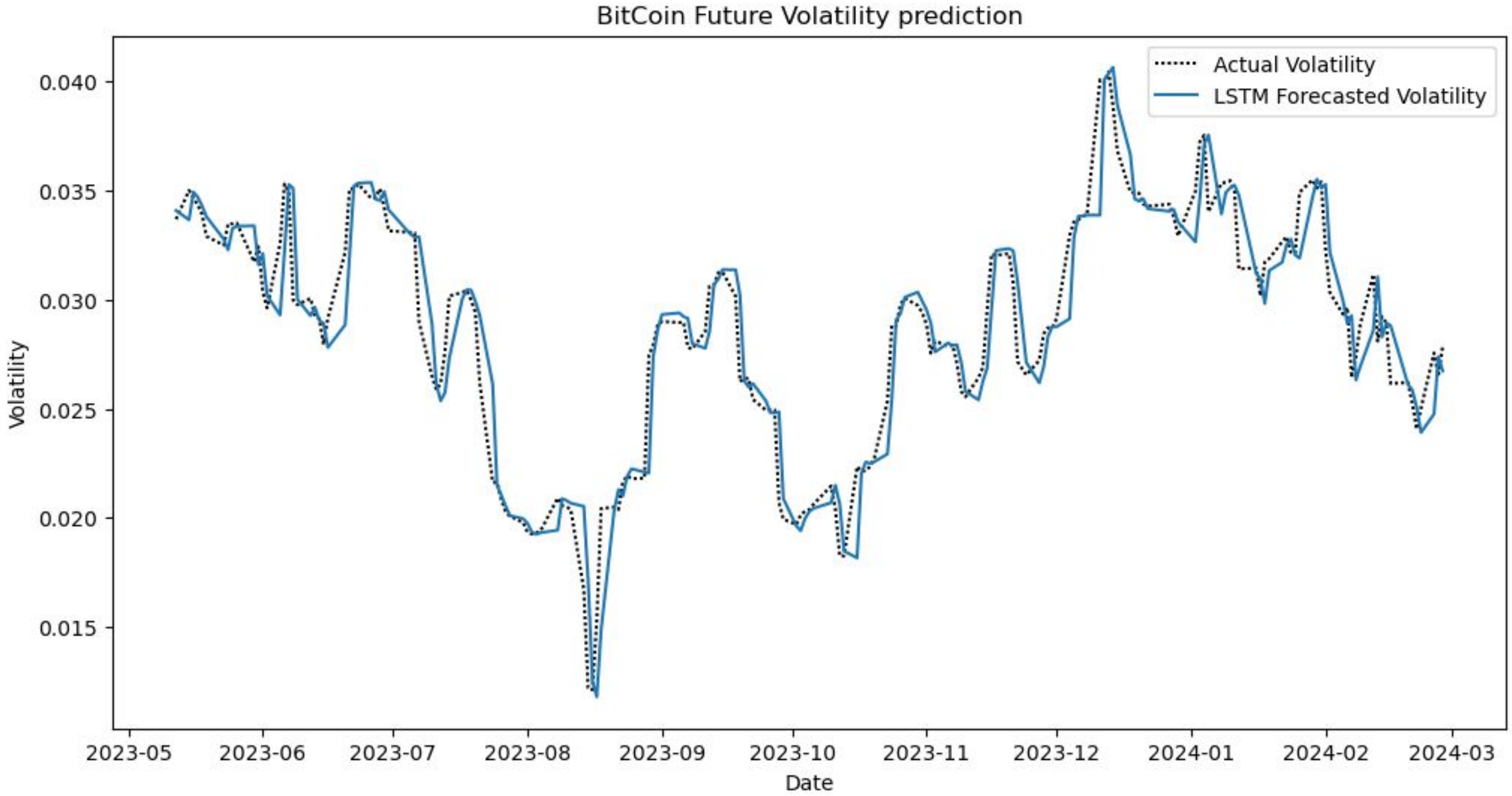
Long-Short Term Memory



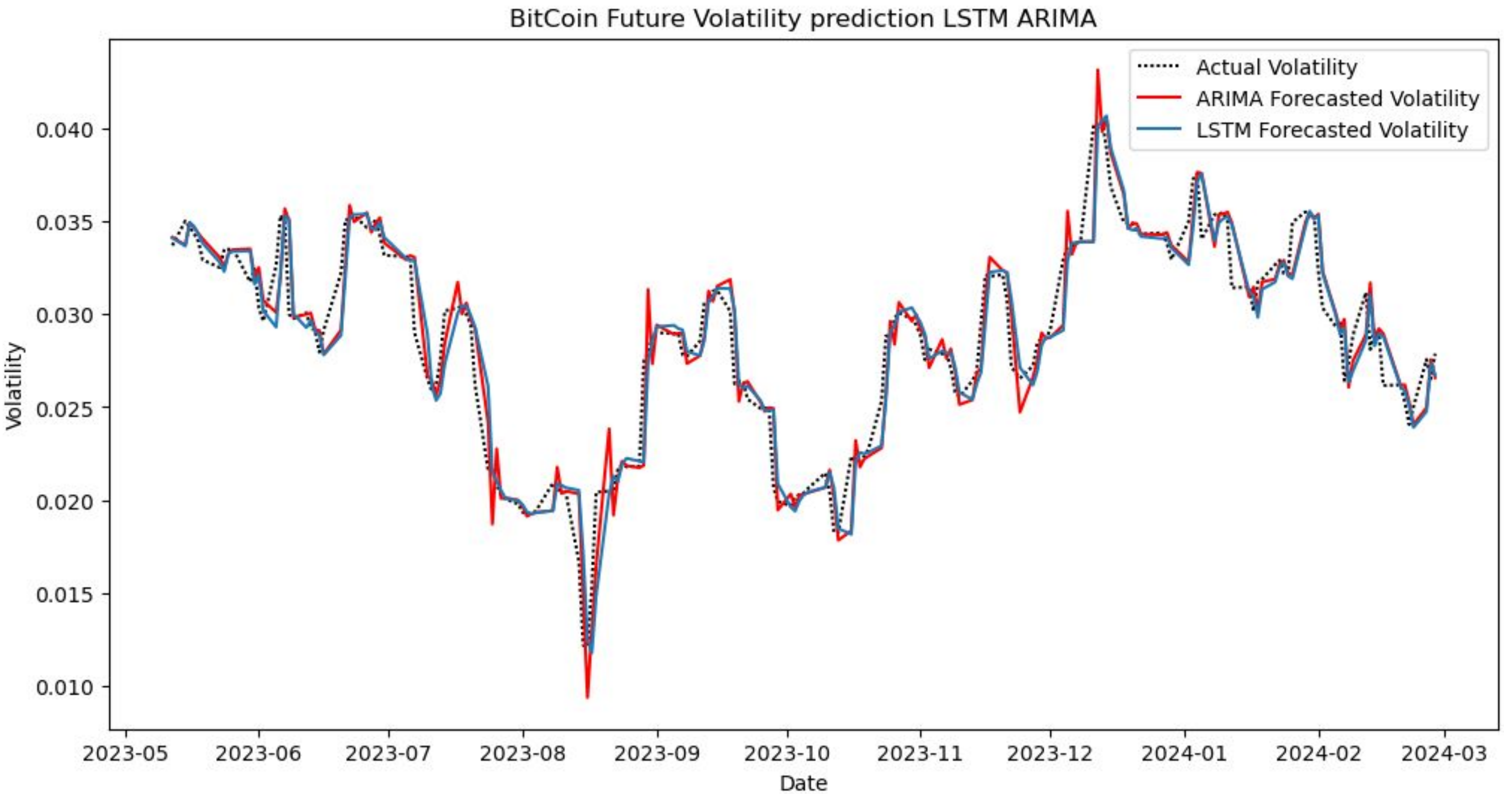
ARIMA Performance



LSTM Performance



ARIMA, LSTM Performance



ARIMA, LSTM Performance

	ARIMA	LSTM
MSE	0.00000283	0.00000394
RMSE	0.0016	0.0019

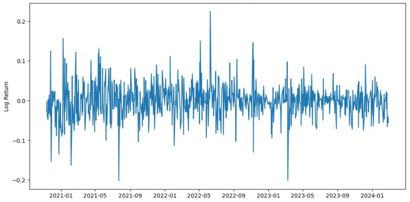
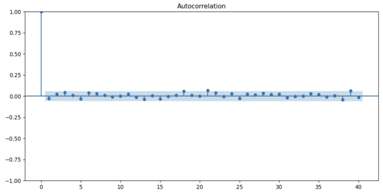
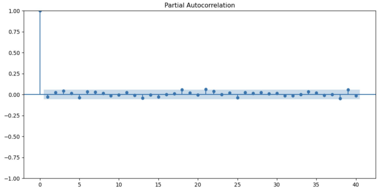
Data Description

Mean	0.00165	**Skewness**	-0.32
Maximum	0.2015	**Kurtosis**	4.45
Minimum	-0.2677	**Standard Deviation**	0.04204

Time Series Analysis

Test	Test Statistics	Result
JB-test	Statistic = 338.12 p-value = 0.00	Different from normal distribution
ADF test	Statistic = -29.18 p-value = 0.00	Stationary
KPSS test	Statistic = 0.24 p-value = 0.10	Stationary

Time Series Analysis

Test	Plot	Result
Log return time series plot		No significant trend
ACF plot		Autocorrelation exists
PACF plot		Partial autocorrelation exists

GARCH (1,1)

Generalized Autoregressive Conditional Heteroskedasticity

The GARCH (1,1) model can be expressed as:

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

where:

σ_t^2 is the forecasted variance for time t

ϵ_{t-1}^2 is the shock (return minus expected return) at time $t-1$

α_0, α_1 and β_1 represent constant variance, reaction to past shocks, persistence of past variances

EGARCH (1,1)

Exponential Generalized Autoregressive Conditional Heteroskedasticity

The EGARCH (1,1) model can be expressed as:

$$\log(\sigma_t^2) = \alpha_0 + \alpha_1 \left(\theta Z_{t-1} + \gamma\left(|Z_{t-1}| - E[|Z_{t-1}|]\right)\right) + \beta_1 \log(\sigma_{t-1}^2)$$

where:

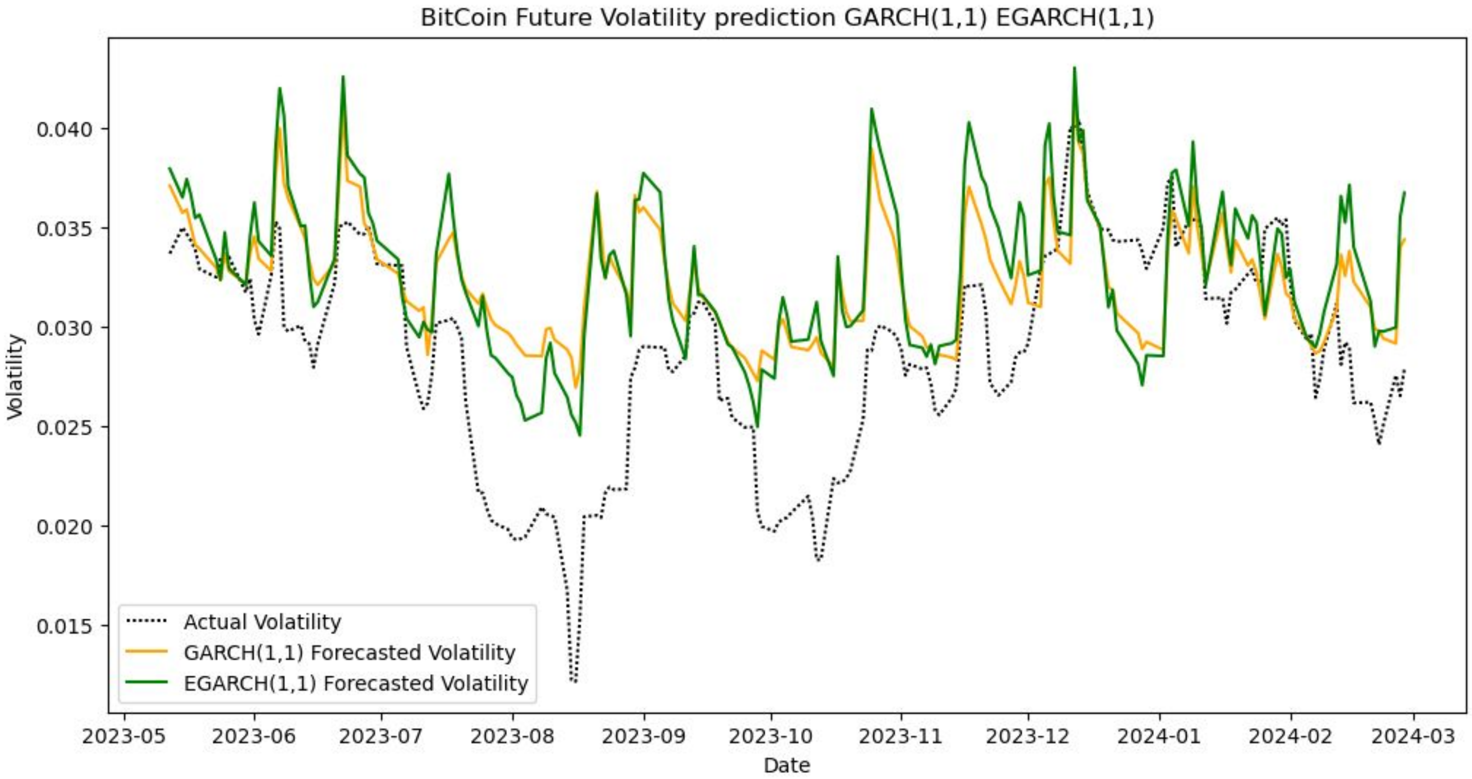
θ represents the effect of past shocks

γ represents the asymmetric impact of shocks on volatility

$Z_{t-1} = \frac{\epsilon_{t-1}}{\sigma_{t-1}}$ is the standard residual from the previous time period,

with ϵ_{t-1} being the error at time $t - 1$

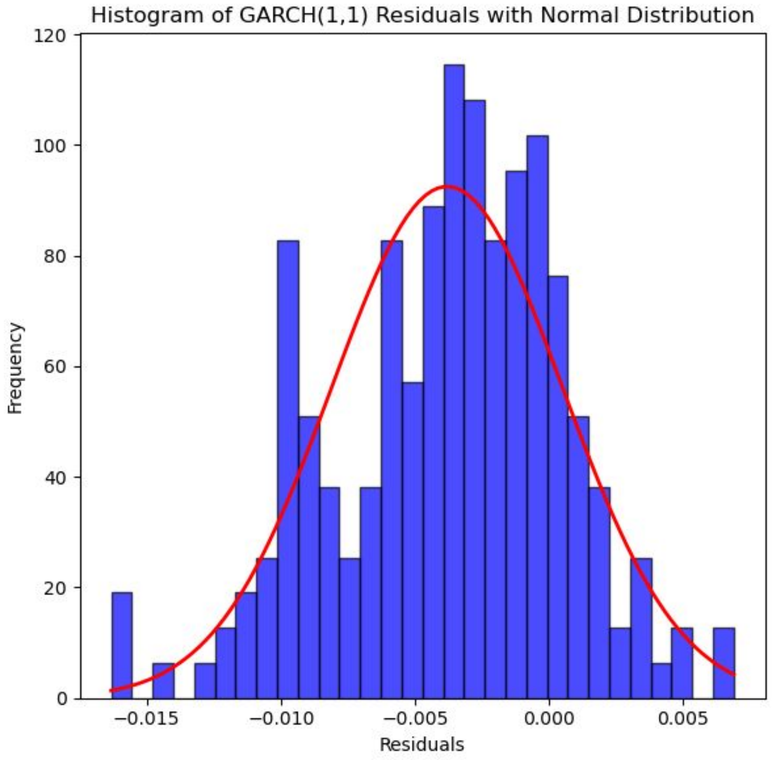
GARCH (1,1), EGARCH (1,1) Performance

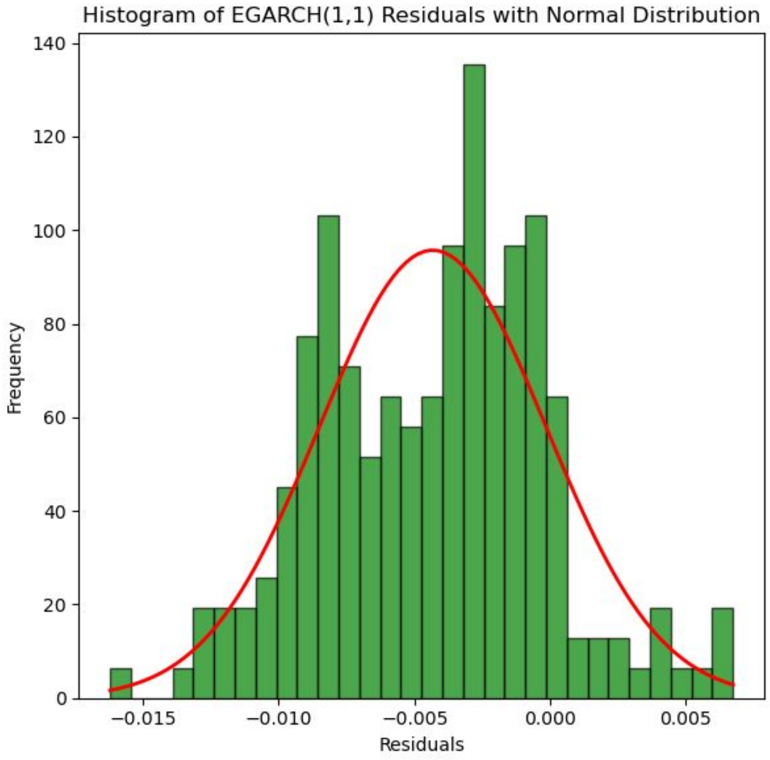


GARCH (1,1), EGARCH (1,1) Performance

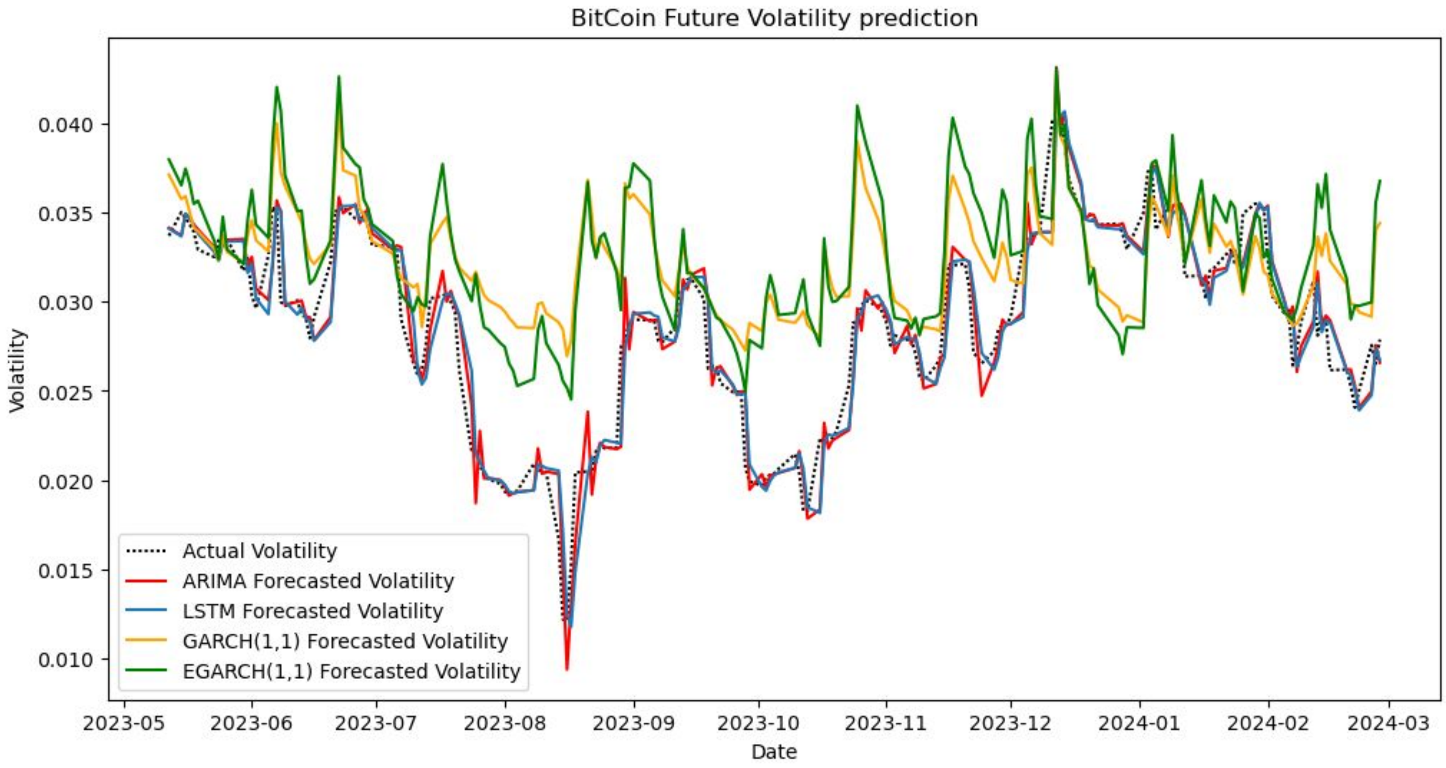
	GARCH (1,1)	EGARCH (1,1)
MSE	0.00009585	0.00010437
RMSE	0.0098	0.0102

GARCH (1,1), EGARCH (1,1) Performance





Performance Comparison



Performance Comparison

	ARIMA	LSTM	GARCH (1,1)	EGARCH (1,1)
MSE	0.00000283	0.00000394	0.00009585	0.00010437
RMSE	0.0016	0.0019	0.0098	0.0102

Future Works

- Analyze what leads to the performance difference
- Try other methods of forecasting: stochastic models, risk metrics models
- Explore more external features: macroeconomic indicators, sentiment analysis from social media

Thank You!
Q&A