Machine learning approaches to forecasting cryptocurrency volatility

International Review of Financial Analysis 90 (2023) 102914

Contents lists available at ScienceDirect

International Review of Financial Analysis

journal homepage: www.elsevier.com/locate/irfa

Machine learning approaches to forecasting cryptocurrency volatility: Considering internal and external determinants

Yijun Wang *, Galina Andreeva, Belen Martin-Barragan

Business School, University of Edinburgh, 29 Buccleuch Place, Edinburgh, EH8 9JS, United Kingdom

ARTICLE INFO

Keywords:
Time-series forecasting
Cryptocurrency volatility forecasting
Machine learning techniques
Deep learning techniques
Determinants

ABSTRACT

Given the volatile nature of cryptocurrencies, accurately forecasting cryptocurrency volatility and understanding its determinants are crucial. This paper applies machine learning (ML) techniques to forecast cryptocurrency volatility using internal determinants (e.g., lagged volatility, previous trading information) and external determinants (e.g., technology, financial, and policy uncertainty factors). Both Random Forest and Long Short-Term Memory (LSTM) networks significantly outperform traditional volatility models such as GARCH. Furthermore, we explore two optimization models—Genetic Algorithm and Artificial Bee Colony—to tune the hyper-parameters of LSTM. Our results indicate that the application of these optimization models substantially improves forecasting performance. Moreover, using SHapley Additive exPlanations, an interpretation method, we find that internal determinants play the most important roles in volatility forecasts. Finally, our results show that models trained with determinants from multiple cryptocurrencies outperform those trained with determinants from a single cryptocurrency, suggesting that considering a broader range of determinants can capture the complex dynamics in the cryptocurrency market.

1. Introduction

Blockchain technology, hailed as a transformative innovation, is revolutionizing the financial industry. A significant by-product of this technology is cryptocurrency, which is rapidly gaining acceptance among consumers, businesses, and even governments. Major FinTech companies such as Revolut and PayPal have embraced this trend by facilitating access to the cryptocurrency market. In December 2021, Visa, a global leader in digital payments, further substantiated this shift by launching the Crypto Advisory. The number of transactions continues to grow annually, with nearly 20,000 cryptocurrencies currently in circulation and a cumulative market value of approximately US $2 trillion (Urquhart & Lucey, 2022). Bitcoin, the largest cryptocurrency, leads the market in capitalization, reaching a peak value of $1.27 trillion in November 2021, according to Coinmarketcap.[1] However, as the cryptocurrency market is decentralized and lacks governmental backing, it faces the risk of high volatility as well as pricing bubbles (Corbet, Lucey, & Yarovaya, 2018). Therefore, accurately forecasting cryptocurrency volatility becomes crucial.

Despite its importance, there has been relatively little focus in the literature on forecasting cryptocurrency volatility. In this paper, we aim to address this gap in the literature by using machine learning (ML) techniques to achieve state-of-art time-series forecasting. We consider both internal determinants (e.g., lagged volatility and trading information) and external determinants (e.g., technology, financial and policy uncertainty) as forecasting variables to forecast cryptocurrency volatility.

Using ML techniques, including Random Forests (RF) and Long short-term memory (LSTM), we provide forecasts of daily, weekly, and monthly return volatility. We use Root Mean Squared Error, Mean Absolute Percentage Error, Normalized Mean Squared Error, and Directional Accuracy to evaluate the performance of the forecasts. We find our ML methods are better than the traditional volatility model, namely Generalized Autoregressive Conditional Heteroskedasticity (GARCH), to forecast volatility. More importantly, we explore two optimization models—Genetic Algorithm and Artificial Bee Colony—to tune the hyper-parameters of LSTM. Our results indicate that the application of these optimization models substantially improves forecasting performance.

To dissect the relative importance of forecasting variables in driving cryptocurrency volatility, we adopt the ML interpretation method, namely SHapley Additive exPlanations (SHAP). We find that the most influential drivers are lagged volatility and moving average volatility.

* Corresponding author.
E-mail address: yijun.wang@ed.ac.uk (Y. Wang).

[1] Coinmarketcap, a cryptocurrency industry utility that reports recently-traded prices for hundreds of cryptocurrencies, https://coinmarketcap.com/currencies/bitcoin/.

https://doi.org/10.1016/j.irfa.2023.102914
Received 30 December 2022; Received in revised form 18 August 2023; Accepted 6 September 2023
Available online 14 September 2023

Among external drivers, technology factors (including Google search volumes and certain blockchain factors) and financial factors (including the adjusted close price of NASDAQ and S&P 500) are also influential determinants.

Furthermore, we compare the universal model, trained on multiple cryptocurrencies, with the cryptocurrency-specific model.[2] This comparison is crucial in determining the appropriate predictors and the optimal data type (single versus multiple cryptocurrencies) to generate the most accurate forecast. We find that the universal model outperforms the cryptocurrency-specific model, suggesting that considering a broader range of determinants can capture the complex dynamics in the cryptocurrency markets.

Our paper relates to the literature that applies time-series forecasting to study financial assets. While ML models have been widely used in forecasting stock returns (Chung & Shin, 2018; Siami-Namini & Namin, 2018; Yun, Yoon, & Won, 2021), bond returns (Bianchi, Büchner, & Tamoni, 2021), and cryptocurrency exchange rates (Chen, Xu, Jia, & Gao, 2021; Gradojevic, Kukolj, Adcock, & Djakovic, 2023), these methods are rarely used in forecasting cryptocurrency volatility. Moreover, academics and practitioners have been increasingly attempting to interpret these "black-box" models. Post-hoc interpretation methods such as SHAP have been used to understand the stock price direction forecasting process (Yun et al., 2021). We contribute to this literature in three ways.

First, to the best of our knowledge, we are the first paper to use both internal and external factors to forecast the return volatility of cryptocurrency. We also compare the predictability of multi-step-ahead cryptocurrency volatility forecasts. ML techniques applied in this study provide statistically optimal forecasts of cryptocurrency volatility, which makes valuable implications for managing cryptocurrencies.

Second, we apply SHAP to examine the relative importance of the internal and external determinants in forecasting cryptocurrency volatility. Understanding how determinants drive cryptocurrency volatility is crucial as it could give us a deeper understanding of the cryptocurrency market and its relations with other financial markets.

Finally, our state-of-the-art forecasts of cryptocurrency volatility have important implications for a diverse range of stakeholders, including investors, financial institutions, policymakers, and academia. For investors, volatility is an essential criterion for determining asset allocation. Our volatility forecasts support more effective risk management and strategic asset allocation by hedging volatility, thereby optimizing portfolio performance. Additionally, financial institutions can leverage insights about blockchain technology and the relations among different crypto-assets to develop innovative financial products, such as providing a fair price for derivatives anchored to cryptocurrency. Moreover, policymakers can use volatility forecasts to formulate strategies to prevent market bubbles and reduce systemic risk. It can also contribute to the future regulation of cryptocurrency derivatives. From an academic perspective, the expanded application and framework of ML techniques foster additional research in financial time-series forecasting. Further research can build on our results to improve the ML algorithms for better predictive accuracy, to investigate a wide range of possible determinants, or to enhance ML interpretation.

The rest of this paper is structured as follows. Section 2 provides a comprehensive review of related literature, focusing on the application of ML in time-series forecasting and the deployment of optimization models and ML interpretation model. Section 3 introduces our Three-Stage framework of volatility forecasting. Section 4 provides details about the dataset and the determinants utilized for forecasting. Section 5 outlines the specifics of time-series forecasting models and discusses two LSTM hyper-parameter optimization algorithms. Section 6 describes the model evaluation and interpretation methods. Section 7 presents the empirical results of our model's forecasting performance and interpretation. Finally, Section 8 concludes the paper.

[2] See Section 3 for the details of the universal model.

2. Related literature

This section reviews papers that are related to our study. Section 2.1 reviews studies that apply ML techniques to time-series forecasting in the financial domain. Section 2.2 reviews the literature on LSTM hyper-parameter optimization models. Section 2.3 discusses the studies that use SHAP to interpret the forecasting process. Section 2.4 identifies the gaps in the existing literature.

2.1. ML time-series forecasting applications

The application of ML and deep learning techniques in the financial field has become popular in recent years, e.g., investment prediction, risk management, and algorithmic trading (van Binsbergen, Han, & Lopez-Lira, 2022; Nti, Adekoya, & Weyori, 2020; Ozbayoglu, Gudelek, & Sezer, 2020). The demonstrated success of these techniques in predicting stock market and commodity futures market performance implies a promising potential for forecasting cryptocurrency volatility. Table 1 encapsulates existing studies that deployed ML techniques for financial forecasting, with a specific focus on cryptocurrency literature presented in the lower part of the Table. Among the several ML techniques that have been used for forecasting tasks, LSTM, a deep learning technique, has risen to prominence because of its excellent prediction performance from its learning ability from massive data and the capability of handling the vanishing gradient problem of long-term time-series data (Chung & Shin, 2018; Fischer & Krauss, 2018; Siami-Namini & Namin, 2018). For instance, Sirignano and Cont (2019) employ LSTM to forecast the direction of stock price movements, comparing asset-specific models trained on specific stocks to universal models trained on all stocks. The results show that the universal model outperforms the asset-specific model, further demonstrating that LSTM offers greater accuracy than linear models in prediction tasks. This prowess of LSTM in capturing complex nonlinear patterns between high-dimension features and outputs has been further substantiated by various studies (Bengio, Goodfellow, & Courville, 2017). Moreover, Lim and Zohren (2021) highlight that deep learning techniques have also gained popularity for time-series forecasting in climate modelling, biological sciences, medicine, and retail areas. Motivated by these insights, we use LSTM to forecast cryptocurrency volatility and compare the universal model with cryptocurrency-specific models.

Up to now, the applications of ML algorithms to cryptocurrency time series analysis have concentrated mainly on predicting cryptocurrency prices and returns and mostly on Bitcoin, as summarized in the lower part of Table 1. For instance, LSTM has been utilized to predict Bitcoin exchange rate (Chen et al., 2021), Bitcoin price (Aggarwal et al., 2019), and Bitcoin price changes (Chen et al., 2020), considering various external factors such as economic and technological elements. Moreover, Random Forest (RF) has been used for forecasting Bitcoin returns, showcasing its prowess in accurate Bitcoin prediction (Gradojevic et al., 2023). Furthermore, Peng et al. (2018) combine the traditional GARCH model with ML approaches (XGboost and SVR) for volatility estimation. As far as we are aware, none of these works has considered using ML techniques and exploring different factors in the cryptocurrency volatility forecasting domain.

Existing research on cryptocurrency volatility relies mainly on economic models such as GARCH (Trucíos, 2019), heterogeneous autoregressive (HAR) models (Shen, Urquhart, & Wang, 2020), and time-varying parameter (TVP) regression (Bianchi, Guidolin, & Pedio, 2022). For example, Bianchi et al. (2022) use TVP regression to predict the weekly returns of cryptocurrencies considering market characteristics, stock market predictors, and sentiment variables. The results show that cryptocurrency is a new asset class, and the returns are less predictable, which is quite different from the traditional asset classes. This study motivates us to explore whether ML techniques could enhance predictability in the cryptocurrency market. Furthermore, Trucíos (2019) compares GARCH-type models for Bitcoin's daily volatility forecasting,

Table 1
Summary of studies using ML techniques in the financial forecasting domain.

Authors (year)	Target variable	Dataset	Input features	Forecasting models	Evaluation metrics
Chung and Shin (2018)	Stock close price	KOSPI from Bloomberg	5 historical values 5 technical indicators	GA-LSTM	MSE MAE MAPE
Siami-Namini and Namin (2018)	Stock adjust close price	Monthly financial time series from Yahoo Finance	historical values	LSTM ARIMA	RMSE
Yun et al. (2021)	Stock movement	KOSPI from Yahoo Finance	5 historical values 7 technical indicators	GA-XGBoost LSTM	Accuracy, F1, Precision, Recall, AUC
Fischer and Krauss (2018)	Stock movement	S&P 500 index constituents from Thomson Reuters	240 day return sequences	LSTM RF NN logistic regression	Mean return Standard deviation Annualized Sharpe ratio Accuracy
Sirignano and Cont (2019)	Stock movement	NASDAQ Level III data from LOBSTER data engine	1000 stocks of order book (supply & demand)	LSTM	Accuracy
Bianchi et al. (2021)	Bond return	U.S Treasury price, macroeconomic dataset	yields-only variables, 128 monthly macroeconomic and financial variables	extreme trees NN	R-square, MSPE
Chen et al. (2021)	Bitcoin exchange rate	Bitcoin exchange rate Economic factors and technology factors	1 historical values, factors	ARIMA, LSTM, SVR, ANFIS	RMSE, MAE, MAPE, DA
Gradojevic et al. (2023)	Bitcoin returns	Hourly Bitcoin exchange rate	20 technical indicators, global spot Bitcoin volume, Google search trend index	ANN, SVM, RF	MSPE, Sign statistic ex-post
Peng, Albuquerque, de Sá, Padula, and Montenegro (2018)	Cryptocurrency volatility, Currency volatility	Bitcoin, Ethereum, Dash	historical values	GARCH family models, SVR	RMSE, MAE
Chen, Li, and Sun (2020)	Change of Bitcoin price	Bitcoin daily price and high-frequency price	Technology factors, Investment and media attention, Gold spot price	XGboost, QDA, SVM, LSTM, Logistic Regression, LDA	Accuracy, Precision, Recall, F1
Aggarwal, Gupta, Garg, and Goel (2019)	Bitcoin price	Bitcoin	historical values, Gold spot price, Twitter sentiment	CNN, LSTM, GRU	RMSE

[a] The lower part of the Table is devoted to crypto literature.
[b] Table B.11 presents the summary of acronyms.

and Shen et al. (2020) use HAR models for Bitcoin volatility forecasting. Their results emphasize the significance of jumps, outliers, and structural breaks in volatility forecasting. They motivate our investigation into the predictive performance of daily, weekly, and monthly volatility and the impact of outliers in forecasting using ML techniques.

Regarding forecasting variables, Catania and Grassi (2022) develop a dynamic model for 606 cryptocurrencies accounting for the long memory, asymmetries, and time-varying skewness and kurtosis in the volatility process. The results show that including time-varying skewness improves the forecasts of volatility. This motivates us to consider internal factors (e.g., lagged volatility) in our forecasting process. Moreover, Conrad, Custovic, and Ghysels (2018) use the GARCH-MIxed Data Sampling (MIDAS) model to extract long- and short-term volatility components of cryptocurrencies and investigate their relationship with the financial market and macroeconomic activity. Their findings demonstrate a strong connection between Bitcoin volatility and global economic activity. Additionally, Jalan, Matkovskyy, Urquhart, and Yarovaya (2022) report a significant positive impact of trust on interest in and adoption of cryptocurrencies. Motivated by these findings, we include various external factors, such as financial and uncertainty factors, in our volatility forecasting.

2.2. LSTM hyper-parameter optimization models

LSTM is a complex model whose performance largely depends on the hyper-parameter settings. Identifying the optimal hyper-parameter set for LSTM is both time and computation-intensive; most current research defaults to subjective approaches based on researchers' experience to determine hyper-parameters. However, systematic approaches and optimization models, like the genetic algorithm (GA) (Bouktif, Fiaz, Ouni, & Serhani, 2018; Chung & Shin, 2018; Li, Li, Li, & Li, 2020), and Artificial Bee Colony (ABC) (Kumar, Kumar, & Kumar, 2021; Yuliyono & Girsang, 2019), have been employed for hyper-parameter tuning. A detailed explanation of optimization methodology is provided in Section 5.2.1. This Section concentrates on the relevant literature.

GA is an intelligent algorithm that simulates the heredity and evolution of natural organisms to adapt to the environment. Given its straightforward algorithmic process, fewer hyper-parameters, quicker optimization speed, and superior results, GA finds a wide range of applications in image processing, function optimization, signal processing, and pattern recognition, among other fields. In finance, GA has been utilized to determine the custom time window and the number of LSTM units for forecasting the daily stock index (Chung & Shin, 2018). The optimization through GA has improved the learning process's efficiency and averted unnecessary computational processes. Similarly, GA-LSTM has superior performance over LSTM in cable joint temperature prediction (Li et al., 2020). There are two main reasons we chose GA for tuning LSTM hyper-parameters. Firstly, GA-LSTM has superior compatibility. Secondly, GA can find global optimal hyper-parameters instead of local optimal hyper-parameters with steps of selection, crossover, and variation (Bouktif et al., 2018).

ABC is a meta-heuristic method that emulates the foraging behaviour of bee colonies. It has been used to optimize hyper-parameters for Bitcoin price prediction (Yuliyono & Girsang, 2019). The results show that ABC-LSTM outperformed LSTM without optimization. Kumar et al. (2021) demonstrate the effectiveness of ABC-LSTM in stock market forecasting, where the ABC algorithm effectively maintains the equilibrium of exploitation and exploration issues.

2.3. Applications of ML interpretation

Given the black-box nature of LSTM, it needs more interpretation regarding its final model results. Therefore, post-hoc interpretation models that deploy simpler, interpretable surrogate models between features and outputs have been employed. These models interpret the trained networks and identify feature attributions based on the surrogate model (Lim & Zohren, 2021). For instance, SHAP utilizes Shapley values from cooperative game theory to discern significant features. More details about this are presented in Section 6.2. Yun et al. (2021) use SHAP to interpret the results of both XGBoost and LSTM models and highlight the critical features of stock prediction. Moreover, Babaei, Giudici, and Raffinetti (2022) propose an explainable portfolio management approach for cryptocurrencies using SHAP. This approach explains portfolio weights based on ML models combined with dynamic Markowitz portfolio optimization. Another study by Fior, Cagliero, and Garza (2022) employs SHAP to interpret the cryptocurrency price prediction of XGBoost, considering blockchain-related, market, and technical features.

2.4. Summary of gaps

In conclusion, to the best of our knowledge, although several studies have addressed time-series forecasting in financial markets, several critical gaps exist in cryptocurrency volatility forecasting. Specifically, no research to date has compared the performance of the traditional volatility model with the most promising ML and deep learning models. Regarding predictors for the Bitcoin exchange rate, Chen et al. (2021) and Gradojevic et al. (2023) have considered various determinants such as blockchain information, macroeconomic, financial, and technical indicators. However, no study has specifically examined the importance of key internal determinants, like lagged volatility, and external determinants, like blockchain information, in predicting cryptocurrency volatility. Moreover, while (Sirignano & Cont, 2019) have demonstrated that universal models outperform asset-specific models in stock movement prediction, no research has drawn this comparison in the cryptocurrency market. Finally, using hyper-parameter optimization and ML interpretation models in cryptocurrency volatility forecasting remains largely unexplored.

3. Three-stage framework of volatility forecasting

This paper aims to conduct a comprehensive exploration of cryptocurrency volatility forecasting. We compare the traditional GARCH model and various ML methods and investigate the determinants that play important roles in the forecasts. When designing the methodology, we consider that unlike exchange rates or prices, volatility is not directly accessible and needs to be estimated. We also consider the need for interpretation as the oft-cited criticism of ML methods being black boxes. Therefore we organize our methodology in a Three-Stage experiment:

- **Stage I Data Processing:** We separately estimate the daily, weekly, and monthly volatility for the chosen cryptocurrencies and prepare the internal and external determinants. Details are given in Section 4.
- **Stage II Time-series Forecasting Model:** This forms the core stage of the comparison experiments. For model fitting, we apply both the traditional GARCH method and ML models, including RF and LSTM, for forecasting. We also compare forecasting accuracy using both internal and external determinants. The models are distinguished into two types:
  - Cryptocurrency-specific model is trained with determinants from **a single cryptocurrency**.
  - Universal model is trained with determinants from **multiple cryptocurrencies**.

  For model tuning, we use methods to find the optimal hyper-parameters of ML models. Details are given in Section 5.
- **Stage III Model Evaluation & Interpretation:** Post-forecasting, we use different evaluation metrics to assess model performance. We employ SHAP to interpret our best-performing models and identify significant determinants. Details are given in Section 6.

The workflow of our Three-Stage experiment is presented in Fig. 1.

4. Dataset

This paper examines the internal and external determinants used by ML approaches for forecasting cryptocurrency volatility, which is the predictive target variable in all models. The internal determinants fall into three categories: lagged volatility, moving average volatility, and previous trading information, detailed further in Section 4.2. We also have three categories of external determinants, namely technology, financial, and policy uncertainty factors, presented in Section 4.6. Table 2 lists the data collected for this study.

4.1. Cryptocurrency dataset

4.1.1. Cryptocurrency price data

To limit our analysis to the most popular and capitalized cryptocurrencies, we collected the daily open, high, low price, trading volume, trading count, and time of Bitcoin (BTC), Ethereum (ETH), Litecoin (LTC) and Ripple (XRP) from Coin-API.[3] The close price of cryptocurrencies is collected every 30 min in one day from November 1, 2017, to July 31, 2022.

4.1.2. Volatility estimation

The volatility analysis is based on cryptocurrency returns instead of the price evolution. The daily volatility index, VOL_1 is calculated as logarithmic percentage change taken from measurements taken at the spot close price every 30 min. The settlement price is calculated from 48 snaps over 24 h. The weekly volatility index, VOL_7, is calculated with the settlement price from 336 snaps over seven days. The monthly volatility index, VOL_30 is calculated with the settlement price from 1440 snaps over 30 days. The calculation equations refer to the BitMEX platform.[4] The calculation equation is presented as follows. P = Last Price (taken at 30 min intervals); STD = Sample Standard Deviation; Ln = Natural Logarithm; Sqrt = Square Root.

$$\mathrm{VOL\_1} = \mathrm{STD}\left(\mathrm{Ln}\left(\boldsymbol{P}_1/\boldsymbol{P}_0\right), \mathrm{Ln}\left(\boldsymbol{P}_2/\boldsymbol{P}_1\right), \dots, \mathrm{Ln}\left(\boldsymbol{P}_{48}/\boldsymbol{P}_{47}\right)\right) * \mathrm{Sqrt}(48) \tag{1}$$

$$\mathrm{VOL\_7} = \mathrm{STD}\left(\mathrm{Ln}\left(\boldsymbol{P}_1/\boldsymbol{P}_0\right), \mathrm{Ln}\left(\boldsymbol{P}_2/\boldsymbol{P}_1\right), \dots, \mathrm{Ln}\left(\boldsymbol{P}_{336}/\boldsymbol{P}_{335}\right)\right) * \mathrm{Sqrt}(336) \tag{2}$$

$$\mathrm{VOL\_30} = \mathrm{STD}\left(\mathrm{Ln}\left(\boldsymbol{P}_1/\boldsymbol{P}_0\right), \mathrm{Ln}\left(\boldsymbol{P}_2/\boldsymbol{P}_1\right), \dots, \mathrm{Ln}\left(\boldsymbol{P}_{1440}/\boldsymbol{P}_{1439}\right)\right) * \mathrm{Sqrt}(1440) \tag{3}$$

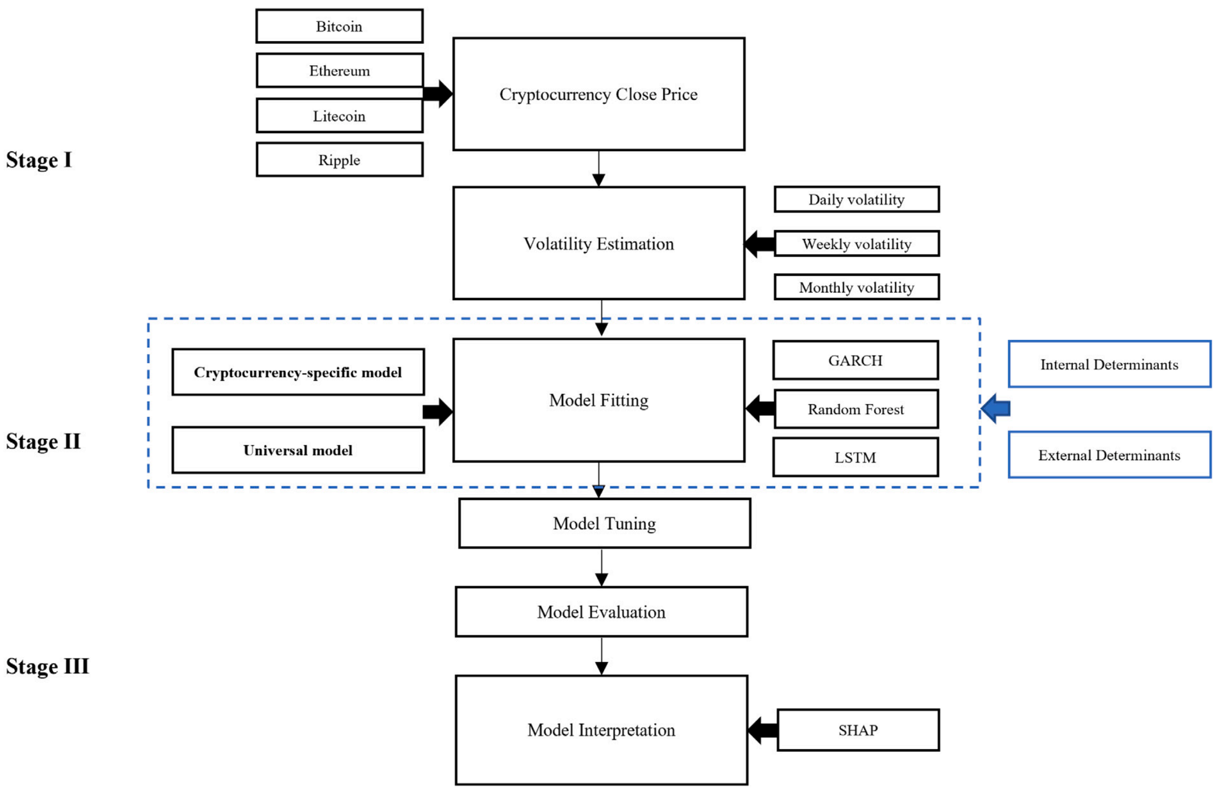


Fig. 1. A Three-Stage experiment workflow.

Table 2
List of data.

Data category		Data	Source	Time frequency
Cryptocurrency data		Close price	Coin-API	Every 30 min
		Open, High, Low price Trading volume, Trading count, Time	Coin-API	Daily
Technology data	Blockchain data	Average block size, Average transactions per block, Average payments per block, Average confirmation time, Hash rate, Difficulty	Blockchain.com	Mixed-frequency
	Google trend data	Cryptocurrency, Blockchain, Bitcoin	Google Trends Data-store	Daily
	Crypto sentiment data	Crypto Fear & Greed Index	Bitcoin fear	Daily
Financial data		Adjust close price and Trading volume for Oil, Gold, Silver, DJI, S&P 500, NASDAQ, Russell 2000	Yahoo YQL Finance API	Daily
		Exchange rate (Yuan-USD, USD-Euro)		Daily
Policy uncertainty data		US daily news index	Economic policy uncertainty	Daily

From the time-series plot presented in Fig. 2, we observe that the daily volatilities of Bitcoin and Ethereum generally lie between the zero line and 0.3. In contrast, the volatilities of Litecoin and Ripple are slightly higher, ranging between the zero line and 0.4. Interestingly, these four cryptocurrencies display their most significant value at around the exact dates, indicating potentially correlated behaviour, except for Ripple, which shows more substantial fluctuations between September 2020 and April 2021 during the Covid-19 pandemic. As we move to weekly and monthly volatility, the volatility generally increases, resulting in less frequent but more substantial fluctuations. This pattern can be particularly evident when comparing the daily with weekly and monthly volatility. The descriptive statistics for the volatility indices of these four cryptocurrencies are presented in Table 3. Bitcoin demonstrates the least average volatility, while Ripple displays the highest average volatility, peaking at a substantial 5.4%. Furthermore, Litecoin's daily volatility exhibits the most extreme behaviour, displaying a kurtosis coefficient of 20.05, which indicates a distribution with heavier tails and sharper peaks compared to a normal distribution. Conversely, the monthly volatilities show relatively less extreme behaviour, except Litecoin.

The Augmented Dickey–Fuller Test (ADF), a well-known unit root test, was employed to assess the stationarity of the volatility time series.

[3] CoinAPI, a platform that provides fast, reliable and unified data APIs to the cryptocurrency market, https://www.coinapi.io/.

[4] BitMEX, an advanced cryptocurrency exchange, and derivative trading platform, https://www.bitmex.com/.

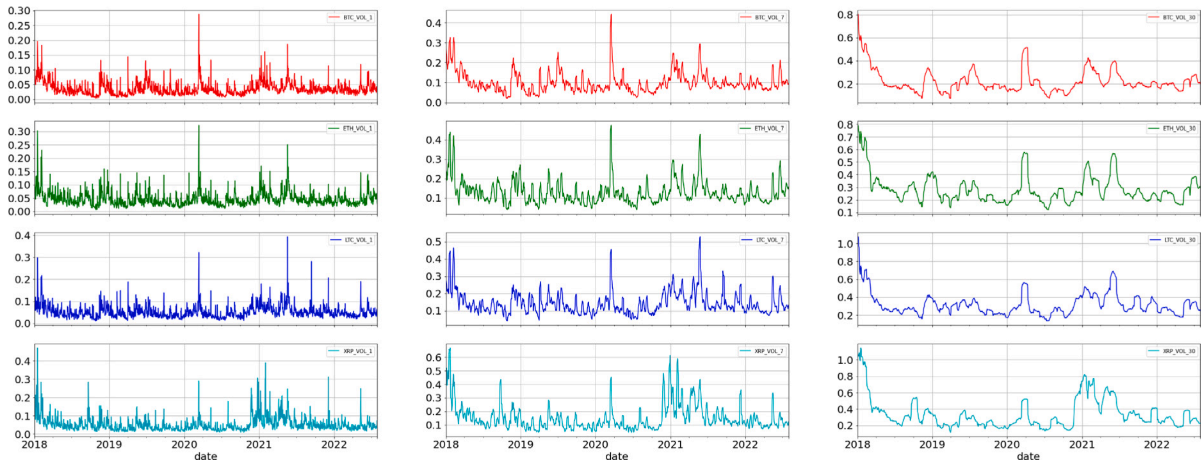


Fig. 2. Time-series of cryptocurrency volatility.

Table 3
Descriptive statistics and ADF test for cryptocurrency volatility.

	Count	Mean	Std	Min	25%	50%	75%	Max	Skewness	Kurtosis	ADF statistic	*P*-value
BTC_VOL_1	1673.0	0.036	0.024	0.003	0.021	0.030	0.044	0.288	2.738	14.710	−7.819	0.00000
BTC_VOL_7	1673.0	0.101	0.053	0.020	0.070	0.089	0.115	0.443	2.002	6.378	−5.469	0.00000
BTC_VOL_30	1673.0	0.222	0.103	0.074	0.160	0.195	0.248	0.805	1.966	5.253	−5.333	0.00000
ETH_VOL_1	1673.0	0.046	0.029	0.006	0.029	0.039	0.056	0.323	2.956	16.230	−10.421	0.00000
ETH_VOL_7	1673.0	0.130	0.062	0.039	0.090	0.115	0.154	0.474	2.089	6.533	−5.955	0.00000
ETH_VOL_30	1673.0	0.284	0.113	0.122	0.210	0.257	0.319	0.791	1.780	3.692	−5.230	0.00001
LTC_VOL_1	1673.0	0.052	0.031	0.011	0.033	0.045	0.063	0.393	3.172	20.050	−6.257	0.00000
LTC_VOL_7	1673.0	0.145	0.066	0.041	0.102	0.128	0.173	0.529	1.995	6.225	−5.263	0.00001
LTC_VOL_30	1673.0	0.318	0.126	0.135	0.243	0.281	0.364	1.077	2.086	6.777	−6.181	0.00000
XRP_VOL_1	1673.0	0.054	0.043	0.010	0.028	0.041	0.063	0.469	3.230	15.953	−6.884	0.00000
XRP_VOL_7	1673.0	0.153	0.099	0.045	0.093	0.123	0.177	0.667	2.156	5.405	−5.492	0.00000
XRP_VOL_30	1673.0	0.340	0.185	0.116	0.216	0.282	0.387	1.145	1.793	3.415	−4.294	0.00046

The null hypothesis of the ADF test posits that the time series is non-stationary, exhibiting a unit root, whereas the alternate hypothesis asserts that the time series is stationary and devoid of a unit root. As reported in Table 3, the results of the ADF test convincingly rejected the null hypothesis at a 1% significance level, affirming that all the time series are stationary and without unit roots.

Two additional diagnostic tools, the Autocorrelation Function (ACF) and the Partial Autocorrelation Function (PACF), are employed to analyse the internal dynamics of the series. The ACF plots (Fig. C.12(a)) exhibit the correlation coefficients between the current value and its lagged values, while the PACF plots (Fig. C.12(b)) depict the partial correlation between the series and its own lags after removing the effects of intermediate lags. All the ACF plots exhibit a tail-off pattern, where the correlation decays to zero slowly, implying a persistent influence of shocks on the time series. Moreover, the PACF plots indicate that the first four lags are significant for daily volatility, whereas the first five lagged volatilities are significant for Ripple. These observations underscore the need to incorporate past lags when modelling the dynamics of these cryptocurrencies' volatilities.

4.2. Summary of internal determinants

According to Liu and Tsyvinski (2021) and Liu, Tsyvinski, and Wu (2022), a significant time-series momentum phenomenon in the cryptocurrency market indicates that past historical data can provide predictive value for the future. Therefore, we incorporate lagged volatility and moving volatility averages as internal predictors. Given the unique nature of cryptocurrencies being traded 24 h a day, seven days a week, unlike traditional financial markets like stocks and bonds, which halt transactions overnight and on weekends, we select a seasonal lag of seven.

For daily volatility (VOL_1), we examine lagged volatilities of the past 28 days (lag_1, lag_2, …, lag_28) and moving average one lagged volatilities of different orders: fast (ma_3, ma_5, ma_7); slow (ma_21, ma_28, ma_35). For weekly volatility (VOL_7), we consider weekly lagged volatilities (lag_7, lag_8, …, lag_14) and moving average seven lagged volatilities of various orders (ma_3, ma_5, ma_7). For monthly volatility (VOL_30), we explore monthly lagged volatilities (lag_30, lag_31, …, lag_37) and moving average thirty lagged volatilities of different orders (ma_3, ma_5, ma_7). Additionally, we include several other determinants in our dataset: daily open price, close price, high price, low price, trading volume, and trading count, all with a lag of one day, seven days, and thirty days respectively, for daily, weekly and monthly volatilities. Table C.13 presents the descriptive statistics of several other internal determinants. Hence, the number of internal determinants for VOL_1, VOL_7, and VOL_30 are 40, 17, and 17, respectively. The correlation analysis results, presented in Table C.12, indicate that lagged volatility, moving average volatility, and trading volume are significantly correlated with the volatility of our chosen cryptocurrencies. Furthermore, compared to Bitcoin and Ethereum, the volatility of Litecoin and Ripple exhibits a stronger correlation with the previous day's trading count.

4.3. Technology dataset

We choose three types of technology predictors, namely Blockchain data, Google Trend data, and Crypto Sentiment data. The difference between cryptocurrencies and other financial products is that they rely on Blockchain technology to make transactions. Besides, as new technology is introduced, the social media attitude towards the cryptocurrency market could play an important role. According to Chen et al. (2021) and Liu and Tsyvinski (2021), these technological factors have the power to predict the cryptocurrency market. Blockchain data presents the adaptation of the cryptocurrency in the whole Blockchain system. Moreover, according to studies by Akyildirim, Corbet, Lucey, Sensoy, and Yarovaya (2020) and Smales (2022), an increase in investor attention correlates with higher volatility. Thus, we examine whether Google Trend data, which provides insights into search frequency trends for specific terms (like cryptocurrency names), and Crypto Sentiment data, which represents public sentiment towards cryptocurrencies, can assist in predicting the cryptocurrency market. Consequently, we have incorporated these three types of technological predictors in our study.

4.3.1. Blockchain data

The following Blockchain predictors are considered in our study, with data collected from Blockchain Charts from January 1, 2018, to July 31, 2022, at mixed frequency intervals.[5] We incorporate the Average Block Size, which refers to the average size of blocks in the Blockchain network measured in megabytes over the past 24 h. The Average Transactions Per Block indicates the average number of transactions per block over the past 24 h, while Average Payments Per Block describes the average number of payments per block during the same period. We also consider the Average Confirmation Time, which denotes the average time taken for a transaction (that includes miner fees) to be included in a mined block and added to the public ledger. The Hash Rate, another predictor used in this study, is the estimated number of terahashes per second the Bitcoin network performs in the last 24 h. Lastly, the Difficulty metric reflects the difficulty encountered in mining a new block for the Blockchain. These variables provide insightful aspects of the Blockchain network's operations.

4.3.2. Google trend data

Google Trend index presents the public attention to the keyword.[6] We collect the worldwide historical search volume considering "Blockchain", "cryptocurrency", and "Bitcoin" as our keywords. This index is from January 1, 2018, to July 31, 2022, accounting for 1673 observations.

4.3.3. Crypto sentiment data

Crypto Fear & Greed Index presents the emotions and sentiments towards the prominent cryptocurrencies from various sources such as social media and dominance.[7] There are five attitudes Extreme Greedy (market correction); Greedy; Neutral; Fear, and Extreme Fear (too worried and a buying opportunity). This index is collected from February 1, 2018, to July 31, 2022, accounting for 1685 observations with 185 Extreme Greedy; 304 Greedy; 150 Neural; 547 Fear, and 453 Extreme Fear.

4.4. Financial dataset

We choose two types of financial predictors, namely macroeconomic data and currency ratio obtained using Yahoo Finance's API using Python package *yahoo-finance*.[8] According to Chen et al. (2021), financial predictors effectively predict the Bitcoin exchange rate. Thus, this research considers the adjusted close price (Adj) and the trading volume (V) for Oil, Gold, Silver, Dow Jones Industrial Average (DJI), S&P 500, NASDAQ, and Russell 2000 as our macroeconomic predictors and exchange rate for Yuan-USD and USD-Euro as currency ratio predictors. According to Alexander and Dakos (2020), the trading volume of the cryptocurrency influenced the investors' decision, which suggested one could explore the trading volume of other financial markets that will affect cryptocurrency volatility.

4.5. Policy uncertainty dataset

Baker, Bloom, and Davis (2016) develop a new index of economic policy uncertainty (EPU) based on newspaper coverage frequency.[9] According to Cheng and Yen (2020) and Yen and Cheng (2021), EPU can predict Bitcoin monthly volatility and returns. We collect the daily news-based EPU from January 1, 2018, to July 31, 2022, accounting for 1673 observations.

4.6. Summary of external determinants

Table C.14 presents the descriptive statistics for external determinants. Fig. C.13 presents the normalization data plot for external determinants. Therefore, the final external determinants are 27.

4.7. Data prepossessing

Before Stage II Forecasting models, data processing is necessary, employing data cleaning, transformation, and division. The linear interpolation method is used to input the missing value in data cleaning and dealing with mixed-frequency data processing. We assume the data is known at two time points $r < t < s$. Then the equation of missing value $\hat{P}_{it}$ is presented below.

$$\hat{P}_{it} = \frac{(t-s)P_{r,t-1} + (r-t)P_{s,t-1}}{r-s} \tag{4}$$

The corresponding date of the missing value is removed. For data transformation, Min–Max normalization is adopted to scale input features between 0 and 1, which avoids the bias from outliers and preserves the relationships among features (Patro & Sahu, 2015). The target variable is not scaled using Min–Max normalization as its values predominantly fall between 0 and 1, except for the maximum monthly volatility for Litecoin and Ripple, which hover around 1.1. The equation is presented as follows:

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \tag{5}$$

For Stage II, we use 80% of the data from February 5, 2018, to August 10, 2021, as the training set. The remaining 20% of the data, from August 11, 2021, to July 31, 2022, is used as the test set. The test set allows us to perform out-of-sample forecasts for cryptocurrency volatility. For ML models, we employ a Time Series Split Cross-Validation technique with five splits to tune the hyper-parameters for the training set optimally. In this method, the validation set is the section after the training set, respecting the temporal order of the data. This approach prevents look-ahead bias, enhancing the generalizability and robustness of our forecasting models.

5. Time-series forecasting models

This section elaborates on the forecasting models used in Stage II of our experiment. We utilize several models, each with its distinct strengths and approaches. We use RF and LSTM to provide a comparison with data-driven techniques. RF, an ensemble learning method renowned for its robustness, has been widely used in time-series forecasts due to its superiority of handling high-dimensional and non-linear data (Alessandretti, ElBahrawy, Aiello, & Baronchelli, 2018; van Binsbergen et al., 2022). LSTM is one of the state-of-the-art models for time-series forecasting due to its ability to learn long-term dependencies. This characteristic has led to success in forecasting tasks such as Bitcoin price prediction (McNally, Roche, & Caton, 2018). To further improve the LSTM model's performance, we employ optimization algorithms to fine-tune the hyperparameters of LSTM. We also consider the GARCH model, which is widely used in the literature as a benchmark for forecasting. The details of the GARCH model are present in Appendix A.

[5] Blockchain Charts: https://www.blockchain.com/en/.

[6] Google Trends database: https://trends.google.com/trends.

[7] Crypto Fear & Greed Index: https://alternative.me/crypto/fear-and-greed-index/.

[8] Yahoo! Finance's API: https://pypi.org/project/yfinance/.

[9] Economic Policy Uncertainty Index: http://www.policyuncertainty.com/index.html.

Table 4
List of hyper-parameters of RF.

Hyper-parameter	Range value	Interval
Max depth of the tree	(4, 20)	2
Min samples splits	(2, 10)	2
The number of trees	(8, 128)	8

5.1. Random forest (RF)

RF is one of the most effective ML models for time-series forecasting tasks and was first proposed by Breiman (2001). RF is an ensemble learning model that builds on multiple decision trees and aggregates their outputs to enhance predictive power. It is designed to mitigate the overfitting problem commonly seen in decision tree models, and its versatility makes it suitable for both regression and classification tasks.

A unique feature of RF is that it uses a method known as "bootstrap aggregating" or "bagging", coupled with random subspace selection. This approach enhances model stability and performance by reducing the variance and preventing overfitting. For time-series data, RF applies a variant of all B bootstrap aggregating called block bootstrap to account for temporal dependencies. The final forecast in RF, as shown in Eq. (6), is obtained by averaging the outputs of each tree (Masini, Medeiros, & Mendes, 2023).

$$\hat{Y}_{t+h|t} = \frac{1}{B}\sum_{b=1}^{B}\left[\sum_{i=1}^{\mathrm{T}_b}\hat{\beta}_{i,b}B_{Ji,b}\left(\boldsymbol{X}_i;\hat{\theta}_{i,b}\right)\right]. \tag{6}$$

Here, $\hat{Y}_{t+h|t}$ represents the final forecast of the target variable, which is the cryptocurrency volatility, at a future time step $t+h$. B represents the number of bootstrap samples. b represents each individual bootstrap sample. T_b represents the size of the subset in the bth sample. $\hat{\beta}_{i,b}$: represents the weight assigned to the ith tree in the bth bootstrap sample. $B_{Ji,b}\left(\boldsymbol{X}i;\hat{\theta}_{i,b}\right)$ is the prediction made by the ith tree in the bth bootstrap sample for the input features $\boldsymbol{X}i$. The prediction is based on the estimated parameters $\hat{\theta}i,b$ and the specific path determined by the parent nodes indexed by J_i. The predictions of all trees in the bth bootstrap sample are weighted by $\hat{\beta}_{i,b}$ and summed up.

To optimize the RF's hyper-parameters, we adopt the Grid Search method (Probst, Wright, & Boulesteix, 2019). The search is performed over a defined space of potential hyper-parameter values and utilizes time-series split cross-validation (with the number of splits set at 5) to evaluate the performance of each combination. The hyper-parameters and their respective range of values are provided in Table 4. These hyper-parameters ensure that the model's complexity is appropriately balanced to avoid overfitting or underfitting, thereby contributing to more accurate and reliable volatility forecasts.

5.2. Long short-term memory (LSTM) networks

LSTM, a type of Recurrent Neural Network (RNN), was introduced by Hochreiter and Schmidhuber (1997) and has been widely used in deep learning, particularly for tasks involving time-series data. Compared to standard RNNs, LSTM is specifically designed to address the challenges of exploding and vanishing gradients during the back-propagation process (Goodfellow, Bengio, & Courville, 2016). This capability is particularly valuable for capturing long-term dependencies and patterns when forecasting cryptocurrency volatility.

The unique architecture of one LSTM block includes the cell state c_t and three types of gates: the input gate, the forget gate, and the output gate. Refer to Fig. 3 for an illustration of the LSTM network structure. The cell state stores information through sequence processing, reducing the impact of short-term memory and acting as a form of "memory" for the network. The input gate controls the flow of new information into the cell state, the forget gate controls the flow of information from the previous cell state, and the output gate provides the activation

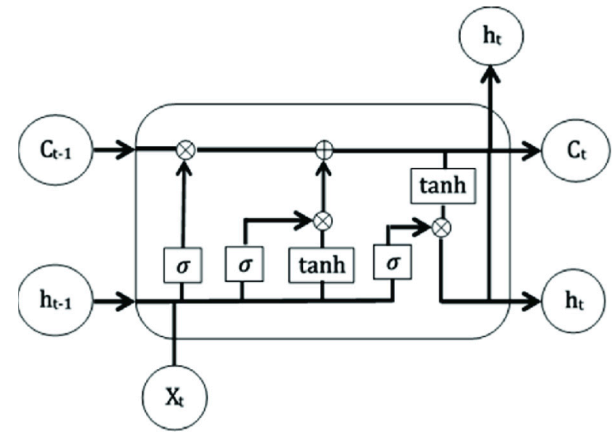


Fig. 3. LSTM network structure.

for the final output of the LSTM block at timestamp t. These gates utilize sigmoid activation functions (i_t, f_t, o_t) to manage the flow of information into, within, and out of the LSTM block, as illustrated in Eq. (7):

$$\begin{aligned} Input\ gate:\ i_t &= \sigma\left(w_i\left[h_{t-1}, x_t\right] + b_i\right) \\ Forget\ gate:\ f_t &= \sigma\left(w_f\left[h_{t-1}, x_t\right] + b_f\right) \\ Output\ gate:\ o_t &= \sigma\left(w_o\left[h_{t-1}, x_t\right] + b_o\right). \end{aligned} \tag{7}$$

Here, σ denotes the sigmoid function; w_i, w_f, and w_o denote the weights of neurons for the respective gate (i, f, o); h_{t-1} denotes the output of the LSTM block at timestamp $t-1$; x_t represents the input at the current timestamp; and b_i, b_f, and b_o denote biases of neurons for the respective gate (i, f, o). The gates modify the candidate cell state, cell state, and the hidden state of the LSTM as follows:

$$\begin{aligned} Candidate\ cell\ state:\ \tilde{c}_t &= \tanh\left(w_c\left[h_{t-1}, x_t\right] + b_c\right) \\ Cell\ state:\ c_t &= f_t * c_{t-1} + i_t * \tilde{c}_t \\ Hidden\ state:\ h_t &= o_t * \tanh\left(c_t\right). \end{aligned} \tag{8}$$

Here, w_c and b_c denote the weights and biases of neurons for the candidate cell state, respectively. The hidden state represents the output of the LSTM block and is passed on to the next timestamp, allowing the LSTM to maintain and update its internal state as it processes the sequence of inputs. During the training process, the LSTM minimizes the loss function, using Mean Squared Error, to find the optimal set of weights and biases. These learned weights enable the LSTM to control its memory process effectively. As a result, the model can remember crucial information and forget irrelevant details over extended periods. Furthermore, multivariate LSTM models can handle more than two input time series at the same timestamp, where we incorporate several determinants to forecast cryptocurrency volatility.

There are several hyper-parameters for LSTM. The number of LSTM neurons within a layer can have a significant impact on accuracy, with a higher number of nodes potentially enhancing accuracy, while a lower number may lead to overfitting issues. The dropout rate is employed to mitigate overfitting problems by reducing sensitivity to specific weights of individual neurons. Additionally, the learning rate defines the update speed of the network hyper-parameters, thereby influencing the overall convergence and stability of the LSTM network.

5.2.1. LSTM hyper-parameters optimization algorithms

As the time complexity of LSTM increases with the dataset's size increase, we choose two optimization algorithms, including GA and ABC, to tune the number of LSTM neurons, dropout, and rate learning rate. Table 5 presents the hyper-parameters and the range value of LSTM in this research.

Table 5
List of optimized hyper-parameters of LSTM.

Hyper-parameter	Range value
Number of LSTM neurons	16, 32, 64, 128, 256
Dropout rate	0.0, 0.1, 0.2, 0.3, 0.4
Learning rate	0.0001, 0.001, 0.01
Batch size	16, 32, 64
Epochs	50, 100
Optimizer	Adam, RMSprop

[a] The lower part of the Table is the hyper-parameters that use the Random search method.
[b] The default LSTM in our research is 128 neurons, followed by a dense layer with 1 neuron, using the Adam optimizer with a mean squared error loss function for 50 epochs and a batch size of 64.

Genetic Algorithm (GA) is a meta-heuristic and stochastic optimization algorithm inspired by the principles of natural evolution, initially proposed by Holland (1992). GA emulates genetic and evolutionary principles to search the solution space, optimizing the target function. The key feature of GA lies in the concept of "chromosomes". Each chromosome represents a potential solution to the problem, typically encoded in binary strings.

The genetic search process for tuning LSTM hyper-parameters can be delineated into six steps: initialization, fitness calculation, termination condition check, selection, crossover, and mutation. For our GA implementation, we set the population size to 50, the number of generations to 10, the crossover rate to 0.8, and the mutation rate to 0.15. Each chromosome in our context encodes different LSTM hyper-parameters, and chromosomes are used to calculate the fitness of the GA. We adopt Root Mean Square Error (RMSE) as the fitness function, where the configuration that results in the smallest RMSE is considered the optimal set of hyper-parameters. The population of chromosomes is initially assigned random values. Selection and recombination operators then search for the optimal solution in this population. If the resulting solution satisfies the termination criteria, it is deemed the optimal solution. Otherwise, the genetic process is repeated until an optimal solution is found, thus allowing the GA to tune the hyper-parameter space of the LSTM model efficiently. We use Python package *deap* to apply GA optimization.

Artificial Bee Colony (ABC) is a meta-heuristic technique inspired by the foraging characteristics of bees, first introduced (Karaboga et al., 2005) to minimize the objective function. There are three types of bees in ABC: employed bees, which are responsible for searching for food sources; onlooker bees, which participate in exploiting food on information received from employed bees in the form of waggle dance; and scout bees, which are responsible for searching for a new food source.

The optimization process of ABC is carried out in the following steps. First, the initial solutions are randomly generated from the sets of LSTM hyper-parameters. Second, train LSTM with these initial hyper-parameters, and evaluate and record the lowest fitness function, which is RMSE in this research. Third, the employed bee considers the RMSE of the initial solution and generates the sets of hyper-parameters to minimize RMSE. Then, onlooker bees choose the best hyper-parameters to minimize RMSE. Lastly, suppose the attempts to generate sets of hyper-parameters are above the abandon limit value. In that case, the employed bee becomes a scout and generates a new hyper-parameter set. The above process is reported until the maximum number of cycles and finds the optimal solution with the lowest RMSE. For our ABC implementation, we set the dimension to 3, the solution number to 10, the population size to 20, the limit to 7, and the maximum cycle number to 15. We use Python package *Hive* to apply ABC optimization.

6. Model evaluation and interpretation

This section elucidates the details of Stage III of our experiment. Section 6.1 introduces the evaluation metrics utilized to evaluate our models. Section 6.2 introduces the model interpretation method, SHAP, to identify significant determinants in forecasting cryptocurrency volatility.

6.1. Model evaluation

To evaluate the performance of the models outlined above, we compute four metrics: Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), Normalized Mean Squared Error (NMSE), and Directional Accuracy (DA). The first three metrics are commonly used in regression tasks, and the lower value suggests a better model fit. NMSE and MAPE are presented in percentage format, indicating the average deviation in percentage terms. Moreover, DA measures the model's ability to predict the direction of changes. A higher value of DA indicates superior forecasting performance. DA presents a perspective on the accuracy of the direction prediction of cryptocurrency volatility, a particularly valuable insight in the context of financial forecasting and investment decisions. The four metrics are defined as follows, $\hat{y}_t$ corresponds to the volatility forecast for time t, y is the actual volatility at time t and n represents the number of forecast time periods:

- RMSE measures the average magnitude of the error, providing a way to quantify the discrepancy between the predicted and actual values:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}\left(y_t - \hat{y}_t\right)^2} \tag{9}$$

- MAPE computes the average absolute percentage difference between the actual and predicted values, providing an indication of relative error:

$$\mathrm{MAPE} = \frac{1}{n}\sum_{t=1}^{n}\frac{\left|y_t - \hat{y}_t\right|}{y_t} \tag{10}$$

- NMSE measures the average squared relative error between the predicted values and actual values:

$$\mathrm{NMSE} = \frac{1}{n}\sum_{t=1}^{n}\left(\frac{y_t - \hat{y}_t}{y_t}\right)^2 \tag{11}$$

- DA measures the model's ability to predict the direction of changes:

$$\begin{aligned} \mathrm{DA} &= \frac{100}{n}\sum_{t=1}^{n} d_t, \\ d_t &= \begin{cases} 1, & \left(y_t - y_{t-1}\right)\left(\hat{y}_t - \hat{y}_{t-1}\right) \geq 0 \\ 0, & \text{otherwise} \end{cases} \end{aligned} \tag{12}$$

6.2. Model interpretation

SHapley Additive exPlanations (SHAP) developed by Lundberg and Lee (2017) is a method to explain individual predictions based on the game theoretically optimal Shapley values. The objective of SHAP is to interpret the prediction of instance x by computing the contribution of each feature to the forecasting process, which computes the Shapley value based on joint game theory. SHAP is the interpretation of Shapley values as an additive feature attribution method, a linear function. SHAP specifies the explanation as:

$$g\left(z'\right) = \phi_0 + \sum_{j=1}^{M}\phi_j z'_j \tag{13}$$

where g is the explanation model, $z' \in \{0,1\}^M$ is the coalition vector, M is the maximum coalition size, $\phi_j \in \mathbb{R}$ is the feature attribution for feature j, and summing the effects of all feature attributions approximate the output $f(x)$ of the original model.

SHAP satisfies three desirable properties: local accuracy, missingness, and consistency.

- Local accuracy: requires the explanation model g to at least match the output of f for the input x', when approximating the original model f for the input x.

$$\hat{f}(x) = g\left(x'\right) = \phi_0 + \sum_{j=1}^{M} \phi_j x'_j \tag{14}$$

- Missingness: a missing feature, where $x'_i = 0$, with no importance.

$$x'_j = 0 \Rightarrow \phi_j = 0 \tag{15}$$

- Consistency: the attribution assigned to the feature will increase or stay the same, even if we change the model. Let $\hat{f}_x\left(z'\right) = \hat{f}\left(h_x\left(z'\right)\right)$ and z'_{-j} indicate that $z'_j = 0$. For any two models f and f' satisfy:

$$\hat{f}'_x\left(z'\right) - \hat{f}'_x\left(z'_{-j}\right) \geq \hat{f}_x\left(z'\right) - \hat{f}_x\left(z'_{-j}\right) \tag{16}$$

for all inputs $z' \in \{0,1\}^M$, then $\phi_j\left(\hat{f}', x\right) \geq \phi_j(\hat{f}, x)$.

SHAP Feature Importance: features with large absolute Shapley values play an important role in forecasting. The global importance is the average of the absolute Shapley values per feature across the data shown as:

$$I_j = \frac{1}{n}\sum_{i=1}^{n}\left|\phi_j^{(i)}\right| \tag{17}$$

SHAP is based on the size of feature attributes and calculates the feature importance by comparing the model predictions with and without the feature, which is a fair process for comparison to show the influence of the input feature in the forecasting process. The impact of the features of the forecasting model is presented as a bar plot to show the global importance of features. Besides, the SHAP summary plot combines feature importance with feature effect, which presents the distribution of the Shapely values of each feature. In this study, we adopt the training set as the background dataset and use Python package *shap* to apply DeepExplainer to compute SHAP values (known as Deep SHAP) that are based on relations between SHAP and Deep Learning Important FeaTures (DeepLIFT) algorithm, proposed by Shrikumar, Greenside, and Kundaje (2017). Deep SHAP combines Shapley values for small components with the whole network through the Deep LIFT multipliers and backwards through the network.

7. Empirical results

As shown in the experiment framework in Section 3, the main objective of our paper is to achieve the most accurate forecasts of cryptocurrency volatility. We consider daily volatility (VOL_1), weekly volatility (VOL_7), and monthly volatility (VOL_30) based on both their internal and external determinants separately. During the forecasting process, we deploy two types of forecasting models. The first is a cryptocurrency-specific model that uses determinants from a single cryptocurrency; these results are discussed in Section 7.1. The second is the universal model that employs determinants from four chosen cryptocurrencies; these results are discussed in Section 7.2. Furthermore, after forecasts, we use SHAP to interpret our models and highlight the significant determinants influencing cryptocurrency volatility. Finally, Section 7.3 discusses how we can hedge cryptocurrency volatility utilizing our model forecasts.

7.1. Cryptocurrency-specific model

This section compares the prediction accuracy of cryptocurrency volatility using internal and external determinants separately for the four chosen cryptocurrencies. The results of internal determinants are discussed in Section 7.1.1. The results of external determinants are discussed in Section 7.1.2.

7.1.1. Internal determinants

In Table 7, we evaluate the performance of different forecasting models using internal determinants. This evaluation is based on metrics such as RMSE, MAPE, NMSE, and DA. The following findings can be made.

Firstly, for daily volatility forecasts (Panel A), when comparing the predictive power of different models, we find that ML techniques such as LSTM and RF models outperform the traditional volatility model, GARCH. Specifically, in Litecoin forecasts, the superiority of ML models is demonstrated. For example, GARCH model yields the lowest MAPE of 31.10%, NMSE of 16.58%, and DA of 63.66%. However, the RF model demonstrates better performance with MAPE of 26.31%, NMSE of 12.86%, and a DA of 41.12%. It is worth noting that DA is not an applicable evaluation metric for GARCH, given that we employ an expanding window strategy for forecasting. In particular, LSTM model, specifically the ABC-LSTM, exhibits significant improvement when its hyper-parameters are fine-tuned using optimization algorithms. This optimization process leads the ABC-LSTM to outperform all other models in out-of-sample Litecoin prediction, achieving the lowest MAPE of 23.45%, NMSE of 8.91%, and a DA of 44.86%. This result can be attributed to the flexibility of ML models and their superior ability to capture complex non-linear relationships between predictors and determinants.

Furthermore, to test the stability of the ML models, we perform seven-day-ahead and 15 day-ahead forecasts (presented in Table D.15), with the seven-day-ahead forecasts demonstrating superior performance, particularly in directional prediction. The best seven-day-ahead forecast is achieved for Litecoin, with a 30.54% MAPE, 15.73% NMSE, and 52.81% DA. Overall, one-day-ahead forecasts perform best, except in directional prediction, where the seven-day-ahead forecasts yield the highest average DA.

Second, for weekly volatility forecasts (Panel B), the RF model performs best for Litecoin, with the lowest MAPE of 27.86% and NMSE of 11.13%. ABC-LSTM also demonstrates superior performance for Litecoin, achieving the lowest MAPE of 22.94% and NMSE of 8.2%. In terms of monthly volatility forecasts (Panel C), using the optimized LSTM model significantly enhances the accuracy of the forecasts compared to the default LSTM and RF models. The most accurate forecast is obtained for Ethereum, with the lowest MAPE of 29.84% and NMSE of 14.38%. The most accurate directional forecast is achieved for Bitcoin, with the highest DA of 57.05%. Regarding the directional forecast accuracy, RF model outperforms other models for both weekly and monthly volatility, achieving the highest DA of 61.12% in Ethereum's weekly forecasts. The ABC-LSTM model demonstrates consistent performance in directional forecasting accuracy, achieving the highest DA of 57.05% in Bitcoin's monthly forecasts.

Finally, the results show that the optimized LSTM models have the most accurate forecasts using internal determinants for weekly cryptocurrency volatility. Daily volatility tends to be excessively noisy due to high-frequency changes in the cryptocurrency market, making it challenging to discern meaningful patterns and accurately capture information related to extreme values. Conversely, monthly volatility is too smooth to capture valuable time points, potentially making the model less sensitive to shorter-term market shifts that can influence the forecast. Therefore, we find that weekly volatility, balancing the granularity of the daily series and the smoothness of the monthly series, offers the optimal temporal scale for predicting cryptocurrency volatility using the models and determinants employed in this study.

Table 6 presents the optimized hyper-parameters of ABC-LSTM for daily volatility. Regarding the number of LSTM neurons, we find that neurons of 6, 32, or 64 outperform models with a higher number of neurons. It indicates that increasing the number of neurons could not improve the model prediction accuracy and increase the model complexity and may lead to the risk of over-fitting problem (Goodfellow et al., 2016).

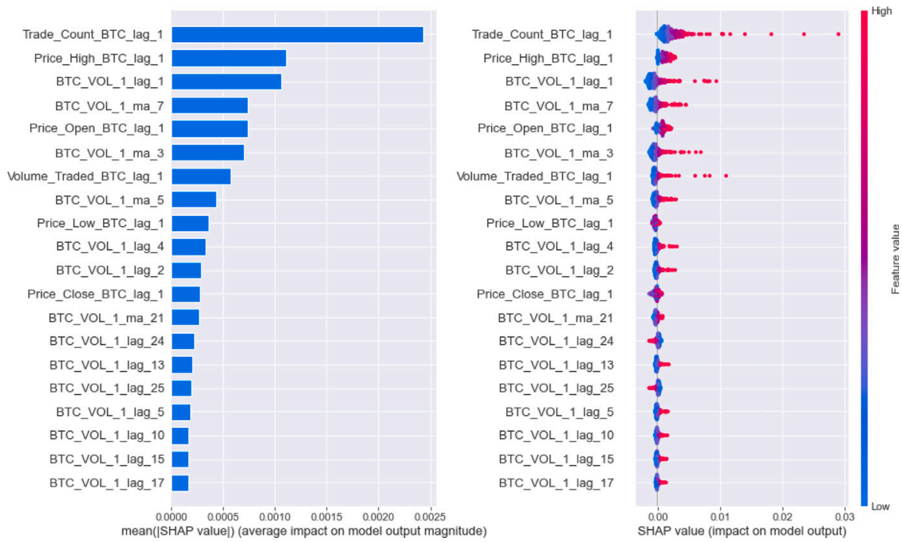


Fig. 4. SHAP summary of VOL_1 for Bitcoin using LSTM with internal determinants.

Table 6
Optimized hyper-parameters of ABC-LSTM for daily volatility.

Hyper-parameters	Optimized value			
	Bitcoin	Ethereum	Litecoin	Ripple
Number of LSTM neurons	16	64	32	16
Dropout rate	0.1	0.2	0.3	0.1
Learning rate	0.001	0.001	0.001	0.001
Batch size	64	32	32	32
Epochs	50	50	50	100
Optimizer	RMSprop	Adam	Adam	Adam

Fig. 4 presents the SHAP summary plot of VOL_1 for Bitcoin using ABC-LSTM considering the internal determinants. The relative importance of each feature is depicted on the left, obtained by averaging the absolute values of the SHAP values in descending order. On the right, each point signifies a row in the dataset, with the gradient colour representing the original value of a feature (high values in red, low values in blue). We find that the leading internal determinants of daily Bitcoin volatility are the previous day's trading count, high price, and volatility. Most internal determinants have a substantial number of red points (indicating high feature values) on the positive SHAP value side. This finding suggests that when internal determinants have high values, they increase the daily Bitcoin volatility for most instances. Furthermore, Figs. E.14 and E.15 present the SHAP summary plot of VOL_7 and VOL_30 for Bitcoin, respectively. We find that the previous day's trading count remains the most influential determinant for both weekly and monthly volatility. This finding is in line with prior findings that indicate a significant time-series momentum phenomenon in the cryptocurrency market (Liu & Tsyvinski, 2021; Liu et al., 2022). However, specific internal determinants of weekly volatility, such as the previous day's low and closing prices, have more negative SHAP values. This implies that when these determinants have high values, they tend to decrease weekly volatility.

7.1.2. External determinants

We use external determinants, including technology, financial, and policy uncertainty factors, to forecast cryptocurrency volatility. Our earlier work has shown that LSTM exhibits superior predictive power over GARCH and RF models when using internal determinants, so we apply LSTM for these volatility forecasts. We also aim to determine whether ABC remains a more sophisticated method than GA in optimizing LSTM's hyperparameters when incorporating external determinants. For each forecast, we consider 27 external determinants. Table 8 presents the out-of-sample performance of forecasts using external determinants, with Panel A, B, and C presenting the daily, weekly, and monthly volatility forecasts, respectively.

Upon comparison, we find that internal determinants exhibit higher predictive power than external ones. When using the external determinants, the best prediction is using ABC-LSTM to predict the Bitcoin daily volatility, achieving 29.58% MAPE, 14.13% NMSE, and 50.43% DA compared with the prediction of 26.35% MAPE, 11.54% NMSE, and 46.73% DA using the internal determinants. However, incorporating external determinants allows us to generate more insights from other financial markets or blockchain platforms, potentially improving the accuracy of directional volatility prediction. Even though the external determinants may yield slightly less accurate predictions overall, they may prove invaluable in predicting the direction of the volatility changes, which can be crucial in investment decision-making and risk management.

Fig. 5 presents a SHAP summary plot of VOL_1 for Bitcoin using ABC-LSTM when considering the external determinants. We find that some financial factors, such as the adjusted close price of NASDAQ and S&P 500, and USD-EURO, are the most influential determinants. Moreover, trading volumes in the financial market emerge as important determinants, implying a direct connection between the financial and cryptocurrency markets. Interestingly, most of these financial factors positively impact daily Bitcoin volatility, except for the adjusted close price of NASDAQ. These observations suggest that specific financial determinants significantly impact daily Bitcoin volatility. Our finding aligns with previous research that suggests the trading volume of the financial market influences volatility and therefore affects investors' decision-making processes (Alexander & Dakos, 2020). Furthermore, the Google search volume for "Bitcoin" and certain blockchain factors, such as mining difficulty and the number of payments per block, also positively influence Bitcoin's daily volatility. This highlights the intricate nature of the cryptocurrency market and affirms previous findings suggesting that technological factors hold predictive power over the cryptocurrency market (Chen et al., 2021).

7.2. Universal model

This section uses the universal models trained with determinants from four chosen cryptocurrencies. Table 9 presents the out-of-sample performance of forecasts using universal RF, GA-LSTM, and ABC-LSTM. The results suggest the following conclusions.

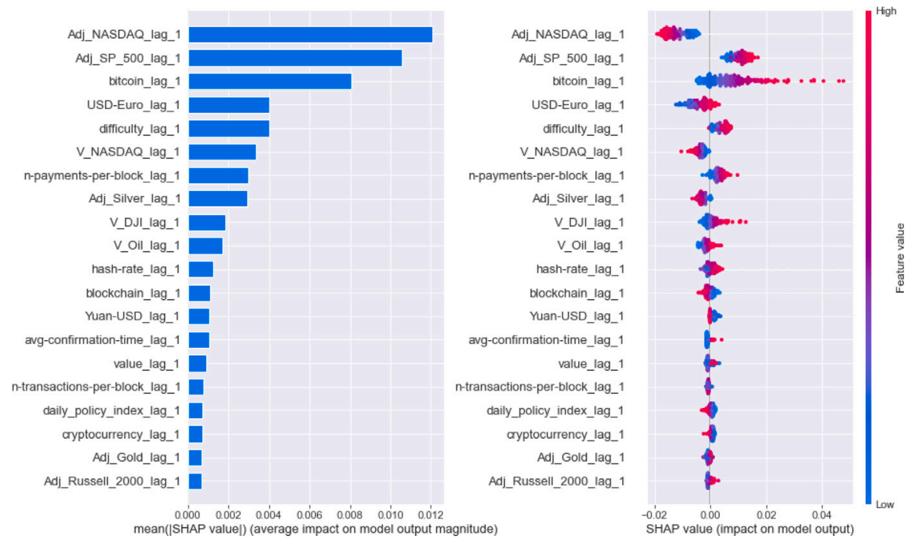


Fig. 5. SHAP summary of VOL_1 for Bitcoin using LSTM with external determinants.

Table 7
Out-of-sample performance of one-day-ahead volatility forecasts using internal features.

	Panel A: Daily volatility forecasts					Panel B: Weekly volatility forecasts				Panel C: Monthly volatility forecasts			
	GARCH	RF	LSTM	GA-LSTM	ABC-LSTM	RF	LSTM	GA-LSTM	ABC-LSTM	RF	LSTM	GA-LSTM	ABC-LSTM
Bitcoin volatility forecasts													
RMSE	0.014	0.012	0.014	0.013	0.013	0.034	0.031	0.031	0.034	0.109	0.117	0.080	**0.065**
MAPE	38.02	28.31	35.81	34.83	26.35	29.21	26.71	26.60	27.80	48.40	59.08	39.86	**32.76**
NMSE	26.02	13.62	21.84	20.76	11.54	13.58	11.18	11.18	12.94	40.47	44.46	21.54	**15.12**
DA	64.51	40.19	42.37	38.44	**46.73**	43.21	40.43	41.67	43.83	56.43	41.69	39.81	**57.05**
Ethereum volatility forecasts													
RMSE	0.017	0.015	0.017	0.016	0.016	0.045	0.055	0.051	0.046	0.146	0.134	0.111	0.075
MAPE	38.61	27.01	34.13	29.01	30.15	31.84	41.91	36.65	29.66	61.65	55.76	46.43	**29.84**
NMSE	24.97	12.85	20.02	14.12	15.86	15.73	25.68	20.35	14.50	50.77	41.33	29.89	**14.38**
DA	63.94	40.50	42.99	45.48	44.24	**61.42**	42.90	44.14	43.52	**58.31**	44.20	45.45	46.71
Litecoin volatility forecasts													
RMSE	0.022	0.023	0.024	0.022	0.023	0.052	0.053	0.051	0.052	0.092	0.083	0.081	0.079
MAPE	**31.10**	**26.31**	36.84	27.33	**23.45**	**27.86**	30.64	27.44	**22.94**	27.17	27.09	26.36	25.17
NMSE	**16.58**	**12.86**	21.92	12.73	**8.91**	**11.13**	13.37	11.03	**8.20**	11.99	11.50	10.83	9.93
DA	63.66	41.12	40.81	45.17	44.86	46.30	42.90	43.52	42.59	57.68	41.69	42.32	41.69
Ripple volatility forecasts													
RMSE	0.021	0.024	0.027	0.026	0.024	0.058	0.062	0.058	0.057	0.097	0.090	0.087	0.087
MAPE	35.28	31.61	45.07	36.93	29.29	32.65	35.93	26.69	27.92	31.20	30.31	29.37	29.13
NMSE	20.38	17.21	33.34	22.72	15.31	17.05	19.73	11.72	12.64	13.79	13.52	12.78	12.31
DA	67.04	39.56	43.61	42.37	43.61	59.26	42.59	45.68	44.14	57.99	40.75	42.63	41.69

[a] Metrics MAPE, NMSE, and DA are presented in the format of percentages.

First, LSTM outperforms RF regarding prediction power, effectively extracting hidden and meaningful information from financial time-series data. Moreover, using an effective hyper-parameter optimization method can further enhance forecasts. In this case, ABC proves superior in identifying the optimal hyper-parameters of LSTM. It demonstrates the high predictive power of ML techniques, particularly deep learning methods, in the cryptocurrency market.

Additionally, the universal model achieves better prediction accuracy than the cryptocurrency-specific model. This finding is in line with previous research in the stock market, such as the study by Sirignano and Cont (2019), where a universal LSTM model outperforms stock-specific models in forecasting the direction of stock price movement. This phenomenon suggests the presence of volatility clustering in the cryptocurrency market, indicating that certain dependencies can be exploited to construct an effective investment portfolio, thereby offsetting risks and maximizing returns. The best prediction is for the Bitcoin monthly volatility, achieving 14.74% MAPE, 3.22% NMSE, and 47.98% DA. This suggests a significant improvement compared to the forecasts using only internal determinants (with 32.76% MAPE and 15.12% NMSE) or external determinants (with 35.63% MAPE and 19.80% NMSE). Therefore, including all internal and external determinants in monthly forecasts offers a more comprehensive and insightful understanding of forecasts.

Furthermore, Fig. 6 provides a SHAP summary plot of VOL_1 for Bitcoin using the universal ABC-LSTM model. Certain financial factors, such as the adjusted close price of NASDAQ and S&P 500, are observed to have a negative impact on daily Bitcoin volatility. Conversely, the previous day's high and open price, as well as the trade count of Bitcoin, positively influence daily Bitcoin volatility. Interestingly, we find that factors, such as the previous day's Ethereum volatility and trading information, also impact daily Bitcoin volatility. These observations suggest a potential correlation between Bitcoin and Ethereum, reinforcing the recognized clustering pattern in the cryptocurrency market where different cryptocurrencies group together due to their temporal similarities (Sigaki, Perc, & Ribeiro, 2019). Moreover, technology factors, including Google search volumes and certain blockchain factors,

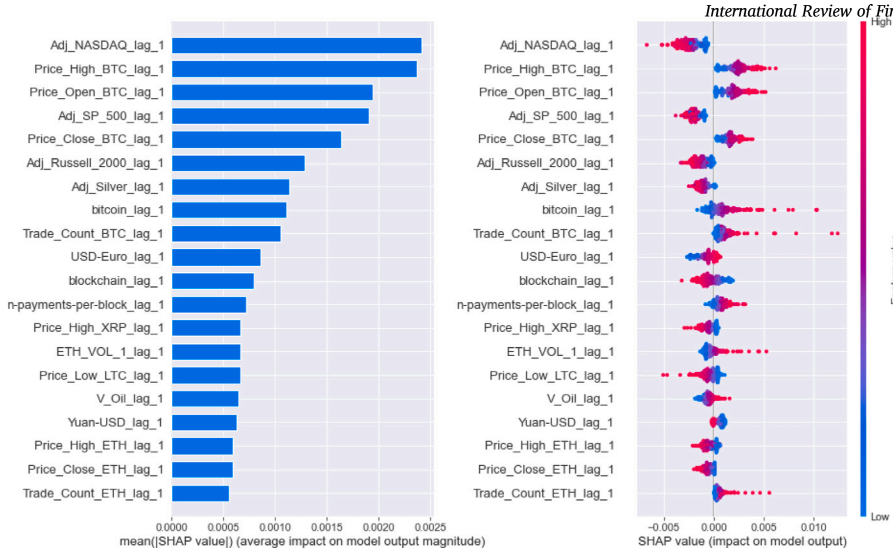


Fig. 6. SHAP summary of VOL_1 for Bitcoin using universal LSTM.

Table 8
Out-of-sample performance of one-day-ahead volatility forecasts using external features.

	Panel A: Daily volatility forecasts			Panel B: Weekly volatility forecasts			Panel C: Monthly volatility forecasts		
	LSTM	GA-LSTM	ABC-LSTM	LSTM	GA-LSTM	ABC-LSTM	LSTM	GA-LSTM	ABC-LSTM
Bitcoin volatility forecasts									
RMSE	0.021	0.018	**0.015**	0.052	0.042	0.042	0.141	0.099	0.077
MAPE	58.71	46.45	**29.58**	48.21	35.09	29.39	70.49	46.38	35.63
NMSE	49.33	34.87	**14.13**	36.70	17.25	15.47	59.95	27.93	19.80
DA	47.86	53.42	**50.43**	51.50	54.51	49.79	49.34	51.09	54.15
Ethereum volatility forecasts									
RMSE	0.023	0.022	0.020	0.090	0.080	0.083	0.167	0.157	0.153
MAPE	45.05	42.73	41.56	77.31	67.32	50.94	67.36	64.11	58.98
NMSE	32.45	28.73	26.87	87.68	65.89	43.92	60.97	55.06	47.86
DA	54.70	52.56	54.70	54.51	52.79	50.21	49.78	48.91	50.22
Litecoin volatility forecasts									
RMSE	0.030	0.027	0.026	0.087	0.083	0.079	0.159	0.149	0.144
MAPE	53.35	40.99	37.08	62.37	58.11	55.06	54.75	51.13	49.42
NMSE	45.39	28.33	23.11	56.24	52.12	47.76	44.77	40.32	38.13
DA	50.00	52.56	**55.56**	51.93	51.50	52.79	48.91	51.09	51.09
Ripple volatility forecasts									
RMSE	0.035	0.028	0.033	0.096	0.092	0.078	0.203	0.195	0.176
MAPE	52.65	48.04	48.65	63.89	56.83	43.09	69.90	67.09	54.38
NMSE	53.20	43.56	45.11	62.75	58.23	37.26	66.31	61.78	46.23
DA	50.85	47.01	50.43	52.36	49.36	53.65	48.03	47.60	45.85

[a] Metrics MAPE, NMSE, and DA are presented in the format of percentages.

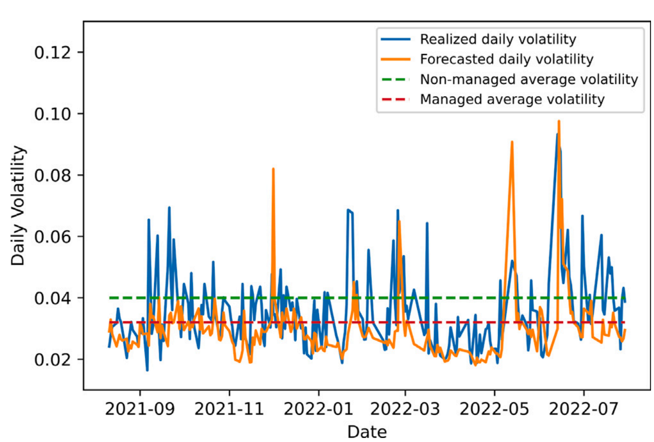


Fig. 7. Forecasted and realized Bitcoin daily volatility.

also emerge as influential determinants. Therefore, to improve the accuracy and reliability of our forecasts, it is crucial to consider the intricacies of the predicted cryptocurrency and account for the influences from the broader cryptocurrency market and external determinants.

7.3. Hedging cryptocurrency volatility

The importance of forecasting volatility is well-recognized, particularly in providing investors with the tools they need to hedge against volatility in the cryptocurrency markets. In the following analysis, we use our forecasted daily Bitcoin volatility as a case study, showing how it can be used to construct a dynamic portfolio that effectively hedges against cryptocurrency volatility.

Fig. 7 shows the out-of-sample forecasted daily volatility (as estimated in the universal model with LSTM) and the realized daily volatility of Bitcoin from August 2021 to July 2022. The two-time series demonstrate a correlation of 0.42, significantly different from zero. This strong correlation suggests a market-timing strategy for hedging volatility. In particular, when the forecasted daily volatility is greater than the historical average volatility, it implies that the realized

Table 9
Out-of-sample performance of one-day-ahead volatility forecasts in universal model.

	RF	GA-LSTM	ABC-LSTM
Panel A: Daily bitcoin volatility			
RMSE	0.011	0.013	0.012
MAPE	22.99	22.53	22.55
NMSE	8.42	9.30	8.54
DA	42.36	50.22	**52.84**
Panel B: Weekly Bitcoin volatility			
RMSE	0.030	0.033	0.032
MAPE	21.49	19.32	17.14
NMSE	8.43	6.19	5.25
DA	50.00	44.74	50.44
Panel C: Monthly Bitcoin volatility			
RMSE	0.043	0.046	0.036
MAPE	18.40	17.33	**14.74**
NMSE	6.33	4.74	**3.22**
DA	49.78	48.88	47.98

[a] Metrics MAPE, NMSE, and DA are presented in the format of percentages.

volatility could also be high the following day. As such, investors can hedge against this volatility risk by not holding any cryptocurrency assets during this period.

The green (red) dash line in Fig. 7 represents the average realized daily volatility when our forecasted volatility is higher (lower) than the historical average. The average daily volatility in the training window (pre-2021 sample period) is used as the historical average volatility, thus avoiding any look-ahead bias when constructing this strategy. Without taking any action to hedge volatility, an investor would face an average volatility of 0.04 in a high volatility period. However, if the investor chooses not to hold any Bitcoin when the forecasted volatility exceeds the historical average, they would face an average volatility of 0.032. This hedging strategy can thus provide investors with around a 25% reduction in volatility, which is economically significant.

While reducing the volatility risk based on the volatility forecast is essential for investors, it is also important to consider (risk-adjusted) returns. This phenomenon is especially true when investors have a mean–variance preference and would like to pursue higher expected returns while lowering the volatility risk. In what follows, we show that the hedging strategy discussed above reduces the volatility risk and improves the returns compared to a passive strategy.

Fig. 8 plots the cumulative daily returns of Bitcoin from August 2021 to July 2022. The Bitcoin market experienced more downward price movements than upward movements during this period. The cumulative returns reach −29.1%, indicating that a one-dollar investment is associated with a loss of $0.291. In this case, it is plausible that staying out of the market in periods of higher volatility can result in higher returns (less losses) and lower volatility, which are mean–variance investors wants ultimately when compared to a passive strategy.

To decide which period investors stay out of the market, we follow the above hedging strategy in which investors stay out of the market when the forecasted daily volatility is greater than the historical average volatility in the training period (pre-August 2021).[10] Since investors also care about the downside (tail) risk in their managed portfolio, as any excessive investments during the down (bear) market state is at risk of substantial losses, it is important to hedge those significant price declines to improve the portfolio performance in terms both of the returns and volatility risks. Therefore, we focus on the period from November 2021, when Bitcoin started to experience the most dramatic price declines.

[10] Note that the strategy here is out-of-sample and in real-time since whether to invest in the Bitcoin market at day $t+1$ is determined by the volatility forecasted at day t.

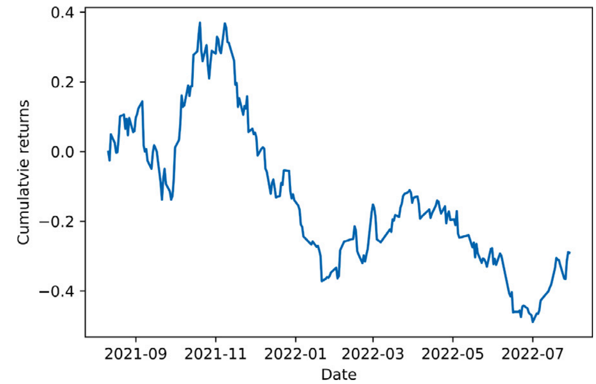


Fig. 8. Cumulative daily returns of Bitcoin.

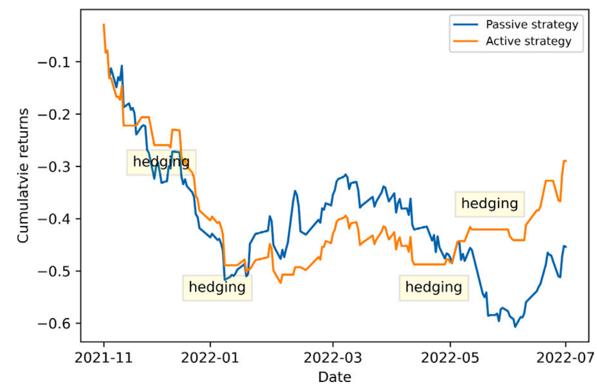


Fig. 9. Cumulative daily returns of hedging and passive strategy.

Fig. 9 plots the cumulative daily returns of the hedging (active) and passive strategy in the Bitcoin market. Compared to the passive strategy, the active strategy hedges many downward movements, as indicated by the straight line (zero returns as no investment is made) in the yellow curve in Fig. 9. As a result, the cumulative daily return of the active strategy (−28.9%) is 16.5% higher than that of the passive strategy (−45.4%), though both are negative since the market was down (bear) state.

Overall, the above findings suggest that our hedging strategy reduces the volatility risk and improves the returns compared to a passive strategy. These findings further reinforce the economic significance of forecasting cryptocurrency volatility.

8. Conclusions

Despite their complex and risky nature, cryptocurrencies have become popular alternative investment tools. In this study, we provide a comprehensive exploration of cryptocurrency volatility forecasts. In particular, we compare the forecasting performance of ML techniques with the traditional GARCH volatility model and explore the determinants of volatility forecasts. Our empirical results demonstrate that ML techniques outperform the traditional method in forecasting cryptocurrency volatility, with ABC-LSTM exhibiting the best prediction performance. Moreover, our universal model outperforms cryptocurrency-specific models, suggesting the presence of volatility clustering in the cryptocurrency market. Furthermore, our SHAP analysis reveals that internal determinants play a significant role in volatility forecasts. Technology factors, including Google search volumes and specific blockchain factors, alongside financial factors, such as the adjusted close prices of NASDAQ and S&P 500, are also influential determinants. This finding suggests that considering a broader range of determinants can help capture the complex dynamics of the cryptocurrency market.

Our research fills gaps in cryptocurrency volatility time-series analysis and provides practical implications. For investors, our state-of-art forecasts of cryptocurrency volatility and deeper understanding of determinants can support more effective investment portfolios of cryptocurrencies and other financial assets, thus mitigating investment risks. For financial institutions and policymakers, our forecasts can support the stable development of the cryptocurrency market, preventing market bubbles and reducing systemic risk. The expanded application and framework of ML techniques can also be applied to other time-series forecasting problems in other financial markets. Future research could enhance the performances of ML algorithms by considering a wider range of potential determinants and refining ML interpretation.

The results of our paper have implications for future studies in several ways. First, the inclusion of more external determinants from sentiment analysis of Twitter and other social media platforms may enhance forecast accuracy. Second, it would be interesting to explore how primary events, such as structural breaks and flash crashes, may affect forecasts. Finally, novel forecasting frameworks that consider correlations among cryptocurrencies and relations between cryptocurrency and other financial markets can be explored. For instance, the graph-neutral network model could provide valuable insights into these relations and improve the forecasts.

Data availability

Data will be made available on request.

Acknowledgments

We are grateful for the helpful comments from Xiao Han, Andrew Urquhart, Yujia Chen and the participants in the 32nd EURO Conference and the Cryptocurrency Research Conference 2022. Yijun Wang acknowledges the best doctoral student paper at the Cryptocurrency Research Conference 2022.

Appendix A. Generalized autoregressive conditional heteroskedasticity (GARCH)

GARCH is a statistical model that is widely employed for forecasting the volatility of financial markets (Agnolucci, 2009; Gökbulut & Pekkaya, 2014; Wang, Ma, Liu, & Yang, 2020). This model was first proposed by Bollerslev (1986) and has since gained widespread acceptance for its versatility and efficacy. A GARCH(p, q) model, where p is the order of the GARCH terms (lagged variances) and q is the order of the ARCH terms (squared residuals), can be defined as shown in Eq. (A.1):

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2, \tag{A.1}$$

where σ_t^2 is the conditional variance of the daily cryptocurrency return at time t, ω is a constant term representing the long-run average variance, α_i are coefficients of the ARCH terms (ϵ_{t-i}^2, the lagged squared residuals from the mean equation), and β_j are coefficients of the GARCH terms (σ_{t-j}^2, the lagged conditional variances). The terms ϵ_{t-i}^2 and σ_{t-j}^2 allow the model to adapt to changes in variance over time, thereby capturing the volatility clustering commonly observed in financial returns. This model's flexibility and effectiveness have made it a tool for financial time series analysis, especially when dealing with market volatility. From ACF and PACF plots of daily return, shown in Fig. A.11, we choose GARCH$(1, 1)$. Additionally, Table A.10 presents the descriptive statistics for daily cryptocurrency return. Moreover, to compare the out-of-sample performance of forecasts with other ML models, we use an expanding window strategy, shown in Fig. A.10. This process continuously adds new data points to the training set to have one-time ahead forecasts.

Appendix B. Summary of acronyms

See Table B.11

Appendix C. Summary statistics of determinants

See Figs. C.12, C.13 and Tables C.12–C.14

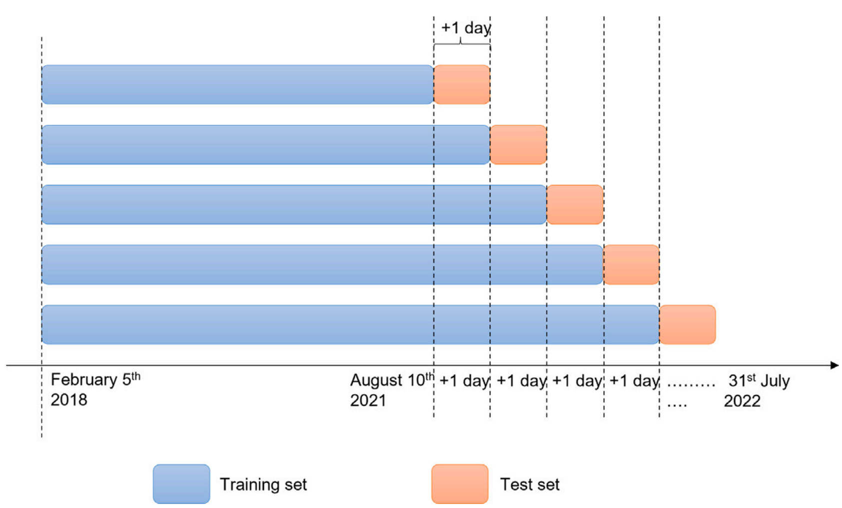
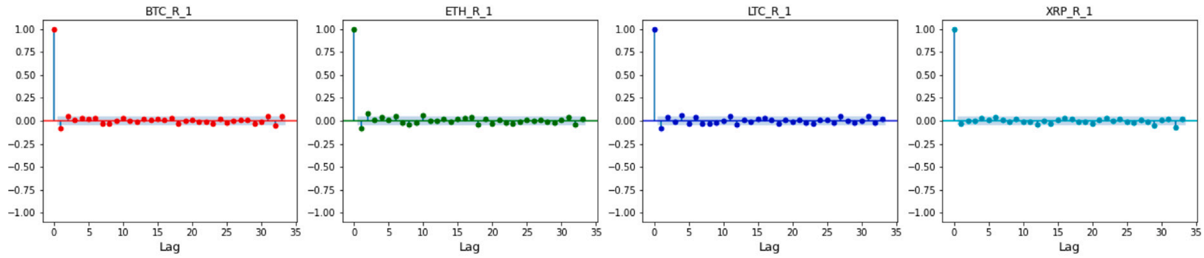


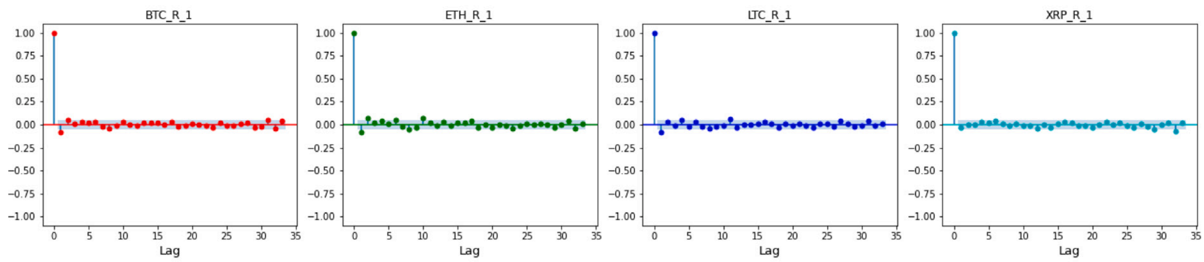
Fig. A.10. Expanding window forecasts.

Table A.10
Descriptive statistics and ADF test for cryptocurrency daily return.

	Count	Mean	Std	Min	25%	50%	75%	Max	Skewness	Kurtosis	ADF Statistic	P-value
BTC_R_1	1673.0	0.000	0.041	−0.491	−0.017	0.001	0.019	0.178	−1.130	14.084	−28.583	0.0
ETH_R_1	1673.0	0.000	0.053	−0.582	−0.024	0.002	0.028	0.235	−1.018	10.715	−12.421	0.0
LTC_R_1	1673.0	−0.001	0.055	−0.478	−0.029	−0.001	0.027	0.288	−0.600	7.991	−12.042	0.0
XRP_R_1	1673.0	−0.001	0.061	−0.538	−0.026	−0.000	0.023	0.450	0.027	11.569	−42.008	0.0

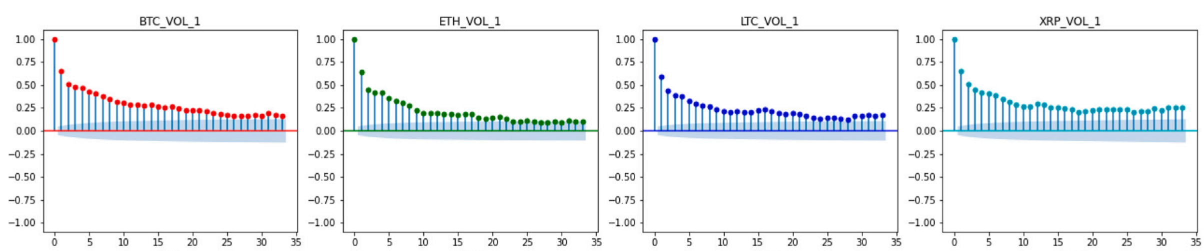


(a) ACF of Daily Return

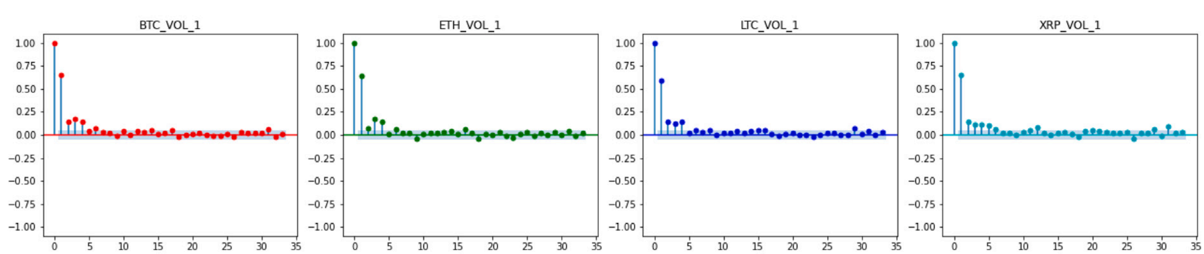


(b) PACF of Daily Return

Fig. A.11. ACF and PACF of daily return.



(a) ACF of Daily Volatility



(b) PACF of Daily Volatility

Fig. C.12. ACF and PACF of daily volatility.

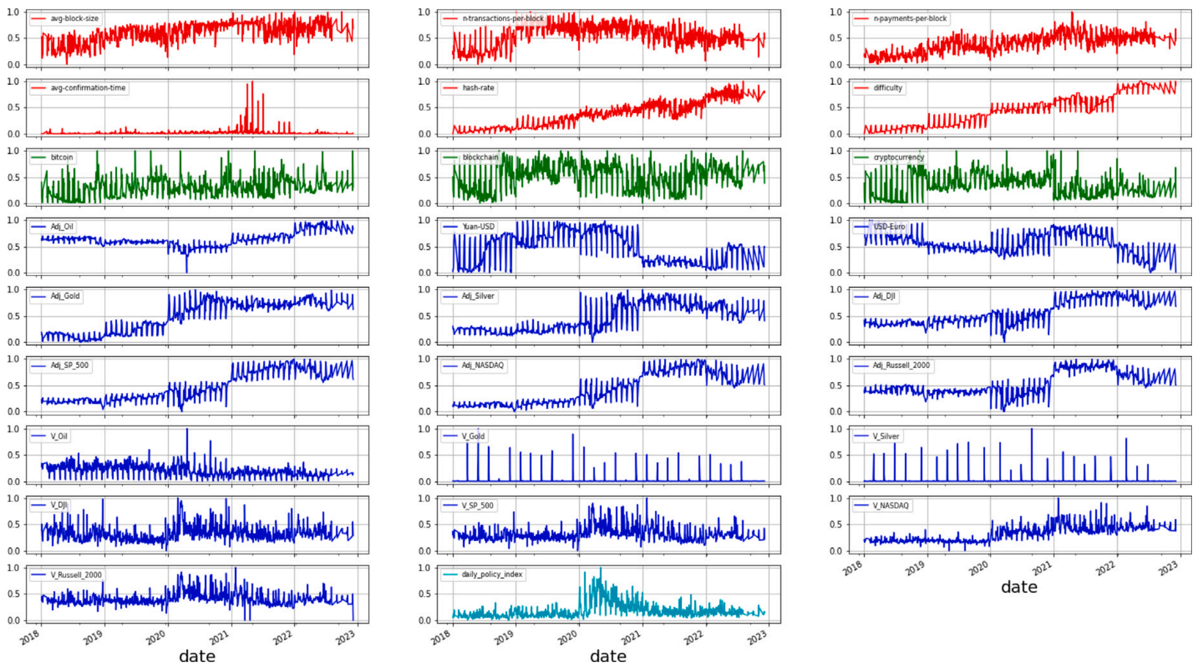


Fig. C.13. The normalization data plot for external determinants.

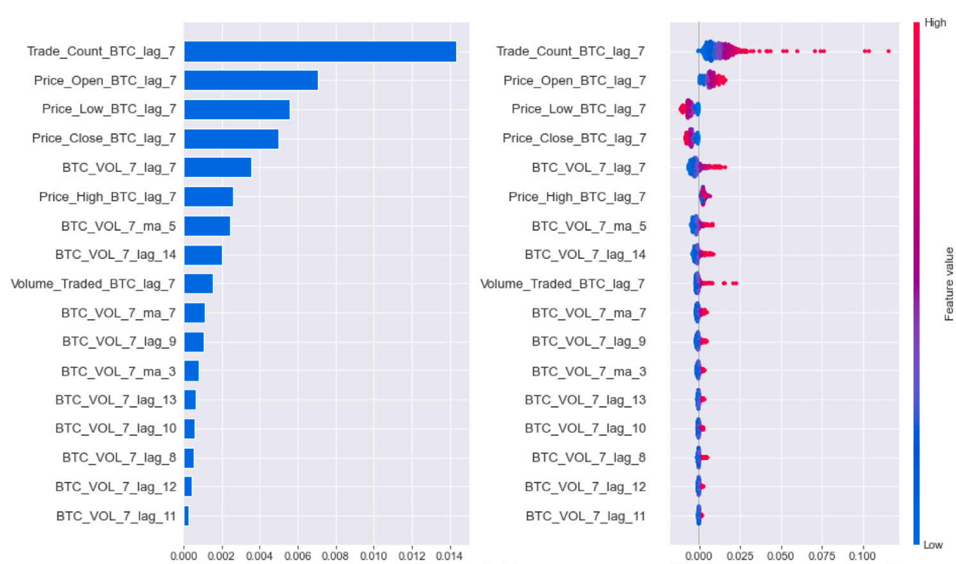


Fig. E.14. SHAP summary of VOL_7 for Bitcoin using LSTM with internal determinants.

Table B.11
Summary of acronyms.

Acronym	Full form
ML	Machine learning
DL	Deep learning
BTC	Bitcoin
ETH	Ethereum
LTC	Litecoin
XRP	Ripple
GARCH	Generalized autoregressive conditional heteroskedasticity
RF	Random forest
LSTM	Long short-term memory
ANN	Artificial neural network
NN	Neural network
ARIMA	Autoregressive integrated moving average
SVR	Support vector regression
SVM	Support vector machine
ANFIS	Adaptive network fuzzy inference system
QDA	Quadratic discriminant analysis
LDA	Linear discriminant analysis
CNN	Convolution neural network
GRU	Gated recurrent unit
KOSPI	Korea stock price index
ABC	Artificial bee colony
GA	Genetic algorithm
SHAP	Shapley additive explanations
RMSE	Root mean squared error
MAE	Mean absolute error
MAPE	Mean absolute percentage error
NMSE	Normalized mean squared error
AUC	Area under the ROC curve
MSPE	Mean squared prediction error
DA	Directional accuracy
ADF	Augmented Dickey–Fuller test
ACF	Autocorrelation function
PACF	Partial autocorrelation function
DJI	Dow Jones industrial average

Table C.12
Correlation analysis.

Panel A: Daily volatility

	lag_1	lag_2	lag_3	lag_4	lag_5	lag_6	lag_7	ma_3	ma_5	ma_7	ma_21	ma_28	ma_35	Open	Close	High	Low	Volume	Count
BTC_VOL_1	0.632	0.481	0.458	0.450	0.392	0.364	0.343	0.615	0.608	0.587	0.487	0.453	0.437	0.125	0.120	0.129	0.112	0.391	0.244
ETH_VOL_1	0.633	0.443	0.423	0.431	0.345	0.300	0.283	0.590	0.581	0.549	0.416	0.376	0.359	0.060	0.055	0.065	0.044	0.425	0.229
LTC_VOL_1	0.574	0.435	0.396	0.399	0.323	0.294	0.278	0.563	0.554	0.531	0.451	0.421	0.417	0.244	0.236	0.252	0.223	0.385	0.339
XRP_VOL_1	0.644	0.496	0.449	0.419	0.375	0.335	0.307	0.619	0.599	0.572	0.485	0.476	0.479	0.210	0.211	0.232	0.191	0.389	0.400

Panel B: Weekly volatility

	lag_7	lag_8	lag_9	lag_10	lag_11	lag_12	lag_13	lag_14	ma_3	ma_5	ma_7	Open	Close	High	Low	Volume	Count
BTC_VOL_7	0.574	0.530	0.497	0.470	0.448	0.433	0.421	0.412	0.542	0.522	0.511	0.113	0.109	0.117	0.101	0.369	0.228
ETH_VOL_7	0.507	0.451	0.409	0.375	0.349	0.331	0.317	0.309	0.464	0.435	0.419	0.070	0.065	0.077	0.055	0.392	0.209
LTC_VOL_7	0.500	0.455	0.422	0.398	0.381	0.370	0.364	0.361	0.468	0.450	0.443	0.379	0.368	0.386	0.354	0.415	0.436
XRP_VOL_7	0.572	0.529	0.498	0.478	0.464	0.453	0.446	0.441	0.541	0.526	0.520	0.366	0.361	0.383	0.329	0.281	0.326

Panel C: Monthly volatility

	lag_30	lag_31	lag_32	lag_33	lag_34	lag_35	lag_36	lag_37	ma_3	ma_5	ma_7	Open	Close	High	Low	Volume	Count
BTC_VOL_30	0.461	0.451	0.443	0.437	0.432	0.427	0.423	0.419	0.453	0.447	0.443	0.132	0.132	0.135	0.128	0.174	0.150
ETH_VOL_30	0.392	0.379	0.370	0.361	0.354	0.347	0.339	0.332	0.381	0.373	0.366	0.065	0.066	0.070	0.059	0.202	0.127
LTC_VOL_30	0.480	0.473	0.467	0.462	0.457	0.452	0.447	0.441	0.474	0.470	0.467	0.421	0.422	0.424	0.415	0.321	0.369
XRP_VOL_30	0.631	0.623	0.616	0.609	0.603	0.596	0.589	0.582	0.624	0.618	0.613	0.331	0.336	0.345	0.323	0.294	0.336

Notes: Open, Close, High, Low, Volume, Count represents lag 1, lag 7 and lag 30 of Open Price, Close Price, High Price, Low Price, Trading Volume, and Trade Count for daily, monthly, and monthly volatility separately.

Table C.13
Descriptive statistics for internal determinants.

	Count	Mean	Std	Min	25%	50%	75%	Max
Price_Open_BTC	1673.00	20 407.71	17 737.32	3180.84	7384.57	10 262.54	36 004.80	67 554.13
Price_Close_BTC	1673.00	20 413.24	17 736.72	3183.00	7384.89	10 262.54	36 018.64	67 554.84
Price_High_BTC	1673.00	20 999.98	18 232.14	3241.00	7588.00	10 490.00	37 593.00	69 060.00
Price_Low_BTC	1673.00	19 707.54	17 129.82	3120.00	7215.00	9899.95	34 381.10	66 250.00
Volume_Traded_BTC	1673.00	33 268.97	24 225.72	4264.68	18 222.42	27 361.88	40 180.87	258 373.35
Trade_Count_BTC	1673.00	287 168.93	268 901.24	37 298.00	96 640.00	167 643.00	431 343.00	2482167.00
Price_Open_ETH	1673.00	1121.13	1250.24	82.92	199.09	430.45	1912.04	4811.89
Price_Close_ETH	1673.00	1121.71	1250.28	82.82	199.13	430.45	1911.98	4811.90
Price_High_ETH	1673.00	1162.40	1291.62	84.98	205.03	443.00	1977.41	4870.51
Price_Low_ETH	1673.00	1071.55	1200.61	80.56	192.42	412.11	1801.94	4697.90
Volume_Traded_ETH	1673.00	339 847.87	273 041.63	28 838.46	168 933.53	271 247.57	412 298.43	2587706.85
Trade_Count_ETH	1673.00	230 999.01	270 115.06	11 268.00	45 291.00	92 041.00	370 713.00	2003052.00
Price_Open_LTC	1673.00	102.92	62.73	22.84	53.31	80.59	142.76	388.13
Price_Close_LTC	1673.00	102.82	62.67	22.88	53.35	80.49	142.72	388.30
Price_High_LTC	1673.00	107.43	66.48	23.49	55.18	83.94	149.48	414.07
Price_Low_LTC	1673.00	97.72	58.61	20.00	51.30	76.01	135.65	345.23
Volume_Traded_LTC	1673.00	422 320.34	364 925.25	33 829.75	196 860.39	312 383.66	511 375.76	3373068.06
Trade_Count_LTC	1673.00	57 800.43	58 666.75	4973.00	18 323.00	40 747.00	76 348.00	514 923.00
Price_Open_XRP	1673.00	0.53	0.36	0.14	0.28	0.40	0.71	2.76
Price_Close_XRP	1673.00	0.53	0.36	0.14	0.28	0.40	0.71	2.78
Price_High_XRP	1673.00	0.56	0.39	0.15	0.29	0.42	0.75	3.35
Price_Low_XRP	1673.00	0.50	0.33	0.11	0.27	0.38	0.67	2.53
Volume_Traded_XRP	1673.00	105728419.80	173551694.57	2397679.61	32212884.93	62083335.71	115498626.50	2965898687.00
Trade_Count_XRP	1673.00	38 025.58	58 539.42	2747.00	14 648.00	23 256.00	39469.00	982 458.00

Table C.14
Descriptive statistics for external determinants.

	Mean	Std	Min	Max	Count
Panel A: Blockchain data					
Average block size	1.15	0.18	0.45	1.53	1254
Average transactions	1944.21	344.22	879.90	2734.44	1254
Average payments per block	3838.01	943.41	1520.09	7236.20	1668
Average confirmation time	99.49	327.09	5.67	5203.91	1398
Hash rate	1.27E+08	5.20E+07	1.51E+07	2.66E+08	1254
Difficulty	1.76E+13	7.11E+12	1.93E+12	3.13E+13	1254
Panel B: Google Trend data					
Blockchain trend	62.07	17.33	22	100	1673
Cryptocurrency trend	36.99	19.26	5	100	1673
Bitcoin trend	34.32	16.02	7	100	1673
Panel C: Financial data					
Adj_Oil	62.90	19.86	−37.63	123.70	1154
Adj_Gold	1595.78	254.33	1176.20	2051.50	1152
Adj_Silver	19.83	4.50	11.73	29.40	1151
Adj_DJI	2.88E+04	4066.57	1.86E+04	3.68E+04	1152
Adj_SP_500	3406.27	681.04	2237.40	4796.56	1152
Adj_NASDAQ	1.04E+04	2900.56	6192.919922	1.61E+04	1152
Adj_Russell_2000	1751.63	316.45	991.16	2442.74	1152
V_Oil	5.06E+05	2.32E+05	6.02E+04	2.29E+06	1154
V_Gold	5761.01	3.22E+04	0.00	3.86E+05	1152
V_Silver	1871.13	1.08E+04	0.00	1.31E+05	1151
V_DJI	3.53E+08	1.13E+08	8.62E+07	9.16E+08	1152
V_SP_500	3.97E+09	1.09E+09	1.30E+09	9.88E+09	1152
V_NASDAQ	3.57E+09	1.55E+09	1.49E+08	1.11E+10	1152
V_Russell_2000	3.96E+09	1.11E+09	0.00E+00	9.88E+09	1152
Yuan-USD	6.69	0.26	6.27	7.18	1193
USD-Euro	1.15	0.05	1.00	1.25	1194
Panel D: Policy Uncertainty Data					
US daily news data	159	119.41	4.05	86.10	1673

Notes: Adj represents the adjusted close price, and V represents the trading volume of financial predictors separately.

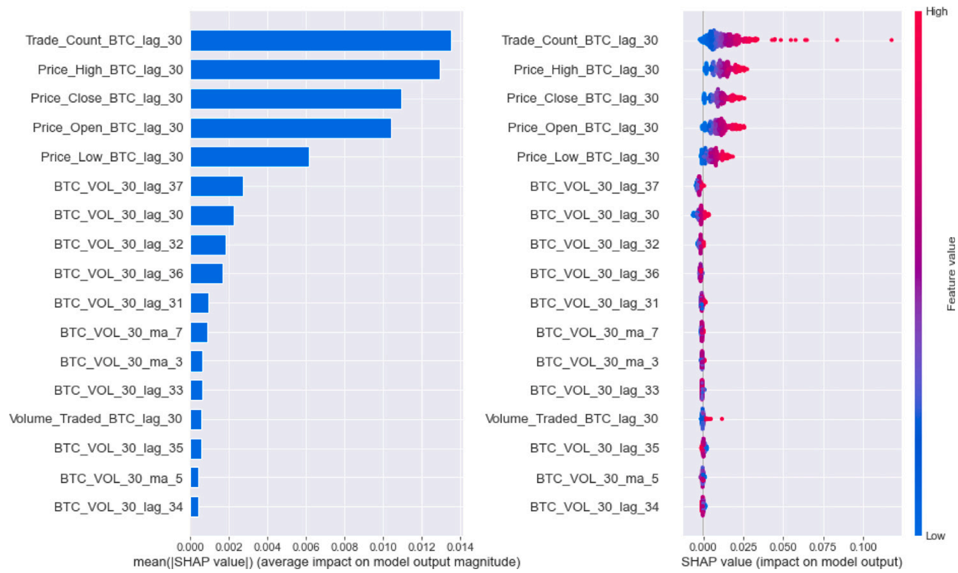


Fig. E.15. SHAP summary of VOL_30 for Bitcoin using LSTM with internal determinants.

Table D.15
Out-of-sample performance of seven-day-ahead and 15 day-ahead daily volatility forecasts using internal determinants.

	Panel A: Seven-day-ahead		Panel B: 15 day-ahead	
Bitcoin volatility forecasts				
	RF	ABC-LSTM	RF	ABC-LSTM
RMSE	0.015	0.016	0.017	0.016
MAPE	37.72	33.65	53.56	45.24
MSPE	24.68	19.70	51.68	34.97
DA	**54.69**	50.94	51.57	48.74
Ethereum volatility forecasts				
RMSE	0.022	0.021	0.036	0.020
MAPE	50.93	43.12	97.19	47.08
MSPE	42.50	32.22	151.35	38.06
DA	49.06	49.38	49.37	46.23
Litecoin volatility forecasts				
RMSE	0.027	0.023	0.024	0.024
MAPE	38.43	**30.54**	35.74	31.40
MSPE	28.90	**15.73**	22.04	17.11
DA	54.37	**52.81**	42.14	47.80
Ripple volatility forecasts				
RMSE	0.025	0.025	0.027	0.025
MAPE	39.60	37.22	52.73	39.02
MSPE	26.37	24.08	48.44	26.56
DA	52.81	50.94	51.57	48.43

Appendix D. Results

See Table D.15

Appendix E. SHAP

See Figs. E.14 and E.15

References

Aggarwal, A., Gupta, I., Garg, N., & Goel, A. (2019). Deep learning approach to determine the impact of socio economic factors on bitcoin price prediction. In *2019 Twelfth international conference on contemporary computing* (pp. 1–5). IEEE.

Agnolucci, P. (2009). Volatility in crude oil futures: A comparison of the predictive ability of GARCH and implied volatility models. *Energy Economics*, *31*, 316–321.

Akyildirim, E., Corbet, S., Lucey, B., Sensoy, A., & Yarovaya, L. (2020). The relationship between implied volatility and cryptocurrency returns. *Finance Research Letters*, *33*, Article 101212.

Alessandretti, L., ElBahrawy, A., Aiello, L. M., & Baronchelli, A. (2018). Machine learning the cryptocurrency market. *Complexity*, *2018*.

Alexander, C., & Dakos, M. (2020). A critical investigation of cryptocurrency data and analysis. *Quantitative Finance*, *20*, 173–188.

Babaei, G., Giudici, P., & Raffinetti, E. (2022). Explainable artificial intelligence for crypto asset allocation. *Finance Research Letters*, Article 102941.

Baker, S. R., Bloom, N., & Davis, S. J. (2016). Measuring economic policy uncertainty. *Quarterly Journal of Economics*, *131*, 1593–1636.

Bengio, Y., Goodfellow, I., & Courville, A. (2017). *Deep learning*: *Vol. 1*, MA, USA: MIT press Cambridge.

Bianchi, D., Büchner, M., & Tamoni, A. (2021). Bond risk premiums with machine learning. *The Review of Financial Studies*, *34*, 1046–1089.

Bianchi, D., Guidolin, M., & Pedio, M. (2022). The dynamics of returns predictability in cryptocurrency markets. *The European Journal of Finance*, 1–29.

van Binsbergen, J. H., Han, X., & Lopez-Lira, A. (2022). Man vs. machine learning: The term structure of earnings expectations and conditional biases. *The Review of Financial Studies*.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *31*, 307–327.

Bouktif, S., Fiaz, A., Ouni, A., & Serhani, M. A. (2018). Optimal deep learning lstm model for electric load forecasting using feature selection and genetic algorithm: Comparison with machine learning approaches. *Energies*, *11*, 1636.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*, 5–32.

Catania, L., & Grassi, S. (2022). Forecasting cryptocurrency volatility. *International Journal of Forecasting*, *38*, 878–894.

Chen, Z., Li, C., & Sun, W. (2020). Bitcoin price prediction using machine learning: An approach to sample dimension engineering. *Journal of Computational and Applied Mathematics*, *365*, Article 112395.

Chen, W., Xu, H., Jia, L., & Gao, Y. (2021). Machine learning model for bitcoin exchange rate prediction using economic and technology determinants. *International Journal of Forecasting*, *37*, 28–43.

Cheng, H. -P., & Yen, K. -C. (2020). The relationship between the economic policy uncertainty and the cryptocurrency market. *Finance Research Letters*, *35*, Article 101308.

Chung, H., & Shin, K. -s. (2018). Genetic algorithm-optimized long short-term memory network for stock market prediction. *Sustainability*, *10*, 3765.

Conrad, C., Custovic, A., & Ghysels, E. (2018). Long-and short-term cryptocurrency volatility components: A GARCH-MIDAS analysis. *Journal of Risk and Financial Management*, *11*, 23.

Corbet, S., Lucey, B., & Yarovaya, L. (2018). Datestamping the bitcoin and ethereum bubbles. *Finance Research Letters*, *26*, 81–88.

Fior, J., Cagliero, L., & Garza, P. (2022). Leveraging explainable AI to support cryptocurrency investors. *Future Internet*, *14*, 251.

Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research*, *270*, 654–669.

Gökbulut, R. I., & Pekkaya, M. (2014). Estimating and forecasting volatility of financial markets using asymmetric GARCH models: An application on Turkish financial markets. *International Journal of Economics and Finance*, *6*, 23–35.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning.* MIT Press.

Gradojevic, N., Kukolj, D., Adcock, R., & Djakovic, V. (2023). Forecasting bitcoin with technical analysis: A not-so-random forest? *International Journal of Forecasting*, *39*, 1–17.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*, 1735–1780.

Holland, J. H. (1992). *Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control, and artificial intelligence*. MIT Press.

Jalan, A., Matkovskyy, R., Urquhart, A., & Yarovaya, L. (2022). The role of interpersonal trust in cryptocurrency adoption. Available at SSRN 4099518.

Karaboga, D., et al. (2005). *An idea based on honey bee swarm for numerical optimization*: *Technical report-tr06*, Erciyes university, engineering faculty, computer engineering department.

Kumar, R., Kumar, P., & Kumar, Y. (2021). Integrating big data driven sentiments polarity and ABC-optimized LSTM for time series forecasting. *Multimedia Tools and Applications*, 1–20.

Li, Z., Li, Z., Li, Z., & Li, Y. (2020). Application of GA-LSTM model in cable joint temperature prediction. In *2020 7th International forum on electrical engineering and automation* (pp. 71–75). IEEE.

Lim, B., & Zohren, S. (2021). Time-series forecasting with deep learning: A survey. *Philosophical Transactions of the Royal Society, Series A*, *379*, Article 20200209.

Liu, Y., & Tsyvinski, A. (2021). Risks and returns of cryptocurrency. *The Review of Financial Studies*, *34*, 2689–2727.

Liu, Y., Tsyvinski, A., & Wu, X. (2022). Common risk factors in cryptocurrency. *The Journal of Finance*, *77*, 1133–1177.

Lundberg, S. M., & Lee, S. -I. (2017). A unified approach to interpreting model predictions. In *Proceedings of the 31st international conference on neural information processing systems* (pp. 4768–4777).

Masini, R. P., Medeiros, M. C., & Mendes, E. F. (2023). Machine learning advances for time series forecasting. *Journal of Economic Surveys*, *37*, 76–111.

McNally, S., Roche, J., & Caton, S. (2018). Predicting the price of bitcoin using machine learning. In *2018 26th Euromicro international conference on parallel, distributed and network-based processing* (pp. 339–343). IEEE.

Nti, I. K., Adekoya, A. F., & Weyori, B. A. (2020). A systematic review of fundamental and technical analysis of stock market predictions. *Artificial Intelligence Review*, *53*, 3007–3057.

Ozbayoglu, A. M., Gudelek, M. U., & Sezer, O. B. (2020). Deep learning for financial applications: A survey. *Applied Soft Computing*, *93*, Article 106384.

Patro, S., & Sahu, K. K. (2015). Normalization: A preprocessing stage. arXiv preprint arXiv:1503.06462.

Peng, Y., Albuquerque, P. H. M., de Sá, J. M. C., Padula, A. J. A., & Montenegro, M. R. (2018). The best of two worlds: Forecasting high frequency volatility for cryptocurrencies and traditional currencies with support vector regression. *Expert Systems with Applications*, *97*, 177–192.

Probst, P., Wright, M. N., & Boulesteix, A. -L. (2019). Hyperparameters and tuning strategies for random forest. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *9*, Article e1301.

Shen, D., Urquhart, A., & Wang, P. (2020). Forecasting the volatility of bitcoin: The importance of jumps and structural breaks. *European Financial Management*, *26*, 1294–1323.

Shrikumar, A., Greenside, P., & Kundaje, A. (2017). Learning important features through propagating activation differences. In *International conference on machine learning* (pp. 3145–3153). PMLR.

Siami-Namini, S., & Namin, A. S. (2018). Forecasting economics and financial time series: ARIMA vs. LSTM. arXiv preprint arXiv:1803.06386.

Sigaki, H. Y. D., Perc, M., & Ribeiro, H. V. (2019). Clustering patterns in efficiency and the coming-of-age of the cryptocurrency market. *Scientific Reports*, *9*, 1–9.

Sirignano, J., & Cont, R. (2019). Universal features of price formation in financial markets: Perspectives from deep learning. *Quantitative Finance*, *19*, 1449–1459.

Smales, L. A. (2022). Investor attention in cryptocurrency markets. *International Review of Financial Analysis*, *79*, Article 101972.

Trucíos, C. (2019). Forecasting bitcoin risk measures: A robust approach. *International Journal of Forecasting*, *35*, 836–847.

Urquhart, A., & Lucey, B. (2022). Crypto and digital currencies—nine research priorities. Nature Publishing Group.

Wang, L., Ma, F., Liu, J., & Yang, L. (2020). Forecasting stock price volatility: New evidence from the GARCH-MIDAS model. *International Journal of Forecasting*, *36*, 684–694.

Yen, K. -C., & Cheng, H. -P. (2021). Economic policy uncertainty and cryptocurrency volatility. *Finance Research Letters*, *38*, Article 101428.

Yuliyono, A. D., & Girsang, A. S. (2019). Artificial bee colony-optimized LSTM for bitcoin price prediction. *Advances in Science, Technology and Engineering Systems Journal*, *4*, 375–383.

Yun, K. K., Yoon, S. W., & Won, D. (2021). Prediction of stock price direction using a hybrid GA-XGBoost algorithm with a three-stage feature engineering process. *Expert Systems with Applications*, *186*, Article 115716.