

Understanding Rate of Return Dynamics of Cryptocurrencies An Experimental Campaign

Koszewski, Krzysztof; Mazumdar, Somnath; Kumar, Anoop S.

Document Version

Final published version

Published in:

Artificial Intelligence Review

DOI:

[10.1007/s10462-023-10629-7](https://doi.org/10.1007/s10462-023-10629-7)

Publication date:

2024

License

CC BY

Citation for published version (APA):

Koszewski, K., Mazumdar, S., & Kumar, A. S. (2024). Understanding Rate of Return Dynamics of Cryptocurrencies: An Experimental Campaign. *Artificial Intelligence Review*, 57(1), Article 8. <https://doi.org/10.1007/s10462-023-10629-7>

[Link to publication in CBS Research Portal](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

If you believe that this document breaches copyright please contact us (research.lib@cbs.dk) providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 04. Jul. 2025





Understanding rate of return dynamics of cryptocurrencies: an experimental campaign

Krzysztof Koszewski¹ · Somnath Mazumdar¹ · Anoop S. Kumar²

Accepted: 17 December 2023
© The Author(s) 2024

Abstract

In recent years, cryptocurrencies have been considered as an asset by public investors and received much research attention. It is a volatile asset, thus predicting its prices is not easy due to the dependence on multiple external factors. Machine learning models are becoming popular for cryptocurrency price predictions, while also considering social media data. In this article, we analyze the rate of return of three cryptocurrencies (Bitcoin, Ether, Binance) from an investor point of view. We also consider three traditional external variables: S&P 500 stock market index, gold price, and volatility index. The rate of return prediction is based on three stages. First, we analyze the correlation between the cryptocurrency returns and the traditional external variables. Next, we focus on the influential social media variables (from Twitter, Reddit, and Wikipedia). Later, we use these variables to improve prediction accuracy. Third, we test how the standard time series models (such as ARIMA and SARIMA) and four machine learning models (such as RNN, LSTM, GRU and Bi-LSTM) predict one-day rate of return. Finally, we also analyze the risk of investing in each cryptocurrencies using value risk statistics. Overall, our result shows no correlation between cryptocurrency returns and three traditional external variables. Second, we found that overall LSTM model is the best, GRU is the second-best prediction model, while the impact of the social media variables varies depending on the cryptocurrencies. Finally, we also found that investment in gold offers better returns than cryptocurrency during Covid-19-like situations.

Keywords Cryptocurrency · Machine learning · Prediction · RoR · Sentiment · VADER · Wavelet

✉ Somnath Mazumdar
sma.digi@cbs.dk

Krzysztof Koszewski
krko17ab@student.cbs.dk

Anoop S. Kumar
askumar@gift.res.in

¹ Department of Digitalization, Copenhagen Business School, Solbjerg Plads 3, 2000 Frederiksberg, Denmark

² Gulati Institute of Finance and Taxation, Thiruvananthapuram 695017, India

1 Introduction

Cryptocurrencies are highly volatile compared to other traditional financial instruments. High volatility of an asset means that an investment is at high risk, but might also offer high reward Damianov and Elsayed (2020). The growing interest in cryptocurrencies has led to many studies that attempt to predict the prices of cryptocurrencies and establish which factors play an essential role in their price fluctuations Khedr et al. (2021), Conlon et al. (2020), Sovbetov (2018), Tschorsch and Scheuermann (2016). The supply of a cryptocurrency is determined by the amount of already existing tokens and those created by validation of transactions Mell and Yaga (2022). Bitcoin is still the most popular cryptocurrency with the highest market capitalization, and many other cryptocurrencies have been created following it, such as Ether Tschorsch and Scheuermann (2016).

Saleemi (2023), Guijarro et al. (2019) shows that microblogging sites have an impact on stock market liquidity. It is possible to detect the effects of public/private opinions on prices by research. Now, social media activity related to cryptocurrencies is a widely studied area. Some of them are Twitter-based Critien et al. (2022), Shen et al. (2019), Reddit-based Seroyizhko et al. (2022), and Wikipedia-based Eisen (2018), ElBahrawy et al. (2019). Our review of the literature revealed that the most relevant factors to consider are Twitter and Reddit with the number of posts and the trend of Wikipedia. For us, it is interesting to investigate how social media data can help investors to navigate challenging times, such as the recent Covid-19 pandemic.

1.1 Research context

Cryptocurrency price prediction is a popular research topic, and most of the work is focused only on Bitcoin price prediction. Existing price prediction works focus either on machine learning (ML) Hansun et al. (2022), Biswas et al. (2021), Raju and Tarif (2020) or time series models Azari (2019), Wirawan et al. (2019). There are many works that combine ML with social media data for better price prediction Critien et al. (2022), Seroyizhko et al. (2022), Shen et al. (2019), ElBahrawy et al. (2019), Eisen (2018). We have not found much that looks into the problem from an investor point of view while focusing on other important external variables (such as the gold price and stock index) to analyze the impact (if any) on price movements. Moreover, very few works consider multiple cryptocurrencies for analysis.

1.2 Research contribution

To fill this gap, our aim is to analyze the rate of return (RoR) of three cryptocurrencies (Bitcoin, Ether, Binance¹) from an investor's point of view. It has led to our primary research question: *How to better understand the price dynamics of the top three cryptocurrencies from an investor perspective?* The primary research question is further subdivided into: (i) How can other factors (such as S&P 500 stock index, gold price, and volatility index) help better understand the price dynamics of the cryptocurrencies? (ii) How standard time series models (ARIMA and SARIMA) and ML models (RNN, LSTM, GRU and

¹ We used the top three cryptocurrencies (Bitcoin, Ether, Binance) based on their asset value at the time of writing of the paper.

Bi-LSTM) are performing while predicting RoR one day ahead? (iii) How can we extract impactful features from social media data (Twitter, Reddit, and Wikipedia) to improve prediction accuracy?

It is the first kind of paper that does not primarily focus on closing price prediction based on social media-related data, but also includes traditional external variables (such as gold and stock index) in one analysis to understand the relationship and RoR impact of the most valuable cryptocurrencies. We also analyze the models' performance during Covid-19. Overall, in this paper, our findings are as follows:

- State-of-the-art time series models ARIMA and SARIMA have mispredicted in all scenarios.
- We found *Relative Positive Twitter sentiment Vol.* and *Reddit Comments* are important social media features for Bitcoin. For Ethereum, *Relative Negative Twitter sentiment* and *Relative Positive sentiment* are important, while *Neutral Twitter sentiment* and *Reddit Score* are important for Binance. It is worth noting that the impact of social media variables varies based on cryptocurrency in the long term.
- We found that the LSTM model is the best, and GRU is the second-best prediction model when predicting RoR using social media data.
- For the Covid-19 scenario, the RNN and Bi-LSTM models perform equally well, followed by the LSTM. At this time, *Relative Negative Twitter Sentiment* and *Reddit comment* is the most impactful social media features.
- We found that investors could withdraw from the cryptocurrency market and invest in gold during turbulent times, such as Covid-19.
- Finally, our analysis also found that Bitcoin is the least risky among the three cryptocurrencies.

2 Background

Cryptocurrencies are intended to be used as forms of exchange, countering traditional currencies. They rely on the blockchain, a distributed ledger technology Kannengießer et al. (2020) that records all transactions on an append-only distributed ledger. Each instance of the cryptocurrency has a unique identifier with all its transactions, since its creation, stored on the distributed ledger, ensuring transparency and security of transactions Tschorsch and Scheuermann (2016). This section will briefly discuss the models used for price analysis. We have used state-of-the-art time series models such as auto-regressive integrated moving averages (ARIMA) and seasonal ARIMA (SARIMA) for the experimental purpose. We recommend the readers to refer to Hyndman and Athanasopoulos (2018) for more information on the ARIMA and SARIMA models. We have also used ML models such as recurrent neural network (RNN), long short-term memory networks (LSTM), gated recurrent unit (GRU) and bidirectional LSTM (Bi-LSTM) Géron (2022). RNNs are developed to solve learning problems where information about the past is linked to future predictions. LSTM is efficient for capturing long-term dependencies across many time stamps. GRU cell is a simplified version of the LSTM cell and can handle sequences much longer than those of simple RNNs. Bi-LSTM is a type of RNN that combines outputs from side-by-side networks to predict the next time step of a sequence having both past and future related information. These time series and ML models are widely used to predict cryptocurrency prices in the literature (see Sect. 3). Below, we will discuss Valence Aware Dictionary and

sEntiment Reasoner (VADER) and Wavelet model used for sentiment analysis and external variable relationship analysis, respectively.

2.1 VADER

VADER is a lexicon and rule-based sentiment analysis tool Hutto and Gilbert (2014). It uses qualitative and quantitative methods to produce and validate a sentiment lexicon tuned for microblogs. Next, it combines these lexical features with consideration for five rules that embody the grammatical and syntactical conventions that humans use. It produces a named vector of VADER results for a single text document. Each word has its sentiment score, which ranges from -4 (most negative) to $+4$ (most positive), and zero represents the neutral sentiment. A text analyzed with VADER yields a vector of positive, neutral, negative, and compound sentiments. Positive, neutral, and negative sentiments add up to one, while compound sentiment ranges from -1 (most negative) to $+1$ (most positive). The compound sentiment will be used to categorize Twitter/Reddit-based sentiments. It is based on normalizing the sum of scores of each VADER-listed word. Here, a compound score above 0.05 will be classified as positive, below -0.05 as negative, and the rest are neutral. We applied the same classification criteria as mentioned in Pano and Kashaf (2020).

2.2 Wavelets

Wavelet transform decomposes time series into its high- and low-frequency components Chui (2014). The attractiveness of the wavelet transform lies in its decomposition properties for time-scale localization, which does not require the assumption of stationarity in the time series. We have used bivariate wavelet coherence (BWC) to detect scale-specific and localized bivariate relationships Hu and Si (2021). BWC measures the correlation between two variables in the time series on different timescales. Wavelet has already been used to find coherence between cryptocurrencies and social media factors Phillips and Gorse (2018). Here, we employ BWC to measure the strength of the association between cryptocurrency returns and other potential variables of interest.

It is important to see whether the cryptocurrency returns are influenced by other factors (such as investor attention, stock market returns, stock market volatility), and other commodities (such as gold price). Wavelet cross-correlation has already been used to find the lead-lag relationship between Bitcoin and a number of representative asset classes (such as gold, oil, and US dollar index) Bhuiyan et al. (2021), Shehzad et al. (2021). We can proceed with the univariate analysis if there is no significant correlation in the short run. If there is a significant short-term correlation, we may employ other potential influencing variables in the forecasting model as exogenous variables.

3 Related work

This section is divided into three subsections. First, we focus on cryptocurrency price prediction that include time series and ML models. We refer the reader to Khedr et al. (2021), which surveys traditional statistical and ML models for cryptocurrency price predictions. Next, we focus on the sentiment analysis of social media data and other online factors. Finally, we list the work that (for better prediction performance) focuses on extended ML models, which include important social media features. A known advantage of ML models

is that most ML models can include additional data that can help to predict better Raju and Tarif (2020). Table 1 compares all the papers studied (based on their applied models and the social media data used).

3.1 Price prediction using standard models

Hansun et al. (2022) have compared the prediction performance of LSTM, GRU and Bi-LSTM using a multivariate approach and found both Bi-LSTM and GRU outperform LSTM. Biswas et al. (2021) have tested multiple neural networks for prediction and found GRU and LSTM work better. Azari (2019) employed multiple ARIMA models to predict the future closing prices of Bitcoin. On the other hand, Wirawan et al. (2019) found that ARIMA models work well for short-term predictions. Ji et al. (2019) showed that negative returns are strongly connected to cryptocurrency price returns and volatility. Yousaf and Ali (2020) examine the links between the pre-Covid-19 and Covid-19 period. The authors found that the conditional correlation between cryptocurrencies is stronger during the Covid-19 period. Conlon et al. (2020) studied whether cryptocurrencies are safe for investors during the Covid-19 period and concluded that Bitcoin and Ethereum are not safe havens. Raju and Tarif (2020) showed that an LSTM model, including the sentiment from Twitter and Reddit posts, is considerably more effective in predicting future prices than an ARIMA model. Sovbetov (2018) focuses on crypto-market factors important for both short and long-term.

3.2 Analysing social media-based features

Online social media activity could greatly influence the prices of cryptocurrencies. Critien et al. (2022) employed an ensemble method and included Twitter sentiment to perform a near-real-time price prediction for Bitcoin. Wołk (2020) included Twitter sentiment in models for short-term cryptocurrency prediction and found that the sentiment tends to be positive regardless of price changes. Valencia et al. (2019) used the VADER polarity score to establish the sentiment for tweets concerning multiple cryptocurrencies. They found that Twitter data could be beneficial in predicting price movement. Shen et al. (2019) have focused on tweet volume and found that the tweet volume from the previous day is a significant driver of the next day's trading volume and realized volatility but not returns. Abraham et al. (2018) found that the total amount of tweets, regardless of sentiment and Google trends, act as better predictors than the sentiment of the tweets, as the latter tends to remain positive regardless of price changes. Lamon et al. (2017) created supervised ML models for Bitcoin and Ethereum price prediction using tweet sentiments and news headlines.

Reddit is another very popular social media that has attracted the attention of many scholars. Applying wavelet decomposition Phillips and Gorse (2018) found that there exists a medium-term correlation between the popularity growth of a given cryptocurrency on Reddit forums, Wikipedia, Google trends, and the price of a given cryptocurrency. Bukovina and Marticek (2016) found that the sentiment of Reddit submission titles can explain a part of Bitcoin's total volatility. Seroyizhko et al. (2022) has found that using too much sentiment data from several subreddits deteriorates the performance of the ML models. ElBahrawy et al. (2019) analyzed the Wikipedia pages of 38 cryptocurrencies and concluded that such data could benefit investors' decision-making. Eisen (2018) employed

Table 1 Literature review table

Paper	Objective	Model	Sentiment Analyzer		Social Media		Studied Cryptocurrencies	
		Time Series	ML		Twitter	Reddit	Wiki	
Seroyzhko et al. (2022)	Prediction, sentiment analysis	No	Yes	VADER	No	Yes	No	BTC
Critien et al. (2022)	Prediction	No	Yes	Yes	Yes	No	No	BTC
Hansun et al. (2022)	Prediction	No	Yes	No	No	No	No	BTC, ETH, ADA, USDT, BNB
Biswas et al. (2021)	Prediction	No	Yes	No	No	No	No	LTC, XMR
Huang et al. (2021)	Prediction	No	Yes	Yes	No	No	No	BTC, ETH, XPR
Khedr et al. (2021)	Survey on price prediction	Yes	Yes	No	No	No	No	BTC
Wolk (2020)	Prediction	No	Yes	VADER	Yes	No	No	BTC, ETH, ETN, XRP, ZEC, XMR
Pano and Kashef (2020)	Sentiment analysis	No	No	VADER	Yes	No	No	BTC
Conlon et al. (2020)	Risk analysis	No	No	No	No	No	No	BTC, ETH, USDT
Yousaf and Ali (2020)	The return and volatility spillover analysis	Yes	No	No	No	No	No	BTC, ETH, LTC
Raju and Tarif (2020)	Prediction	Yes	Yes	Yes	Yes	Yes	No	BTC
Rouhani and Abedin (2020)	sentiment analysis	No	Yes	Yes	Yes	No	No	BTC, ADA, ETH, LTC, XRP
ElBahrawy et al. (2019)	Prediction	No	No	No	No	No	Yes	BTC, ETH, Ripple
Ji et al. (2019)	Return and volatility spillover analysis	No	No	No	No	No	No	BTC, ETH, LTC, Dash, XRP, XML
Burnie and Yilmaz (2019)	Online factors influencing cryptocurrency prices	No	No	VADER	No	Yes	No	BTC
Valencia et al. (2019)	Prediction	No	Yes	VADER	Yes	No	No	BTC, ETH, XRP, LTC
Shen et al. (2019)	Online factors influencing cryptocurrency prices	Yes	No	No	Yes	No	No	BTC
Azari (2019)	Prediction	Yes	No	No	No	No	No	BTC
Wirawan et al. (2019)	Prediction	Yes	No	No	No	No	No	BTC
Abraham et al. (2018)	Prediction	No	No	VADER	Yes	No	No	BTC, ETH
Phillips and Gorse (2018)	Online factors influencing cryptocurrency prices	No	No	No	No	Yes	Yes	BTC, ETH, XMR, LTC

Table 1 (continued)

Paper	Objective	Model		Sentiment Analyzer		Social Media			Studied Cryptocurrencies
		Time Series	ML			Twitter	Reddit	Wiki	
Sovbetov (2018)	Online factors influencing cryptocurrency prices	No	No	No		No	No	No	BTC, ETH, Dash, LTC, XMR
Eisen (2018)	Prediction	No	Yes	No		No	No	Yes	BTC, Dash, DOGE, ETH, LTC, XMR, NEM, XRP, XLM
Lamon et al. (2017)	Prediction	No	Yes	Yes		Yes	No	No	BTC, ETH, LTC
Bukovina and Marticek (2016)	Volatility and sentiment analysis	No	No	Yes		No	Yes	No	BTC

an LSTM model and added Wikipedia pageviews to predict Bitcoin prices. It has shown a strong relationship between the Wikipedia pageview and Bitcoin's closing price.

3.3 Sentiment analysis approaches

There are a few ways in which sentiment could be analyzed. Rouhani and Abedin (2020) first used wordnet to produce a test sample and then used several classification techniques to develop the sentiment of the rest of the data. They found that the support vector machine model can predict the sentiment of the tweets more accurately. Pano and Kashef (2020) analyzed different preprocessing strategies to improve the sentiment scores and found that removing Twitter-specific tags tends to improve the correlation of sentiment scores only over shorter timespans. Burnie and Yilmaz (2019) analyze the price dynamics of given words from the Reddit submissions and built a data-driven phasic word identification methodology. They concluded that the growing popularity of certain words follows the change in price dynamics. In addition to that, Xia et al. (2023) found that global economic policy uncertainty has negative effects, and cryptocurrency uncertainty has a positive long-term impact on Bitcoin volatility.

In this analysis, we have focused more on ML models over time series models as ML models predict better than standard time series models, and RNN, LSTM, GRU and Bi-LSTM are the most popular ML models for cryptocurrency price prediction Hansun et al. (2022), Khedr et al. (2021), Biswas et al. (2021), Raju and Tarif (2020). Apart from that, the most studied online factors that can help explain the closing prices of cryptocurrencies are the trends on Twitter, Reddit, and Wikipedia. Additionally, VADER is the most widely used tool for sentiment analysis. Following these studies, we have selected the models for our analysis. It is worth noting that we have predicted the RoR, which is different from the closing price prediction focused works.

4 Methodology

The literature review section shows that VADER is one of the most common tools implemented for the sentiment of social media posts. Therefore, the text of the acquired tweets and Reddit submissions need to undergo natural language processing (NLP) preprocessing to perform sentiment analysis. This section explains how the data has been collected and processed. Important characteristics are selected in two ways: (i) each feature variable is compared with the price of the corresponding cryptocurrency using Pearson's correlation value r to the price, and (ii) random forest feature extraction is used to find important features. The processing stages performed are depicted in Fig. 1.

From Fig. 1, it can be seen that after collecting the data (prices and three social media data), NLP preprocessing and VADER has been applied on Twitter and Reddit data. Next, the RoR of each cryptocurrencies is also predicted. We also collected external variables such as S&P 500 (represents the 500 largest companies in the United States), gold, and the CBOE volatility index (VIX) for wavelet analysis. Later, we further analyzed the VADER results and passed the VADER output to prediction models for better prediction accuracy.

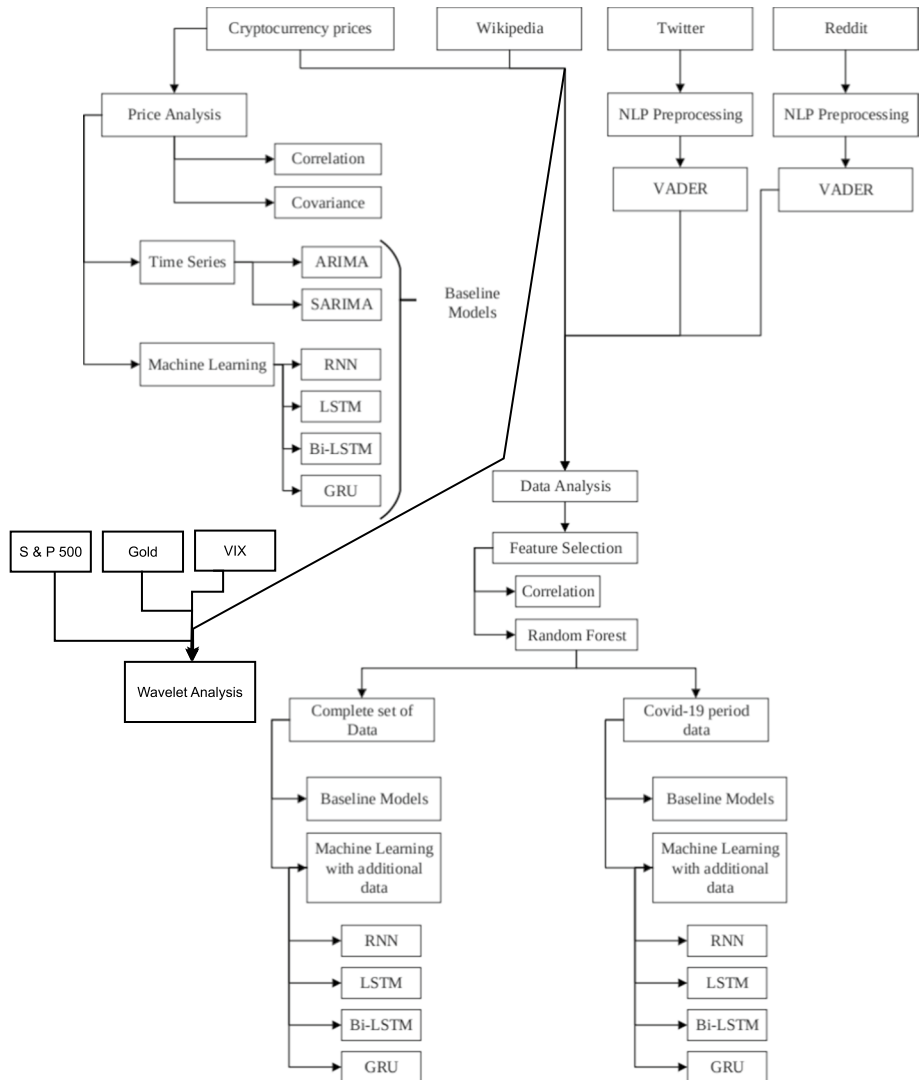


Fig. 1 Representation of applied methodology for the experimental campaign

4.1 Data collection

The raw data has been gathered from Twitter (raw tweets), Reddit (raw Reddit submissions), Wikipedia (pageviews), and Yahoo! Finance and Coinmarketcap² (historical closing prices for each cryptocurrency) Vidal-Tomás (2022). The historical daily closing prices for Bitcoin are acquired from 17–09–2014 to 5–10–2022, and the closing prices of Ethereum and Binance are from 9–11–2017 to 5–10–2022. Similarly, we have collected data for S & P

² <https://coinmarketcap.com/>

500, gold, and VIX. It is worth noting that, unlike others, we have calculated the RoR from the daily closing prices to work on further.

4.2 Twitter data collection

Twitter data are acquired using *snsraper*,³ which is a scraper for social networking services. On Twitter, user posts are visible to one's followers or accessible via a search for a specific hashtag symbol ('#') followed by specific words. The hashtags are symbols and names of the chosen cryptocurrencies: '#Bitcoin', '#btc', '#Ethereum', '#eth', '#Binance', and '#bnb'. Selected tweets have at least one like and one retweet and are written in English. Twitter data contains 15 columns, but only six are kept. They are 'Datetime', 'Replies Count', 'Retweet Count', 'Like Count', 'Quote Count', and 'Tweet'. This process ensures that the collected tweets have at least the minimum engagement and is applicable for sentiment analysis using VADER. The dataset of tweets related to Bitcoin with hashtags '#Bitcoin' and '#btc' has a total of 3,853,299 tweets, Ether with hashtags '#Ethereum' and '#eth' has a total of 1,640,045 tweets, and Binance with hashtags '#Binance' and '#bnb' has a total of 359,855 tweets.

4.3 Reddit data collection

Reddit offers APIs for downloading raw submissions. However, *Pushshift*⁴ offers better search capabilities to search for Reddit comments and submissions. For our work, we have used a multi-thread *Pushshift* API wrapper.⁵ It also offers a larger limit than Reddit's API. Reddit can be navigated through different subreddits, and three subreddits are considered 'r/bitcoin' with 4.7 million members, 'r/EthTrader' with 2.2 million members, and 'r/binance' with 0.88 million members. A total of six variables are chosen for the analysis: upvote ratio, awards, score, number of comments and the 'selftext', which is the actual text of the submission and 'created utc'.⁶ A total of 252,568 valid submissions are pulled from the r/Bitcoin, 40,861 from r/EthTrader and 53,005 from r/Binance.

The sentiment is based on the compound score, and the sentiment analysis variables 'pos' (positive), 'neg' (negative), and 'neu' (neutral) for Twitter and Reddit will be discarded. The sentiment of each tweet should be classified. As mentioned, we classify tweets or submissions with a compound sentiment of above 0.05 as *positive*, below -0.05 as *negative* and the rest as *neutral*.

4.4 NLP preprocessing

Raw texts have to be pre-processed to perform sentiment analysis on Tweets and Reddit submissions. There is a total of 24 variables selected for three cryptocurrencies. The preprocessing includes removing the stopwords (including & symbol), URLs, and text normalization techniques stemming and lemmatizing. As VADER is a rule and lexicon-based sentiment analyzer, the text will be first stemmed and then lemmatized. VADER can deal with UTF-8

³ <https://github.com/JustAnotherArchivist/snsraper>

⁴ <https://github.com/pushshift/api>

⁵ <https://github.com/mattpodolak/pmaw>

⁶ The UTC timestamp had to be converted into a standard date format that does not include hours.

encoded emojis and treats punctuation and capitalization as important metrics of the sentiment polarity score. Punctuation and emojis are kept as they are important for analyzing social media data with the VADER sentiment analyzer.

4.5 Time series data preprocessing

For an ARIMA model to produce a forecast, the time series must first be stationary, meaning that their properties cannot depend on the time at which the series is observed Hyndman and Athanasopoulos (2018). The augmented Dickey-Fuller test is performed to determine whether the data are stationary. A custom function is developed to perform logarithmic and differencing transformations for the given data.

Finally, all features are converted to the same scale so that ML models can work on it.

5 Results

This section is split into social media feature analysis and RoR analysis. During feature analysis, important social media feature variables are extracted by correlation, VADER, and random forest. The RoR analysis presents the correlation and covariance analysis together with the performance of selected time series and ML models. Later, an analysis specific to the Covid-19 period is also presented. As the research question has been constructed from an investor perspective, we have analyzed the VaR of each cryptocurrency.

We used the RoR mentioned in Eq. 1 for the following analysis.

$$\text{Rate of return} = \left[\frac{\text{Current Closing Price} - \text{Previous Closing Price}}{\text{Previous Closing Price}} \right] \times 100 \quad (1)$$

5.1 Prediction error measures

The performance of the time series and ML models are compared using two prediction accuracy measures: mean absolute percentage error (MAPE) and root mean squared error (RMSE). MAPE (see Eq. 2) is a sum of all prediction errors divided by the sum of actual values. MAPE is calculated as the unsigned percentage error and is not scale dependent. Where, n is the number of fitted points, y_i is the actual value, and \hat{y}_i is the predicted value.

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (2)$$

RMSE is defined as the square root of the quotient of the squared sum of errors and the prediction length (see Eq. 3). Importantly, RMSE is scale-dependent.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (3)$$

Table 2 The covariance between the RoRs of BTC, ETH, BNB

Cryptocurrency	BTC	ETH	BNB
BTC	16.31733	15.82546	15.55615
ETH	15.82546	26.02875	19.27642
BNB	15.55615	19.27642	37.75252

Table 3 The correlation between the RoRs of BTC, ETH, BNB

Cryptocurrencies	Pearson's r	p -value
BTC–ETH	0.76790	0
BTC–BNB	0.62676	4.43E-196
ETH–BNB	0.61493	8.79E-187

5.2 Covariance and correlation

There are multiple works which detected cross-correlation relationship among cryptocurrency prices Stosic et al. (2018), cryptocurrency price and stock market index Caferra and Vidal-Tomás (2021), Zhang et al. (2018) and also cryptocurrency-based token price dynamics Vidal-Tomás (2023). To examine the RoR-related relationship between cryptocurrencies, covariance and correlation are calculated. Covariance measures the direction of the relationship, whereas Pearson's r -correlation coefficient measures the strength of the relationship. We have reported the covariance and correlation of three cryptocurrencies in Tables 2 and 3, respectively. We will use BTC, ETH, and BNB to represent the Bitcoin, Ethereum, and Binance cryptocurrency, respectively.

From Table 2, we can see a positive covariance, which means that these three cryptocurrencies tend to move together. However, to know more about the strength of the relationship, we need to calculate their correlation.

Generally, the correlation ranging from +1, presents a positive correlation, while ranging to -1 , shows a negative correlation. From Table 3, we can see that Pearson's r correlation coefficient between the BTC and the ETH is 0.76790, which is the strongest among all. All three Pearson's r coefficients are positive and very high, indicating a strong positive correlation between the RoR. Sadly, none of the p -values are significant enough.

5.3 Feature importance

The compound sentiment is the output of the VADER sentiment analyzer. A compound score above 0.05 will be classified as positive, below -0.05 as negative, and the rest are neutral (similar to Pano and Kashef (2020)). VADER model shows that Twitter-based features are much more influential than Raddit features. This finding aligns with other related works Valencia et al. (2019), Shen et al. (2019), Abraham et al. (2018). From Table 4, we can see that the negative and neutral sentiments are very important.

The random forest model searches for the best feature among a random subset of features. It offers an easy way to measure the relative importance of each feature (weighted average). The random forest regression process makes it possible to know how useful each variable is. The method to obtain that information is the Gini importance measure, whose

Table 4 Sentiment analysis of social media data by VADER model and important features extracted by Random forest model (side-by-side comparison)

Cryptocurrency	Top 3 features by VADER	Top 3 features by random forest
BTC	Relative Twitter Negative Sentiment	Relative Twitter Negative Sentiment
BTC	Relative Twitter Neutral Sentiment	Relative Twitter Neutral Sentiment
BTC	Volume of Reddit Neutral Sentiment	Reddit Comments
ETH	Relative Twitter Negative Sentiment	Reddit Comments
ETH	Relative Twitter Neutral Sentiment	Relative Twitter Negative Sentiment
ETH	Relative Reddit Positive Sentiment	Relative Twitter Neutral Sentiment
BNB	Wiki_trend	Reddit Score
BNB	Relative Twitter Positive Sentiment	Wiki_trend
BNB	Reddit Compound Sentiment	Relative Twitter Positive Sentiment

output value ranges from zero to one in terms of importance, so that the importance of all variables is summed up to one.

We have applied the random forest model to the Twitter, Reddit, and Wikipedia data. The result is shown on the right side of Table 4. The most significant Twitter variables are relative negative and neutral sentiment, while Reddit comments and scores are significant. Lastly, the daily trend of Wikipedia has a very low correlation. The results of the extraction of random forest features suggest that these variables and the trends of Wikipedia are more important in price prediction than pure sentiments. The variables in question measure user engagement on social media and can explain why the volume of a sentiment performs better than the relative sentiment. Here, the volume of a sentiment contains both the sentiment and the volume. Twitter data seems to be much more correlated with closing prices than Reddit data. Daily sentiments and volume are also significantly correlated with the RoR.

Overall, we found that many non-sentiment variables (such as the daily number of likes, retweets, the daily number of all Tweets, the upvote ratio, and the number of comments on Reddit) are also correlated with price fluctuations. Each cryptocurrency dataset (BTC, ETH, BNB) had 24 variables. Correlation analysis and random forest feature extraction are used to determine which variables had the strongest relationship with RoR. However, due to the non-existence of a strong relationship, all 24 variables are used (one variable at a time) with RNN, LSTM, GRU, and Bi-LSTM models and tested to improve the prediction ability.

5.4 Wavelet-base coherence analysis

Before we begin the forecasting exercise, it is important to understand whether any other variables influence the variable of interest. If there is a significant influence, the univariate forecasting methods will provide suboptimal forecasts as we ignore the contribution of other significant factors that influence cryptocurrency returns.

In addition, it is important to see the nature of the influence of other variables on cryptocurrency returns. The wavelet coherence approach has already been applied to detect the movement of cryptocurrencies and stock markets during the Covid-19 pandemic Caferra and Vidal-Tomás (2021). Here, we employ the bivariate wavelet coherence (BWC) to know how each factor influences the cryptocurrency returns. Using wavelet coherence, we can see the strength of the correlation between cryptocurrency returns and other significant

variables in the short, medium and long run. We aim to see if any of these factors influence cryptocurrency returns in the short run. If there is significant coherence in the short run, we may include these variables as exogenous factors in univariate forecasting models. We can proceed with the univariate analysis if there is no significant influence in the short run. In the following paragraphs, we provide a comprehensive discussion of the results (see Fig. 2).

5.4.1 Bitcoin analysis

We analyze the BWC between Bitcoin returns and other potential determinants of its returns. We analyze the coherence between BTC/Wiki, BTC/S&P 500, BTC/VIX, and BTC/Gold (see left side of Fig. 2). First, we see the coherence plot between BTC/Wiki. Here, Wikipedia search interest is used as a proxy for investor attention. We can observe significant coherence during 2018 around the scale of 32 days, coinciding with the cryptocurrency boom and the resultant crash. Later, we found considerable coherence during 2019, when the cryptocurrency market underwent a turbulent phase. Interestingly, we do not see Wikipedia queries affecting BTC returns during the first phase of Covid-19, possibly due to its safe haven property. However, we see a significant correlation on the 32-day scale during 2021.

There is medium-term (32–64 days) coherence between BTC and S&P 500. We employ S&P 500 as a proxy to measure the influence of stock market returns on cryptocurrency returns. During 2018–19, we observed isolated coherence between BTC and S&P 500. There is little fluctuation to speak of. This could be due to the potential hedging nature of BTC. However, during Covid-19, we see strong coherence between BTC and S&P 500, around 128 days. It is possible that investors from the stock market flocked to BTC to use *BTC as a safe haven instrument*. A safe haven instrument is a financial instrument that retains its value or gains in value during financial turbulence. Traditionally, gold and other precious metals are used as safe haven instruments.

Next, we employ the VIX as a proxy for stock market volatility. The coherence pattern between BTC/VIX is almost similar compared to the coherence between BTC/S&P 500. There are two significant periods. First, during 2018–19 and later, during Covid-19, both from the 32-day scales and above. Here, too, the increased coherence could be attributed to the safe haven aspect of BTC.

Looking at the coherence between BTC and gold, it is evident that there are instances of isolated coherence during 2018–19 around the scale of 64 days. Like BTC, gold is also used as a safe haven asset. Therefore, *investors might withdraw themselves from the BTC market and invest in gold during times of turbulence* in the BTC market (our findings similar to Shehzad et al. (2021), Conlon et al. (2020)). The same would be reflected in the returns.

5.4.2 Ethereum analysis

Next, we examine the pairwise coherence between Ethereum (ETH) and the variables of interest (see the middle part of Fig. 2). Here, we include BTC as one of the potential determinants of the ETH returns. As BTC is the market leader in the cryptocurrency market, this is a logical assumption. The following figure exhibits the pairwise coherence between ETH and the variables under study. Looking into the coherence plot between ETH/Wiki,

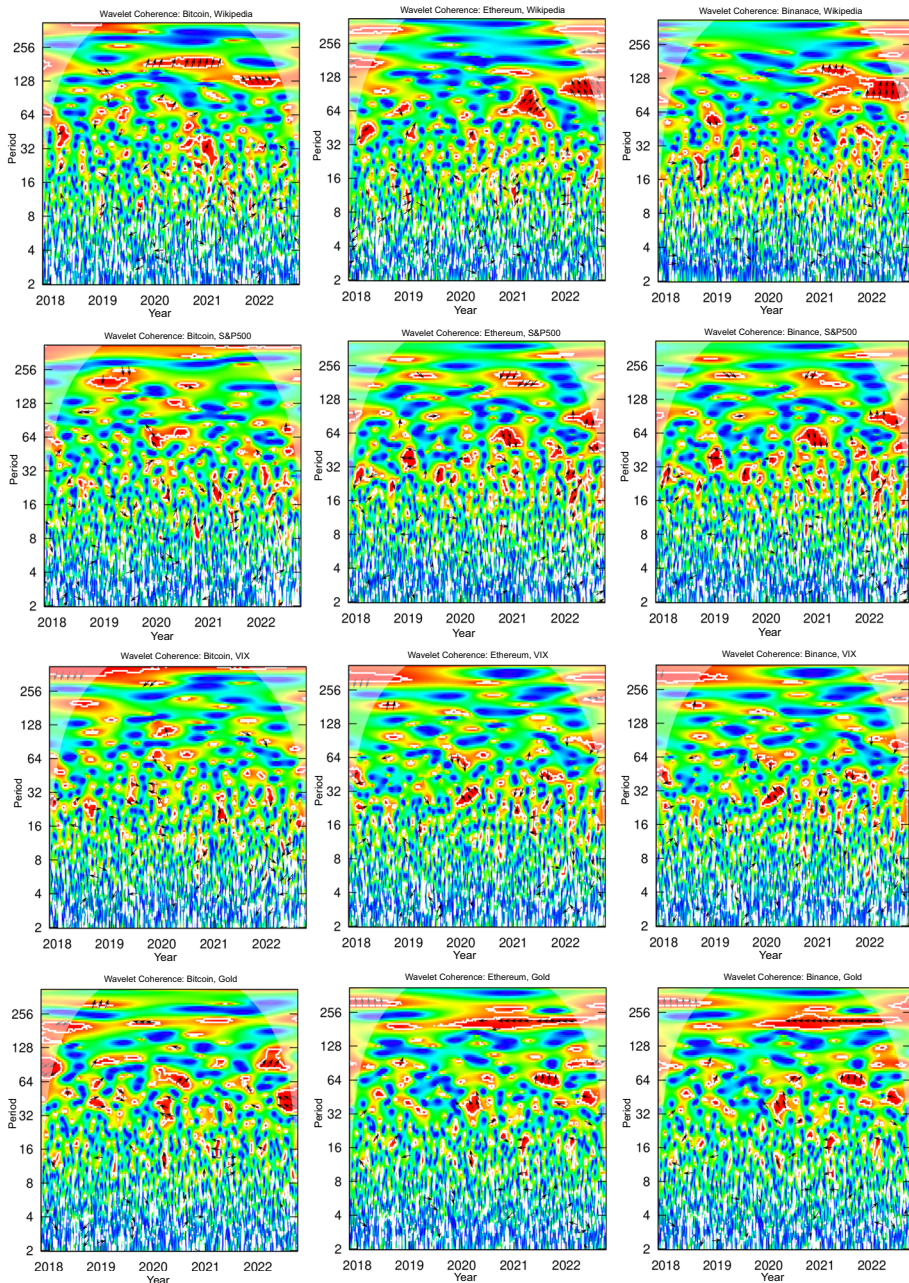


Fig. 2 Wavelet coherence result of three cryptocurrencies based on external variables

we can observe isolated instances of significant coherence during 2018 over 64 days, coinciding with the cryptocurrency boom and the resultant crash.

Table 5 Time series models performance

Cryptocurrency	Model	RMSE	MAPE
BTC	ARIMA(0,0,0)	0.86605	100
BTC	SARIMA(0,0,0)(1,0,0,7)	0.82675	95.46186
ETH	ARIMA(2,0,0)	0.81336	119.26423
ETH	SARIMA(2,0,0)(0,0,0,7)	0.81336	119.26423
BNB	ARIMA(0,0,2)	0.90177	123.65503
BNB	SARIMA(1,0,1)(0,0,2,7)	0.74154	101.68395

We observe isolated medium-term (32–64 days) coherence between ETH and S&P 500 during 2018–19. During Covid-19, we observed strong coherence between ETH and S&P 500 for up to 64 days. It could be due to the potential safe haven property of ETH. The coherence pattern between ETH/VIX is almost similar to that of ETH/S&P 500. We observe isolated instances of significant coherence on a scale of 32 to 64 days during 2018–19. Furthermore, we observe significant coherence during the 2020 Covid-19 crisis in 32–64 days.

Like in the case of BTC and gold, there are instances of isolated coherence between ETH/Gold during 2018–19 and the Covid-19 crisis in 32–128 days. The coherence can be attributed to investors switching between ETH and gold as safe haven assets.

5.4.3 Binance analysis

We also examine the case of BNB (see the right side of Fig. 2). From the BNB/Wiki coherence plot, we can identify significant coherence during the 2018 cryptocurrency crisis across 64 days. However, there is no significant coherence between BNB/Wiki during the Covid-19 crisis.

We observe isolated instances of medium-term coherence (32–64 days) between BNB and S&P 500 during 2018–19. However, there is a strong coherence between BNB/S&P 500 during the Covid-19 period. We observe a coherence pattern between BNB/VIX that is similar to BNB/S&P 500, indicating the potential safe haven property of BNB. The coherence between BNB/Gold is similar to that of the previous cases.

From the wavelet coherence results, we can infer certain patterns. First, there is no correlation between the cryptocurrency returns and the other potential variables of interest. In the short run, the returns dynamics of cryptocurrency are mostly endogenous, that is, determined by factors related to the cryptocurrency market alone. However, we find a significant correlation in the medium to long run (32 days and above). As there is no significant correlation in the short run, we can proceed with the univariate forecasting analysis, as our objective is the short-term prediction of cryptocurrency returns.

5.5 Time series models' based prediction

Table 5 compares the performance of the ARIMA and SARIMA models for each cryptocurrency, where the number of observations per seasonal cycle is set to seven. Following the RMSE and MAPE measures, we can see that both models performed poorly in all cases. For accurate predictions following the RMSE measure, the model should achieve a

low RMSE value. Ideally, a close-to-zero MAPE value is preferred, while the model is usually mispredicted in all cases.

In general, the ARIMA and SARIMA models have failed to predict a one-day RoR for each cryptocurrency. A *perfect* model should have an RMSE value of zero, but we can see that it is invalid in these cases. Here in comparison, SARIMA is less worse than ARIMA.

5.6 ML models' prediction performance

For each model, below three tables (Tables 6, 7 and 8) includes three different set results. We have demonstrated standard RNN (i.e., called Baseline RNN) and its variants' (LSTM, GRU and Bi-LSTM) performance on the left-hand side. Next, we added the Wikipedia pageview count feature to improve the baseline model's performance. We also added the best time series model's performance for comparison.

On the left-hand side (of all three above-mentioned tables), we added the top three Twitter features that offer the best prediction result (by comparing among all Twitter feature variables) and compared all four ML models. Like Twitter features, we added the top three Reddit features to the selected models in the middle part of the tables and compared them.

5.6.1 Bitcoin

Table 6 shows the prediction performance of the four ML models while predicting the RoR of BTC one day in advance. We can see that the GRU model has performed best among all models when adding Twitter features, while the LSTM model outperforms others when adding Reddit features. It can also be seen that adding Wikipedia features to the baseline model also improves the model's prediction performance, except for RNN and LSTM. We also have found that *Relative Positive Twitter sentiment Vol.* and *Reddit Comments* are influential variables for Bitcoin.

5.6.2 Ethereum

Table 7 shows the prediction performance of the four ML models while predicting ETH's RoR one day in advance. In this scenario, we can see that the LSTM model with the Twitter feature has performed best among all four models, while it is also true for the prediction scenario based on Reddit features. However, it is interesting that the standard GRU model has outperformed LSTM in both scenarios. We found that *Relative Negative Twitter sentiment* and *Relative Positive Reddit sentiment* are influential variables for the ETH case. Unlike the previous case, Wikipedia improved the performance of the LSTM and Bi-LSTM models.

5.6.3 Binance

Table 8 shows the prediction performance of the four ML models while predicting BNB's RoR one day in advance. In this scenario, we can see that the GRU model with the Twitter feature has performed best among all four models. LSTM model has performed best when using the Reddit feature. Interestingly, basic GRU model has also performed well. We found that *Relative Neutral Twitter sentiment* and *Reddit Score* are influential variables for BNB. Unlike the ETH case, Wikipedia feature improves the RNN model's performance.

Table 6 Four ML models' prediction performance and added social media features for predicting BTC's RoR

Model	Twitter's top 3 features			Reddit's top 3 features			Baselines & Wiki		
	Variable name	RMSE	MAPE	Variable name	RMSE	MAPE	Model name	RMSE	MAPE
RNN	Retweets	0.08799	10.15957	Submissions vol.	0.21111	24.37652	Baseline RNN	0.65914	76.10934
	Replies	0.18498	21.35858	Positive sentiment vol	0.21260	24.54835	RNN + Wikipedia	3.89680	449.95168
	Quotes	0.24073	27.79680	Comments	0.28223	32.58853	SARIMA	0.82675	95.46186
GRU	Relative positive sentiment	0.05248	6.05971	Neutral sentiment vol.	0.11134	12.85619	Baseline GRU	0.41989	48.48296
	Relative neutral sentiment	0.08728	10.07813	Reddit submission vol.	0.15079	17.41097	GRU + Wikipedia	0.13744	15.86966
	Negative sentiment Vol.	0.14898	17.20263	Positive sentiment vol.	0.15468	17.86030	SARIMA	0.82675	95.46186
LSTM	Positive sentiment Vol.	0.11430	13.19753	Comments	0.06235	7.19894	Baseline LSTM	0.09071	10.47402
	Compound sentiment	0.13922	16.07568	Compound sentiment	0.08321	9.60799	LSTM + Wikipedia	0.11830	13.65932
	Volume	0.16204	18.71026	Reddit score	0.10842	12.51879	SARIMA	0.82675	95.46186
Bi-LSTM	Relative positive sentiment	0.15419	17.80428	Reddit score	0.09254	10.68555	Baseline Bi-LSTM	0.16120	18.61335
	Relative neutral sentiment	0.23891	27.58587	Negative sentiment vol.	0.10431	12.04472	Bi-LSTM + Wikipedia	0.11805	13.63114
	Positive sentiment vol	0.36046	41.62137	Positive sentiment vol.	0.10814	12.48633	SARIMA	0.82675	95.46186

Table 7 Four ML models' prediction performance and added social media features for predicting ETH's RoR

Model	Twitter's top 3 features			Reddit's top 3 features			Baselines & Wiki		
	Variable name	RMSE	MAPE	Variable name	RMSE	MAPE	Variable name	RMSE	MAPE
RNN	Quotes	0.62264	91.29950	Neutral Sentiment Vol.	0.50450	73.97568	Baseline RNN	2.24236	328.80197
	Negative Sentiment Vol.	1.31513	192.84057	Negative Sentiment Vol.	1.17827	172.77156	RNN + Wikipedia	3.24372	475.63283
	Retweets	1.32774	194.68914	Reddit Score	1.30409	191.22113	ARIMA	0.81336	119.26423
GRU	Relative Negative Sentiment	0.37674	55.24229	Comments	0.23112	33.89001	Baseline GRU	0.02637	3.86680
	Relative Neutral Sentiment	0.49026	71.88778	Compound Sentiment	0.25057	36.74164	GRU + Wikipedia	0.35571	52.15856
	Relative Positive Sentiment	0.61491	90.16474	Positive Sentiment Vol.	0.31229	45.79142	ARIMA	0.81336	119.26423
LSTM	Relative Negative Sentiment	0.03915	5.74063	Relative Positive Sentiment	0.04992	7.32019	Baseline LSTM	0.83958	123.10873
	Relative Neutral Sentiment	0.05934	8.70109	Relative Negative Sentiment	0.41237	60.46626	LSTM + Wikipedia	0.64336	94.33754
	Quotes	0.18839	27.62330	Reddit Comments	0.69016	101.19899	ARIMA	0.81336	119.26423
Bi-LSTM	Negative Sentiment Vol.	0.05731	8.40282	Relative Neutral Sentiment	0.18920	27.74298	Baseline Bi-LSTM	0.68192	99.99087
	Volume of Tweets	0.06138	8.99978	Upvote Ratio	0.31185	45.72708	Bi-LSTM + Wikipedia	0.37994	55.71076
	Neutral Sentiment Vol.	0.08888	13.03324	Reddit Score	0.33184	48.65867	ARIMA	0.81336	119.26423

Table 8 Four ML models' prediction performance and added social media features for predicting BNB's RoR

Model	Twitter's top 3 features			Reddit's top 3 features			Baselines & Wiki		
	Variable name	RMSE	MAPE	Variable name	RMSE	MAPE	Variable name	RMSE	MAPE
RNN	Relative Negative Sentiment	0.16326	22.38720	Compound Sentiment	0.14704	20.16319	Baseline RNN	0.52137	71.49253
	Replies	0.25425	34.86444	Awards	0.19422	26.63202	RNN + Wikipedia	0.31859	43.68674
	Quotes	0.53966	74.00009	Neutral Sentiment Vol.	0.22659	31.07077	SARIMA	0.74154	101.68395
GRU	Relative Neutral Sentiment	0.02163	2.96655	Reddit Score	0.02482	3.40341	Baseline GRU	0.06010	8.24086
	Likes	0.15740	21.58376	Neutral Sentiment Vol.	0.02638	3.61744	GRU + Wikipedia	0.06042	8.28528
	Relative Negative Sentiment	0.16588	22.74558	Submissions Volume	0.03019	4.14039	SARIMA	0.74154	101.68395
LSTM	Relative Positive Sentiment	0.18830	25.81990	Reddit Score	0.00960	1.31690	Baseline LSTM	0.38285	52.49809
	Relative Neutral Sentiment	0.21443	29.40333	Awards	0.19043	26.11241	LSTM + Wikipedia	0.43504	59.65477
	Quotes	0.95645	131.15254	Relative Neutral Sentiment	0.67339	92.33786	SARIMA	0.74154	101.68395
Bi-LSTM	Retweets	0.11153	15.29304	Neutral Sentiment Vol.	0.23633	32.40600	Baseline Bi-LSTM	0.44089	60.45731
	Compound Sentiment	0.17512	24.01312	Compound Sentiment	0.26958	36.96570	Bi-LSTM + Wikipedia	1.63324	223.95733
	Relative Neutral Sentiment	0.18995	26.04715	Submissions Vol.	0.27297	37.43025	SARIMA	0.74154	101.68395

Table 9 Covid-19 period results of time series models

Cryptocurrency	Model	RMSE	MAPE
BTC	ARIMA(1,0,1)	1.96396	1438.13754
	SARIMA(1,0,1)(1,0,1,7)	2.00306	1466.77291
ETH	ARIMA(1,0,1)	2.72357	587.66805
	SARIMA(1,0,1)(1,0,1,7)	2.38645	514.92754
BNB	ARIMA(1,0,1)	3.54513	79.12633
	SARIMA(1,0,1)(1,0,1,7)	4.11502	91.84625

Going forward, we also want to see whether a similar trend stays the same during special situations such as Covid-19.

5.7 Analysis of Covid-19 period

SARS-Cov-2 outbreak was declared a pandemic on the 11–03-2020.⁷ This article defines the period of Covid-19 from 11–03-2020 to 26–01-2021. Saleemi (2021) studies the risk of pre- and post-Covid-19 market liquidity associated with Bitcoin trading. Shehzad et al. (2021) found that gold has robust safe haven properties compared to bitcoin during Covid-19. Here, we wanted to see whether the earlier trend persisted during the Covid-19 time.

5.7.1 Times series models

From Table 9, we can see that following the RMSE score ARIMA model outperforms in the BTC case while following the MAPE score, the SARIMA model outperforms in the BNB case. Interestingly, both time series models have failed to forecast. Thus, we may infer that the investment is risky when applying these models. Next, we look at whether the earlier trend of ML models holds further or not.

5.7.2 Bitcoin

Table 10 shows that the Bi-LSTM model with *Relative Negative Twitter Sentiment* has shown the best MAPE and RMSE performance compared to the GRU model (which is the best predictor for Bitcoin for long trend). Strangely, this is the first time the Bi-LSTM model has outperformed other three models. Following both error measures, we found that Bi-LSTM/LSTM behaves the same while using three Reddit features. The baseline Bi-LSTM model with no added features also outperforms others. We also found that adding Wikipedia pageview does not improve any model's performance. ARIMA model performs better than SARIMA, which is the best time series model for Bitcoin for the long trend. Overall, we can state that the Bi-LSTM model overall outperforms, followed by LSTM.

⁷ World Health Organization.

Table 10 Results of ML models during the Covid-19 period for Bitcoin

Model	Twitter's top 3 features			Reddit's top 3 features			Baselines & Wiki		
	Variable	RMSE	MAPE	Variable	RMSE	MAPE	Variable	RMSE	MAPE
RNN	Relative Negative Sentiment	0.61698	451.79546	Relative Neutral Sentiment	2.61070	1911.72565	Baseline RNN	3.32115	2431.96851
	Relative Positive Sentiment	3.04205	2227.59302	Relative Negative Sentiment	2.76079	2021.63143	RNN + Wikipedia	24.98680	18296.98596
	Relative Neutral Sentiment	3.19943	2342.83198	Relative Positive Sentiment	2.94832	2158.95802	ARIMA	1.96396	1438.13754
GRU	Relative Neutral Sentiment	2.64814	1939.14400	Relative Neutral Sentiment	2.61070	1911.72560	Baseline GRU	2.99108	2190.26649
	Compound Sentiment	3.10793	2275.83449	Relative Negative Sentiment	2.76079	2021.63140	GRU + Wikipedia	5.40100	3954.96683
	Relative Positive Sentiment	3.19646	2340.65978	Relative Positive Sentiment	2.94832	2158.95800	ARIMA	1.96396	1438.13754
LSTM	Relative Positive Sentiment	2.07696	1520.88440	Comments	0.37538	274.87771	Baseline LSTM	2.38925	1749.56807
	Relative Negative Sentiment	3.43083	2512.27903	Compound Sentiment	0.47691	349.22747	LSTM + Wikipedia	6.06689	4442.57978
	Quotes	4.03461	2954.40780	Positive Sentiment Vol.	0.81801	599.00419	ARIMA	1.96396	1438.13754
Bi-LSTM	Relative Negative Sentiment	0.02398	17.56095	Comments	0.37538	274.87771	Baseline Bi-LSTM	1.45608	1066.23509
	Relative Neutral Sentiment	0.30809	225.60034	Compound Sentiment	0.47691	349.22747	Bi-LSTM + Wikipedia	1.66056	1215.97288
	Replies	0.88239	646.14088	Positive Sentiment Vol.	0.81801	599.00419	ARIMA	1.96396	1438.13754

5.7.3 Ethereum

From Table 11, we can see that the RNN model with *Relative positive Twitter Sentiment* has shown the best MAPE and RMSE performance compared to the LSTM model (which is the best predictor for ETH for long trend). While Bi-LSTM is the second-best model, following both error measure scores. It is worth noting that the RNN model with *Relative positive Reddit Sentiment* also has shown the best MAPE and RMSE performance. Again, Bi-LSTM has performed well with Reddit features. Next, the Bi-LSTM model with the Wikipedia pageview also performs well, followed by the SARIMA model. In this case, we have seen that RNN and Bi-LSTM consistently perform best among the four selected models.

5.7.4 Binance

From Table 12, we can see that, like the ETH case, RNN again outperforms all other variants following both error measures while using Twitter feature *Relative Negative Sentiment*. However, the Bi-LSTM model performs well on both measurement scales while using *Reddit Comments*. However, we can also see that the baseline Bi-LSTM model with the Wikipedia pageview feature performs poorly compared to the ARIMA model. Overall, we found that for the Covid-19 scenario, the RNN and Bi-LSTM model is the winner, followed by LSTM while predicting three cryptocurrencies. At this time, *Relative Negative Twitter Sentiment* and *Reddit comment* is the most impactful social media feature.

5.8 VaR: value at risk analysis

Finally, to showcase how risky an investment in cryptocurrencies is, the value at risk (VaR) of each cryptocurrency is calculated (refer to Table 13). VaR is a statistical technique used to measure potential losses to a given asset over a period. We have calculated VaR, following the historical method, for which the VaR value is calculated by creating a histogram of historical returns and choosing the confidence interval from there.

From Table 13, we can see that an investor in Bitcoin has a 90% confidence level that their losses will not exceed 3.83%, 95% confidence that their losses will not exceed 6.01%, and 99% confidence that their losses will not exceed 10.56%. An Ether investor, for a one-day investment in Ether, has a 90% confidence level that their losses will not exceed 5.36%, a 95% confidence level that their losses will not exceed 7.69% and a 99% confidence level that their losses will not exceed 13.63%. A Binance investor has a confidence level 90% that their losses will not exceed 5.52%, a confidence level 95% that their losses will not exceed 7.67% and a confidence level 99% that their losses will not exceed 13.68% for a 1-day investment in Binance. We can say that *Bitcoin is less risky than others considering more than 10% daily losses*.

Table 11 Prediction measures during the Covid-19 period for ETH

Model	Twitter's top 3 features			Reddit's top 3 features			Baselines & Wiki		
	Variable	RMSE	MAPE	Variable	RMSE	MAPE	Variable	RMSE	MAPE
RNN	RelativePositive Sentiment	0.30678	66.19460	RelativePositive Sentiment	0.03945	8.51241	Baseline RNN	4.18801	903.65348
	RelativeNegative Sentiment	1.39109	300.15824	Awards	4.63984	1001.14548	RNN + Wikipedia	31.96918	6898.03643
	Quotes	5.10454	1101.41380	Score	6.67584	1440.45480	SARIMA	2.38645	514.92754
GRU	Relative Negative Sentiment	2.82783	610.16399	Relative Positive Sentiment	3.30486	713.09546	Baseline GRU	3.41530	736.92387
	Relative Neutral Sentiment	3.28933	709.74461	Relative Neutral Sentiment	3.48331	751.59913	GRU + Wikipedia	6.36455	1373.28743
	Quotes	3.47562	749.93996	Awards	3.67019	791.92273	SARIMA	2.38645	514.92754
LSTM	Quotes	3.59206	775.06301	Awards	3.79463	818.77347	Baseline LSTM	2.66975	576.05636
	Relative Neutral Sentiment	3.96580	855.70682	Relative Negative Sentiment	3.81221	822.56623	LSTM + Wikipedia	7.44878	1607.23352
	Replies	4.44524	959.15561	Score	3.84846	830.38694	SARIMA	2.38645	514.92754
Bi-LSTM	Retweets	0.52860	114.05695	Compound Sentiment	0.75503	162.91449	Baseline Bi-LSTM	3.43728	741.66778
	Likes	0.72127	155.62867	Positive Sentiment Vol.	0.99740	215.21071	Bi-LSTM + Wikipedia	0.97269	209.87832
	Relative Neutral Sentiment	1.21529	262.22460	Negative Sentiment Vol.	1.07446	231.83765	SARIMA	2.38645	514.92754

Table 12 Prediction measures during the Covid-19 period for binance

Model	Twitter's top 5 features			Reddit's top 5 features			Baselines & Wiki		
	Variable	RMSE	MAPE	Variable	RMSE	MAPE	Variable	RMSE	MAPE
RNN	Relative Negative Sentiment	0.50182	11.20045	Relative Positive Sentiment	8.13997	181.68195	Baseline RNN	8.84559	197.43139
	Relative Positive Sentiment	5.73477	127.99849	Awards	9.45633	211.06297	RNN + Wikipedia	26.72972	596.60045
	Quotes	11.40558	254.56972	Relative Negative Sentiment	9.83978	219.62144	ARIMA	3.54513	79.12633
GRU	Relative Neutral Sentiment	8.44661	188.52614	Awards	7.60900	169.83095	Baseline GRU	8.63232	192.67123
	Negative Sentiment Vol.	9.08036	202.67130	Relative Neutral Sentiment	7.64998	170.74569	GRU + Wikipedia	12.07562	269.52477
	Compound Sentiment	9.36315	208.98315	Relative Negative Sentiment	7.67741	171.35781	ARIMA	3.54513	79.12633
LSTM	Relative Negative Sentiment	7.95301	177.50909	Awards	8.26755	184.52958	Baseline LSTM	7.53057	168.08041
	Negative Sentiment Vol.	8.75238	195.35089	Relative Negative Sentiment	8.61315	192.24340	LSTM + Wikipedia	10.01734	223.58451
	Relative Neutral Sentiment	8.99481	200.76180	Neutral Sentiment Vol.	9.33380	208.32810	ARIMA	3.54513	79.12633
Bi-LSTM	Neutral Sentiment	3.67484	82.02153	Comments	4.22368	94.27142	Baseline Bi-LSTM	6.58761	147.03375
	Volume	4.43065	98.89105	Score	4.33168	96.68196	Bi-LSTM + Wikipedia	3.81667	85.18701
	Relative Neutral Sentiment	4.52620	101.02372	Awards	7.02830	156.86989	ARIMA	3.54513	79.12633

Table 13 VaR analysis for three cryptocurrencies

Confidence level (%)	BTC (%)	ETH (%)	BNB (%)
90	-3.83	-5.36	-5.52
95	-6.01	-7.69	-7.67
99	-10.56	-13.63	-13.68

6 Discussion and limitations

Cryptocurrencies are decentralized virtual assets, which are very volatile. It presents an opportunity for large profits and the risk of large losses Damianov and Elsayed (2020). Cryptocurrencies are typically not backed by any physical assets, which, along with their volatility, makes them extremely challenging for price prediction. It led scholars to try to find factors that may explain the behavior and nature of cryptocurrency prices. Generally, cryptocurrency price predictions are based on two approaches. They are traditional statistical methods and ML models Khedr et al. (2021). This work implemented the most popular methods from the survey Khedr et al. (2021) for the prediction of the price of the top three cryptocurrencies.

We found that the features listed in Table 4 do not have a similar impact when using multiple ML models. During Covid-19, the negative sentiment is a bit prevalent, but the same trend is not seen in a long trend. Time series models have not been able to outperform any ML models. Adding Wikipedia features to the models does not always have a positive impact. In general, we found that *Relative Positive Twitter sentiment Vol.* and *Reddit Comments* are important for Bitcoin. For ETH, *Relative Negative Twitter sentiment* and *Relative Positive sentiment* are important, while *Neutral Twitter sentiment* and *Reddit Score* are important for BNB for long trend scenarios. We found that the LSTM model is the best and GRU is the second best model while predicting RoR using social media features with long-term data. However, the trend changes during Covid-19, where the RNN and Bi-LSTM models perform quite well. Finally, the VaR results show that Bitcoin, Ether, and Binance are very volatile assets with a risk of substantial losses even for short-term investments. However, Bitcoin is the least risky among the three.

The results could have been expanded by including more than one online variable in an ML model so that all the features of social media would be combined and tested.

7 Conclusion

We have analyzed the cryptocurrency RoRs from an investor perspective. To address the primary research question, this study examined the influence of Twitter, Reddit, and Wikipedia pageview data on the prices of the three important cryptocurrencies. It also examines how important the social media features are within each social media channel. We have also analyzed the potential correlation between the cryptocurrency returns and other potential variables such as gold and stock index. The Covid-19 period is also examined to observe how the relationship changes in a volatile economic scenario. Lastly, the VaR of each cryptocurrency is calculated to give a perspective on how risky investing in cryptocurrencies

is. The results are consistent with those of the existing literature. ML models outperformed traditional time series models.

Here, we used wavelet coherence to see if the cryptocurrency returns are influenced by other exogenous factors such as investor sentiment, stock market returns, stock market volatility, and commodities such as gold. The results showed that its own market dynamics influence cryptocurrency returns in the short run. We found that both the ARIMA and SARIMA models outperformed each other (case-by-case) but failed in all cases compared to the ML models. It means that these two time series models cannot capture strong seasonal trends in RoR data. VADER analyzes Twitter and Reddit sentiment data. We found that Twitter data is somewhat correlated with the RoR, much more than Reddit data. Furthermore, all three sentiments, positive, negative, and neutral, are among the most correlated features among all three cryptocurrencies. Wikipedia pageviews are of very little importance in explaining the RoR of cryptocurrencies. Therefore, the findings prove that the online variables do not have a strong relationship with the RoR of Bitcoin, Ether, and Binance. However, there is no clear pattern to tell which sentiment has the most significance.

Our analysis shows that various exogenous factors do not influence cryptocurrency returns. Preliminary analysis shows that cryptocurrency returns could be partially explained in the long run by variables such as social media attention and other financial markets. This issue needs to be studied in detail and is left for future research. As our future work, we also want to replace the current ML models with reinforcement learning models to maximize the RoR of an investor while adding more exogenous factors such as crude oil price, trading volumes of top cryptocurrency spot exchanges. It could be interesting to pass all variables together to the model and compare models' performance.

Author contributions KK: Software (Validation and Verification), Writing (Original Draft) and Visualisation. SM: Conceptualization, Software (Verification), Writing (Original Draft, Review and Editing). AS: Software (Validation and Verification), Writing (Editing) and Visualisation.

Funding Open access funding provided by Copenhagen Business School.

Data availability Not applicable.

Declarations

Competing interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Ethical approval Not applicable.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abraham J, Higdon D, Nelson J et al (2018) Cryptocurrency price prediction using tweet volumes and sentiment analysis. *SMU Data Sci Rev* 1(3):1
- Azari A (2019) Bitcoin price prediction: an arima approach. arXiv preprint [arXiv:1904.05315](https://arxiv.org/abs/1904.05315)
- Bhuiyan RA, Husain A, Zhang C (2021) A wavelet approach for causal relationship between bitcoin and conventional asset classes. *Resour Policy* 71:101971
- Biswas S, Pawar M, Badole S, et al (2021) Cryptocurrency price prediction using neural networks and deep learning. In: 2021 7th international conference on advanced computing and communication systems (ICACCS), IEEE, pp 408–413
- Bukovina J, Marticek M et al (2016) Sentiment and bitcoin volatility. University of Brno
- Burnie A, Yilmaz E (2019) Social media and bitcoin metrics: which words matter. *R Soc Open Sci* 6(10):191068
- Caferra R, Vidal-Tomás D (2021) Who raised from the abyss? A comparison between cryptocurrency and stock market dynamics during the COVID-19 pandemic. *Financ Res Lett* 43:101954
- Chui CK (2014) An introduction to wavelets, vol 1. Academic Press
- Conlon T, Corbet S, McGee RJ (2020) Are cryptocurrencies a safe haven for equity markets? An international perspective from the COVID-19 pandemic. *Res Int Bus Financ* 54:101248
- Critien JV, Gatt A, Ellul J (2022) Bitcoin price change and trend prediction through twitter sentiment and data volume. *Financ Innovation* 8(1):1–20
- Damianov DS, Elsayed AH (2020) Does bitcoin add value to global industry portfolios? *Econ Lett* 191:108935
- Eisen AM (2018) Prediction of cryptocurrency price using wikipedia page views. PhD thesis, Ben-Gurion University of the Negev
- ElBahrawy A, Alessandretti L, Baronchelli A (2019) Wikipedia and cryptocurrencies: interplay between collective attention and market performance. *Front Blockchain* 2:12
- Géron A (2022) Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow. O'Reilly Media Inc
- Guijarro F, Moya-Clemente I, Saleemi J (2019) Liquidity risk and investors' mood: linking the financial market liquidity to sentiment analysis through twitter in the s &p500 index. *Sustainability* 11(24):7048
- Hansun S, Wicaksana A, Khaliq AQ (2022) Multivariate cryptocurrency prediction: comparative analysis of three recurrent neural networks approaches. *J Big Data* 9(1):1–15
- Hu W, Si B (2021) Improved partial wavelet coherency for understanding scale-specific and localized bivariate relationships in geosciences. *Hydrol Earth Syst Sci* 25(1):321–331
- Huang X, Zhang W, Tang X, et al (2021) Lstm based sentiment analysis for cryptocurrency prediction. In: Database systems for advanced applications: 26th International Conference, DASFAA 2021, Taipei, Taiwan, April 11–14, 2021, Proceedings, Part III 26, Springer, pp 617–621
- Hutto C, Gilbert E (2014) Vader: a parsimonious rule-based model for sentiment analysis of social media text. *Proc Int AAAI Conf Web Social Media* 8(1):216–225
- Hyndman RJ, Athanasopoulos G (2018) Forecasting: principles and practice. OTexts
- Ji Q, Bouri E, Lau CKM et al (2019) Dynamic connectedness and integration in cryptocurrency markets. *Int Rev Financ Anal* 63:257–272
- Kannengießer N, Lins S, Dehling T et al (2020) Trade-offs between distributed ledger technology characteristics. *ACM Comput Surv (CSUR)* 53(2):1–37
- Khedr AM, Arif I, El-Bannany M et al (2021) Cryptocurrency price prediction using traditional statistical and machine-learning techniques: a survey. *Intell Syst Account Finance Manage* 28(1):3–34
- Lamon C, Nielsen E, Redondo E (2017) Cryptocurrency price prediction using news and social media sentiment. *SMU Data Sci Rev* 1(3):1–22
- Mell P, Yaga D (2022) Understanding stablecoin technology and related security considerations. Tech. rep, National Institute of Standards and Technology
- Pano T, Kashef R (2020) A complete Vader-based sentiment analysis of bitcoin (BTC) tweets during the era of COVID-19. *Big Data Cognitive Comput* 4(4):33
- Phillips RC, Gorse D (2018) Cryptocurrency price drivers: wavelet coherence analysis revisited. *PLoS ONE* 13(4):e0195200
- Raju S, Tarif AM (2020) Real-time prediction of bitcoin price using machine learning techniques and public sentiment analysis. arXiv preprint [arXiv:2006.14473](https://arxiv.org/abs/2006.14473)
- Rouhani S, Abedin E (2020) Crypto-currencies narrated on tweets: a sentiment analysis approach. *Int J Ethics Syst* 36(1):58–72

- Saleemi J (2021) Covid-19 uncertainty and bitcoin market, linking the liquidity cost to the cryptocurrency yields. *Finance Markets Valuation* 7(1):1–11
- Saleemi J (2023) Microblogging perceptive and pricing liquidity: exploring asymmetric information as a risk determinant of liquidity in the pandemic environments. *Econ Anal Lett* 2(1):1–9
- Seroyizhko P, Zhexenova Z, Shafiq MZ, et al (2022) A sentiment and emotion annotated dataset for bitcoin price forecasting based on reddit posts. In: *Proceedings of the fourth workshop on financial technology and natural language processing*, pp 50–56
- Shehzad K, Bilgili F, Zaman U et al (2021) Is gold favourable than bitcoin during the COVID-19 outbreak? comparative analysis through wavelet approach. *Resour Policy* 73:102163
- Shen D, Urquhart A, Wang P (2019) Does twitter predict bitcoin? *Econ Lett* 174:118–122
- Sovbetov Y (2018) Factors influencing cryptocurrency prices: evidence from bitcoin, ethereum, dash, bitcoin, and monero. *J Econ Financ Anal* 2(2):1–27
- Stosic D, Stosic D, Ludermitr TB et al (2018) Collective behavior of cryptocurrency price changes. *Physica A* 507:499–509
- Tschorsch F, Scheuermann B (2016) Bitcoin and beyond: a technical survey on decentralized digital currencies. *IEEE Commun Surv Tutor* 18:2084–2123
- Valencia F, Gómez-Espinoza A, Valdés-Aguirre B (2019) Price movement prediction of cryptocurrencies using sentiment analysis and machine learning. *Entropy* 21(6):589
- Vidal-Tomás D (2023) Blockchain, sport and fan tokens. *J Econ Stud*
- Vidal-Tomás D (2022) Which cryptocurrency data sources should scholars use? *Int Rev Financ Anal* 81:102061
- Wirawan IM, Widiyaningtyas T, Hasan MM (2019) Short term prediction on bitcoin price using arima method. In: *2019 international seminar on application for technology of information and communication (iSemantic)*, IEEE, pp 260–265
- Wolk K (2020) Advanced social media sentiment analysis for short-term cryptocurrency price prediction. *Expert Syst* 37(2):e12493
- Xia Y, Sang C, He L et al (2023) The role of uncertainty index in forecasting volatility of bitcoin: fresh evidence from Garch-midas approach. *Financ Res Lett* 52:103391
- Yousaf I, Ali S (2020) Discovering interlinkages between major cryptocurrencies using high-frequency data: new evidence from COVID-19 pandemic. *Financial Innovation* 6(1):1–18
- Zhang W, Wang P, Li X et al (2018) The inefficiency of cryptocurrency and its cross-correlation with DOW jones industrial average. *Physica A* 510:658–670

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.