Research

DLCP²F: a DL-based cryptocurrency price prediction framework

Abdussalam Aljadani[1]

Abstract

Cryptocurrencies are distributed digital currencies that have emerged as a consequence of financial technology advancement. In 2017, cryptocurrencies have shown a huge rise in their market capitalization and popularity. They are now employed in today's financial systems as individual investors, corporate firms, and big institutions are heavily investing in them. However, this industry is less stable than traditional currency markets. It can be affected by several legal, sentimental, and technical factors, so it is highly volatile, dynamic, uncertain, and unpredictable, hence, accurate forecasting is essential. Recently, cryptocurrency price prediction becomes a trending research topic globally. Various machine and deep learning algorithms, e.g., Neural Networks (NN), Gated Recurrent Unit (GRU), Long Short-Term Memory (LSTM), and Bidirectional LSTM (BiLSTM) were utilized to analyze the factors influencing the prices of the cryptocurrencies and accordingly predict them. This paper suggests a five-phase framework for cryptocurrency price prediction based on two state-of-the-art deep learning architectures (i.e., BiLSTM and GRU). The current study uses three public real-time cryptocurrency datasets from "Yahoo Finance". Bidirectional Long Short-Term Memory and Gated Recurrent Unit deep learning-based algorithms are used to forecast the prices of three popular cryptocurrencies (i.e., Bitcoin, Ethereum, and Cardano). The Grid Search approach is used for the hyperparameters optimization processes. Results indicate that GRU outperformed the BiLSTM algorithm for Bitcoin, Ethereum, and Cardano, respectively. The lowest RMSE for the GRU model was found to be 0.01711, 0.02662, and 0.00852 for Bitcoin, Ethereum, and Cardano, respectively. Experimental results proved the significant performance of the proposed framework that achieves the minimum MSE and RMSE values.

Keywords Bidirectional Long Short-Term Memory · Cryptocurrency · Deep learning · Gated Recurrent Unit · Recurrent Neural Network · Time series

1 Introduction

Traditionally, economic systems depend only on third-party financial organizations (e.g., banks). These organizations are the intermediates between the parties, having complete control over the transactions and process of exchanging funds. These traditional systems allow limited money transactions and lack security, trust, flexibility, and transparency [1]. To address these issues, a system that can minimize the role of financial intermediaries. Satoshi Nakamoto, a researcher, wrote a paper [2] in 2008, describing the notion of peer-to-peer (P2P) electronic payment transmission without the involvement of any intermediary financial organizations. In financial systems, cryptocurrency is a decentralized virtual currency [3]. Cryptography is used to secure this currency, making it impossible to be double-spent or faked [4]. A central authority (i.e., central banks) is not used to issue cryptocurrencies. It is created using the blockchain technology [5]. This

✉ Abdussalam Aljadani, ajadani@taibahu.edu.sa | [1]Department of Management, College of Business Administration in Yanbu, Taibah University, 41411 Al-Madinah Al-Munawarah, Kingdom of Saudi Arabia.

Springer

technology is incredibly complicated and seeks to store data in a way that makes hacking, altering, or defrauding the system impossible or difficult. The blockchain is composed of two essential modules, namely, transaction and block. The transaction defines the participant's initiated action, while the block records the transaction and additional details (e.g., creation timestamp and correct sequence) [6].
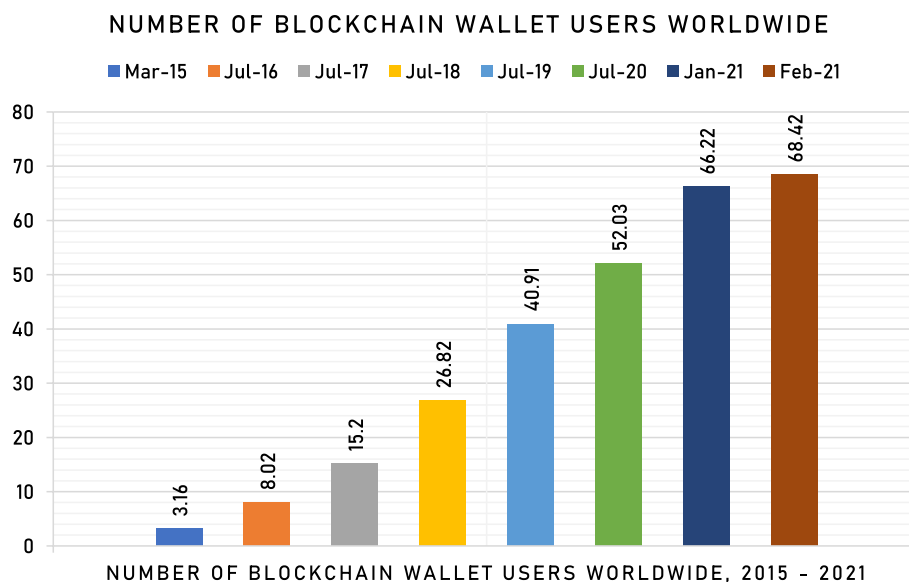
For more than 2 years, Bitcoin was the first and only blockchain-based cryptocurrency [7]. Beginning from the first cryptocurrency suggested by Satoshi Nakamoto, more than 5200 cryptocurrencies, such as Bitcoin, Ethereum, Cardano, Ripple, Monero, Stellar, Litecoin, and Dash, are being traded nowadays [8]. Because of its essential nature of mixing monetary units and encryption technology, cryptocurrencies have recently gained a lot of attention in the domains of cryptography, computer science, and economics [7]. Generally, cryptocurrencies are grouped into three major categories (i.e., currency, platform, and application). The currency domain serve as an exchange medium (i.e., payment method). Platform cryptocurrencies enable the development of a wide range of blockchain-based apps. Finally, cryptocurrencies classified as application domains are used in specific sectors [1].

Since the cryptocurrency's inception, there has been a significant increase in usage, particularly in the previous 5 years. According to Blockchain.com, the the number of blockchain wallet users had increased from 2015 to 2021, as illustrated in Fig. 1 [9, 10]. Bitcoin went from having no value in 2008 to reaching the highest recorded price of $20,089.00 in 2017. Since then, the price of bitcoin has not fallen below $3000. In mid-April, 2021, Bitcoin prices reached all-time highs of more than $60,000 as Coinbase (i.e., a cryptocurrency exchange) went public [11]. On Nov. 10, 2021, Bitcoin again reached an all-time high of $68,789 before closing at 64,995$ [12]. On June 13, 2021, the prices of the cryptocurrencies dropped, thus, Bitcoin price lowered below $23,000 for the first time since December 2020 [13]. The total market capitalization of all cryptocurrencies was approximately 19 billion USD as of February 2017. According to [14], the top 15 currencies amounted to more than 97% of the market, while seven accounted for 90% of market capitalization [14].

As mentioned before, cryptocurrencies gained broad market acceptance and accelerated development. Many financial institutions have included cryptocurrency-related assets in their trading strategies. Cryptocurrency trading is concerned with the act of selling and buying cryptocurrencies to make a profit. Kyriazis [15] surveyed the predictability of the pricing of cryptocurrencies. So, the efficient market hypothesis is rejected and speculation is possible through trading. Additionally, in Fang et al. [16], a comprehensive survey of cryptocurrency trading research, covering various aspects (e.g., cryptocurrency trading systems, prediction of volatility and return, and technical trading) was proposed. Makarova and Schoar [17] studied the arbitrage and price formation in the market of cryptocurrencies. Chava et al. [18] used Google's Search Volume Index as a proxy for retail investor attention. They found that celebrity endorsements of crypto tokens, initial coin offerings (ICOs), non-fungible tokens (NFTs), and crypto platforms providing high yields to investors cause greater attention from areas concerning with higher lottery sales per capita.

The bitcoin price has gained the interest of scholars all over the world. Unfortunately, cryptocurrency prices are volatile and dynamic. It is determined by several elements (for example, popularity, mining difficulty, transaction cost, market trends, price of alternate coins, sentiments, stock markets, and some legal issues) [19]. Additionally, small

Fig. 1 Global blockchain wallet usership from 2015 to 2021

NUMBER OF BLOCKCHAIN WALLET USERS WORLDWIDE

■ Mar-15 ■ Jul-16 ■ Jul-17 ■ Jul-18 ■ Jul-19 ■ Jul-20 ■ Jan-21 ■ Feb-21

NUMBER OF BLOCKCHAIN WALLET USERS WORLDWIDE, 2015 – 2021

cryptocurrencies with a small market share become a source of shocks that can benefit or harm other cryptocurrencies. These factors make the prices of cryptocurrencies unstable, change rapidly over time, and are difficult to be predicted. As a result, scams, suspected hacks, or other hidden problems lead to dramatically drop in cryptocurrency prices [20]. For example, on June 26th, 2019, more than 10% of the Bitcoin price was lost in a few minutes because of the crashes of the Coinbase digital exchange. Consequently, price prediction has become a critical responsibility for researchers [21]. Clients use thousands of coins around the world. The focus of this study is on three of the most popular cryptocurrencies (i.e., Bitcoin, Ethereum, and Cardano). Bitcoin is leading at $960.79 billion, followed by Ethereum at $189.98 billion, Binance Coin at $39.91 billion, Tether at $35.96 billion, and Cardano at $33.36 billion in market capitalization [22]. It is worth mentioning that these values are regularly updated as the market changes.

Different machine and deep learning architectures have been proposed to perform the task of cryptocurrency price prediction (e.g., support vector machines SVM [23], Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM) [24], and Deep Neural Networks (DNN) [25]). This study proposes a deep learning-based system to predict cryptocurrency values, employing 2 RNN algorithms, specifically, Bidirectional Long Short-Term Memory (BiLSTM) and Gated Recurrent Unit (GRU). Three datasets, namely, Bitcoin USD (BTC-USD) [26], Ethereum USD (ETH-USD) [27], and Cardano USD (ADA-USD) [28], including past cryptocurrency prices, are utilized for training the algorithms to anticipate cryptocurrency values. Figure 2 shows the flow for a cryptocurrency price prediction system.

The cryptocurrency price is suffering from instability and dynamism; hence, prediction is a crucial task. In the current study, deep learning-based algorithms (i.e., Bidirectional LSTM and GRU), are used to predict the prices of three of the most used cryptocurrencies (i.e., Bitcoin, Ethereum, and Cardano). This method seeks to uncover hidden patterns in data, integrate them, and generate more accurate forecasts. The current study's contributions can be summarised as follows:

– Utilizing deep learning-based algorithms (i.e., BiLSTM and GRU) to predict the three cryptocurrencies' prices (i.e., Bitcoin, Ethereum, and Cardano).
– Utilizing the Grid Search approach for the hyperparameters optimization processes.
– Evaluating the performance of the proposed models utilizing evaluation metrics such as MSE, RMSE, MAE, and MAPE.

The rest of this paper is structured as follows: Sect. 2 introduces the related work in this field. Section 3 discusses the methodology, data acquisition phase, data pre-processing phase, classification and optimization phase, and performance evaluation phase. Section 4 explores the experimental results. Section 5 presents the current study limitations. Section 5 concludes the paper and presents the future work.

Fig. 2 The flow for a cryptocurrency price prediction system
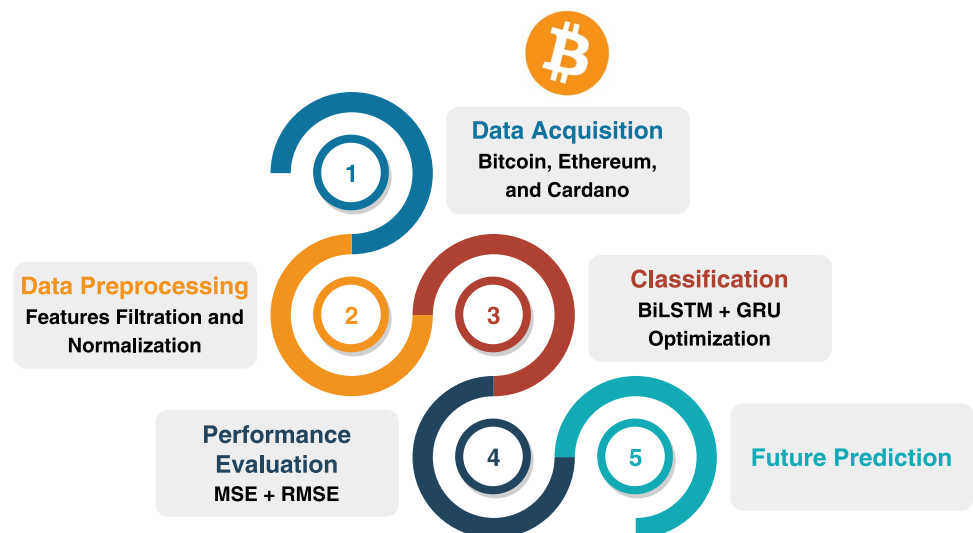
2 Related studies

The prediction of the cryptocurrency price is a time series problem [29]. A time series is a sequence of variable measurements made over time. Usually, these measurements are made at equally spaced times. Time series problems are completely different from other problems because of the following reasons. First, time series problems are time-dependent. Hence, the basic hypothesis of a linear regression model can not be held in this case. Secondly, most of these problems have some form of seasonality trends along with a decreasing or increasing trend (i.e. particular time frame has specific variations) [30]. Time series prediction is used to predict the future values based on the prior observed values over time. Some methods include Autoregressive (AR) [31], Autoregressive Integrated Moving Average (ARIMA) [32], Seasonal Autoregressive Integrated Moving Average (SARIMA) [33], Exponential Smoothing (ES) [34], DeepAR [35], and N-BEATS [36]. Machine learning (ML) models used to solve time series problems include K-Nearest Neighbor (KNN), Classification and Regression Trees (CART), Decision Trees, and Support Vector Regression (SVR), while deep learning (DL) algorithms include Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), XGBoost, AdaBoost, and Convolutional Neural Network (CNN). The most significant advantage of ML-based approaches is that they are computationally inexpensive for implementing online models. DL-based models show several advantages over other predicting models as it not only produces a result that is almost or exactly the same as the actual result but also enhance the result accuracy [37].

Deep learning (DL) is a type of artificial intelligence that uses previous data to predict the future. DL-based models offer several advantages over other prediction models in that they not only produce a result that is nearly or exactly the same as the true result, but they also enhance the result's accuracy [37]. DNN, RNN, LSTM, and Deep Belief Networks are examples of DL.

Derbentseva et al. [38] presented two machine learning (ML) techniques for time-series data forecasting: random forests (RF) and stochastic gradient boosting machine (SGBM). They forecasted the prices of three of most valuable cryptocurrencies: Bitcoin, Ethereum, and Ripple. Their reported results show that the ML ensemble technique can be used to forecast bitcoin values. Three types of recurrent neural network techniques are proposed in Hamayel, and Owda's [39] to predict the prices of three cryptocurrencies (i.e., Bitcoin,itecoin, and Ethereum). Their proposed methods produce accurate forecasts based on the mean absolute percentage error. According to the results, the GRU model outperformed the LSTM and bidirectional LSTM (BiLSTM) models in prediction for the three cryptocurrencies. GRU provides MAPE percentages of 0.2454%, 0.2116%, and 0.8267% for Bitcoin, Litecoin, and Ethereum, respectively.

Pour et al. [40] proposed a hybrid model for Bitcoin price prediction that uses Long Short-Term Memory and Bayesian Optimization. Their model was validated using MSE, RMSE, and NRMSE. Their reported results showed that the proposed model had good predictive power. Patel et al. [1] proposed a hybrid cryptocurrency prediction method based on LSTM and GRU, concentrating on Litecoin and Monero. According to reported results, the proposed algorithm predicted the prices with great accuracy, and it may be used in various applications based on the bitcoin price prediction. Finally, Kim et al. [41] proposed a multiple LSTM based on self-attention. Their proposed model contains numerous LSTM modules for on-chain variable groups as well as the attention mechanism of the prediction model. Experiments based on real Bitcoin price data were utilized to assess the suggested framework, which resulted in MAE, RMSE, MSE, and MAPE values of 0.3462, 0.5035, 0.2536, and 1.3251, respectively.

Miura et al. [42] utilized several ML models to predict future values based on prior samples, including multilayer perceptron, GRU, LSTM, SVM, and regression. The results demonstrated that the proposed system accurately predicts prices, thus, the technology might be utilized to anticipate prices for other cryptocurrencies. Yiying and Yeze [43] proposed a sophisticated artificial intelligence approach of fully connected Artificial Neural Network (ANN) and LSTM to study Bitcoin, Etherum, and Ripple price fluctuations. Their stated results revealed that ANN was more dependent on long-term history. LSTM, on the other hand, depended more on short-term dynamics, indicating that it extracts information from historical memory more efficiently than ANN.

A machine learning-based strategy was proposed in SHAHBAZI and BYUN [44] to perform cryptocurrency price predictions (i.e., Litecoin and Monero) for a financial institution. Their proposed method included a reinforcement Learning algorithm to assess and forecast cryptocurrency prices and a blockchain infrastructure to ensure a secure transaction environment. Their findings revealed that the proposed approach predicted prices more correctly than previous state-of-the-art algorithms. KALARIYA et al. [45] proposed a stochastic neural network model for cryptocurrency price prediction based on random walk theory. The Multi-Layer Perceptron (MLP) and LSTM models were trained, and the experimental results for Bitcoin, Ethereum, and Litecoin were released. The model outperformed the deterministic models, according to their findings.

Uras et al. [46] have used statistical methodologies and ML algorithms to forecast the prices of Bitcoin, Litecoin, and Ethereum. The Simple Linear Regression was used to forecast univariate series using only price data, whereas the Multiple Linear Regression was used to forecast multivariate series utilizing both price and volume data. Multilayer Perceptron and Long Short-Term Memory were the deep learning methods employed. In addition, a deep-learning-based hybrid model employing GRU and LSTM was proposed by TANWAR et al. [47] to forecast the price of Litecoin and Zcash. Their proposed model has the potential to be utilized in real-time applications. According to their published results, the model forecasted prices with great accuracy compared to existing models.

Related studies comparison: Table 1 presents a comparison between the mentioned literature and the current study. They are organized in descending order.

Research Gap: In the light of the mentioned literature, the studies addressing Bitcoin are more than those that address the other altcoins as seen in Table 1. For traders, security and privacy are critical problems while trading to gain more trust. Hence, the possibility of using ML- and DL-based algorithms to address the anonymity, security, and privacy level of other cryptocurrencies are needed to be explored. ML Ensemble algorithms are not examined that much in the field of price prediction of cryptocurrencies. Furthermore, not enough focus was performed on optimizing ML techniques to enhance the accuracy.

3 The suggested DLCP^{2}F

As mentioned earlier, Cryptocurrency popularity increased in 2017 as its market value increased rapidly for several months in a row. Prices peaked at around $800 billion in January 2018 [48]. The current study suggests a framework for the Cryptocurrency Price Prediction that utilizes state-of-the-art deep learning architectures. The proposed framework is presented in five phases: (1) data acquisition, where the data is acquired from a public source, (2) data preprocessing phase to prepare the dataset for the next phase, (3) classification phase to learn and optimize the models, (4) performance evaluation phase, and (5) future prediction phase. It is summarized graphically in Fig. 3.

3.1 Data acquisition phase

The current study depends on three public real-time cryptocurrency datasets retrieved from "Yahoo Finance". Their "Historical Prices" daily records are retrieved until the date "August 9, 2022". The first dataset is named Bitcoin USD (BTC-USD) and consists of 2885 daily records from "September 17, 2014" [26]. The second dataset is named Ethereum USD (ETH-USD) and consists of 1735 daily records from "November 9, 2017" [27]. The third dataset is named Cardano USD (ADA-USD) and consists of 1735 daily records from "November 9, 2017" [28]. The three datasets consist of 7 columns: "Date", "Open", "High", "Low", "Close", "Adj Close, and "Volume". The "Open" and "Close" prices represent the currency market's open and closed prices on a specific "Date". The "High" and "Low" prices represent the currency market's maximum and minimum prices on a specific "Date". The "Volume" is the amount of money in circulation on a particular "Date". Table 2 summarizes the details of the datasets. Figure 4 shows the close prices summarization for the three datasets. From it, the close prices are low in the initial period then takes an incremental slope. After that, the prices change but in the high region. From that, the datasets show a recognizable challenge to forecasting the cryptocurrency prices using the given trading features. Statistics on the three datasets are reported in Table 3. Skew is concerned with the measurement of symmetry. A distribution (i.e., dataset) is symmetric if the right and left sides look the same from the center point. Kurtosis measures whether the data is heavy- or light-tailed when compared to a normal distribution. Thus, datasets with high kurtosis (i.e., heavy-tailed) tend to contain outliers. while datasets with low kurtosis (i.e., light-tailed) lack outliers [49, 50]. Table 3 shows that the last column has a very high standard deviation compared to other columns.

3.2 Data preprocessing phase

The data is arranged chronologically and recorded at regular intervals (i.e., 1 day). It is considered a time series data that requires special treatment with the used models (i.e., BiLSTM and GRU). The first step is to filter the features. The current study uses the "Open", "Close", "Adj Close", and "Volume" features while the other features are dropped. As the target of the current study is to predict the price of the cryptocurrency, it just depends on the selected columns. After that, the features will be squished using the min-max scaler (Eq. 1) where X_i is the input record and X_o is the scaled output record. This will facilitate the optimization algorithm to converge faster.

Table 1 Comparison between related studies and the current study

Publication	Year	Approach	Used metrics	Cryptocurrencies
Miura et al. [42]	2019	Multilayer Perceptron, GRU, LSTM, SVM, and regression	MSE and RMSE	Bitcoin
Yiying and Yeze [43]	2019	Fully-connected Artificial Neural Network (ANN) and LSTM	MSE and Pearson correlation	Bitcoin, Etherum, and Ripple
Patel et al. [1]	2020	Hybrid algorithm using LSTM and GRU	MSE, RMSE, MAE, and MAPE	Litecoin and Monero
KALARIYA et al. [45]	2020	Multi-Layer Perceptron and Long Short-Term Memory models	MAPE, MAE, RMSE, and MSE	Bitcoin, Ethereum, and Litecoin
Uras et al. [46]	2020	Multilayer Perceptron and Long short-term memory	MAPE and relative RMSE	Bitcoin, Litecoin, and Ethereum
Hamayel and Owda [39]	2021	GRU, LSTM, and bidirectional LSTM	MAPE and RMSE	Bitcoin, Litecoin, and Ethereum
SHAHBAZI and BYUN [44]	2021	A reinforcement Learning algorithm with blockchain framework	MSE, RMSE, MAE, and MAPE	Litecoin and Monero
TANWAR et al. [47]	2021	Hybrid model using GRU and LSTM	MSE	Litecoin and Zcash
Derbentseva et al. [38]	2021	Two ML algorithms (i.e, random forests and stochastic gradient boosting machine)	MAPE and RMSE	Bitcoin, Ethereum, and Ripple
Kim et al. [41]	2022	Self-attention-based multiple long short-term memory	MAE, RMSE, MSE, and MAPE	Bitcoin
Pour et al. [40]	2022	Hybrid model using Long Short-Term Memory and Bayesian Optimization	MSE, RMSE, and NRMSE	Bitcoin
Current study	2022	Bidirectional LSTM and GRU	MSE, RMSE, MAE, MAPE, and R2 Score	Bitcoin, Ethereum, and Cardano

Fig. 3 The suggested framework for the cryptocurrency price prediction
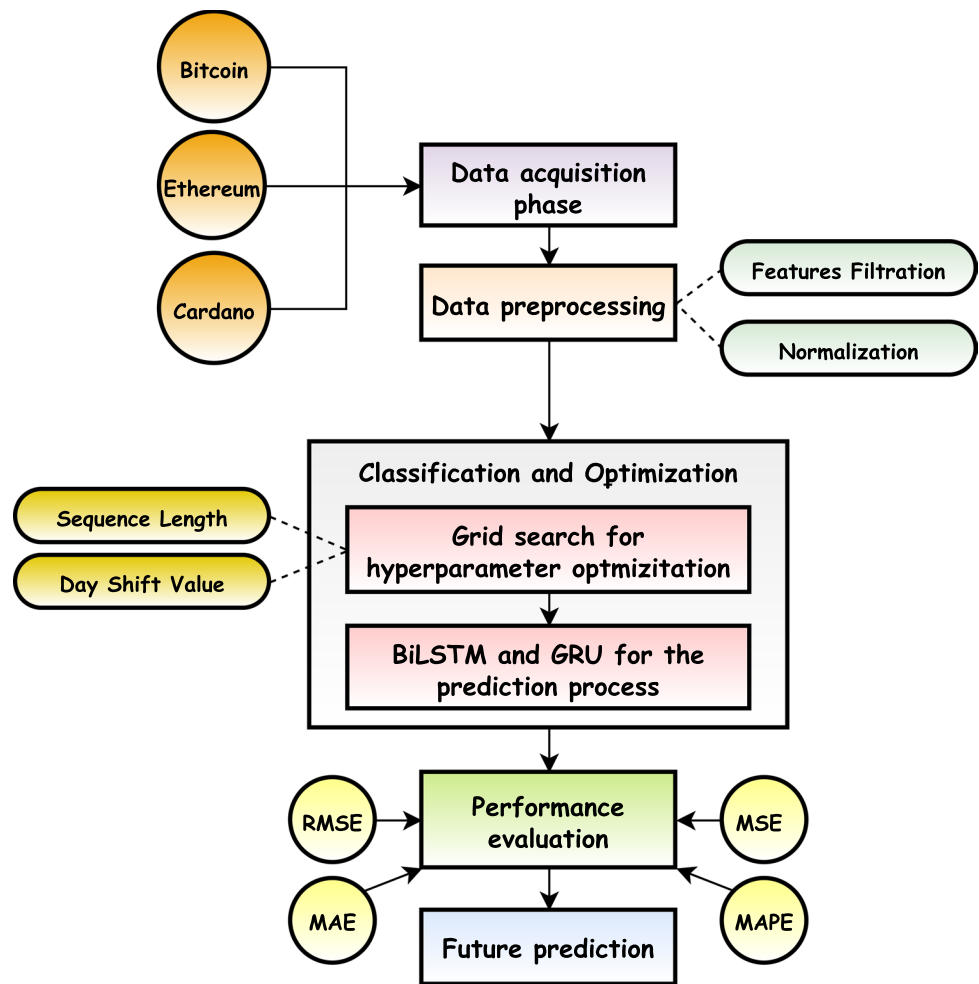
**Table 2** The different used datasets in the current study

Dataset	No. of records	From	To
Bitcoin USD (BTC-USD) [26]	2781	September 17, 2014	August 9, 2022
Ethereum USD (ETH-USD) [27]	1632	November 9, 2017	August 9, 2022
Cardano USD (ADA-USD) [28]	1632	November 9, 2017	August 9, 2022

$$X_o = \frac{X_i - \min(X_i)}{\max(X_i) - \min(X_i)} \tag{1}$$

The last step in the preprocessing phase is to make data sequences. Building sequences begin with creating a sequence of a specific length (i.e., window size) at position 0. Then a new sequence is created by shifting one position to the right. This is continued until all of the available positions have been utilized. Finally, the inputs and outputs are created using the same approach. The only difference between the inputs and outputs is that specified value shifts the outputs, namely "days shift".

The models are controlled by two variables (i.e., days shift and sequence length). The days shift is concerned with the time gap between input (i.e., features) and output (i.e., close price). For example, if the value of the days shift is 5 and the first 10 days were taken as an input, the output will be from the 5th to the 15th day. How this will affect the prediction? When 10 days are entered as an input (i.e., from the 1st to the 10th day) and the value of the days shift is 3, it is supposed to predict the output from the 3rd to the 13th days. Since the goal is to predict future data, the last three elements on the predicted output values are the future values. The lowest value of the days shift is 1, hence, the future data of the next day will be predicted along with with the previous days. The sequence length is concerned with how

Fig. 4 The close prices summarization for the three datasets (i.e., BTC-USD, ETH-USD, and ADA-USD) from the initiated dated until "August 9, 2022"
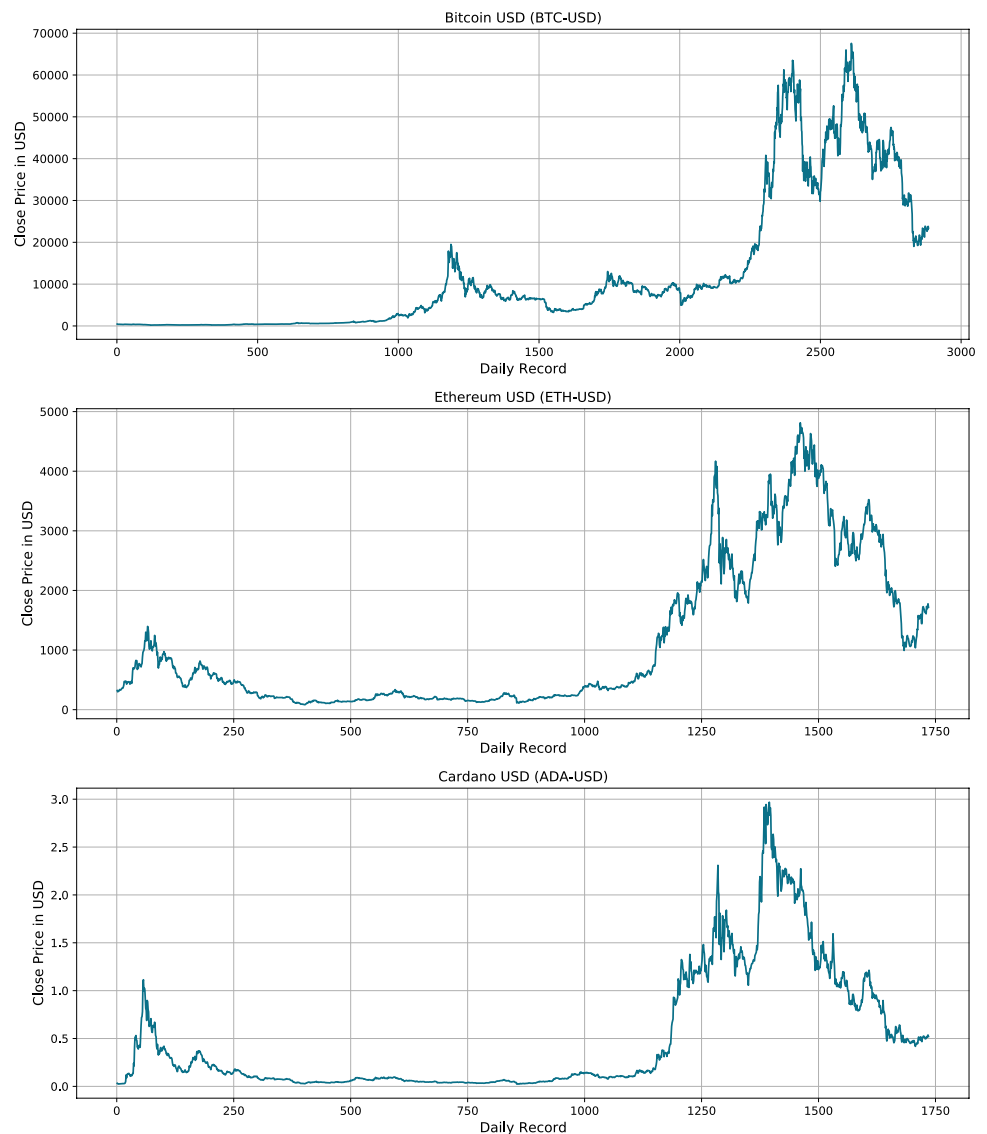
the data are passed to the model. When the value of the sequence length is 10, the input will be divided into groups, each group consisting of 10 records, and each group is treated as one record. For example, the input X consists of 100 records, and the sequence length is set to 25, hence, four sequences will be generated, and each is treated as one record by the model. When the value sequence length is higher, the performance will be better. This happens because each record contains more information, however, the complexity of the training time will be increased. Thus, the current work aims to determine the best value for both days shift and sequence length through the use of the grid search approach.

Figure 5 shows a graphical sample of the training and testing inputs and outputs process. In this example, a 1000-record dataset is split into 900 for training and 100 for testing where the days shift value is set to 5. The training inputs start from 0 while the training outputs start from 5 (i.e., the days shift value). Hence, the input X is the first 850 records and the output Y is the last 850 records. This means that the prediction will be the forecast for the next 5 days based on the current inputs.

3.3 Classification and optimization phase

The current phase works on creating two state-of-the-art deep learning models (i.e., BiLSTM and GRU) and optimizing them based on the input data. Long-Short Term Memory (LSTM) works by allowing each internal layer to use certain gates to access information from both previous and current layers. After going through several gates (for example, they forget and input gates) and many activation functions, the data is delivered via the LSTM cells (such as the Tanh function

Table 3 Statistics on the used datasets in the current study

Dataset	Column	Mean	Min	Max	Standard deviation	Skew	Kurtosis
Bitcoin USD (BTC-USD) [26]	Open	12547.40934	176.897003	67549.73438	16521.66801	1.57690	1.28320
	High	12873.51768	211.731003	68789.625	16948.13805	1.57007	1.24927
	Low	12183.19076	171.509995	66382.0625	16028.0046	1.58399	1.32005
	Close	12553.72666	178.102997	67566.82813	16517.67463	1.57544	1.27915
	Adj Close	12553.72666	178.102997	67566.82813	16517.67463	1.57544	1.27915
	Volume	15,592,140,812	5,914,570	3.50968E+11	19,876,925,724	2.89389	29.46061
Ethereum USD (ETH-USD) [27]	Open	1106.827381	84.279694	4810.071289	1232.748038	1.21861	0.21451
	High	1143.456172	85.342743	4891.70459	1270.587528	1.20543	0.17437
	Low	1065.079037	82.829887	4718.039063	1188.849134	1.23450	0.26245
	Close	1107.384122	84.308296	4812.087402	1232.181209	1.21579	0.20645
	Adj Close	1107.384122	84.308296	4812.087402	1232.181209	1.21579	0.20645
	Volume	12,890,400,981	621,732,992	84,482,912,776	11,033,194,348	1.60069	3.95134
Cardano USD (ADA-USD) [28]	Open	0.498241584	0.023954	2.96639	0.648616123	1.55453	1.62294
	High	0.520439048	0.025993	3.099186	0.676057896	1.54390	1.57649
	Low	0.474530443	0.01913	2.907606	0.619377364	1.57031	1.68900
	Close	0.498527944	0.023961	2.968239	0.648597112	1.55506	1.62681
	Adj Close	0.498527944	0.023961	2.968239	0.648597112	1.55506	1.62681
	Volume	1,195,124,546	2,930,550	19,141,980,288	2,119,527,238	3.23058	13.56876

Fig. 5 A graphical sample of the process for the training and testing inputs and outputs


and ReLU function). The main advantage is that each LSTM cell can recall patterns for a specific time. It is important to note that LSTM can remember important information while forgetting irrelevant information. Furthermore, an LSTM's default behavior is remembering information for a long time [51].

Bidirectional LSTM (BiLSTM) is a recurrent neural network (RNN) that is commonly used to process natural language. In contrast to typical LSTM, the input flows in both directions and can use data from both sides. In short, BiLSTM adds another LSTM layer, reversing the data flow. In a nutshell, the input sequence flows backward in the additional LSTM layer. The outputs from both LSTM layers are combined in various methods, including average, sum, multiplication, and concatenation [52]. The suggested BiLSTM network consists of: (1) an input LSTM layer with several units equal to the sequence length and Tanh activation function, (2) a 50% dropout layer, (3) a BiLSTM layer with 256 units, (4) another 50% dropout layer, and (5) an output layer with a linear activation function. Figure 6 presents the hierarchy of the BiLSTM model using a sequence length of 50. The "None" keyword means to accept any value.

GRU (Gated Recurrent Unit) is an RNN that seeks to tackle the vanishing gradient problem. GRU might be regarded as a variant of the LSTM. It employs the so-called update gate and reset gate to overcome the vanishing gradient problem of a regular RNN. Two vectors determine what data should be sent to the output. They are unique in that they can be trained to retain knowledge from the past without having to wash it away over time or delete information unrelated to the forecast [53]. The suggested GRU network consists of: (1) an input GRU layer with a number of 50 units and Tanh activation function, (2) a 25% dropout layer, (3) another GRU layer with 100 units and Tanh activation function, (4) another 25% dropout layer, and (5) an output layer with a linear activation function. Figure 7 presents the hierarchy of the GRU model using a sequence length of 50. The "None" keyword means to accept any value.

For both networks, the AdaGrad parameters' optimizer [54] is used. It has several advantages: (1) it eliminates the need to manually regulate the learning rate, (2) it achieves faster and more reliable convergence than the basic SGD when the weight scaling is unequal, and (3) it is not sensitive to the size of the step. It uses the update rule in Eq. 2 where η is the

Fig. 6 A graphical presentation of the hierarchy of the BiLSTM model using a sequence length of 50. The graph is generated from TensorFlow
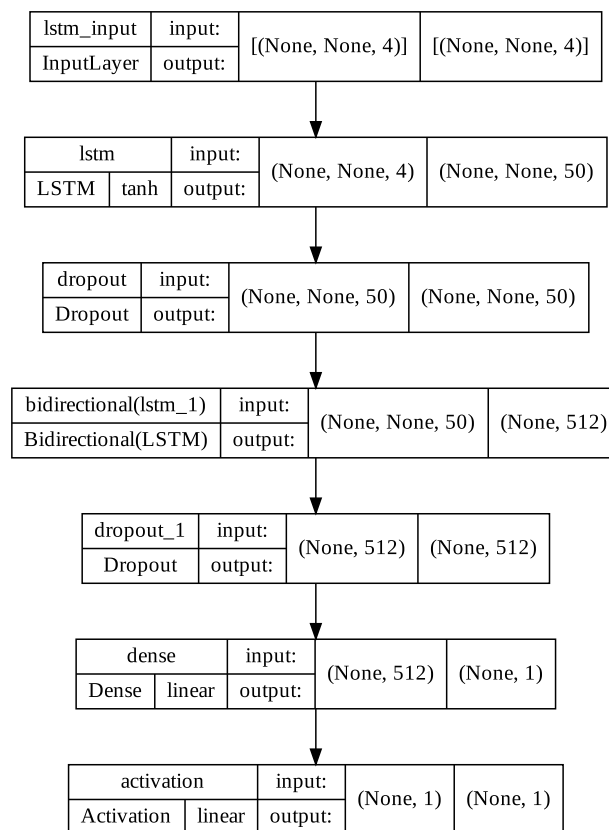
Fig. 7 A graphical presentation of the hierarchy of the GRU model using a sequence length of 50. The graph is generated from TensorFlow
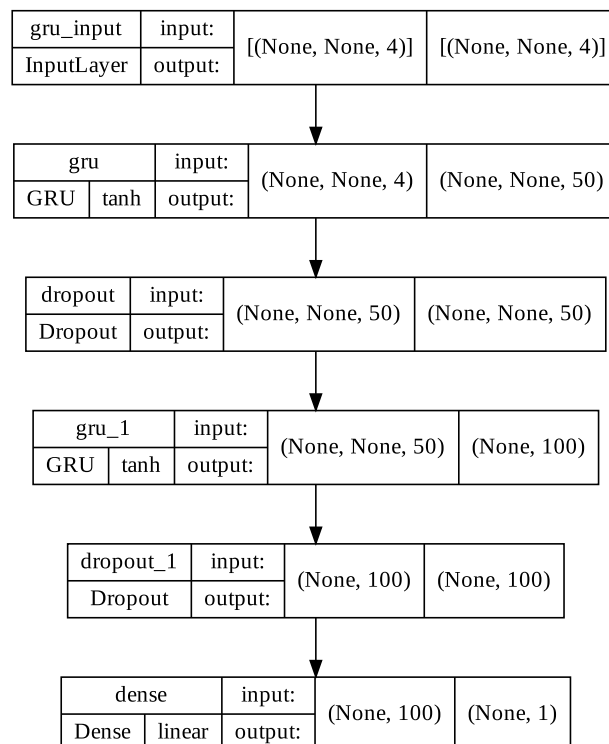
learning rate, g_t is the partial derivative of the objective function, and G_t is a diagonal matrix. ε is added to avoid any divisions by zeros. A model's hyperparameter is a model's feature that is independent of the model and whose value cannot be calculated from data. Before the learning process can begin, the hyperparameter's value must be determined. The grid search (GS) is used to identify the model's optimal hyperparameters that produce the most optimistic predictions [55]. The target is the GS approach to find the best combination between the sequence length and day shift value. The sequence length range is [10, 20, 30, 40, 50] and the days shift range is [1, 2, 3, 4, 5].

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{G_t + \varepsilon}} \times g_t \tag{2}$$

3.4 Performance evaluation phase

For each epoch, the performance is evaluated. The current study applies 100 epochs with the early stopping of 10. The dataset is split into training, testing, and validation. The testing size is set to 100. The validation size is set to 10% of the remaining data. The mean squared error is used as the loss and evaluation function. The less the value, the better the model. It is equated in Eq. 3 where N is number of records, y_i is the actual value, and y_i^* is the predicted value. Also, the root mean squared error, mean absolute error, mean absolute percentage error, and R2 score are calculated, and their equations are shown from Eqs. 4 to 6. It is worth mentioning that both Eqs. 3 and 4 can be derived from one another.

$$\text{MSE} = \frac{1}{N} \times \sum_{i=1}^{N} (y_i - y_i^*)^2 \tag{3}$$

$$\text{RMSE} = \sqrt{\text{MSE}} = \sqrt{\frac{1}{N} \times \sum_{i=1}^{N} (y_i - y_i^*)^2} \tag{4}$$

$$\text{MAE} = \frac{1}{N} \times \sum_{i=1}^{N} |y_i - y_i^*| \tag{5}$$

$$\text{MAPE} = \frac{1}{N} \times \sum_{i=1}^{N} \left| \frac{y_i - y_i^*}{y_i} \right| \tag{6}$$

4 Experiments and discussions

The current section reports the executed experiments and their discussions. The experiments' configurations are reported in Table 4.

4.1 The "Bitcoin USD (BTC-USD)" dataset experiments

Table 5 shows reported performance metrics for the "Bitcoin USD (BTC-USD)" dataset using the two models (i.e., GRU and BiLSTM). The minimum MSE, RMSE, MAE, MAPE, and R2 are 0.00029, 0.01711, 0.02214, 0.07036, and − 2.89528 respectively for the GRU and 0.00210, 0.04582, 0.03358, 0.10942, and − 10.71676 respectively for the BiLSTM. The maximum MSE, RMSE, MAE, MAPE, and R2 are 0.00264, 0.05136, 0.07599, 0.21793, and 0.74536 respectively for the GRU and 0.00793, 0.08907, 0.08577, 0.25511, and 0.44163 respectively for the BiLSTM. For the GRU model, the best sequence length and days shift concerning the MSE (and RMSE) are 50 and 1 respectively, concerning the MAE are 50 and 4 respectively, and concerning the MAPE are 50 and 4 respectively. For the BiLSTM model, the best sequence length and days shift concerning the MSE (and RMSE) are 50 and 1 respectively, concerning the MAE are 50 and 1 respectively, and concerning the MAPE are 50 and 1 respectively. From them, the majority voted sequence length and days shift are are 50 and

**Table 4** The experiments' configurations in the current study

Configuration	Value
Datasets	Described in Table 2
Models	BiLSTM and GRU
Epochs	100
Batch size	10
Sequence length range	[10, 20, 30, 40, 50]
Days shift range	[1, 2, 3, 4, 5]
Parameters optimizer	AdaGrad Optimizer
Hyperparameters optimizer	Grid search (GS)
Early stopping patience	10
Metrics	MSE, RMSE, MAE, MAPE, and R2 Score
Scripting language	Python

Table 5 The reported performance metrics for the "Bitcoin USD (BTC-USD)" dataset using the two models (i.e., GRU and BiLSTM)

Sequence length	Days shift	GRU					BiLSTM				
		MSE	RMSE	MAE	MAPE	R2	MSE	RMSE	MAE	MAPE	R2
10	1	0.00096	0.03094	0.06991	0.19828	0.74536	0.00210	0.04582	0.07519	0.21610	0.44163
10	2	0.00161	0.04011	0.07272	0.20781	0.57209	0.00331	0.05751	0.08013	0.23421	0.12045
10	3	0.00176	0.04198	0.07221	0.20572	0.53137	0.00355	0.05961	0.08208	0.23776	0.05475
10	4	0.00214	0.04631	0.07408	0.21188	0.42967	0.00456	0.06749	0.08577	0.25063	− 0.21153
10	5	0.00254	0.05040	0.07599	0.21793	0.32440	0.00439	0.06623	0.08421	0.24577	− 0.16681
20	1	0.00116	0.03400	0.06532	0.18823	0.65493	0.00364	0.06036	0.07482	0.22167	− 0.08781
20	2	0.00140	0.03744	0.06552	0.18914	0.58163	0.00458	0.06771	0.07942	0.23646	− 0.36864
20	3	0.00166	0.04072	0.06549	0.18886	0.50502	0.00595	0.07712	0.08496	0.25511	− 0.77565
20	4	0.00184	0.04294	0.06426	0.18565	0.44957	0.00540	0.07348	0.08167	0.24341	− 0.61185
20	5	0.00213	0.04618	0.06704	0.19352	0.36340	0.00620	0.07876	0.08502	0.25495	− 0.85191
30	1	0.00121	0.03475	0.05499	0.16153	0.51620	0.00579	0.07610	0.07615	0.23445	− 1.32094
30	2	0.00152	0.03898	0.05640	0.16644	0.39115	0.00538	0.07334	0.07341	0.22504	− 1.15523
30	3	0.00179	0.04227	0.05672	0.16734	0.28418	0.00484	0.06958	0.07020	0.21437	− 0.94027
30	4	0.00222	0.04717	0.05823	0.17303	0.10857	0.00630	0.07936	0.07724	0.23732	− 1.52363
30	5	0.00242	0.04920	0.05854	0.17305	0.03011	0.00722	0.08495	0.08186	0.25221	− 1.89176
40	1	0.00123	0.03512	0.03399	0.10665	− 0.82155	0.00468	0.06843	0.05183	0.16626	− 5.91519
40	2	0.00170	0.04126	0.03628	0.11411	− 1.51382	0.00638	0.07990	0.06227	0.19989	− 8.42895
40	3	0.00198	0.04453	0.03744	0.11769	− 1.92786	0.00569	0.07543	0.05655	0.18122	− 7.40382
40	4	0.00242	0.04916	0.03944	0.12417	− 2.56972	0.00793	0.08907	0.07160	0.22960	− 10.71676
40	5	0.00264	0.05136	0.04016	0.12591	− 2.89528	0.00679	0.08237	0.06208	0.19902	− 9.02049
50	1	0.00029	0.01711	0.02257	0.07210	0.33537	0.00213	0.04616	0.03358	0.10942	− 3.83673
50	2	0.00032	0.01800	0.02220	0.07070	0.26494	0.00215	0.04634	0.03451	0.11250	− 3.87318
50	3	0.00037	0.01927	0.02230	0.07099	0.15753	0.00344	0.05866	0.04419	0.14358	− 6.80990
50	4	0.00041	0.02032	0.02214	0.07036	0.06265	0.00285	0.05339	0.03783	0.12289	− 5.47078
50	5	0.00060	0.02457	0.02400	0.07686	− 0.37024	0.00365	0.06042	0.04272	0.13861	− 7.28534
Min. values		0.00029	0.01711	0.02214	0.07036	− 2.89528	0.00210	0.04582	0.03358	0.10942	− 10.71676
Max. values		0.00264	0.05136	0.07599	0.21793	0.74536	0.00793	0.08907	0.08577	0.25511	0.44163

1 respectively. Figures 8 and 9 summarize the reported RMSE results graphically using BiLSTM and GRU, where the x-axis represents the day shifts range (i.e., [1, 2, 3, 4, 5] and the y-axis represents RMSE values). From them, the GRU and 50-sequence-length report the best RMSE values while the BiLSTM and 50-sequence-length report the best RMSE values.

Fig. 8 A graphical summarization of the reported results using BiLSTM and "Bitcoin USD (BTC-USD)" dataset

GRAPHICAL SUMMARY OF THE RMSE VALUES USING BILSTM + BTC-USD
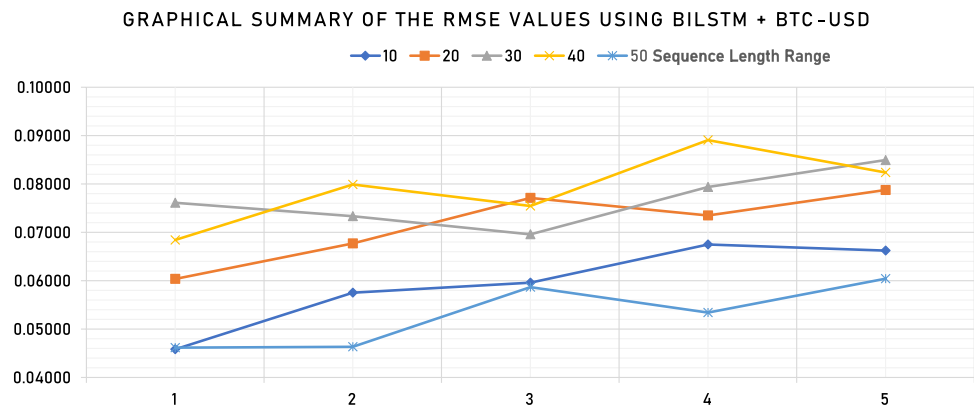
Fig. 9 A graphical summarization of the reported results using GRU and "Bitcoin USD (BTC-USD)" dataset

GRAPHICAL SUMMARY OF THE RMSE VALUES USING GRU + BTC-USD
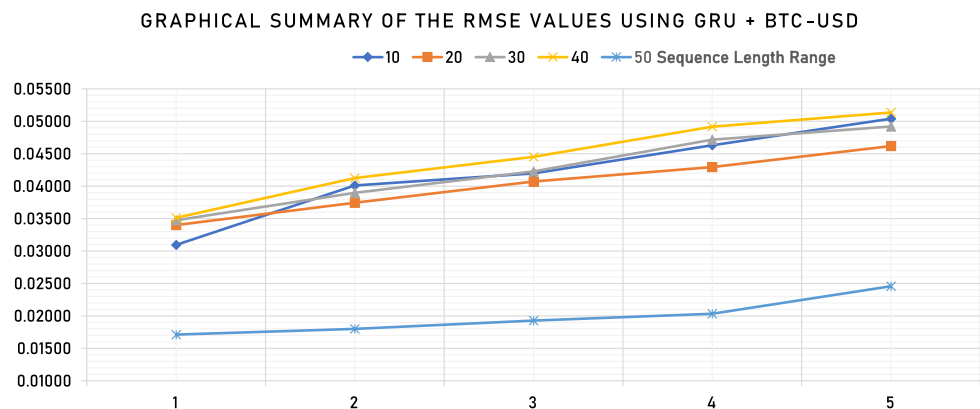
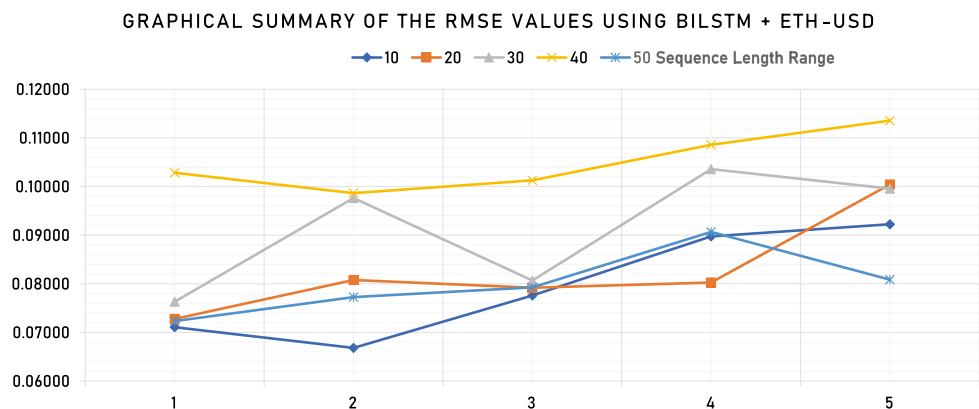## 4.2 The "Ethereum USD (ETH-USD)" dataset experiments

Table 6 shows reported performance metrics for the "Ethereum USD (ETH-USD)" dataset using the two models (i.e., GRU and BiLSTM). The minimum MSE, RMSE, MAE, MAPE, and R2 are 0.00071, 0.02662, 0.05350, 19.67071, and − 0.48171 respectively for the GRU and 0.00446, 0.06681, 0.05710, 0.23379, and − 3.91045 respectively for the BiLSTM. The maximum MSE, RMSE, MAE, MAPE, and R2 are 0.00411, 0.06414, 0.09458, 33.75066, and 0.73784 respectively for the GRU and 0.01289, 0.11355, 0.10970, 0.40123, and 0.11536 respectively for the BiLSTM. For the GRU model, the best sequence length and days shift concerning the MSE (and RMSE) are 50 and 1 respectively, concerning the MAE are 50 and 5 respectively, and concerning the MAPE are 50 and 5 respectively. For the BiLSTM model, the best sequence length and days shift concerning the MSE (and RMSE) are 10 and 2 respectively, concerning the MAE are 50 and 1 respectively, and concerning the MAPE are 50 and 1 respectively. From them, the majority voted sequence length and days shift are are 50 and 1 respectively. Figures 10 and 11 summarize the reported RMSE results graphically using BiLSTM and GRU, where the x-axis represents the day shifts range (i.e., [1, 2, 3, 4, 5] and the y-axis represents RMSE values). From them, the GRU and 50-sequence-length report the best RMSE values while the BiLSTM and 10-sequence-length report the best RMSE values.

4.3 The "Cardano USD (ADA-USD)" dataset experiments

Table 7 shows reported performance metrics for the "Cardano USD (ADA-USD)" dataset using the two models (i.e., GRU and BiLSTM). The minimum MSE, RMSE, MAE, MAPE, and R2 are 0.00007, 0.00852, 0.00971, 0.06256, and − 2.59725 respectively for the GRU and 0.00031, 0.01752, 0.01599, 0.10515, and − 8.22629 respectively for the BiLSTM. The maximum MSE, RMSE, MAE, MAPE, and R2 are 0.00068, 0.02604, 0.02152, 0.13393, and 0.24750 respectively for the GRU and 0.00165, 0.04058, 0.03424, 0.21699, and − 0.20519 respectively for the BiLSTM. For the GRU model, the best
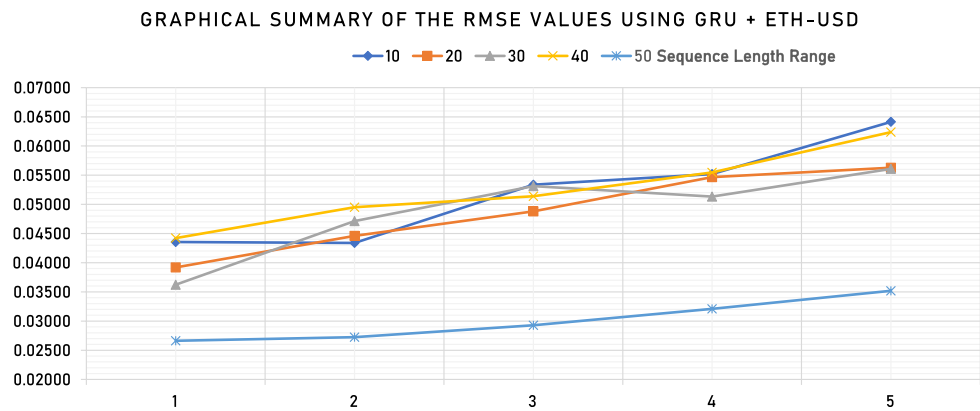
Table 6 The reported performance metrics for the "Ethereum USD (ETH-USD)" dataset using the two models (i.e., GRU and BiLSTM)

Sequence length	Days shift	GRU					BiLSTM				
		MSE	RMSE	MAE	MAPE	R2	MSE	RMSE	MAE	MAPE	R2
10	1	0.00190	0.04354	0.08720	0.31097	0.62423	0.00505	0.07106	0.10070	0.36715	− 0.00086
10	2	0.00188	0.04339	0.08710	0.30829	0.62688	0.00446	0.06681	0.09778	0.34999	0.11536
10	3	0.00285	0.05338	0.09064	0.32345	0.43517	0.00602	0.07760	0.10348	0.37361	− 0.19344
10	4	0.00305	0.05519	0.09090	0.32083	0.39635	0.00806	0.08976	0.10777	0.39804	− 0.59694
10	5	0.00411	0.06414	0.09458	0.33751	0.18467	0.00851	0.09224	0.10970	0.40123	− 0.68629
20	1	0.00154	0.03919	0.07524	0.27613	0.63226	0.00530	0.07277	0.08746	0.32926	− 0.26769
20	2	0.00199	0.04460	0.07622	0.28021	0.52387	0.00652	0.08077	0.09121	0.34340	− 0.56185
20	3	0.00238	0.04880	0.07645	0.28256	0.43000	0.00627	0.07918	0.08956	0.33641	− 0.50098
20	4	0.00299	0.05467	0.07807	0.28917	0.28461	0.00644	0.08026	0.08908	0.33190	− 0.54213
20	5	0.00317	0.05628	0.07881	0.29098	0.24176	0.01008	0.10042	0.10121	0.38293	− 1.41393
30	1	0.00131	0.03623	0.06574	0.24598	0.61101	0.00582	0.07631	0.07516	0.29116	− 0.72540
30	2	0.00222	0.04715	0.06757	0.25666	0.34134	0.00953	0.09764	0.08749	0.34874	− 1.82491
30	3	0.00282	0.05312	0.06876	0.26228	0.16395	0.00651	0.08068	0.07612	0.29516	− 0.92854
30	4	0.00264	0.05134	0.06759	0.25535	0.21913	0.01073	0.10357	0.08922	0.35264	− 2.17852
30	5	0.00314	0.05608	0.06821	0.25849	0.06828	0.00991	0.09955	0.08660	0.34064	− 1.93643
40	1	0.00196	0.04422	0.05783	0.22634	0.25545	0.01057	0.10283	0.07552	0.31423	− 3.02686
40	2	0.00245	0.04950	0.05846	0.22918	0.06697	0.00973	0.09863	0.07324	0.30286	− 2.70458
40	3	0.00264	0.05140	0.05773	0.22381	− 0.00610	0.01026	0.10128	0.07381	0.30502	− 2.90628
40	4	0.00308	0.05548	0.05801	0.22457	− 0.17217	0.01179	0.10857	0.07716	0.31863	− 3.48851
40	5	0.00389	0.06238	0.05964	0.23339	− 0.48171	0.01289	0.11355	0.08008	0.33086	− 3.91045
50	1	0.00071	0.02662	0.05520	0.20288	0.73784	0.00523	0.07232	0.05710	0.23379	− 0.93497
50	2	0.00074	0.02723	0.05452	0.20170	0.72561	0.00597	0.07726	0.06022	0.24912	− 1.20843
50	3	0.00086	0.02929	0.05488	0.20454	0.68261	0.00628	0.07926	0.06022	0.24741	− 1.32369
50	4	0.00103	0.03209	0.05441	0.20232	0.61896	0.00822	0.09066	0.06573	0.27388	− 2.04079
50	5	0.00124	0.03519	0.05350	0.19671	0.54193	0.00654	0.08085	0.06062	0.24625	− 1.41803
Min. values		0.00071	0.02662	0.05350	0.19671	− 0.48171	0.00446	0.06681	0.05710	0.23379	− 3.91045
Max. values		0.00411	0.06414	0.09458	0.33751	0.73784	0.01289	0.11355	0.10970	0.40123	0.11536

Fig. 10 A graphical summarization of the reported results using BiLSTM and "Ethereum USD (ETH-USD)" dataset

GRAPHICAL SUMMARY OF THE RMSE VALUES USING BILSTM + ETH –USD

sequence length and days shift concerning the MSE (and RMSE) are 50 and 1 respectively, concerning the MAE are 50 and 4 respectively, and concerning the MAPE are 50 and 4 respectively. For the BiLSTM model, the best sequence length and days shift concerning the MSE (and RMSE) are 30 and 1 respectively, concerning the MAE are 40 and 3 respectively, and concerning the MAPE are 40 and 3 respectively. From them, the majority voted sequence length

Fig. 11 A graphical summarization of the reported results using GRU and "Ethereum USD (ETH-USD)" dataset


GRAPHICAL SUMMARY OF THE RMSE VALUES USING GRU + ETH–USD

Table 7 The reported performance metrics for the "Cardano USD (ADA-USD)" dataset using the two models (i.e., GRU and BiLSTM)

| Sequence length	Days shift	GRU					BiLSTM				
		MSE	RMSE	MAE	MAPE	R2	MSE	RMSE	MAE	MAPE	R2
10	1	0.00036	0.01896	0.01907	0.11607	− 0.38361	0.00071	0.02667	0.02300	0.14320	− 1.73886
10	2	0.00038	0.01962	0.01882	0.11631	− 0.48161	0.00079	0.02816	0.02385	0.14859	− 2.05359
10	3	0.00053	0.02305	0.02116	0.12907	− 1.04552	0.00165	0.04058	0.03424	0.21699	− 5.34154
10	4	0.00052	0.02291	0.02041	0.12478	− 1.02035	0.00074	0.02716	0.02211	0.13697	− 1.83944
10	5	0.00068	0.02604	0.02152	0.13393	− 1.61146	0.00119	0.03444	0.02757	0.17284	− 3.56650
20	1	0.00026	0.01609	0.01756	0.10937	0.00983	0.00077	0.02777	0.02470	0.15746	− 1.95009
20	2	0.00026	0.01619	0.01691	0.10321	− 0.00310	0.00068	0.02604	0.02274	0.14452	− 1.59391
20	3	0.00027	0.01656	0.01650	0.10173	− 0.04979	0.00106	0.03257	0.02794	0.17830	− 3.05924
20	4	0.00031	0.01757	0.01704	0.10474	− 0.18095	0.00066	0.02563	0.02171	0.13721	− 1.51413
20	5	0.00030	0.01741	0.01633	0.10044	− 0.15923	0.00125	0.03541	0.03050	0.19488	− 3.79748
30	1	0.00019	0.01384	0.01691	0.10324	0.24750	0.00031	0.01752	0.01773	0.11142	− 0.20519
30	2	0.00024	0.01540	0.01674	0.10227	0.06843	0.00048	0.02198	0.02133	0.13633	− 0.89745
30	3	0.00025	0.01592	0.01644	0.10179	0.00420	0.00047	0.02156	0.02058	0.13110	− 0.82658
30	4	0.00026	0.01609	0.01572	0.09600	− 0.01735	0.00053	0.02309	0.02192	0.14002	− 1.09465
30	5	0.00035	0.01878	0.01717	0.10825	− 0.38537	0.00069	0.02633	0.02483	0.15934	− 1.72254
40	1	0.00017	0.01310	0.01265	0.08124	− 0.79895	0.00064	0.02532	0.02117	0.13941	− 5.72381
40	2	0.00026	0.01615	0.01413	0.09148	− 1.73523	0.00055	0.02341	0.01956	0.12889	− 4.74776
40	3	0.00026	0.01599	0.01386	0.08995	− 1.68271	0.00043	0.02072	0.01599	0.10515	− 3.50270
40	4	0.00025	0.01566	0.01356	0.08774	− 1.57180	0.00064	0.02524	0.02024	0.13318	− 5.68265
40	5	0.00034	0.01852	0.01536	0.10059	− 2.59725	0.00072	0.02679	0.02142	0.14072	− 6.52794
50	1	0.00007	0.00852	0.00985	0.06439	0.23136	0.00053	0.02298	0.01985	0.13158	− 4.58963
50	2	0.00011	0.01055	0.01071	0.07081	− 0.17865	0.00055	0.02342	0.01910	0.12668	− 4.80525
50	3	0.00010	0.00995	0.01022	0.06713	− 0.04824	0.00047	0.02164	0.01742	0.11567	− 3.95760
50	4	0.00009	0.00950	0.00971	0.06256	0.04382	0.00058	0.02414	0.01967	0.13029	− 5.16724
50	5	0.00010	0.01013	0.00998	0.06447	− 0.08678	0.00087	0.02952	0.02536	0.16708	− 8.22629
Min. values		0.00007	0.00852	0.00971	0.06256	− 2.59725	0.00031	0.01752	0.01599	0.10515	− 8.22629
Max. values		0.00068	0.02604	0.02152	0.13393	0.24750	0.00165	0.04058	0.03424	0.21699	− 0.20519

and days shift are are 50 and 1 respectively. Figures 12 and 13 summarize the reported RMSE results graphically using BiLSTM and GRU, where the x-axis represents the day shifts range (i.e., [1, 2, 3, 4, 5] and the y-axis represents RMSE values). From them, the GRU and 50-sequence-length report the best RMSE values while the BiLSTM and 30- and 40-sequence-length report the best RMSE values.

Fig. 12 A graphical summarization of the reported results using BiLSTM and "Cardano USD (ADA-USD)" dataset

GRAPHICAL SUMMARY OF THE RMSE VALUES USING BILSTM + ADA–USD
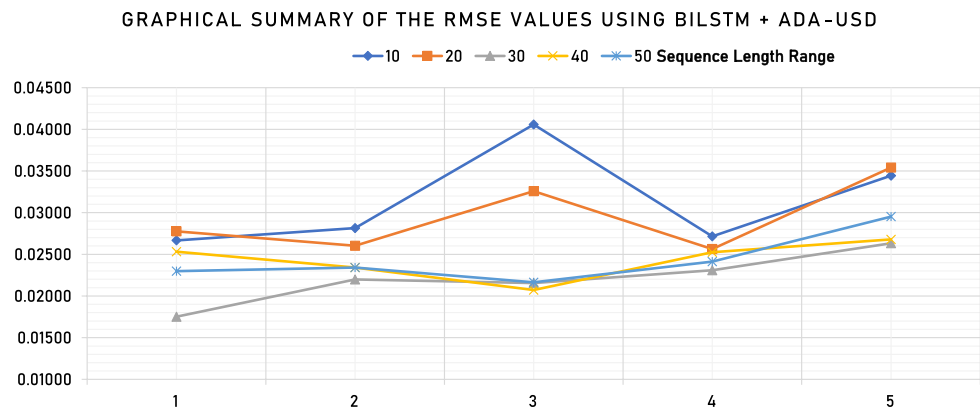
Fig. 13 A graphical summarization of the reported results using GRU and "Cardano USD (ADA-USD)" dataset

GRAPHICAL SUMMARY OF THE RMSE VALUES USING GRU + ADA–USD
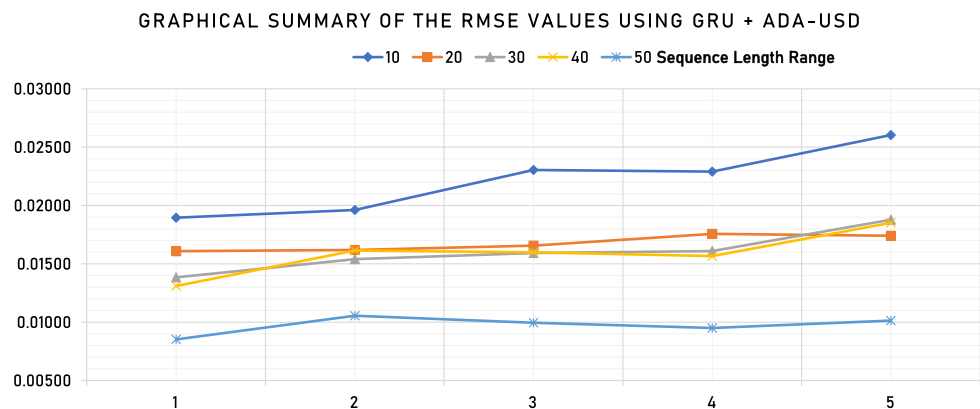
5 Limitations

One of the main limitations of this work is that each cryptocurrency was treated independently neglecting its potential relations with other cryptocurrencies. Additionally, the non-stationarity and the complexity of cryptocurrency time-series data are not considered. Additionally, all the coins considered in the study are high market capitalization, thus, their behavior is different from newly ICO coins that have low market capitalization. The proposed approach suffers from un-instantaneity. As the training of the models is the most time-consuming stage. Moreover, only 2 RNN architectures are used.

6 Conclusion and future work

A precise prediction of cryptocurrency prices methodology is critical in digital financial markets. Due to variances in forecasting capability per coin, artificial intelligence and machine learning approaches are appealing. This study introduced a framework based on two types of deep learning algorithms (i.e., BiLSTM and GRU). They are utilized to predict the prices of three of the most famous types of cryptocurrency (i.e., Bitcoin, Ethereum, and Cardano). The framework consists of five main phases. First, data is retrieved from a public real-time cryptocurrency source from "Yahoo Finance." Second, data is preprocessed to prepare the dataset for the next phase via filtering and squished features. Third, the classification via BiLSTM and GRU and model optimization. Fourth, performance evaluation for each epoch, and finally, the future prediction phase. The evaluation metrics such as MSE, RMSE, MAE, MAPE, and R2 are applied to test the accuracy of the used models. For the "Bitcoin USD (BTC-USD)" dataset, for the GRU model, the minimum MSE, RMSE, MAE, and MAPE are 0.00029, 0.01711, 0.02214, and 0.07036, respectively. Similarly, for the BiLSTM model, the minimum MSE, RMSE, MAE, and MAPE are 0.00210, 0.04582, 0.03358, and 0.10942, respectively. For

the "Ethereum USD (ETH-USD)" dataset, for the GRU model, the minimum MSE, RMSE, MAE, and MAPE are 0.00071, 0.02662, 0.05350, and 19.67071, respectively. Similarly, for the BiLSTM model, the minimum MSE, RMSE, MAE, and MAPE are 0.00446, 0.06681, 0.05710, and 0.23379, respectively. For the "Cardano USD (ADA-USD)" dataset, for the GRU model, the minimum MSE, RMSE, MAE, and MAPE are 0.00007, 0.00852, 0.00971, and 0.06256, respectively. Similarly, for the BiLSTM model, the minimum MSE, RMSE, MAE, and MAPE are 0.00031, 0.01752, 0.01599, and 0.10515, respectively. The results indicate that GRU outperformed the BiLSTM algorithm for Bitcoin, Ethereum, and Cardano, respectively. The R2 for the GRU model was found to be -2.89528, -0.48171, and -2.5972 for Bitcoin, Ethereum, and Cardano, respectively. Relying on these results, the GRU model is more efficient and reliable in predicting the prices of cryptocurrencies than BiLSTM but the two algorithms deliver excellent results.

In future work, other factors that affect the cryptocurrency market will be investigated. Autoencoder-based time series neural networks will be applied to perform predictions of the time-series data. Moreover, various hyperparameter tuning algorithms such as random or metaheuristic optimization algorithms (e.g., Genetic algorithm, Particle Swarm algorithm, and Bayesian optimization algorithm) can be applied instead of grid search. Additionally, the effect that social media can have on the price and trading volume of cryptocurrencies will be focused on. Hence, sentiment analysis and natural language processing techniques will be used to analyze posts and tweets to extract insights.

Author contributions The authors certify that the work they have submitted for publication is entirely new, has never been published before, and is not presently being considered for publication elsewhere. We certify that all mentioned authors have reviewed the paper and that no other individuals who meet the requirements for authorship but are not listed have contributed to the work. We further reaffirm that we all approved of the order in which the authors are listed in the text. We recognize that the corresponding author serves as the exclusive point of contact for the editorial procedure. He is in charge of informing the other writers of his progress, their submission of corrections, and his final acceptance of the proofs. The author read and approved the final manuscript.

Data availibility The datasets, if existing, that are used, generated, or analyzed during the current study (A) if the datasets are owned by the authors, they are available from the corresponding author on reasonable request, (B) if the datasets are not owned by the authors, the supplementary information including the links are included in this published article.

Declarations

Ethics approval and consent to participate The current study does not contain any studies with human participants and/or animals performed by any of the authors. There is no informed consent for the current study.

Consent for publication Not applicable.

Competing interests The authors declare that they have no competing interests.

References

1. Patel MM, Tanwar S, Gupta R, Kumar N. A deep learning-based cryptocurrency price prediction scheme for financial institutions. J Inf Security Appl. 2020;55:102583.
2. Nakamoto S. Bitcoin: a peer-to-peer electronic cash system. Decentralized Business Review. 2008:21260.
3. Mukhopadhyay U, Skjellum A, Hambolu O, Oakley J, Yu L, Brooks R. A brief survey of cryptocurrency systems. In: 2016 14th annual conference on privacy, security and trust (PST), IEEE. 2016:745–52.
4. Rose C, et al. The evolution of digital currencies: Bitcoin, a cryptocurrency causing a monetary revolution. Int Bus Econ Res J. 2015;14(4):617–22.
5. Eyal I. Blockchain technology: transforming libertarian cryptocurrency dreams to finance and banking realities. Computer. 2017;50(9):38–49.
6. Adams R, Kewell B, Parry G. Blockchain for good? digital ledger technology and sustainable development goals. In: Leal Filho W, Marans R, Callewaert J, editors. Handbook of sustainability and social science research. Berlin: Springer; 2018. p. 127–40.

7.   Saad M, Choi J, Nyang D, Kim J, Mohaisen A. Toward characterizing blockchain-based cryptocurrencies for highly accurate predictions. IEEE Syst J. 2019;14(1):321–32.
8.   Jang H, Lee J. An empirical study on modeling and prediction of bitcoin prices with Bayesian neural networks based on blockchain information. IEEE Access. 2017;6:5427–37.
9.   Number of blockchain wallet users 2022/2023: breakdowns, timelines, and predictions. https://financesonline.com/number-of-block chain-wallet-users/. Accessed 24 June 2022.
10. Blockchain.com wallets. https://www.blockchain.com/bg/charts/my-wallet-n-users. Accessed 16 Sept 2022.
11. Coinbase announces effectiveness of registration statement and anticipated listing date of its class a common stock on the nasdaq global select market. https://blog.coinbase.com/coinbase-announces-effectiveness-of-registration-statement-and-anticipated-listing-date-of-its-1509b281f760. Accessed 16 Sept 2022.
12. Coinmarketcap. "bitcoin: Historical data". https://coinmarketcap.com/currencies/bitcoin/historical-data/. Accessed 16 Sept 2022.
13. Blockchain.com. "bitcoin btc: Jan. 1, 2022 to June 13, 2022". https://www.blockchain.com/explorer/assets/BTC?from=1641056400 &to=1655136000 &timeSpan=custom &scale=0 &style=line. Accessed 16 Sept 2022.
14. Chan S, Chu J, Nadarajah S, Osterrieder J. A statistical analysis of cryptocurrencies. J Risk Financ Manag. 2017;10(2):12.
15. Kyriazis NA. A survey on efficiency and profitable trading opportunities in cryptocurrency markets. J Risk Financ Manag. 2019;12(2):67.
16. Fang F, Ventre C, Basios M, Kanthan L, Martinez-Rego D, Fan W, Li L. Cryptocurrency trading: a comprehensive survey. Financ Innov. 2022;8(1):1–59.
17. Makarov I, Schoar A. Trading and arbitrage in cryptocurrency markets. J Financ Econ. 2020;135(2):293–319.
18. Chava S, Hu F, Paradkar N. Gambling on crypto tokens?. SSRN 4149937. 2022.
19. Sovbetov Y. Factors influencing cryptocurrency prices: Evidence from bitcoin, ethereum, dash, litcoin, and monero. J Econ Financ Anal. 2018;2(2):1–27.
20. Fantazzini D, Zimin S. A multivariate approach for the simultaneous modelling of market risk and credit risk for cryptocurrencies. J Ind Bus Econ. 2020;47(1):19–69.
21. Fosso Wamba S, Kala Kamdjoug JR, Epie Bawack R, Keogh JG. Bitcoin, blockchain and fintech: a systematic review and case studies in the supply chain. Prod Plan Control. 2020;31(2–3):115–42.
22. Coinmarketcap. "cryptocurrency prices, charts, and market capitalizations". https://coinmarketcap.com/. Accessed 17 Sept 2022.
23. Binaghi E, Omodei M, Pedoia V, Balbi S, Lattanzi D, Monti E. Automatic segmentation of MR brain tumor images using support vector machine in combination with graph cut. In: IJCCI (NCTA). 2014:152–7.
24. Schuster M, Paliwal KK. Bidirectional recurrent neural networks. IEEE Trans Signal Process. 1997;45(11):2673–81.
25. Ciregan D, Meier U, Schmidhuber J. Multi-column deep neural networks for image classification. In: 2012 IEEE conference on computer vision and pattern recognition, IEEE. 2012:3642–9.
26. Bitcoin usd (btc-usd). https://finance.yahoo.com/quote/BTC-USD/history?p=BTC-USD. Accessed 09 Aug 2022.
27. Ethereum usd (eth-usd). https://finance.yahoo.com/quote/ETH-USD/history?p=ETH-USD. Accessed 09 Aug 2022.
28. Cardano usd (ada-usd). https://finance.yahoo.com/quote/ADA-USD/history?p=ADA-USD. Accessed 09 Aug 2022.
29. Mudassir M, Bennbaia S, Unal D, Hammoudeh M. Time-series forecasting of bitcoin prices using high-dimensional features: a machine learning approach. Neural Comput Appl. 2020:1–15.
30. Gupta M, Asthana A, Joshi N, Mehndiratta P. Improving time series forecasting using mathematical and deep learning models. In: International conference on big data analytics, Springer. 2018:115–25.
31. Lustick LS, Saltzberg B, Buckley JK, Heath RG. Autoregressive model for simplified computer generation of EEG correlation functions. In: Proceedings of the IEEE annual conference on engineering in medicine and biology. 1968;10:78–94.
32. Nelson K. Time series analysis using autoregressive integrated moving average (arima) models. Acad Emerg Med. 1998;5(7):739–44.
33. Chang X, Gao M, Wang Y, Hou X. Seasonal autoregressive integrated moving average model for precipitation time series. J Math Stat. 2012;8(4):500–5.
34. Hyndman RJ, Athanasopoulos G. Forecasting: principles and practice. Heathmont: OTexts; 2018.
35. Salinas D, Flunkert V, Gasthaus J, Januschowski T. Deepar: Probabilistic forecasting with autoregressive recurrent networks. Int J Forecast. 2020;36(3):1181–91.
36. Oreshkin BN, Carpov D, Chapados N, Bengio Y. N-beats: neural basis expansion analysis for interpretable time series forecasting. 2019;arXiv preprint arXiv:1905.10437.
37. Hitam NA, Ismail AR. Comparative performance of machine learning algorithms for cryptocurrency forecasting. Ind J Electr Eng Comput Sci. 2018;11(3):1121–8.
38. Derbentsev V, Babenko V, Khrustalev K, Obruch H, Khrustalova S. Comparative performance of machine learning ensemble algorithms for forecasting cryptocurrency prices. Int J Eng. 2021;34(1):140–8.
39. Hamayel MJ, Owda AY. A novel cryptocurrency price prediction model using GRU, LSTM and bi-LSTM machine learning algorithms. AI. 2021;2(4):477–96.
40. Pour ES, Jafari H, Lashgari A, Rabiee E, Ahmadisharaf A. Cryptocurrency price prediction with neural networks of LSTM and Bayesian optimization. Eur J Bus Manage Res. 2022;7(2):20–7.
41. Kim G, Shin D-H. Jae GC, Sunghoon L. IEEE access: a deep learning-based cryptocurrency price prediction model that uses on-chain data; 2022.
42. Miura R, Pichl L, Kaizoji T. Artificial neural networks for realized volatility prediction in cryptocurrency time series. In: International symposium on neural networks. Springer; 2019:165–72.
43. Yiying W, Yeze Z. Cryptocurrency price analysis with artificial intelligence. In: 2019 5th international conference on Information Management (ICIM), IEEE. 2019:97–101.
44. Shahbazi Z, Byun Y-C. Improving the cryptocurrency price prediction performance based on reinforcement learning. IEEE Access. 2021;9:162651–9.
45. Jay P, Kalariya V, Parmar P, Tanwar S, Kumar N, Alazab M. Stochastic neural networks for cryptocurrency price prediction. IEEE Access. 2020;8:82804–18.

46. Uras N, Marchesi L, Marchesi M, Tonelli R. Forecasting bitcoin closing price series using linear regression and neural networks models. PeerJ Comput Sci. 2020;6:e279.
47. Tanwar S, Patel NP, Patel SN, Patel JR, Sharma G, Davidson IE. Deep learning-based cryptocurrency price prediction scheme with inter-dependent relations. IEEE Access. 2021;9:138633–46.
48. Alessandretti L, ElBahrawy A, Aiello LM, Baronchelli A. Machine learning the cryptocurrency market. SSRN 3183792. 2018.
49. Fang F, Chung W, Ventre C, Basios M, Kanthan L, Li L, Wu F. Ascertaining price formation in cryptocurrency markets with machine learning. Eur J Finance. 2021. https://doi.org/10.1080/1351847X.2021.1908390.
50. Alanyali M, Moat HS, Preis T. Quantifying the relationship between financial news and the stock market. Sci Rep. 2013;3(1):1–6.
51. Greff K, Srivastava RK, Koutník J, Steunebrink BR, Schmidhuber J. Lstm: a search space odyssey. IEEE Trans Neural Netw Learn Syst. 2016;28(10):2222–32.
52. Siami-Namini S, Tavakoli N, Namin AS. The performance of LSTM and bilstm in forecasting time series. In: 2019 IEEE international conference on Big Data (Big Data), IEEE. 2019:3285–92.
53. Dey R, Salem FM. Gate-variants of gated recurrent unit (GRU) neural networks. In: 2017 IEEE 60th international midwest symposium on circuits and systems (MWSCAS), IEEE. 2017:1597–600.
54. Lydia A, Francis S. Adagrad—an optimizer for stochastic gradient descent. Int J Inf Comput Sci. 2019;6(5):566–8.
55. Liashchynskyi P, Liashchynskyi P. Grid search, random search, genetic algorithm: a big comparison for nas. 2019. arXiv preprint arXiv:1912.06059.

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