AN ANALYSIS FOR CRYPTOCURRENCY PRICE PREDICTION USING LSTM, GRU, AND THE BI- DIRECTIONAL IMPLICATIONS

Chapter · March 2022

Sule Ozturk Birim
Manisa Celal Bayar University

AN ANALYSIS FOR CRYPTOCURRENCY PRICE PREDICTION USING LSTM, GRU, AND THE BI-DIRECTIONAL IMPLICATIONS

Şule ÖZTÜRK BİRİM[1]

1. INTRODUCTION

Digital currencies experienced rapid development in the last decade. This is one of the most prominent novelties that occurred in the modern global economy. Bitcoin was the first cryptocurrency ever transacted on the internet. Recently, around 2000 kinds of currencies with huge amounts of transactions are observed. Cryptocurrencies have a working principle of encryption solving to create unique hashes which the users exchange during transactions as if they exchange fiat currencies (DeVries, 2016). Cryptocurrencies are different from fiat currencies since they are not generated by countries, do not follow any country rules, and do not require a third party for the interaction (Ergun & Karabiyik, 2021).

The demand for the cryptocurrency market increased rapidly over time. Now for only bitcoin, more than 10 000 transactions occur per hour (BitInfoCharts, 2022). Bitcoin is the market leader with a value of 793 billion dollars (Coinmarketcap, 2022). Bitcoin price has shown a rapid movement during the last decade. In January 2017, one Bitcoin was worth less than 1000 dollars while in January 2020 the value of one Bitcoin was more than 7200 dollars. In January 2022, it was over 43 000 dollars compared to 66 000 dollars in September 2021 (Coinmarketcap, 2022). As the numbers reveal, the price of a cryptocurrency shows huge amounts of variations and these fluctuations take the attention of the investors (Saad et al., 2020). Fluctuations can give a vast amount of profit to the buyer as shown in the Bitcoin case. Additional to the potential gain, bitcoin, and other digital currencies are independent of government rules or restrictions. Due to this independence, digital currencies are regarded as being protected from

[1] Asst. Prof. Dr., Manisa Celal Bayar University, Faculty of Salihli Economics and Administrative Sciences, Department of Business, sule.ozturk@cbu.edu.tr, ORCID: 0000-0001-7544-8588.

inflation or other economical situations (Magro, 2017). Therefore, digital currencies are seen as safe to invest in, as they are not directly connected to inflation or the economic conditions of a country (DeVries, 2016). This pattern can show itself in the price increase of Bitcoin overtime where inflation can negatively affect the value of a national currency.

Although the digital currency market is attractive due to its unique characteristics such as potential high return, no association with classical financial assets and governmental restrictions, it should also be regarded as volatile and risky. In the literature, previous studies aimed to predict stock market prices and propose profitable transaction approaches based on those predictions (Sun et al., 2020). Since digital currencies are significant investment tools for buyers, it is necessary to apply forecasting methods in this new evolving market (Valencia et al., 2019). Considering the high popularity of cryptocurrency investment and the high risk associated with them, proposing well-working prediction methods for digital currencies become an important issue. Time series data about the prices of cryptocurrencies are characterized by high volatile, non-linear, and non-normal distributions including abnormal and extreme movements (Catania & Grassi, 2017). Additionally features that are related to the abnormal and extreme movements cannot be understood and identified (Selmi et al., 2018). Due to the highly volatile nature of digital currencies, traditional statistical approaches and econometric models used to predict traditional financial assets are proven to be ineffective when used for cryptocurrency prediction (Derbentsev et al., 2021). It is necessary to develop smart systems that do not require pre-determined assumptions about the data. Deep learning algorithms are based on neural networks which aim to extract underlying hidden patterns, integrate them and elaborate the valuable featural information by eliminating noise in the input data (Livieris et al., 2020). Several studies showed that deep learning algorithms provide highly effective results in forecasting financial assets data (Dautel et al., 2020; Derbentsev et al., 2021; Hamayel & Owda, 2021; Park & Ryu, 2021). This paper aims to predict the values of four chosen cryptocurrencies using various deep learning algorithms. The aims of this paper are presented as follows:

- Presenting a framework to model and predict prices of Bitcoin (BTC), Ethereum (ETH), Ripple (XRP) and Monero (XMR)

- Applying deep learning algorithms including LSTM (Long-Short Term Memory), GRU (Gated Recurrent Unit), Bidirectional LSTM and Bidirectional GRU techniques.
- Evaluating the prediction performance of the proposed deep learning algorithms using metrics of RMSE (Root Mean Squared Error) and MAPE (Mean Absolute Percentage Error).
- Identifying the best model to forecast prices of the chosen cryptocurrencies

2. LITERATURE REVIEW

Machine learning (ML) is an artificial intelligence tool that uses past data to predict the future. From this aspect ML techniques are suitable to be used in time series prediction. As the previous research showed, ML techniques have advantages over the traditional forecasting methods with providing higher prediction accuracy and showing lower error rates (Hamayel & Owda, 2021; Hitam & Ismail, 2018; Zhang et al., 2021). Additionally, as opposed to traditional time series forecasting models, ML methods are more powerful in analyzing non-linear multivariate data and have more robustness to noise in the data (Zhang et al., 2021). There are several machine learning techniques including, linear regression (LR), support vector machines (SVM), random forest (RF), decision trees (DT), k-nearest neighbor (KNN), naïve bayes (NB), artificial neural networks (ANN). There is also a specific type of ML which is called deep learning (DL). Deep learning utilizes several embedded layers of neural networks while using past data to predict the future. In this paper literature review about the studies exploiting traditional machine learning algorithms and studies related to deep learning algorithms in predicting cryptocurrency prices are given separately in the following sections.

2.1. Traditional Machine Learning in Cryptocurrency Prediction

In the literature, several studies applied and showed the encouraging performance of ML algorithms in cryptocurrency price forecasting. Hitam and Ismail (2018) predicted the prices of six digital currencies which are Bitcoin, Ethereum, Litecoin, Nem, Ripple and Stellar. The authors compared the performance of ML techniques which are SVM, ANN and DL and found that SVM outperformed the other ML techniques with higher accuracy. In another study, the authors used several features as predictors of the Bitcoin and Ethereum price and determined the best predictors with correlation analysis (Saad et al., 2020). With the selected features Saad et al. (2020) predicted digital currency prices with linear regression (LR), SVM and RF. The results showed that LR outperformed the other methods. In the same study, the authors used a specific type of DL which is Long-short term memory in price prediction of Bitcoin and Ethereum. The results showed that LSTM showed minimum prediction error for Bitcoin.

Derbentsev et al. (2019) constructed a ML model for short-term forecasting of the prices for Bitcoin, Ethereum and Ripple. They applied BART (Binary Auto-Regressive Tree), a specific type of DT, which combines the features of DT and ARIMA, a popular autoregressive statistical technique. The authors conducted short-term forecasts with the proposed BART model and the ARIMA model alone. They found that the proposed BART model revealed more accurate results than the ARIMA-only model (Derbentsev et al., 2019). Ergun & Karabiyik (2021) applied PATSOS to predict Monero prices. PATSOS is a hybrid system that combines two Adaptive Neuro-Fuzzy Inference System (ANFIS) sub-models (Atsalakis et al., 2019). Ergun and Karabiyik (2021) compared PATSOS performance with a single ANFIS technique and found that PATSOS revealed more accurate predictions for Monero prices.

Another study compared the prediction performance of LR and SVM for Bitcoin daily closing prices (Karasu et al., 2018). The prediction performance of SVM is found to be better than LR. On the other hand, Rathan et al. (2019) found that LR shows higher accuracy in price prediction of Bitcoin than the DT technique.

To predict nine different cryptocurrencies, Chowdhury et al. (2020) used several ML algorithms including ANN, KNN, gradient boosted trees, and an ensemble model which combines various ML techniques. Among the proposed models, the ensemble learning model showed the least error in the predictions. Derbentsev et al. (2021) also used an ensemble model consisting of Gradient Boosting Machine (GBM) and RF. They calculated the Mean Absolute Percentage Error (MAPE) for three digital currencies which are Bitcoin, Ethereum, and Ripple. The results showed that MAPE values for the ensemble model change between 0.92-2.61 %.

2.2. Deep Learning in Cryptocurrency Prediction

Deep learning (DL) is a specific type of machine learning consisting of several processing layers and using these layers to learn from the sample data. DL discovers hidden internal patterns in large data by utilizing a backpropagation algorithm and forms a representation based on the structure of the previous layer (Lecun et al., 2015). Deep learning achieved revolutionary successes in image processing, speech recognition, computer vision, and natural language processing (Hamayel & Owda, 2021; Lecun et al., 2015). Deep learning, especially the types of Recurrent Neural Networks (RNNs) are regarded as effective tools to conduct time series prediction (Ayoobi et al., 2021; Dautel et al., 2020; Shahid et al., 2020), while Convolutional Neural Networks are mostly used for computer vision missions (Dautel et al., 2020; Lecun et al., 2015). Although deep learning is used in time series prediction, studies which apply DL methods in cryptocurrency prediction are still infrequent in the literature.

Among the type of Recurrent Neural Networks (RNNs), LSTM and GRU are commonly used in time series prediction. Chen et al. (2021) developed a two-stage prediction methodology for Bitcoin prices. First, the authors used ANN and RF to determine the features used in price prediction. In the second stage, they applied LSTM with the selected features to predict Bitcoin prices. Results showed that the LSTM model showed superior performance when compared to ARIMA and SVM (Chen et al., 2021). In another paper, a hybrid methodology based on LSTM and GRU networks is used in the price prediction of Litecoin and Monero (Patel et al., 2020). The proposed method was compared with the LSTM only method. The findings showed that the hybrid model showed higher accuracy than the single LSTM model in price prediction of the selected currencies.

Wu et al. (2019) integrated the autoregressive (AR) feature to the LSTM network and used this methodology in predicting daily Bitcoin prices. Authors compared their proposed model with the traditional LSTM method and found that LSTM-AR showed lower error rates computed with mean squared error (MSE), root mean square error (RMSE), mean absolute percentage error (MAPE) and mean absolute error (MAE).

A specific type of deep learning called stacked denoising autoencoder (SDAE) is used in a cryptocurrency price prediction study (Liu et al., 2020). In that study feature selection for the predictors of Bitcoin is applied. Then, SDAE is exploited to predict Bitcoin prices. The performance of the proposed method was compared with a backpropagation neural network and an SVM model. Findings show that SDAE outperformed the other models with the best performance indicator scores measured with MAPE and RMSE (Liu et al., 2020).

A recent study applied a methodology based on one-dimensional CNN to predict Bitcoin trend (Cavalli & Amoretti, 2021). Bitcoin trend is resembled as whether the price has increased or declined as compared to the previous period, making trend prediction a classification problem. One-dimensional CNN was firstly applied to select predictors of Bitcoin prices among various datasets including social media, financial indicators, and transaction history. The proposed 1D CNN model is also applied to predict Bitcoin trend and compared with the performance of the LSTM model. The results showed that 1D CNN showed higher classification accuracy than the LSTM model (Cavalli & Amoretti, 2021).

Livieris et al. (2020) applied deep learning methods for both trend and price prediction of selected currencies. Bitcoin, Ethereum, and Ripple hourly prices are used as data. The authors proposed an ensemble learning method that consists of LSTM, Bi-directional LSTM and CNN. Based on their analysis, authors concluded that ensemble learning and deep learning methodologies can be effectively used together to bring accurate, reliable and powerful predictions (Livieris et al., 2020). Sun et al. (2020) proposed a Light Gradient Boosting Machine (LightGBM) to predict trend in 42 different cryptocurrencies. Authors used economic indicators as features for prediction. The results are compared with the performance of Gradient Boosting Decision Tree (GBDT) indicating that Light GBM was more robust results than the GBDT method. Lahmiri and Bekiros (2019) applied LSTM in price prediction of Bitcoin, Ripple and Digital Cash and compared with generalized regression neural networks (GRNN). The authors not only found LSTM produced superior prediction results than GRNN, but also learn chaotic patterns better in the utilized data.

Zhang et al. (2021) proposed a hybrid deep learning methodology that combines various features of GRU and CNN in the different layers of the network. They compared the proposed model with several conventional methods such as ARIMA, SVR, RF and other deep learning methods of LSTM, GRU and CNN. The results showed that the proposed hybrid model provided more accurate results than the compared techniques by showing the least prediction error. (Livieris et al., 2021) also proposed a new methodology in which different cryptocurrency data are blended in a way to be used in prediction of Bitcoin, Ethereum, and Ripple prices independently.

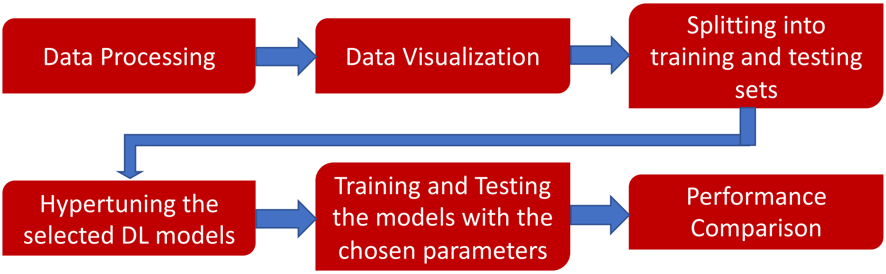
The predictions are conducted with a CNN-LSTM combined network. The results showed that the combined network effectively utilized mixed data with reduced overfitting and computational cost. Hamayel and Owda (2021) used three different types of RNNs to predict the prices of Bitcoin, Litecoin and Etherium. The proposed deep learning methods were GRU, LSTM and Bi-LSTM. The results showed that GRU model was superior to the other two DL models in all of the currency prices based on prediction accuracy.

Investors aim to earn revenue from their investments. Prediction of the future movement of an investment tool is a significant topic for an investor, since the investor may base her decision on the result of the prediction. Cryptocurrency has increasing attention as an investment. This makes forecasting Cryptocurrency prices an important issue. The literature shows that deep learning methodologies provide superior performance over traditional forecasting techniques. Therefore, in this study various deep learning techniques explained in the next section are applied in digital currency price prediction.

3. METHODOLOGY

To reach the goals of this study four different deep learning models including LSTM (Long-Short Term Memory), GRU (Gated Recurrent Unit), Bidirectional LSTM and Bidirectional GRU, CNN (Convolutional Neural Network) and a combination of CNN-LSTM are used to predict prices for four different cryptocurrencies that are Bitcoin (BTC), Ethereum (ETC), Ripple (XRP) and Monero (XMR). To evaluate the performance of the proposed DL techniques the methodology was applied in several stages including: (1) Data gathering for the historical prices of the selected digital currencies, (2) Visualization of the historical patterns for each currency, (3) Splitting each dataset into training and testing datasets (4) Hypertuning the DL models to find the best parameters (5) Training and testing the four DL models (6) Calculating and comparing the performance of each DL method based on the chosen metrics. Methodological framework of this study is shown in Figure 1.

Figure 1. **Methodological Framework**



3.1. Data Description and Preprocessing

Data for the selected digital currencies are taken from Yahoo Finance. Daily data for BTC, ETC, XRP and XMR are obtained from November 9th 2017 to January 7th 2022. The dataset for each digital currency contains 1521 observations. Dataset includes date, opening price, the highest price, the lowest price, the closing price and the volume of transactions conducted at that day. In this study, closing price is used as the time series data for all the currencies.

Data is preprocessed first to make it proper to be analyzed with deep learning tools. Missing data check is conducted, and it is seen that there was no missing value in the dataset. After that, normalization of the data is realized because it helps DL algorithm to converge faster (Singh et al., 2021). Minimum-maximum scaler is used for normalization which converts the values into range 0 to 1. After the scaling, the dataset was shaped in a way ready for supervised learning. Data is converted in a form to predict the cryptocurrency price at the current day based on the previous two days. In other words, time steps used in this study is two to predict the following value. With this conversion, the data was set to be used in supervised time series forecasting.

3.2. Data Visualization

Before analyzing the data, it is beneficial to understand the behavior of the data by observing its pattern. Figure 2a, 2b, 2c and 2d indicate the time-dependent values of BTC, ETC, XRP and XMR respectively and illustrate the distribution of the digital currency values between November 9th 2017 to January 7th 2022.

Figure 2a. BTC/USD closing prices for the selected days

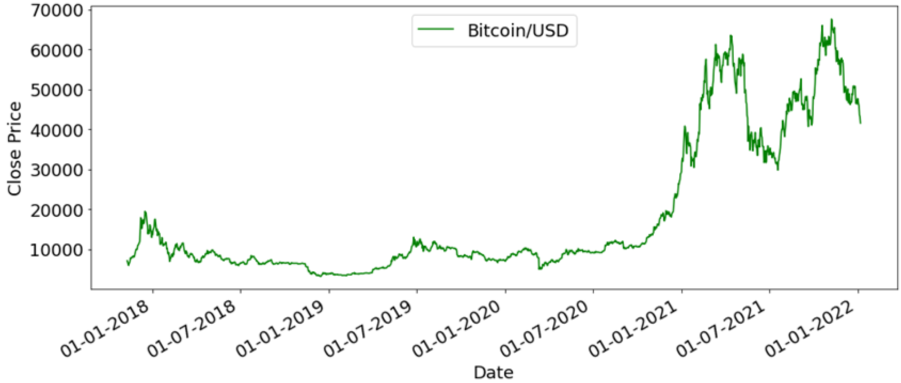


Figure 2b. ETH/USD closing prices for the selected days

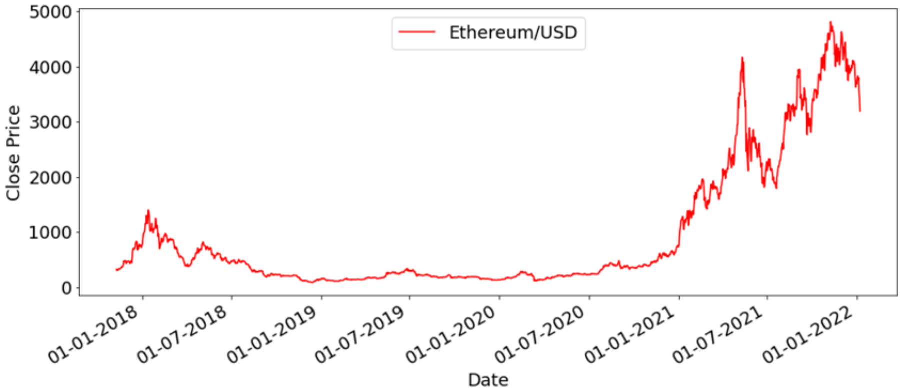


Figure 2c. XRP/USD closing prices for the selected days

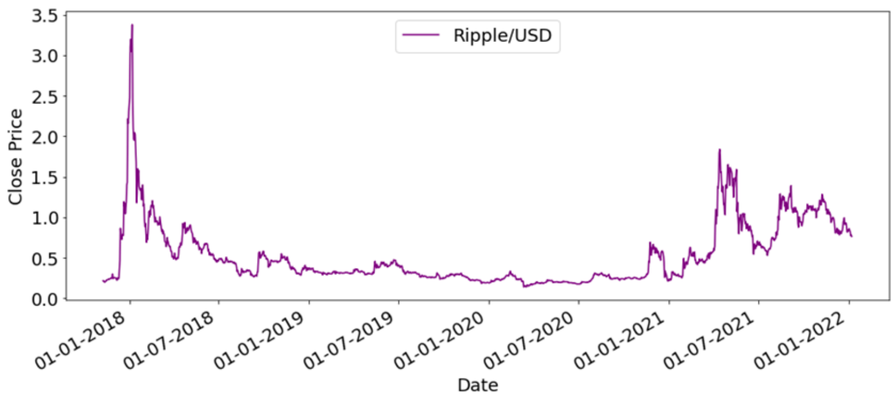


Figure 2d. XMR/USD closing prices for the selected days

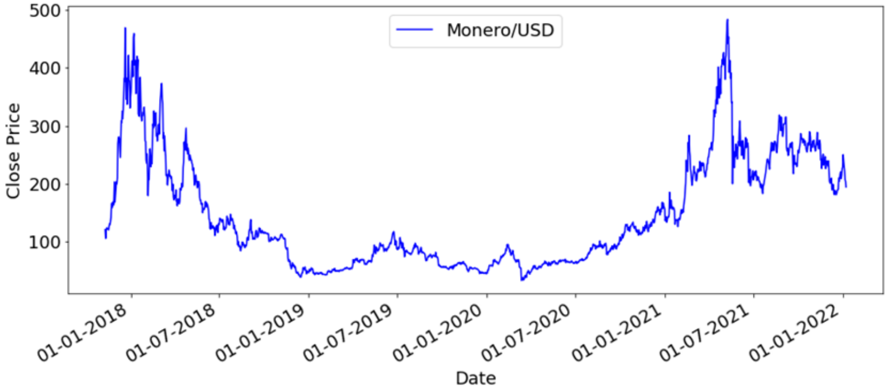


Figure 2a and 2b show that price of BTC and ETH has an increasing trend in the selected interval despite they experienced drops and rises. Figure 2c and 2d show that XRP and XMR do not show an increasing trend, but they had two important peaks on the beginning of 2018 and in mid-2021. Additionally, Figures 2a, 2b, 2c and 2c show that between the end of 2018 and the end of mid-2020 all the currencies illustrated a steady period in which the currencies did not show any remarkable trend or peaks.

3.3. Training and Testing Sets

To split the datasets into training and testing ratio of 80:20 is used. For each currency dataset, the first 80 percent of the data is selected as the training set while the remaining 20 percent is selected as the testing set. For the four cryptocurrencies, the training data is composed of the first 1216 days while the test data consists of the last 305 days of the datasets.

3.4. Deep Learning Algorithms

This section gives information about the utilized deep learning algorithms in this study that are LSTM, Bi-LSTM, GRU and Bi-GRU.

3.4.1. Long-Short Term Memory (LSTM)

Earlier forms of RNNs have weakness of not capturing long-term temporal dependencies of sequential data. LSTM was emerged as a type of RNN to overcome this weakness. LSTM networks are effective in handling long-term temporal dependencies in sequential data while do not experience the optimization difficulties which the traditional RNNs face (Greff et al., 2017).

An LSTM cell has three gates which are input, output and forget gates. LSTM also has a cell state which stores the temporal state and control the gates (Shahid et al., 2020). Forget gate determines the amount of the information which should be thrown away from the cell state. Input gate regulates the amount of new information stored in the cell state while the output gate determines the information that should be sent as the output from the cell state (Ayoobi et al., 2021). Equations for LSTM are as follows (Ayoobi et al., 2021; Hamayel & Owda, 2021; Shahid et al., 2020):

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_t) \quad (1)$$

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (2)$$

$$c_t = f_t * c_{t-1} + i_t + \tanh(W_c[h_{t-1}, x_t] + b_c) \quad (3)$$

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (4)$$

$$h_t = o_t * \tanh(c_t) \quad (5)$$

Where, x_t is the input, b is the bias, w is the weights, c_t is the cell state, f_t is the forget gate, i_t is the input gate, o_t is the output gate and h_t is the hidden layer. While σ is the sigmoid function, tanh is the hyperbolic tangent function.

3.4.2. Gated Recurrent Unit (GRU)

A GRU cell includes two gates as opposed to an LSTM cell. These two gates are an update gate and a reset gate. The weights of these gates are adaptively updated in the learning phase of the algorithm (Dey & Salem, 2017). GRU has achieved remarkable success on various tasks including computer vision and sequential data. The success of GRU lies in the simpler structure than LSTM with two gates instead of three (Yang et al., 2020).

GRU can be regarded as a variant of LSTM. The aim of the GRU is the same as LSTM that is to provide improvement of a traditional RNN in predicting sequential data. GRU can even be more successful and faster than LSTM in low dimensional sequential data due to its smaller number of gates and parameters. Additionally, training in LSTM can be longer since the network structure is more complex while LSTM can get better results in high dimensional data (Lai et al., 2021). Equations handled for a GRU unit are as follows (Dey & Salem, 2017; Lai et al., 2021):

$$u_t = \sigma(W_u[h_{t-1}, x_t]) \quad (6)$$

$$r_t = \sigma(W_r[h_{t-1}, x_t]) \quad (7)$$

$$h_t = (1 - u_t) * h_{t-1} + u_t * tanh(W[r_t * h_{t-1}, u_t) \quad (8)$$

In the above equations, u_t is the update gate and r_t is the reset gate.

3.4.3. Bi-directional Applications of Recurrent Neural Networks

Bi-directional application is a different way of stacking RNNs. Bi-directional approach to RNN was first introduced by Schuster and Paliwal (1997). Bi-RNN not only uses past information while training but also utilizes future information of a specific time frame. This is handled by splitting neurons of a RNN into two parts; one is responsible for future states, the other part is responsible for past states (Schuster & Paliwal, 1997). Each group of neurons constructs the forward and backward layers, and these two layers are linked to each other. Various merging methods are used to combine the outputs of the forward and backward layers to produce a final result (Althelaya et al., 2018).

When the two layers of LSTM one of which flows in forward direction and the other one flows in backward direction are stacked together, a bidirectional LSTM (Bi-LSTM) structure is formed. Similarly, if two opposite direction layers of GRU network are stacked then a Bidirectional GRU (Bi-GRU) network is formed. Bidirectional application can be beneficial since the network does not only use the past information. The future information of the current time is valuable to extract relations inside the data and solve the problem of unawareness of the network (Yang et al., 2020).

Bi-LSTM and Bi-GRU have been applied in time series predictions and illustrated encouraging results. Lai et al.(2021), Singh et al.(2021) and Yang et al.(2020) applied Bi-GRU to a time series data and achieved successful results. Althelaya et al. (2018), Ayoobi et al. (2021), Hamayel and Owda (2021) and Shahid et al. (2020) applied Bi-LSTM to the time series data and indicated the powerful functionality of LSTM.

3.4.4. Performance Metrics

To evaluate the performance of the proposed DL methods RMSE and MAPE values are used in this study. RMSE and MAPE are calculated based on the following equations:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(A_i - F_i)^2}{n}} \tag{9}$$

$$MAPE = \frac{\sum_{i=1}^{n}(\frac{|A_i - F_i|}{A_i})}{n} \tag{10}$$

In above equations n is the number of observations, A_i is the actual observations while and F_i is the forecasted values based on the model execution.

3.4.5. Hypertuning of the Parameters

Hyperparameter tuning is an important application that affects the machine learning algorithm's performance significantly. With the best hyperparameters, algorithm's performance may enhance remarkably and the model can conduct more accurate predictions. If the model run with the best-known parameters, the algorithm can learn tasks faster than the non-tuned model (Amirabadi et al., 2020). Therefore, before the final run of the DL algorithm, it is necessary to identify the best values for the parameters. In this study, number of neurons in each layer, epoch size and batch size are used as the parameters to be tuned. One complete forward and backward run of the entire dataset during the execution of the model is called an epoch. The whole dataset must be divided into groups during the execution and the number of these divided groups are called batch size. Table 1 shows the alternative values of the parameters used for every DL method, while Table 2 shows the chosen values of the parameters which demonstrated the best performance in the test set based on the RMSE.

Table 1. Alternative Values for the Chosen Parameters

Parameter	Values
Number of neurons	[16, 32, 64, 128]
Epoch size	[100, 250, 500, 1000]
Batch Size	[100, 250, 500]

Table 2. Best Parameters for the Models

Currency	Model	Number of Neurons	Epoch Size	Batch Size
BTC	LSTM	128	500	500
	GRU	16	1000	500
	biLSTM	32	250	500
	biGRU	16	500	250
ETH	LSTM	128	250	500
	GRU	128	1000	250
	biLSTM	32	500	100
	biGRU	16	1000	100
XRP	LSTM	128	500	500
	GRU	16	500	100
	biLSTM	64	500	500
	biGRU	128	1000	250
XMR	LSTM	64	1000	250
	GRU	64	500	250
	biLSTM	32	1000	250
	biGRU	16	500	100

4. RESULTS

To implement the proposed deep learning models Sklearn, Keras and Tensorflow libraries in Python were used. The algorithms of the models were written in Python 3.8. They are executed on the Jupyter Lab platform, on a Mac computer that has 2.6 GHz Dual-Core Intel Core i5 processor and 8 GB 1600 MHz DDR3 memory. The performance results of BTC, ETH, XRP and XMR based on the DL models are listed in Table 3.

The model that shows the smallest RMSE and MAPE values is determined as the best model. Figures 3-7 illustrate the comparisons between the actual and the predicted results of the four DL models for the currencies BTC, ETH, XRP and XMR respectively. Figures show that the pattern of the forecasted results are similar with true values of the digital currencies with some differences. These differences between the models are observed in the performance metrics shown in Table 3.

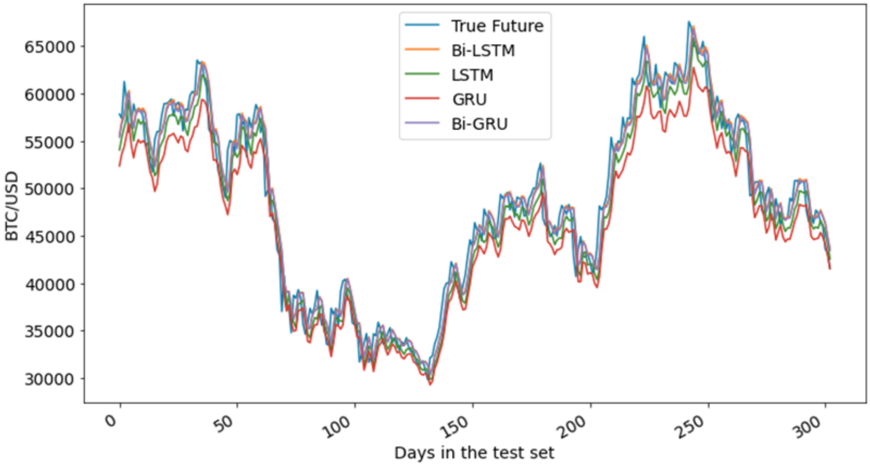
4.1. Results for BTC

The performance metric results are illustrated in Table 3. According to the BTC section of Table 3, Bi-LSTM shows the lowest RMSE and MAPE values with 1992.8766 and 3.28% respectively. Therefore Bi-LSTM is more successful in predicting BTC prices than the LSTM, GRU and the Bi-GRU model. Figure 3 shows the actual and predicted prices of BTC for the four proposed DL methods. The graph shows that the line of Bi-LSTM almost overlaps with the line of actual prices throughout the test set interval. The second-best model for BTC is bi-GRU with RMSE of 1994.1151 and MAPE value of 3.30%. These results indicate bidirectional implications of RNN networks show superior performance over the traditional implications for BTC price prediction.

Table 3. Performance Results of the Proposed Models

Currency	Model	RMSE	MAPE
BTC	LSTM	2350.5276	3.97%
	GRU	3223.0107	5.33%
	biLSTM	**1992.8766**	**3.28%**
	biGRU	1994.1151	3.30%
ETH	LSTM	183.8398	4.69%
	GRU	181.0269	4.68%
	biLSTM	**168.6**	**4.23%**
	biGRU	170.4354	4.28%
XRP	LSTM	0.0979	6.33%
	GRU	0.1042	7.21%
	biLSTM	**0.0785**	**4.84%**
	biGRU	0.0787	4.84%
XMR	LSTM	18.1997	4.76%
	GRU	18.1259	4.92%
	biLSTM	**16.5341**	**4.13%**
	biGRU	16.7791	4.30%

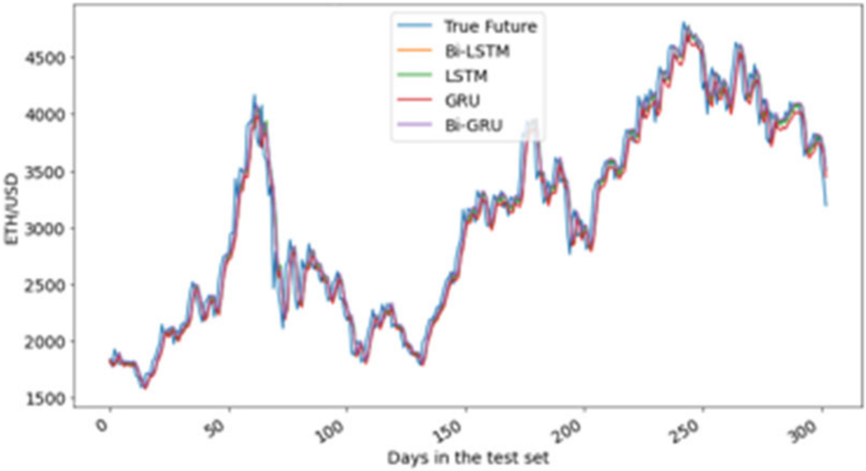
Figure 3. Actual vs Predicted Values for BTC



4.2. Results for ETH

According to ETH section of Table 3, Bi-LSTM shows the lowest RMSE and MAPE values with 168.6 and 4.23% respectively. Based on this result, Bi-LSTM is found to be more capable in forecasting BTC prices than the LSTM, GRU and the Bi-GRU models. Figure 4 illustrates the actual and predicted prices of ETH for the four proposed DL methods. The graph shows that difference between the true future and the forecasted price of ETH for the bi-LSTM model throughout the test set interval are almost negligible. The second-best model for ETH is bi-GRU with RMSE of 170.4354 and MAPE value of 4.28%. Similar to BTC, for ETH bidirectional RNN networks predicted the prices more accurately than the traditional LSTM and GRU networks.

Figure 4. Actual vs Predicted Values for ETH



4.3. Results for XRP and XMR

For the currencies XMR and XRP bi-LSTM showed the best performance compared to the other models. As Table 3 shows, RMSE values for XRP is 0.0785 and XMR is 16.5341 while MAPE values are 4.84 and 4.13 for XRP and XMR respectively. Success of bidirectional networks can also be observed for XRP and XMR, since the second-best model was Bi-GRU for both XRP and XMR. Figures 5 and 6 shows the graphs of actual and predicted values of the XRP and XMR for the four DL methods. Both figures show that the difference between the actual values and the values of Bi-LSTM is almost non-existent.

Figure 5. Actual vs Predicted Values for XRP

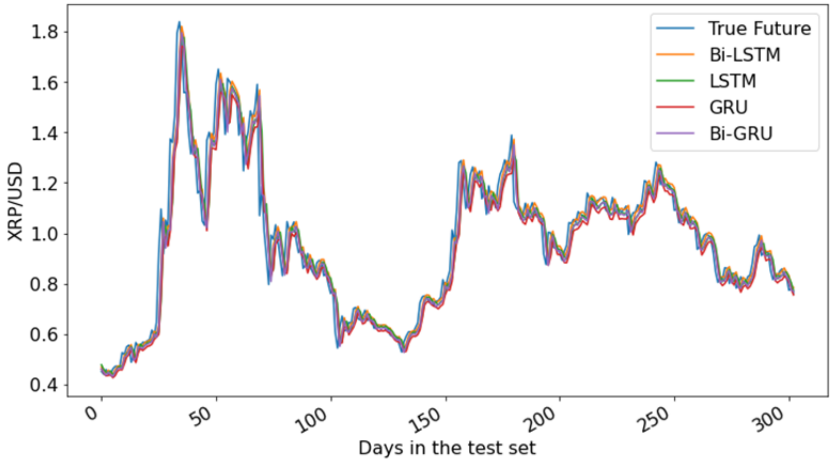
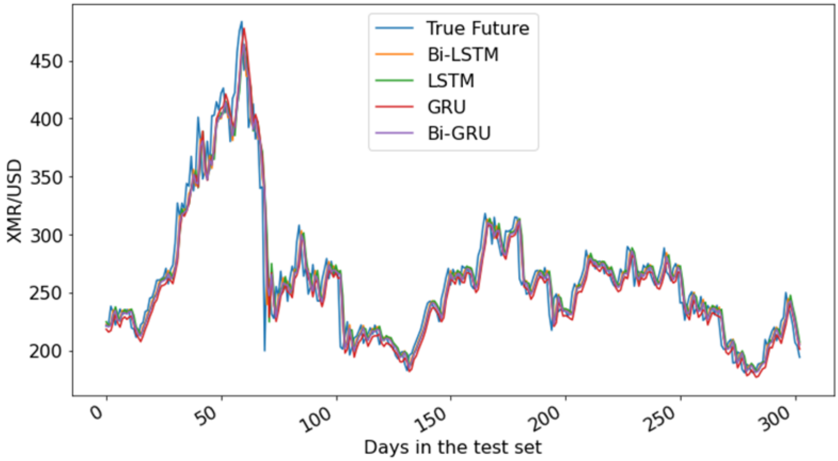


Figure 6. Actual vs Predicted values for XMR



5. CONCLUSION

In this study four types of deep learning techniques – LSTM, GRU, Bi-LSTM and Bi-GRU - are constructed and applied to the real dataset to predict the prices of four currencies that are Bitcoin (BTC), Ethereum (ETH), Ripple (XRP) and Monero (XMR). Performance scores - RMSE and MAPE - were calculated for every currency to test the accuracy of the proposed models. Results of this study indicated that Bi-LSTM provided more accurate results than LSTM, GRU and Bi-GRU when forecasting prices of all the selected currencies. Based on the findings, Bi-LSTM can be considered as the most efficient algorithm for the targeted digital currencies. The second-best model for all the currencies was observed as Bi-GRU. This finding indicates the superior performance of bi-directional implications of LSTM and GRU in time series prediction. Bi-directional models showed more accurate results than the traditional LSTM and GRU models in this study. This result indicated the success of forward and backward flow combination of bi-directional implication while performing predictions for time series data. Conclusions for this study can be summarized as follows:

- Deep learning algorithms are acceptable and efficient in cryptocurrency prediction.
- Bi-LSTM method can predict cryptocurrency prices better than LSTM, GRU and Bi-GRU.
- Overall bidirectional implications of LSTM and GRU indicate outstanding prediction results.

In future studies, different predictors of digital currency will be a topic of interest. More specifically, how tweets of people and the sentiments in the tweets will affect the prices of cryptocurrencies is aimed to be analyzed with proper machine learning techniques.

REFERENCES

Althelaya, K. A., El-Alfy, E. S. M., & Mohammed, S. (2018). Stock Market Forecast Using Multivariate Analysis with Bidirectional and Stacked (LSTM, GRU). 21st Saudi Computer Society National Computer Conference, NCC 2018. https://doi.org/10.1109/NCG.2018.8593076

Amirabadi, M. A., Kahaei, M. H., & Nezamalhosseini, S. A. (2020). Novel suboptimal approaches for hyperparameter tuning of deep neural network [under the shelf of optical communication]. Physical Communication, 41, 101057. https://doi.org/10.1016/j.phycom.2020.101057

Atsalakis, G. S., Atsalaki, I. G., Pasiouras, F., & Zopounidis, C. (2019). Bitcoin price forecasting with neuro-fuzzy techniques. European Journal of Operational Research, 276(2), 770–780. https://doi.org/10.1016/j.ejor.2019.01.040

Ayoobi, N., Sharifrazi, D., Alizadehsani, R., Shoeibi, A., Gorriz, J. M., Moosaei, H., Khosravi, A., Nahavandi, S., Gholamzadeh Chofreh, A., Goni, F. A., Klemeš, J. J., & Mosavi, A. (2021). Time series forecasting of new cases and new deaths rate for COVID-19 using deep learning methods. Results in Physics, 27(March), 104495. https://doi.org/10.1016/j.rinp.2021.104495

BitInfoCharts. (2022). Bitcoin (BTC) price stats and information. https://bitinfocharts.com/bitcoin/

Catania, L., & Grassi, S. (2017). Modelling Crypto-Currencies Financial Time-Series. 15(8).

Cavalli, S., & Amoretti, M. (2021). CNN-based multivariate data analysis for bitcoin trend prediction. Applied Soft Computing, 101, 107065. https://doi.org/10.1016/j.asoc.2020.107065

Chen, W., Xu, H., Jia, L., & Gao, Y. (2021). Machine learning model for Bitcoin exchange rate prediction using economic and technology determinants. International Journal of Forecasting, 37(1), 28–43. https://doi.org/10.1016/j.ijforecast.2020.02.008

Chowdhury, R., Rahman, M. A., Rahman, M. S., & Mahdy, M. R. C. (2020). An approach to predict and forecast the price of constituents and index of cryptocurrency using machine learning. Physica A: Statistical Mechanics and Its Applications, 551, 124569. https://doi.org/10.1016/j.physa.2020.124569

Coinmarketcap. (2022). Today's Cryptocurrency Prices by Market Cap. www.coinmarketcap.com

Dautel, A. J., Härdle, W. K., Lessmann, S., & Seow, H.-V. (2020). Forex exchange rate forecasting using deep recurrent neural networks. Digital Finance, 2(1–2), 69–96. https://doi.org/10.1007/s42521-020-00019-x

Derbentsev, V., Babenko, V., Khrustalev, K., Obruch, H., & Khrustalova, S. (2021). Comparative performance of machine learning ensemble algorithms for forecasting cryptocurrency prices. International Journal of Engineering, Transactions A: Basics, 34(1), 140–148. https://doi.org/10.5829/IJE.2021.34.01A.16

Derbentsev, V., Datsenko, N., Stepanenko, O., & Bezkorovainyi, V. (2019). Forecasting cryptocurrency prices time series using machine learning approach. SHS Web of Conferences, 65, 1–6.

DeVries, P. (2016). An Analysis of Cryptocurrency, Bitcoin, and the Future. International Journal of Business Management and Commerce, 1(2), 1–9.

Dey, R., & Salem, F. M. (2017). Gate-Variants of Gated Recurrent Unit (GRU). ArXiv:1701.05923.

Ergun, Z. C., & Karabiyik, B. K. (2021). Forecasting Monero Prices with a Machine Learning Algorithm. Eskişehir Osmangazi Üniversitesi İİBF Dergisi, 16(3), 651–663. https://doi.org/10.17153/oguiibf.932839

Greff, K., Srivastava, R. K., Koutnik, J., Steunebrink, B. R., & Schmidhuber, J. (2017). LSTM: A Search Space Odyssey. IEEE Transactions on Neural Networks and Learning Systems, 28(10), 2222–2232. https://doi.org/10.1109/TNNLS.2016.2582924

Hamayel, M. J., & Owda, A. Y. (2021). A Novel Cryptocurrency Price Prediction Model Using GRU, LSTM and bi-LSTM Machine Learning Algorithms. Ai, 2(4), 477–496. https://doi.org/10.3390/ai2040030

Hitam, N. A., & Ismail, A. R. (2018). Comparative performance of machine learning algorithms for cryptocurrency forecasting. Indonesian Journal of Electrical Engineering and Computer Science, 11(3), 1121–1128. https://doi.org/10.11591/ijeecs.v11.i3.pp1121-1128

Karasu, S., Altan, A., Sarac, Z., & Hacioglu, R. (2018). Zaman Serisi Verilerini Kullanarak Makine Öğrenmesi Yöntemleri ile Bitcoin Fiyat Tahmini. 26th IEEE Signal Processing and Communications Applications Conference, SIU 2018, 1–4.

Lai, S., Ye, C., & Zhou, H. J. H. (2021). Chinese stock trend prediction based on multi-feature learning and model fusion. 18–23. https://doi.org/10.1109/smds53860.2021.00013

Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436–444. https://doi.org/10.1038/nature14539

Liu, M., Li, G., Li, J., Zhu, X., & Yao, Y. (2020). Forecasting the price of Bitcoin using deep learning. Finance Research Letters, 40(April 2020), 101755. https://doi.org/10.1016/j.frl.2020.101755

Livieris, I. E., Kiriakidou, N., Stavroyiannis, S., & Pintelas, P. (2021). An advanced CNN-LSTM model for cryptocurrency forecasting. Electronics (Switzerland), 10(3), 1–16. https://doi.org/10.3390/electronics10030287

Livieris, I. E., Pintelas, E., Stavroyiannis, S., & Pintelas, P. (2020). Ensemble Deep learning models for forecasting cryptocurrency time-series. Algorithms, 13(5), 1–21. https://doi.org/10.3390/A13050121

Magro, P. (2017). What Greece can learn from bitcoin adoption in Latin America. Ibtimes.Co.Uk, 2–4.

Patel, M. M., Tanwar, S., Gupta, R., & Kumar, N. (2020). A Deep Learning-based Cryptocurrency Price Prediction Scheme for Financial Institutions. Journal of Information Security and Applications, 55(May), 102583. https://doi.org/10.1016/j.jisa.2020.102583

Rathan, K., Sai, S. V., & Manikanta, T. S. (2019). Crypto-currency price prediction using decision tree and regression techniques. Proceedings of the International Conference on Trends in Electronics and Informatics, ICOEI 2019, 2019-April(Icoei), 190–194. https://doi.org/10.1109/icoei.2019.8862585

Saad, M., Choi, J., Nyang, D., Kim, J., & Mohaisen, A. (2020). Toward characterizing blockchain-based cryptocurrencies for highly accurate predictions. IEEE Systems Journal, 14(1), 321–332. https://doi.org/10.1109/JSYST.2019.2927707

Schuster, M., & Paliwal, K. K. (1997). Bidirectional Recurrent Neural Networks. IEEE Transactions On Signal Processing, 45(11), 2673–2681. https://doi.org/10.1016/s1634-6939(13)59289-1

Selmi, R., Tiwari, A. K., & Hammoudeh, S. (2018). Efficiency or speculation? A dynamic analysis of the Bitcoin market Refk. Economics Bulletin, 38(4), 2037–2046.

Shahid, F., Zameer, A., & Muneeb, M. (2020). Predictions for COVID-19 with deep learning models of LSTM, GRU and Bi-LSTM. Chaos, Solitons and Fractals, 140, 110212. https://doi.org/10.1016/j.chaos.2020.110212

Singh, A., Kumar, A., & Akhtar, Z. (2021). Bitcoin Price Prediction: A Deep Learning Approach. 1053–1058. https://doi.org/10.1109/spin52536.2021.9565988

Sun, X., Liu, M., & Sima, Z. (2020). A novel cryptocurrency price trend forecasting model based on LightGBM. Finance Research Letters, 32(December 2018). https://doi.org/10.1016/j.frl.2018.12.032

Valencia, F., Gómez-Espinosa, A., & Valdés-Aguirre, B. (2019). Price movement prediction of cryptocurrencies using sentiment analysis and machine learning. Entropy, 21(6). https://doi.org/10.3390/e21060589

Wu, C. H., Lu, C. C., Ma, Y. F., & Lu, R. S. (2019). A new forecasting framework for bitcoin price with LSTM. IEEE International Conference on Data Mining Workshops, ICDMW, 2018-Novem, 168–175. https://doi.org/10.1109/ICDMW.2018.00032

Yang, N., Shi, D., & Hua, Y. (2020). Bidirectional Gated Recurrent Unit Neural Networks for Relation Extraction of Chinese Enterprises. Proceedings of 2020 IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference, ITNEC 2020, Itnec, 1539–1543. https://doi.org/10.1109/ITNEC48623.2020.9084718

Zhang, Z., Dai, H. N., Zhou, J., Mondal, S. K., García, M. M., & Wang, H. (2021). Forecasting cryptocurrency price using convolutional neural networks with weighted and attentive memory channels. Expert Systems with Applications, 183(June), 115378. https://doi.org/10.1016/j.eswa.2021.115378