The Potential for Forecasting Cryptocurrency Price Movements Using the LSTM Method

1st Dino Arnaut
Contemporary Business and IT Management
International Business-Information Academy Tuzla
Tuzla, Bosnia and Herzegovina
arnaut.dino@gmail.com

2nd Damir Bećirović
Contemporary Business and IT Management
International Business-Information Academy Tuzla
Tuzla, Bosnia and Herzegovina
damirbeci@hotmail.com

***Abstract*—The Bitcoin (BTC) is a decentralized cryptocurrency that represents a new and unique medium currency, and it is often considered as the currency of the future. Therefore, it is important to understand how Bitcoin is valued and how different factors influence its extreme volatility. Several studies had discussion whether BTC prices are predictable, but the time periods of these studies were limited by data. Therefore, we believe that a current study is necessary when we consider the volume of the BTC price movements in 2021. This research is assessing the potential for forecasting cryptocurrency price movements using machine learning approach of LSTM method and closing price of BTC in five-year period. The results show that it is possible to forecast the BTC price and even potential rise and fall using LSTM.**

***Keywords*—Forecasting, Cryptocurrency, Bitcoin, Machine Learning, LSTM, Google Colab.**

I. Introduction

Transformation of economies and financial systems in digital era is in motion and it is happening at a fast rate. New digital economy size is estimated to be 25% in 2025 which is around 23 trillion USD [1]. Current global development is overflown by Information Technologies (IT) innovations, especially in business and financial sector. The most interesting and controversial technology for creating and spending digital assets is the concept of distributed ledger technology (DLT), popularly known in its most well-recognized applicative form as Bitcoin [2].

Bitcoin (BTC) is the most popular cryptocurrency, a digital currency, that represents the alternative to the real currency. It was created by a cryptologist whose real identity is still unknown and the world knows him only as *Satoshi Nakamoto* [3]. Bitcoin became a unique digital currency with the potential to change the very nature of the everyday digital transactions, by enabling consumers person to person electronic transaction without the need for any intermediary [4].

An issue about the digital assets, such as cryptocurrencies, is price volatility. BTC price value is formed when the group of buyers and sellers that are exchanging the currency come to an agreed-upon value of traded cryptocurrency [5]. The price of BTC for the five-year period of May 2016, to May 2021, can be seen in Fig. 1. We can see that the value of a BTC in November 2016 was around 750 USD, and, still, the currency was not stable at the time. We can see spikes prices as high as 15,000 USD in 2020, as well as low as 360 USD in 2016. As we can see BTC prices have exhibited extreme volatility in previous five-year period as the price increased 1900% in 2017, after it lost 72% of its value in 2018 [6]. BTC prices exhibit extraordinary and extreme volatility, but BTC as a digital asset shows resilience as it can regain its value even after suffering significant drops, and even during COVID-19 pandemic when the uncertainty is generally high [7]. Therefore, a true concern for consumers is whether to make a purchase of an asset that had so much variation in its value.

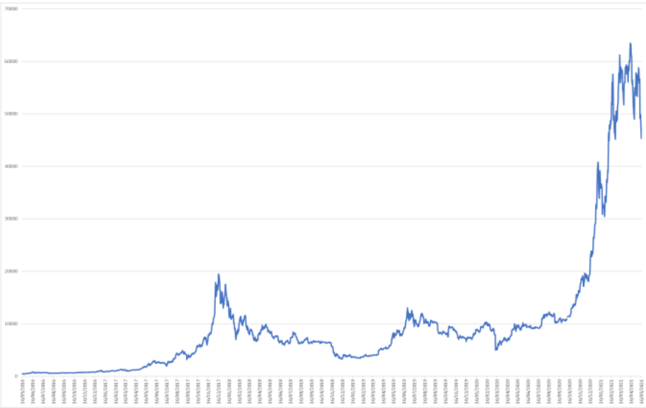

Fig. 1. BTC price - May 2016 - May 2021

Regarding its volatility the price of BTC has been an area of interest where researchers have displayed their efforts for price forecast. Studies by Liu [8] and Huang et al. [9] discussed whether BTC prices are predictable and demonstrated significant return predictability. Other studies applied various machine learning methods for end-of-day price forecast and price increase/decrease forecasting [10, 11, 12]. Achieved accuracy ranged from 63% for forecasting of increase or decrease [11] to 98% success rate for daily price forecast [12].

Felizardo et al. [13] presented a comparative study of price prediction performance among several machine learning (ML) models including long short-term memory (LSTM), WaveNet, support vector machine (SVM), and random forest (RF). The results indicated that for time-series data, the LSTM model performs better than other ML models. Tandon et al. [14] came to a similar conclusion.

These studies were all limited by the data that was available. Therefore, considering the price movements in 2021 we believe that this type of study is necessary. Best approach to use for forecasting cryptocurrencies time-series data, according to previous studies is long short-term memory (LSTM).

II. Machine Learning and Cryptocurrency Markets

A. ML and Cryptocurrency Studies

One of the first work that addressed forecasting of cryptocurrencies by using ML techniques was by Madan et al. [15], which used Binomial logistic regressions (BLR) and random forest (RF). They used five years period of data in 10 minute and 10 second frequency. Their research indicated that 10-minute data give a much better sensitivity and specificity ratio than the 10-second data.

TABLE I. Studies on Cryptocurrency price predictions using Machine Learning

Article	**Used models and main findings**		
	Year	***Model*** [a]	**Findings**
Jang & Lee [16]	2018	BLR & RF	10-min data give a better sensitivity and specificity ratio than the 10-sec data.
Nakano et al. [17]	2018	ANN	Higher performance of the ANN strategy, and results are extremely sensitive to the model specification and input data.
Atsalakis et al. [18]	2019	PATSOS	PATSOS approach outperforms other competing methods and yields a return that is significantly higher than the Buy-and-Hold strategy.
De Souza et al. [19]	2019	ANN & SVM	SVM provides conservative returns on the risk adjusted basis, & ANN generates abnormal profits during short run bull trends.
Ji et al. [20]	2019	CNN, DNN, LSTM, ResNet, CRNN & their combinations	Performances of the prediction models were comparable, LSTM is the best prediction model, DNN models are the best classification models, & classification models were more effective for trading.
Lahmiri & Bekiros [21]	2019	LSTM & GRNN	Predictability of LSTM is significantly higher than GRNN
Chen et al. [22]	2020	LR, LDA, RF, XGBoost, SVM, & LSTM	For 5-min data ML achieved better accuracy than LR and LDA, with LSTM achieving the best result (67% accuracy). For daily data, LR and LDA are better (average accuracy 65%)
Sun et al. [23]	2020	LightGBM, SVM & RF	LightGBM outperforms SVM and RF, and the accuracy is higher for 2 weeks predictions.

[a] AAN: Artificial Neural Networks
BLR: Binomial Logistic Regressions
CNN: Convolutional Neural Network
CRNN: Convolutional Recurrent Neural Network - combination of CNNs and RNNs
DNN: Deep Neural Network
GRNN: Generalized Regression Neural Networks
LDA: Linear Discriminant Analysis
LightGBM: Light Gradient Boosted Machine
LR: Logistic Regression
LSTM: Long Short-Term Memory
PATSOS: Hybrid neuro-fuzzy model
ResNet: Deep Residual Network
RF: Random Forest
SVM: Support Vector Machine
XGBoost: eXtreme Gradient Boosting

Since there is an increased interest in profiting from cryptocurrencies by using forecasting via ML techniques many papers addressed this issue. Since the list of these studies is vast, we will summarize only several studies from the last three years in which interest for this type of forecasting has gained in popularity (Table 1).

All studies from the Table 1 point out that ML models indicate higher levels of accuracy and compared to usual competing forecasting models such Exponential Moving Average and Autoregressive Integrated Moving Average (ARIMA), when it comes to predicting prices and returns of cryptocurrencies their efficiency is significantly higher. Most of studies in the last three years also compare trading strategies performance based on ML models and passive buy-and-hold strategies.

When comparing different ML models there is no clear and unambiguous winner. The only conclusion is that strategies for forecasting cryptocurrency prices that are based on ML are better.

Recurrent neural networks (RNN) are neural networks that have a temporal dimension. RNN functions by taking the current input as well as previously perceived inputs into account. Characteristics of the RNN is that it can take into account information from far back in a sequence known as long-term dependencies [24] which have hidden states that hold data from previous sequential steps.

Feed forward networks learn by using backpropagation and traditional gradient descent, but since RNN have a temporal component, they utilize a different training algorithm known as Back-Propagation-Through-Time [24]. RNN can be viewed as multiple copies of the same network that simply passes the message to a successor.

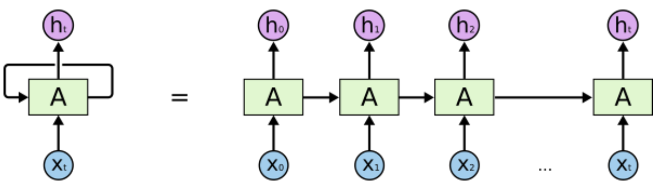

Fig. 2. RNN & vanishing gradient problem simplified

RNN represents a class of artificial neural networks where connections between the nodes form a directed graph along a temporal sequence, and, thus allowing it to exhibit temporal dynamic behavior and making them ideal for time series analysis, as well as speech recognition, grammar learning, literal compositions, etc.

If we unroll the functioning of RNN we see its problem of handling large sequences and thus vanishing gradient problem occurs (Fig. 2) which requires vast computational resources. To solve this problem LSTM RNN were created where LSTMs contain extra cells that preserve information from any time step in the sequence, that is by integrating memory units to enable learning of long temporal dynamics [25].

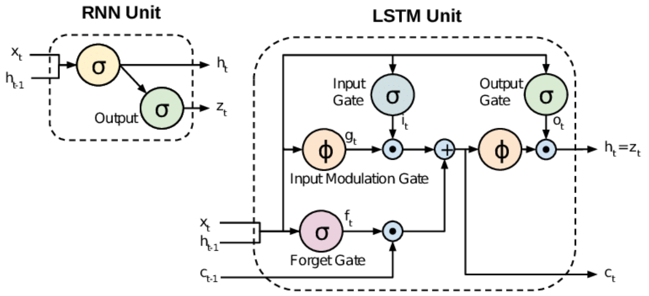

Fig. 3. RNN vs. LSTM cell

The LSTM cell represents a variant architecture of RNN cell which incorporates memory units that enable it to learn complex and long-term temporal dynamics of long sequences as it can be seen in Fig. 3.

The ability of LSTM to classify sequential data efficiently and to maintain long-term dependencies makes it well suited for price prediction of financial data and its historic data can be treated as a supervised learning problem and parsed into sequences with classification labels (buy, sell or hold patterns), or labelled with time steps in the future to train the network to classify or forecast novel sequences [26].

B. *Bitcoin Market Efficiency*

Cryptocurrency markets have become interesting for investment in the last ten years, both for institutional investors and individual investors. Cryptocurrency markets show certain characteristics very similar to other markets, primarily vulnerability to speculative bubbles. Expressed volatility and bubbles condition that cryptocurrencies can hardly be considered a true form of store of value [27]. Also, it is interesting to see how these markets, especially the Bitcoin market, behave in relation to the Efficient Market Hypothesis because the choice of an adequate investment strategy depends on it. Efficient Market Hypothesis (EMH) is one of the key theories in finance. It was developed by Fama [28], and market efficiency refers to how much the prices of financial assets reflect all available information. According to EMH, in efficient markets, prices reflect all available information, and there is little hope that the market can be won. Urquhart [29] stated that the Bitcoin market is still inefficient, but it is expected to become more efficient in the future by involving more investors. As it is a new investment asset, the Bitcoin market is similar to emerging capital markets. The Bitcoin market is highly unregulated and volatile and therefore not conducive to hedging. On the other hand, there is strong evidence against EMH [30]. The partial inefficiency of the Bitcoin market was also shown by Hirano et al. [31]. They tested EMH in the Bitcoin market from two points of view: the existence of a profitable arbitrage spread among Bitcoin exchanges and the ability to predict Bitcoin prices, i.e., the direction of daily prices (up or down) using machine learning.

On the other hand, Bartos [32] states that the price of Bitcoin follows the hypothesis of efficient markets and reacts immediately to publicly disclosed information. The price of Bitcoin is higher on days with positive events, and lower during the days with negative events compared to other days without any events. Also, Mingoti et al. [33] show that for Bitcoin prices, following a random walk cannot be ruled out indicating the existence of a low degree of efficiency.

Thus, depending on the methodology used, the size of the time series and the periods taken, there are no unique results when it comes to the efficiency of the Bitcoin market.

III. METHODOLOGY

A. *LSTM*

ML algorithms interpret the data, so there is no need to perform initial decomposition as with classical modeling. The main purpose of our study is to determine the ability of LSTM method to effectively analyze the time series data of cryptocurrencies, in this case the post popular and widely spread BTC, and to identify the patterns that form the basis for the qualitative forecasts. LSTM is an artificial recurrent neural network (RNN) architecture that we can use in the deep learning field, where we can process an entire sequence of data. Unlike the traditional neural network LSTM has a feedback connection that helps it remember preceding information, thus making it the perfect model for time series analysis.

B. *Data and Analysis Tools*

In this study we used a BTC prices dataset from Yahoo finance, that contains prices data for last five years, that is from May 16, 2016, to May 16, 2021. The data contains information about the BTC such as High, Low, Open, Close, Adjacent close and Volume. Only the day-wise closing price of the stock has been used for forecasting. We used *Google Colab* because it is a simple and powerful tool that allows us to write and execute Python. Colab enables us the full power of all popular Python libraries that we used to analyze and visualize the data.

Number of trading days in our sample were 1823 days. To make it as simple as possible we used only one variable which is the *Close* price. We decided to use *reshape* which allow us to add dimensions or change the number of elements in each dimension. In our case using *reshape (-1, 1)* was logical because we had just one dimension in our array, so *numpy* could create the same number of our rows and add one more axis, that is 1 to be the second dimension.

The normalized data was split into 80% training, 20% testing data sets. We used *MinMaxScaler* to scale our data between zero and one. Afterwards, we created the function that enabled us to create the datasets. For the features (x), we append the last 50 prices, and for the label (y), we append the next price. Then we used *numpy* to convert it into an array. Then we created our training and testing data, and we reshaped our data to make it a 3D array in order to use it in LSTM layer.

C. *Model Building*

The next step was to build our model. Finding the right model represent an art and it takes several tries and testing to find the right layers. The model we ended up with is relatively simple and it is standard for this kind of ML problem solving.

We initialized our model as a sequential one with 96 units in the output's dimensionality. We used *return_sequences=True* to make the LSTM layer with three-dimensional input and *input_shape* to shape our dataset. Making the dropout fraction 0.2 drops 20% of the layers. Also, we added a dense layer with a value of 1 because we want to output one value. After that, we reshaped our feature for the LSTM layer, because it is *sequential_3*, which is expecting 3 dimensions, not 2. We compiled our model, and we used *loss='mean_squared_error'* because it is a regression problem, and the *adam optimizer* to update network weights iteratively based on training data. In data training every *epoch* referred to one cycle through the full training dataset, and *batch size* refers to the number of training examples utilized in one iteration.

Finding the right hyper-parameters is also considered to be an art. It is important that the parameter *shuffle* was set to *False* because our analysis depends on the order of the information, so if we changed the order our results would make no sense at all.

IV. Results and Implications

In this study LSTM model was used to analyze and understand the mechanics of the BTC market. This research presents a potential way to predict Bitcoin market price. Our results show that it is possible to forecast the BTC price and even potential rise and fall using ML (Fig. 4). The theoretical and practical implications of our research are:

- Modeling results of short-term cryptocurrency dynamics and application using LSTM demonstrated the effectiveness of using ML approach, especially RRN, for forecasting;
- Conducted model simulation have confirmed the feasibility of using the ML for the short-term forecasting of financial time series;
- Constructed models have potential basis for creating algorithms for automated trading systems.

We recommend combing different ML prediction models to create more accurate cryptocurrencies price forecasts.

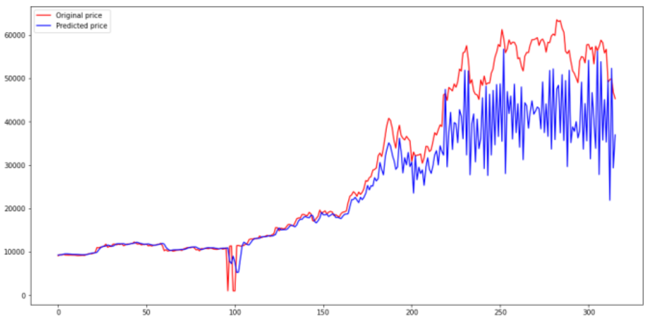

Fig. 4. Original vs. predicted BTC price

It can be difficult to determine whether LSTM model is performing well on given sequence prediction problem like price of Bitcoin. Therefore, it is important to know whether tested model is a good fit for our used data or if it is underfit or overfit and would it do better with a different configuration.

LSTM models are trained by using the *fit()* function. This function returns a variable *history* that contains a trace of the loss as well as any other metrics that were specified during the compilation of the model. These scores are recorded at the end of each epoch. *Keras* allowed us to specify a separate validation dataset while fitting our model. We done this by setting the *validation_split* argument on *fit()* function to use a 10% of the training data as a validation dataset.

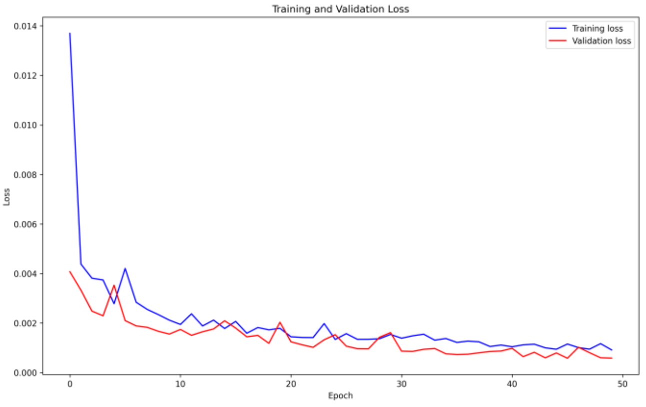

Fig. 5. Training and validaiton loss

A good model fit represents a case where the performance of the model is good on both the train and validation sets. This can be diagnosed from a plot where the train and validation loss both decrease and stabilize around the same point as we can see from Fig. 5. As we can see our model demonstrated an LSTM model with a good fit.

Future research should investigate the use of artificial intelligence for modeling the price of any cryptocurrency which would enable easier measuring of the risk factor for the justification of financial usage of blockchain technology.

Also, it would be necessary to consider adding external data inputs that are related to global events to detect price anomalies/spikes and to use ML to predict and assess the stability of cryptocurrencies.

Overall, there is ample empirical evidence that at the current stage the Bitcoin market is still not complying with EMH. From the investor's point of view, this means that it is possible to identify undervalued or overvalued assets in the market using technical analysis. This indicates that an active investment strategy in the Bitcoin market is possible, although many investors today view it as a means of preserving value and opt for passive investing. Machine learning can be seen as an instrument of technical analysis that can help investors make better investment decisions.

References

[1] W. Xu, A. Cooper, "Digital spillover: measuring the true impact of the digital economy," Huawei and Oxford Economics.

[2] B. Imran, "Mastering Blockchain: Distributed Ledger Technology, Decentralization, and Smart Contracts Explained," 2nd Edition, 2018.

[3] N. Satoshi. "Bitcoin: A Peer-to-Peer Electronic Cash System," 2008.

[4] J. Brito, "Bitcoin: Examining the Benefits and Risks for Small Business," 2014 [Statement from Jerry Brito].

[5] P. Pagnottoni, and T. Dimpfl, "Price discovery on Bitcoin markets," Digital Finance 1: 139–61, 2019.

[6] D. Z. Morris, "Bitcoin Hits a New Record High, But Stops Short of $20,000," Fortune, 2017.

[7] S. Shukla, and S. Dave, "Bitcoin beats coronavirus blues," The Economic Times, 2020.

[8] L. Liu, "Are Bitcon returns predictable?: Evidence from technical indicators," 2019.

[9] J. Z. Huang, W. Huang, and J. Ni, "Predicting Bitcoin returns using

high-dimensional technical indicators," J. Finance Data Sci.,5(3):140–155, 2018.

[10] Z. Chen, C. Li, and W. Sun, "Bitcoin price prediction using machine learning: an approach to sample dimension engineering," J. Comput. Appl. Math., 365:112395, 2020.

[11] R. Adcock, and N. Gradojevic, "Non-fundamental, non-parametric Bitcoin forecasting," Physica A: Stat. Mech. Appl. 531:121727, 2019.

[12] D. C. Mallqui, and R. A. Fernandes, "Predicting the direction, maximum, minimum and closing prices of daily Bitcoin exchange rate using machine learning techniques," Appl. Soft. Comput. 75:596–60, 2019.

[13] L. Felizardo, R. Oliveira, E. Del-Moral-Hernandez, and F. Cozman, "Comparative study of Bitcoin price prediction using WaveNets, Recurrent Neural Networks and other Machine Learning Methods," 2019 [Paper presented at 6th International Conference on Behavioral, Economic and Socio-Cultural Computing (BESC), Beijing, China, October 28–30.]

[14] S. Tandon, S. Tripathi, P. Saraswat, and C. Dabas, "Bitcoin Price Forecasting using LSTM and 10-Fold Cross validation," 2019 [Paper presented at the 2019 International Conference on Signal Processing and Communication (ICSC), Noida, India, March 7–9.]

[15] I. Madan, S. Saluja S, and A. Zhao, "Automated bitcoin trading via machine learning algorithms," 2015.

[16] H. Jang, and J. Lee, "An empirical study on modeling and prediction of Bitcoin prices with Bayesian neural networks based on blockchain information," 2018.

[17] M. Nakano, A. Takahashi, and S. Takahashi, "Bitcoin technical trading with artificial neural network," Phys A. 510:587–609, 2018.

[18] G.S. Atsalakis, I.G. Atsalaki, F. Pasiouras, and C. Zopounidis, "Bitcoin price forecasting with neuro-fuzzy techniques," Eur J Oper Res 276(2):770–780, 2019.

[19] M.J.S. De Souza, F.W. Almudhaf, B.M. Henrique, A.B.S. Negredo, D.F.G. Ramos, V.A. Sobreiro, and H. Kimura, "Can artificial intelligence enhance the Bitcoin bonanza," J Finance Data Sci 5(2):83–98, 2019.

[20] S. Ji, J. Kim, and H. Im, "A comparative study of bitcoin price prediction using deep learning," Mathematics 7(10):898, 2019.

[21] S. Lahmiri,and S. Bekiros, "Cryptocurrency forecasting with deep learning chaotic neural networks," Chaos Solitons Fractals 118:35–40, 2019.

[22] Z. Chen, C. Li, and W. Sun, "Bitcoin price prediction using machine learning: an approach to sample dimension engineering," J Comput Appl Math 365:112395, 2020.

[23] X. Sun, M. Liu, and Z. Sima, "A novel cryptocurrency price trend forecasting model based on LightGBM," Finance Res Lett 32:101084, 2020.

[24] H. Salehinejad, S. Sankar, J. Barfett, E. Colak, and S. Valaee, "Recent advances in recurrent neural networks," arXiv preprint arXiv: 1801.01078, 2017.

[25] S. Hochreiter, "Long Short-Term Memory," Neural Computation, 9(8), 1735-1780, 1997.

[26] J. Sirignano, and R. Cont, "Universal Features of Price Formation in Financial Markets: Perspectives From Deep Learning," 2018.

[27] E.T. Cheah, and J. Fry, "Speculative bubbles in Bitcoin markets? An empirical investigation into the fundamental value of Bitcoin," Economics Letters, 130, pp. 32-36, 2015.

[28] E.F. Fama, "Efficient capital markets: a review of theory and empirical work," Journal of Finance, 25(2), pp. 383-417, 1970.

[29] A. Urquhart, "The inefficiency of Bitcoin. Economics Letters," 148. pp. 80-82, 2016.

[30] S. Lahmiri, S. Bekiros, and A. Salvi, "Long-range memory, distributional variation and randomness of bitcoin volatility," Chaos solitons & fractals, 107, pp. 43-48, 2018.

[31] Y. Hirano, L. Pichl, E. Cheoljun, and T. Kaizoji, "Analysisi of bitcoin market efficiency by using machine learing," 2018 [CBU International conference on innovations in science and education, Prague.]

[32] J. Bartos, "Does Bitcoin folow the hypothesis of efficient market?," International Journal of Economic Sciences, 4(2), pp. 10-23, 2015.

[33] N. Mingotti, L. Rosa, and J. Romo, "A Random Walk Test for Functional Time Series," DES - Working Papers. Statistics and Econometrics. WS, Universidad Carlos III de Madrid. Departamento de Estadística, 2015.