Bitcoin price prediction and machine learning features: New financial scenarios

Tiziana Ciano, Department of Economics and Political Science, University of Aosta Valley, Aosta, Italy

Key points

- Technological advancements have led to a shift toward a digital economy, with cryptocurrencies being introduced into payment systems and various algorithms used for price prediction.
- Bitcoin's value varies like financial assets, but its dynamic process innovates market characteristics.
- Machine learning technology is suggested to predict bitcoin price
- This study provides forecasting analysis to help investors predict future prices and review their portfolios.

Abstract

Technological advances have reshaped our society, leading to a shift toward a more digital economy. In this work, a machine learning model is developed using neural networks to make predictions about the price of Bitcoin. A training dataset is created that contains closing price values from the last 60 days to predict the closing price value of day 61. This short lapse time was chosen to point out the high volatility of this kind of market as well. Performing an in-depth study on the correlation between cryptocurrencies strengthens the analysis for driving the read into this framework related to prediction for a complex financial scenario.

Introduction

Technological evolution over the years has made "disruptive" innovation advances that are redefining the structural characteristics of our society. An evolution that pushes the world economy to move toward an increasingly digital future. Over the course of the 21st century, we have witnessed several changes in the economic and financial sector. One of the most relevant and impactful is certainly the one that has introduced a digital innovation in payment systems, namely the creation of "cryptocurrencies." They are, therefore, "digital representations of value" that are not subject to issuance, guarantee or control by Central Banks or Public Authorities. These are currencies typically issued by private issuers that use highly specialized software and, generally, blockchain technologies. Their management is usually done through virtual wallets called "e-wallets." Being based on Blockchain technology, which is incredibly complicated and tries to store data in such a way that it is difficult and beyond expectations to hack and modify it. "Encryption" further protects these currencies, making it difficult to create fraudulent cryptocurrencies. In general, cryptocurrencies can be converted into "fiat currencies" at exchange rates that vary over time, but they should not be confused with electronic payment systems. The system in the cryptocurrency market is quite complex and quite difficult to understand, even for industry players and researchers carrying out studies in this field (Fry and Cheah, 2016).

There are a lot of cryptocurrencies flowing into the digital market. Among the various cryptocurrencies, Bitcoin is the most well-known and is influenced by external variables such as social sites, digital data, market analysis, and so on. It was first introduced by Nakamoto in 2008 (Giudici et al., 2020). Peer-to-peer transactions are possible thanks to blockchain technology. This technology, which underpins the Bitcoin cryptocurrency system, is largely critical to ensuring greater security and privacy in a variety of sectors, including IoT (Arias-Oliva et al., 2019).

Bitcoin is the first digital currency in the world to have used the blockchain platform. Böhme et al. (2015) suggested that cryptocurrency, particularly Bitcoin, is more of a payment platform rather than a currency due to its real-time convertibility into a conventional fixed-value currency. This cryptocurrency is different from other assets in terms of portfolio analysis, risk management, and sentiment analysis (Dyhrberg, 2016).

There have been many researchers who have revealed the advantages of Bitcoin such as security (Bariviera et al., 2017), low transaction cost (Kim, 2017), high yield (Ciaian et al., 2016; Kristoufek, 2013; Hong, 2017) and as an alternative tool for a country's bailout mechanism (Bouri et al., 2017) and use for employee wages (Angel and McCabe, 2015). Despite this, there are also researchers who point out the risks and disadvantages of using this digital currency, in terms of lack of regulation (Cheung et al., 2015; Böhme et al., 2015), high electricity bills due to energy consumption (Hayes, 2017; Vranken, 2017), lack of security (Bradbury, 2013; Conte De Leon et al., 2017) and other issues such as anonymity (Androulaki et al., 2013) and the cost of switching (Luther, 2016) to predict cryptocurrency market movements.

The relationship between machine learning algorithms and cryptocurrency is considered a new field with limited research studies. However, the selection of the appropriate machine for learning algorithms depends on the frequency and structure of the data, otherwise the complexity of such models could lead to overfitting. Previous studies have implemented well-known machine learning algorithms such as neural and recurrent neural network (RNN), long-term memory (LSTM), support vector machines (SVMs), and random forest models.

Most research in this field focuses only on accuracy when using machine learning algorithms while ignoring sample size (Chen et al., 2020; McNally et al., 2018). Some studies on Bitcoin have predicted its price using two approaches: empirical analysis and analysis of machine learning algorithms.

In this note, a machine learning model is developed using an Artificial Neural Network (ANN), i.e., an LSTM, to predict the closing price of bitcoins using the price of bitcoins over the last 60 days. The analysis is strengthened by carrying out a careful study of the "correlation" between cryptocurrencies. The aim of this training to connect the reader to this frame of research useful to financial practitioners.

The rest of this work is organized as follows: in **Cryptocurrency price prediction** section there is an overview of the machine learning models and techniques applied to the world of cryptocurrencies. Sections **Bitcoin price predictions: A case study model** and **"Correlation" between different cryptocurrencies: A modeling through Python** describe the analysis of our forecasting model related to the analysis of the correlation between cryptocurrencies.

Cryptocurrency price prediction

The value of Bitcoin changes just like a financial asset. However, in this case, this dynamic process presents some insubstantial peculiarities such as to profoundly innovate the characteristics of the market that observes these "dynamics." There are several algorithms used on stock market data for price prediction. On the other hand, the parameters that influence Bitcoin are so different. Therefore, you need to predict the value of Bitcoin so that you can make correct investment decisions. The price of Bitcoin does not depend on corporate events or government intervention unlike the stock market. Therefore, in order to predict its value, we believe it is advisable to leverage machine learning technology to predict the price of Bitcoin. After the "boom" and fall in cryptocurrency prices in recent years, Bitcoin has increasingly been viewed as an "investment asset." Due to its highly volatile nature, there is a need for good forecasts on which to base investment decisions. Although existing studies have leveraged machine learning for more accurate prediction of Bitcoin prices, few have focused on the feasibility of applying different modeling techniques to samples with different data structures and characteristics. Due to price fluctuations and instability, cryptocurrency prices are difficult to predict. This gap in the field is filled by comparing different machine learning models to predict market movements of the most relevant cryptocurrency. In fact, many researchers have studied various factors that affect the price of Bitcoin and the patterns underlying its fluctuations using various machine learning methods. By implementing time-series analytical mechanisms with daily Bitcoin data, Patel et al. (2015) conducted an empirical analysis of the determinants of Bitcoin's price within a Barro model (Cortes and Vapnik, 1995). Their findings revealed that among the traditional determinants and specific factors of digital currencies, market forces and investor attractiveness most determine Bitcoin's price, disproving those financial developments at the macro level are Bitcoin's drivers for a long time.

Using a nonparametric quantile causality test, revealed that volume can be a predictor of Bitcoin returns, suggesting the importance of modeling nonlinearity and accounting for tail behavior (Balcilar et al., 2017), Kim et al. (2016) analyzed user comments in online cryptocurrency communities for Bitcoin, Ethereum, and Ripple to predict price fluctuations and transaction numbers. Mai et al. (2015) examined the dynamic interactions between the value of Bitcoin and social media based on the information system and financial literature. Shah and Zhang (2014) used a Bayesian regression "latent source model" created by Chen et al. (2020), designed to leverage binary classification to predict Bitcoin price changes. Using a supporting vector machine algorithm, Georgoula et al. (2015) examined the relationship between Bitcoin's price and determinants, including economic variables, technological factors, and sentiment. Madan et al. (2015) applied machine learning algorithms to predict the price of Bitcoin with an accuracy of 98.7% for the daily price and 50%—55% for the high-frequency price. McNally et al. (2018) compared the accuracy of Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), and autoregressive Integrated Moving Average (ARIMA) models for Bitcoin price prediction and showed that LSTM achieves the highest accuracy (52%). Chen et al. (2020) explores the use of machine learning techniques to predict Bitcoin's daily prices, focusing on extending feature sizes and evaluating different techniques. It identifies higher-dimensional features for predicting daily Bitcoin prices, classifying Bitcoin price data by range, and evaluating different machine learning techniques.

Due to recent advances in machine learning, many deep learning-based prediction models have been proposed for the price of Bitcoin (McNally et al., 2018; Saad et al., 2019; Jang and Lee, 2017; Nakano et al., 2018; Rebane et al., 2018; Huisu et al., 2018;

Shintate and Pichl, 2019). Many researchers have studied various factors influencing the price of Bitcoin and the patterns behind its fluctuations using various analytical and experimental methods; for example, see the works of the authors of Alessandretti et al. (2018), Corbet et al. (2019). Kim et al. (2016) study and compare various state-of-the-art deep learning methods, such as DNNs, long-term memory models (LSTMs) (Hochreiter and Schmidhuber, 1997) CNN, ResNet, a combination of CNN and RNN (CRNN) (Wang et al., 2016), and their ensemble models for Bitcoin price prediction. Specifically, they developed both regression and classification models by leveraging information on the Bitcoin blockchain and comparing their prediction performance in various settings. McNally et al. (2018) proposed two prediction models based on recurrent neural networks (RNNs) and long-term memory (LSTMs) and compared them with an autoregressive integrated moving average (ARIMA) model (Box et al., 2015), which is a time series traditionally widely used as a prediction model. They developed ranking models using Bitcoin price information, which predicts whether Bitcoin's next price will go up or down based on previous prices. In the work of the authors of McNally et al. (2018), the RNN and LSTM models proved to be better than the ARIMA model. In addition to price information, Saad et al. (2019) analyzed information about the Bitcoin blockchain, such as the number of Bitcoin wallets and unique addresses, block mining difficulty, hash rates, etc., and used those features that are highly correlated with the Bitcoin price to build prediction models. They studied several regression models based on linear regression, random forests (Ho, 1995), gradient enhancement and (Freund and Schapire, 1997) and neural networks.

In addition to blockchain information, Jang and Lee (2017) further considered macroeconomic factors such as the S&P 500, Euro stoxx 50, DOW 30, and NASDAQ, and exchange rates between major fiat currencies. They studied three prediction models based on a Bayesian neural network (BNN) (Bishop and Nasrabadi, 2006), linear regression and supporting vector regression (SVR) (Cortes and Vapnik, 1995), and showed that the BNN model outperformed the other two prediction models. In their follow-up study, Huisu et al. (2018) proposed a moving window LSTM model and showed that it outperformed prediction models based on linear regression, SVR, neural networks, and LSTM. Similarly, Shintate and Pichl (2019) proposed a deep learning-based random sampling model and showed that it outperformed LSTM-based models. Meanwhile, Kim et al. (2016) and Li et al. (2019) took social data into account to predict Bitcoin price fluctuations. Rathore et al. (2022) in their study apply the Fbprophet model, which is better in terms of functionality and can handle "seasonal data." This approach provides a methodology for predicting the future price of bitcoin, addressing seasonality in historical data, and reducing the difference between predicted and actual values even with seasonal data. The work of Ranjan et al. (2023) aims to predict the price of Bitcoin using daily data and high-frequency data of accessibility, like stock market forecasting (Madan et al., 2015).

Bitcoin price predictions: A case study model

The goal of this section is to predict the price of Bitcoin using a machine learning technique called Long Short-Term Memory (LSTM). You want to predict the closing price of bitcoins using the bitcoin price of the last 60 days. In receiving the bitcoin price quote, we use the company stock exchange ticker BTC-USD from January 1, 2018, to November 20, 2023.[1]

Fig. 1 shows the dates, the opening price, the highest price, the lowest price, the closing price, the adjusted close, and finally the bitcoin volume for the period considered.

Date	Open	High	Low	Close	Adj Close	Volume
2018-01-01	14112.200195	14112.200195	13154.700195	13657.200195	13657.200195	10291200000
2018-01-02	13625.000000	15444.599609	13163.599609	14982.099609	14982.099609	16846600192
2018-01-03	14978.200195	15572.799805	14844.500000	15201.000000	15201.000000	16871900160
2018-01-04	15270.700195	15739.700195	14522.200195	15599.200195	15599.200195	21783199744
2018-01-05	15477.200195	17705.199219	15202.799805	17429.500000	17429.500000	23840899072
...	...	...	...	...	...	...
2023-11-15	35548.113281	37964.894531	35383.781250	37880.582031	37880.582031	27365821679
2023-11-16	37879.980469	37934.625000	35545.472656	36154.769531	36154.769531	26007385366
2023-11-17	36164.824219	36704.484375	35901.234375	36596.683594	36596.683594	22445028430
2023-11-18	36625.371094	36839.281250	36233.312500	36585.703125	36585.703125	11886022717
2023-11-19	36585.765625	37509.355469	36414.597656	37386.546875	37386.546875	12915986553

2149 rows × 6 columns

Fig. 1 Bitcoin price quote.

[1] https://it.finance.yahoo.com/quote/BTC-USD?p=BTC-USD&.tsrc=fin-srch.

```
Import Math
Importing pandas_datareader as Web
Import numpy as NP
Import panda as PD
from sklearn.preprocessing import MinMaxScaler
by keras.models import Sequential
from keras.layers import Denso, LSTM
Import matplotlib.pyplot as PLT
plt.style.use('fivethirtyeight')


Import yfinance as yf
From Yfinance Import Download
download=yf.download("BTC-USD", start='2018-01-01', end=   '20-11-2023')
unload
```

Our dataset consists of 2149 rows and 6 columns. **Graph 1** shows the closing price history of Bitcoin.

It was creating a new data frame with only the closing price and convert it to an array. So, let's create a variable to store the length of the training dataset. So, I establish that the training dataset contains about 80% of the data.

```
data = downloads. filter(['Close'])
dataset = data.values
training_data_len = math.ceil( len(dataset) *.8)
```

The dataset was scaled so that it is values into 0 and 1 inclusive, before giving them to the neural network.

```
scaler = MinMaxScaler(feature_range=(0, 1))
scaled_data = scaler.fit_transform(dataset)
```

Next, a training dataset was arranged that contains the closing price values from the last 60 days that it is used to use to predict the value of the 61st closing price. Then, the first column in dataset "x_train" contains the values from the dataset from index 0 to index 59 (60 values in total), and the second column contains the values from dataset from index 1 to index 60 (60 values), and so on. The dataset "y_train" contains the 61st value located at index 60 for its first column and the 62nd value located at index 61 of the dataset for its second value, and so on.

Graph 1 Bitcoin closing prices.

```
train_data = scaled_data[0:training_data_len , : ]
x_train=[]
y_train = []
For the range(60,len(train_data)):
x_train.append(train_data[i-60:i,0])
y_train.append(train_data[i,0])
```

You need to convert the independent dataset "x_train" and the dependent dataset "y_train" into numpy arrays so that they can be used for training the LSTM model.

```
x_train, y_train = np.array(x_train), np.array(y_train)
```

At this step it must be reshaping in a three-dimensional way in the form: number of samples, number of time steps, number of features. This step is pivotal because the LSTM model expects a three-dimensional dataset:

```
x_train = np.reshape(x_train, (x_train.shape[0],x_train.shape[1],1))
```

The LSTM model was arranged to work by two "layers" equipped with 50 neurons and two "dense" layers, one with 25 neurons and the other with 1 neuron output.

```
model = Sequential()
model.add(LSTM(units=50, return_sequences=True,input_shape=(x_train.shape[1],1)))
model.add(LSTM(units=50, return_sequences=False))
model.add(Dense(units=25))
model.add(Dense(unit=1))
```

Hence the model by using the mean square error loss function and "Adam optimization approach":

```
model. compile(optimizer='Adam', loss='mean_squared_error')
```

And after by using training datasets. Batch size is the total number of training samples present in a single batch, and epoch is the number of iterations in which an entire dataset is passed back and forth through the neural network.

```
model.fit(x_train, y_train, batch_size=1, epoche=10)
```

We get a test dataset:

```
test_data = scaled_data[training_data_len - 60: , : ]
x_test = []
y_test = dataset[training_data_len : , : ]
For the in range(60,len(test_data)):
x_test.append(test_data[i-60:i,0])
```

Then the independent test dataset "x_test" was converted into a numpy array so that it can be used to test the LSTM model.

```
x_test = np.array(x_test)
```

Reshape the data in a three-dimensional way in the form [number of samples, number of time steps, and number of features.

```
x_test = np.reshape(x_test, (x_test.shape[0],x_test.shape[1];1))
```

This results in the predicted values of the model using the data from the test set.

```
Predictions = model.predict(x_test)
Predictions = scaler.inverse_transform(Tips)
```

To measure the accuracy of the model, the mean square error (RMSE) was considered. A value of 0 indicates that the values predicted by the models perfectly match the actual values of the test dataset. It was gained a value of 0.669. In Graph 2 a qualitative analysis of the prices training.

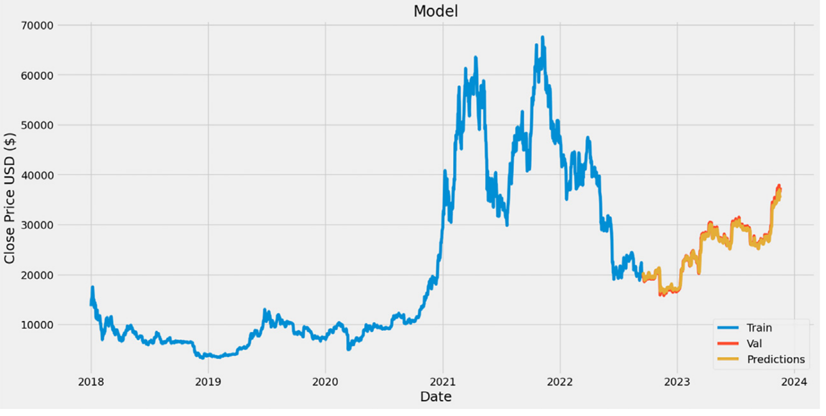
```
train = data[:training_data_len]
valid = data[training_data_len:]
valid['Predictions'] = predictions
#Visualize the data
plt.figure(figsize=(16,8))
plt.title('Template')
plt.xlabel('Data', fontsize=18)
plt.ylabel('USD ($)' closing price, fontsize=18)
plt.plot(train['Close'])
plt.plot(valid[['Close',  'Predictions']])
plt.legend(['Train', 'Val',  'Forecast'], loc='bottom right')
plt.show()
```

Fig. 2 shows the present (closing) and expected (forecast) price values.

The model is tested on a daily basis in order to get the value of Bitcoin's predicted closing price for November 21, 2023. Therefore, the data was converted into an array that contains only the closing price. Then the closing price of the last 60 days was got and scaling the data so that are values between 0 and 1 inclusive. After that an empty list was introduced and adding the price of the last 60 days, then convert itself to a jumpy array and reshape it so the data can be inserted into the model. Finally, the expected price equal to $ 36,526,207:

```
download_btc = yf.download('BTC-USD', Start  = '2018-01-01', End =  '20-11-2023')
new_download = download_btc. filter(['Close'])
last_60_days = new_download[-60:].values
last_60_days_scaled = scaler.transform(last_60_days)
X_test = []
X_test.append(last_60_days_scaled)
X_test = np.array(X_test)
X_test = np.reshape(X_test, (X_test.shape[0], X_test.shape[1], 1))
pred_price = model.predict(X_test)

pred_price = scaler.inverse_transform (pred_price)
Press (pred_price)
```



Graph 2 Predicted value visualizations.

Date	Close	Predictions
2022-09-17	20127.576172	19944.712891
2022-09-18	19419.505859	20247.564453
2022-09-19	19544.128906	19700.478516
2022-09-20	18890.789062	19729.210938
2022-09-21	18547.400391	19197.214844
...	...	...
2023-11-15	37880.582031	34891.515625
2023-11-16	36154.769531	36934.226562
2023-11-17	36596.683594	35452.035156
2023-11-18	36585.703125	35827.632812
2023-11-19	37386.546875	35806.289062

429 rows × 2 columns

Fig. 2 Actual and forecast price values.

It was observed that the price of bitcoin on November 21st is equal to $ 35813.81265.

```
download_btc2 = yf.download('BTC-USD', start='2023-11-21', end='2023-11-22')
print(download_btc2['Close'])
```

The analysis shows a convergent forecast.

"Correlation" between different cryptocurrencies: A modeling through python

In the cryptocurrency's framework, correlation analysis is a very significant "metric" that can support investors in minimizing risk and maximizing returns. The correlation of cryptocurrencies is a central topic in the world of digital asset investing. Many people believe that investing in assets that move synchronously with others produces some benefits. By investing in a portfolio of correlated assets, investors can potentially reduce their overall risk, while maintaining a safe return profile.

Technically, it is usually known that the correlation is based on a scale of 1 to −1. The closer the correlation coefficient is to 1, the higher it is, i.e., positive. On the other hand, with a negative correlation approaching −1, it means that the trend of the two underlying assets goes in different directions.

Starting from this premise and after the analysis of Bitcoin price predictions, we considered it essential to highlight that, even though Bitcoin "leads" the world of cryptocurrencies in most situations, this does not exclude that there are cryptocurrencies, which sometimes move in an uncorrelated way from its trend. Therefore, a comparison based on the correlation coefficient between Bitcoin and the main cryptocurrencies on the market was reported in this work.

I have considered the older cryptocurrencies and, I have downloaded the series stories from yahoo finance related to the following cryptocurrencies: Bitcoin USD, Ethereum USD, Binance Coin USD, Litecoin USD, Cardano USD and XRP USD considering a 5-year time horizon, i.e., from January 01, 2018, to November 20, 2023 (see Fig. 3).

Only the modified closure is considered, so we save it in another data frame and display it (see Fig. 4).

In finance practices, correlation is conducted not on prices understood as an absolute value but on returns, i.e., on the price changes in percentage that security has made. For example, the change of 1 on a cryptocurrency that costs 35 versus a change of 1 on a cryptocurrency that costs 1200 must have a different percentage value.

To do this, we use python's *pct.change* function. This method gives us the difference between one row and the previous one in percentage terms. Let's visualize this new data frame and see that, apart from the first day when we don't have any previous rows to compare with, for all the others, we got the performance of cryptocurrencies day by day (see Fig. 5).

So, let's calculate the correlation between them thanks to a *corr* function of pandas. The correlation can be made in different ways but we use the Pearson index understood as the covariance between two variables divided by the product of the standard deviation of the two variables and is a value ranging from −1 to 1, where −1 means that the variables are inversely proportional, that is, as one increases, the other decreases; 1 means that they are perfectly correlated, as one grows, the other grows as well, and zero which means that the two variables are unrelated to each other (there is no relationship). Hence, let's visualize the new data frame in Fig. 6.

	Adj Close						Close				...	Open				Volume				
	ADA-USD	BNB-USD	BTC-USD	ETH-USD	LTC-USD	XRP-USD	ADA-USD	BNB-USD	BTC-USD	ETH-USD	...	BTC-USD	ETH-USD	LTC-USD	XRP-USD	ADA-USD	BNB-USD	BTC-USD	ETH-USD	LTC-USD
Date																				
2018-01-01	0.728657	8.414610	13657.200195	772.640991	229.033005	2.391030	0.728657	8.414610	13657.200195	772.640991	...	14112.200195	755.757019	231.666000	2.296020	150186000	66422800	10291200000	2595760128	633142016
2018-01-02	0.782587	8.837770	14982.099609	884.443970	255.684006	2.480900	0.782587	8.837770	14982.099609	884.443970	...	13625.000000	772.346008	228.990005	2.369480	289712000	104789000	16846600192	5783349760	1237949952
2018-01-03	1.079660	9.535880	15201.000000	962.719971	245.367996	3.105370	1.079660	9.535880	15201.000000	962.719971	...	14978.200195	886.000000	255.695007	2.464100	657398016	108852000	16871900160	5093159936	3215280128
2018-01-04	1.114120	9.213990	15599.200195	980.921997	241.369995	3.196630	1.114120	9.213990	15599.200195	980.921997	...	15270.700195	961.713013	245.475006	3.117340	593430016	158819008	21783199744	6502859776	3481550080
2018-01-05	0.999559	14.917200	17429.500000	997.719971	249.270996	3.048710	0.999559	14.917200	17429.500000	997.719971	...	15477.200195	975.750000	241.033997	3.300810	508100000	341504992	23840899072	6683149824	1710599936
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
2023-11-15	0.379292	253.755341	37880.582031	2060.408447	74.080322	0.649648	0.379292	253.755341	37880.582031	2060.408447	...	35548.113281	1979.472656	70.635925	0.629726	400207664	921052035	27365821679	12626326991	335516203
2023-11-16	0.371766	242.758881	36154.769531	1960.881592	71.179054	0.612168	0.371766	242.758881	36154.769531	1960.881592	...	37879.980469	2059.965820	74.080063	0.649609	784296833	950453378	26007385366	14651619483	343373015
2023-11-17	0.366609	244.745361	36596.683594	1961.280762	70.182335	0.613717	0.366609	244.745361	36596.683594	1961.280762	...	36164.824219	1961.867554	71.149132	0.612199	419332542	788884171	22445028430	11881648738	406903529
2023-11-18	0.377694	244.950760	36585.703125	1963.285034	69.893806	0.611189	0.377694	244.950760	36585.703125	1963.285034	...	36625.371094	1961.671265	70.191620	0.613738	350054655	679630905	11886022717	8064677046	243237870
2023-11-19	0.384525	246.587784	37386.546875	2013.204468	70.640312	0.627499	0.384525	246.587784	37386.546875	2013.204468	...	36585.765625	1963.180054	69.895370	0.611187	308083348	618053586	12915986553	7716048818	226121115

2149 rows × 36 columns

Fig. 3 Cryptocurrency datasets.

	ADA-USD	BNB-USD	BTC-USD	ETH-USD	LTC-USD	XRP-USD
Date						
2018-01-01	0.728657	8.414610	13657.200195	772.640991	229.033005	2.391030
2018-01-02	0.782587	8.837770	14982.099609	884.443970	255.684006	2.480900
2018-01-03	1.079660	9.535880	15201.000000	962.719971	245.367996	3.105370
2018-01-04	1.114120	9.213990	15599.200195	980.921997	241.369995	3.196630
2018-01-05	0.999559	14.917200	17429.500000	997.719971	249.270996	3.048710
...	...	...	...	...	...	...
2023-11-15	0.379292	253.755341	37880.582031	2060.408447	74.080322	0.649648
2023-11-16	0.371766	242.758881	36154.769531	1960.881592	71.179054	0.612168
2023-11-17	0.366609	244.745361	36596.683594	1961.280762	70.182335	0.613717
2023-11-18	0.377694	244.950760	36585.703125	1963.285034	69.893806	0.611189
2023-11-19	0.384525	246.587784	37386.546875	2013.204468	70.640312	0.627499

2149 rows × 6 columns

Fig. 4 Adjusted crypto closed.

	ADA-USD	BNB-USD	BTC-USD	ETH-USD	LTC-USD	XRP-USD
Date						
2018-01-01	NaN	NaN	NaN	NaN	NaN	NaN
2018-01-02	0.074013	0.050289	0.097011	0.144702	0.116363	0.037586
2018-01-03	0.379604	0.078992	0.014611	0.088503	-0.040347	0.251711
2018-01-04	0.031917	-0.033756	0.026196	0.018907	-0.016294	0.029388
2018-01-05	-0.102826	0.618973	0.117333	0.017125	0.032734	-0.046274
...	...	...	...	...	...	...
2023-11-15	0.059860	0.047435	0.065928	0.041108	0.048773	0.031723
2023-11-16	-0.019842	-0.043335	-0.045559	-0.048304	-0.039164	-0.057693
2023-11-17	-0.013872	0.008183	0.012223	0.000204	-0.014003	0.002530
2023-11-18	0.030237	0.000839	-0.000300	0.001022	-0.004111	-0.004119
2023-11-19	0.018086	0.006683	0.021890	0.025426	0.010681	0.026686

2149 rows × 6 columns

Fig. 5 Percentage values of cryptocurrencies.

	ADA-USD	BNB-USD	BTC-USD	ETH-USD	LTC-USD	XRP-USD
ADA-USD	1.000000	0.560602	0.687810	0.750432	0.709487	0.644637
BNB-USD	0.560602	1.000000	0.640732	0.642729	0.631388	0.494809
BTC-USD	0.687810	0.640732	1.000000	0.819461	0.786640	0.588521
ETH-USD	0.750432	0.642729	0.819461	1.000000	0.809947	0.642601
LTC-USD	0.709487	0.631388	0.786640	0.809947	1.000000	0.633703
XRP-USD	0.644637	0.494809	0.588521	0.642601	0.633703	1.000000

Fig. 6 Correlation between cryptocurrencies.

It can be observed that on the main diagonal there are all one, this is because in the main diagonal the single cryptocurrency meets itself. It was shown that no negative values work, but to the contrary higher and lower values. To better visualize this data a heat map can be arranged (see Fig. 7).

It emerges that the values below the main diagonal are mirrored to those above the main diagonal. The lowest values are those between XRP-USD and BNB-USD with a value of 0.49, followed by BTC-USD and XRP-USD with a value of 0.59, while the highest values are between BTC-USD and ETH-USD with a value of 0.82, followed by ETH-USD and LTC-USD with a value of 0.81. This means that they are strongly correlated and move together.

In general, the positive correlation between Bitcoin's performance and the "microcosm" of Altcoins clearly arises. As it was emerged among those with a high historical correlation, Ethereum and Litecoin have often traveled (for structural time frames) often hand in hand with Bitcoin. Frequently, "uncorrelated" movements with respect to BTC concern events and news in the life of individual cryptocurrencies, which fortunately give trading opportunities in contrast to Bitcoin.

There are several reasons why investors might want to consider cryptocurrency correlation. First, when asset prices move in tandem, they can provide some measure of stability and predictability. Firstly, this can be especially useful in times of volatility, when markets experience a lot of ups and downs. Secondly, correlation can also help investors diversify their portfolios. By spreading their investments across several different assets, investors can reduce their exposure to a particular asset class. This diversification can help mitigate losses if an asset class experiences a sudden drop in value.

Finally, many investors believe that there is a positive correlation between the price of Bitcoin and the overall health of the cryptocurrency market. While this relationship is not always linear, there is evidence to indicate that when Bitcoin rises, the rest of the market tends to follow. For this reason, some investors are using Bitcoin as a barometer of the health of the cryptocurrency market. However, this does not mean that investing in cryptocurrencies does not come with risks. Correlation is just one of the factors being considered.

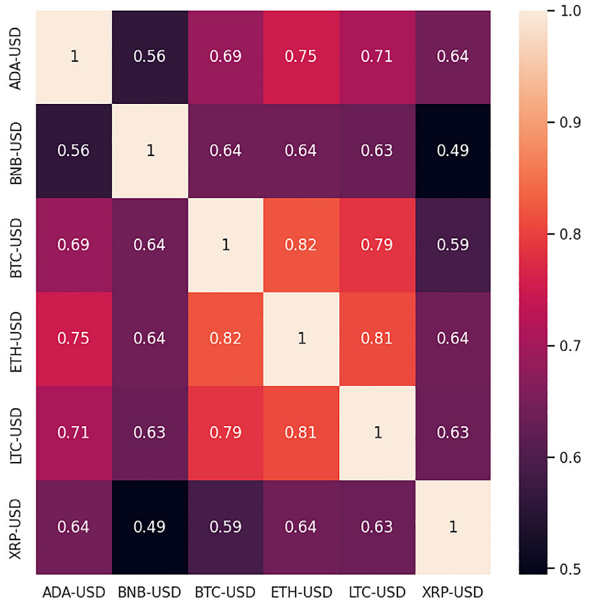


Fig. 7 Correlation matrix.

Conclusion

Due to its uncertain volatility, Bitcoin is a complex and diversified asset. Many algorithms are used to predict prices, but the parameters that affect Bitcoin differ. Therefore, machine learning technology is needed to accurately predict the value of Bitcoin. The studies focused on using a variety of modeling techniques to combat price instability and fluctuations. Therefore, researchers have used a variety of machine learning techniques to study a variety of factors that affect the price of Bitcoin, as well as its patterns based on price changes. In the world of digital asset investing, the issue of cryptocurrency correlation is much discussed, it can help investors maximize returns while also reducing risk.

Bitcoin's price and the global market for Altcoins such as Ethereum and Litecoin have a positive correlation, which are often traded alongside Bitcoin. This correlation can provide stability and predictability during volatile times and help investors diversify their portfolios by spreading investments across various assets. Many investors believe that there is a positive correlation between the price of Bitcoin and the complex health of the cryptocurrency market, and some use Bitcoin as a benchmark for the health of the cryptocurrency market. However, that doesn't mean that investing in cryptocurrencies is risk-free. Starting from this assumption, the forecasting analysis developed in this work makes it possible to help investors predict the future price and therefore to be able to review their portfolio but above all, thanks to the study conducted on the correlation between the various cryptocurrencies, to give the investor the opportunity to diversify his portfolio by investing in alternative cryptocurrencies, increasing returns.

References

Alessandretti, L., ElBahrawy, A., Aiello, L.M., Baronchelli, A., 2018. Anticipating cryptocurrency prices using machine learning. Complexity 2018, 1—16.

Androulaki, E., Karame, G.O., Roeschlin, M., Scherer, T., Capkun, S., 2013. Evaluating user privacy in Bitcoin. In: Financial Cryptography and Data Security: 17th International Conference, FC 2013, Okinawa, Japan, April 1—5, 2013. Springer Berlin Heidelberg, pp. 34—51. Revised Selected Papers 17.

Angel, J.J., McCabe, D., 2015. The ethics of payments: paper, plastic, or Bitcoin? J. Bus. Ethics 132, 603—611.

Arias-Oliva, M., Pelegrín-Borondo, J., Matías-Clavero, G., 2019. Variables influencing cryptocurrency use: a technology acceptance model in Spain. Front. Psychol. 10, 475.

Balcilar, M., Bouri, E., Gupta, R., Roubaud, D., 2017. Can volume predict Bitcoin returns and volatility? A quantiles-based approach. Econ. Modell. 64, 74—81.

Bariviera, A.F., Basgall, M.J., Hasperué, W., Naiouf, M., 2017. Some stylized facts of the Bitcoin market. Phys. Stat. Mech. Appl. 484, 82—90.

Bishop, C.M., Nasrabadi, N.M., 2006. Pattern Recognition and Machine Learning, vol. 4. Springer, New York, p. 738. No. 4.

Böhme, R., Christin, N., Edelman, B., Moore, T., 2015. Bitcoin: economics, technology, and governance. J. Econ. Perspect. 29 (2), 213—238.

Bouri, E., Gupta, R., Tiwari, A.K., Roubaud, D., 2017. Does Bitcoin hedge global uncertainty? Evidence from wavelet-based quantile-in-quantile regressions. Finance Res. Lett. 23, 87—95.

Box, G.E., Jenkins, G.M., Reinsel, G.C., Ljung, G.M., 2015. Time Series Analysis: Forecasting and Control. John Wiley & Sons.

Bradbury, D., 2013. The problem with Bitcoin. Comput. Fraud Secur. 2013 (11), 5—8.

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

Cheung, A., Roca, E., Su, J.J., 2015. Crypto-currency bubbles: an application of the Phillips—Shi—Yu (2013) methodology on Mt. Gox Bitcoin prices. Appl. Econ. 47 (23), 2348—2358.

Ciaian, P., Rajcaniova, M., Kancs, D.A., 2016. The economics of BitCoin price formation. Appl. Econ. 48 (19), 1799—1815.

Conte de Leon, D., Stalick, A.Q., Jillepalli, A.A., Haney, M.A., Sheldon, F.T., 2017. Blockchain: properties and misconceptions. Asia Pac. J. Innovat. Entrepreneursh. 11 (3), 286—300.

Corbet, S., Lucey, B., Urquhart, A., Yarovaya, L., 2019. Cryptocurrencies as a financial asset: a systematic analysis. Int. Rev. Financ. Anal. 62, 182—199.

Cortes, C., Vapnik, V., 1995. Support-vector networks. Mach. Learn. 20, 273—297.

Dyhrberg, A.H., 2016. Hedging capabilities of Bitcoin. Is it the virtual gold? Finance Res. Lett. 16, 139—144.

Freund, Y., Schapire, R.E., 1997. A generalization of the decision theory of online learning is an application to boo. J. Calculate. Syst. Sci. 55, 119—139.

Fry, J., Cheah, E.T., 2016. Negative bubbles and shocks in cryptocurrency markets. Int. Rev. Financ. Anal. 47, 343—352.

Georgoula, I., Pournarakis, D., Bilanakos, C., Sotiropoulos, D., Giaglis, G.M., 2015. Using Time-Series and Sentiment Analysis to Detect the Determinants of Bitcoin Prices. Available at SSRN 2607167.

Giudici, G., Milne, A., Vinogradov, D., 2020. Cryptocurrencies: market analysis and perspectives. J. Ind. Business Econ. 47, 1—18.

Hayes, A.S., 2017. Cryptocurrency value formation: an empirical study leading to a cost of production model for valuing Bitcoin. Telematics Inf. 34 (7), 1308—1321.

Ho, T.K., August 1995. Random decision forests. In: Proceedings of 3rd International Conference on Document Analysis and Recognition, vol. 1. IEEE, pp. 278—282.

Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. Neural Comput. 9 (8), 1735—1780.

Hong, K., 2017. Bitcoin as an alternative investment vehicle. Inf. Technol. Manag. 18, 265—275.

Huisu, J., Lee, J., Ko, H., Lee, W., August 2018. Predicting Bitcoin prices by using rolling window LSTM model. In: Proceedings of the KDD Data Science in Fintech Workshop, pp. 19—23. London, UK.

Jang, H., Lee, J., 2017. An empirical study on modeling and prediction of Bitcoin prices with Bayesian neural networks based on blockchain information. IEEE Access 6, 5427—5437.

Kim, T., 2017. On the transaction cost of Bitcoin. Finance Res. Lett. 23, 300—305.

Kim, Y.B., Kim, J.G., Kim, W., Im, J.H., Kim, T.H., Kang, S.J., Kim, C.H., 2016. Predicting fluctuations in cryptocurrency transactions based on user comments and replies. PLoS One 11 (8), e0161197.

Kristoufek, L., 2013. BitCoin meets Google trends and Wikipedia: quantifying the relationship between phenomena of the Internet era. Sci. Rep. 3 (1), 3415.

Li, T.R., Chamrajnagar, A.S., Fong, X.R., Rizik, N.R., Fu, F., 2019. Sentiment-based prediction of alternative cryptocurrency price fluctuations using gradient boosting tree model. Front. Phys. 7, 98.

Luther, W.J., 2016. Cryptocurrencies, network effects, and switching costs. Contemp. Econ. Pol. 34 (3), 553—571.

Madan, I., Saluja, S., Zhao, A., 2015. Automated Bitcoin Trading Via Machine Learning Algorithms. URL. http://cs229.stanford.edu/proj2014/Isaac%20.

Mai, F., Bai, Q., Shan, J., Wang, X.S., Chiang, R.H., 2015. The Impacts of Social Media on Bitcoin Performance. Thirty Sixth International Conference on Information Systems.

McNally, S., Roche, J., Caton, S., March 2018. Predicting the price of Bitcoin using machine learning. In: 2018 26th Euromicro International Conference on Parallel, Distributed and Network-Based Processing (PDP). IEEE, pp. 339—343.

Nakano, M., Takahashi, A., Takahashi, S., 2018. Bitcoin technical trading with artificial neural network. Phys. Stat. Mech. Appl. 510, 587—609.

Patel, J., Shah, S., Thakkar, P., Kotecha, K., 2015. Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques. Expert Syst. Appl. 42 (1), 259—268.

Ranjan, S., Kayal, P., Saraf, M., 2023. Bitcoin price prediction: a machine learning sample dimension approach. Comput. Econ. 61 (4), 1617—1636.

Rathore, R.K., Mishra, D., Mehra, P.S., Pal, O., Hashim, A.S., Shapi'i, A., et al., 2022. Real-world model for Bitcoin price prediction. Inf. Process. Manag. 59 (4), 102968.

Rebane, J., Karlsson, I., Papapetrou, P., Denic, S., 2018. Seq2Seq RNNs and ARIMA Models for Cryptocurrency.

Saad, M., Choi, J., Nyang, D., Kim, J., Mohaisen, A., 2019. Toward characterizing blockchain-based cryptocurrencies for highly accurate predictions. IEEE Syst. J. 14 (1), 321—332.

Shah, D., Zhang, K., September 2014. Bayesian regression and Bitcoin. In: 2014 52nd Annual Allerton Conference on Communication, Control, and Computing (Allerton). IEEE, pp. 409—414.

Shintate, T., Pichl, L., 2019. Trend prediction classification for high frequency Bitcoin time series with deep learning. J. Risk Financ. Manag. 12 (1), 17.

Vranken, H., 2017. Sustainability of Bitcoin and blockchains. Curr. Opin. Environ. Sustain. 28, 1—9.

Wang, J., Yang, Y., Mao, J., Huang, Z., Huang, C., Xu, W., 2016. CNN-RNN: a unified framework for multi-label image classification. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2285—2294.