

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/339797568>

# Machine Learning for Bitcoin Pricing — A Structured Literature Review

Conference Paper · March 2020

DOI: 10.30844/wi\_2020\_b4-jaquart

CITATIONS

3

READS

977

3 authors, including:



Patrick Jaquart

Karlsruhe Institute of Technology

6 PUBLICATIONS 5 CITATIONS

[SEE PROFILE](#)



David Dann

Karlsruhe Institute of Technology

17 PUBLICATIONS 259 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Data Science Process Model (DASC-PM) [View project](#)



**This is the author's version of a work that was published in the following source**

*Patrick Jaquart, David Dann, and Carl Martin (2020): "Machine Learning for Bitcoin Pricing — A Structured Literature Review," in WI 2020 Proceedings.*

**Please note: Copyright is owned by the author and / or the publisher.  
Commercial use is not allowed.**



**Institute of Information Systems and Marketing (IISM)**  
Kaiserstraße 89-93  
76133 Karlsruhe – Germany  
<http://iism.kit.edu>



**Karlsruhe Service Research Institute (KSRI)**  
Kaiserstraße 89  
76133 Karlsruhe – Germany  
<http://ksri.kit.edu>



© 2020. This manuscript version is made available under the CC-BY-NC-ND 4.0 license <http://creativecommons.org/licenses/by-nc-nd/4.0/>

# Machine Learning for Bitcoin Pricing — A Structured Literature Review

Patrick Jaquart<sup>1</sup>, David Dann<sup>1</sup>, and Carl Martin<sup>1</sup>

<sup>1</sup> Karlsruhe Institute of Technology, Institute of Information Systems and Marketing (IISM),  
Karlsruhe, Germany  
{patrick.jaquart, david.dann}@kit.edu, carl.martin@student.kit.edu

**Abstract.** Bitcoin, as the most popular cryptocurrency, has received increasing attention from both investors and researchers over recent years. One emerging branch of the research on bitcoin focuses on empirical bitcoin pricing. Machine learning methods are well suited for predictive problems, and researchers frequently apply these methods to predict bitcoin prices and returns. In this study, we analyze the existing body of literature on empirical bitcoin pricing via machine learning and structure it according to four different concepts. We show that research on this topic is highly diverse and that the results of several studies can only be compared to a limited extent. We further derive guidelines for future publications in the field to ensure a sufficient level of transparency and reproducibility.

**Keywords:** bitcoin, cryptocurrency, asset pricing, machine learning, prediction

## 1 Introduction

Bitcoin is a digital peer-to-peer cash system introduced by Nakamoto in 2008 [1]. Its underlying technology blockchain is referred to as “trust machine” [2] due to its three central properties: secure transfer of information through a cryptographic protocol, a distributed database, and a decentralized consensus mechanism [3]. Benefiting from the stated innovative properties (cf. traditional currencies and payment systems), bitcoin quickly gained relevance in academic literature as well as for the global financial system [4]. As of November 2019, Bitcoin has a market capitalization of over 155 billion US dollars, which corresponds to more than 66 percent of the whole cryptocurrency market [5].

Researchers are investigating a variety of topics in connection with bitcoin markets, for instance asset type, asset pricing, hedging, and market efficiency. Within the research branch of bitcoin pricing, there are several different smaller streams of research. Some researchers [6–9] work on creating and validating theoretical economic models, while other researchers concentrate on empirical asset pricing. In this work, we analyze and structure the body of literature on empirical bitcoin pricing via machine learning. Thereby, we use the term “bitcoin pricing” for the forecasting of target values based on the bitcoin price (e.g., price, absolute price change, or return). For empirical

bitcoin pricing, return predictive features might consist of priced risk factors (which might be identified through reviewing theoretical economic models [6–9]) or other factors based on possible market inefficiencies. Empirical bitcoin pricing is of explicit economic relevance, as accurate prediction models enable the employment of profitable trading strategies. Advances in computing technology, open-source implementation tools, and a skyrocketed bitcoin price, boosted the scientific community’s interest in the employment of machine learning methods for bitcoin pricing. Searching google scholar for the terms “bitcoin” and “machine learning” yields over 8200 results.

Due to the novelty of the Bitcoin technology, the research on predictive features for the bitcoin price is still in its early stages, and findings of several researchers indicate that bitcoin might represent a new asset class [10–12]. Therefore, classical return predictive signals from other asset classes (e.g., stocks [13]) are only partly applicable to bitcoin pricing. Most machine learning approaches demonstrate the ability to flexibly incorporate a large number of features (e.g., [14–18]). Together with the availability of large amounts of multidimensional data, this flexibility might render machine learning methods suitable for bitcoin pricing. This flexibility is especially important, since the stream of research on bitcoin pricing is still young and there exists limited guidance in the scientific literature about the nature of the bitcoin price formation process. For the course of this work, we adopt the definition of machine learning from Gu et al. [19], who apply machine learning to predict excess returns of stocks. They use the term machine learning “to describe (i) a diverse collection of high-dimensional models for statistical prediction, combined with (ii) so-called ‘regularization’ methods for model selection and mitigation of overfitting, and (iii) efficient algorithms for searching among a vast number of potential model specification” [11, pp. 2f]. From the wide range of definitions, the chosen one stands out due to its broad scope, which allows us to consider a large variety of approaches (e.g., linear models). Since the spectrum of employed machine learning methods and models used is rather broad, analyzing and comparing the different approaches remains a challenging task.

Against this backdrop, we argue it is time to take a step back and evaluate the current status quo. With our literature review, we provide an overview of current research on bitcoin pricing via machine learning. In so, we identify common methods, types of analysis, and findings. To our best knowledge, there is no comprehensive overview examining the diverse branch of research in the context of machine learning for bitcoin pricing. Therefore, we seize the opportunity to take this step back, assess the current state of research in this field, and outline potentials for future research.

Doing so, our contribution is threefold. First, we provide researchers in this field an overview of already existing work, identify recurring patterns and remaining niches to be occupied. Second, we identify which methods appear promising for the bitcoin pricing problem based on the evaluated body of literature. Third, we develop reporting guidelines for future research to enhance transparency and accelerate scientific progress.

The remainder of this work is structured as follows. Section 2 introduces the employed methodology for our structured literature review and provides summary statistics. In Section 3, we analyze the existing body of literature. Section 4 discusses

prevalent shortcomings, theoretical and practical implications, and paves the way for future research. Eventually, Section 5 concludes this work.

## 2 Methodology

Our literature search follows the suggestions by Webster and Watson [20] and vom Brocke et al. [21]. We build our initial literature base by querying a broad set of interdisciplinary research databases (i.e., *ACM Digital Library*, *AIS eLibrary*, *Business Source Premier*, *emerald insight*, *IEEE*, *ProQuest*, *SAGE Journals*, *ScienceDirect/Scopus*, *Taylor & Francis Online*, *Web of Science*). We query those databases for matching our search term<sup>1</sup> in title, abstract, or keywords [21]. We adopt the machine learning definition by Gu et al. [19], which is rather broad as it also includes linear models (e.g., linear regressions). By April 2019, this yields an initial set of 101 publications for further review. Analyzing each publication's title and abstract, we exclude 76 papers, which do not explicitly match the scope of our literature review. This may be due to (i) papers, employing methods not matching the machine learning definition of Gu et al. [19], (ii) papers not focusing on the prediction of bitcoin price/return (e.g., volatility), (iii) papers not being available in English, or (iv) papers not employing a prediction task (e.g., not using a time lag between predictive variables and target). Subsequent forward and backward search with the remaining relevant papers yields additional eight articles resulting in a total of 33 papers for in-depth review. Table 1. documents the number of identified articles for each database.

**Table 1.** Machine learning on bitcoin pricing research corpus

<i>Data Base</i>	<i># Paper</i>
ACM Digital Library	3
AIS eLibrary	1
Business Source Premier	2
emerald insight	0
IEEE	10
ProQuest	0
SAGE Journals	0
ScienceDirect/Scopus	8
Taylor & Francis Online	0
Web of Science	1
Forward / Backward Search	8
Total	33

<sup>1</sup> “(Pric\* OR Return\*) AND (Bitcoin\* OR Cryptocurrenc\*) AND ((Machine Learning) OR (Attention Model\*) OR Bayes\* OR Boost\* OR (Classification Tree\*) OR (Regression\*) OR (Deep Learning) OR (Discriminant Analysis) OR (Encoder-Decoder\*) OR (Feedforward Net\*) OR (Genetic Algorithm\*) OR (K-Nearest Neighbo\*) OR (LSTM\*) OR (Neural Net\*) OR (Random Forest\*) OR (Support Vector\* ))” *Note:* \* represents one or more wildcard characters.

Next, we derive key concepts for paper categorization. Therefore, we initially screened a set of 10 papers consisting of the most recent peer-reviewed conference proceedings and journal papers, as we assumed that recent papers incorporate findings of previous works. Three researchers independently reviewed each of these papers and developed an initial set of concepts for classification. Throughout the entire paper screening process, we evaluated these initial concepts and adapted them as required. Subsequent discussion and synthesis of all identified concepts led to a final set of four distinct concepts:

- *Method* (i.e., *multilayer perceptrons*, *recurrent neural networks*, *regression-based models*, *support vector machines*, *tree-based models*)
- *Features* (i.e., *technical*, *blockchain-based*, *sentiment- and interest-based*, *asset-based*)
- *Prediction Interval* (i.e., *second*, *minute*, *hour*, *day*, *week*)
- *Prediction Type* (i.e., *classification*, *regression*)

The subsequent paper classification process to one or more of the identified concepts followed a similar process — the classification was initially conducted independently and discrepancies were discussed afterward. The categorization guidelines for each researcher allowed a non-exclusive categorization (i.e., each article can be assigned to multiple categories). Table 1 summarizes all reviewed papers and specifies the assigned concepts. All reviewed papers have been published within the last five years 2019 (5), 2018 (17), 2017 (2), 2016 (4), 2015 (4), and 2014 (1).

### 3 Machine Learning and Bitcoin in Research

Bitcoin, representing the most popular crypto asset [5], has received a considerable amount of research attention since its inception in 2009. We use four different concepts to analyze and structure the literature, namely predictive *features*, *type* of prediction problem, prediction *intervals*, and machine learning *methods*. These identified concepts are rather broad and are applicable to a multitude of prediction tasks. However, some of the concept characteristics (e.g., blockchain-based features) are specific to the bitcoin pricing problem. Since the models analyze different time horizons, have different parameter specifications, and are evaluated using different evaluation metrics, it remains infeasible to compare them *across* different papers. Yet, comparing different machine learning models *within* the same paper remains possible, since they, among other aspects, use the same data. However, even the comparison of models *within* the same paper remains only valid under the assumptions that (i) all models are equally optimally tuned, and (ii) the selected time window is representative of bitcoin's price formation process. Table 2 provides an overview of the analysis of the different papers and concepts.

#### 3.1 Machine Learning Methods

The analyzed body of literature leverages a multiplicity of different machine learning methods. We group the literature into five categories based on the introduced models.

We differentiate *multilayer perceptrons*, *recurrent neural networks*, *regression-based models*, *support vector machines*, and *tree-based models*.

*Multilayer perceptrons* represent a type of feedforward neural network and consist of one input layer, one or more hidden layers, and one output layer [14, 15]. In feedforward networks, information only flows into one direction. *Multilayer perceptrons* with a non-linear and differentiable activation function can approximate any non-linear function, rendering them universal approximators [22]. Eight of the reviewed papers use *multilayer perceptrons*.

*Recurrent neural networks* drop the requirement for acyclic graphs from *multilayer perceptrons*, allowing for arbitrary feedback connections of the network [23]. Hammer [16] shows that *recurrent neural networks* with a sufficient number of hidden nodes and non-linear activation function also satisfy the requirements of a universal approximator. Ten papers use *recurrent neural networks*, and they remain the best reported model in all papers that use benchmarked scenarios.

*Regression-based models* refer to models based on linear regressions (e.g., logistic regressions, lasso regressions, and vector autoregressions). 16 papers employ regression-based models, which often serve as a reference point for more sophisticated machine learning methods.

The underlying idea of *support vector machines* consists of minimizing generalization error through constructing a (set of) hyperplane(s) in a high-dimensional space [17, 24, 25]. Six of the reviewed publications employ *support vector machines*.

Last, seven papers employ *tree-based models*. In these models, the outcomes are cuboid regions with axis-aligned edges [26]. A frequently used implementation of the methodology is the random forest, which constitutes an ensemble of imperfectly correlated trees to reduce the variance of forecasts [18].

Eleven papers employ methods that are part of none of the five major categories (e.g., fuzzy-systems [27]).

### 3.2 Return-predictive Features

Literature on bitcoin pricing via machine learning uses a multiplicity of return predictive signals. While, for instance, technical features (e.g., historical returns) are used in the literature on pricing traditional financial assets [28, 29], blockchain-based features (e.g., mining difficulty) are specifically related to cryptocurrencies — in particular bitcoin. Unlike stocks, bonds or other financial assets, bitcoins exhibit no fundamental value in a typical sense as they do not promise future cash flows, are not backed by a central bank, and cannot be utilized physically. Due to these different characteristics of bitcoin, it is not possible to use the same feature categorizations as for other financial assets. Based on the reviewed literature, we categorize return predictive features into *technical*, *blockchain-based*, *sentiment-* and *interest-based*, and *asset-based*.

*Technical features* include past data of the bitcoin market, for instance, historical prices or trading volumes. *Technical* features are the most frequently used features in the reviewed literature (27 models).

*Blockchain-based features* refer to data from the Bitcoin blockchain, for instance, mining difficulty or the number of transactions per block. Nine papers use *blockchain-based* features.

*Sentiment- and interest-based features* relate to social media sentiment and internet search volume, for instance, twitter sentiment or google trends data. Ten papers employ this type of feature.

*Asset-based features* relate to prices and returns of commodities and financial assets other than bitcoin, for instance, oil or stock market prices. *Asset-based* features are used in nine papers.

Features not covered by one of the presented categories are categorized as *other* features. Among these, Demir et al. [30] use economic policy uncertainty, Aysan et al. [31] use geopolitical risks, Hotz-Behofsits et al. [32] use GPU prices from Amazon's bestseller lists. Phaladisailoed and Numnonda [33], as well as Mallqui and Fernandes [34], use timestamps. Demir et al. [30] and Aysan et al. [31] conclude that bitcoin may serve as a hedge against policy uncertainty and geopolitical risks, respectively.

### 3.3 Prediction Interval

The authors in the reviewed literature use different prediction intervals to price bitcoin. The term "prediction interval" thereby denotes the frequency at which a model makes new predictions. The prediction intervals in the reviewed literature range from five seconds up to one week. Based on the prediction intervals, we group the models into five categories — *second*, *minute*, *hour*, *day*, and *week*.

*Second* includes models with prediction intervals of less than a minute (3 papers), *minute* between a minute and less than an hour (5 papers), *hour* between one hour and less than a day (3 papers), *day* between one day and less than one week (26 papers), and *week* includes models with prediction intervals of one week or longer (1 paper). Smuts [35] tests multiple models with prediction intervals ranging from one hour to one week and finds that the model with the highest prediction accuracy for bitcoin prices has a prediction accuracy of one week. Madan et al. [36] directly compare prediction intervals of ten seconds and ten minutes and find a slightly higher prediction accuracy for the prediction interval of ten minutes.

### 3.4 Prediction Types

There are several options to set up the prediction problem for bitcoin pricing. First, we distinguish between prediction problems formulated as a *regression* or *classification* problem. Bitcoin prices and returns are numerical and continuous variables. Hence, it is possible to formulate a *regression* model, which tries to predict the exact values of these target variables. However, one can reduce the *regression* problem into a *classification* problem by creating classes based on the target variable. In this case, the prediction model attempts to predict class affiliations based. Second, we distinguish the literature based on whether *absolute* bitcoin price levels or *relative* price changes are predicted. Traditional financial literature on other financial assets (e.g., on stocks [13]) usually analyzes *relative* price changes.



The reviewed literature formulates the bitcoin pricing problem 14 times as a *classification* problem and 21 times as a *regression* problem. Some scholars create multiple models and set up the prediction problem as both a *classification* problem and a *regression* problem [34, 37]. For *classification* problems, nine of 14 cases formulate it as a binary *classification* problem, predicting the sign (i.e., positive or negative) of the bitcoin price change. In contrast, three papers split the bitcoin price change into three classes (i.e., positive, neutral, negative). Beyond that, Nakano et al. [38] create four target classes based on the price change quantiles, and Huang et al. [39] create 21 classes based on different bitcoin return intervals. All papers that use *classification* models create target classes based on *relative* price changes, while 17 of the 21 papers that use *regression* models predict *absolute* bitcoin price levels and only four of these papers predict *relative* price changes.

## 4 Discussion

Overall, the research on bitcoin pricing via machine learning is not at a mature state yet. This may be due to the novelty of the protocol itself [1], and that machine learning techniques require a substantial amount of data to learn relationships between features and target variables. An explicit limitation of the reviewed work is that none of the papers is published in a top-rated finance or information systems journal [40]. Furthermore, a considerable amount of available literature barely meets academic standards in terms of transparent documentation of applied method and results. This includes, for instance, studies reporting unlikely  $R^2$  values for four different methods within the range of .991 and .992 [33]. An  $R^2$  of this magnitude is fairly unusual compared to the rest of the reviewed literature and might indicate setup problems (e.g., the use of unlagged features or a high similarity between features and target). In so, further shortcomings in the documentation render it impossible to reproduce and verify the empirical analyses at all. These include, not explicitly reporting the analyzed time range [41, 42], data split [43], or machine learning setup (e.g., layer structure, activation function, loss function, learning function) [44–46]. Furthermore, inconsistencies in the reporting prohibit reproducing the empirical tests. These inconsistencies can stem from reporting to optimize the number of units in a hidden layer of a multilayer perceptron within a specific range and using a number outside of that range in the final model [47] or setting up a regression problem, but using the accuracy metric for model evaluation without further explanation [48].

Throughout the literature, the machine learning models are built and evaluated on rather short time periods and small data samples. A choice of longer prediction intervals, (e.g. weekly intervals [35]) in combination with advanced machine learning models and a large number of features might result in an insufficient number of data points in the sample [49]. Furthermore, test splits of three percent or less, corresponding to 60 observations or less, limit the generalizability of the reported results [41, 46].

#### 4.1 Theoretical Implications

Researchers apply a wide variety of methods and underlying architectures with alternating success, such as *artificial neural networks*, *recurrent neural networks*, *regressions-based* models, *tree-based* methods, and *support vector machines*. Their main objectives are accurately predicting the bitcoin price (absolute or relative) using *classification* and *regression* approaches. The models embody a broad spectrum of features, which relate to *technical*, *blockchain-based*, *sentiment-* and *interest-based*, and *assets-based* aspects. Most researchers use *technical* features for their models. Only few authors [50–52] use features from all four categories. Since 2017, scholars begin to consider features beyond these main categories (e.g., economic policy uncertainty [30]).

Researchers formulate *regression* and *classification* problems equally often until the end of 2017, while from 2018 onwards there is a slight shift towards a higher share of *regression* problems. Consequently, researchers in the field mostly (i.e., 60%) utilize *regression-based* methods in total.

The majority (i.e., 79%) of models are set up with *daily* prediction intervals. The relative share of these *daily* models further increased after 2017. However, varying time horizons and model specifications limited the comparability of methods *across* different papers. Importantly, this resonates with limited options to validate any trading strategies applied. To ensure a certain level of comparability (e.g., uniform time horizons), we focus on comparisons of different methods *within* the same paper. Nevertheless, as they are based on several assumptions (e.g., representative time windows and equally optimal tuning states of different models), these comparisons are limited. None of the authors have published their machine learning model, which would allow future researchers to train the model on new data and compare the performance to other methods. Additionally, there are no widely established guidelines or best practices in this research stream for reporting machine learning models.

Given these limitations, we find that recurrent neural networks, and in particular long-short term memory neural networks, perform well in the bitcoin pricing problem compared to other methods [33, 34, 43–46, 53]. Interestingly, even though long short-term memory neural networks were published in 1997 already [54], the first paper [53] taking these into account is from 2018.

#### 4.2 Practical Implications

Based on the finding that complex network architectures such as recurrent neural networks yield promising results [33, 34, 43–46, 53], future research should evaluate further sophisticated network architectures for this particular problem. This may include assessing the effectiveness of ordinary convolutional neural networks [55], as well as dilated convolutional neural networks [56]. The latter has proven to provide promising results in forecasting S&P 500 stock market index already [57]. However, more sophisticated models require more data [49], which might be achieved, for instance, by considering shorter prediction intervals.

Beyond identifying appropriate modeling architecture, the process of model reporting demands for refinement and harmonization. Contrasting research from (bio)medical research [58] or psychology [59], the analyzed research follows no established guidelines for uniformly reporting machine learning results. We recommend the following reporting standards for future research in the field of bitcoin pricing via machine learning and machine learning projects in general. First, we propose that researchers are required to reveal and document the entire model configuration (i.e., hyperparameters) in a structured manner. This may include a distinct table providing information about the number of a multilayer perceptron's hidden layers, number of units per layer, activation/loss functions, or optimizers. Second, we propose that researchers publish all reported models and data to enhance the comparability amongst them. Thereby, future scholars may fall back on previously validated modeling approaches. Since all major machine learning frameworks (e.g., TensorFlow, Keras, PyTorch) provide distinct functions to save and export trained models, we argue that publishing model and data to an open research repository (e.g., CORE [60], Open Research Library ANU [61]) is a reasonable and necessary step to ensure a sufficient level of transparency. Third, we propose that researchers who publish new modeling approaches, benchmark their models against other existing and published models from the field on the same dataset. Currently, there is no established benchmarking dataset. However, researchers commonly use benchmarking datasets (e.g., MNIST for handwritten digits) in other machine learning fields. Overall, the guidelines were developed due to shortcomings in the existing bitcoin pricing literature and are therefore of particular importance in this specific field. However, they are applicable to empirical machine learning studies of various domains.

### **4.3 Limitations**

There are three main limitations of the presented analysis. First, machine learning and bitcoin pricing are two fast-evolving research disciplines. Therefore, our work reflects a quick blink in time of the literature in this field, and future analysis may yield different results. Moreover, the scope of our literature search is limited, as there exists no unique and widespread acceptance of the term “machine learning.” Additionally, this review suffers from the low quality (insufficient documentation and data samples) from part of the bitcoin pricing literature. Furthermore, we may speculate about the existence of more accurate machine learning models, which are exploited monetarily rather than contributed to the scientific body of literature.

### **4.4 Future Research**

We encourage future researchers in the field to evaluate advanced machine learning models (e.g., dilated convolutional neural networks [56]) for time series forecasting, which are not considered by contemporary research in this field. Theoretical economic models for bitcoin prices [6–9] might help to guide the search for further predictive features. To enable and accelerate scientific progress in the field, we propose that future

researchers report all model configurations in a structured way, report and publish their model and data, and benchmark new models against other reported models.

## 5 Conclusion

Bitcoin has received a considerable amount of interest from researchers and investors since its inception in 2008. The research on bitcoin pricing via machine learning constitutes a relevant and emerging topic. We review the existing body of literature of this research branch based on the guidelines of Webster and Watson [20] and vom Brocke et al. [21]. We structure and analyze the body of literature according to four different concepts, namely *method*, *feature*, *prediction interval*, and *prediction type*. A comparison of methods *within* the same paper indicates that *recurrent neural networks* might be well suited for the prediction problem. Most researchers use features from four categories, namely *technical*, *blockchain-based*, *sentiment- and interest-based*, and *asset-based*. Across the reviewed literature, we find a lack of transparency and comparability, limiting options to validate and reproduce model results and eventually applied trading strategies.

Based on these issues we propose that future researchers (i) reveal all relevant model configurations in a structured way, (ii) publish and upload their model and data to an open research repository, and (iii) benchmark their model against other published models.

## Appendix

**Table 2.** Literature overview. Best method marked with bold cross (based on accuracy or lowest error). For papers using classification and regression: <sup>A</sup>: best method for the classification problem, <sup>B</sup>: best method for the regression problem. For papers in which an ensemble consisting of multiple methods achieves the best results: <sup>C</sup>: methods applied.

Source	Method						Features					Interval	Type	
	Multilayer Perceptrons	Recurrent Neural Networks	Regression-based	Support Vector	Tree-based	Other	Technical	Blockchain-based	Sentiment-/interest-based	Asset-based	Other	S: Secondly, M: Minutely, H: Hourly, D: Daily, W: Weekly,	Classification	Regression
[47]	×						×					D	×	
[62]			×		×	×	×					S	×	
[41]	×					×	×					D		×
[31]			×								×	D		×
[63]		×					×		×			M, D		×

[51]			x				x	x	x	x		D		x
[30]			x				x				x	D		x
[52]			x				x	x	x	x		D		x
[64]			x				x			x		D		x
[37]	x <sup>A</sup>		x <sup>B</sup>	x			x	x				H	x	x
[65]			x		x	x	x					M	x	
[32]						x			x	x	x	D		x
[39]					x		x					D	x	
[66]			x						x			H		x
[67]	x		x	x			x	x		x		D		x
[46]		x	x				x					D		x
[44]		x				x	x					D		x
[68]						x	x		x			D	x	
[45]	x	x					x					D		x
[36]				x	x	x	x	x				S, M, D	x	
[34]	x	x <sup>A</sup>		x <sup>B</sup>	x <sup>A</sup>	x	x	x		x	x	D	x	x
[53]		x	x				x	x				D		x
[38]	x						x					M	x	
[48]		x					x		x			D		x
[33]		x	x				x				x	D		x
[50]						x	x	x	x	x		D		x
[42]			x	x	x	x			x			D	x	
[69]			x				x					S		x
[70]	x						x	x				D	x	
[35]		x					x		x			H, D, W	x	
[71]				x	x					x		D	x	
[72]						x				x		M	x	
[43]		x <sup>C</sup>	x <sup>C</sup>				x					D		x
Σ	8	10	16	6	7	11	27	9	10	9	5	S:3, M:5, H:3, D:26, W:1	14	21

## References

1. Nakamoto, S.: Bitcoin: A peer-to-peer electronic cash system, <https://bitcoin.org/bitcoin.pdf>.
2. The Economist: The trust machine, <https://www.economist.com/leaders/2015/10/31/the-trust-machine>.
3. Beck, R., Avital, M., Rossi, M., Thatcher, J.B.: Blockchain Technology in Business and Information Systems Research. Bus. Inf. Syst. Eng. 59, 381–384 (2017). <https://doi.org/10.1007/s12599-017-0505-1>.

4. Böhme, R., Christin, N., Edelman, B., Moore, T.: Bitcoin: Economics, technology, and governance. *J. Econ. Perspect.* 29, 213–238 (2015).
5. Coinmarketcap, <https://coinmarketcap.com/>.
6. Bolt, W., Van Oordt, M.R.C.: On the value of virtual currencies. *J. Money, Credit Bank.* (2019).
7. Pagnotta, E., Buraschi, A.: An equilibrium valuation of bitcoin and decentralized network assets. *Work. Pap.* (2018).
8. Biais, B., Bisiere, C., Bouvard, M., Casamatta, C., Menkveld, A.J.: Equilibrium bitcoin pricing. *Work. Pap.* (2018).
9. Schilling, L., Uhlig, H.: Some simple bitcoin economics. *J. Monet. Econ.* (2019).
10. Glaser, F., Zimmermann, K., Haferkorn, M., Weber, M.C., Siering, M.: Bitcoin-asset or currency? revealing users' hidden intentions. In: *ECIS 2014 Proceedings*. pp. 1–15 (2014).
11. Dyhrberg, A.H.: Bitcoin, gold and the dollar – A GARCH volatility analysis. *Financ. Res. Lett.* 16, 85–92 (2016).
12. Burniske, C., White, A.: Bitcoin: Ringing the bell for a new asset class, [https://research.ark-invest.com/hubfs/1\\_Download\\_Files\\_ARK-Invest/White\\_Papers/Bitcoin-Ringing-The-Bell-For-A-New-Asset-Class.pdf](https://research.ark-invest.com/hubfs/1_Download_Files_ARK-Invest/White_Papers/Bitcoin-Ringing-The-Bell-For-A-New-Asset-Class.pdf).
13. Green, J., Hand, J.R.M., Zhang, X.F.: The superview of return predictive signals. *Rev. Account. Stud.* 18, 692–730 (2013).
14. Cybenko, G.: Approximation by superpositions of a sigmoidal function. *Math. Control. signals Syst.* 2, 303–314 (1989).
15. Hornik, K., Stinchcombe, M., White, H.: Multilayer feedforward networks are universal approximators. *Neural networks.* 2, 359–366 (1989).
16. Hammer, B.: On the approximation capability of recurrent neural networks. *Neurocomputing.* 31, 107–123 (2000).
17. Cortes, C., Vapnik, V.: Support-vector networks. *Mach. Learn.* 20, 273–297 (1995).
18. Breiman, L.: Random forests. *Mach. Learn.* 45, 5–32 (2001).
19. Gu, S., Kelly, B., Xiu, D.: Empirical asset pricing via machine learning. *Work. Pap.* (2018).
20. Webster, J., Watson, R.T.: Analyzing the past to prepare for the future: Writing a literature review. *MIS Q.* xiii--xxiii (2002).
21. Vom Brocke, J., Simons, A., Niehaves, B., Riemer, K., Plattfaut, R., Cleven, A., others: Reconstructing the giant: on the importance of rigour in documenting the literature search process. In: *ECIS 2009 Proceedings*. pp. 2206–2217 (2009).
22. Hornik, K.: Approximation capabilities of multilayer feedforward networks. *Neural networks.* 4, 251–257 (1991).
23. Rumelhart, D.E., Hinton, G.E., Williams, R.J., others: Learning representations by back-propagating errors. *Nature.* 323, 533–536 (1988).
24. Vapnik, V.: The nature of statistical learning theory. Springer science & business media (1995).
25. Drucker, H., Burges, C.J.C., Kaufman, L., Smola, A.J., Vapnik, V.: Support

- vector regression machines. In: *Advances in neural information processing systems*. pp. 155–161 (1997).
26. Breiman, L., Friedman, J.H., Olshen, R., Stone, C.J.: *Classification and Regression Trees*. Routledge (1984).
  27. Atsalakis, G.S., Valavanis, K.P.: Forecasting stock market short-term trends using a neuro-fuzzy based methodology. *Expert Syst. Appl.* 36, 10696–10707 (2009).
  28. Jegadeesh, N.: Evidence of predictable behavior of security returns. *J. Finance*. 45, 881–898 (1990).
  29. Jegadeesh, N., Titman, S.: Returns to buying winners and selling losers: Implications for stock market efficiency. *J. Finance*. 48, 65–91 (1993).
  30. Demir, E., Gozgor, G., Lau, C.K.M., Vigne, S.A.: Does economic policy uncertainty predict the Bitcoin returns? An empirical investigation. *Financ. Res. Lett.* 26, 145–149 (2018).
  31. Aysan, A.F., Demir, E., Gozgor, G., Lau, C.K.M.: Effects of the geopolitical risks on Bitcoin returns and volatility. *Res. Int. Bus. Financ.* 47, 511–518 (2019).
  32. Hotz-Behofsits, C., Huber, F., Zörner, T.O.: Predicting crypto-currencies using sparse non-Gaussian state space models. *J. Forecast.* 37, 627–640 (2018).
  33. Phaladisailoed, T., Numnonda, T.: Machine learning models comparison for bitcoin price prediction. In: *Proceedings of 2018 International Conference on Information Technology and Electrical Engineering: Smart Technology for Better Society* (2018). <https://doi.org/10.1109/ICITEED.2018.8534911>.
  34. Mallqui, D.C.A., Fernandes, R.A.S.: Predicting the direction, maximum, minimum and closing prices of daily Bitcoin exchange rate using machine learning techniques. *Appl. Soft Comput. J.* 75, 596–606 (2019). <https://doi.org/10.1016/j.asoc.2018.11.038>.
  35. Smuts, N.: What Drives Cryptocurrency Prices?: An Investigation of Google Trends and Telegram Sentiment. *ACM SIGMETRICS Perform. Eval. Rev.* 46, 131–134 (2019).
  36. Madan, I., Saluja, S., Zhao, A.: Automated bitcoin trading via machine learning algorithms. *Work. Pap.* (2015).
  37. Greaves, A., Au, B.: Using the bitcoin transaction graph to predict the price of bitcoin. *Work. Pap.* (2015).
  38. Nakano, M., Takahashi, A., Takahashi, S.: Bitcoin technical trading with artificial neural network. *Physica A*. 510, 587–609 (2018). <https://doi.org/10.1016/j.physa.2018.07.017>.
  39. Huang, J.-Z., Huang, W., Ni, J.: Predicting Bitcoin Returns Using High-Dimensional Technical Indicators. *J. Financ. Data Sci.* (2018).
  40. VHB, <https://vhbonline.org/en/service/jourqual/vhb-jourqual-3/complete-list-of-the-journals/>.
  41. Atsalakis, G.S., Atsalaki, I.G., Pasiouras, F., Zopounidis, C.: Bitcoin price forecasting with neuro-fuzzy techniques. *Eur. J. Oper. Res.* 276, 770–780 (2019). <https://doi.org/10.1016/j.ejor.2019.01.040>.
  42. Rahman, S., Hemel, J.N., Junayed Ahmed Anta, S., Muhee, H. Al, Uddin, J.:

- Sentiment analysis using R: An approach to correlate cryptocurrency price fluctuations with change in user sentiment using machine learning. In: Proceedings of 2018 Joint International Conference on Informatics, Electronics and Vision and International Conference on Imaging, Vision and Pattern Recognition. pp. 492–497. IEEE (2019). <https://doi.org/10.1109/ICIEV.2018.8641075>.
43. Wu, C.H., Lu, C.C., Ma, Y.F., Lu, R.S.: A new forecasting framework for bitcoin price with LSTM. In: 2019 Proceedings of IEEE International Conference on Data Mining Workshops. pp. 168–175 (2019). <https://doi.org/10.1109/ICDMW.2018.00032>.
  44. Khaldi, R., El Afia, A., Chiheb, R., Faizi, R.: Forecasting of Bitcoin Daily Returns with EEMD-ELMAN based Model. In: Proceedings of 2018 International Conference on Learning and Optimization Algorithms: Theory and Applications. pp. 1–6 (2018). <https://doi.org/10.1145/3230905.3230948>.
  45. Lahmiri, S., Bekiros, S.: Cryptocurrency forecasting with deep learning chaotic neural networks. *Chaos, Solitons & Fractals*. 118, 35–40 (2019).
  46. Karakoyun, E.S., Cibikdiken, A.O.: Comparison of ARIMA Time Series Model and LSTM Deep Learning Algorithm for Bitcoin Price Forecasting. In: Proceedings of 2018 Multidisciplinary Academic Conference. pp. 171–180 (2018).
  47. Almeida, J., Tata, S., Moser, A., Smit, V.: Bitcoin prediction using ANN. *Neural networks*. 1–12 (2015).
  48. Pant, D.R., Neupane, P., Poudel, A., Pokhrel, A.K., Lama, B.K.: Recurrent Neural Network Based Bitcoin Price Prediction by Twitter Sentiment Analysis. In: Proceedings of 2018 IEEE International Conference on Computing, Communication and Security. pp. 128–132 (2018).
  49. Arnott, R., Harvey, C.R., Markowitz, H.: A backtesting protocol in the era of machine learning. *J. Financ. Data Sci.* 1, 64–74 (2019).
  50. Poyser, O.: Exploring the dynamics of Bitcoin’s price: a Bayesian structural time series approach. *Eurasian Econ. Rev.* 9, 29–60 (2019).
  51. Ciaian, P., Rajcaniova, M., Kancs, D.: The economics of BitCoin price formation. *Appl. Econ.* 48, 1799–1815 (2016).
  52. Georgoula, I., Pournarakis, D., Bilanakos, C., Sotiropoulos, D., Giaglis, G.M.: Using Time-Series and Sentiment Analysis to Detect the Determinants of Bitcoin Prices. *Work. Pap.* (2015). <https://doi.org/10.2139/ssrn.2607167>.
  53. McNally, S., Roche, J., Caton, S.: Predicting the Price of Bitcoin Using Machine Learning. In: Proceedings of 2018 Euromicro International Conference on Parallel, Distributed, and Network-Based Processing. pp. 339–343 (2018). <https://doi.org/10.1109/PDP2018.2018.00060>.
  54. Hochreiter, S., Schmidhuber, J.: Long short-term memory. *Neural Comput.* 9, 1735–1780 (1997).
  55. LeCun, Y., Bengio, Y., others: Convolutional networks for images, speech, and time series. *Handb. brain theory neural networks*. 3361, 1–14 (1995).
  56. Yu, F., Koltun, V.: Multi-Scale Context Aggregation by Dilated Convolutions. In: Proceedings of 2016 International Conference on Learning Representations.



- pp. 1–13 (2016).
57. Borovykh, A., Bohte, S., Oosterlee, C.W.: Conditional time series forecasting with convolutional neural networks. Work. Pap. (2017).
  58. Luo, W., Phung, D., Tran, T., Gupta, S., Rana, S., Karmakar, C., Shilton, A., Yearwood, J., Dimitrova, N., Ho, T.B., others: Guidelines for developing and reporting machine learning predictive models in biomedical research: a multidisciplinary view. *J. Med. Internet Res.* 18, e323 (2016).
  59. Wilkinson, L.: Statistical methods in psychology journals: Guidelines and explanations. *Am. Psychol.* 54, 594 (1999).
  60. CORE, <https://core.ac.uk/>.
  61. Open Research Library - Australian National University, <https://openresearch-repository.anu.edu.au/>.
  62. Amjad, M., Shah, D.: Trading bitcoin and online time series prediction. In: *NIPS 2016 Proceedings*. pp. 1–15 (2017).
  63. Cerda, G.C., Reutter, J., Maza, D. La: Bitcoin Price Prediction Through Opinion Mining. In: *Proceedings of 2019 World Wide Web Conference*. pp. 755–762 (2019).
  64. Giudici, P., Abu-Hashish, I.: What determines bitcoin exchange prices? A network VAR approach. *Financ. Res. Lett.* 28, 309–318 (2019).
  65. Hegazy, K., Mumford, S.: Comparative automated bitcoin trading strategies. Work. Pap. (2016).
  66. Jain, A., Tripathi, S., Dwivedi, H.D., Saxena, P.: Forecasting Price of Cryptocurrencies using Tweets Sentiment Analysis. In: *Proceedings of 2018 International Conference on Contemporary Computing*. pp. 2–4 (2018).
  67. Jang, H., Lee, J.: An Empirical Study on Modeling and Prediction of Bitcoin Prices with Bayesian Neural Networks Based on Blockchain Information. *IEEE Access*. 6, 5427–5437 (2017). <https://doi.org/10.1109/ACCESS.2017.2779181>.
  68. Kim, Y. Bin, Kim, J.G., Kim, W., Im, J.H., Kim, T.H., Kang, S.J., Kim, C.H.: Predicting fluctuations in cryptocurrency transactions based on user comments and replies. *PLoS One*. 11, 1–17 (2016).
  69. Shah, D., Zhang, K.: Bayesian regression and Bitcoin. In: *Proceedings of 2014 Annual Allerton Conference on Communication, Control, and Computing*. pp. 409–414 (2014). <https://doi.org/10.1109/ALLERTON.2014.7028484>.
  70. Sin, E., Wang, L.: Bitcoin price prediction using ensembles of neural networks. In: *Proceedings of 2018 International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery*. pp. 666–671 (2018). <https://doi.org/10.1109/FSKD.2017.8393351>.
  71. Sun, X., Liu, M., Sima, Z.: A novel cryptocurrency price trend forecasting model based on LightGBM. *Financ. Res. Lett.* (2018). <https://doi.org/10.1016/j.frl.2018.12.032>.
  72. Tupinambás, T.M., Leão, R.A., Lemos, A.P.: Cryptocurrencies transactions advisor using a genetic mamdani-type fuzzy rules based system. In: *Proceedings of 2018 IEEE International Conference on Fuzzy Systems*. pp. 1–7 (2018). <https://doi.org/10.1109/FUZZ-IEEE.2018.8491619>.