Finance Research Letters 40 (2021) 101755

Contents lists available at ScienceDirect

Finance Research Letters

journal homepage: www.elsevier.com/locate/frl

Forecasting the price of Bitcoin using deep learning

Mingxi Liu [a], Guowen Li [a,b], Jianping Li [a,c], Xiaoqian Zhu [a,*], Yinhong Yao [a,b]

[a] Institutes of Science and Development, Chinese Academy of Sciences, Beijing 100190, China
[b] School of Public Policy and Management, University of Chinese Academy of Sciences, Beijing 100049, China
[c] School of Economics and Management, University of Chinese Academy of Sciences, Beijing 100190, China

A R T I C L E I N F O

Keywords:
Bitcoin price prediction
Stacked denoising autoencoders
Feature learning
Deep extraction

A B S T R A C T

After constructing a feature system with 40 determinants that affect the price of Bitcoin considering aspects of the cryptocurrency market, public attention, and the macroeconomic environment, a deep learning method named stacked denoising autoencoders (SDAE) is utilized to predict the price of Bitcoin. The results show that compared with the most popular machine learning methods, such as back propagation neural network (BPNN) and support vector regression (SVR) methods, the SDAE model performs better in both directional and level prediction, measured using commonly used indicators, i.e., mean absolute percentage error (MAPE), root mean squared error (RMSE), and directional accuracy (DA).

1. Introduction

Given its innovative features, such as its decentralization and traceability, Bitcoin has attracted much attention from the media and investors (Fry and Cheah, 2016; Urquhart, 2016). Unlike sovereign currency, Bitcoin is a kind of decentralized digital currency that is not backed by any government's credit, and so the volatility of the price of Bitcoin is extremely high (Sun et al., 2020). Its price increased from zero value at the time of its inception in 2009 to around $13 per Bitcoin in January 2013 and subsequently shot up to around $20 thousand per Bitcoin in December 2017. In December 2019, although the price of Bitcoin experienced two years of overall large shocks that pushed it down, the total market capitalization is still more than $130 billion.

The accurate prediction of the Bitcoin price can not only provide decision support for investors but also provide a reference for governments to make regulatory policies (Matkovskyy and Jalan, 2019). In recent years, many scholars have attempted to find potential qualitative relationships between the price of Bitcoin and a number of factors, and they have mainly utilized econometrics methods such as the error correction model (ECM) (Li and Wang, 2017), the vector autoregressive model (VAR) (Demir et al., 2018), and the autoregressive integrated moving average model (ARIMA) (Bakar and Rosbi, 2017).

Very few studies focus on the quantitative prediction of the Bitcoin price. Greaves and Au (2015) construct a Bitcoin transaction network and predict the price of Bitcoin one hour later through some traditional machine learning methods, but they only considered a few Bitcoin market-related indicators. McNally et al. (2018) suggest that the prediction ability of the Bitcoin price using the Long Short-Term Memory Network is better than that using the ARIMA model. Indera et al. (2018) accurately predict the price of Bitcoin by constructing a Multi-Layer Perceptron-based Non-Linear Autoregressive with Exogeneous Input model using the opening, closing, minimum, and maximum past prices together with moving average technical indicators.

To predict the price of Bitcoin, after studying several potential determinants and constructing a feature system for Bitcoin's price, a

* Corresponding author.
E-mail address: zhuxq@casipm.ac.cn (X. Zhu).

https://doi.org/10.1016/j.frl.2020.101755
Received 19 April 2020; Received in revised form 31 August 2020; Accepted 4 September 2020
Available online 6 September 2020
1544-6123/

Table 1
The determinants of the price of Bitcoin.

No.	Determinant	Category	No.	Determinant	Category
1	Average hashrate	Cryptocurrency market	21	Dow Jones industrial average	Macroeconomic environment
2	Bitcoin market capitalization		22	New York Stock Exchange Gold futures	
3	Bitcoin trading volume		23	HuShen 300 index	
4	Litecoin market capitalization		24	Korea composite index	
5	Litecoin price		25	Singapore Straits Times index	
6	Litecoin trading volume		26	Standard & Poor's 500 index	
7	Mining award		27	Exchange rate of US dollar to Australian dollar	
8	Mining difficulty		28	Exchange rate of US dollar to British pound	
9	Number of addresses		29	Exchange rate of US dollar to Canadian dollar	
10	Volume of blocks		30	Exchange rate of US dollar to Danish Krone	
11	Ripple market capitalization		31	Exchange rate of US dollar to Japanese yen	
12	Ripple price		32	Exchange rate of US dollar to New Zealand dollar	
13	Ripple volume		33	Exchange rate of US dollar to Norwegian krona	
14	Commission charges		34	Exchange rate of US dollar to Swedish krona	
15	Bitcoin mining profitability		35	Exchange rate of US dollar to Euro	
16	Baidu index	Public attention	36	Exchange rate of US dollar to Swiss franc	
17	Google trends index		37	FTSE 100 index	
18	Brent crude futures	Macroeconomic environment	38	Tokyo's Nikkei index	
19	CBOE volatility index		39	US dollar index	
20	NYSE Copper futures		40	WTI oil future	

deep learning method named the Stacked Denoising Autoencoders (SDAE) is implemented to predict the price of Bitcoin one-day-later. To verify the effectiveness of the SDAE method, some benchmark methods are used for comparison, such as the Back Propagation Neural Network (BPNN) (Narendra and Parthasarathy, 1990) and the Support Vector Regression (SVR) (Drucker et al., 1996). These benchmark methods are widely recognized to be among the best performing machine learning methods and are very popular in the forecasting domain (Hong, 2011; Huang et al., 2010; Li et al., 2018; Lu and Wu, 2011; Sezer et al., 2020; Succurro et al., 2019).

The rest of this paper is organized as follows. Section 2 describes the indicator system built for predicting the price of Bitcoin. Section 3 presents the deep learning method used in this paper. Section 4 compares and discusses the experimental results. Concluding remarks are given in Section 5.

2. Determinants of the Price of Bitcoin

Previous studies have tried to find the determinants of the price of Bitcoin. Alstyne and Marshall (2014) find that the Dow Jones index, the euro-dollar exchange rate, and the oil price have significant impacts on the price of Bitcoin. Kristoufek (2015) verifies that there is a positive correlation relationship between search engine queries for the word "Bitcoin" (showing the degree of public interest in Bitcoin), the hashrate, the mining difficulty, and the price of Bitcoin. Ciaian et al. (2016) find that Bitcoin's attractiveness for investors has a significant impact on its price. Bouoiyour et al. (2016) discover that the anxieties over the outcomes of Brexit and the 2016 U.S. presidential election contributed positively to the price of Bitcoin. Similarly, Mai et al. (2018) verify that social sentiment is an important determinant.

Although increasingly more potential determinants of the Bitcoin price have been identified, previous studies have not given an overview of all the possible factors. This paper provides a comprehensive list of the determinants. They are divided into three categories: (1) the cryptocurrency market, such as Bitcoin's daily turnover, Bitcoin's total market capitalization, Bitcoin's hashrate, and the market data of other cryptocurrencies; (2) public attention, such as the search volume of Bitcoin-related keywords from different search engines; and (3) the macroeconomic environment, such as stock market indexes, the crude oil price, exchange rates, etc. A total of forty determinants are shown in Table 1.

3. Methodology

The Stacked Denoising Autoencoders (SDAE) is a popular deep learning model due to its prediction performance. It is built based on the stacking layers of Denoising Autoencoders (DAE) (Vincent et al., 2010). To illustrate the DAE, the Autoencoder (AE) should be introduced first.

An AE is a one hidden layer neural network where its input and output sizes are equal. It maps an input vector $\mathbf{x} \in [0, 1]^d$ to a hidden representation $\mathbf{y} \in [0, 1]^{d'}$ based on Eq. (1), and then it decodes the representation $\mathbf{y}$ back into a vector $\mathbf{z} \in [0, 1]^d$ in the input space based on Eq. (2), given below.

$$\mathbf{y} = f_\theta(\mathbf{x}) = s(\mathbf{W}\mathbf{x} + \mathbf{b}) \tag{1}$$

$$\mathbf{z} = g_{\theta'}(\mathbf{y}) = s(\mathbf{W}'\mathbf{y} + \mathbf{b}') \tag{2}$$

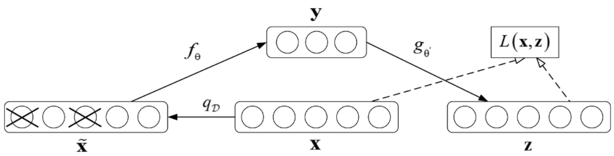

Fig. 1. The architecture of Denoising Autoencoders.

The deterministic mapping f_θ is called the encoder with the parameter set $\theta = \{\mathbf{W}, \mathbf{b}\}$, where $\mathbf{W}$ is a $d' \times d$ weight matrix and $\mathbf{b}$ is an offset vector with a dimension d'. The mapping $g_{\theta'}$ is called the decoder with parameters $\theta' = \{\mathbf{W}', \mathbf{b}'\}$. Each training $\mathbf{x}^{(i)}$ is mapped to a corresponding $\mathbf{y}^{(i)}$ and a reconstruction $\mathbf{z}^{(i)}$. By minimizing the average reconstruction error shown in Eq. (3), the optimal parameters θ^* and θ'^* can be derived.

$$\theta^*, \theta'^* = \underset{\theta,\theta'}{\mathrm{argmin}} \frac{1}{n} \sum_{i=1}^{n} L\left(\mathbf{x}^{(i)}, \mathbf{z}^{(i)}\right) = \underset{\theta,\theta'}{\mathrm{argmin}} \frac{1}{n} \sum_{i=1}^{n} L\left(\mathbf{x}^{(i)}, g_{\theta}\left(f_{\theta}\left(\mathbf{x}^{(i)}\right)\right)\right) \tag{3}$$

where L is a loss function that can be either the traditional squared error or reconstruction cross-entropy, and n is the number of training samples.

Training a common AE is unable to guarantee the extraction of useful features that constitute a high-level representation (Vincent et al., 2010; Zhao et al., 2017). The DAE changes the reconstruction criteria by denoising and reconstructs a clean input from a corrupted version. The reconstruction process of the DAE is as follows. The initial input $\mathbf{x}$ is corrupted into $x^{\sim}$ by a stochastic mapping $x^{\sim} \sim q_{\mathcal{D}}(x^{\sim}|\mathbf{x})$. Corrupted input $x^{\sim}$ is then mapped, as with the basic autoencoder, to a hidden representation $\mathbf{y} = f_\theta(x^{\sim}) = s(\mathbf{W}x^{\sim} + \mathbf{b})$ from which $\mathbf{z} = g_{\theta'}(\mathbf{y})$ is constructed. Fig. 1 shows a schematic representation of the procedure. With the training set, the parameters θ and θ' can be optimized by minimizing the average reconstruction error between the uncorrupted input $\mathbf{x}$ and $\mathbf{z}$.

Through stacking several DAEs, the deep neural network SDAE can be constructed (Hinton and Salakhutdinov, 2006). The whole training process of the SDAE network is composed of two steps, which are the step-by-step unsupervised pre-training step and the supervised fine-tuning step (Hinton and Salakhutdinov, 2006; Zhao et al., 2017).

To comprehensively evaluate the prediction accuracy of the models from different aspects (i.e., the directional prediction and level prediction), three widely-used indicators are adopted in this paper, which are the Mean Absolute Percentage Error (MAPE), the Root Mean Squared Error (RMSE), and the Directional Accuracy (DA) (Yu and Xu, 2014; Zhao et al., 2017). These indicators are defined as Eqs. (4) to (6), respectively.

$$MAPE = \frac{1}{m} \sum_{t=1}^{m} \left| \frac{y(t) - \hat{y}(t)}{y(t)} \right| \tag{4}$$

$$RMSE = \sqrt{\frac{1}{m} \sum_{t=1}^{m} \left(y(t) - \hat{y}(t)\right)^2} \tag{5}$$

$$DA = \frac{1}{m} \sum_{t=1}^{m} a(t) \times 100\% \tag{6}$$

where m is the number of samples in the test set; $y(t)$ and $\hat{y}(t)$ are the actual and predicted values at the time t, respectively; and $a(t) = 1$ if $(y(t+1) - y(t))(\hat{y}(t+1) - y(t)) \geq 0$ and $a(t) = 0$ otherwise. The smaller the MAPE and RMSE and the larger the DA are, the more accurate the predicted results.

4. Experiment results

4.1. Data description and processing

The price of Bitcoin and the cryptocurrency market determinant data are collected from professional Bitcoin trading websites and platforms such as 'www.coindesk.com' and 'BTC.com'. The search volume per day about Bitcoin comes from two world-famous search engines, i.e., Baidu and Google, and is used to quantify the degree of public attention. The macroeconomic environment data are obtained from the Wind financial database and the Choice financial database, two leading financial data providers in China (Zhu et al., 2020); and the famous Thomson Datastream database (Li et al., 2020c). The time period of these data spans from July 2013 to December 2019 because the data of some determinants such as the search data of Baidu are not available before July 2013. After deleting the missing data, 1356 days of data for all 40 determinants and the price of Bitcoin are obtained. Finally, these data are normalized to [0, 1] to eliminate the impact of the magnitude of data.

Table 2
The Bitcoin price prediction accuracies of different models.

Competing models	MAPE	RMSE	DA
BPNN	0.3736	390.07	0.5457
PCA-SVR	0.2278	318.62	0.5728
SVR	0.1783	248.24	0.5750
SDAE	**0.1019**	**160.63**	**0.5985**

Note: 1. The results are the average of 30 runs. 2. The best result for each indicator is in bold.

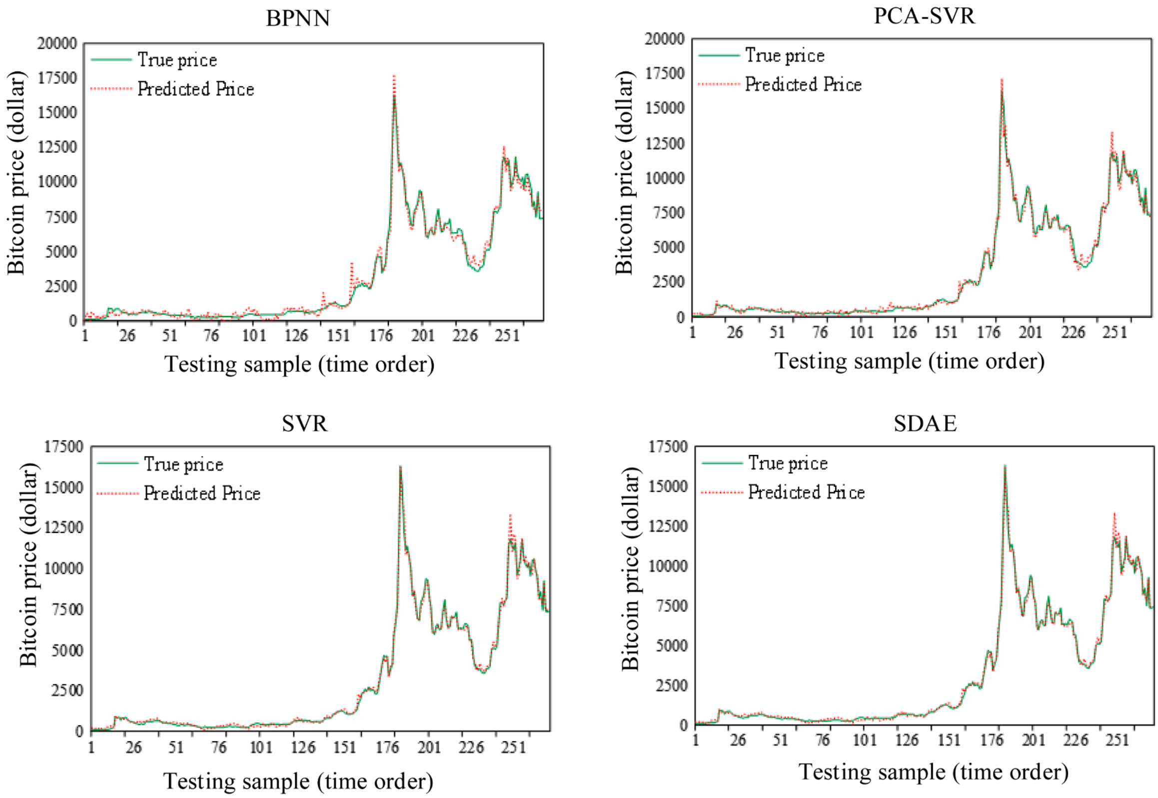

Fig. 2. Comparison of the true price of Bitcoin and predicted price based on different models.

4.2. *Hyper-parameter optimization*

The parameters for the SDAE and the comparison models are determined by fully referring to the existing studies. For the SDAE, the network structure is set as [40, 40, 20, 1]. The activation functions of the encoder and decoder are sigmoid and linear functions respectively. The learning rate is set as 0.1, the batch size is 50, and the number of the epoch is 20 (Vincent et al., 2010; Zhao et al., 2017). For the BPNN, the number of hidden layers is 1 with the network structure of [40, 20, 1], the activation function is sigmoid function, and the learning rate is 0.1 (Godarzi et al., 2014). The loss function for SDAE and BPNN is mean square error. For the SVR and principal component analysis-based support vector regression (PCA-SVR), the most widely-used Radial Basis Function kernel (also called a Gaussian kernel) is employed. The parameter spaces for the penalty parameter and kernel parameter are (0.01, 0.1, 1, 2, 4, 8) and (0.125, 0.25, 0.5, 1, 2, 4), respectively (Li et al., 2020a). Grid search and 5-fold validation are used to optimize all the parameters (Zhu et al. 2013). Python 3.7 is used to perform all the computations in this paper.

4.3. *Prediction power comparison*

In line with most of the studies, the dataset is randomly divided into training data (80%) and testing data (20%) (Dong et al., 2018; Li et al., 2020a; Liu et al., 2020; Zhao et al., 2017). To ensure the stability of prediction results, generally the training and testing process should be repeated many times. The number of repetitions is set as 30 in this paper because it is the minimum number that can support the statistical test in the robustness analysis (Li et al., 2020a; Wang et al., 2020). The BPNN, SVR, and PCA-SVR are chosen as the comparison models because they are among the most popular traditional machine learning models that are widely recognized for

Table 3
Results of the *t*-tests for different models under three evaluation criteria.

Models	MAPE BPNN	PCA-SVR	SVR	SDAE	RMSE BPNN	PCA-SVR	SVR	SDAE	DA BPNN	PCA-SVR	SVR	SDAE
BPNN		13.38***	18.83***	27.01***		6.88***	15.56***	25.08***		-4.14***	-4.28***	-5.91***
PCA-SVR			10.06***	29.83***			7.57***	16.94***			-0.37	-2.88**
SVR				29.12***				11.12***				-2.50*
SDAE												

Note: 1. $p < 0.05$(*), $p < 0.01$(**), and $p < 0.001$(***). 2. Positive t-statistics indicate that the values in the row are smaller than the values in the column, and vice versa.

their excellent prediction performance (Hong, 2011; Huang et al., 2010; Li et al., 2018; Lu and Wu, 2011; Sezer et al., 2020; Succurro et al., 2019). Then, the SDAE and the comparison models are used to predict the price of Bitcoin. The prediction accuracies of these models are shown in Table 2.

Table 2 shows that the SDAE achieves the lowest MAPE and RMSE of 0.1019 and 160.63, respectively, and the highest DA of 0.5985. This means that it performs the best in forecasting the price of Bitcoin in both the directional prediction and level prediction. The SVR performs the second best and the BPNN performs the worst. To intuitively show their prediction performances, the true price of Bitcoin and the predicted prices derived from these four models are drawn in Fig. 2. It can be seen that generally all of these models perform relatively well. For the SDAE and SVR, the lines denoting the true price of Bitcoin and the predicted price highly coincide with each other, even in the area where the price of Bitcoin fluctuates violently. For the PCA-SVR, the prediction accuracy is acceptable in the stable area but is relatively lower in the area with wild fluctuations. The BPNN cannot predict the sharp rise and fall of the Bitcoin price well, which limits the accuracy of the model.

In conclusion, SDAE can predict the price of Bitcoin and its fluctuation effectively. The possible reason is that the SDAE can learn useful information from features by abstracting multiple levels of representation and this thus leads to superior predictive ability. Although the prediction accuracy of the SVR is also competitive, it costs much more time to predict because the optimization of its parameters is very time-consuming. In comparison, the SDAE implements an unsupervised feature learning process that can lead to a better initial value of all the parameters and prevent them from falling into the local optimization. This is another advantage of the deep learning method.

4.4. Robustness tests

To further examine the robustness of the SDAE and the three benchmark methods, t-tests are carried out to statistically test the differences of the prediction results between pairwise models. It is a common method used in the field of machine learning to compare the performance of two prediction methods. Generally, the smaller the p value of the t-statistics is, the greater the difference between the two models. The t-test results of these models on the three evaluation indicators (MAPE, RMSE, and DA) are shown in Table 3.

As shown in Table 3, in terms of the MAPE and RMSE, the t-statistic values of the SDAE are all positive and significant at different levels, which means that the MAPE and RMSE of the SDAE are indeed significantly smaller than those of the other three models. Furthermore, the t-statistic values of the SDAE in terms of the DA are all negative and significant, which means that the DA of the SDAE is indeed significantly larger than those of the other three models. All of these results demonstrate that the SDAE model used in this paper is accurate and robust at predicting the price of Bitcoin.

5. Conclusions

In this study, a deep learning method named the SDAE is utilized to predict the price of Bitcoin. The empirical results show that based on these determinants, the SDAE performs very well in both directional prediction and level prediction, and it is better than the traditional popular machine learning methods such as the SVM and BPNN.

The major conclusions and implications are two-fold. Firstly, the feature system constructed using aspects of the cryptocurrency market, public attention, and the macroeconomic environment is effective and provides some further insights on what explains the wild fluctuation of Bitcoin's price. Secondly, the deep learning method is an effective tool for forecasting the price of Bitcoin. The main reason is that compared with the traditional machine learning methods that only use shallow features, it can learn more information from Bitcoin's sophisticated features by abstracting multiple levels of representation, leading to a superior predictive ability with out-of-sample data.

These conclusions provide decision support for investors and a reference for the governments to design better regulatory policies. Considering the similarities and correlations between different types of digital currencies, the determinant system, and the deep learning model can be used to predict the prices of other cryptocurrencies. At last, it is noteworthy that some recent studies such as (Li et al., 2020b), (Li et al., 2020a; Wei et al., 2019a), and (Wei et al., 2019b) have used the text mining methods to identify the influential factors from textual data such as the massive news , which might provide an important way to expand the determinants system of Bitcoin price.

CRediT authorship contribution statement

Mingxi Liu: Methodology, Software, Investigation, Writing - original draft, Funding acquisition. **Guowen Li:** Methodology, Data curation, Writing - original draft. **Jianping Li:** Conceptualization, Formal analysis, Supervision, Funding acquisition. **Xiaoqian Zhu:** Conceptualization, Project administration, Validation, Writing - review & editing, Funding acquisition. **Yinhong Yao:** Investigation, Visualization.

Declaration of Competing Interest

None.

Acknowledgements

This study was supported by grants from the National Natural Science Foundation of China (71971207, 71901204, 71425002, 71601178) and the Youth Innovation Promotion Association of Chinese Academy of Sciences (2017200).

References

Alstyne, V., Marshall, 2014. Why Bitcoin has value. Commun. ACM 57 (5), 30–32. https://doi.org/10.1145/2594288.
Bakar, N.A., Rosbi, S., 2017. Autoregressive integrated moving average (ARIMA) model for forecasting cryptocurrency exchange rate in high volatility environment: A new insight of bitcoin transaction. International Journal of Advanced Engineering Research and Science 4 (11), 130–137. https://doi.org/10.22161/ijaers.4.11.20.
Bouoiyour, J., Selmi, R., Tiwari, A.K., Olayeni, O.R., 2016. What drives Bitcoin price? Econ. Bull. 36 (2), 843–850. https://hal-univ-pau.archives-ouvertes.fr/hal-01879673.
Ciaian, P., Rajcaniova, M., Kancs, D., 2016. The economics of Bitcoin price formation. Appl. Econ. 48 (19), 1799–1815. https://doi.org/10.1080/00036846.2015.1109038.
Demir, E., Gozgor, G., Lau, C.K.M., Vigne, S.A., 2018. Does economic policy uncertainty predict the Bitcoin returns? An empirical investigation. Finance Res. Lett. 26, 145–149. https://doi.org/10.1016/j.frl.2018.01.005.
Dong, W., Liao, S., Zhang, Z., 2018. Leveraging financial social media data for corporate fraud detection. J. Manage. Inf. Syst. 35 (2), 461–487. https://doi.org/10.1080/07421222.2018.1451954.
Drucker, H., Burges, C.J.C., Kaufman, L., Smola, A., Vapnik, V., 1996. Support vector regression machines. In: Proceedings of the 9th International Conference on Neural Information Processing Systems, pp. 155–161. https://dl.acm.org/doi/10.5555/2998981.2999003.
Fry, J., Cheah, E., 2016. Negative bubbles and shocks in cryptocurrency markets. Int. Rev. Financ. Anal. 47, 343–352. https://doi.org/10.1016/j.irfa.2016.02.008.
Godarzi, A.A., Amiri, R.M., Talaei, A., Jamasb, T., 2014. Predicting oil price movements: a dynamic artificial neural network approach. Energy Policy 68, 371–382. https://doi.org/10.1016/j.enpol.2013.12.049.
Greaves, A., Au, B., 2015. Using the Bitcoin Transaction Graph to Predict the Price of Bitcoin. Stanford University. Working paper.
Hinton, G.E., Salakhutdinov, R.R., 2006. Reducing the dimensionality of data with neural networks. Science 313 (5786), 504–507. https://doi.org/10.1126/science.1127647.
Hong, W., 2011. Traffic flow forecasting by seasonal SVR with chaotic simulated annealing algorithm. Neurocomputing 74 (12-13), 2096–2107. https://doi.org/10.1016/j.neucom.2010.12.032.
Huang, S., Chuang, P., Wu, C., Lai, H., 2010. Chaos-based support vector regressions for exchange rate forecasting. Expert Syst. Appl. 37 (12), 8590–8598. https://doi.org/10.1016/j.eswa.2010.06.001.
Indera, N.I., Yassin, I.M., Zabidi, A., Rizman, Z.I., 2018. Non-linear autoregressive with exogeneous input (NARX) Bitcoin price prediction model using PSO-optimized parameters and moving average technical indicators. J. Fundam. Appl. Sci. 9 (3S), 791–808. https://doi.org/10.4314/jfas.v9i3s.61.
Kristoufek, L., 2015. What are the main drivers of the Bitcoin price? Evidence from wavelet coherence analysis. PLoS One 10 (4), e0123923. https://doi.org/10.1371/journal.pone.0123923.
Li, J., Li, G., Liu, M., Zhu, X., Wei, L., 2020a. A novel text-based framework for forecasting agricultural futures using massive online news headlines. Int. J. Forecast. in press https://doi.org/10.1016/j.ijforecast.2020.02.002.
Li, Jingyu, Li, Jianping, Zhu, Xiaoqian, 2020b. Risk dependence between energy corporations: A text-based measurement approach. Int. Rev. Econ. Finance 68, 33–46. https://doi.org/10.1016/j.iref.2020.02.009.
Li, J., Li, J., Zhu, X., Yao, Y., Casu, B., 2020c. Risk spillovers between FinTech and traditional financial institutions: evidence from the U.S. Int. Rev. Financ. Anal. 71, 101544. https://doi.org/10.1016/j.irfa.2020.101544.
Li, S., Chen, T., Wang, L., Ming, C., 2018. Effective tourist volume forecasting supported by PCA and improved BPNN using Baidu index. Tour. Manage. 68 (10), 116–126. https://doi.org/10.1016/j.tourman.2018.03.006.
Li, X., Wang, C., 2017. The technology and economic determinants of cryptocurrency exchange rates: the case of bitcoin. Decis. Support Syst. 95, 49–60. https://doi.org/10.1016/j.dss.2016.12.001.
Liu, Z., Jiang, P., Zhang, L., Niu, X., 2020. A combined forecasting model for time series: Application to short-term wind speed forecasting. Appl. Energy 259, 114137. https://doi.org/10.1016/j.apenergy.2019.114137.
Lu, C., Wu, J., 2011. An efficient CMAC neural network for stock index forecasting. Expert Syst. Appl. 38 (12), 15194–15201. https://doi.org/10.1016/j.eswa.2011.05.082.
Mai, F., Shan, Z., Bai, Q., Wang, X., Chiang, R.H.L., 2018. How does social media impact Bitcoin value? A test of the silent majority hypothesis. J. Manage. Inf. Syst. 35 (1), 19–52. https://doi.org/10.1080/07421222.2018.1440774.
Matkovskyy, R., Jalan, A., 2019. From financial markets to Bitcoin markets: aA fresh look at the contagion effect. Finance Res. Lett. 31, 93–97. https://doi.org/10.1016/j.frl.2019.04.007.
McNally, S., Roche, J., Caton, S., 2018. Predicting the price of bitcoin using machine learning. In: The 26th Euromicro International Conference on Parallel, Distributed and Network-based Processing, pp. 339–343. https://doi.org/10.1109/PDP2018.2018.00060.
Narendra, K.S., Parthasarathy, K., 1990. Identification and control of dynamical systems using neural networks. IEEE Trans. Neural Netw. 1 (1), 4–27. https://doi.org/10.1109/72.80202.
Sezer, O.B., Gudelek, M.U., Ozbayoglu, A.M., 2020. Financial time series forecasting with deep learning: a systematic literature review: 2005-–2019. Appl. Soft Comput. 90, 106181. https://doi.org/10.1016/j.asoc.2020.106181.
Succurro, M., Arcuri, G., Costanzo, G.D., 2019. A combined approach based on robust PCA to improve bankruptcy forecasting. Rev. Account. Finance 18 (2), 296–320. https://doi.org/10.1108/RAF-04-2018-0077.
Sun, X., Liu, M., Sima, Z., 2020. A novel cryptocurrency price trend forecasting model based on LightGBM. Finance Res. Lett. 32, 101084. https://doi.org/10.1016/j.frl.2018.12.032.
Urquhart, A., 2016. The inefficiency of Bitcoin. Econ. Lett. 148, 80–82. https://doi.org/10.1016/j.econlet.2016.09.019.
Vincent, P., Larochelle, H., Lajoie, I., Bengio, Y., Manzagol, P., 2010. Stacked denoising autoencoders: learning useful representations in a deep network with a local denoising criterion. J. Mach. Learn. Res. 11 (12), 3371–3408. https://dl.acm.org/doi/10.5555/1756006.1953039.
Wang, J., Zhou, H., Hong, T., Li, X., Wang, S., 2020. A multi-granularity heterogeneous combination approach to crude oil price forecasting. Energy Econ. in press https://doi.org/10.1016/j.eneco.2020.104790.
Wei, Lu, Li, Guowen, Zhu, Xiaoqian, Li, Jianping, 2019a. Discovering bank risk factors from financial statements based on a new semi-supervised text mining algorithm. Accounting Finance 53 (3), 1525–1558. https://doi.org/10.1111/acfi.12453.
Wei, Lu, Li, Guowen, Zhu, Xiaoqian, Sun, Xiaolei, Li, Jianping, 2019b. Developing a hierarchical system for energy corporate risk factors based on textual risk disclosures. Energy Econ. 80, 452–460. https://doi.org/10.1016/j.eneco.2019.01.020.
Yu, F., Xu, X., 2014. A short-term load forecasting model of natural gas based on optimized genetic algorithm and improved BP neural network. Appl. Energy 134, 102–113. https://doi.org/10.1016/j.apenergy.2014.07.104.

Zhao, Y., Li, J., Yu, L., 2017. A deep learning ensemble approach for crude oil price forecasting. Energy Econ. 66, 9–16. https://doi.org/10.1016/j.eneco.2017.05.023.
Zhu, X., Li, J., Wu, D., Wang, H., Liang, C., 2013. Balancing accuracy, complexity and interpretability in consumer credit decision making: a C-TOPSIS classification approach. Knowl.-Based Syst. 52, 258–267. https://doi.org/10.1016/j.knosys.2013.08.004.
Zhu, X., Wei, L., Li, J., 2020. A two-stage general approach to aggregate multiple bank risks. Finance Res. Lett. in press https://doi.org/10.1016/j.frl.2020.101688.