

Research Article

DLI: A Deep Learning-Based Granger Causality Inference

Wei Peng 

College of Economics and Management, Shandong University of Science and Technology, Qingdao 266590, China

Correspondence should be addressed to Wei Peng; pengweisd@foxmail.com

Received 31 March 2020; Revised 1 May 2020; Accepted 9 May 2020; Published 9 June 2020

Guest Editor: Lei Xie

Copyright © 2020 Wei Peng. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Integrating autoencoder (AE), long short-term memory (LSTM), and convolutional neural network (CNN), we propose an interpretable deep learning architecture for Granger causality inference, named deep learning-based Granger causality inference (DLI). Two contributions of the proposed DLI are to reveal the Granger causality between the bitcoin price and S&P index and to forecast the bitcoin price and S&P index with a higher accuracy. Experimental results demonstrate that there is a bidirectional but asymmetric Granger causality between the bitcoin price and S&P index. And the DLI performs a superior prediction accuracy by integrating variables that have causalities with the target variable into the prediction process.

1. Introduction

Time series is a series of observation values of a variable arranged in a chronological order, which reflects the change of a phenomenon itself with time if there are no exogenous variables. Generally speaking, time series analysis focuses more on predicting the future based on the existing historical data [1–3] than interpreting the causalities which may exist among the variables. Exploring the causalities among financial time series can be important for portfolio management [4]. As a decentralized cryptocurrency, bitcoin has attracted more and more investors and traders owing to high-investment returns in recent years [5]. From January 1, 2014, to December 31, 2018, bitcoin price jumped from \$771 to \$3742 (USD), which made bitcoin a promising investment cryptocurrency. Interestingly, Yermack [6] asserted that bitcoin was not a currency as it performs poorly as a unit of account and as a store of value. And Corbet et al. [7] supported the conclusion of Yermack that bitcoin was a speculative asset rather than a currency. Moreover, Dyhrberg [8] proved that bitcoin can serve as a hedge against the stock market, and it is a helpful tool for both portfolio diversification and risk management. Therefore, it is of great importance for investors and traders to forecast the bitcoin price and investigate the causes of its volatility.

In most circumstances, causality inference among financial time series is based on the Granger causality [9]. As a predictive causality, the Granger causality refers to that a time series x Granger-causes y if x 's values provide statistically significant information about future values of y , i.e., predictions of y based on its prior values, and the prior values of x are better than predictions of y based only on its prior values. Some traditional approaches for Granger causality inference mainly include vector autoregression (VAR) [10], vector error correction model (VECM) [11], and their variants [12, 13]. VAR and VECM are valid mostly when the input is stationary data. However, the results of some unit root test methods, such as ADF [14], showed that most economic time series are not stationary, while they may be stationary after preprocessing. Hence, traditional Granger causality inference for nonstationary time series needs to preprocess the input to reach a stationary sequence, which may bring pretesting distortions. The Wald test [15] has attracted much attention because there is no pretesting distortion, and it is based on a standard asymptotical distribution, irrespective of the unit roots and the cointegrating properties of the data [16]. However, the Wald test method may be inefficient since it intentionally overfits the VAR. Moreover, those aforementioned approaches are not good at capturing the complex representation of the input data.

Deep learning-based architecture could learn more abstract representation from the input data without data stationarity requirement. Chong et al. [17] proposed a deep learning-based stock market forecasting model to examine the ability of three unsupervised feature extraction methods of predicting future market behaviour. Based on a deep learning model, Chen et al. [18] built a computer-aided diagnosis and decision-making system for medical data from MR images. Long et al. [19] proposed a multi-filter neural network that integrated convolutional and recurrent neurons for feature extraction on economic time series samples and price volatility prediction. And the aforementioned deep learning-based forecasting models achieved promising forecasting performances. Lahmiri and Bekiros [20] employed LSTM for cryptocurrency prediction, which proved deep learning was highly efficient in predicting the inherent chaotic dynamics of cryptocurrency markets. Those aforementioned deep learning-based models are prone to perform better than traditional econometric methods, which suggest the deep learning-based architecture is more potent in dealing with financial time series data.

In this paper, we construct a deep learning-based Granger causality inference architecture, named DLI, which consists of AE, CNN, and LSTM. The two contributions of our work are exploring the Granger causality between the bitcoin price and S&P index and predicting the bitcoin price and S&P index with a higher accuracy.

The remainder of this paper is organized as follows. Available datasets we employed are presented in Section 2. The proposed DLI is depicted in Section 3. Experiments and results are introduced in Section 4. Our contributions and future work are summarized in Section 5.

2. Data

We took the bitcoin price¹ and S & P index² as experimental datasets. Both of them can be downloaded from the Yahoo website, and their relative prices are in US dollars. Without loss of generality, we take the daily closing price as the day's price. The descriptive statistics for the bitcoin price and S&P index covering the period from January 1, 2014, to December 31, 2018, can be found in Table 1. The sample of the bitcoin price and S&P index contains 1,826 and 1,258 data points, respectively. Since stock markets are usually closed for holidays or other reasons, we employed AE to remove the data noise caused by default values.

To obtain a desirable model, we divide the experimental data into three parts: 70% training dataset, 10% validation dataset, and 20% test dataset. The training dataset is to reach a sound model, the validation dataset is to further determine the parameters of the whole network, and the test dataset is to test the generalization ability of the model.

3. Model Development

Autoencoder is a simple but powerful unsupervised deep learning model. A typical AE consists of three layers: input layer, hidden layer, and output layer, as shown in Figure 1.

TABLE 1: Descriptive statistics for the bitcoin price and S&P index.

	Bitcoin price	S&P index
Mean	2,588.19	2,255.94
Standard error	83.06	8.81
Median	638.92	2,124.25
Mode	236.15	1,920.03
Standard deviation	3,549.46	312.58
Sample variance	1.26E + 07	97,706.86
Kurtosis	6.11	2.03
Skewness	1.82	0.57
Range	19,319.3	1,188.86
Minimum	178.1	1,741.89
Maximum	19,497.4	2,930.75
Count	1,826	1,258

And its output layer is an approximate reconstruction of the input layer, which can be used for filtering and representation learning. In the proposed DLI, we adopt AE as a filter to denoise the origin input, which is helpful for improving prediction accuracy.

Long short-term memory is a widely used deep learning model, which focuses on processing sequence data, such as time series data and speech. It is an extension of the recurrent neural network by adding the gate mechanism, which shows a better performance in long-term prediction. In the proposed DLI, we hope it can achieve a long-term accurate prediction by introducing the LSTM model.

Convolutional neural network is also a widely used deep learning model [21], which focuses on processing time series data (1D CNN), image (2D CNN), and video or medical image (3D CNN). CNN includes the convolution layer and pooling layer, as shown in Figure 1. And it can greatly reduce the amount of parameters and speed up training by local receptive fields and shared weights. Moreover, LeCun and Bengio [22] showed that time series have a strong 1D structure: variables that are spatially or temporally nearby are highly correlated, and CNN can effectively extract the spatial feature of time series. Therefore, CNN is introduced into the proposed DLI to extract the spatial feature and to speed up training.

Figure 1 shows the graphic illustration of the DLI which consists AE, CNN, and LSTM. We assume that both S&P index (X) and bitcoin price (Y) are time series of length T ,

$$\text{where } X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_t \\ \vdots \\ x_T \end{pmatrix} \in \mathbb{R}^{T \times 1} \text{ and } Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_t \\ \vdots \\ y_T \end{pmatrix} \in \mathbb{R}^{T \times 1}. \text{ Let } x_t$$

be the S&P index at time t and y_t be the bitcoin price at time t .

The DLI consists of three processing stages: denoising, feature extracting, and forecasting. As described in Section 2, since stock markets are usually closed for holidays or other reasons, the S&P index time series has many default values. Therefore, at the denoising stage, AE is firstly used for data filtering to remove the noises in the S&P index. At the feature extracting stage, the denoised S&P index and bitcoin

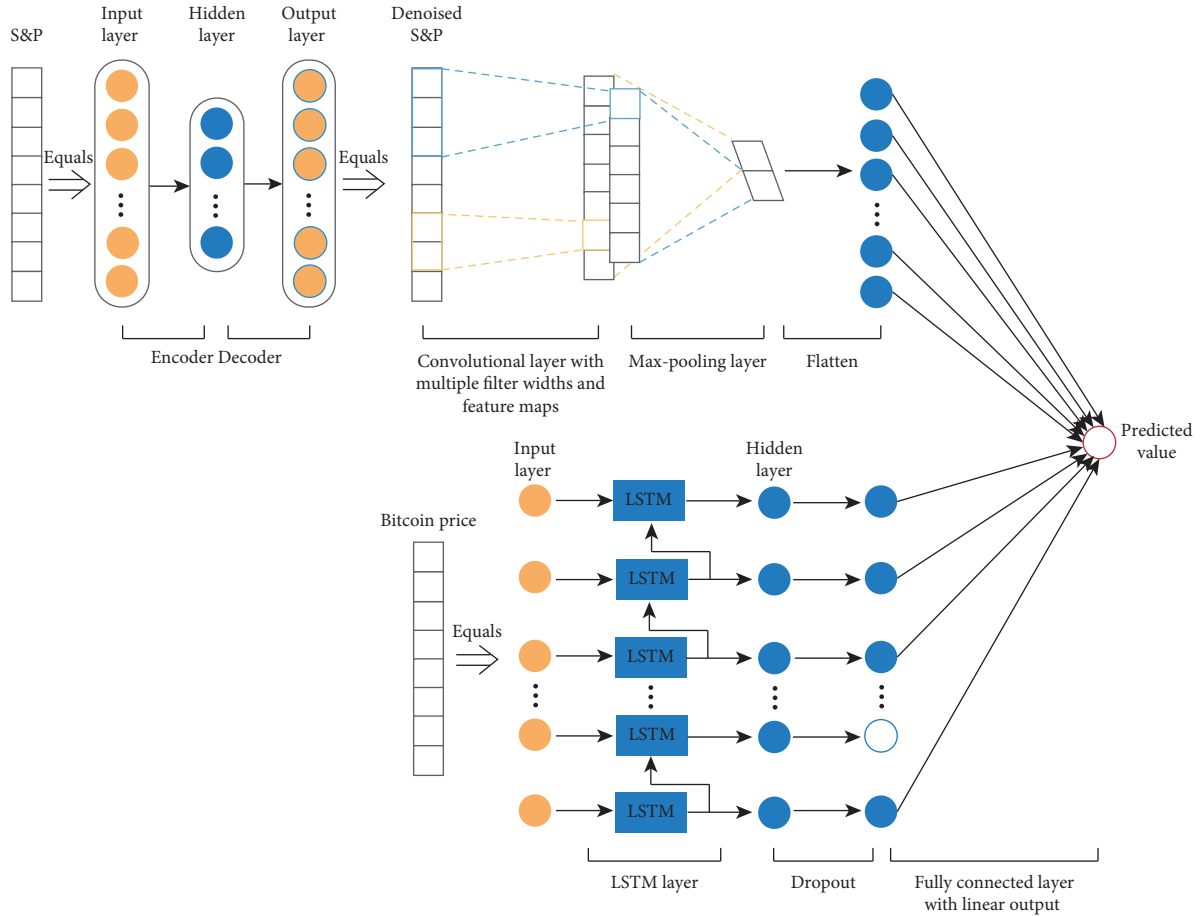


FIGURE 1: The architecture of the proposed DLI.

price would be taken as the inputs of CNN and LSTM to extract deep representations, respectively. At the forecasting stage, we would obtain the bitcoin price prediction through a fully connected layer.

The optimization of the DLI model is to minimize the reconstruction error of AE and the training error of the whole model. At the denoising stage, the output of AE is an approximate copy of the input. Therefore, we have to minimize the reconstruction error between the input and the output, which could maintain the economic significance of the S&P index. The reconstruction error of AE is defined as follows:

$$e_{AE} = \min \sum \|X - g(W' \times f(WX + b) + b')\|^2, \quad (1)$$

where $f(\cdot)$ and $g(\cdot)$ are activation functions, W and W' are weights, and b and b' are biases.

It is necessary for obtaining a sound model to minimize the training error of the whole model. The objective function of the whole model can be described as

$$e_{DLI} = \min \sum \|y_t - \hat{y}_t\|^2, \quad (2)$$

where \hat{y}_t denotes the predicted value.

4. Empirical Results

In this part, we will explore the Granger causality between the bitcoin price and S&P index. To investigate whether the S&P index Granger-causes the bitcoin price, we firstly predict the bitcoin price without considering the S&P index, as shown in Figure 2. Then, for comparison, we take the S&P index as auxiliary information to predict the bitcoin price, as shown in Figure 3. In the same way, to investigate whether the bitcoin price Granger-causes the S&P index, we firstly predict the S&P index without considering the bitcoin price, as shown in Figure 4. Then, for comparison, we take the bitcoin price as auxiliary information to predict the S&P index, as shown in Figure 5. In addition, we employ the traditional approach ARIMA to demonstrate the superiority of the proposed model. Owing to the continuous value prediction, we employ the root mean squared errors (RMSEs) as the forecasting performance indicator. The smaller the RMSE value, the better the prediction performance. And the corresponding prediction RMSEs are shown in Table 2.

From Table 2, we can see that the bitcoin price prediction RMSE of the DLI decreases by 92.10% and 23.32% compared with that of the ARIMA and LSTM, respectively. And the S&P index prediction RMSE of the DLI significantly

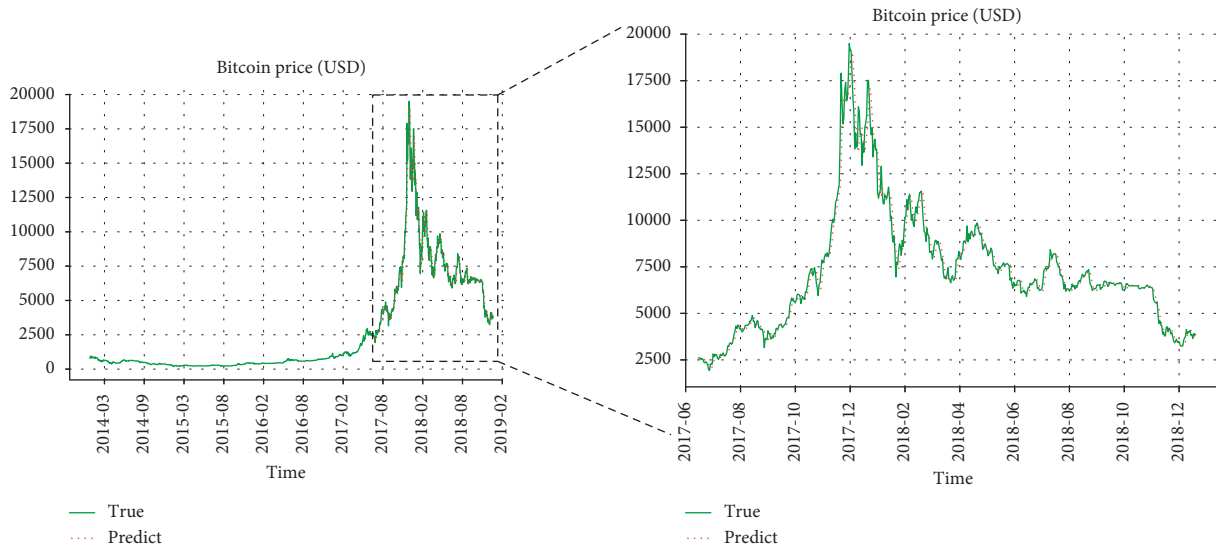


FIGURE 2: The bitcoin price prediction without consideration of the S&P index.

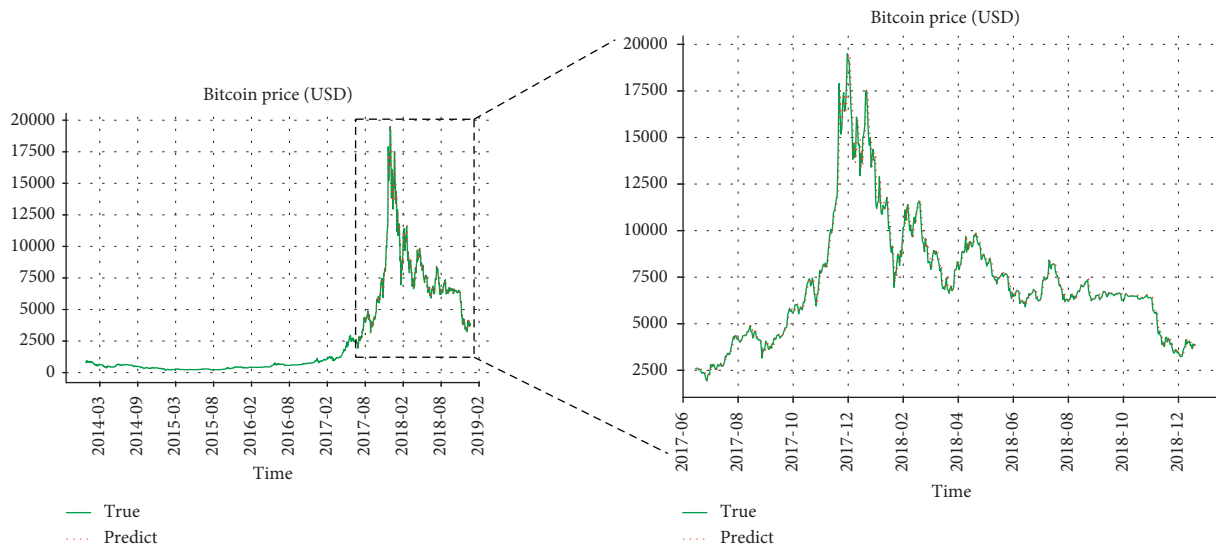


FIGURE 3: The bitcoin price prediction with consideration of the S&P index.

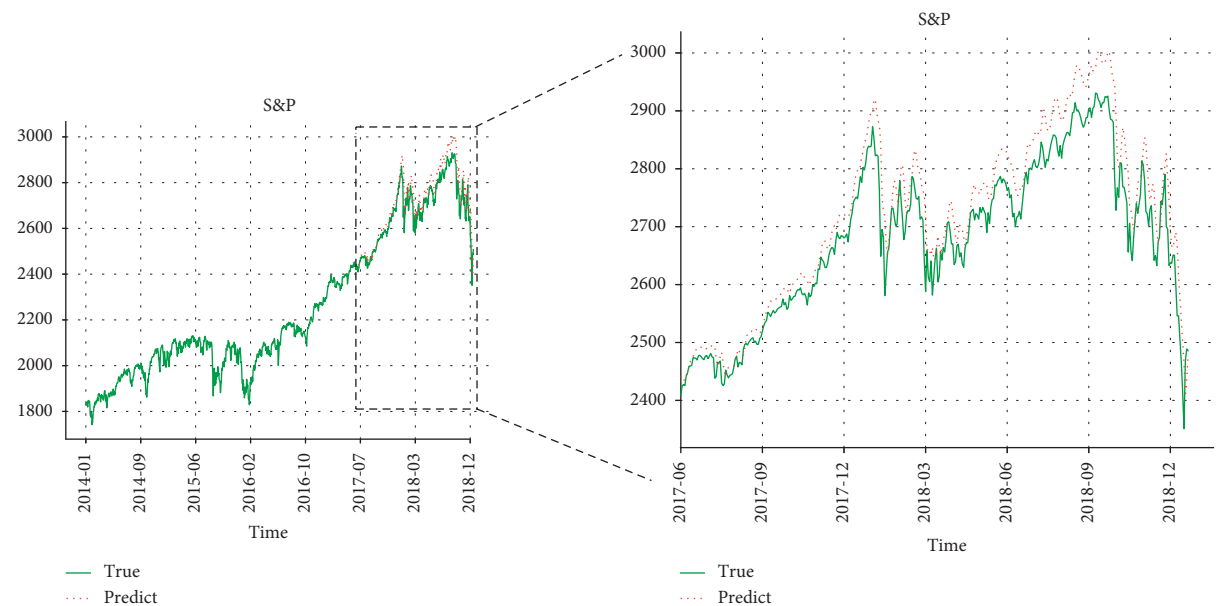


FIGURE 4: The S&P index prediction without consideration of the bitcoin price.

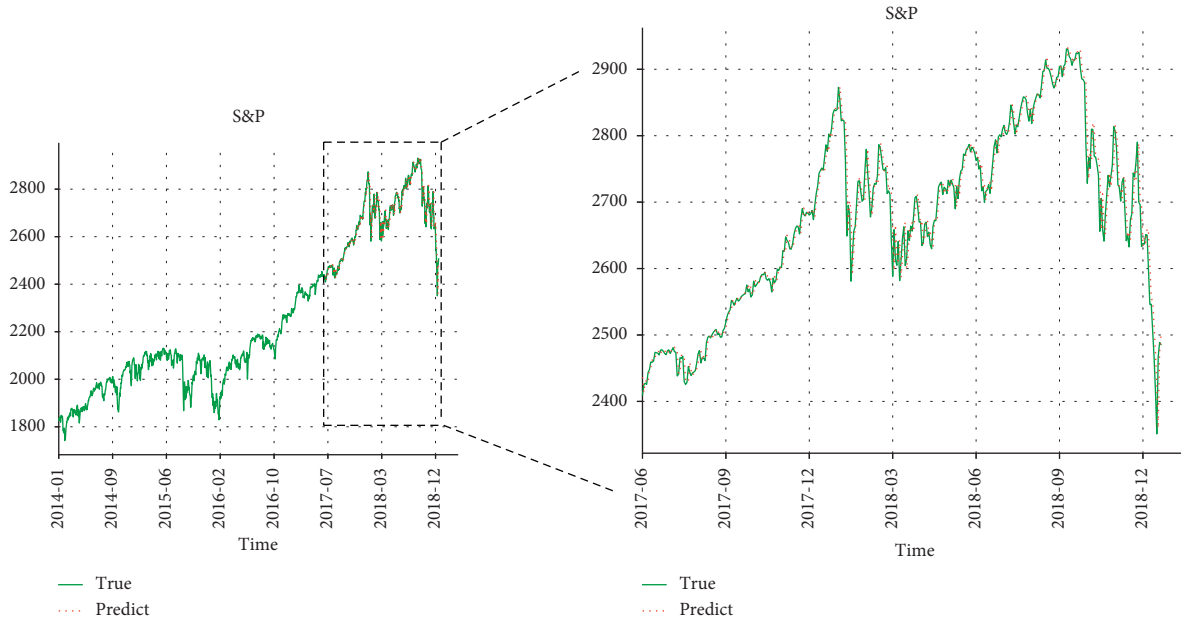


FIGURE 5: The S&P index prediction with consideration of the bitcoin price.

TABLE 2: RMSEs of the prediction.

Prediction task	Whether or not to consider auxiliary information	Approaches	RMSE
Bitcoin price prediction	Without consideration of the S&P index	ARIMA	5,932.96
		LSTM	611.47
	With consideration of the S&P index	DLI	468.87
		Without consideration of the bitcoin price	ARIMA
S&P index prediction	Without consideration of the bitcoin price	LSTM	49.63
		DLI	24.34
	With consideration of the bitcoin price	DLI	24.34

decreases by 98.06% and 50.96% compared with that of the ARIMA and LSTM, respectively. The above results demonstrate that both bitcoin price and S&P index prediction performances would be enhanced with consideration of the S&P index and bitcoin price, respectively. And the prediction performance improvement of the S&P index is more significant than that of the bitcoin price. Therefore, we can conclude that there is a bidirectional but asymmetric Granger causality between the bitcoin price and S&P index.

5. Conclusions

In this paper, we proposed an interpretable deep learning-based Granger causality inference architecture by integrating AE, CNN, and LSTM, named DLI. The proposed DLI, as a deep learning-based model, one of its advantages compared with traditional econometric models is that it can process big data efficiently and retain its original economic significance of variables after data preprocessing.

Our two contributions are exploring the Granger causality between the bitcoin price and S&P index and predicting the bitcoin price and S&P index with a higher accuracy. Our experiments reveal a bidirectional but asymmetric Granger causality between the bitcoin price and S&P index. And the DLI performs a superior prediction

accuracy by integrating variables that have causalities with the target variable into the prediction process.

In future work, the proposed DLI can be extended to some other economic variables to provide a reasonable reference for portfolio management, or it can be used for prediction in other scientific fields. Moreover, the DLI can also be extended from two variables to multivariables to determine causalities among the multitime series.

Data Availability

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation, to any qualified researcher.

Conflicts of Interest

The author declares that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The author acknowledges the National Natural Science Foundation of China (Grant no. 11801060) and the Innovation Program of Shandong University of Science and Technology (no. SDKDYC190114).

References

- [1] Z. Zhou, T. Ren, H. Xiao, and W. Liu, "Time-consistent investment and reinsurance strategies for insurers under multi-period mean-variance formulation with generalized correlated returns," *Journal of Management Science and Engineering*, vol. 4, no. 2, pp. 142–157, 2019.
- [2] K. Wang, K. Li, L. Zhou et al., "Multiple convolutional neural networks for multivariate time series prediction," *Neurocomputing*, vol. 360, pp. 107–119, 2019.
- [3] A. Yadav, C. K. Jha, and A. Sharan, "Optimizing LSTM for time series prediction in Indian stock market," *Procedia Computer Science*, vol. 167, pp. 2091–2100, 2020.
- [4] M. Brière, K. Oosterlinck, and A. Szafarz, "Virtual currency, tangible return: portfolio diversification with bitcoin," *Journal of Asset Management*, vol. 16, no. 6, pp. 365–373, 2015.
- [5] G. Wang, Y. Tang, C. Xie, and S. Chen, "Is bitcoin a safe haven or a hedging asset? evidence from China," *Journal of Management Science and Engineering*, vol. 4, no. 3, pp. 173–188, 2019.
- [6] D. Yermack, "Chapter 2-is bitcoin a real currency? an economic appraisal," in *Handbook of Digital Currency*, D. Lee Kuo Chuen, Ed., pp. 31–43, Academic Press, Cambridge, MA, USA, 2015.
- [7] S. Corbet, B. Lucey, M. Peat, and S. Vigne, "Bitcoin futures-what use are they?" *Economics Letters*, vol. 172, pp. 23–27, 2018.
- [8] A. H. Dyhrberg, "Hedging capabilities of bitcoin. is it the virtual gold?" *Finance Research Letters*, vol. 16, pp. 139–144, 2016.
- [9] C. W. J. Granger, "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica Journal of the Econometric Society*, vol. 37, no. 3, pp. 424–438, 1969.
- [10] X. Hou, S. Li, W. Li, and Q. Wang, "Bank diversification and liquidity creation: panel granger-causality evidence from China," *Economic Modelling*, vol. 71, pp. 87–98, 2018.
- [11] X. Meng and J. Han, "Roads, economy, population density, and CO₂: a city-scaled causality analysis," *Resources, Conservation and Recycling*, vol. 128, pp. 508–515, 2018.
- [12] T. Chang, F. Gatwabayege, R. Gupta, R. Inglesi-Lotz, N. C. Manjezi, and B. D. Simo-Kengne, "Causal relationship between nuclear energy consumption and economic growth in G6 countries: evidence from panel granger causality tests," *Progress in Nuclear Energy*, vol. 77, pp. 187–193, 2014.
- [13] Y. Zhao, S. A. Billings, H. Wei, F. He, and P. G. Sarrigiannis, "A new NARX-based granger linear and nonlinear causal influence detection method with applications to EEG data," *Journal of Neuroscience Methods*, vol. 212, no. 1, pp. 79–86, 2013.
- [14] D. A. Dickey and W. A. Fuller, "Distribution of the estimators for autoregressive time series with a unit root," *Journal of the American Statistical Association*, vol. 74, no. 366a, pp. 427–431, 1979.
- [15] H. Y. Toda and T. Yamamoto, "Statistical inference in vector autoregressions with possibly integrated processes," *Journal of Econometrics*, vol. 66, no. 1-2, pp. 225–250, 1995.
- [16] R. S. Hacker and A. Hatemi-J, "Tests for causality between integrated variables using asymptotic and bootstrap distributions: theory and application," *Applied Economics*, vol. 38, no. 13, pp. 1489–1500, 2006.
- [17] E. Chong, C. Han, and F. C. Park, "Deep learning networks for stock market analysis and prediction: methodology, data representations, and case studies," *Expert Systems with Applications*, vol. 83, pp. 187–205, 2017.
- [18] A. Chen, L. Zhu, H. Zang, Z. Ding, and S. Zhan, "Computer-aided diagnosis and decision-making system for medical data analysis: a case study on prostate MR images," *Journal of Management Science and Engineering*, vol. 4, no. 4, pp. 266–278, 2019.
- [19] W. Long, Z. Lu, and L. Cui, "Deep learning-based feature engineering for stock price movement prediction," *Knowledge-Based Systems*, vol. 164, pp. 163–173, 2019.
- [20] S. Lahmiri and S. Bekiros, "Cryptocurrency forecasting with deep learning chaotic neural networks," *Chaos, Solitons & Fractals*, vol. 118, pp. 35–40, 2019.
- [21] Y. Zheng, X. L. Xu, and L. Y. Qi, "Deep CNN-assisted personalized recommendation over big data for mobile wireless networks," *Wireless Communications & Mobile Computing*, vol. 2019, p. 6082047, 2019.
- [22] Y. LeCun and Y. Bengio, "Convolutional networks for images, speech, and time series," *The Handbook of Brain Theory and Neural Networks*, vol. 3361, no. 10, 1995.