Comparison of the Predictions of Convolutional Neural Networks with Image Arguments and Long Short-Term Memory Neural Networks with Time-Series Arguments for Cryptocurrency Markets

A Misnik[1], S Krutalevich[1], S Prakapenka[1], P Borovykh[2], M Vasiliev[2]

[1]Inter-state educational institution of higher education “Belarusian-Russian university”, Mogilev, Belarus
[2]RoninAI Lab, New York, USA

Abstract. Convolutional neural networks are currently very popular in a wide variety of applications. The aim of this article is to verify the reasonableness for using convolutional neural networks in cases when time-series data is available. The predictions of a convolutional neural network, that analyzes the graphical representation of a time-series are compared with the predictions of long short-term memory network, that analyzes time-series data in numerical representation. We show how the accuracy of cryptocurrency predictions could be improved with time-series data inputs versus image arguments.

1 Introduction

First of all, we have to decide which data to use in our study. For our study, we need data that can be presented as numerical time-series as well as images.

We have chosen cryptocurrency market prices due to the rising popularity of this field. In particular, we used market prices of Bitcoin (BTC), the most popular cryptocurrency, to analyze the difference in predictions between different types of networks with different ways of acquiring inputs.

Results were tested, based on neural networks created by RoninAI Lab. RoninAI uses various neural networks for cryptocurrency rate prediction, lending hands-on data and analysis to this study.

This study was based on two neural network architectures: Convolutional Neural Networks and Long Short-Term Memory neural networks.

Convolutional Neural Network (CNN) is a Deep Learning algorithm that can take in an input image and assign importance (learnable weights and biases) to various aspects/objects in the image.

Long Short-Term Memory (LSTM) is a type of recurrent neural network (RNN). LSTMs excel in learning, processing, and classifying sequential data. [1,2]

The aim of this study is to figure out if the use of CNN on a given time-series data is reasonable when compared to LSTM networks.

In: P. Sosnin, V. Maklaev, E. Sosnina (eds.): Proceedings of the IS-2019 Conference, Ulyanovsk, Russia, 24-27 September 2019, published at http://ceur-ws.org

Market prices of cryptocurrencies are subject to great volatility, partially due to the absence of agreed-upon fundamentals to back up their price value. Thus, the variety of psychological factors and emotional perception of the market situation by the trader is very important when he decides whether to buy or sell a cryptocurrency.

Every cryptocurrency exchange or analytical service has several basic elements, shown on their index pages. The main one – is a price chart. An example of this chart is shown in Figure 1. We believe that the appearance of this chart has a great influence on a trader's decision to buy or sell a cryptocurrency. [5]

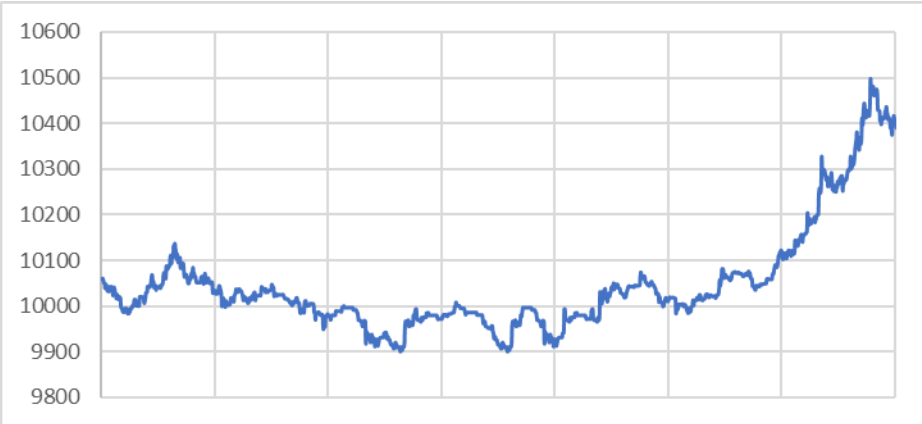

Figure 1. 24-hour BTC Price Chart

For this study, we will compare prediction results made by CNN with chart images used as input data to LSTM with time-series data used as input data. The output will be the change in BTC market price for the next minute.

We will use a statistical F-test to determine the quality of predictions. [9]

2 Experimental Setup

For our setup, we used a server equipped with Intel® Core™ i7-6700 Quad-Core CPU, 64 GB DDR4 RAM, 2 x 500 GB SATA SSD hard drives and GTX 1080 Ti GPU. This setup with a powerful GPU allows for rapid model compilation and training of the model with the data.

We used the Keras library and the TensorFlow library to operate with neural network models.

The data itself is stored and grouped into CSV files. Every CSV file contains two columns and 1440 rows. 1440 rows correspond to 1440 minutes in a 24-hour timeframe. The first column includes the closing market price for each minute and the second column represents a minute change of the market price.

We had 525600 files with training examples in total, matching every minute situation during the year. Thus, the critical F-test value using alpha=0.05 is 1.83.

3 CNN

First of all, we prepared training samples for CNN. To do it, we generated an image for every CSV file. The example of such image is shown in Figure 2.

Figure 2. Learning example for CNN

Our CNN consists of three convolutional layers and Multilayer perceptron.
The experimental results, using trained CNN are shown in Figure 3.

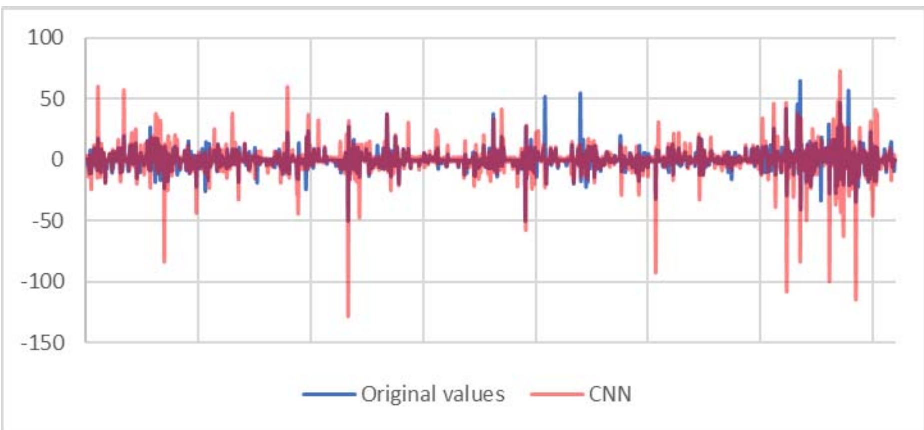

Figure 3. Trained CNN results

Using trained CNN results, the calculated F-test value turned out to be -2.9. Although the value of -2.9 is above critical, this level is still low.

4 LSTM

To train LSTM network we just fed the CSV files.

Our LSTM network consists of one LSTM layer and three direct distribution layers.

The experimental results using trained LSTM are shown in Figure 4.

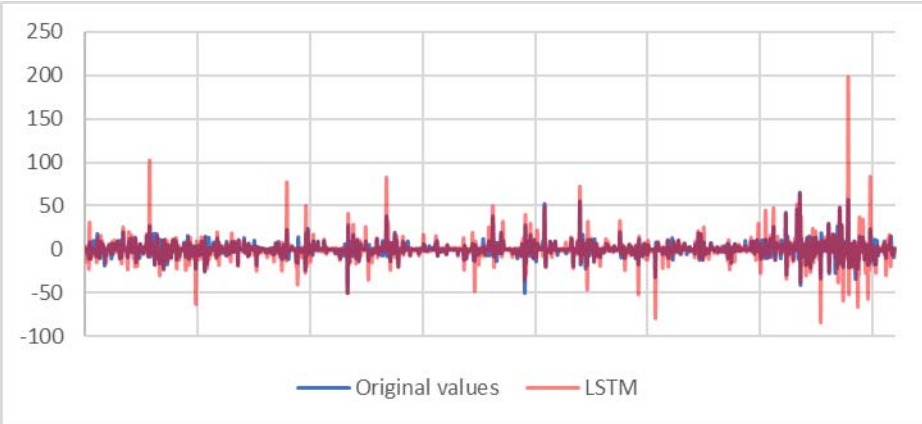

Figure 4. Trained LSTM results

Using trained LSTM results, the calculated F-test value turned out to be -1.5. The value of -1.5 is below critical.

5 Evolving LSTM

Let's consider an algorithm that allows us to identify some complex indicators of a numerical series, based on which we can predict a trader's subjective assessment of chart appearance (Figure 1).

In general, any series of data can be considered as a sum of linear and harmonic components.

The purpose of further research is to investigate the algorithm for isolating these components and their normalization. [3,4,5]

The proposed algorithm assumes the following stages.

1. Definition of the minimal element E_{min} of the series E.

2. Carrying out the subtraction operation $E' = E_i - E_{min}$.

3. Approximation of the series E' by a polynomial

$$E_1 = a(0) + a(1) * n, \quad (1)$$

where a(0) and a(1) are the approximation coefficients; n-discrete values of the time axis.

4. Calculation of the coefficient of relative change in the linear component over the period T by the formula:

$$E_L = a\,(1) * T * 100 / E_{min}, \tag{2}$$

The value of E_L is relative and does not depend on the absolute value of the series. If the quantity $E_L > 0$ then the linear component increases.

5. To estimate the harmonic component of the series, we perform a Fourier transform (FT) for the series E'.

6. Determine the moduli of the oscillation amplitudes in the frequency domain A(w). Carry out the filtering of frequencies according to the amplitude values.

To analyze the efficiency of the proposed algorithm, we used MATLAB environment.

Step 1. The linear component is constant. The harmonic component is absent. The results are shown in Figure 5. The value of the coefficient EL displayed on the second chart. In this case, $E_L = 0$.

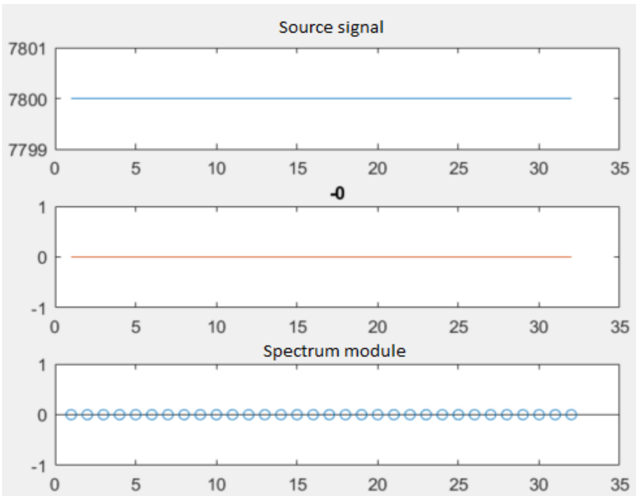

Figure 5. Step 1 of the Analysis

Step 2. The linear component grows. The harmonic component is absent. The results are shown in Figure 5. The coefficient EL = 20%. The spectrum modules have values in the low-frequency range (1-2 Hz.). It should be remembered that the main frequency band for the module of the real sequence lies in the interval 0<k <= N / 2-1.

Therefore, frequencies above 15 Hz in our example should be ignored. If it is necessary to analyze higher frequencies, we will increase N - the number of sampling points, Figure 6.

Step 3. The linear component decreases. The harmonic component is absent. The results are shown in Figure 7.

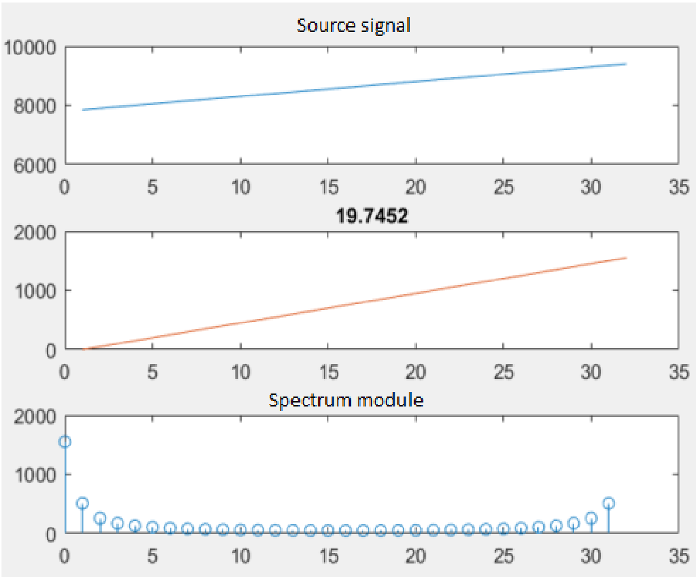

Figure 6. Step 2 of the Analysis

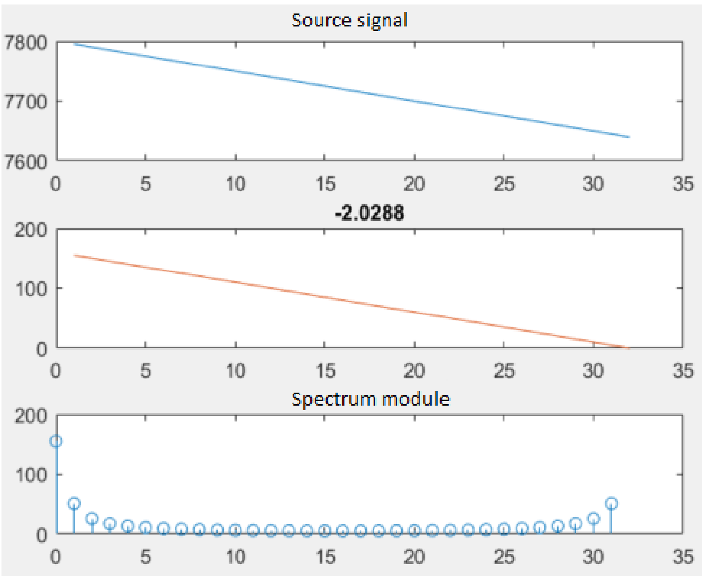

Figure 7. Step 3 of the Analysis

Step 4. The linear component is absent. The harmonic component is present at a frequency of 3 Hz - amplitude 20. The results are shown in Figure 8.

Step 5. The linear component increases. The harmonic component is present at a frequency of 3 Hz - an amplitude of 400. The results are shown in Figure 9.

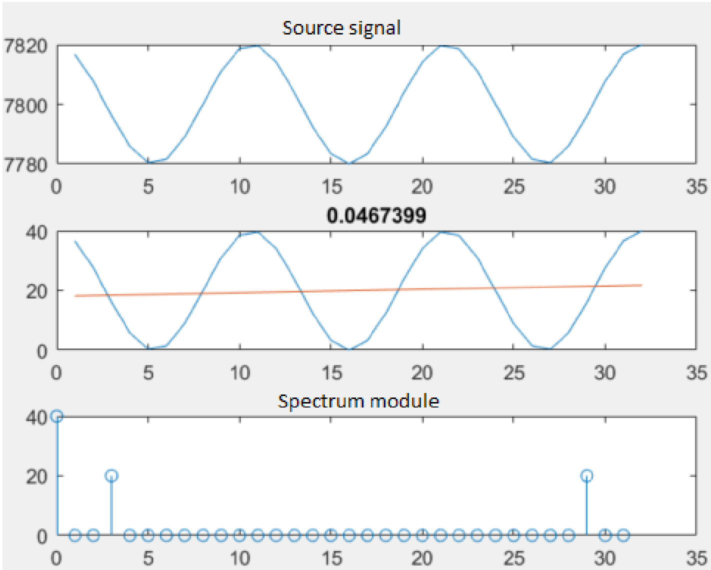

Figure 8. Step 4 of the Analysis

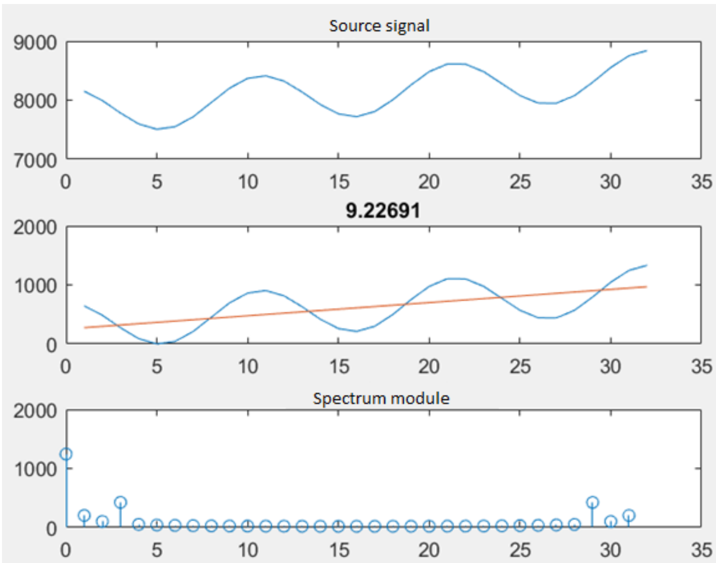

Figure 9. Step 5 of the Analysis

We applied the proposed algorithm to the data presented in Figure 1. The linear component predicted an uptrend with a rate of 10.2% per period. Modules of amplitudes of the harmonic components had an influence in frequencies of up to 10 Hz.

The proposed algorithm distinguishes the linear and harmonic components of the numerical series.

The proposed coefficient E_L - can act as a measure of the trend of changes in the values of a numerical series.

Moduli of the amplitude of the oscillations in the frequency domain after the Fourier transformation can act as a measure of the estimation of the vibrational component of a series of data.

In our case, we've extended inputs of LSTM neural network with 4 new parameters: coefficient E_L, and 3 harmonic component ranges: 0 to 10Hz, 10 to 30Hz, 30Hz and higher.

The experimental results using trained LSTM with additional inputs are shown in Figure 10.

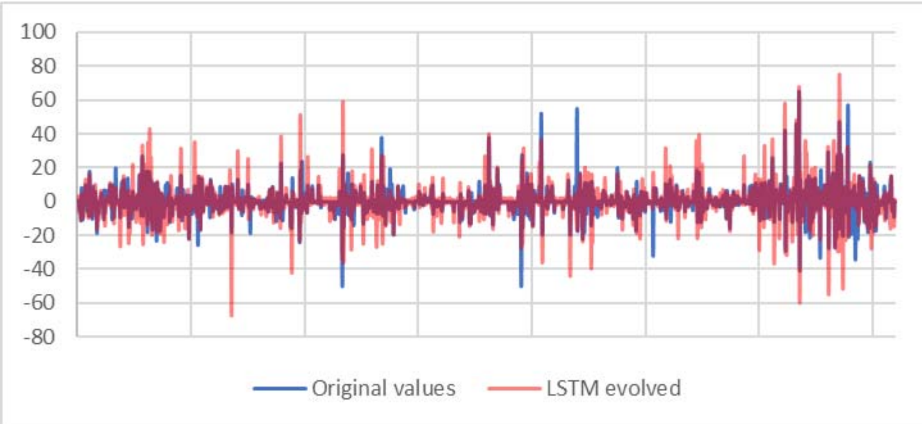

Figure 10. Trained LSTM results with additional inputs

Using trained LSTM with additional inputs results we conducted the F-test resulting in the F-test value of -5.8. This value is above critical, and better than CNN.

7 Conclusion

In this paper we compared CNN and LSTM efficiency to process time-series data in numerical and image form.

In our experiment, CNN shows non-critical results in terms of F-test, while LSTM with only one input – BTC market price, shows under critical results. But the extension of inputs, based on series data analysis can improve LSTM performance to the non-critical level, prevailing results produced by CNN.

References:

1. S. Haykin, Neural Networks and Learning Machines (3rd Edition), Prentice Hall, 2009
2. Tax, N.; Verenich, I.; La Rosa, M.; Dumas, M. (2017). "Predictive Business Process Monitoring with LSTM neural networks". Proceedings of the International Conference on Advanced Information Systems Engineering (CAiSE): 477–492. arXiv:1612.02130 Freely accessible. doi:10.1007/978-3-319-59536-8_30.
3. A. Misnik Neural network modeling of heat supply systems // "Neurocomputers: development, application" - 2012. №2. - p 62-73.
4. V. Borisov, A. Misnik Combined neural network modeling for the operational management of complex systems / / Information technology. - 2012. - No. 7. - P. 69-72.
5. Anton Misnik, Sergey Krutolevich, Siarhei Prakapenka, Max Vasilyeu Neural Network Approximation Precision Change Analysis on Cryptocurrency Price Prediction p. 96-101 Proceedings of the II International Scientific and Practical Conference “Fuzzy Technologies in the Industry – FTI 2018”
6. Wierstra, Daan; Schmidhuber, J.; Gomez, F. J. (2005). "Evolino: Hybrid Neuroevolution/Optimal Linear Search for Sequence Learning". Proceedings of the 19th International Joint Conference on Artificial Intelligence (IJCAI), Edinburgh: 853–858.
7. Stein, Elias; Shakarchi, Rami (2003), Fourier Analysis: An introduction, Princeton University Press, ISBN 978-0-691-11384-5.
8. Rahman, Matiur (2011), Applications of Fourier Transforms to Generalized Functions, WIT Press, ISBN 978-1-84564-564-9.
9. Abarbanel, Henry (Nov 25, 1997). Analysis of Observed Chaotic Data. New York: Springer. ISBN 978-0387983721.