

Cryptocurrency Forecasting using α -Sutte Indicator, ARIMA, and Long Short-Term Memory

Apriliyanus Rakhmadi Pratama¹, Sigit Nugroho², Ketut Sukiyono³
pratama.gokilz@gmail.com¹, snugroho@unib.ac.id², ksukiyono@unib.ac.id³

Department of Statistics, University of Bengkulu, Bengkulu 38371, Indonesia¹, Department of Statistics, University of Bengkulu, Bengkulu 38371, Indonesia², Department of Agricultural Economics, University of Bengkulu, Bengkulu 38371, Indonesia³

Abstract. The purpose of these studies are to obtain bitcoin price predictions using three different approach in forecasting methods : ARIMA model, α -sutte indicator and LSTM algorithm, and to find out the accuracy level of the three methods in forecasting bitcoin's price as well. Bitcoin closing's price each day taken from website of coin market starting from April 29 2013 to February 06 2019 was analyzed using R and Python softwares. Based on the smallest value of MSE, MAPE, and MAD, the LSTM algorithm gave the best prediction, followed by the α -Sutte indicator and the ARIMA model.

Keywords: bitcoin, forecasting, a-Sutte indicator, ARIMA, LSTM.

1 Introduction

The instability of cryptocurrency bitcoin's price in the last few years encourages investor to find the most effective and efficient ways to make right the decision to achieve maximum profits. The uncertainty of bitcoin's price reflects that cryptocurrency has different system from the electronic money (e-money) that comes from the conventional banks. Unlike centralized conventional banks, cryptocurrency use decentralized control, which means the prices are not regulated by a single system rather than works through distributed ledger technology, typically a block chain, which serves as a public financial transaction database. It causes bitcoin's price being very volatile, dynamic and probabilistic. One way that can be used to overcome the problems is the forecasting approach. The forecasting approach theoretically offers many methods including Time Series Analysis, Neural Network, and α -Sutte Indicator.

ARIMA models are one of the model in Time Series Analysis, stands for AutoRegressive Integrated Moving Average, with (p) as an Auto Regressive (AR) operator, the Integrated word refers to degrees of differencing (d) and (q) as a Moving Average (MA) operator [1]. ARIMA models are fitted to time series data either to better understand the data or to predict future points in the time series (forecasting). They are applied in some cases where data show evidence of non-stationery, where an initial differencing step can be applied one or more times to eliminate the non-stationery. Long Short Term-Memory algorithm is one of Neural Network algorithm, developed from Recurrent Neural Network (RNN) algorithm and popularized by Hochreiter and Schmidhuber [2]. The development is to overcome RNN algorithm problems that cause inaccurate result when analyzing long term data [3]. The α -sutte indicator is the development of the sutte indicator by Ahmar [4] in 2017. Sutte indicator was

originally developed based on the modification of the MA indicator by considering prices at the opening, closing, lowest and highest [5]. The α -sutte indicator aims to increase the accuracy of forecasting and can be used to predict time series data.

2Notation

2.1 ARIMA Model

The following is the general form of ARIMA models:

$$\phi_p(B)(1-B)^d Z_t = \theta_q(B)\alpha_t \quad (1)$$

where Z_t is actual bitcoin's price at time t , ϕ_p is the p -th degree of AR operator, θ_q is the q -th degree of MA operator, α_t is the error at time t , while B is the backshift operator.

2.2LSTM Algorithm

The LSTM algorithm is composed of an input layer, one or more hidden layers, and an output layer. The structure of LSTM memory cell is illustrated in Figure 1. Each memory cell has three types of gates, the forget gate, the input gate, and the output gate to maintain and adjust its cell states s_c^t . The forget gate b_f^t defines which information is deleted from memory (cell state), the input gate b_i^t specifies which information is added to the memory (cell state), the output gate b_o^t specifies which information from memory (cell state) to be used as output information [5]. At the t -th step, LSTM decide what information will remove from the cell state through the decision by a transformation function (*sigmoid* and *tanh*) in forget layer. Every step to the forget layer, LSTM will decide what information will be removed or stored from the cell state. The forget layer will produce output in the form of numbers between 0 and 1, where number 1 is interpreted as "completely keep this", while the number 0 is interpreted as "completely forget this" [5][6].

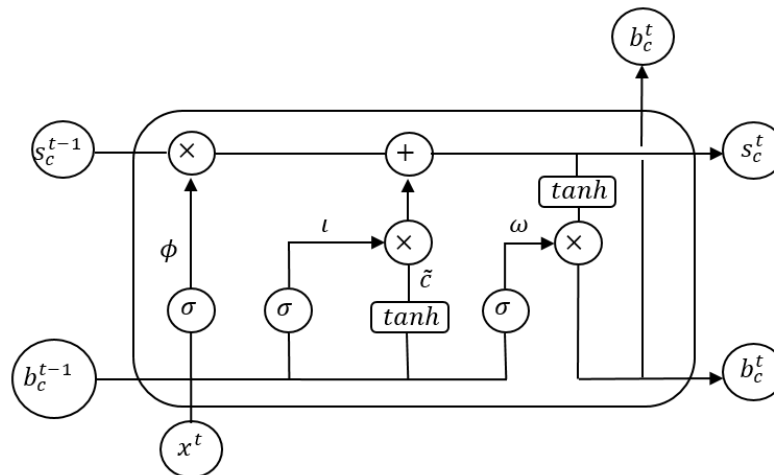


Fig.1. Memory Cell of LSTM algorithm

2.3 Alpha-Sutte Indicator

The α -sutte indicator is the development of the sutte indicator by Ahmar in 2017 which aims to increase the accuracy of forecasting and can be used to predict time series data. The α -sutte indicator has equation as followed:

$$\alpha_t = (\alpha(2\Delta x / (\alpha + \delta)) + \beta(2\Delta y / (\beta + \alpha)) + \gamma(2\Delta z / (\gamma + \beta))) / 3 \quad (2)$$

where a_t = observation at time t, $\delta = a_{t-4}$, $\alpha = a_{t-3}$, $\beta = a_{t-2}$, $\gamma = a_{t-1}$, $\Delta x = \alpha - \delta$, $\Delta y = \beta - \alpha$, and $\Delta z = \gamma - \beta$.

2.4 Performance evaluation indicators

Mean Square Error (MSE), Mean Percentage Error (MAPE), and Mean Absolute Deviation are used to evaluate the forecasting performances.

$$MSE = (\sum_{t=1}^n Z_t) / n \quad (3)$$

$$MAPE = \sum_{t=1}^n |(Z_t - \hat{Z}_t) / Z_t| 100\% / n \quad (4)$$

$$MAD = \sum_{t=1}^n |Z_t - \hat{Z}_t| / n \quad (5)$$

where Z_t is the actual bitcoin's price and \hat{Z}_t is predicted price of bitcoin.

3 Methods

The data are the bitcoin closing's price each day starting from the period 29 April 2013 to 06 February 2019 obtained from the site www.coinmarketcap.com. They are divided into training data to form a model and testing data to test the accuracy of predictions from the model built. The number of training and testing data are 2104 and 7, respectively. From Figure 2 we can see that in the first 4 years bitcoin's price were pretty stable and low, then in the early 2017 bitcoin's price started to increase up until late 2017 reached its highest price at \$19,497.4. But then they continued to fall sharply.

3.1 ARIMA Model Procedures

- Plot ACF and PACF on training data. The result shows in Figure 3. From the ACF plot, it shows a dies-down; while the PACF plot shows cut-off at lag-2. Therefore, the training data indicates an AR(2) model.
- Stationary test was held on training data, the result of ADF's test indicate that the training data are not stationary.
- Figure 4 shows that training data are not stationary in variances and mean. To overcome instability in variance, log transformation was held on the training data.
- Then, transformed data will be differencing 1 time.
- Redo stationary test on differenced data. From the ADF's test, it is said that the differenced data are stationary, as seen in Figure 5.
- The plot of ACF and PACF stationary data are shown in Figure 6. As we can see, the plot show no significant lag period, so we cannot evaluate the number of periods of

p and q . Then, Python's syntax is used to find more candidate models with value of p, d, q within range 0 – 4.

- g) Then, do white-noise test and residual test on best candidate model by smallest AIC's value.
- h) Forecast using best ARIMA model.

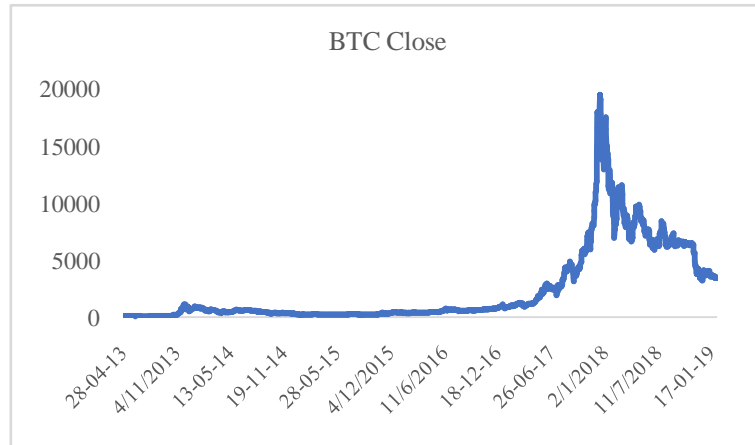


Fig.2. Plot of bitcoin's price each day from period April 29 2013 to February 06 2019

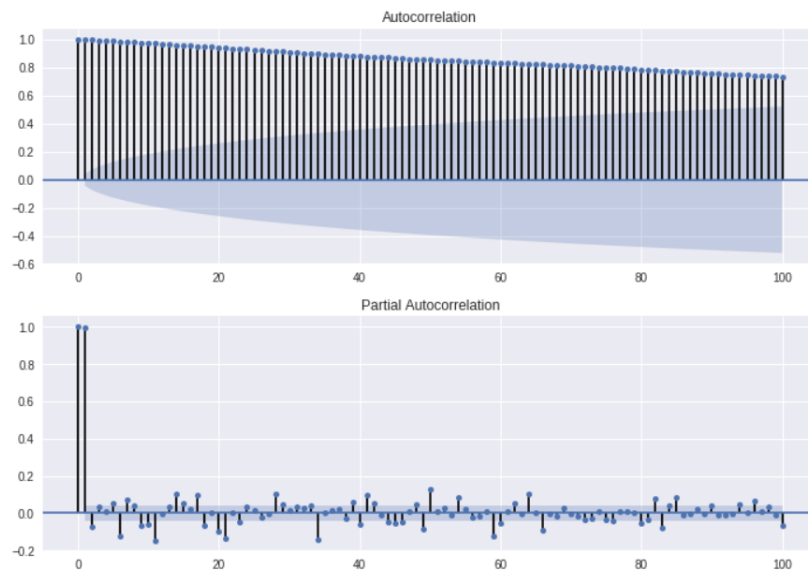


Fig.3. Plot ACF and PACF on training data



Fig.4. Plot Rolling Mean and Rolling Variance in 7 days of training data

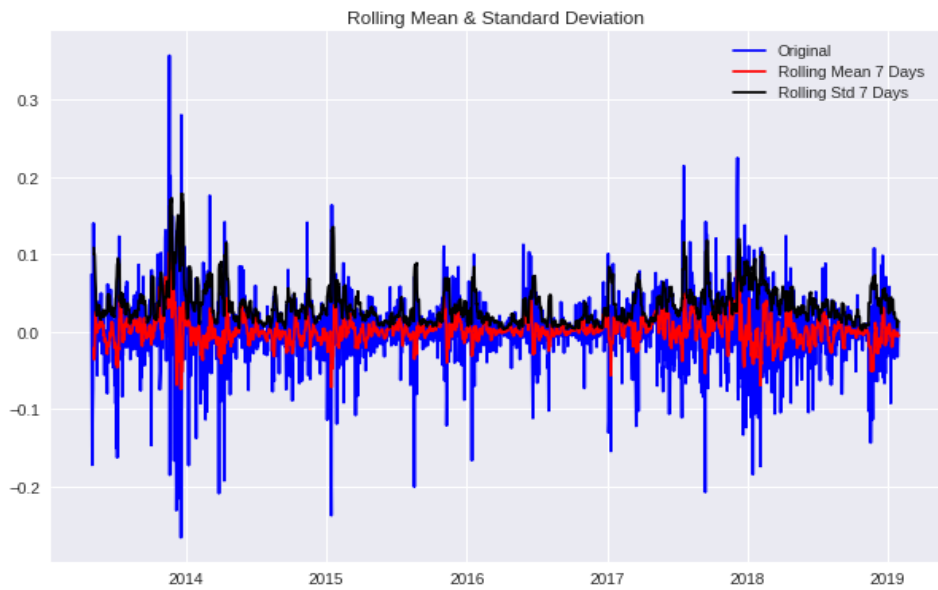


Fig.5. Plot Rolling Means and Rolling Variances in 7 days of differenced data

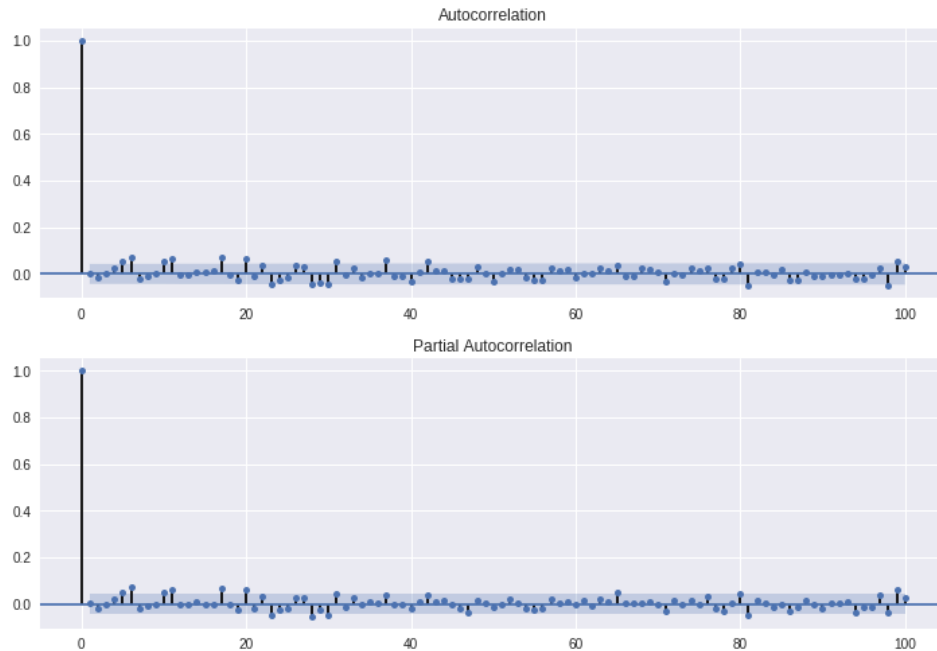


Fig.6. Plot ACF and PACF of differenced data

3.2LSTM Algorithm Procedure

- a) Training data will be normalize within range $[0,1]$ for easier computation.
- b) Train the model using normalized training data, training process was held with help Keras Syntax.
- c) Forecast using the obtained model.
- d) Denormalize the forecasting results.

3.3 α -Sutte Indicator Procedure

- a) Calling library `sutteForecastR`.
- b) Inputting training data on R program
- c) Running syntax's program from library `sutteForecastR` on inputted data.

4Result and Discussion

The accuracy of forecasting results from the three methods can be seen by comparing MSE, MAPE, and MAD criterion of each method. The higher the accuracy of forecasting, the lower the value of each criterion. These comparisons are shown in Table 1.

Table 1. Accuracy level comparison of each criterion

Method	MSE	MAPE	MAD
ARIMA(2, 1, 2)	2532.718	1.266 %	43.761
LSTM	*421.686	*0.592 %	*20.535
α -Sutte Indicator	1118.008	0.761 %	26.364

Table 1 show that LSTM algorithm has the best accuracy level in forecasting bitcoin's price, followed by α -Sutte Indicators and ARIMA models (2,1,2).The forecasting results of each methods and its plot are shown in Table 2 and Figure 7.

Table 2. Forecasting result of three methods with full length data

Date	Price	Predicted Price		
		(ARIMA)	(LSTM)	(α -Sutte indicator)
2019-01-31	3457.79	3485.713027	3437.243408	3454.286
2019-02-01	3487.95	3478.178058	3467.436768	3453.702
2019-02-02	3521.06	3479.668616	3500.587402	3501.379
2019-02-03	3464.01	3494.809816	3443.469727	3532.821
2019-02-04	3459.15	3511.329616	3438.604492	3466.334
2019-02-05	3466.36	3515.425258	3445.822021	3449.759
2019-02-06	3413.77	3508.963757	3393.181396	3448.286

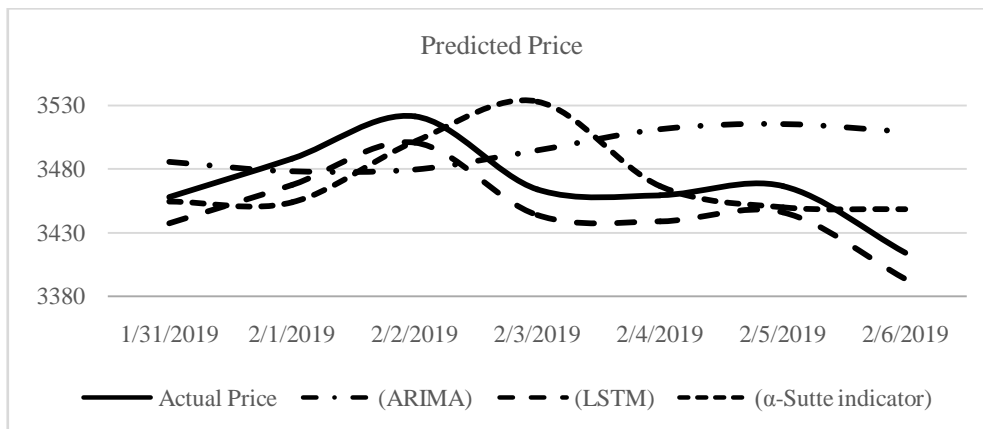


Fig.7. Plot of forecasting results from each method

As we can see from Figure 6, LSTM predicted prices are able to follow the real price movement with a slight variance, while for α -Sutte indicator, they are also trying to follow real price movement but with slightly off track on the 4th day, and finally for the ARIMA model, they are failed to follow real price movement completely.

As seen in Figure 1, bitcoin's price has high volatility level. In order to overcome the volatility problem, training data will be cut and only used data starting from July 1 2017 until January 30 2019. The same forecasting process was carried out. Plot of new training data length is shown in Figure 8.

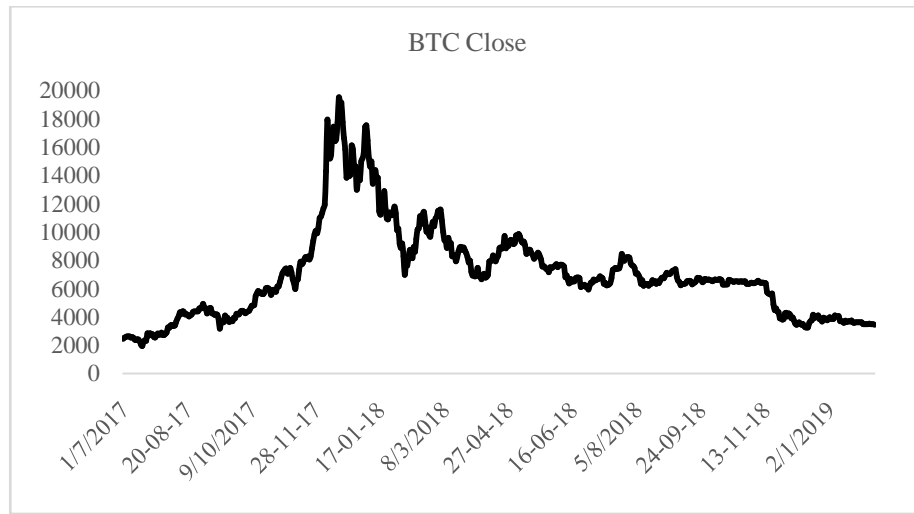


Fig.8. Plot of new length of Training Data

The accuracy of forecasting results from the three methods with new data can be seen in Table 3.

Table 3. Accuracy level comparison of each criterion with new length of Bitcoin's price data

Method	MSE	MAPE	MAD
ARIMA(3,1,2)	1755.761	1.006 %	34.705
LSTM	*5.549	*0.060 %	*2.077
Indicator α -Sutte	1118.008	0.761 %	26.364

Table 3 still shows that LSTM algorithm has the best accuracy level in forecasting bitcoin's price, also followed by α -Sutte Indicators and ARIMA model (3,1,2).

The forecasting results of each methods and its plot are shown in Table 4 and Figure 9.

Table 4. Forecasting result of three methods with new length of Bitcoin's price data

<i>Date</i>	<i>Price</i>	<i>Predicted Price</i>		
		<i>(ARIMA)</i>	<i>(LSTM)</i>	<i>(Indikator α-Sutte)</i>
2019-01-31	3457.79	3482.602722	3456.049561	3454.286
2019-02-01	3487.95	3491.978792	3485.100342	3453.702
2019-02-02	3521.06	3494.214721	3517.014893	3501.379
2019-02-03	3464.01	3492.224930	3462.039062	3532.821
2019-02-04	3459.15	3498.696435	3457.359131	3466.334
2019-02-05	3466.36	3499.281838	3464.302490	3449.759
2019-02-06	3413.77	3500.337364	3413.683350	3448.286

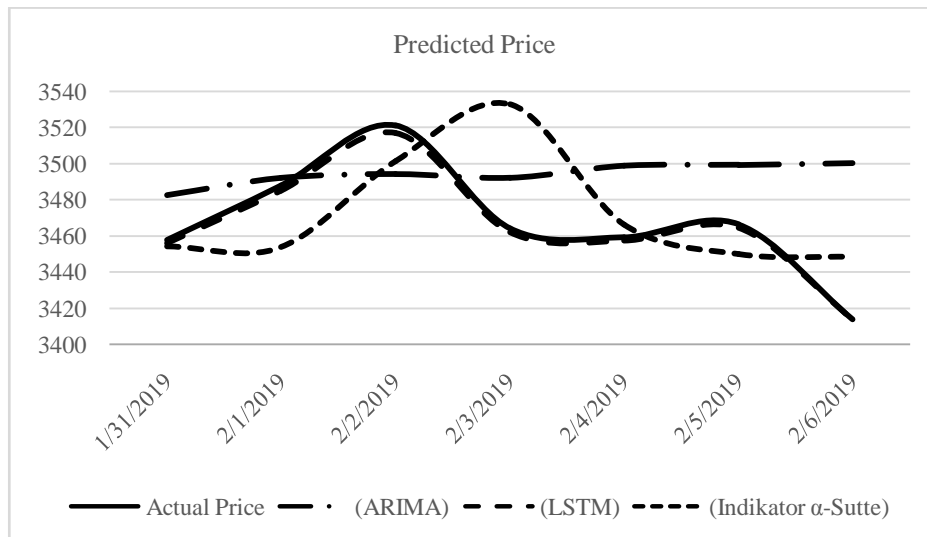


Figure 9. Plot of forecasting results from each method with new length of training data

As we can see at Figure 8, LSTM predicted price also able to follow real price's movement with much better variance than the last result, while α -Sutte indicator still has same forecast result as before; and for the ARIMA model, it is still failed to follow real price movement, but having slightly better accuracy than the last result.

5Conclusion

It is undeniable that the LSTM algorithm has a higher accuracy in predicting bitcoin's price compared to α -sutte indicator and ARIMA model. Either using full length data or after cutting length of data, LSTM algorithm still has higher accuracy than the other methods by

considering the MSE, MAPE, and MAD. This result indicates that the LSTM algorithm is more suitable for forecasting bitcoin's price compare other two methods. LSTM algorithm and ARIMA are sensitive to the length of data, while α -Sutte Indicator is not.

References

- [1] Box, G.E.P and Jenkins, G.M., *Time Series Analysis: Forecasting and Control*. Third Edition. Prentice Hall: New Jersey. (1994)
- [2] Hochreiter, S., and Schmidhuber, J., *Long Short-Term Memory*, Neural Computation. (1997)
- [3] Gers, F.A., Schmidhuber, J., and Cummins, F., Learning to Forget: Continual Prediction for LSTM, *Neural Computation*, 12(10),2451-2471, MIT Press, Doi: 10.1162/089976600300015015. (2000)
- [4] Ahmar, A. S., Rahman, A., and Mulbar, U., α -Sutte: A New Method for Timer Series Forecasting, *J.Phys. Conf. Ser.* 1040, 1 p.012018. (2017)
- [5] T. Fischer and Krauss C., "Deep learning with long short-term memory networks for financial market predictions," *European Journal of Operational Research*, vol. 270, no. 2, pp. 654-669. (2018)
- [6] Olah, C., "Understanding LSTM Networks," <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>, 2015.