Gaussian Processes for Bitcoin Prediction; Does forcing weekly periodicity improve the prediction?

Tinka Valentijn & Janne Holopainen
Aalto University
May 9, 2018

***Abstract*—Bitcoin is a popular cryptocurrency and hence there is a interest in predicting the future price. Among cryptocurrency enthusiasts there is a belief that the price of Bitcoin follows a weekly trend, where the price goes up during weekdays and drops during weekends. This research aims to discover if forcing such periodicity would yield evidence of the trend or if it would improve the prediction accuracy of the Bitcoin price. For predictions, Gaussian Processes are used. This method is chosen since it is well suited for time series prediction and can easily be forced to include a periodicity.**

We conclude that adding periodicity doesn't improve the prediction significantly. The results suggest that this is due to a lack of a strong periodicity in the data.

I. Introduction

Predicting time series is a popular problem in machine learning. One of those time series are stock prices. People have tried to solve the problem of predicting stock prices for the last 20 years using machine learning and it is still an active area of research [1], [2]. It appears to be a difficult problem, because of the unpredictability and the large number of factors that influence the stock prices. A new time series problem comparable to stock market prediction, is the prediction of the Bitcoin price. Bitcoin is a cryptocurrency that has been around for 10 years now, with a market cap of around 160 billion USD dollars [3]. The development of the price was different than most expected, so the question arises if we can predict the future price. Some people have tried to do this, but without too much success [4], [5]. Some even question if there is any predictive value in the historical data alone [5].

There is a belief among cryptocurrency enthusiasts that the price of Bitcoin follows a trend, where the price goes up during weekdays and drops during weekends. In this research we try to figure out if there is really such a trend and if using this periodic trend can be used small fluctuations of the Bitcoin price. In order to do so we use Gaussian Processes (GPs). This is a machine learning method which finds the best fitting gaussian distribution defined by mean and covariance matrix. This method is very suitable for time series prediction and thus might have potential for predicting the Bitcoin price. Moreover, this method enables you to define different kernels, and hence it is easy to place a weekly periodicity in here. To the best of our knowledge, GPs have never been used for Bitcoin prediction. They have however been used to predict stock-market prices [6].

Instead of trying to fully predict the Bitcoin price, our main research question is to see if forcing a periodic trend in the Gaussian Process improves the prediction.

II. Data analysis

As data we use the daily closing price of the Bitcoin. We retrieve the data from CoinDesk, which is a news outlet specialized in cryptocurrencies [7]. We have data from 2010 to 2018, where price ranges from \$0.05 to \$19343. We normalize all data by subtracting the mean from the data. This results in zero-mean centered data, which is assumed by GPs.

We experiment with different periods of data, since some have suggested that prediction by GPs works better when only recent data is used [6]. See Fig. 1 for a plot of the development of the Bitcoin price. As can be seen, the development is rather unique with a very steep spike, which can cause problems, when trying to find trends from the data. The data also suffers from heteroscedasticity, which can make GP convergence unstable.

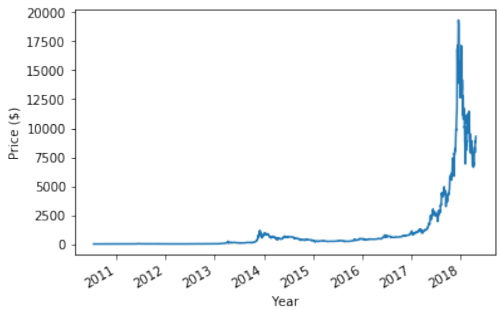


Fig. 1: Development of the Bitcoin price over time

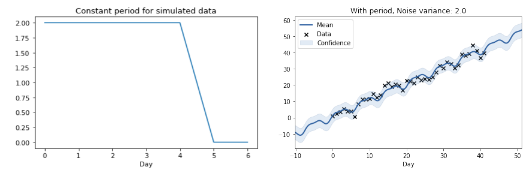


Fig. 2: Left: Constant period used in creation of the simulated data. Right: GP fit to simulated data - consisting of linear, periodic and noise components.

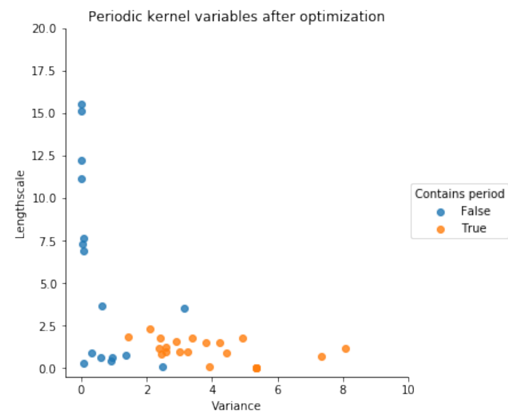


Fig. 3: Variables defining the periodic kernel after GPy optimized the fit for the data. Points with lengthscale < 0.1 excluded.

III. Experiments

For our experiments we used GPy, which is a Gaussian Process framework for Python [8]. Our hypothesis was, that if a weekly price-shift is present in the data - fitting a GP with a periodic kernel, with constant period of 7 days, would yield a reasonable variance and lengthscale to the periodic kernel. We assessed the fit of the GP by marginal log likelihood, mean log posterior predictive density (MLPPD) and mean squared error (MSE).

A. Discovering the period from simulated data

Before doing experiments on the Bitcoin data, we wanted to determine whether GP could reliably identify the underlying periodicity in data, with presence of varying amounts of noise. We set up an experiment with simulated data. This data consisted of linear, constant periodic and constant noise components. The simulated period followed the shape of our hypothesis, where business days have a value of 2, and weekends have a value of 0. See Fig.2 for the shape of the periodicity induced in the toy data.

In our test, we added different amounts of noise to the signal, and collected the parameter values of the kernel. Our goal was to see if the resulting parameters show a clear division between data with and without period. Fig.3 depicts the results where the variance and the lengthscale of the periodic kernel are plotted. From this we can see that the GP reliably assigns larger variance values and a smaller lengthscale to the periodic kernel, when the underlying signal contains a weekly period.

Based on the results, we concluded that GP should be able to identify the period from the price data, if such a period exists. We note however that, the simulated data does not fully represent the Bitcoin data, as the variance does not change over time and the tested signal is quite simple.

B. Discovering the period from actual data

We attempted to fit a GP to following datasets: whole data - normal scale, whole data - log-scale, partial data - log-scale. The partial data consists of the data points from the last year. We use 80% of the data for training and 20% for testing. After initial experiments, we discovered that a GP has a hard time finding a trend on the whole data in normal scale, because of the spike, and thus gave bad overall performance. To smooth out the affects of the price spike of the late 2017, we converted the data to log-scale. Before adding any periodicity, we first tried to find the best combination of kernels that fits the general trend. Our experiments showed that a linear plus RBF kernel captures this trend the best, see Fig. 4 for the fit of the whole data and one year data.

	Marginal Log Likelihood	MLPPD	MSE
One year, RBF+LIN	**348.78**	**-0.299**	0.0357
All data, RBF+LIN	2311.38	-2.602	7.709
One year, RBF+LIN+PERIODIC	350.60	-0.333	**0.0208**
All data, RBF+LIN+PERIODIC	2311.38	-2.602	7.709

TABLE I: Error measurements for the different experiments on Bitcoin data

After that we added periodicity. In most cases adding periodicity, doesn't change the results much. It predicts a bit differently, but doesn't seem to find a clear 7 day periodicity. The GP assigns either a rather large lengthscale or a small variance to the periodic kernel. In the case of a one year data period, adding the periodic kernel makes the MSE and MLPPD decrease a bit while the marginal log likelihood increases a bit. This indicates that the predictions get a little better, but the uncertainty becomes larger. See Table I for the error measurements of all experiments and Fig. 5 for the fit and prediction on the one year data with a periodic, RBF and linear kernel. In the case of the full data, there is no difference in the error measurements when adding a periodic kernel.

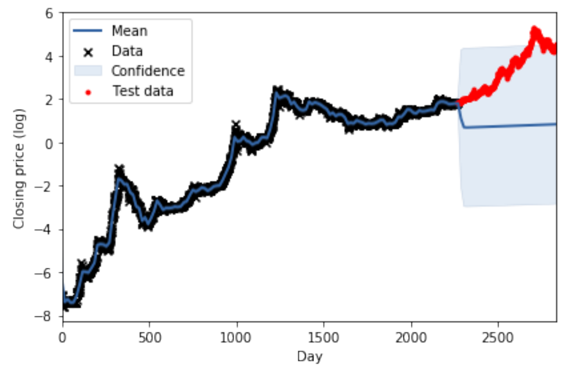
We did several other experiments, such as fixing the period of the kernel, adding several periodic kernels, and multiplying the kernel with a RBF or linear kernel. None of these made a significant difference in the results.

One interesting result is that the MLPPD and test error are lower when only using data from one year instead of the whole data. Apparently the more current data contains more information about future points and by using all data the GP cannot find a clear trend and goes back to zero. From a GP perspective it is reasonable that using only the current data gives better results, but it is contradictory to the common belief that more data is always better in machine learning.

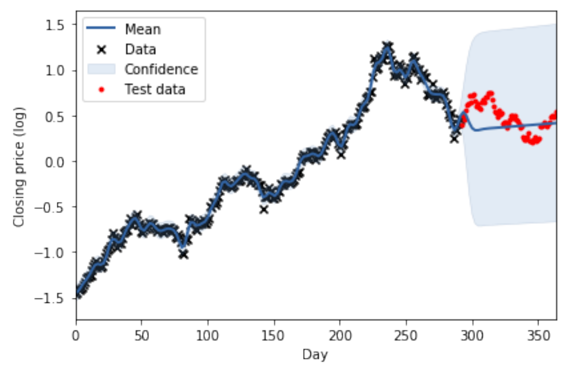
C. *Fourier analysis of the data*

Functions of time can be decomposed to periodic components. This kind of decomposition is called Fourier transform, and it can be used to analyze affects of different frequency components. We will refer to amplitude of a frequency - this represents the contribution of periodic component with given frequency to the original signal.

We performed Fourier transform to the Bitcoin data. Should the data contain a clear weekly period, we should be able to observe it from the frequency-amplitude plot. The plot is depicted in Fig.6, and frequency of 7 days, does not stand out.



(a) All data



(b) One year data

Fig. 4: The fitting and prediction of a Gaussian Process consisting of a RBF and linear kernel

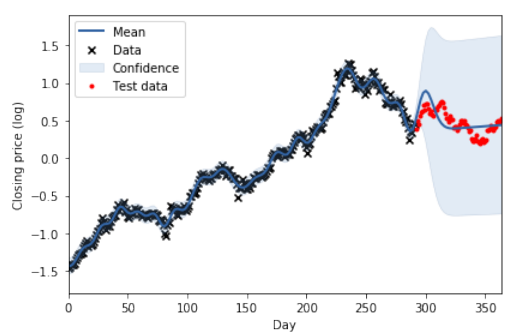


Fig. 5: The fitting and prediction of a Gaussian Process consisting of the sum of a RBF, linear and periodic kernel on data of a one year period.

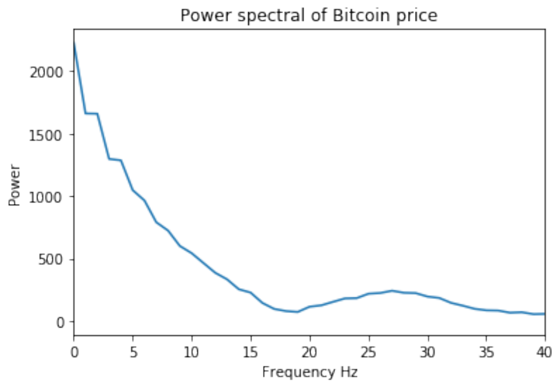


Fig. 6: Power/Amplitude spectral of the Bitcoin prices.

However, interestingly between 20 and 35 days, there is an increase in the amplitude.

IV. Conclusions

Based on our tests, we did not find any evidence to support the claim for weekly trend in the price. The reason can partly be because of the heteroscedasticity of the data, for which there are proposed approaches, like Wang et al. [9]. Alternatively pre-processing data to be percentage change between days, could also make the data better behaving, and subsequently yielding better results. Moreover, the prediction of the market price in general is a rather hard problem, containing a lot of variables which can cause unexpected changes in the price. Hence, it is possible that only by looking historical data, finding a periodic change in the data can be almost impossible.

However, when looking at the data by eye, no clear weekly trend could be discovered and hence it is not a very unlikely result, that the GP did not discover a weekly trend.

Our experiments with the simulated data indicate that GP should be able to identify a weekly trend in noisy data. Thus we conclude with reasonable confidence that Bitcoin price does not contain weekly period.

References

[1] I. Kaastra and M. Boyd, "Designing a neural network for forecasting financial and economic time series," *Neurocomputing*, vol. 10, no. 3, pp. 215–236, 1996.

[2] J. Patel, S. Shah, P. Thakkar, and K. Kotecha, "Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques," *Expert Systems with Applications*, vol. 42, no. 1, pp. 259 – 268, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0957417414004473

[3] Coinmarketcap, "Crypotocurrency price monitoring," accessed: 07.5.18. [Online]. Available: https://coinmarketcap.com/

[4] S. McNally, "Predicting the price of bitcoin using machine learning," Ph.D. dissertation, Dublin, National College of Ireland, 2016.

[5] A. Greaves and B. Au, "Using the bitcoin transaction graph to predict the price of bitcoin," *No Data*, 2015.

[6] M. T. Farrell and A. Correa, "Gaussian process regression models for predicting stock trends," *Relation*, vol. 10, p. 3414, 2007.

[7] Coindesk, "Bitcoin historical prices," accessed: 24.4.18. [Online]. Available: https://www.coindesk.com/price/

[8] GPy, "GPy: A gaussian process framework in python," http://github.com/SheffieldML/GPy, since 2012.

[9] C. Wang and R. M. Neal, "Gaussian Process Regression with Heteroscedastic or Non-Gaussian Residuals," *ArXiv e-prints*, Dec. 2012.