Sensors & Transducers, Vol. 244, Issue 5, September 2020, pp. 44-47

www.sensorsportal.com

Sensors & Transducers

Published by IFSA Publishing, S. L., 2020
http://www.sensorsportal.com

Cryptocurrency Price Prediction: A Machine Learning Approach

[1] Rohit Chivukula and [2,*] T. Jaya Lakshmi

[1] School of Computing and Engineering, University of Huddersfield, Huddersfield, United Kingdom
[2] Department of Computer Science and Engineering, SRM University, AP, Andhra Pradesh, India
[*] E-mail: jaya.phd.hcu@gmail.com

Received: 20 July 2020 /Accepted: 28 August 2020 /Published: 30 September 2020

Abstract: Cryptocurrency is used worldwide for digital payment or simply for investment purposes. Bitcoin price prediction is an interesting research problem in current scenario. In this paper, we have studied the application of machine learning approach in predicting the future price of bitcoin. Many dynamic factors effect Bitcoin prices and accurate predictions form strong base for investment decisions. In this study, we have collected the live data corresponding to bitcoin from quindle.com containing 8 features. Then we have compared the prediction performance of 11 regression algorithms. It is found that Lasso regression with a combination of generalized linear regression outperformed others with an improvement of 9 % accuracy over other regression algorithms.

Keywords: Machine Learning, Regression, Bitcoin price prediction.

1. Introduction

Bitcoin is a novel digital currency system which functions autonomously without any central governing authority. Bitcoin is used worldwide for digital payment and for investment purposes. Bitcoin is decentralized i.e. it is not owned by anyone. Bitcoin trades enable individuals to sell/purchase bitcoins utilizing various currencies. The largest Bitcoin exchange is Mt Gox [13].

Exchanges made by bitcoins are simple as they are not attached to any nation. The record of all the transactions, the timestamp data is stored in a specific kind of distributed ledger called Blockchain. Each record in a blockchain is called a block. Every block stores digital information pertaining to the transaction viz. date, time and amount involved in transaction. Every block is connected to its previous block to form blockchain. During transactions the user's identity is maintained private strictly, but only their wallet ID is made public. Payments are processed by a peer-to-peer network of users over the web. Bitcoins can be exchanged with other currencies in the exchange office, where all transactions are stored on the order book. The Bitcoin's value changes dynamically. The parameters affecting Bitcoin price prediction are different from those used in stock market price prediction. Therefore, it is necessary to predict the value of Bitcoin so that correct investment decisions can be made. Bitcoin price can be predicted efficiently using Machine learning algorithms.

In this work, we focus on the short-term price prediction of bitcoin from machine learning perspective.

2. Literature

Given a training set $\{(X_1,y_1), (X_2,y_2)\ldots(X_m, y_m)\}$, where each $X_i \in R^n$ representing the input space and $y_i \in R$ denoting the target value, the aim of the regression problem is to fit a function that can approximate the value of y for an X not in the training set. Bitcoin price prediction can use this model efficiently in which X_i representing the vector of features affecting price of bitcoin, based on which the future price y_i can be predicted.

Ciaian, *et al.* [1] studied the conventional as well as specific factors that contribute to the future Bitcoin price formation such as market forces and Bitcoin attractiveness for investors and users. Their framework is rooted on the popular Barro model. The authors extract daily data for a period of five years and apply time-series based mechanisms. McNally, *et al*. [2] have quantified the accuracy by considering Bitcoin Price Index using Bayesian recurrent neural network (RNN) [15] and Long Short-Term Memory (LSTM) network [11]. Another model based on deep learning found in literature is ARIMA model. The non-linear deep learning methods outperform the ARIMA. Madan, *et al.* [4] used machine-learning approach to predict Bitcoin price. The authors predict the sign of the daily price change using the data with 25 features corresponding to the Bitcoin price as well as payment network for a duration of five years. Further, the authors focus on the Bitcoin price data alone during different time durations. Katsiampa, *et al*. [3] have studied optimal conditional hetero-skedasticity model with regards to goodness-of-fit to Bitcoin price data. They found that the AR-CGARCH model emphasized the need of both short-run as well as long-run component of the conditional variance in price prediction.

Dyhrberg, *et al*. have explored the hedging capabilities of bitcoin [5]. Their work applies asymmetric GARCH methodology. The outcomes show that bitcoin can be utilized as a support against stocks in the Financial Times Stock Exchange Index. In their investigation, they have found that Bitcoin and gold stocking have similar hedging abilities. Some more works on bitcoin price prediction can be found in [6-8] and [9]. Alessandretti, *et al*. [16] use XGBoost [14] as well as LSTM [11] for prediction and achieve better prediction performance. A sophisticated analysis addressing the problem has been given in [17]. The summary of literature is tabulated in Table 1.

Table 1. Summary of literature.

Name/Author/Year	Techniques Used	Limitations
Predicting the price of Bitcoin using machine learning, S. McNally, 2016	Price prediction using recurrent neural networks (RNNs) and long short-term memory (LSTM)	A machine trained only with Bitcoin price index and transformed prices exhibits poor predictive performance
The economics of Bitcoin price formation, P. Ciaian, M. Rajcaniova, D. Kancs, 2016	Linear model by considering related information that is categorized into several factors of market forces, attractiveness for investors, and global macro-financial factors	They assume that market forces and attractiveness for investors influence Bitcoin prices but with variation over time
Automated Bitcoin trading via machine learning algorithms, I. Madan, S. Saluja, A. Zhao, Dept. Comput. Sci., Stanford Univ., Stanford, CA, USA, Tech. Rep., 2015	Price prediction problem as a binary classification task	Does not explore or disclose the relationship between Bitcoin price and other features in the space, such as market capitalization
Volatility estimation for Bitcoin: A comparison of GARCH models, P. Katsiampa, Econ. Lett., Vol. 158, Sep. 2017, pp. 3–6	Optimal conditional hetero skedasticity model with regards to goodness-of-fit to Bitcoin price data	Considers only volatility and not price prediction
Hedging capabilities of Bitcoin. Is it the virtual gold ?, A. H. Dyhrberg, Finance Res. Lett., Vol. 16, Feb. 2016, pp. 139-144	Hedging capabilities of bitcoin by applying the asymmetric GARCH methodology used in investigation of gold	Only hedging is discussed, no price prediction is done

3. Proposed Approach

The proposed approach acquires time-series data recorded daily for five certain time period at different time instances, and normalizes and smoothens. Then from the preprocessed data, the features and parameters are extracted. The accuracy is compared with different models after the final prediction.

The execution is carried out in the following steps.

3.1. Dataset Collection and Preprocessing

As Bitcoin is a kind of stock traded in stock market, dataset is available in plenty with all time intervals. We have collected live data from quandl.com during the period 2011 to till date. This provided us the most comprehensive bitcoin price in date wise data. Dataset is extracted to CSV file. The features Timestamp, Open, High, Low, Close, Volume_btc, Volume_currency, Weighted_price are extracted and used in the prediction task.

3.2. Split Dataset as Train and Test Set

10-fold cross validation is used for evaluating the prediction performance. 80 % of the of data is taken as training input for constructing the machine learning model and the remaining 20 % of data is considered as test set for performance prediction. The process is repeated for 10 times and the average of 10 iterations are reported.

3.3. Machine Learning Algorithms used to Predict Price of Bitcoin

The standard regression techniques of Linear regression (LR), Logistic regression (LGR), Support Vector Machine (SVM), Multivariate Regression and Multiple Linear Regression (MLR), Partial Least Squares (PLS), Support Vector Regression (SVR), Back-Propagation Neural Network (BPNN), K Nearest Neighbors (kNN), Decision Trees (DT), Gradient Boosted Machine (GBM), Random Forest (RF) and LOSSO are applied on train data and evaluated on test set. The Lasso Regression algorithm introduces the L1 regularization term that plays a vital role in feature selection. The Grid Search logic is applied in this model for hyperparameter tuning in order to achieve optimum performance. A maximum of 100000 iterations are computed using the LASSO regression model. The model performed well with high accuracy.

3.4. Evaluation Measures

It is a common practice to evaluate the prediction performance of a regression problem using MAE, RMSE and Accuracy [10, 12]. Let y and ŷ be vector of actual and predicted values and n be the number of samples. The measures are given below.

3.4.1. RMSE

Root Mean Square Error is normalized distance between predicted and actual values.

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{j=1}^{n}(y_j - \hat{y}_j)^2} \quad (1)$$

3.4.2. MAE

Mean Absolute Error is the average of the absolute difference between the predicted values and observed value.

$$\mathrm{MAE} = \frac{1}{n}\sum_{j=1}^{n}|y_j - \hat{y}_j| \quad (2)$$

3.4.3. Accuracy

Accuracy is another evaluation measure to evaluate the performance. Accuracy is defined as follows:

$$Accuracy = \frac{Number\ of\ Correct\ predictions}{Total\ number\ of\ predictions\ made} \quad (3)$$

The proposed approach is shown in the Fig. 1.

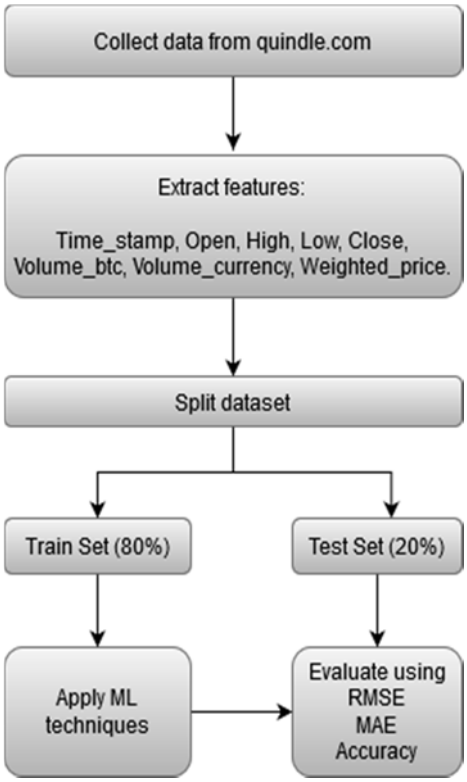


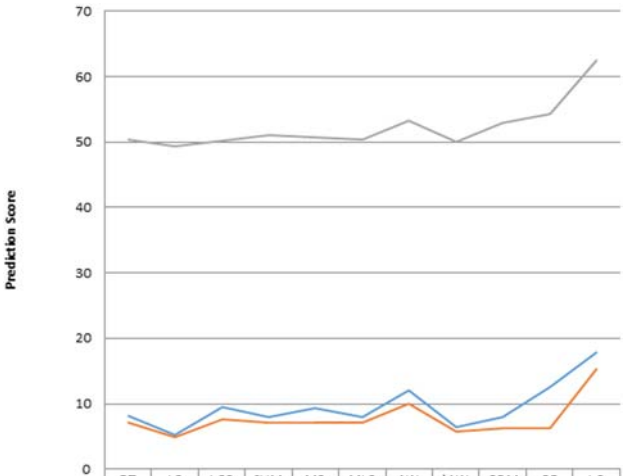
Fig. 1. Proposed Approach.

4. Results

The proposed work is implemented in Python 3.6.4 with libraries scikit-learn, pandas, matplotlib and other mandatory libraries. We have downloaded dataset from quandl.com with necessary authentication keys. The data downloaded contains up-to date data. The dataset is 80 % considered as train set and 20 % considered as test set. Five days forecast price prediction is done. The experimental results are shown in Fig. 2. It is observed that all the regression algorithms produced a minimum accuracy of 50 %. LASSO in combination with GLM has given highest prediction accuracy with an improvement in terms of RMSE, MAE as well as accuracy. Neural Network based algorithms performed well over other methods.

5. Conclusion

Bitcoin is a popular cryptocurrency, and it has been studied in depth in financial and computer science fields. In this work, we analyze the time series of Bitcoin price with various regression models for forecasting price for five days. The experimental results show that Lasso regression in combination with generalized linear regression outperforms the other by high accuracy on price prediction.



	DT	LR	LGR	SVM	MR	MLR	NN	kNN	GBM	RF	LG
RMSE	8,12	5,23	9,54	8,03	9,33	7,98	12,07	6,34	7,98	12,54	17,84
MAE	7,16	4,85	7,55	7,08	7,13	7,06	9,98	5,78	6,22	6,3	15,34
Accuracy	50,32	49,34	50,23	51,02	50,7	50,33	53,24	50,02	53	54,23	62,43

Fig. 2. Results of Crypto currency Price Prediction.

References

[1]. Ciaian Pavel, Miroslava Rajcaniova, d'Artis Kancs, The economics of BitCoin price formation, *Applied Economics,* Vol. 48, Issue 19, 2016, pp. 1799-1815.

[2]. McNally Sean, Jason Roche, Simon Caton, Predicting the price of bitcoin using machine learning, in *Proceedings of the IEEE 26th Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP)*, 2018, pp. 339-343.

[3]. Katsiampa Paraskevi, Volatility estimation for Bitcoin: A comparison of GARCH models, *Economics Letters,* Vol. 158, Issue C, 2017, pp. 3-6.

[4]. Madan Isaac, Shaurya Saluja, Aojia Zhao, Automated bitcoin trading via machine learning algorithms, *URL:* http://cs229. stanford. edu/proj2014/Isaac% 20Madan 20, 2015.

[5]. A. H. Dyhrberg, Hedging capabilities of Bitcoin. Is it the virtual gold?, *Finance Research Letters*, Vol. 16, Issue C, Feb. 2016, pp. 139-144.

[6]. D. Shah, K. Zhang, Bayesian regression and Bitcoin, in *Proceedings of the 52nd Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 2015, pp. 409-415.

[7]. Jang Huisu, Jaewook Lee, An empirical study on modeling and prediction of bitcoin prices with Bayesian neural networks based on blockchain information**,** *IEEE Access,* Vol. 6, 2017, pp. 5427-5437.

[8]. F. Andrade de Oliveira, L. Enrique Zãrate, M. de Azevedo Reis, C. Neri Nobre, The use of artificial neural networks in the analysis and prediction of stock prices, in *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, 2011, pp. 2151-2155.

[9]. M. Daniela, A. Butoi, Data mining on Romanian stock market using neural networks for price prediction, *Informatica Economica*, Vol. 17, Issue 3, 2013, pp. 125-136.

[10]. Willmott, Cort J., et al., Statistics for the evaluation and comparison of models, *Journal of Geophysical Research: Oceans* 90.C5, 1985, pp. 8995-9005.

[11]. Xingjian, S. H. I., et al., Convolutional LSTM network: A machine learning approach for precipitation nowcasting, *Advances in neural information processing systems*. 2015, pp. 802-810.

[12]. Chai Tianfeng, Roland R. Draxler, Root mean square error (RMSE) or mean absolute error (MAE)?–Arguments against avoiding RMSE in the literature, *Geoscientific Model Development,* Vol. 7, 2014, pp. 1247-1250.

[13]. Emily M. Little, Bitcoin, *The Investment Lawyer*, 21, 5, 2014, pp. 22-27.

[14]. Chen Tianqi, Carlos Guestrin, XGBoost: A scalable tree boosting system, in *Proceedings of the 22nd ACM Sigkdd International Conference on Knowledge Discovery and Data Mining*, 2016.

[15]. Lipton Zachary C., John Berkowitz, Charles Elkan, A critical review of recurrent neural networks for sequence learning, arXiv preprint arXiv:1506.00019, 2015.

[16]. Alessandretti Laura, *et al*., Anticipating cryptocurrency prices using machine learning, *Complexity,* Vol. 2018, 2018, pp. 1-16.

[17]. Abhishek Muvvala, Rohit Chivukula, T. Jaya Lakshmi, Comparative Analysis of Various Machine Learning Approaches for Bitcoin Price Prediction, in *Proceedings of the 2nd International Conference on Advances in Signal Processing and Artificial Intelligence (ASPAI'2020)*, Berlin, Germany, 2020, pp. 155-159.