International Journal of Future Generation Communication and Networking
Vol. 13, No. 1, (2020), pp.745-752

USING NONLINEAR MACHINE LEARNING ALGORITHMS TO PREDICT THE PRICE OF CRYPTOCURRENCIES

Saad Ali Alahmari
Department of Computer Science, Shaqra University
Saudi Arabia

Abstract

The potential growing for profit in virtual currency has made the prediction of cryptocurrency's price an appealing research topic. Numerous research has already been conducted to predict future prices of a specific virtual currency using a machine-learning model. However, very few involved many cryptocurrency using various machine-learning model in their studies. This study applies a three non-linear algorithm: Decision Tree Regressor, (DTR) and the K-Nearest Neighbor (KNN) models for forecasting the three big cryptocurrency prices: Bitcoin, XRP and Ethereum using bivariate time series method where the cryptocurrency (daily-Closed Price) is the continuous dependent variable and the Morgan Stanley Capital International (MSCI) All Country World Index (MSCI-ACWI)-(daily-Closed Price) is the predictor variables. The results demonstrate that (DTR) outperforms the K-Nearest Neighbor (KNN) in terms of Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R-Squared (R^2).

1. Introduction

The paper concerns novel techniques. The prediction of cryptocurrency prices has become an important research because of the rapid volatility and the significant profit for investors. Cryptocurrencies are digital currencies, which are exchanged anonymously over a decentralized network. Over 2000 cryptocurrencies are in use, most of which use the Block-chain technology. This analysis focussed on the three big cryptocurrencies, Bitcoin, XRP and Ethereum, representing the largest market capitalization share. In 2009 Satoshi Nakmoto launched Bitcoin, the first crypto-currency. It uses Block-chain technology as its transaction platform. XRP, the Ripple payment system's virtual currency, was developed in 2013 as an open source Internet protocol. Ethereum is a fully autonomous, decentralized system with considerable capabilities as a whole networkLike Bitcoin and Ethereum, XRP is not using BlockChain technology because it has its own Ripple Protocol Consensus Algorithm (RPCA) technology.

Several research papers have experimented various financial and economic indices to predict Cryptocurrencies prices. The previous research focus on the study of specific indices such as Cryptocurrencies market only [1, 2, 3, 4], Cryptocurrencies market and commodity markets [5], Cryptocurrencies market and relative technologies of block-chains [6, 7]. These specific indices are not limited to specific market and factors. In this analysis, in addition to market data from Cryptocurrencies, we are analysing the MSCI-ACWI international indicator of equity market performance worldwide maintained by MSCI. The MSCI- is a model index strategy capturing sources of equity returns in 49 stock markets [8].

Predictions of the time series are used in different fields, including economics, education, and marketing. Nonetheless, the key challenge facing researchers when forecasting cryptocurrencies prices is the extreme price uncertainty [3, 5, 9]. In using three non-linear algorithms: Decision Tree Regressor, (DTR) and the K-Nearest Neighbor (KNN) are performed, researchers have found that non-linear algorithms outperform linear algorithms [1]. These models apply bivariate time series method where we experiment the relationship between several crypto-currencies (daily-Closed Price) and the MSCI ACWI Indexes (daily-Closed Price).

ISSN: 2233-7857 IJFGCN

The paper is organized as follows: Section 2 discusses previous research related work on the estimation of cryptocurrencies prices. Section 3 describes the theoretical context of the models DTR, KNN and SVR. Section 4 deals with the experimental assessment, presets methods for data collection and extraction of data features. The experimental results are discussed in Section 5. Section 6 deals with findings and observations, suggesting drawbacks to this method and recommendations for future research.

2. Related Work

Several researchers have tried to predict the price of cryptocurrencies using variety of statistical and machine learning algorithms. Most of the research were mainly performed in statistical analysis and limited number of non-linear and linear algorithms. However, Researchers have found that non-linear algorithms outperform linear algorithms [1]. In this research, three non-linear algorithms: Decision Tree Regressor, KNN and SVR are investigated and compared. These algorithms apply bivariate time series method where crypto-currencies (daily-Closed Price) and the MSCI ACWI Indexes (daily-Closed Price) are continuous. We decided to study the dataset of three cryptocurrencies (Bitcoin, XRP and Ethereum) with the highest market capitalization value compared to the large MSCI ACWI indexes.

Reference [1] used nonlinear regression, Neural Networks (NNs), and Classification and Regression Tree (CART) models to compare prediction accuracy in three ways. The results of the research could not show obvious distinctions between NNs and CART models, however they outperform non-linear regression model in the prediction accuracy. Reference [2] performed a prediction algorithm of K-Nearest Neighbor (KNN) based on a non-parametric regression model. Reference [9] used ARIMA and seq2seq recurrent deep multi-layer neural network (seq2seq) to forecast Bitcoin pricing, but their models showed that recurrent neural networks (RNNs) only outperformed ARIMA with additional input sources over the long term. Additionally, the significant volatility in the Bitcoin datasets over the study period has resulted in varied results.

Reference [3] uses a number of statistical tests (e.g. exponential weighted moving average (EMWA)) and machine learning models (e.g. random forests and Gaussian processes) to analyze the short-term estimation of volatility for Bitcoin and US dollars using an hourly time series. We found that among other prediction models, the extreme gradient boosting (XGT) and the elastic-net (ENET) achieved the highest precision. Reference [6] centred on the Bitcoin method, using the Bayesian neural networks (BNNs) to consider Blockchain and macroeconomic features. To predict the Bitcoin pricing process Reference [10] applied a Bayesian optimized RNN and a long short-term memory (LSTM) network. The results showed that the LSTM network had a higher 52 percent classification accuracy and a lower 8 percent regression error. They then compared findings without regressors to that of the ARIMA model. Reference [11] used network Blockchain features combined with simple Bitcoin features and included multiple regression and classification models (linear, logistic, SVM, and neural network). Although with the neural network classification, the best accuracy results obtained were 55 percent, the authors found that using the network-based features had no significant effect on Bitcoin price predictions.

Some research has investigated the extensions and companied of traditional machine learning models. Reference [12] examines the estimation of three classical currencies ' volatility against three cryptocurrencies by integrating conventional Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model with machine learning support vector regression (SVR). Reference [12, 13] evaluate the Support Vector Regression GARCH against GARCH, EGARCH and GJR models based on volatility forecast, they find SVR- GARCH yields to better accuracy in the volatility forecast. Reference [14] proposes Binary Auto Regressive Tree (BART) which is a

ISSN: 2233-7857 IJFGCN

combination of classic algorithm classification and regression trees (C&RT) and ARIMA autoregressive models. The research results in better RMSE relative to classic ARIMA-ARFIMA variants. Reference [5] experiments multiple models of Regression and Deep Learning using Bitcoin trading data for a period of six years at 1-minute intervals. Those models are Theil-Sen Regression and Huber Regression, LSTM and Gated Recurrent Unit (GRU). The GRU model produces MSE's best results at 0.00002 and R2's at 0.992, followed by LSTM model results. The author of [15] uses a classic ARIMA model to predict the course of future prices of major cryptocurrencies in different time-scales.

It should be noted that preceding work has investigated the prediction of various cryptocurrencies using different models with different features from different domains. However, there are limited focus on applying and comparing the non-linear algorithms: Decision Tree Regressor, KNN regression and SVR on such regression problem of price prediction of cryptocurrencies with features of MSCI ACWI Indexes (daily-Closed Price). This work considered three models of DTR, KNN regression and SVR equipped with cryptocurrencies datasets (Bitcoin, XRP and Ethereum), based on training the data set on daily time series to predict future cryptocurrencies prices (Bitcoin, XRP, and Ethereum).

3. Approaches Employed

In this section, we briefly describe various regressions algorithms used for predicting. Regression research focuses on predicting the outcome of a dependent variable given a set of independent variables.

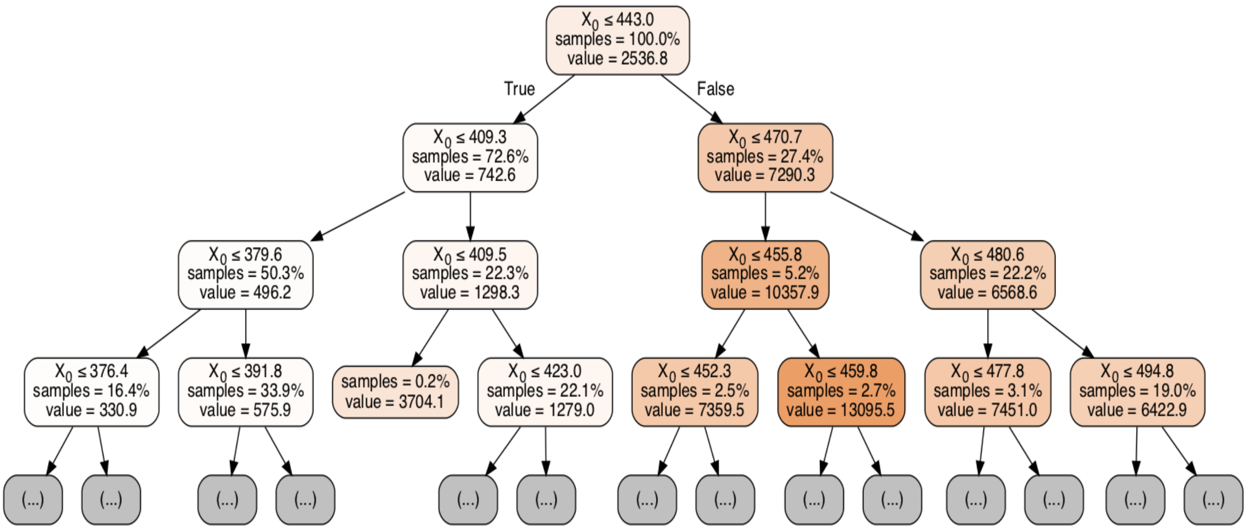

Figure SEQ Figure * ARABIC 1 Decision Tree Structure of the Bitcoin Cryptocurrency

3.1 Decision trees regression

The DTR is based on binary Recursive Partitioning (RP) algorithm [16]. It splits iteratively the data into partitions using continuous values of the predictors to certain thresholds. Decision tree regression notices selected features and trains a model based on the tree structure rules to predict target variables (Cryptocurrency Prices). The (RP) algorithm works as follow:

1. 1. Selecting the rules on splitting. (Example of the minimum weighted fraction of the total weights needed for a leaf node).
2. Determining the termination criterion for tree nodes (such as maximum depth of the leaf node).
3. Assigning value to each terminal node.

ISSN: 2233-7857 IJFGCN

Fig. 1 shows a decision tree structure of the Bitcoin training dataset. We construct the decision trees in a top-down approach using the training dataset, the “Close Price” as main feature which is found to be the most high- importance feature. Thus, the entire training dataset is spitted subsequently to further subdivided around the target feature of “Close Price” until reaches the stopping rules which is the defined maximum depth. The maximum depth parameter controls at a leaf node to prevent overfitting. The parameter of maximum depth is assigned to ‘3’ in fig. 1.

3.2 K-Nearest-Neighbors (KNN)

The KNN is a straightforward to implement machine learning classification and regression problems [17]. KNN is based on similarity function and finds the K most similar instances in the training dataset for a new data instance. A mean or median target variable is taken as prediction from the K neighbors. The K-NN regression works as follow:

1. Initialize K number of neighbors (we assumed k=100).
2. Compute K of the nearest neighbors.
3. Calculate the distance between the test samples and the training samples.
4. Sort the training data based on distances.
5. Find the labels of the selected K entries and assign it as prediction value.

4. Experimental Evaluation

In this section, we describe the data collection method. It shows how the features of the dataset extracted and applied analysis.

4.1 Data collection

The cryptocurrency datasets used in this analysis included 200 digital cryptocurrencies and were obtained from the CoinMarketcap[19] online data source[18] with notifications derived from it. A subset was extracted for the three cryptocurrencies (Bitcoin, XRP, and Ethereum), as the focus of the research was on them. The time series chosen for the data sets differed between the three cryptocurrencies. Unlike the conventional monetary markets, during these selected times there were dramatic changes in the price of the cryptocurrencies. Through dataset row represented regular market information for the worldwide cryptocurrencies and defined eight features: date, currency, open, high, low close, volume and market capital. The "Near" function was specifically selected as a dependent variable because it indicated the day's closing price and the next day's open price.
The original dataset included 84,080 submissions. After the sub-datasets relating to the three cryptocurrencies (Bitcoin, XRP, and Ethereum) were removed, the number of observations was reduced to 5,251.

4.2 Feature extraction

The researchers noted major variances in comparing mean values for cryptocurrency data sets and proposed reducing variances by using different time frequencies, such as hourly[8]. Nonetheless, choosing a particular frequency of the time series (e.g. daily, weekly, monthly) observation is a feature-engineering activity that can lead to different insights. The datasets were down-sampled in daily base for this study in order to obtain greater insights. The model was trained and the "Close Price" for each of the three cryptocurrencies (Bitcoin, XRP and Ethereum) was extracted within a particular date as follows:

1. Bitcoin:(28/04/2013-To-24/06/2019).

ISSN: 2233-7857 IJFGCN

2. XRP :(04/08/2013-To- 24/06/2019).
3. Ethereum :(07/08/2015-To-24/06/2019).

The data sets plotted in Figure (2) for the selected time series reflect the historical datasets for Bitcoin, XRP, and Ethereum. The enormous volatility in Bitcoin prices can be clearly seen, beginning in mid-2017 and extending through the end of the yearThis rise has positively affected the prices for XRP and Ethereum, but these rises are difficult to see because the improvements are very small compared to the Bitcoin prices, only hitting \$3.38 and \$1396.42, respectively.

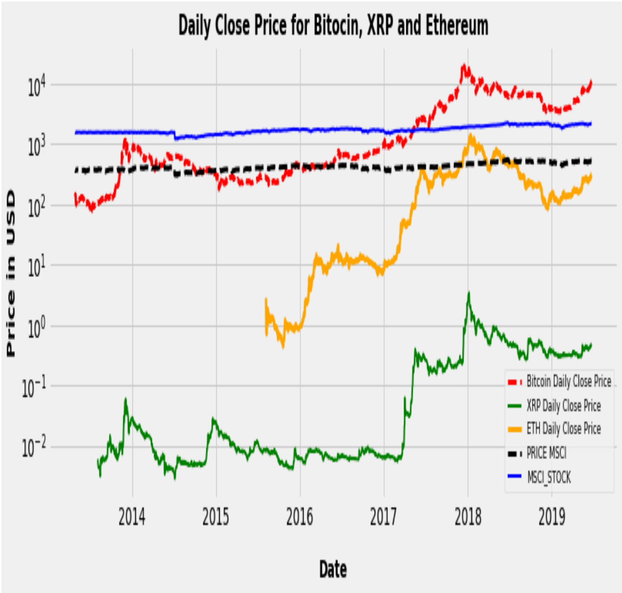

Figure 2 Historical datasets of Bitcoin, XRP and Ethereum

Table (1) lists descriptive statistics for those data sets. It should be remembered that there are significant variances in the prices of the cryptocurrencies examined. Nonetheless, significant price spikes in major cryptocurrencies in the third and fourth quarters of 2017 have greatly distorted the datasets.

Table 1 Summary Statistics

Statistics	Bitcoin	XRP	Ethereum
#Observ.	2249	2151	1418
mean	2545.61	0.18	205.842
std	3425.18	0.33	258.46
min	68.43	0.002	0.434
max	19497.4	3.380	1396.42

Depending on the date feature the datasets were partitioned; historical data was used to train and the latest data was used to check. The partitioned datasets were then divided into a training set of 80% and a test set of 20 per cent. For prediction precision, refer to the "(5), (6), (7), (8)" measures: mean absolute error (MAE), mean square error (MSE), root mean square error (RMSE), and the determination-R-squared (R^2) coefficient. They are defined as follows:

ISSN: 2233-7857 IJFGCN

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N} |d_i - z_i| \qquad (5)$$

$$\text{MSE} = \frac{1}{N}\sum_{i=1}^{N} (d_i - z_i)^2 \qquad (6)$$

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} (d_i - z_i)^2} \qquad (7)$$

$$R^2 = 1 - \frac{\sum (d_i - z_i)^2}{\sum (d_i - \underline{z_i})^2} \qquad (8)$$

Where $\boldsymbol{N}$ denotes the total number of samples forecasted, $\boldsymbol{d_i}$ denotes the actual value of the sample, $\boldsymbol{z_i}$ denotes the forecasting value of the sample and $\underline{\boldsymbol{z_i}}$ refers to the mean value of the sample.

4.3 Model analysis

We used the best random split in the Decision Tree Regressor algorithm which it results in unbalanced training datasets and makes probably biased dataset. There was a class domination for low cryptocurrency prices which represents most values comparing to the short-period of dramatic increase of prices, from fourth quarter of 2017 to the second quart of 2018. The initial split divides the decision tree training dataset into 20 percent and 80 percent which misrepresents data set classification. We use the weight-based pre-pruning criterion to void the biased datasets and determine the minimum weighted fraction of the total weights needed to be at a leaf node. As result of optimization without impact the accuracy, the split of the training datasets becomes 58% and 42% respectively. After several experiments using the KNN model, we found that k=10 is the optimal based on the outcomes of MAE, MSE RMSE and $\boldsymbol{R^2}$. The selected K value leads to the best performance.

5. Results and discussions

This section describes the results of using the non-linear algorithms: decision Tree Regressor, KNN and SVR to forecast prices for the cryptocurrencies Bitcoin, XRP and Ethereum. The experimental results of the training failures and evaluation errors are summarized in Table (2).

We found that using the KKN model outperforms DTR models in terms of MAE, MSE, RMSE, and R^2 to forecast Bitcoin, XRP, and Ethereum cryptocurrency prices with high degree of errors and variances. We find KNN outperforming the DTR and SVR with low-dimensional space in those datasets.

Fig. 3-5(a) and Fig.3-5(b) gives the error graph of the actual MSCI-ACWI and MSCI Stock values (on the x-axis) against the predicted Cryptocurrency Closed Price values, in US dollars; (on the y-axis). It is observed from table (2) that the DTR yields better prediction performance than the KNN model. The DTR will effectively project as a figure in Cryptocurrency Closed Value. 3(a), 4(a) and 5(a) provide a stronger link between the values predicted and the real values. The detailed Fig. 3, 4 and 5 are shown in Appendix A. \

ISSN: 2233-7857 IJFGCN

Test	Bitcoin		XRP		Ethereum	
	DT	KNN	DT	KNN	DT	KNN
MAE	612.43	670	0.056	0.07	75.9	89.4
MSE	1601417	1751049	0.027	0.03	24863	21629.9
RMSE	1265	1323	0.164	0.175	157.7	147
R^2	**0.92**	**0.89**	**0.98**	**0.86**	**0.95**	**0.82**

6. Conclusion

This research paper provides compares and assesses tow machine-learning methods (DTR and KNN) for predicting cryptocurrency prices using the datasets of three-major cryptocurrencies: Bitcoin, XRP and Ethereum. The dependent variable is time series continuous and predictor variables are all continuous. Many experimental studies have been performed in the past, however, very few compared more than one model such as DRT and KNN with several cryptocurrencies. It is noteworthy that the Decision Tree Regressor (DTR) outperforms the K-Nearest Neighbor (KNN) in terms of MAE, MSE, RMSE and R^2. Although the dramatic changes in prices of cryptocurrencies price, as well as the variance of space problem of the selected datasets, the KNN model leads to better performance. This might conclude the selected features are relevant and they present a low dimensional space.

In terms of future research, focusing on advanced instance-based algorithms such as Learning Vector Quantization (LVQ) and Self-Organizing Map (SOM) and comparing their result with this research results should present accurately best match instance-based algorithm.

Indeed, substantial datasets using non-economical features are required for further evaluation of the selected method.

References

1. Razi, M.A. and Athappilly, K., A comparative predictive analysis of neural networks (NNs), nonlinear regression and classification and regression tree (CART) models. Expert Systems with Applications, 29(1), 2005 , pp.65-74.
2. Razi, M.A. and Athappilly, K., A comparative predictive analysis of neural networks (NNs), nonlinear regression and classification and regression tree (CART) models. Expert Systems with Applications, 29(1), 2005 , pp.65-74.
3. Wang, S., Li, R. and Guo, M., Application of nonparametric regression in predicting traffic incident duration. Transport, 33(1), 2018. pp.22-31.
4. Catania, Leopoldo, Stefano Grassi, and Francesco Ravazzolo. "Forecasting cryptocurrencies under model and parameter instability." International Journal of Forecasting 35.2 (2019): 485-501.
5. Thearasak Phaladisailoed and Thanisa Numnonda. Machine learning models comparison for bitcoin price prediction. In 2018 10th International Conference on Information Technology and Electrical Engineering (ICITEE), pages 506– 511. IEEE, 2018.
6. Greaves, Alex, and Benjamin Au. "Using the Bitcoin transaction graph to predict the price of Bitcoin. 2015." Quoted 3 (2015): 22.
7. Jang, Huisu, and Jaewook Lee. "An empirical study on modeling and prediction of bitcoin prices with bayesian neural networks based on blockchain information." IEEE Access 6 (2017): 5427-5437.

ISSN: 2233-7857 IJFGCN

8. S. McNally, J. Roche and S. Caton, "Predicting the Price of Bitcoin Using Machine Learning, " 2018 26th Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP), Cambridge, 2018, pp. 339-343. doi: 10.1109/PDP2018.2018.00060.
9. Peng, Y., Albuquerque, P.H.M., de Sá, J.M.C., Padula, A.J.A. and Montenegro, M.R., 2018. The best of two worlds: Forecasting high frequency volatility for cryptocurrencies and traditional currencies with Support Vector Regression. , Expert Systems with Applications, 97, pp.177-192.
10. Sun, H. and Yu, B., 2019. Forecasting Financial Returns Volatility: A GARCH-SVR Model. Computational Economics, pp.1-21.
11. Derbentsev, V., Datsenko, N., Stepanenko, O. and Bezkorovainyi, V., 2019. Forecasting cryptocurrency prices time series using machine learning approach. In SHS Web of Conferences (Vol. 65, p. 02001). EDP Sciences
12. Alahmari, S. A. (2019). Using Machine Learning ARIMA to Predict the Price of Cryptocurrencies. ISeCure, 11(3).
13. Xu, M., Watanachaturaporn, P., Varshney, P.K. and Arora, M.K., 2005. Decision tree regression for soft classification of remote sensing data. Remote Sensing of Environment, 97(3), pp.322-336.
14. Imandoust, S.B. and Bolandraftar, M., 2013. Application of k-nearest neighbor (knn) approach for predicting economic events: Theoretical background. International Journal of Engineering Research and Applications, 3(5), pp.605-610.
15. Kaggle , view 15 June 2018. <http:// www.kaggle.com. >
16. Coinmarktcap , viewed 18 Dec 2018. <http:// www. coinmarketcap.com. >.
    Jawadi, Fredj, Zied Ftiti, and Mouna Hdia. "Assessing efficiency and investment opportunities in commodities: A time series and portfolio simulations approach." Economic Modelling 64 (2017): 567-588.
17. Abu Bakar, N. and Rosbi, S. (2017), Autoregressive Integrated Moving Average (ARIMA) Model for Forecasting Cryptocurrency Exchange Rate in HighVolatility Environment: A New Insight of Bitcoin Transaction, International Journal of Advanced Engineering Research and Science, Vol. 4 (11), pp. 130-137
18. T. G. a. N. Antulov-Fantulin, "An experimental study of Bitcoin fluctuation using machine learning methods," in Guo2018AnES, New York, NY, USA, 2018.

ISSN: 2233-7857 IJFGCN