



PREDICTING THE PRICE OF CRYPTOCURRENCY USING SUPPORT VECTOR REGRESSION METHODS

Saad Ali. Alahmari

Department of Computer Science, Shaqra University, Saudi Arabia.

S.alahmari@su.edu.sa

<https://doi.org/10.26782/jmcs.2020.04.00023>

Abstract

The rising profit potential in virtual currency has made forecasting the prices of crypto currency a fascinating subject of study. Numerous studies have already been conducted to predict future prices of a specific virtual currency using a machine-learning model. However, very few have focused on using different kernels of a "Support Vector Regression" (SVR) model. This study applies the Linear, Polynomial and "Radial Basis Function" (RBF) kernels to predict the prices of the three major crypto currencies, Bitcoin, XRP and Ethereum, using a bivariate time series method employing the cryptocurrency (daily-Closed Price) as the continuous dependent variable and the "Morgan Stanley Capital International" (MSCI) World Index (MSCI-WI) and the (daily-Closed Price) as the predictor variable. The results demonstrated that 'RBF' outperforms most other kernel methods in predicting cryptocurrency prices in terms of "Mean Absolute Error" (MAE), "Mean Squared Error" (MSE), "Root Mean Squared Error" (RMSE) and R-squared (R^2).

Keywords: Support Vector Regression, Cryptocurrency, Machine Learning, Time-series Analysis. Non-linear.

1. Introduction

The predictability of cryptocurrency prices is a highly investigated topic in the digital market. Many authors have attempted to predict future directions of different digital currencies by using various features such as changes of daily prices of a specific digital currency, relevant currency exchange rates and prices of economic goods.

The digital currencies are processed in anonymous way through a decentralized network. There are a number of cryptocurrencies with the largest shares of market capitalization such as Bitcoin, XRP and Ethereum. Satoshi Nakamoto launched the first crypto-currency, Bitcoin, in 2009. Ethereum is autonomous with significant capabilities as a whole network. The Ripple payment network launched the currency XRP, which was developed in 2013 as a transparent-source Internet protocol. Several researchers have used various financial and non-financial features to predict prices of cryptocurrencies based on cryptocurrencies market only [XV, XIX, XII, XIV], cryptocurrencies market and commodity markets [XX], and relative technologies of block-chains [IV, VIII]. In this research, the goal was to predict cryptocurrencies prices

using the MSCI-WI, which represents 23 developed market countries with 1,650 constituents [VI]. And applied a bivariate time-series method in which the dependent variable is the Daily-Closed Price of cryptocurrency and the daily price of the MSCI World Index is the predictor.

The main challenge in forecasting the price of “cryptocurrencies” is the significant instability in prices [XII, XX, IX]. Machine learning models have been used in price predictions, particularly the non-linear algorithms that outperform the linear algorithms [XV]. SVR is adopted, which is an extended technique of the ‘Support Vector Machine’ (SVM) [VII]. It is important to note that using SVR kernel methods provides the ability to experiment with linear and nonlinear problems. In this paper, the linear, polynomial and ‘RBF’ kernels are experimented with to identify the optimal kernel function for predicting prices of cryptocurrencies.

Section 2 in this paper discusses similar past work on cryptocurrencies price prediction. The theoretical history of SVR models is outlined in section 3. Section 4 deals with the experimental assessment, presenting the methods for data collection and extraction of the data function. The experimental findings are examined in section 5. Section 6 deals with assumptions and findings, explains the drawbacks of this method and proposes future studies.

II. Related Work

Many researchers have tried forecasting the price of cryptocurrencies using a number of “machine learning” algorithms. Most of the studies were mainly performed in statistical analysis and a limited number of non-linear and linear algorithms. Few researchers have experimented with the SVR model [IV, V]. In this research, experimented with the SVR algorithm in a bivariate time series model in which crypto-currencies (daily-Closed Price) and the MSCI World Indexes (daily-Closed Price) are continuous. And collected a dataset of three cryptocurrencies (“Bitcoin, XRP, and Ethereum”) with the highest market capitalization values relative to the MSCI-WI Wide Index.

The authors of [XV] used nonlinear regression, “Neural Networks”, and “Classification and Regression Tree” models to make a three-way comparison of prediction accuracy. The results of the research could not show obvious distinctions between NNs and CART models, although they outperform non-linear regression models in prediction accuracy. The authors of [XIX] conducted a K Nearest Neighbour (KNN) prediction algorithm based on a non-parametric regression model. The authors of [XII] used a variety of statistical tests and “machine learning” models to analyse the short-term estimation of instability for Bitcoin and U.S. dollars by means of an hourly time series. They initiate that “extreme gradient boosting” and “elastic-net” achieved the best accuracy.

The authors of [XX] used ARIMA and a recurrent seq2seq deep multi-layer neural network (seq2seq) to predict Bitcoin pricing, but their models showed that “Recurrent Neural Networks” (RNN) only beaten ARIMA with additional input sources over the long term. Additionally, the major variability in the Bitcoin datasets over the time of research has resulted in diverse results. The authors of [IV] concentrated on the Bitcoin method, taking “BlockChain” and macroeconomic features into consideration and

using the “Bayesian Neural Networks” (BNN).

Some research has investigated the extensions and including traditional machine learning models. The authors of [XXII] examined the valuation of the instability of three standard currencies against three “cryptocurrencies” by combining aold-style“Generalized Autoregressive Conditional Heteroskedasticity” (GARCH) model with the “machine learning” SVR. They evaluated the Support Vector Regression GARCH against GARCH, EGARCH and GJR models based on volatility forecast, finding that SVR-GARCH yields better accuracy in the volatility forecast. The authors of [IX] propose Binary Auto Regressive Tree (BART), which combines the classic algorithm “Classification and Regression Trees” (C&RT) and autoregressive models ARIMA. The research leads to better RMSE compared to the classic ARIMA-ARFIMA models. The authors of [XX1] experiment with several “Regression” and “Deep Learning” models using 1-minute interval Bitcoin trading data for a six-year period. These models are “Theil-Sen Regression” and “Huber Regression”, “Long Short-Term Memory” (LSTM) and “Gated Recurrent Unit” (GRU). The GRU model gives the best results of MSE at 0.00002 and R2 at 0.992, followed by the results of the LSTM model.

The authors of [XVIII] applied a Bayesian optimized RNN and a LSTM network to forecast the Bitcoin pricing method. The results showed that the LSTM network had a higher 52 percent classification accuracy and a lower 8 percent regression error. They then contrasted those findings without regressors with that of the ARIMA model. The authors of [XXI] used the architecture of the BlockChain network combined with simple Bitcoin architecture and used multiple regression and classification models. They are Linear, Logistic, SVM, and Neural Network. Though the most reliable result using the Neural Network classification was 55 percent, the authors found that the use of the network-based features had no significant effect on Bitcoin price predictions. The authors of [XVII] applied a traditional ARIMA model to various cryptocurrencies to determine the directions of future prices.

Previous research has investigated the prediction of various cryptocurrencies using different models with different features from various domains. However, there is a limited focus on applying SVR algorithms with features of MSCI World Indexes (daily-Closed Price). This research considered three kernel methods of an SVR model trained with datasets from cryptocurrencies (Bitcoin, XRP and Ethereum), Focused on training the regular time series data collection to forecast potential cryptocurrential prices (Bitcoin, XRP, and Ethereum).

III. Support Vector Regression (SVR)

The SVR is an extended technique of the “Support Vector Machine” (SVM) [VII]. Researchers have found that SVM offers excellent performance of time series predictions [XIII]. The SVM finds an optimal hyper-plane to separate two classes of patterns. SVR is formulated as an optimization problem that tries to minimize the prediction error. Refer to “(1)” as the formulation of the optimization function of SVR:

$$\min \frac{1}{2} \|w\|^2 + \gamma \sum_{i=1}^n (\xi_i + \xi_i^*) \quad (1)$$

assume that $\xi_i, \xi_i^* \geq 0$ and w is the undetermined parameter vector.

Copyright reserved © J. Mech. Cont.& Math. Sci.
Saad Ali. Alahmari

One advantage of using the SVR is the ability to apply different kernel methods which they form different formula [X]. Refer to “(2), (3), (4)” defines the applied kernels methods as follow:

$$\text{Linear} = k(x_i, y_j) = x_i \cdot y_j \quad (2)$$

where x_i, y_j are datasets.

$$\text{Polynomial} = k(x_i, y_j) = ((x_i \cdot y_j) + p)^q \quad (3)$$

Where p and q are the kernel parameters and satisfy the condition $p \geq 0, q \in \mathbb{N}$.

$$\text{Gauss Kernl(RBF)} = \exp(- \|x_i - x_j\|^2 / Q) \quad (4)$$

where $Q \geq 0$.

This research applies the three kernels with relevant parameters. However, reaching the optimal parameters requires experimenting with different parameter values. In RBF and Polynomial kernel, it is important to set the optimal values of the ‘Gamma’ and ‘C’ parameters because otherwise, the default values are assigned automatically when the SVR function is called. The ‘C’ parameter shows the degree of the empirical error penalty in the datasets. The ‘Gamma’ parameter, for non-linear hyperplanes, refers to the ‘spread’ of the kernel around the data points [XVI]. The lower the value of Gamma, the broader the decision region, and vice versa.

IV. Experimental Evaluation

IV.i. Data Collection

The datasets used in this analysis includes 200 digital cryptocurrencies and were gathered from online data sources [III] with notifications derived from the “CoinMarketcap” [I]. However, the focus of this research is the largest market capitalisation: Bitcoin, XRP and Ethereum. Thus, extracting a subset for the three cryptocurrencies that represents details of daily markets. The details of market include the following features: open, close, high, low, currency, date, market capital and volume. This was deliberately chosen as a dependent variable as it reflected the day's closing price and the next day's opening price. Also extracted the MSCI World Index dataset from online resources [II] in which the closing daily price was selected as the predictor variable.

IV.ii. Feature Extraction

The datasets were down-sampled every day to gain more insight. The model earned training and extracted within the specific date of the feature “Closed Price” of the MSCI-WI and selected cryptocurrencies as follows:

- 1) Cryptocurrency of Bitcoin:(2013-04-28/ 2019-06-24).
- 2) Cryptocurrency of XRP:(2013-08-04/ 2019-06-24).
- 3) Cryptocurrency of Ethereum:(2015-08-07/ 2019-06-24).

4) MSCI World Index:(2013-04-28/ 2019-06-24).

The captured data sets in Fig. 1 denotes the chronological datasets for MSCI-WI and the three cryptocurrencies (Bitcoin, XRP and Ethereum) for the selected time series. From the mid of 2017 to the end of the year the Bitcoin prices can be clearly seen its tremendous instability. This rise has thus positively influenced XRP and Ethereum prices. It is also obvious that there is a steady increase in the MSCI World Index during the specified time series. The datasets were split into exercise and test sets, with 80 percent of the data being given to the exercise set and 20 percent to the test set.

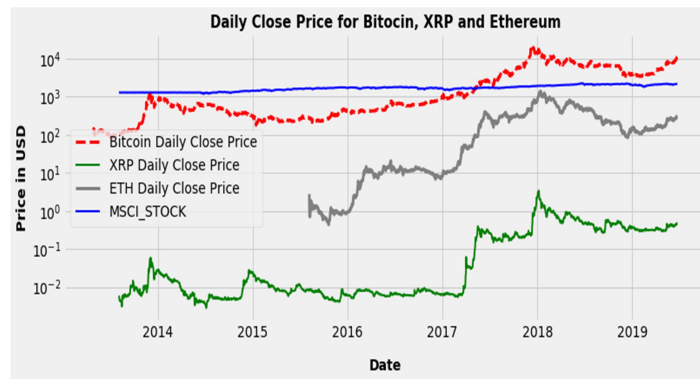


Fig.1: Historical datasets of MSCI-WI, Bitcoin, XRP and Ethereum

Table I shows the summary statistics for these datasets. It should be noted that there are large variances among the prices of the cryptocurrencies being investigated. Substantial price rises in major cryptocurrencies in the third and fourth quarters of 2017 are evident, but the datasets have positively skewed.

Table I: 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

Based on the date feature, the datasets were partitioned; historical data sets were used for exercise, and the latest data were used for testing. The separated datasets were then split into a exercise set of 80% and a test set of 20 per cent. For prediction precision, refer to the steps "(5), (6), (7), (8)": "Mean Absolute Error" (MAE), "Mean Square Error" (MSE), "Root Mean Square Error" (RMSE), and the "coefficient of

determination-R-squared” (R^2). They are defined as follows:

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

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

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

$$R^2 = 1 - \frac{\sum (d_i - z_i)^2}{\sum (d_i - \bar{z}_i)^2} \quad (8)$$

V. Results and Discussion

After several experiments, that observed the ‘Gamma’ and ‘C’ parameters that most affect the SVR model using our datasets.

In the ‘Polynomial’ and ‘RBF’ kernels, optimal values of the Gamma parameter are set equal to 0.00001 and 0.00005, which led to better results with more points of the datasets constraining the model. However, the lower Gamma value may refer to enlarging the decision region.

Consequently, the optimal empirical error penalties ‘C’ for the ‘Polynomial’ and ‘RBF’ kernels are $1e2$ and $1e3$. Thus, the model is highly penalised for the misclassified data with the high values of the ‘C’ parameter, which means low bias. In addition, the SVR is based on using various loss functions that can be controlled by a parameter called ‘epsilon’. It yields the optimal value of epsilon equals 0.2 because its default value is 0.1.

The experimental results of training errors and test errorssummarizes in (see Table I). In terms of MAE, MSE, RMSE and R^2 , the ‘RBF’ kernels outperform those of the ‘Linear’ and ‘Polynomial’ kernels using the Bitcoin and XRP and Ethereum datasets. By the observation, the large size of the datasets, the higher proportion of correlation between the daily-Closed Price (predictor variable) and the MSCI-WI (dependent variable), e.g. the Bitcoin’s dataset as the largest dataset has $R^2 = 78\%$, whileEthereum as the smallest dataset has $R^2 = 56\%$. The ‘RBF’ kernel has the advantage over other kernels of modelling non-linearly distributed problems.

Fig.5(a, b, c) show the model of three kernels using the selected datasets of three cryptocurrencies, respectively. The x-axis stands for MSCI-WI data and the y-axis reflects price changes in US dollars. Every figure compares the learned model of Linear, Polynomial and ‘RBF’ kernels with a selected cryptocurrency. Forinstance, in (5a) using the datasets of ‘Bitcoin’, the ‘RBF’ achieves $R^2 = 78\%$ comparing to $R^2 = 48\%$ and 43% in Linear and Poly. Kernels accordingly. Although the dataset is non-linearly separable differently, the ‘RBF’ kernel performs better than other methods. The detailed results for thesekernel methods are given in Appendix (shown in Table II).

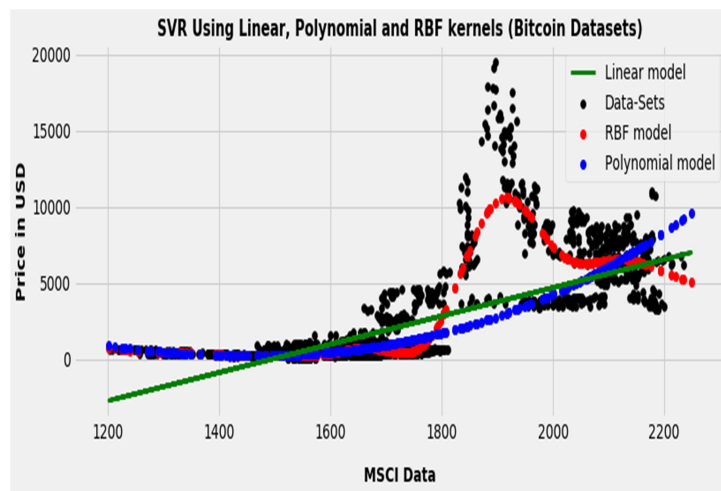


Fig.2:(A): SVR Prediction for Bitcoin Datasets

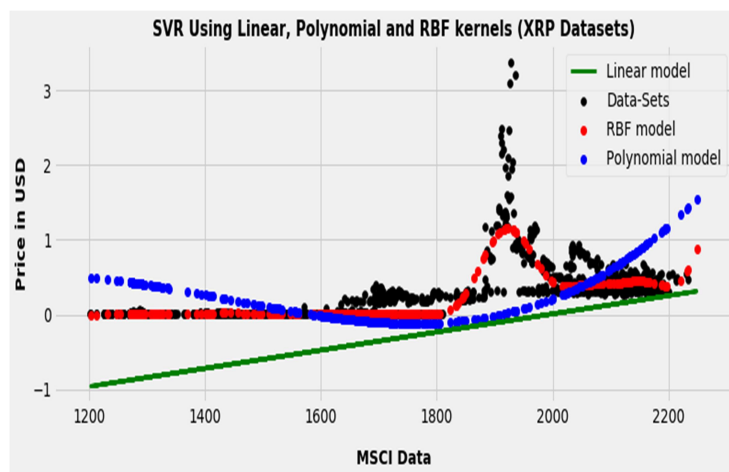


Fig. 2:(B): SVR Prediction for XRPDatasets

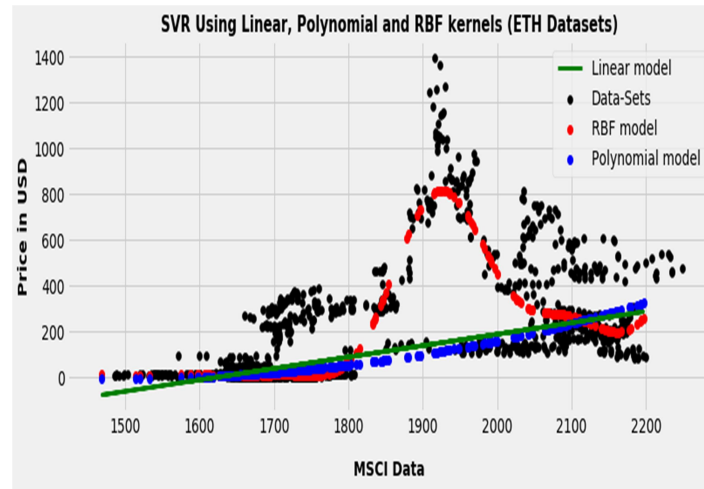


Fig. 2:(C): SVR Prediction for Ethereum Datasets

VI. Conclusion

Cryptocurrency prediction is a research topic with considerable potential, given the significant market capitalisation. This research aims to produce a potential SVR model with the best kernel method for future price prediction for cryptocurrency. Experiments were carried out with Linear, Polynomial and 'RBF' kernels training datasets of real datasets from three major cryptocurrencies and MSCI-WI.

It results that the 'RBF' kernel method outperforms other kernel methods in terms of MAE, MSE, RMSE and R^2 using the datasets for Bitcoin, XRP and Ethereum cryptocurrencies. It can be concluded that the 'RBF' kernel method might solve the problems of dataset variances and high dimensional data comparing to that of Linear, Polynomial methods. However, this high accuracy results in the experiments could be result of being able to optimise SVR hyperparameters. Thus, it is recommended to apply other non-linearly techniques such as Random Forest, K-Nearest Neighbors and Bayesian Network and then compare the results among those techniques. Considering more related highly multi-economic and financial features will influence the prediction results and require applying more advanced machine learning techniques.

References

- I. "Coinmarketcap," <http://www.coinmarketcap.com> (accessed 18 Dec. 2018).
- II. "Investing," <https://www.investing.com/indices/msci-world-stock-historical-data> (accessed 15 June 2018).
- III. "Kaggle," <http://www.kaggle.com> (accessed 15 June 2018).

- IV. B. Alex Greaves, "Using the Bitcoin transaction graph to predict the price of Bitcoin."
- V. C. Giakloglou and P. Newbold, "Empirical evidence on Dickey-Fuller-type tests," *Journal of Time Series Analysis*, vol. 13, pp. 471–483, 1992, doi:10.1111/j.1467-9892.1992.tb00121.x.
- VI. Das, Debojyoti, and KannadhasanManoharan. "Emerging stock market co-movements in South Asia: wavelet approach." *International Journal of Managerial Finance* 15, no. 2 (2019): 236-256.
- VII. . H. Drucker, C. Burges, L. Kaufman, A. Smola, and V. Vapnik, "Support vector regression machines," in M. Mozer, M. Jordan, and T. Petsche, Eds., *Advances in Neural Information Processing Systems 9*, Cambridge, MA, USA: MIT Press, 1997, pp. 155–161.
- VIII. H. Drucker, C. Burges, L. Kaufman, A. Smola, and V. Vapnik, "Support vector regression machines," in M. Mozer, M. Jordan, and T. Petsche, Eds., *Advances in Neural Information Processing Systems 9*, Cambridge, MA, USA: MIT Press, 1997, pp. 155–161.
- IX. H. Sun and B. Yu, "Forecasting financial returns volatility: A GARCH-SVR model," *Computational Economics*, pp. 1–21, 2019.
- X. H. Wang and D. Xu, "Parameter selection method for support vector regression based on adaptive fusion of the mixed kernel function," *Journal of Control Science and Engineering*, 2017.
- XI. J. Huisu and J. Lee. "An empirical study on modeling and prediction of bitcoin prices with bayesian neural networks based on blockchain information." *IEEE Access*, 6 ,pp. 5427-5437.2017.
- XII. J. Rebane, I. Karlsson and P. Papapetrou, "Seq2Seq RNNs and ARIMA models for cryptocurrency prediction: A comparative study," in *Proceedings of SIGKDD Workshop on Fintech (SIGKDD Fintech'18)*, London, UK, Association for Computing Machinery (ACM), 2018, article id 4.
- XIII. K. Müller, A. Smola, G. Rätsch, B. Schölkopf, J. Kohlmorgen, and V. Vapnik, "Predicting time series with support vector machines," in *International Conference on Artificial Neural Networks*, Berlin, Heidelberg: Springer, pp. 999–1004, 1997.
- XIV. L. Catania, S. Grassi, and F. Ravazzolo, "Forecasting cryptocurrencies under model and parameter instability," *International Journal of Forecasting*, vol. 35, no. 2, pp. 485–501, 2019.
- XV. M. Razi and K. Athappilly, "A comparative predictive analysis of neural networks (NNs), nonlinear regression and classification and regression tree (CART) models," *Expert Systems with Applications*, vol. 29, no. 1, pp. 65–74, 2005.

- XVI. M. Suganyadevi and C. K. Babulal, "Support vector regression model for the prediction of loadability margin of a power system," *Applied Soft Computing*, vol. 24, pp. 304–315, 2014.
- XVII. S. Alahmari, "Using machine learning ARIMA to predict the price of cryptocurrencies," *The ISC International Journal of Information Security*, vol. 11, no. 3, pp. 139–144, 2019, doi: 10.22042/isecure.2019.11.0.18.
- XVIII. S. McNally, J. Roche and S. Caton, "Predicting the price of Bitcoin using machine learning," in *2018 26th Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP)*, Cambridge, 2018, pp. 339–343, doi: 10.1109/PDP2018.2018.00060.
- XIX. S. Wang, R. Li, and M. Guo, "Application of nonparametric regression in predicting traffic incident duration," *Transport*, vol. 33, no. 1, pp. 22–31, 2018.
- XX. T. Phaladisailoed and T. Numnonda, "Machine learning models comparison for Bitcoin price prediction," in *2018 10th International Conference on Information Technology and Electrical Engineering (ICITEE)*, IEEE, 2018, pp. 506–511.
- XXI. V. Derbentsev, N. Datsenko, O. Stepanenko, and V. Bezkorovainyi, "Forecasting cryptocurrency prices time series using machine learning approach," in *SHS Web of Conferences*, vol. 65, pp. 02001, 201.
- XXII. Y. Peng, P. Albuquerque, J. de Sá, A. Padula, and M. Montenegro, "The best of two worlds: Forecasting high frequency volatility for cryptocurrencies and traditional currencies with Support Vector Regression, 2018. Expert Systems with Applications, 97, pp. 177–192.