

Comparative Performance of Machine Learning Algorithms for Cryptocurrency Forecasting

Nor Azizah Hitam, Amelia Ritahani Ismail*

Department of Computer Science,
International Islamic University Malaysia (IIUM),
Kuala Lumpur, Malaysia
e-mail: ezajaan032@gmail.com, amelia@iium.edu.my*

Abstract

Machine Learning is part of Artificial Intelligence that has the ability to make future forecastings based on the previous experience. Methods has been proposed to construct models including machine learning algorithms such as Neural Networks (NN), Support Vector Machines (SVM) and Deep Learning. This paper presents a comparative performance of Machine Learning algorithms for cryptocurrency forecasting. Specifically, this paper concentrates on forecasting of time series data. SVM has several advantages over the other models in forecasting, and previous research revealed that SVM provides a result that is almost or close to actual result yet also improve the accuracy of the result itself. However, recent research has showed that due to small range of samples and data manipulation by inadequate evidence and professional analyzers, overall status and accuracy rate of the forecasting needs to be improved in further studies. Thus, advanced research on the accuracy rate of the forecasted price has to be done.

Keywords: Artificial Intelligence; Machine Learning; Support Vector Machines, Neural Networks; Deep Learning;

1.0 Introduction

Forecasting future values or price of experimental time series plays a vital role in almost all fields of studies including economics, science and engineering, finance, business, meteorology and telecommunication [20]. Cryptocurrency, an alternative medium of exchange consisting of over 1441 (as of January 2018) decentralized crypto coin types. Relating machine learning algorithms to cryptocurrency is considered as a new field with limited research studies. In general, system can be used to any directive machine learning problem, in return the system will provide a description relevant to samples both in and out of the dataset.

There are numerous type of cryptocurrency including Bitcoin, Litecoin, Ethereum, Nem, Ripple, Iota, Stellar and others. The cryptographic foundation of each crypto coin makes them vital. Considering the exchange rates of cryptocurrencies are notorious for being volatile, we attempt to model an algorithm that can be used in trading of numerous cryptocurrencies. In order to show the accuracy rate of the predicted price of the proposed methodology, two different data are used as explanatory examples. The comparative cryptocurrencies are Litecoin and Ethereum, Bitcoin, Stellar, Ripple and Nem. This paper uses the root mean square error (RMSE) to evaluate the proposed models.

The outline of this paper is as follows. Section 1 introduces some basic notions of cryptocurrencies and machine learning algorithms. Section 2 discusses the type of cryptocurrency and two largest alternative blockchain technologies, Litecoin (LTC) and Ethereum (XRP) and the purposes of each development. Section 3 presents about machine learning algorithms and three most widely used algorithms, Artificial Neural Networks (ANN) and Support Vector Machines (SVM) and Deep Learning. Section 4 explains the experiments and results of experiments using all models.

1.1 Cryptocurrency

Litecoin (LTC) and Ethereum (XRP) are among the largest alternative blockchain technologies, known as altcoins and were invented after Bitcoin (BTC). Altcoins may have different purposes of

development but are using general methodology based on decentralized P2P network, with the assumption of no network failure and no Internet interruption [5], [6], [7], [8]. Research on the cryptocurrency field is still limited. Mostly, research in this field is focusing on a single cryptocurrency rather than broader areas such as technological advancement, government participation in market regulations as well as market development [32]. This section will focus on six types of cryptocurrency begins with Bitcoin, Ethereum, Litecoin, Nem, Ripple followed by Stellar. In the succeeding section, we focus the review of previous studies on Machine Learning, Support Vector Machines (SVM), Artificial Neural Networks (ANNs) and Deep Learning applied in forecasting.

A peer to peer (p2p) payment cash system, non regulated digital currency and introduced in 2008 with no legal status tendered is known as Bitcoin. It is called as one type of cryptocurrencies with its cryptographic function in its security of creation and money transfer. In recent years, bitcoin turns out to be the most well known currency in the area of volume trading, thus makes a Bitcoin as the most potential financial medium for investors [30]. It locks the transaction as the individualities of the sender, receiver and the volume of transaction are all encrypted [32].

Ethereum (XRP) is a decentralized block-chain based technology that runs Turing-complete to build and execute smart contracts or circulated systems [2], [3]. The value of its coin is called *ether*. It was introduced by Vitalik Buterin in 2013 and funded a year later amounted US\$18 million worth of bitcoins, raised through online public crowd sale [2]. Ether has no boundaries on its circulation, can be traded in cryptocurrency exchanges, not to be one of the payment system but it's intention is merely to be used in the Ethereum network [3], [4].

Litecoin (LTC) was released in October 2011 using a similar technology to Bitcoin, and invented by Charles Lee. The block generation time is decreased as much as 4 times per block (from 10 minutes to 2.5 minutes per block) 84 million of maximum limit, it is equivalent to 4 times higher than Bitcoin and has adopted a different hashing algorithm [1], [2]. Litecoin is considered as the 'silver standard' of crypto coin and turn into a second most accepted by both miners and exchanges [1]. It uses Scrypt encryption algorithm and contradicts to SHA-256 and developed to bid the Bitcoin network transaction confirmation speed and uses an algorithm that was resilient to the advancement of hardware mining technologies.

NEM is a blockchain notarization also known as a peer-to-peer platform that provides services like online payment and messaging system. Having a conjointly owned notarization, it then makes NEM to become as the first public/private blockchain combination [2].

Ripple, an open source digital currency, produced by Jed McCaleb and partner, Chris Larsen, a distributed peer-to-peer network payment medium controlled and managed by a single organization and offers another medium of security mechanism [2], [32]. The development of Ripple is based on Byzantine Consensus Protocol and maximum number of Ripple is 100 million [2].

Stellar, like Ripple offers and entire substitute of security instrument and implemented based on Byzantine Consensus Protocol. Stellar has implemented a new technology to process the financial transactions including open source, scattered and unlimited ownership [32], [33].

1.2 Machine Learning

To succeed on trading, mastering analysis is very important. Future value can be analyzed in two different ways, technical analysis and fundamental analysis. Technical analysis uses trading information from the market information, such as price, trading volume to forecast future price while other uses the information outside the market like economic situation, interest rate and geopolitical issues to forecast future direction [21]. Many investors focus on technical while some focus fundamental. However, there are some investors who focus on overlaps between fundamental as well as technical. This paper will present about technical analysis by applying the machine learning algorithms. Machine learning has been established as a serious model in classical statistics in the forecasting world for over more than two decades [16], [17]. Two most

widely used algorithms for forecasting price movement are known as Artificial Neural Networks (ANNs) and Support Vector Machine (SVM) and both has own patterns of learning [15], [21]. ANNs has been widely used for prediction in securities. Number of issues in ANNs has been discussed by researchers including the selection of parameters and training set [18]. According to [19], the embedding formulation recommends that when a historical dataset S is available, the one-step forecasting can be considered as supervised learning. Supervised learning is the task of deriving a function from training data consist of a set of training dataset. It comes in a set of *input* and *output* variables that is also considered as dependent on the inputs. One-step forecasting can be applied when a mapping model is exist [19]. In one-step forecasting, the previous values of the series, n are available, thus forecasting can be performed as a generic regression problem as Figure 1 below. General approach to model an input/output sense, relies on the accessibility of experimental pairs and denoted as *training set*. Training set is initiated by the historical series S by creating the $[(N - n - 1) \times n]$ input data matrix.

In one step forecasting, the approximator \hat{f} returns the prediction of the value of the time series at time $t + 1$ as a function of the n previous values (the rectangular box containing z^{-1} represents a unit delay operator, i.e., $y_{t-1} = z^{-1} y_t$) [19].

And the $[(N - n - 1) \times 1]$ output vector

$$Y = \begin{bmatrix} y_N \\ y_{N-1} \\ \vdots \\ y_{n+1} \end{bmatrix} \quad (1)$$

For the sake of simplicity, a is assume as $d = 0$ lag time. Henceforth, in this chapter we will refer to the i^{th} row of X , which is essentially a temporal pattern of the series, as to the (reconstructed) *state* of the series at time $t - i + 1$.

1.2.1 Support Vector Machine (SVM)

Support Vector Machine (SVM) method or classifier was introduced as an induction principle that can avoid over-fitting the data at the assimilation of the training dataset [25] and is known as the most flexible technique to construct the explicit and accurate boundaries [23], [24]. SVM works very well in various applications, provide fast training result and easy to use [31]. Eventually, SVM has been invented to answer pattern recognition problems to fault diagnosis problems [25], [26]. It gives nonlinear and solid solution by applying kernel functions to map the input space into a higher dimensional feature [22]. There are many benefits of the SVM including outperforms in generalization model and perform well with small datasets. SVM creates a lot of benefits in many fields including pattern classification problem [27]. As SVM is to produce a classification hyper-plane that differentiate two classes of data with maximum margin.

Standard SVM model is as follows:

$$\begin{aligned} \min_{w, b, \xi_i} \quad & \frac{1}{2} \|w\|^2 + C \sum_{i=1}^l \xi_i \\ \text{s.t.} \quad & y_i (< w, \phi(x_i) > + b) + \xi_i \geq 1, i = 1, 2, \dots, l \\ & \xi_i \geq 0, i = 1, 2, \dots, l, \end{aligned} \quad (2)$$

Another important point of discussion is the options offered by type of SVM. SVM offers linear and nonlinear type of models. Linear SVMs outperforms the nonlinear in terms of speed and execution time, but underperform dealing with complex datasets contains many training examples but less features. While nonlinear SVMs although losing its explanatory power, seems to perform

steadily across various problems, and becomes most preferred choice compared to linear SVMs [31].

1.2.2 Artificial Neural Networks (ANNs)

A common neural network that is doing the deep learning at its hidden layers is called an artificial neural networks [35]. It is an extremely similar system consisting interrelated and interacting processing nodes or neurons [28, 29], works like a human brain and process the information by interacting with a numbers of straightforward processing features [28]. There are input and output neurons in this environment where input neurons will be triggered upon instruments sensing the environment. While other neurons trigger through weighted connections from neurons which was activated earlier, some neurons could effect the environment by activating actions [29]. Depending on the issue and how neurons are linked, such behavior may need a long connecting chains of computational phases where each phase revises the aggregate activation of the network.

1.2.3 Deep Learning (DL)

Deep Learning is mainly to get the most precise result across many phases [29]. Models explained in this section applies a non-linear function on the hidden units and enables a more lavish model that is capable to learn more abstract illustrations to form a deep network when modules are arranged on top of each other [34]. The goal of deep network is to design structures at the lower layers that will separate the variation factors in the input data ad chain the representations at the higher layers, but the drawbacks of the training with multiple hidden layer units lies in the event of the error signal being backpropagated [34].

2.0 Proposed Methodology

In this paper, we consider time series data based on 5 years of daily history, as inputs for all models and may vary based on the availability of datasets from the source. The data is prepared from daily open, close, high and low price of a daily trading for all total of six types of cryptocurrencies and are downloaded from the market capitalization database and range from 2013 through 2018.

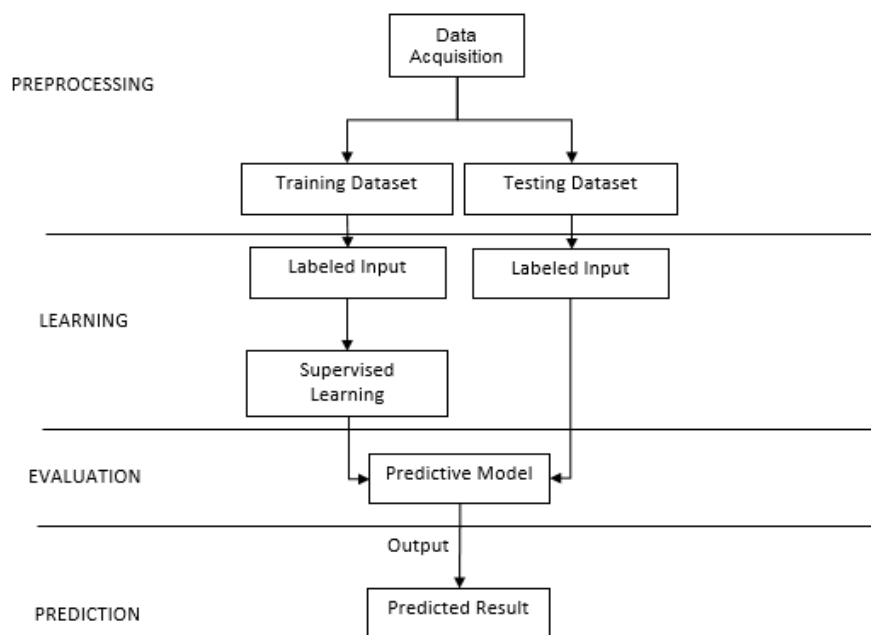


Figure 1: Proposed Methodology

Variable	Description
Open Price	The first price of a given cryptocurrency in a daily trading
Close Price	The price of the last transaction for a given cryptocurrency at the end of a daily trading
High Price	The highest price that was paid for a cryptocurrency during a daily trading
Low Price	The lowest price of a cryptocurrency reached in a daily trading

Table 1: Variable Description

2.1 Data Description

Our main purpose of this paper is to get the most accurate forecasting price, based on the above mentioned methods. Bitcoin, BTC is the first digital currency in market capitalization list and begins since March 2013 through January 2018. Training data for bitcoin starts from 28th March 2013 to 16th until January 2017, followed by Ethereum from 7th August, 2015 to 16th January, 2017, Litecoin from 28th April 2013 to 16th January 2017, Nem 1st April 2015 to 16th January 2017, Ripple 4th August, 2015 through 16th January, 2017 and Stellar from 4th August 2013 to 16th January 2017. While testing data starts for all selected type of cryptocurrencies start from 17th January, 2017 through 16th January 2018 subsequently (Figure 2).

Figure 2: The training and testing dataset in our time series data. The first part is the training set (number of values as per #Observations) in the first segment, accordingly. Several classifiers are then used to predict the test data (number of values in the testing set is = 364) in the second segment.

Cryptocurrency Name	Training Data			Test Data		
	From	To	#Observations	From	To	#Observations
Bitcoin, BTC, XBT	28-Mar-13	16-Jan-17	1388	17-Jan-17	16-Jan-18	364
Ether or "Ethereum", ETH	7-Aug-15	16-Jan-17	526	17-Jan-17	16-Jan-18	364
Litecoin, LTC	28-Apr-13	16-Jan-17	1358	17-Jan-17	16-Jan-18	364
Nem, XEM	1-Apr-15	16-Jan-17	657	17-Jan-17	16-Jan-18	364
Ripple, XRP	4-Aug-13	16-Jan-17	1262	17-Jan-17	16-Jan-18	364
Stellar, XLM	5-Aug-14	16-Jan-17	896	17-Jan-17	16-Jan-18	364

3.0 Results and Analysis

The result section begins by showing performance measures for each cryptocurrency types according to classifiers. These serve as a control for the rest of the discussion.

The analysis is separated into **two** different experiments: i) Performance measures by various classifiers ii) Forecasted cryptocurrency value by machine learning algorithms vs actual value

Figure 3 shows the performance accuracy in correspondence to four classifiers on the cryptocurrency market capitalization. The maximum value is 95.5%, which means that any alphas over 95.5% have p-value of 0.01 or less.

Figure 3: Performance Measures by various classifiers

Classifiers	Performance Accuracy (%)					
	Bitcoin	Ethereum	Litecoin	Nem	Ripple	Stellar
SVM	78.90	95.50	82.40	47.70	70.00	58.70
ANNs	79.40	78.00	75.80	77.80	81.40	89.80
DL	61.90	69.40	62.80	57.20	60.90	70.70
BoostedNN	81.20	81.60	72.20	77.40	81.50	92.80

Several different classifiers were trained with the same set of features. In this case, the datasets were evaluated using classification accuracy. The comparison of all classifiers generated by different methods are based on the same dataset. Thus it will be fair for all classifiers to perform the testing and training.

The results for the classifiers with the best performance on the test set are testified. The results show that SVM classifier works well for Ethereum followed by Litecoin. While, ANN is seen works best for Bitcoin followed by Nem. Ripple and Stellar has the best performance accuracy for BoostedNN. However, among all, SVM classifier performs the best compared to the other classifiers with the performance accuracy of 95.5%.

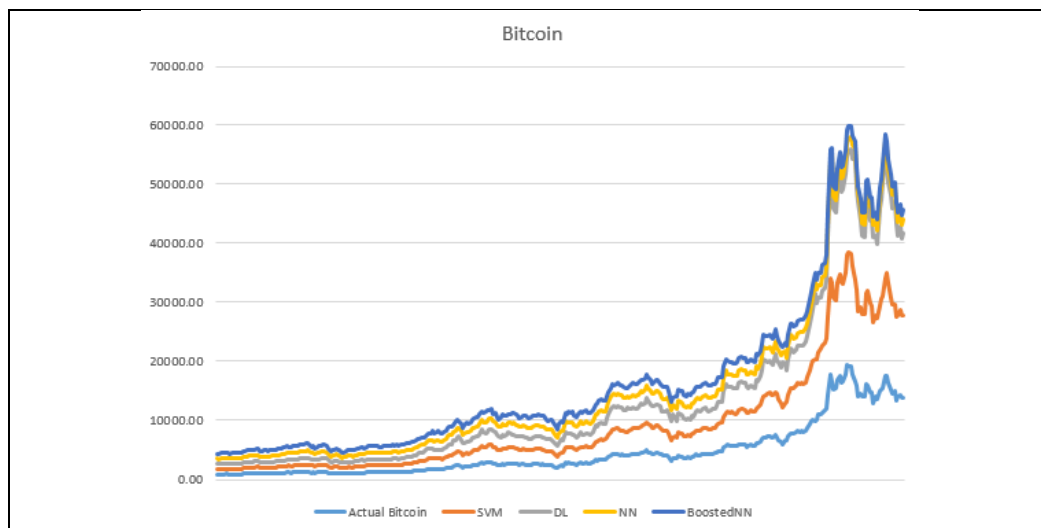


Figure 4: SVM value is comparable to actual Bitcoin for the period from 17/1/2017 to 16/1/2018.

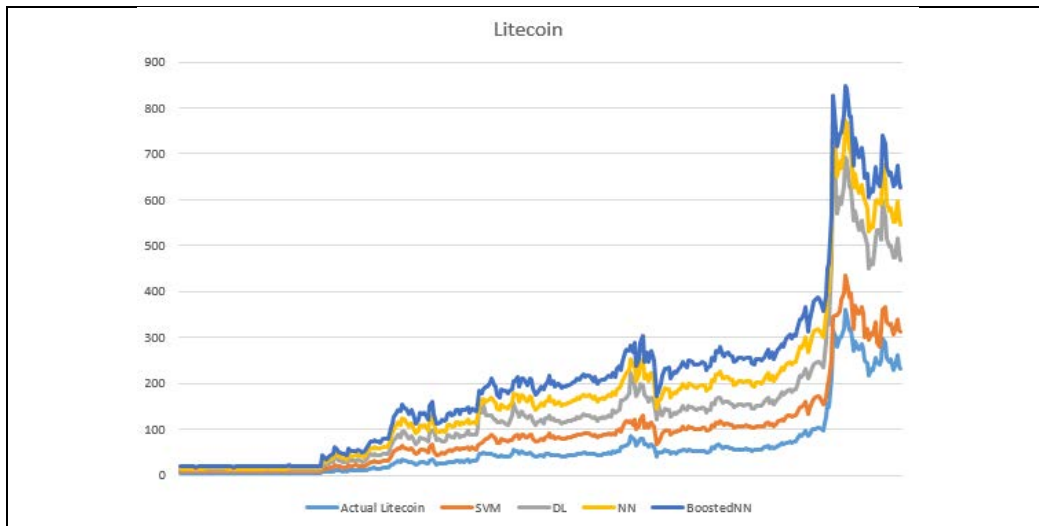


Figure 5: SVM value is comparable to actual Litecoin for the period from 17/1/2017 to 16/1/2018.

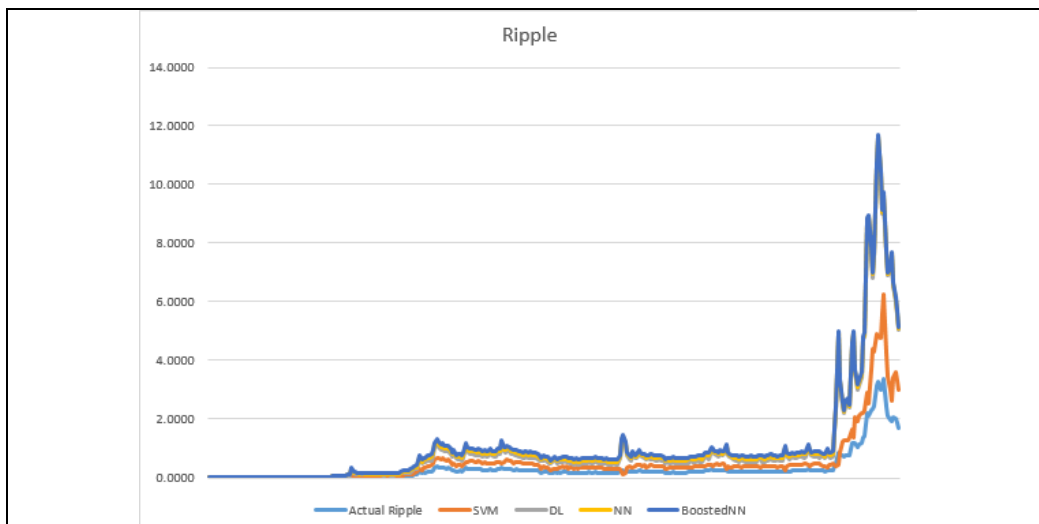


Figure 6: SVM value is comparable to actual Ripple for the period from 17/1/2017 to 16/1/2018.

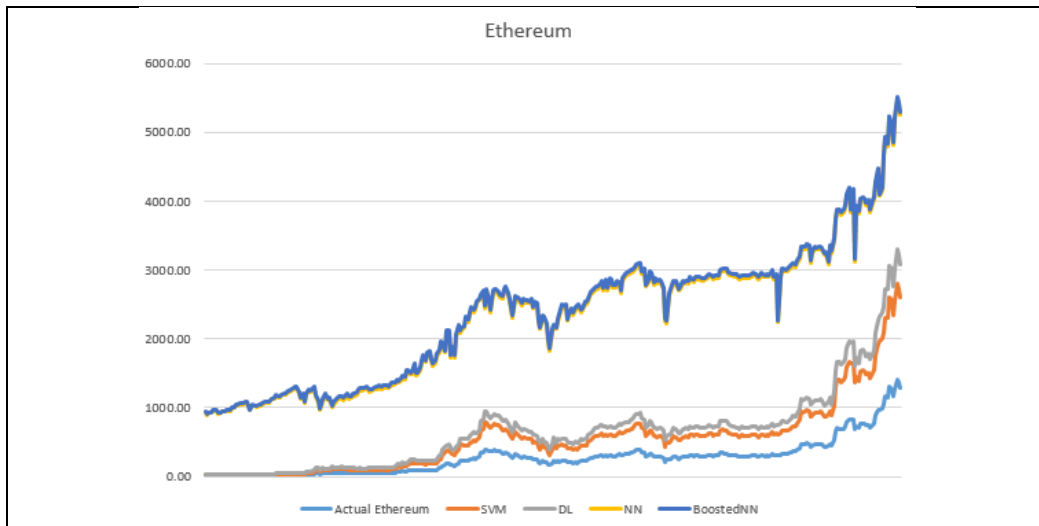


Figure 7: SVM value is comparable to actual Ethereum for the period from 17/1/2017 to 16/1/2018.

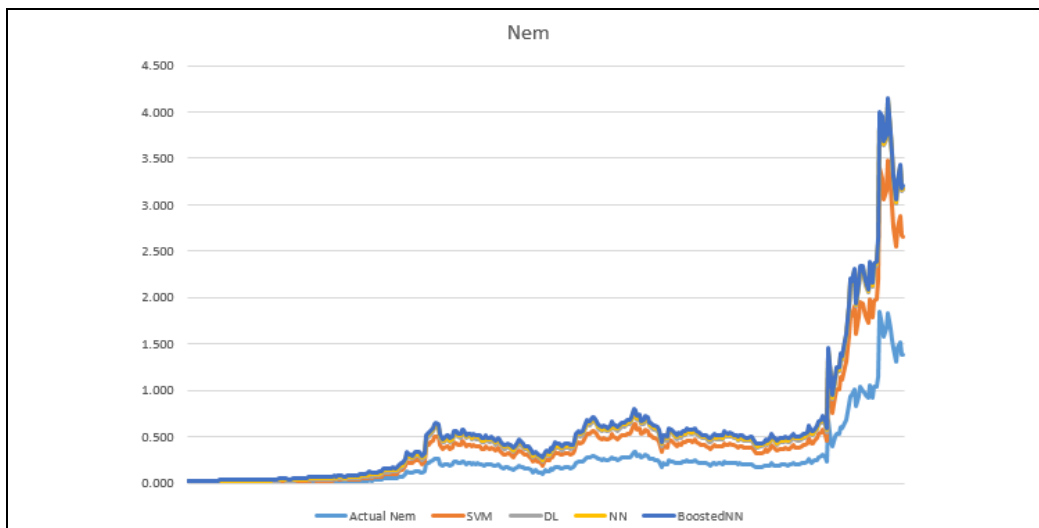


Figure 8: SVM value is comparable to actual Nem for the period from 17/1/2017 to 16/1/2018.

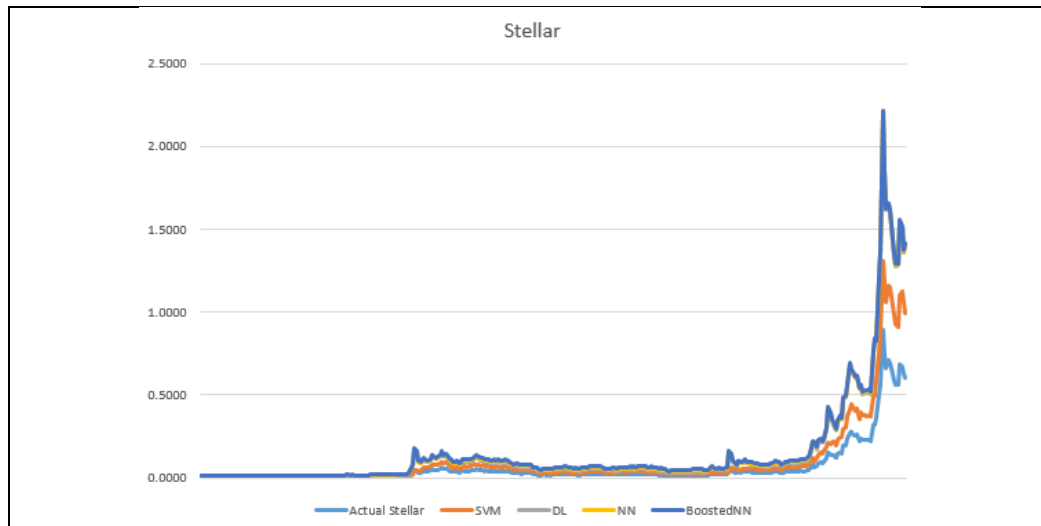


Figure 9: SVM value is comparable to actual Stellar for the period from 17/1/2017 to 16/1/2018.

For comparability, same data sets and period of 364 days were chosen for all classifiers. Performance can be seen in Figure 4 – Figure 9 above. The SVM significantly outperformed the other classifiers.

This result is further explored using mean absolute percentage error (MAPE) calculation. SVM mean absolute percentage error is 0.31% and is the lowest MAPE. Thus, the SVM is considered as reliable forecasting model for these six selected cryptocurrency.

4.0 Conclusion

The paper is highly focuses on the comparative performance of machine learning algorithms of six cryptocurrencies. To begin with, the review of cryptocurrency has covered six major cryptocurrency, there are Bitcoin, Ethereum, Litecoin, Nem, Ripple and Stellar. Further, previous studies on Machine Learning, Support Vector Machines (SVM), Artificial Neural Networks (ANNs) and Deep Learning forecasting has been explored.

Firstly, the performance measures were done to get the accuracy of classifiers over the selected cryptocurrency and obtained the result as in Figure 3. Result shows that SVM outperformed other classifiers with the accuracy of 95.5%. It is realized, that the quality of training data and population of dataset plays an important role for a successful prediction.

Secondly, the forecasted cryptocurrency value by Machine Learning vs actual value of cryptocurrency were then analyzed. From the comparative analysis done in this section, SVM has a comparable values for all cryptocurrency for the period from 17/1/2017 to 16/1/2018.

Moreover, the result is further explored using mean absolute percentage error (MAPE) calculation. The results show that SVM has the lowest value of MAPE. Thus, the SVM is considered as a reliable forecasting model for the selected cryptocurrency.

In future, the algorithm will be improved on the accuracy rate of the forecasted price. Besides, with the power of SVM, future work will be done to further optimize the SVM to get the most accurate result as per actual value of cryptocurrency.

5.0 References

- [1] Heid, A. (2013). Analysis of the Cryptocurrency Marketplace. *Retrieved February, 15, 2014.*
- [2] Lee, D., Chuen, K., Guo, L., Wang, Y., & Chian, L. K. (2017). Cryptocurrency: A New Investment Opportunity?, 1–54.
- [3] Economics, F. O. F. (2016). FACULTY OF ECONOMICS AND BUSINESS The Evolution of Bitcoin Price Drivers : Moving Towards Stability ?
- [4] Buterin, & Vitalik. (2014). Ethereum White Paper: A Next Generation Smart Contract & Decentralized Application Platform. *Etherum*, (January), 1–36. Retrieved from <https://github.com/ethereum/wiki/wiki/White-Paper>
- [5] Huckle, S., & White, M. (2016). Socialism and the blockchain. *Future Internet*, 8(4). <https://doi.org/10.3390/fi8040049>
- [6] Bitcoin. Bitcoin Developer Guide. Available online: <https://bitcoin.org/en/developer-guide#block-chain> (accessed on 24 January 2018).
- [7] Ethereum. Ethereum Project. Available online: <https://www.ethereum.org/> (accessed on 24 January 2018).
- [8] Litecoin. Litecoin—Open Source P2P Digital Currency. Available online: <https://litecoin.org/> (accessed on 24 January 2018).
- [9] Bontempi, G., Taieb, S. Ben, & Borgne, L. (2013). Machine Learning Strategies for Time Series Forecasting, 62–77.
- [10] Coglianese, C. (n.d.). Regulating by Robot: Administrative Decision Making in the Machine-Learning Era.
- [11] Krollner, B., Vanstone, B., & Finnie, G. (2010). Financial time series forecasting with machine learning techniques: A survey. Retrieved from http://works.bepress.com/bruce_vanstone/17/%5Cnpapers2://publication/uuid/CCB43489-A51D-4316-8CD1-0B369F4A07B1
- [12] Dingli, A., & Fournier, K. S. (2017). FINANCIAL TIME SERIES FORECASTING – A MACHINE LEARNING APPROACH, 4(1), 11–27. <https://doi.org/10.5121/mlaij.2017.4302>
- [13] Roche, J., & McNally, S. (n.d.). Predicting the price of Bitcoin using Machine Learning Sean McNally Supervisor :
- [14] Gerlein, E. A., McGinnity, M., Belatreche, A., & Coleman, S. (2016). Evaluating machine learning classification for financial trading: An empirical approach. *Expert Systems with Applications*, 54, 193–207. <https://doi.org/10.1016/j.eswa.2016.01.018>
- [15] Patel, J., Shah, S., Thakkar, P., & Kotecha, K. (2015). Predicting stock and stock price index movement using Trend Deterministic Data Preparation and machine learning techniques. *Expert Systems with Applications*, 42(1), 259–268. <https://doi.org/10.1016/j.eswa.2014.07.040>
- [16] Bontempi, G., Taieb, S. Ben, & Borgne, L. (2013). Machine Learning Strategies for Time Series Forecasting, 62–77.
- [17] Ahmed, N. K., Atiya, A. F., El Gayar, N., & El-Shishiny, H. (2010). An empirical comparison of machine learning models for time series forecasting. *Econometric Reviews*, 29(5), 594–621. <https://doi.org/10.1080/07474938.2010.481556>
- [18] Kongsilp, W., Mateus, C., Huang, M., Ting-ting, Z., Wan-yi, C., Maita, A. R. C., ... de Carvalho, A. F. (2015). Prediction of Stock Trading Signal Based on Support Vector Machine. *Engineering Computations*, 32(1), 445–463. <https://doi.org/10.1108/02644401311286099>
- [19] Bontempi, G., Taieb, S. Ben, & Borgne, L. (2013). Machine Learning Strategies for Time Series Forecasting, 62–77.
- [20] Bontempi, G., Taieb, S., & Borgne, Y. Le. (2013). Machine learning strategies for time series forecasting. *Business Intelligence*, 62–77. Retrieved from http://link.springer.com/chapter/10.1007/978-3-642-36318-4_3
- [21] Chaigusin, S. (2014). An Application of Decision Tree for Stock Trading Rules : A Case of the Stock Exchange of Thailand Proceedings of Eurasia Business Research Conference, (June).

-
- [22] Hacib, T., Acikgoz, H., Bihan, Y. Le, Mekideche, M. R., Meyer, O., & Pichon, L. (2010). Support vector machines for measuring dielectric properties of materials. *COMPEL: The International Journal for Computation and Mathematics in Electrical and Electronic Engineering*, 29(4), 1081–1089. <https://doi.org/10.1108/03321641011044497>
- [23] Basudhar, A. and Missoum, S. (2010), “An improved adaptive sampling scheme for the construction of explicit boundaries”, *Structural and Multidisciplinary Optimization*, Vol. 42 No. 4, pp. 1-13.
- [24] Lin, K., Basudhar, A., & Missoum, S. (2012). Parallel construction of explicit boundaries using support vector machines. *Engineering Computations*, 30(1), 132–148. <https://doi.org/10.1108/02644401311286099>
- [25] Zhang, L., & Wang, J. (2015). Optimizing parameters of support vector machines using team-search-based particle swarm optimization. *Engineering Computations*, 32(5), 1194–1213. <https://doi.org/10.1108/EC-12-2013-0310>
- [26] Baccarini, L.M.R., Rocha e Silva, V.V., de Menezes, B.R. and Caminhas, W.M. (2011), “SVM practical industrial application for mechanical faults diagnostic”, *Expert Systems with Applications*, Vol. 38 No. 6, pp. 6980-6984.
- [27] Kongsilp, W., Mateus, C., Huang, M., Ting-ting, Z., Wan-yi, C., Maita, A. R. C., de Carvalho, A. F. (2015). Prediction of Stock Trading Signal Based on Support Vector Machine. *Engineering Computations*, 32(1), 445–463. <https://doi.org/10.1108/02644401311286099>
- [28] Lu, C. J. (2010). Integrating independent component analysis-based denoising scheme with neural network for stock price prediction. *Expert Systems with Applications*, 37(10), 7056–7064. <https://doi.org/10.1016/j.eswa.2010.03.012>
- [29] The, S., Ai, S., Dalle, I., & Galleria, S. (2014). Deep Learning in Neural Networks : An Overview, 1–88.
- [30] Krause, D. (2017). Bitcoin – a favourable instrument for diversification ? A quantitative study on the relations between Bitcoin Authors :
- [31] Huerta, R., Corbacho, F., & Elkan, C. (2013). Nonlinear support vector machines can systematically identify stocks with high and low future returns. *Algorithmic Finance*, 2(1), 45–58. <https://doi.org/10.3233/AF-13016>
- [32] Farrell, R. (2015). An Analysis of the Cryptocurrency Industry. *Wharton Research Scholars Journal. Paper*, 130. Retrieved from http://repository.upenn.edu/wharton_research_scholars%0Ahttp://repository.upenn.edu/wharton_research_scholars/130
- [33] Application, F. A., & Guidelines, G. (2013). Ashesi University College. Office, 1–4.
- [34] M. L. Å., Karlsson, L., & Loutfi, A. (2014). A review of unsupervised feature learning and deep learning for time-series modeling. *Pattern Recognition Letters*, 42(C), 11–24. <https://doi.org/10.1016/j.patrec.2014.01.008>
- [35] Nivetha, R. Y. (2017). Developing a Prediction Model for Stock Analysis, 4–6. <https://doi.org/10.1109/ICTACC.2017.11>
-