Journal of the Korea Industrial Information Systems Research Vol. 23 No. 4, Aug. 2018 : 81-92
http://dx.doi.org./10.9723/jksiis.2018.23.4.081 ISSN:1229-3741

Bitcoin Price Forecasting Using Neural Decomposition and Deep Learning

Adyan Marendra Ramadhani[1)], Kim Na Rang[2)*], Lee Tai Hun[3)], and Ryu Seung Eui[4)]

Abstract Bitcoin is a cryptographic digital currency and has been given a significant amount of attention in literature since it was first introduced by Satoshi Nakamoto in 2009. It has become an outstanding digital currency with a current market capitalization of approximately $60 billion. By 2019, it is expected to have over 5 million users. Nowadays, investing in Bitcoin is popular, and along with the advantages and disadvantages of Bitcoin, learning how to forecast is important for investors in their decision-making so that they are able to anticipate problems and earn a profit. However, most investors are reluctant to invest in bitcoin because it often fluctuates and is unpredictable, which may cost a lot of money. In this paper, we focus on solving the Bitcoin forecasting prediction problem based on deep learning structures and neural decomposition. First, we propose a deep learning-based framework for the bitcoin forecasting problem with deep feed forward neural network. Forecasting is a time-dependent data type; thus, to extract the information from the data requires decomposition as the feature extraction technique. Based on the results of the experiment, the use of neural decomposition and deep neural networks allows for accurate predictions of around 89%.

Key Words : Neural Decomposition, Deep Neural Networks, Deep Learning, Bitcoin, Forecasting

1. Introduction

Bitcoin is a cryptographic digital currency and has been given attention in literature since it was first introduced by Satoshi Nakamoto in 2009. Bitcoin has become an outstanding digital currency, owning a market capitalization of approximately $60 billion today. Bitcoin and blockchain technology have begun to shape and define new aspects of computer science and information technology[1]. Bitcoin is a decentralized system, which means that it is not regulated by any party other than the regulator, and it is applied as a form of peer-to-peer payment. However, bitcoin supply is limited because of the nature of cryptocurrency itself[2]. In the last few years, there has been a rising demand for bitcoin,

* Corresponding Author : whitecoral@hanmail.net
+ This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2015S1A3A2046781)
Manuscript received May 27, 2018 / revised June 22, 2018 / accepted August 17, 2018
1) Dept. of Management Information Systems, College of Business Administration, Dong-A University, First author
2) Dept. of Management Information Systems, College of Business Administration, Dong-A University, Corresponding author
3) Dong-A University Research Foundation for Industry-academy Cooperation
4) Dept. of Management Information Systems, College of Business Administration, Dong-A University

because of its character and the limitations of its supply. Based on these conditions, it has become a commodity. It is popularly used as an investment product and people trade in it the same manner as they trade in foreign exchange or stock markets[2]. Cryptocurrency trade is now a popular type of currency investment. Bitcoin is one of the investment products in cryptocurrency exchange, because when treated the same as other currencies, it can be volatile. Therefore, investors are still using the same basic principle of investment of "buy low, sell high." With this principle, investors do not blindly invest without calculating the risks. One of the common methods of calculating investment risks is technical market analysis.

Market technical analyses identify the trends of the market in certain periods by using historical market prices. This technique requires knowledge to analyze trends. Mostly used in currency investments, market technical analyses mostly do not take into consideration knowledge skills for analyzing trends. Occasionally, these trends become difficult to analyze when they become too complex or the data is too extensive. Machine learning is one method that is used to resolve the knowledge problem. It provides the capability of producing a prediction model that can estimate trends more accurately without expert knowledge or other skills. In the stock and forex domains, the neural network is one of the popular methods used in predicting future trends.

A neural network is a machine learning method that is loosely inspired by its biological counterparts (Biological Neural Network)[3]. Several works have used neural networks as a machine learning method for predicting future trends[4,5,6]. Among these works, all have proven that this method is sufficiently accurate for predicting future trends. Future trends are a type of time series data; they need to be decomposed first. Neural decomposition (ND) is a new type of time series data that involves decomposition using the neural network technology, neural network for the analysis, and the extrapolation of time series data[7]. Units with a sinusoidal activation function are used to perform a Fourier-like decomposition of training samples into a sum of sinusoids, augmented by units with nonperiodic activation functions to capture linear trends and other nonperiodic components[7]. Neural decomposition outperforms popular time series forecasting techniques including LSTM, echo state networks, ARIMA, SARIMA, SVR with a radial basis function, and Gashler and Ashmore's model.

Deep learning (Deep Neural Networks, DNN) is distinguished from the more commonplace single-hidden-layer neural networks by the fact that the number of node layers through which data passes in a multistep process of pattern recognition[8]. Deep learning is currently popular because of its efficient and effective algorithms and architecture. It performs feature extraction and classification simultaneously, which implies that it only needs one model[9]. There are several works that use deep learning in several cases of foreign exchange data forecasting and other time series problems[10,11,12]. These works use deep learning methods such as deep neural networks and convolutional neural networks for predicting future forecasting. Based on their results, all the studies have proven that this method is sufficiently accurate for forecast prediction.

The previous study cases focus mostly on foreign currency exchange rate prediction as the main problem in the research, and Bitcoin exchange was excluded as a main problem;

this was because bitcoin was controversial in certain countries, but there are several countries that allow it in currency exchange markets, such as Japan and the United States. Most previous studies comparing the methods did not select the best features; therefore, the models that are compared may not show their best performance.

The purpose of this study is to find an optimum method and algorithm to predict Bitcoin forecasting using a combination of DNN and neural decomposition. In order to achieve this, the neural network and neural decomposition methods will be applied in our experiment.

To achieve the purpose of this study, the remainder of this paper is structured in the following manner: Section I presents the introduction; Section II introduces related work and other research in this field; Section III presents the model of the study; Section IV presents the results and conclusion.

2. Related Works

There have been several studies that focused on currency forecasting and another time series data problem using the machine learning method [19,20,21]. Wang used a neural network to predict an exchange rate [4]; Sespajayadi used one to predict the Euro and US dollar exchange rate[5]; and Gill used one to predict the Indian currency exchange rate. Mostly, they all used the neural network as a method of predicting currency exchange rates. Other works used deep learning architecture for predicting currency exchange rates [10,11,12].

Forecasting uses time-dependent data and so it is categorized as the time series data type. Time series data can hold a lot of obvious and hidden information, such as seasonal etc.; decomposition is required to extract the time series data. Decomposition provides a useful general abstract model for thinking about time series and for better understanding of problems during time series analysis and forecasting. Time series decomposition is a mathematical procedure that transforms a time series into multiple time series (seasonal, trending, and random) decomposition[13]. There are several studies that use decomposition, such as Godfrey. Godfrey used the neural decomposition technique to decompose the time series data type[7].

Recently, the research world has witnessed many different algorithms and different types of neural networks, and the deep learning method was developed as a result. Finding the perfect method for one's research is a mammoth task, requiring considerable study and analysis. Our review also required us to study time series decomposition and deep learning.

Based on all related work, mostly prior research employing neural networks, the Convolutional Neural Network (CNN) combined with another method, such as ARITMA, to get better performance has a different feature extraction. Most prior research struggles with forecasting data because the data is highly time-variant and are normally in a nonlinear pattern. Deep-feed forward neural networks and neural decomposition are both suitable methods to resolve this. Based on prior research, feed-forward neural networks yield better results as does neural decomposition. This research adds to the deep learning architecture on a feed forward neural network as the machine learning and combining it with neural decomposition to get a better and optimum performance for Bitcoin forecasting prediction.

3. Research Methodology

The design and method of the proposed forecasting prediction model is experimental design. Experimental design is a scientific approach to research that manipulates one or more variables, and controls and measures any change in other variables. Fig. 1. illustrates the proposed experimental design for bitcoin forecasting. The proposed experiment uses neural decomposition and a DNN (Deep Feed Forward Neural Network) to predict daily Bitcoin forecasting. The model comprises data preprocessing, neural decomposition, data division, DNN training, deep neural network tests, and predicted results. This research uses the python programming and keras library platform as the basis of the programming.

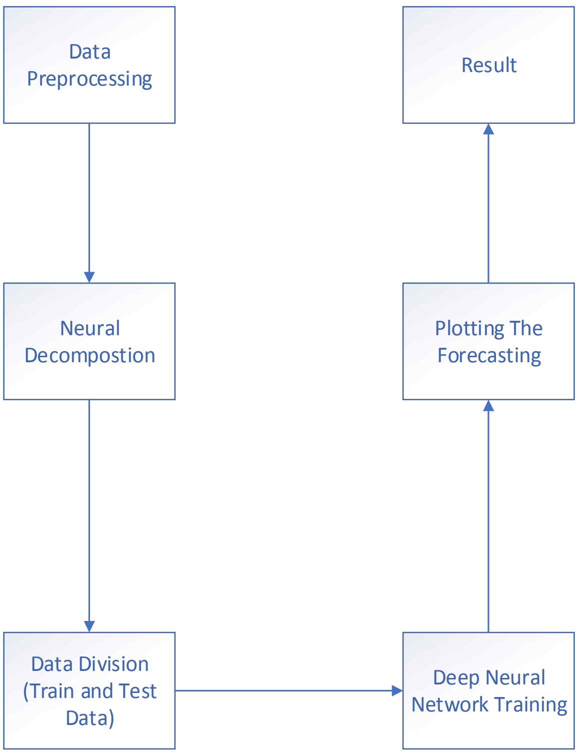

Fig. 1 Research Methodology

The following computer specification is used in this research:

- Intel Core i5 Series (4570 Series)
- RAM 16 GB
- Windows 10 64-bit system
- GPU Nvidia GeForce GT 625 (OEM)

A. Dataset

The Bitcoin dataset was retrieved and extracted from the Kaggle dataset Bitcoin Historical Data. This data included USD and JPY exchange rate histories to and from Bitcoin. Table 1 explains the Bitcoin forecasting dataset used in this research. For training and testing purposes, the Bitcoin to USD exchange rate was selected because of the research limitation.

Table 1 Kaggle Bitcoin Historical Dataset

No	Dataset	Information
1	USD	Bitcoin to USD (US Dollars)

B. Data Preprocessing

Data preprocessing is one of the most critical steps in machine learning and data mining process which deals with the preparation of and transformation of the initial dataset. The following data preprocessing methods were used in this research:

- Data Cleaning
- Data Transformation

Data Cleaning deals with the inconsistent, incomplete, and noisy dataset. In this research, the dataset was check using the interpolation function for incomplete and inconsistent data (Missing Data Value).

The interpolation formula is described in equation (1).

$$y_2 = \frac{(x_2 - x_1)(y_3 - y_1)}{(x_3 - x_1)} + y_1$$

Where :

x_1 and x_3	: First Coordinates
y_1 and y_3	: Values to be interpolated
x_2	: Target X Coordinate
y_2	: Interpolated Y coordinate

Data transformation deals with the transformation or appropriate representation of the dataset. This research uses data transformation for changing the dataset grouping from yearly to daily.

C. Neural Decomposition

The Neural Decomposition (ND) technique was developed by L.B. Godfrey and Michael S. Gashler. Neural decomposition decomposes a set of training samples into a sum of sinusoids, inspired by the Fourier transformation, augmented with additional components to enable our model to generalize and extrapolate beyond the input set. Each component of the resulting signal is trained so that it can find a simpler set of constituent signals[7].

The following ND procedure was used in this research:

- Seasonal decomposition (data checking)[13]
- Seasonal : Patterns that repeat within a fixed period of time.
- Trend : The underlying trend of the metrics.
- Random : The "noise," "irregular," or "remainder." These are the residuals of the original time series after the seasonal and trend series are eliminated.
- Set sin layer kernel to $2\pi[n/2]$
- Set sin layer bias to $\pi/2(1+k(\mathrm{mod}2))$
- Set linear layer kernel to ones with some small noise (Graph)
- Set all other weights to small values close to zero

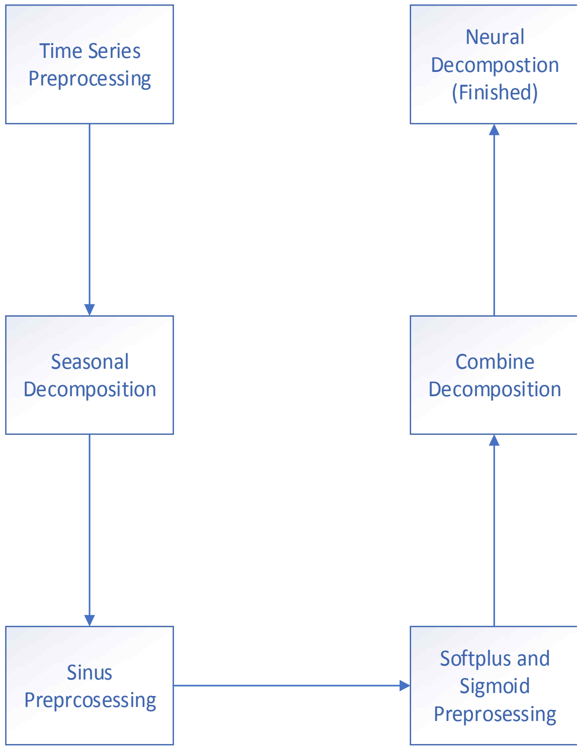

Fig. 2 Neural Decomposition Proposed Method

D. Deep Neural Network

DNNs use deep learning that contains many hidden neuron layers [14]. In other words, deep learning is a sub-field of machine learning, which is based on several levels of representations, corresponding to a hierarchy of features, factors, or concepts, where higher-level concepts are defined from lower-level ones; similarly, lower-level concepts can help to define many higher-level ones[15]. There are several DNN types, such as RBM, deep feed forward neural networks, deep convolutional neural networks, and deep belief networks.

In this research, a deep feed forward neural network consisted of one input and output layer and three hidden layers. All the functions were optimized using deep learning architecture. The result data of Bitcoin neural decomposition is used for CNN training. Fig. 3 explains the details of the deep feed forward neural network architecture that was used in this research. Each of the layers apply the sigmoid and ReLU calculation, and the output applies the MSE and Adam optimizer in the keras system.

- Input layer, comprising four hidden layers
- Layer I, comprising four hidden layers
- Layer II, comprising four hidden layers
- Layer III, comprising sixteen hidden layers
- Output layer, comprising one layer

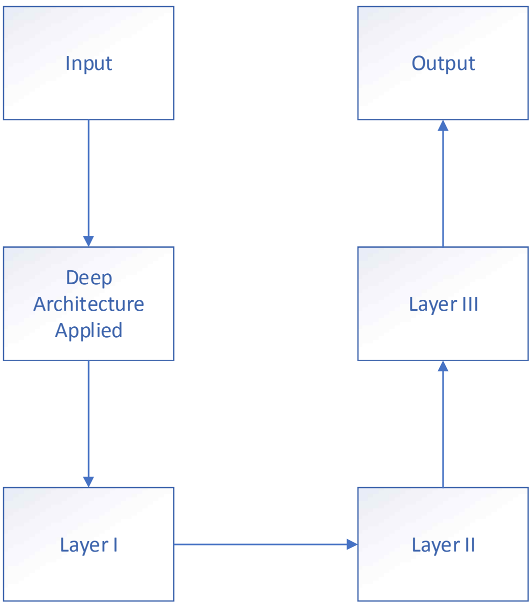

Fig. 3 Deep Feed forward Neural Network Architecture

The sigmoid activation formula is

$$s(x) = \frac{1}{1+\exp(-x)}$$

where x $\rightarrow$ $\pm\infty$.

The ReLU activation formula is

$$f'(x) = \begin{cases} 0 & \text{for } x < 1 \\ 1 & \text{for } x \geq 0 \end{cases}$$

where x is the ReLU data input.

The formula for the mean square unit is :

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y^i - Y_i)$$

Where the Y^i is the vector of n prediction and Y_i is the vector of observed values.

Gradient descent is a way to minimize an objective function $J(\theta)$ by a model's parameters $\theta \epsilon RD$ by updating the parameters in the opposite direction of the gradient of the objective function $\nabla_{\theta}J(\theta)$ with respect to the parameters[16].

The stochastic gradient descent formula is[17]:

$$\omega \leftarrow \omega - \eta(\alpha\frac{\delta R(\omega)}{\delta\omega} + \frac{\delta L(\omega^T x_i + b, y_i}{\delta\omega})$$

Where is η the learning rate that controls the step size in the parameter space. The intercept b is updated similarly but without regularization[17].

4. Results and Conclusion

A. Neural Decomposition

After data preprocessing was applied to the Bitcoin forecasting data, the next step was neural decomposition. These data must always be first decomposed before the training step in a deep feed forward neural network. Fig. 4 illustrates the Bitcoin exchange rate group by time period (days, months, quarters, and years). The data does not offer very much information other than the exchange means. In order to retrieve more information on time series data, decomposition is necessary. The decomposition of time series is a statistical task that deconstructs a time series into several components, each representing one of the underlying category patterns. In this research, we used seasonal decomposition and neural decomposition to retrieve more information about the time series data.

Fig. 5 explains the first seasonal decomposition graph. The graph displays the observed trends and the seasonal and residual Bitcoin time series data. It offers minimal information on the seasonal trends and random time series decomposition and, thus, needs to be analyzed further. Fig. 6 presents the final results of the seasonal trends and random time series decomposition for the Bitcoin data. After the final results of the seasonal trends and random time series decomposition, neural decomposition was applied. The final observed data illustrated in Fig. 6 was used as input for neural decomposition. The neural decomposition used a combination of four different activation functions (sin, softplus, softmax, and linear). The initial weights are based on Fourier transformation. Fig. 7 demonstrates the final neural decomposition for Bitcoin forecasting data from 2013 to 2018. This data shows the final observed Bitcoin forecasting information that will be used as the input for data training. Fig. 8 explains the error rate of neural decomposition, since neural decomposition is based on neural networks. Based on Fig. 8, it can be said that the process of neural decomposition went well, because the graph obtained a convergence state between the 500th and 3000th epochs.

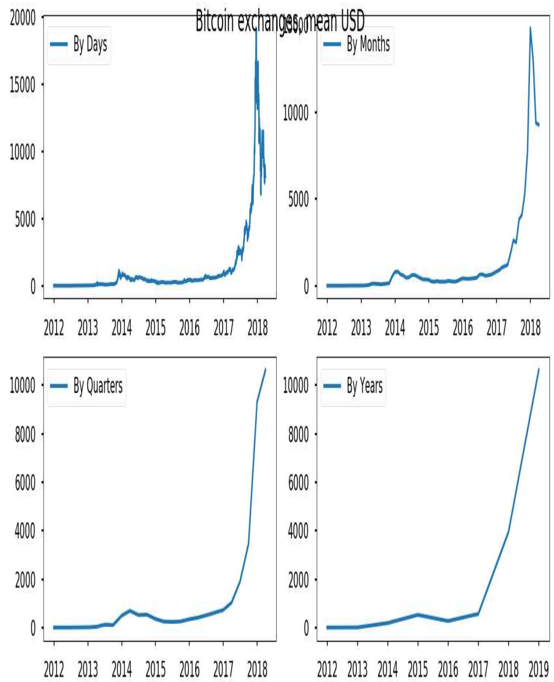

Fig. 4 Bitcoin Forecasting, Grouped by Time Periods

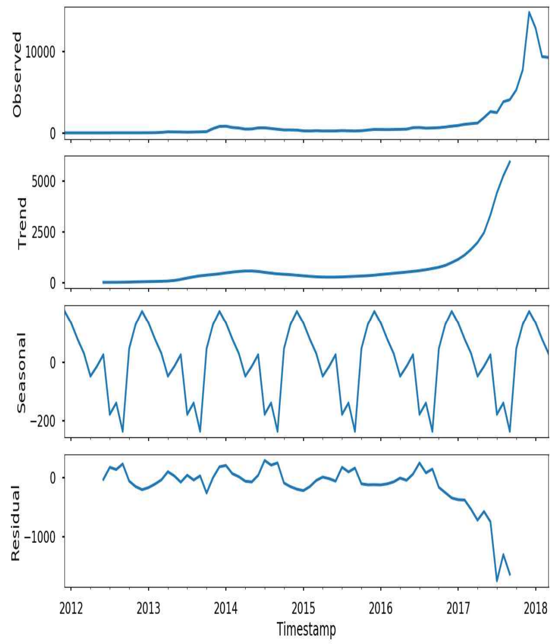

Fig. 5 First Seasonal Decomposition of Bitcoin Data

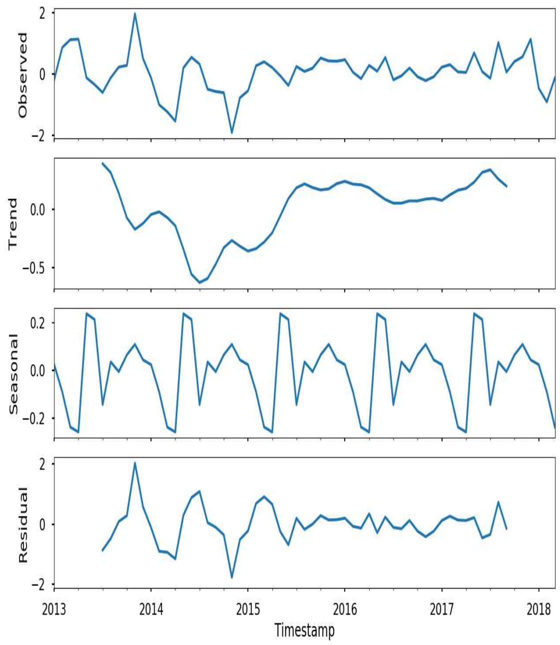

Fig. 6 Final Seasonal Decomposition of Bitcoin Data

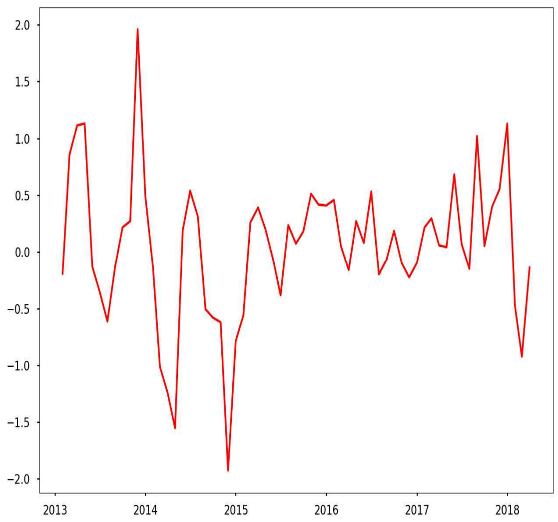

Fig. 7 Neural Decomposition Results After STL Decomposition

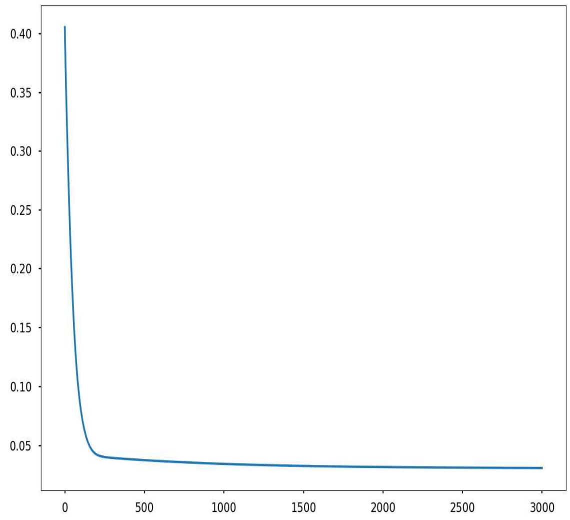

Fig. 8 Loss graph on Neural Decomposition

B. Deep Neural Network

After neural decomposition, a DNN was applied for training and testing. Fig. 9 explains the results of the Bitcoin forecasting using a deep feed forward neural network and neural decomposition data. The deep neural network applied in this research uses feed forward neural network and deep learning architecture. Each of the hidden layers use the sigmoid and ReLU activation, and the output uses MSE and Adam optimizer in keras library system. The red line represents the real Bitcoin price (testing data), while the blue one represents the predicted Bitcoin price. Based on Fig. 9, the prediction is almost accurate. The pattern and value of the blue line is almost the same as the actual data from the Bitcoin price. The prediction accuracy value is 89%. Fig. 10 explain about the loss or error rate in deep neural network. based on Fig. 10 it shows the convergence state between 10^{th} until 100^{th} epoch, so the train phase goes well.

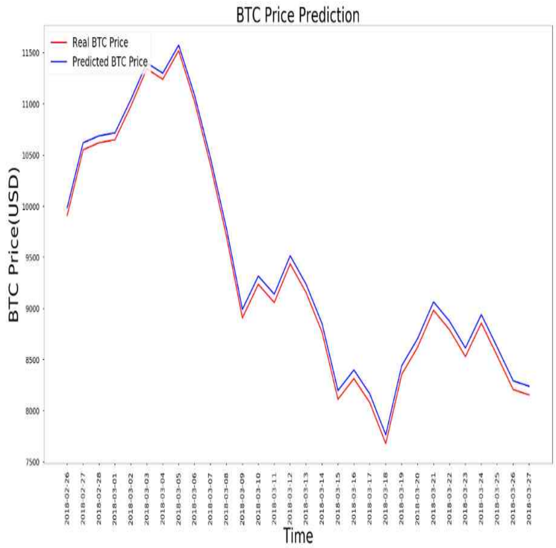

Fig. 9 Bitcoin Prediction Results on Deep Feed forward Neural Network

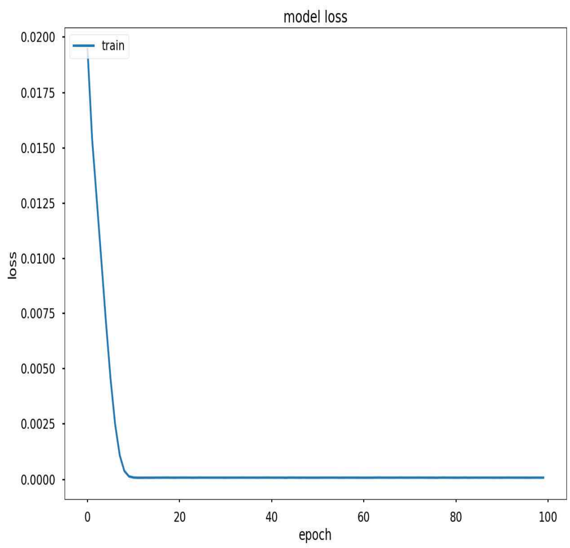

Fig. 10 Loss Graph on Deep Feed forward Neural Network

Based on the results, the combination of the seasonal, neural decomposition, and deep feed forward neural network shows an optimum result for predicting Bitcoin forecasting data. The decomposition helps the deep feed forward neural network predict the Bitcoin forecasting data.

```
Predicted BTC Price: [[ 9998.371 ]
 [10626.85  ]
 [10694.126 ]
 [10722.742 ]
 [11045.302 ]
 [11399.05  ]
 [11299.462 ]

Real BTC Data: [ 9909.97437026 10552.23171617 10621.09253024 10650.38757629
 10980.88168421 11343.89373053 11241.63840684 11520.95055753
```

Fig. 11 Comparison between Predicted and Real USD BTC Exchange Rate

It was found that the deep forward neural network technique is useful in predicting exchange rates. Because of the high volatility, complexity, and noise market environments, neural network techniques are prime candidates for prediction purposes. The neural decomposition and DNN is considered to exhibit better performance in the prediction of exchange rates on a daily basis. Therefore, it is evident that the predictor follows the actual result very closely throughout the daily test period time. Fig. 11 explains and shows the predicted result and the actual BTC data.

Exchange rate data is highly time-variant and is normally in a nonlinear pattern; therefore, predicting the future exchange rate between any two currencies is highly challenging. The noisy feature of the time series is due to the unavailability of complete information from the past history of financial markets to capture the full dependency between future and past prices.

The purpose of this research was to obtain optimum results for income and earning prediction, and to determine an efficient and effective method when faced with a substantial amount of data. However, due to limitations in hardware and data length, this experiment did not handle any more than 35,000 sets of data in the income dataset.

The USD dataset was used because the USD is widely used by investors worldwide, with an interest to invest in Bitcoin, and the research focuses on USD and BTC forecasting.

Based on the current limitations and conditions, this research suggests that several considerations are needed for the future. The first consideration is dataset limitations. For future research, collecting and employing more datasets and testing using another currency such as JPY (Japanese Yen) and Euro will lead to better information for forecasting prediction. The second is hardware limitations; the current research hardware was trained in the model with over two hours, while testing only takes four minutes using high specification hardware for better performance and fast processing. The last consideration is that using this method with another method is optimal for comparing predictions and performance.

References

[1] Vujičić, D., Jagodić, D., and Ranđić, S., "Blockchain Technology, Bitcoin, and Ethereum: A Brief Overview," pp. 1-6, 2018. doi:10.1109/INFOTEH.2018.8345547

[2] Radityo, A., Munajat, Q., and Budi, I., "Prediction of Bitcoin Exchange Rate to American Dollar Using Artificial Neural Network Methods," Bali, Indonesia, pp. 433-438, 2017.

[3] Gerven, M., and Bohte, S., "Artificial Neural Networks as Models of Neural Information Processing," Frontiers Research Topics, 2018.

[4] Wang, S., Tang, Z., and Chai, B., "Exchange Rate Prediction Model Analysis Based on Improved Artificial Neural Network Algorithm," Coimbatore, pp. 1-5, 2016. doi:10.1109/CESYS.2016.7889912

[5] Sespajayadi, A., Indrabayu, and Nurtanio, I., "Technical Data Analysis for Movement Prediction of Euro to USD Using Genetic Algorithm-Neural Network," 2015 International Seminar on Intelligent Technology and Its Applications (ISITIA), Surabaya, pp. 23-26, 2015. doi: 10.1109/ISITIA.2015.7219947

[6] Gill, S. S., Gill, A. K., and Goel, N., "Notice of Retraction of Indian Currency Exchange Rate Forecasting Using Neural Networks," Chengdu, pp. 196-199, 2010. doi:10.1109/ICAMS.2010.5553258

[7] Godfrey, L. B. and Gashler, M. S., "Neural Decomposition of Time-Series Data for Effective Generalization," in IEEE Transactions on Neural Networks and Learning Systems, 2017. doi:10.1109/TNNLS. 2017.2709324

[8] Deeplearning4j.org. "Introduction to Deep Neural Networks (Deep Learning) - Deep learning4j: Open-source, Distributed Deep Learning for the JVM," [online], Available at: https://deeplearning4j.org/neuralnet-overview [Accessed May 8, 2018], 2018.

[9] Sief, G., "I'll tell you why deep learning is so popular and in demand," [online], Medium, Available at: https://medium.com/swlh/ill-tell-you-why-deep-learning-is-so-popular-and- in-demand-5aca72628780 [Accessed May 8, 2018], 2018.

[10] Galeshchuk, S. and Demazeau, Y., "Forecasting Hungarian Forint Exchange Rate with Convolutional Neural Networks," Krakow, pp. 1-3, 2017.

[11] Khare, K., Darekar, O., Gupta, P., and Attar, V. Z., "Short-Term Stock Price Prediction Using Deep Learning," 2017 2nd IEEE International Conference on Recent Trends in Electronics, Information and Communication Technology (RTEICT), Bangalore, pp. 482-486, 2017.

[12] Korczak, J. and Hemes, M., "Deep

Learning for Financial Time Series Forecasting in A-Trader System," 2017 Federated Conference on Computer Science and Information Systems (FedCSIS), Prague, pp. 905-912, 2017.
[13] Anomaly., "Seasonal Trend Decomposition in R," [Online], Available: https://anomaly.io/seasonal-trend-decomposition-in-r/. [Accessed: May 16, 2018], 2018.
[14] Miachel, R., "3 Steps of Text Mining," [Online], Available at http://www2.cs.man.ac.uk/~raym8/comp38212/main/node203.html [Accessed May 20, 2017], 2012.
[15] Goodfellow, I., Bengio, Y., and Courville, B., "Deep Learning," MIT Press, [Online], Available at http://www.deeplearningbook.org [Accessed May 20, 2017], 2016.
[16] En.wikipedia.org., "Deep Learning," [Online] , Available at: https://en.wikipedia.org/wiki/ Deep_learning. [Accessed: May 16, 2018], 2018.
[17] Ruder, S., "An Overview of Gradient Descent Optimization Algorithms," [Online], Available at http://sebastianruder.com/optimizing-gradient-descent/ [Accessed May 20, 2017], 2016.
[18] Scikit Learn Team., "Stochastic Gradient Descent," [Online], Available at http://scikit-learn.org/stable/modules/sgd.html [Accessed 20 May 2017], 2016.
[19] Jin. S. S., "An Idea, Strategy of Congestion Pricing for Differentiated Services and Forecasting Probability of Access Using Logistic Regression Model," Journal of the Korea Industrial Information Systems Research, Vol. 10 No. 1, pp. 9-15, 2005.
[20] Kang, K. H. and Kim, P. K., "Study of Forecasting Game Usage Hours Using Time Series Analysis," Journal of the Korea Industrial Information Systems Research, Vol. 15 No. 5, pp. 63-69, 2010.
[21] Na, J. H. and Kim, E. S., "Forecasting Unemployment Rate Using Social Media Information," Journal of the Korea Industrial Information Systems Research, Vol. 18 No. 6, pp. 95-101, 2013.

마 렌 드 라 (Adyan Marendra Ramadhani)

- 정회원
- Indonesia University of Education 자연과학대학 컴퓨터 과학 교육학과 학사
- Bandung Institute of Technology 전기 공학 및 정보학대학 전기공학 석사
- 부경대학교 IT융합응용공학과 대학원정보시스템협동과정 석사
- 동아대학교 경영대학 경영정보학과 박사 과정
- 관심분야 : 인공지능, 정보시스템, 경영정보시스템, 블록체인

김 나 랑 (Kim Na Rang)

- 정회원
- 부산대학교문헌정보학과 문헌정보학사
- 동아대학교경영정보학과 경영학석사
- 동아대학교경영정보학과 경영학박사
- 관심분야 : 지역혁신, Co-creation, 경영정보시스템, SNA

이 태 헌 (Lee Tai Hun)

- 정회원
- 동아대학교도시계획학과 공학사
- 리츠메이칸대학교정책과학 연구학 이학석사
- 리츠메이칸대학교정책과학 연구학 이학박사
- 관심분야 : 지역혁신, 마을만들기, 지역정책, 지리 정보시스템

유 승 의 (Ryu Seung Eui)

- 정회원
- 동명정보대학교경영정보학과 경영학사
- Texas A&M International University 경영학석사
- University of Texas at El Paso 경영학박사
- 관심분야 : 경영정보시스템, 빅데이터, 오피니언 마이닝