Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 167 (2020) 788–798

Procedia Computer Science

www.elsevier.com/locate/procedia

International Conference on Computational Intelligence and Data Science (ICCIDS 2019)

Analysis of look back period for stock price prediction with RNN variants: A case study on banking sector of NEPSE

Arjun Singh Saud[a]*, Subarna Shakya

[a]*Tribhuvan University, IOST, Kathmandu, Nepal*
[b]*Tribhuvan University, IOE, Lalitpur, Nepal*

Abstract

Stock market prediction is an attempt of determining the future value of a stock traded on a stock exchange. Stock market investors try to predict the stock's future price to make trading decisions such that optimum profit can be earned. Deep learning models are found most successful in predicting stock prices. This paper has performed a novel analysis of the parameter look-back period used with recurrent neural networks and also compared stock price prediction performance of three deep learning models: Vanilla RNN, LSTM, and GRU for predicting stock prices of the two most popular and strongest commercial banks listed on Nepal Stock Exchange (NEPSE). From the experiments performed, it is found that GRU is most successful in stock price prediction. In addition, the research work has suggested suitable values of the look-back period that could be used with LSTM and GRU for better stock price prediction performance.

Keywords: Stock Price Prediction; Deep Learning; LSTM; GRU; NEPSE;

* Corresponding Author. Tel.: 977-1-4332619;
E-mail address: arjunsaud@cdcsit.edu.np

1877-0509
10.1016/j.procs.2020.03.419

1. Introduction

The stock price is the highest price that a buyer is ready to pay for the stock or it is the lowest price on which a seller is ready to sell the stock. Supply and demand is the major factor that may change stock prices. If many people want to buy a stock then there will be high demand and stock price goes up and if many people want to sell a stock then there will be high supply and stock price goes down. Though it is easy to understand the relationship between demand supply and stock price, it is difficult to determine the exact factors that contribute to increased demand and supply. Stock market prediction is an attempt of determining the future value of a stock traded on a stock exchange. The two main schools of thought used by financial analysts for analyzing and predicting stock markets are technical analysis and fundamental analysis.

It has been the aim of investors since the beginning of the stock market to predict stock prices. Technical analysts forecast the future value of stock prices by looking at stock charts and identifying patterns and trends. It is difficult for financial analysts to analyze and forecast the market by churning huge volumes of past data generated by stock markets. Therefore, many Artificial Intelligence (AI) techniques have been invented to predict stock prices automatically. Probably the first research on automatic prediction of stock prices dates back to 1994, in which a comparative study of machine learning regression models was performed [1]. Since then, many attempts are made to devise strategies for forecasting the price of stocks automatically. Automated methods of forecasting the stock prices are becoming feasible with the advent of AI, big data and increased computational capabilities.

The major problem with machine learning algorithms is that their performance depends heavily upon the representation of data they are given [2]. Time series data generated by a stock market can be best described as a random walk which makes feature engineering of stock data much harder. Therefore stock price prediction using machine learning algorithms is much harder. Since deep learning models do not require feature engineering to be performed separately, these models are widely used methods for prediction stock prices or stock price trends from a large amount of past data. This research work has compared the prediction performance of recurrent neural network models: Vanilla Recurrent Neural Network (VRNN), Long Short-term Memory (LSTM), and Gated Recurrent Unit (GRU). The rationale behind choosing RNN variants is that these are the best and widely used models for sequence prediction [3]. In addition, this paper has performed a novel analysis of the look-back period used with recurrent neural networks (RNNs) for predicting stock prices. Two banking stocks that are analyzed in this research work are Nepal Investment Bank (NIB) and Nabil Bank Limited (NABIL).

2. RNN Variants

Three RNN models used in this research work for stock price prediction are: Vanilla RNN (VRNN), Long Short-term Memory (LSTM), and Gated Recurrent Unit (GRU).

2.1 Vanilla Recurrent Neural Networks

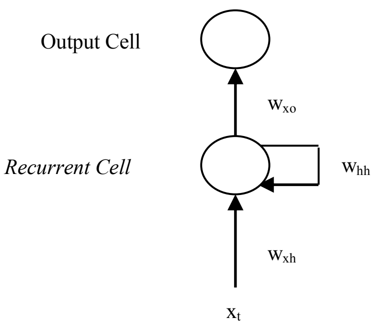

Figure 1: Architecture of VRNN

It is the most basic form of a recurrent neural network. The simplest form of VRNN can be represented as in the above figure. Normally, the tanh activation function is used in the hidden recurrent layer and the activation function for the output layer is selected according to the need of the problem to be solved. The operation of vanilla RNN can be expressed mathematically as below.

$$
\begin{aligned}
&H_t = f(W_{xh}X_t + W_{hh}H_{t-1}) \\
&O_t = g(W_{xo}S_t) \\
&\textit{Where}, X_t \textit{ is input at time } t, H_t \textit{ is state } \text{info} \textit{ at time } t \\
&\quad, W_{xh}, W_{hh}, \textit{and } W_{xo} \textit{ are weight matrices}
\end{aligned}
\tag{1}
$$

The main problem with Vanilla RNN is that the gradient of the loss function decays exponentially with time and the model stops learning or takes way too long. This problem is called the vanishing gradient problem. When Back-propagation is done the gradient tends to get smaller and smaller as we keep on moving backward in the Network [4].

2.2 Long Short-term Memory Network

The LSTM is a special kind of RNN capable of learning long-term dependencies. It is explicitly designed to avoid the vanishing/exploding gradient problem. An LSTM is well-suited to classify and/or predict time-series data. There are several architectures of LSTM units. A common architecture is composed of a memory cell, an input gate, an output gate and a forget gate [5]. In addition, every LSTM cell computes new values of hidden state and cell sate. The mathematical formulation of the LSTM cell is given below.

$$
\begin{aligned}
&f_t = \sigma(x_t W_f + H_{t-1} U_f) \qquad i_t = \sigma(x_t W_i + H_{t-1} U_i) \\
&o_t = \sigma(x_t W_o + H_{t-1} U_o) \qquad H_t' = \tanh(x_t W_g + H_{t-1} U_g) \\
&S_t = \sigma(S_{t-1} \times f_t + i_t \times H_t') \qquad H_t = \tanh(S_t) \times o_t \\
&\textit{where}, i, f, \textit{and } o \textit{ are input, forget, and output gates respectively} \\
&H \textit{ and } S \textit{ are hidden state and memory state recpectively}
\end{aligned}
\tag{2}
$$

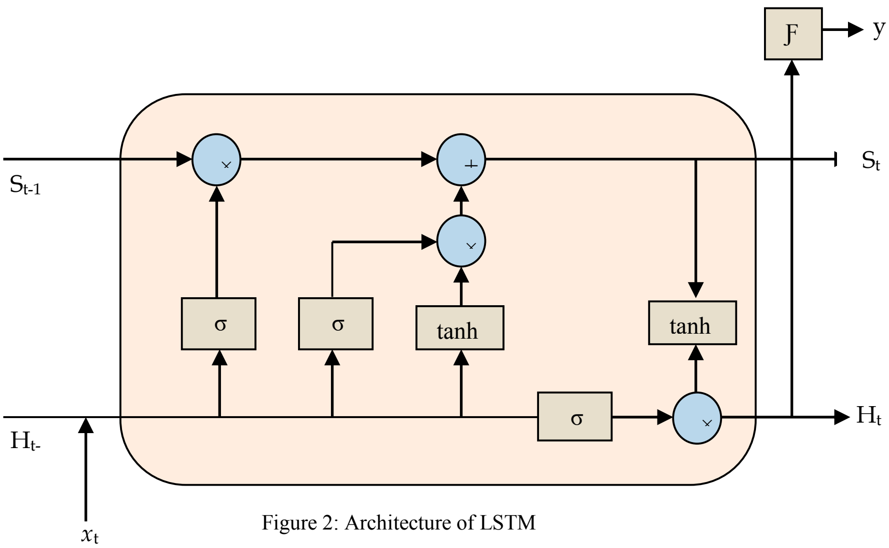

Figure 2: Architecture of LSTM

2.3 Gated Recurrent Unit Network

GRU is also a type of RNN architecture and aims to solve the vanishing gradient problem. GRU is similar to the LSTM but simplified in structure. GRU uses two gates: update gate and reset gate [6]. GRU has fewer parameters and thus may train a bit faster or need fewer data to generalize. But, with large data, the LSTMs with higher expressiveness may lead to better results [7]. The mathematical formulation for GRU is given below.

$$
\begin{aligned}
&Z_t = \sigma(W_z x_t + U_z H_{t-1}) \qquad R_t = \sigma(W_r x_t + U_r H_{t-1}) \\
&H_t^{'} = \tanh(W_h x_t + (R_t \times H_{t-1}) U_h) \qquad H_t = (Z_t \times H_t^{'}) + ((1 - Z_t) \times H_{t-1}) \\
&\textit{where, } Z \textit{ and } R \textit{ are update and reset gates,} \\
&H_t^{'} \textit{ and } H_t \textit{ are candiadte hidden state and hidden states}
\end{aligned}
\tag{3}
$$

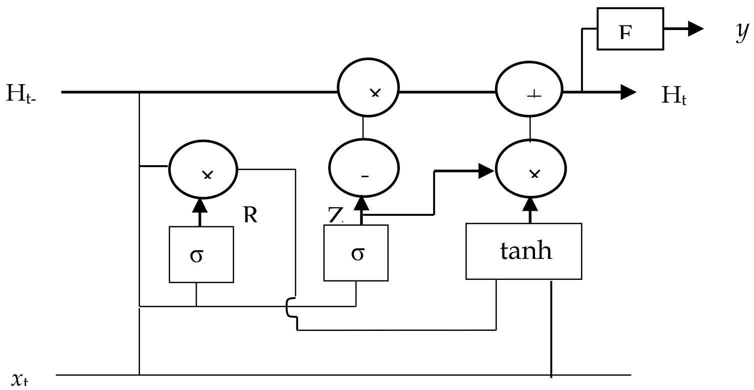

Figure 3: Architecture of GRU

3. Related Works

Minh et al. proposed a two-stream gated recurrent unit (TGRU) network and a sentimental word embedding model Stock2Vec for short term stock trend prediction. The authors have conducted two experiments: the first experiment for predicting S&P 500 index stock price directions using the historical S&P 500 prices and the articles from Reuters and Bloomberg, and the second experiment for forecasting the price trends of VN-index using Viet Stock news and stock prices from cophieu68. The experiments showed that TGRU is able to outperform GRU and LSTM [8]. Hiransha et al. used four deep learning architectures and one linear prediction model ARIMA for the stock price prediction of stocks listed on the National Stock Exchange of India (NSE) and the New York Stock Exchange (NYSE). The researchers have trained four networks MLP, RNN, LSTM, and CNN with the stock price of TATA MOTORS from NSE and the model was used for predicting stocks from both NSE and NYSE. The authors have observed that the models are capable of identifying the patterns existing in both the stock markets and concluded that there exist an underlying dynamics, common to both the stock markets. Results showed that Linear models like ARIMA were not capable of identifying underlying dynamics within various time series and deep learning models were able to outperform the ARIMA model. Besides that CNN has performed better than the other three deep learning architectures [9].

Selvin et al. proposed an overlapping sliding window based approach with RNN, LSTM, and CNN. The window size was fixed to be 100 minutes with an overlap of 90 minute's information and the prediction was made for 10 minutes in the future. The paper showed that CNN gives more accurate results than the other two models [10]. Khare et al. forecasted these short-term prices of 10 unique stocks recorded on NYSE using MLP and LSTM. The research showed that the LSTM model is successful in predicting the future price trend almost at all the points. But the model fails to predict the exact price with the required accuracy. On the other hand, the MLP model is able to capture the future trends as well as it is able to predict the prices with an extremely high level of accuracy as compared to the LSTM model [11]. Huynh et al. proposed a deep neural network architecture named bidirectional gated recurrent unit (BGRU) for predicting stocks listed on the S&P 500 index and compared performance of BGRU with LSTM and GRU networks. To assess the impact of the financial news on the price of the stock over time, the researchers examined the prediction method on many time intervals. The experiments showed that the highest accuracy is obtained in the first 24 hours and accuracy achieved with BGRU is higher than accuracy achieved from LSTM and GRU [12].

Ding et al. proposed a deep learning method for event-driven stock market prediction. The input to the neural tensor network was word embedding and the output was event embedding. Experiments are then conducted to predict the S&P 500 index and its individual stocks. The experiments showed that events are better features than words for stock market prediction [13]. Vargas et al. proposed the RCNN model to forecast the intra-day directional movements of the S&P 500. The model used a set of seven technical indicators and financial news titles published the day before the prediction day. The results showed that sentence embedding is better than word embedding, RCNN is better than CNN, and the influence of technical indicators leads to better performance [14]. Singh and Srivastava predicted stock price from NASDAQ by using a deep neural network with 2D2PCA and compared it with Radial Basis Function Neural Network with 2D2PCA. The authors have conducted experiments for varying window sizes and concluded that window size 20 performs the best [15]. Gunduz et al. predicted the daily movement directions of three frequently traded stocks GARAN, THYAO, and ISCTR on the Borsa Istanbul Stock Exchange (BIST) using CNN. The authors performed two experiments, one using only stock price as input and another using stock price and gold price as input. The experiments showed that the second experiment leads to more accurate prediction results [16].

Zhuge et al. proposed a prediction model for predicting the opening price of stocks listed on the Shanghai Stock Composite Index (SCI). The proposed model consists of two parts: the Emotional analysis model and the LSTM time series learning model. The authors concluded that when stock data and emotional data both are given as input to the LSTM, it can provide better forecasts of opening price [17]. Yang et al. proposed a deep neural network ensemble to predict stock prices of SCI and SZSE composite index. The authors predicted low and high index with an accuracy of 74.15% each for SCI and with an accuracy of 72.34% and 73.95% for SZSE. On the other hand authors are not able to achieve satisfactory results with predictions on close indices. The accuracy of trend predictions on the close index was 49.9% for SCI and 51.1% for SZSE. The authors also showed that the relative error of an ensemble is lower than that of its all component networks [18].

4. Methodology

Historical stock data used in this research work is from Aug 8, 2007 to Nov 5, 2018. This data is pre-processed before using it for training and testing deep recurrent neural networks. The prediction attribute "Next day's Close Price" is added to the data. It is simply Close price shifted back by one position. Then, the data is scaled by using z-score normalization. Configuration of DRNN's used in this research work is 14×50×50×50×50×1. The train/test split of the dataset is done at a 9:1 ratio. The dataset is prepared such that "Open", "High", "Low", "Close", "Trade Volume", "Trade Amount", "High-Low", "Close-Open", "3day MA", "10day MA", "30day MA", "Standard Deviation", "Relative Strength Index", and "William %R" will be used as input attribute and "Next Day's Close Price" will be used as the prediction attribute. Finally, stock prices are predicted and the predicted stock prices are converted back to its original form by using inverse z-score transformation.

5. Result Analysis

5.1 Predicted Stock Prices

In this research work, stock price prediction is made for 7 different look-back periods and for each look back-period 10 experiments are conducted, which results in 210 prediction curves in total. Therefore, only the best-fitted curves are presented here for each DRNN studied in the research work.

Figure 4: Best Fitted Prediction of NIB with MAPE=1.69

Figure 5: Best Fitted Prediction of NIB with MAPE=1.53

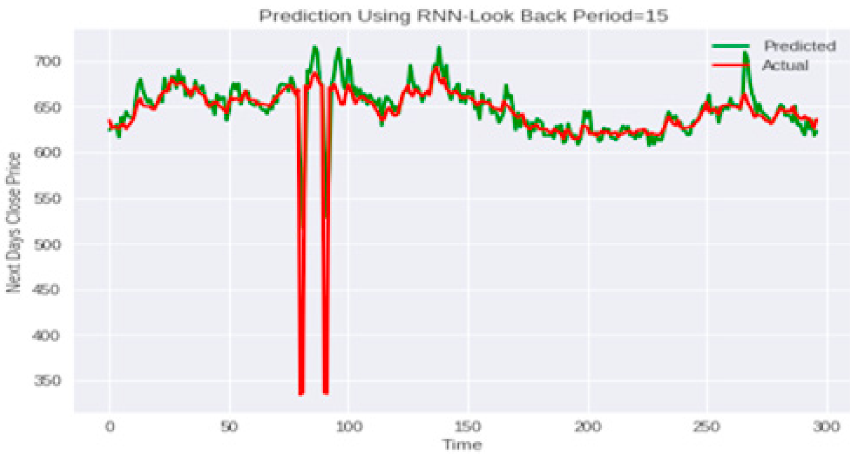

Figure 6: Best Fitted Prediction of NIB with MAPE=1.73

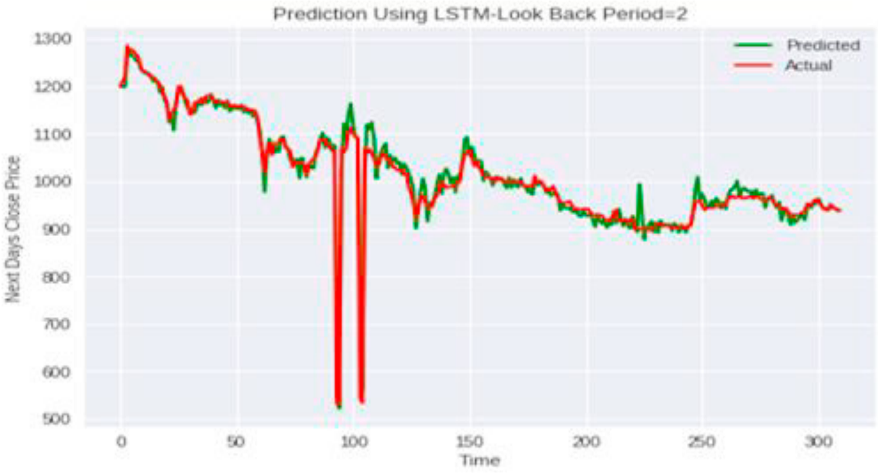

Figure 7: Best Fitted Prediction of NABIL with MAPE=1.04

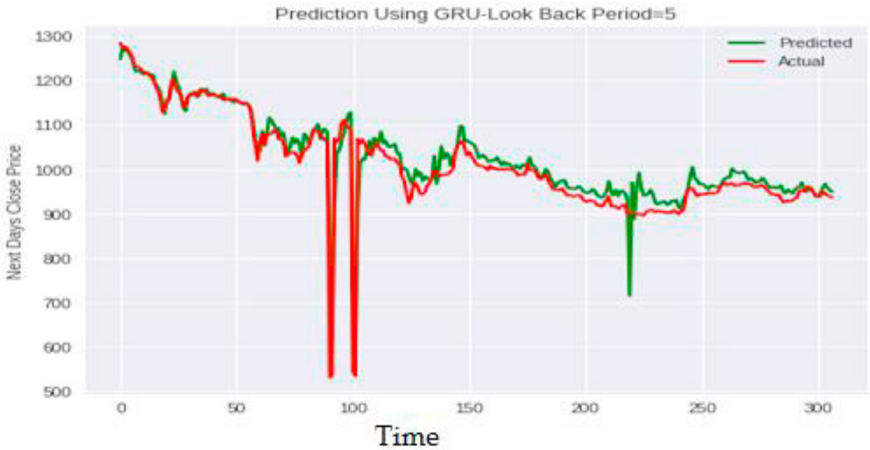

Figure 8: Best Fitted Prediction of NABIL with MAPE=1.67

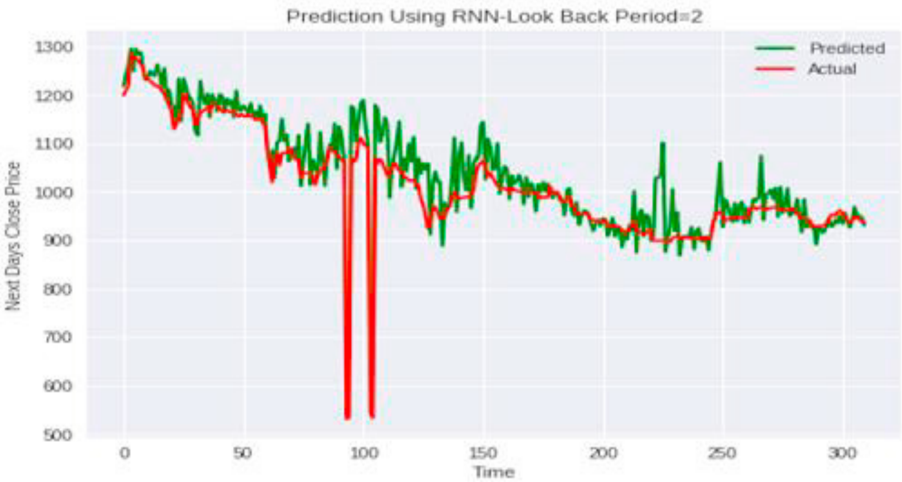

Figure 9: Best Fitted Prediction of NABIL with MAPE=2.42

From figure 4 to figure 9, we can see that Vanilla RNN is not able to predict stock prices of NIB and NABIL as accurately as LSTM and GRU. The best value of Mean Absolute Percentage Errors (MAPEs) obtained with VRNN is 1.73 and 2.42 for NIB and NABIL respectively. On the other hand, LSTM is able to predict stock prices of NIB and NABIL with MAPEs 1.69 and 1.04 respectively and GRU is able to predict stock prices of NIB and NABIL with MAPEs 1.53 and 1.67 respectively. There are some spikes of falling stock prices in the actual stock price of both stocks. It seems like due to inaccurate data. But few incorrect data in the training set or test set cannot affect the prediction model largely or only has the effect that can be ignored. Further analysis of MAPEs obtained during stock price prediction is presented below.

5.2 Result Analysis

In this research work, the value of the look-back period is varied from 2 to 30. Ten experiments are conducted for each value of look-back periods used with deep RNN models and MAPE's are captured.

Figure 10: Minimum MAPEs Obtained for NIB

Figure 11: Maximum MAPEs Obtained for NIB

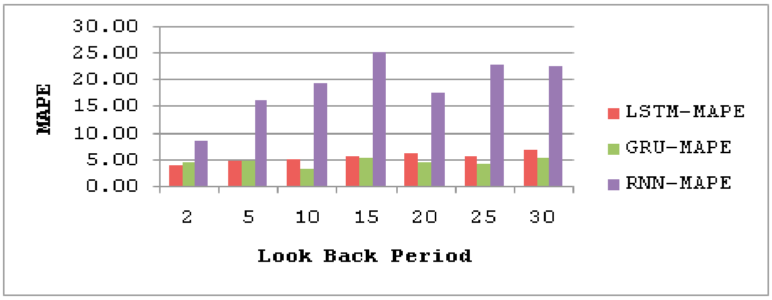

Figure 12: Average MAPEs Obtained for NIB

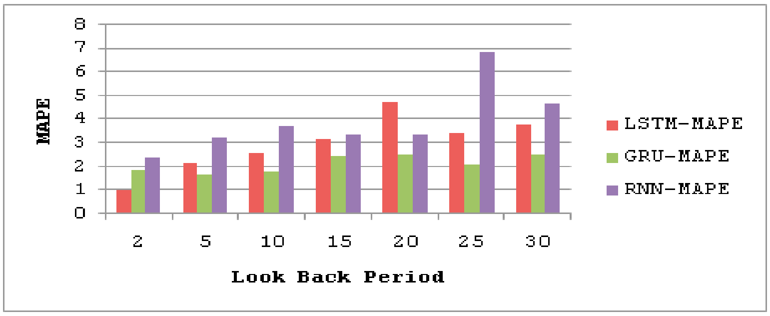

Figure 13: Minimum MAPEs Obtained for NABIL

\

Figure 14: Maximum MAPEs Obtained for NABIL

Figure 15: Average MAPEs Obtained for NABIL

If we look at figure 10, we can see that the minimum values of MAPE obtained while predicting the stock price of NIB using GRU are lower than the minimum values of MAPE obtained with LSTM and VRNN. A similar type of pattern can also be observed in figure 13 while predicting stock prices of NABIL. From the figures, we can see that the lowest values of MAPE while predicting stock prices of NIB using GRU, LSTM, and VRNN are captured at look-back period 10, 2, and 15 respectively. On the other hand, the minimum values of MAPE while predicting stock price NABIL is captured at look-back period 5 for GRU and at 2 for both LSTM and VRNN.

Again, if we look at figure 11, we can see that the maximum values of MAPE obtained while predicting the stock price of NIB using GRU and LSTM are much lower than the maximum values of MAPE obtained while predicting stock price using VRNN. A similar type of pattern can also be observed in figure 14 while predicting the stock price of NABIL. From the above two figures, we can also observe that the highest values of MAPE for predicting stock prices of NIB and NABIL using GRU, LSTM, and VRNN is obtained for look-back periods more than 15.

Finally, if we look at figure 12, we can see that the average values of MAPE obtained while predicting the stock price of NIB using GRU and LSTM is much lower than the average values of MAPE obtained with VRNN. A similar type of pattern can also be observed in figure 15 while predicting stock prices of NABIL. From these two figures, we can also observe that the lowest values of average MAPE for GRU is achieved at look back period 10 for NIB and at look back period 5 for NABIL. But, the lowest values of average MAPE for LSTM are achieved at look back period 2 for both NIB and NABIL. Another interesting pattern that can be seen from figures 13 and 16 is that the average values of MAPE obtained with LSTM are less than average values of MAPE obtained with GRU for look back periods 2 and 5. But, for the look back periods more than 5, the average values of MAPE obtained with GRU is lower.

6. Conclusion

This research work has shown that GRU and LSTM are able to outperform VRNN in predicting stock prices. From the analysis of average MAPEs obtained from GRU and LSTM, it can be concluded that GRU performs slightly better than LSTM because the average value of average MAPEs obtained with GRU for predicting the stock price of NIB and NABIL is 4.74 and 4.71 respectively but it is 5.58 and 5.06 respectively for LSTM. The reason behind the better performance of GRU over LSTM is the moderate amount of data available for training the prediction model. This fact is further supported by the research work performed by Weiss et al. [7]. Another fact that is in favour of GRU over LSTM is that GRU has fewer parameters than LSTM and hence it is faster to train and predict results with GRU than LSTM [8]. Besides this, the research work has performed a novel analysis of the look-back period and identified its best value that is suitable for LSTM and GRU. From the analysis of MAPEs obtained with LSTM and GRU, it can be concluded that the suitable value of the look-back period used with LSTM networks should be less than 5 and suitable value of the look-back period used with GRU networks should be between 5 to around 10. The results showed that using the look-back period value more than 15 is just wastage of model training time and prediction time. In addition, using larger values of the look-back period may result in poor stock price prediction performance.

There are many ways in which this research work can be further extended. Analysis of the look-back period can be done for long term stock price prediction. There are many fundamental technical indicators that can be included in the stock data sets for better stock price prediction. PE ratio, PB ratio, base rate, exponentially weighted moving averages (EMWA), rate of change (ROC), momentum indicators, etc. are some of the indicators among them. It may also be possible to include Elliot wave principles in predicting stock prices, which may help to improve the prediction performance further.

References

[1] Refenes, A. N., Zapranis, A., and Francis, G.(1994) "Stock performance modeling using neural networks: a comparative study with regression models." Neural Networks 7(2): 375-388.

[2] Goodfellow, I., Bengio, Y., and Courville, A. (2016) "Deep Learning (Adaptive Computation and Machine Learning series). Ebook." The MIT Press, Cambridge.

[3] Gao, X., Zhang, J., and Wei, Z. (2018) "Deep Learning for Sequence Pattern Recognition. IEEE 15th International Conference on Networking, Sensing and Control (ICNSC)." IEEE, Zhuhai 1-1.

[4] Haykin, S. (2008) "Neural Networks and Learning Machines." Third Edition, Pearson Education , New York.

[5] Hochreiter, S., and Schmidhuber, J. (1997) "Long short-term memory." Neural Computation 9(8) , 1735-1780.

[6] Cho, K., Bart, V. M., Caglar, G., Dzmitry, B., Bougares, F., Schwenk, H., and Bengio, Y. (2014) "Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation. Conference on Empirical Methods in Natural Language Processing (EMNLP)." ACL, Doha, 1724-1734.

[7] Weiss, G., Goldberg, Y., and Yahav, E. (2018) "On the Practical Computational Power of Finite Precision RNNs for Language Recognition. 56th Annual Meeting of the Association for Computational Linguistics." ACL, Melbourne, 740-745.

[8] Minh, D. L., Sadeghi-Niaraki, A., Huy, H. D., Min, K., Hyeonjoon, M. (2018) "Deep Learning Approach for Short-Term Stock Trends Prediction Based on Two-Stream Gated Recurrent Unit Network." IEEE Access 6, 55392-55404.

[9] Hiransha, M., Gopalakrishnan, E., Menon, V., and Soman, K. (2018) "NSE Stock Market Prediction Using Deep-Learning Models. International Conference on Computational Intelligence and Data Science (ICCIDS), Procedia Computer Science." Haryana, 1351-1362.

[10] Selvin, S., Vinayakumar, R., Gopalakrishnan, E., Menon, V., and Soman, K. (2017) "Stock price prediction using LSTM, RNN and CNN-sliding window model. International Conference on Advances in Computing, Communications and Informatics (ICACCI)." IEEE, Udupi, 1643-1647.
[11] Khare, K., Darekar, O., Gupta, P., and Attar, A. Z. (2017) "Short Term Stock Price Prediction Using Deep Learning. IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)." IEEE, Bangalore, 482-486.
[12] Huynh, H. D., Dang, L., and Duong, D. (2017) "A New Model for Stock Price Movements Prediction Using Deep Neural Network. Eighth International Symposium on Information." ACM, Nha Trang, 57-62.
[13] Ding, X., Zhang, Y., Liu, T., and Duan, J. (2015) "Deep Learning for Event-Driven Stock Prediction. Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence." ACM, Singapore, 2327-2333.
[14] Vargas, M., Lima, B., and Evsukoff, A. (2017) "Deep learning for stock market prediction from financial news articles. In: IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA)." IEEE, Annecy, 60-65.
[15] Singh, R., and Srivastava, S. (2017) "Stock prediction using deep learning." Multimedia Tools and Applications 76(18), 18569-18584.
[16] Gunduz, H., Cataltepe, Z., and Yaslan, Y. (2017) "Stock market direction prediction using deep neural networks. 25th Signal Processing and Communications Applications Conference (SIU)." IEEE, Turkey, 1-4.
[17] Zhuge, Q., XU, L., and Zhang, G. (2017) "LSTM Neural Network with Emotional Analysis for prediction of stock price." Engineering Letters 25(2), 167-175.
[18] Yang, B., Gong, Z., and Yang, W. (2017) "Stock Market Index Prediction Using Deep Neural Network Ensemble. 36th Chinese Control Conference." IEEE, Dalian, 3882-3887.