Proceedings of 11th IOE Graduate Conference
Peer Reviewed
ISSN: 2350-8914 (Online), 2350-8906 (Print)
Year: 2022 Month: March Volume: 11

A CNN-BGRU Method for Stock Price Prediction

Janak Kumar Lal [a], Arun Kumar Timalsina [b]

[a, b] *Department of Electronics and Computer Engineering, Pulchowk Campus, IOE, TU, Nepal*
[a] 075msice009.janak@pcampus.edu.np, [b] t.arun@ioe.edu.np

Abstract

With the rapid development in technology and digitization of financial market in recent years, an increasing number of people have begun to invest in the stock market. Accurately anticipating the change in stock price can reduce stock investors' investment risk while also increasing their investment return. Stock price prediction is frequently a nonlinear time series prediction due to the stock market's volatility.

This paper presents a CNN-BGRU (Convolutional Neural Network - Bidirectional Gated Recurrent Unit) model for stock price prediction. The presented model includes a 1-D CNN followed by bidirectional GRU to predict stock price. This model is used to predict stock price for short-term (next trading day), mid-term (closing price after 15 days) and long term (closing price after 30 days). The proposed model has been compared with unidirectional RNN (Recurrent Neural Network) model such as: LSTM (Long Short Term Memory) and GRU and bidirectional models such as: BLSTM (Bidirectional Long Short Term Memory) and BGRU in terms of MAPE (Mean Absolute Percentage Error) and training time.

Keywords

CNN, BGRU, Stock price prediction

1. Introduction

The ability to forecast the stock market remains a crucial component of a successful investing strategy that generates above-average returns. Accurate stock price predictions are critical for building trading strategies and hedging against probable market risks, allowing speculators and arbitrageurs to profit from stock index trading.

This paper presents a CNN-BGRU model for stock price prediction. The present model includes a 1-D CNN followed by bidirectional GRU to predict stock price. The 1-D CNN helps to deeply mine the data features while bidirectional GRU can be trained to keep long term information through time and remove the irrelevant information. This model is used to predict stock price for short-term (next trading day), mid-term (close price after 15 days) and long term (closing price after 30 days). The experiments have been conducted for short term, mid-term and long term so that the proposed model can be useful for the investors looking for high frequency trading as well as investors looking for long term investment.

The proposed model has been compared with unidirectional RNN models such as: LSTM and GRU and bidirectional models such as: BLSTM and BGRU in terms of MAPE and training time. The dataset used is the closing price of 26 commercial banks listed in NEPSE (Nepal Stock Exchange) from 2018-01-01 to 2021-12-29. This dataset is publicly available on Kaggle.com. The stock price of 26 commercial banks have different ranges, mean, standard deviation and they fluctuate differently over time. The experiment on this larger and diverse data helps us to analyze how the proposed model performs against other RNN models in general and confirms that the proposed model is better suited for stock price forecast. Effect of different look-back periods have also been analysed for short term, mid term and long term prediction.

The organization of the paper is as follows. In section 2, related works in the field of stock price forecasting using deep neural networks has been presented. In section 3, the methodology used is described. The result on the performance of the models are presented in section 4 which is followed by conclusion in section 5.

2. Literature Review

Deep neural networks have been used to forecast the index price of the Singapore stock market using the

FTSE (Financial Times Stock Exchange) Straits Time Index (STI) by [1]. The effect of varying the number of steps of the forecast model was investigated in this paper. Analysis of look back period for stock price prediction with rnn variants has been conducted by [2]. Evaluation of bidirectional LSTM for short and long term stock market prediction was done in [3]. Where authors have shown that BLSTM networks demonstrated better performance and convergence for both short- and long-term predictions. LSTMs followed by BGRUs proved to be the best models in predicting the African stock markets as shown by [4]. Online financial news and historical data are used as an input of BGRU model in [5] and the given model has outperformed LSTM and GRU models. The CNN-BiLSTM-AM method proposed by [6] shows that a hybrid of CNN and RNN improves the accuracy of prediction.

3. Methodology

3.1 CNN-BGRU Model

This section outlines the CNN-BGRU model that has been proposed.

CNN is extensively employed in feature engineering because it has the property of focusing on the most obvious feature in the line of sight. BGRU is extensively used in time series analysis since it has the property of learning according to time sequence. A stock forecasting model based on CNN-BGRU is developed based on the properties of CNN and BGRU.

The proposed CNN-BGRU model along with its parameters is shown in Figure 1. This model a one-dimensional CNN with 100 kernels each of size 2 units. The padding of one-dimensional CNN has been set such that the output of CNN has the same dimension as its input. The 'relu'(Rectified Linear Unit) has been used as the activation function for one-dimensional CNN. The MaxPool layer has the pool size of two and stride of one. The BGRU is a combination of two GRU each with 50 units. The TanH is used as activation function here. The output of BGRU is sent to final layer with single neuron.

3.2 The CNN-BGRU training process

The training process of CNN-BRGU model is shown the Figure 2. The main steps are as follows:

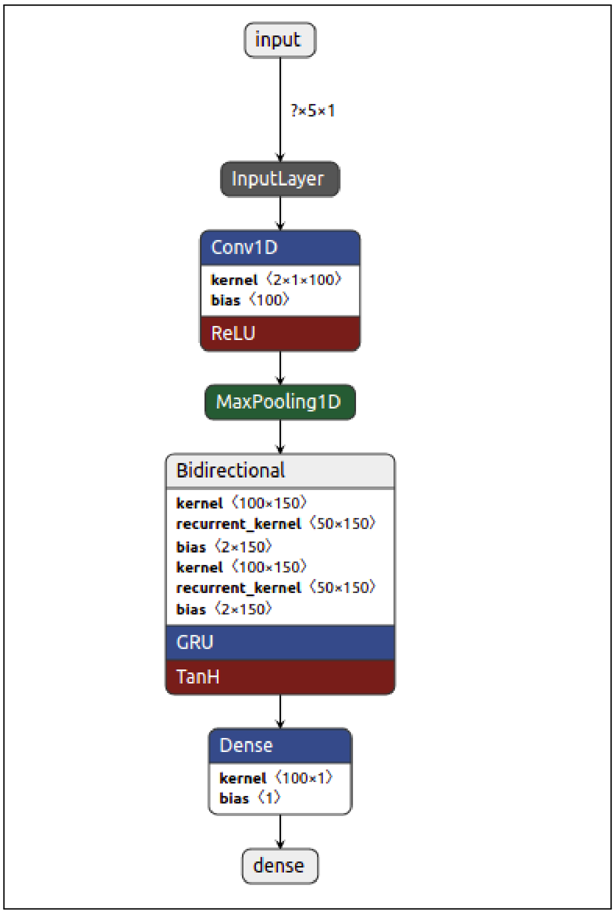

Figure 1: The CNN-BGRU model

1. Input data: The data of close price of commercial banks listed in NEPSE is used. The data from 2018-01-01 to 2021-12-29 (904 trading days) is used. This data is used both as input and target output. The target output is chosen based on the type of prediction such as: short term (next day), mid term (close price after 15 days) and long term (close price after 30 days).

2. Data standardization: As there is large gap in the input data, in order to better train the model, the z-score standardization method is adopted to standardize the input data as shown below:

   The data are standardized according to the following formula.

$$y_i = \frac{x_i - \bar{x}}{s} \tag{1}$$

   where y_i is the standardized value, x_i is the input data, $\bar{x}$ is the average of the input data, and s is the standard deviation of the input data. This method of standardization was instead of min-max standardization as z-score standardization does not limit the output of forecast value between minimum and maximum of data points.

3. After standardization the input data is shaped according to the look-back period.
4. Network initialization: The weights and biases of each layer of CNN-BGRU are initialized.
5. CNN layer calculation: The input data is passed through 1-dimensional CNN layer, the feature extraction of the input data is carried out and the output value is obtained.
6. BGRU layer calculation: The output of CNN layer is fed as an input to BGRU layer and the output is obtained.
7. Finally output of BGRU is fed to a single neuron to get the final value of our model.
8. Since it is a regression task we are using 'mean-squared-error' as loss function. The 'adam' [7] optimizer is being used in the training phase as it is straightforward to implement, is computationally efficient, has little memory requirements, is invariant to diagonal re-scaling of the gradients, and is well suited for problems that are large in terms of data and/or parameters. The learning rate of 'adam' optimizer is set to 0.001, the exponential decay rate for the 1st moment estimates is set to 0.9 and the exponential decay rate for the 2nd moment estimates is set to 0.999. The model is trained for one hundred epochs with the batch size of 64.

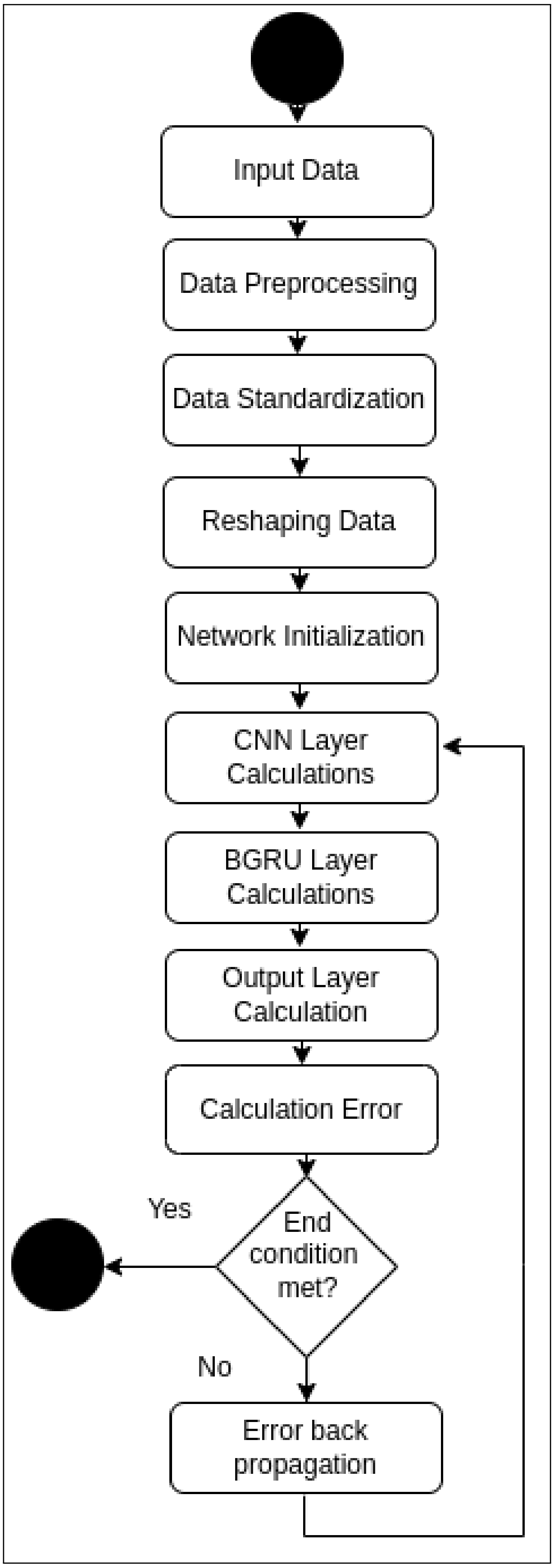

Figure 2: The CNN-BGRU Training Process

3.3 The CNN-BGRU forecast process

The pre-condition for CNN-BGRU forecast is that CNN-BGRU has completed its training. The forecast process is depicted in Figure 3.

The main steps in CNN-BGRU forecast are as follows:

1. Input data: The input data required for the prediction are fed to the network in a form suitable to the given network, which involves standardization and reshaping the data.
2. Forecast: CNN-BGRU model gives a certain output data after receiving the input.

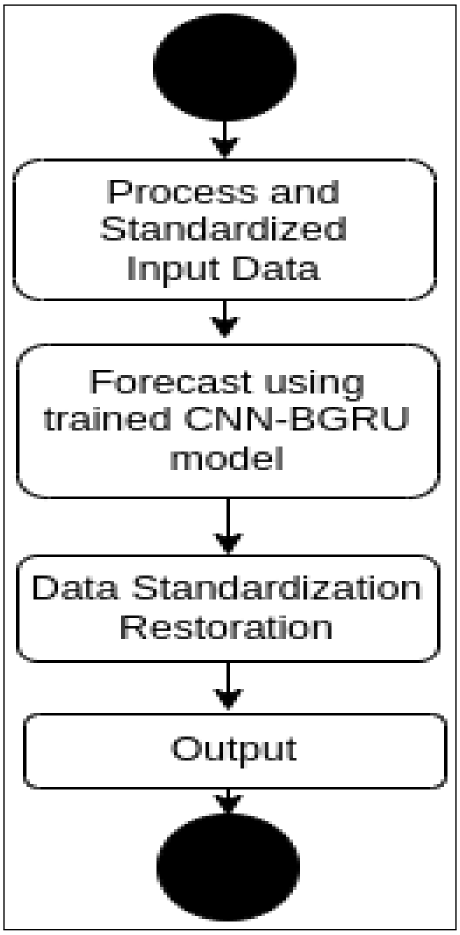

Figure 3: The CNN-BGRU Forecast Process

3. Standardization Restoration: The output value obtained from the CNN-BGRU is standardized value. The standardized value is restored using the following formula.

$$x_i = y_i * s + \bar{x} \quad (2)$$

4. Output result: The restored results are outputted to complete the forecast process.

Calculation of MAPE:

1. After receiving the forecast value it is compared with the actual value and MAPE is calculated according to the following formula.

$$MAPE = \frac{100\%}{n} \sum_{t=1}^{n} \left| \frac{A_t - F_t}{A_t} \right| \quad (3)$$

   Where A_t is actual value and F_t is forecast value.

2. The MAPE obtained after comparing the actual value with the predicted value for various lookback period, to predict the short term, mid term and long term closing price of all commercial banks listed in NEPSE are stored and analysed. The result and the analysis are presented in the following section.

3.4 Experimental tools and techniques

In order to prove the effectiveness of the proposed CNN-BGRU model, it is compared with LSTM, GRU, BLSTM and BGRU using same data and platform. The performance of various models with different look-back period, for different nature of forecast was compared on the basis of MAPE and training time. Total number of models trained for four different RNN models were 936. This included 3 look-back period for four RNN models for three different look-back periods and for stock price of 26 commercial banks listed in NEPSE. To compare the performance of proposed CNN-BGRU model with other RNN models and to gain insight of how look-back periods effect different nature of prediction, the proposed model was trained with 6 different look-back periods for three different scenarios with data of 26 commercial banks. The average of MAPE and training time were calculated for each model for different nature of prediction with different look-back periods.

All the experiments are carried out on Google Colaboratory. It has the following specifications:

- Ubuntu 18.04.5 LTS
- RAM: 12.69 GB
- CPU Model name: Intel(R) Xeon(R) CPU @ 2.20GHz
- Cores : 2

The programming language, libraries and their versions used in this project are listed below:

- Python 3.7.12
- Numpy 1.19.5
- Pandas 1.3.5
- Sklearn 1.0.2
- Tensorflow 2.7.0
- Kears 2.7.0

4. Results

4.1 Short term prediction

The average MAPE obtained for next day prediction after training all the models using various look-back periods for all the commercial banks listed in NEPSE is shown in Figure 4.

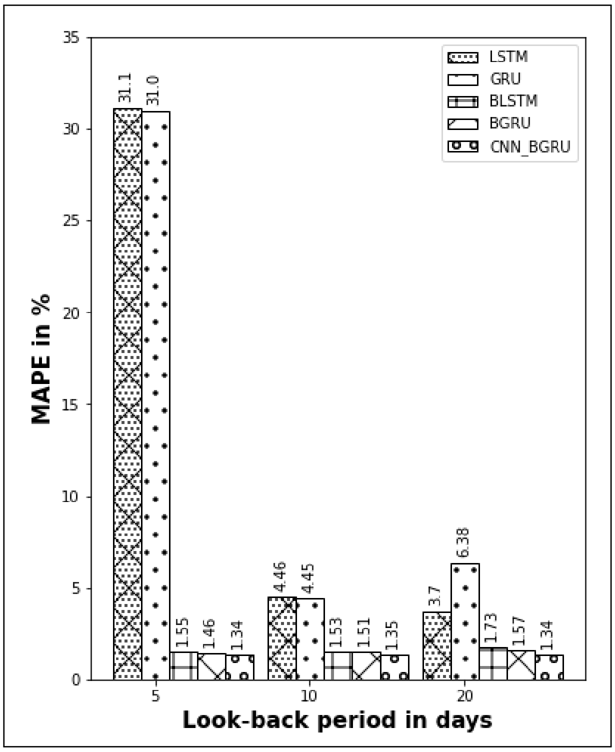

Figure 4: Comparison of various models for short-term prediction for different look-back periods

It is clear from the Figure 4 that bidirectional RNN models perform better than unidirectional RNN models. Our CNN-BRGU model outperforms other four models for all three look-back periods. The best MAPE obtained for short-term forecast is 1.34% for CNN-BGRU model for look-back period of 5 days and 20 days. It can be concluded that the increase in look-back period do not have high impact for short-term prediction, for bidirectional RNN models and our hybrid CNN-BGRU model.

Table 1: Average Training time in seconds for short term forecast using different models for different look-back period

	Models				
Lookback	LSTM	GRU	BLSTM	BGRU	CNN-BGRU
5	10.09	10.49	12.73	14.50	19.38
10	14.36	15.49	19.680	20.896	28.26
20	21.50	22.53	32.667	33.48	54.47

From the Table 1, it can be concluded that the proposed CNN-BGRU model takes more time to train. The look-back period does not have a significant impact on MAPE in this case but longer the look-back period results in longer training time.

4.2 Mid term prediction

The average MAPE obtained for mid-term forecast after training all the models using various look-back periods for all the commercial banks listed in NEPSE is shown in Figure 5.

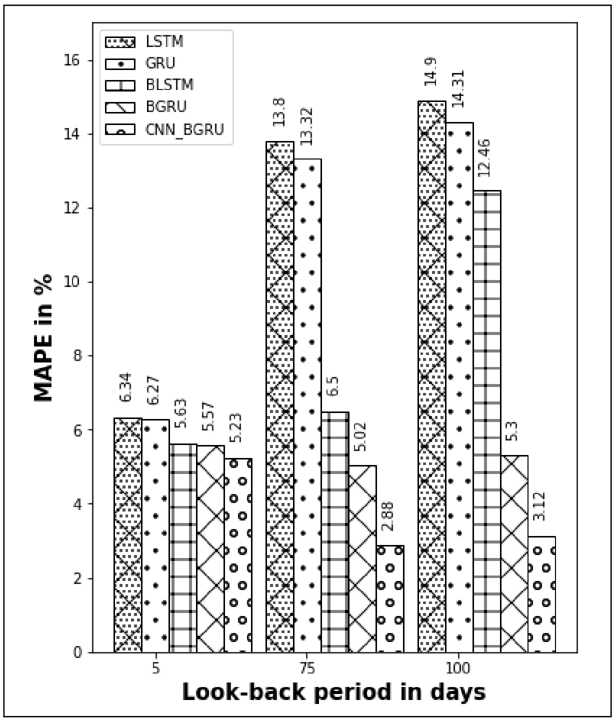

Figure 5: Comparison of various models for mid-term prediction for different look-back periods

The increase in look-back period has different effect on different models. The best result is obtained for mid-term forecast when CNN-BGRU model is used with the look-back period of 75 days. There is decrease in average MAPE for CNN-BGRU model with increase look-back period till 75 days after which average MAPE period seems to increase.

Table 2: Average Training time in seconds for mid term forecast using different models for different look-back period

	Models				
Look-back	LSTM	GRU	BLSTM	BGRU	CNN-BGRU
5	12.29	13.07	15.217	16.78	18.14
75	50.77	52.336	85.796	87.999	133.41
100	63.208	68.260	107.246	109.616	166.69

For mid-term prediction, LSTM model with look-back period of 5 days has least training time while CNN-BGRU model has longest training time for all three different look-back periods used in the experiment.

4.3 Long term prediction

The average MAPE obtained for long-term forecast after training all the models using various look-back periods for all the commercial banks listed in NEPSE is shown in Figure 6.

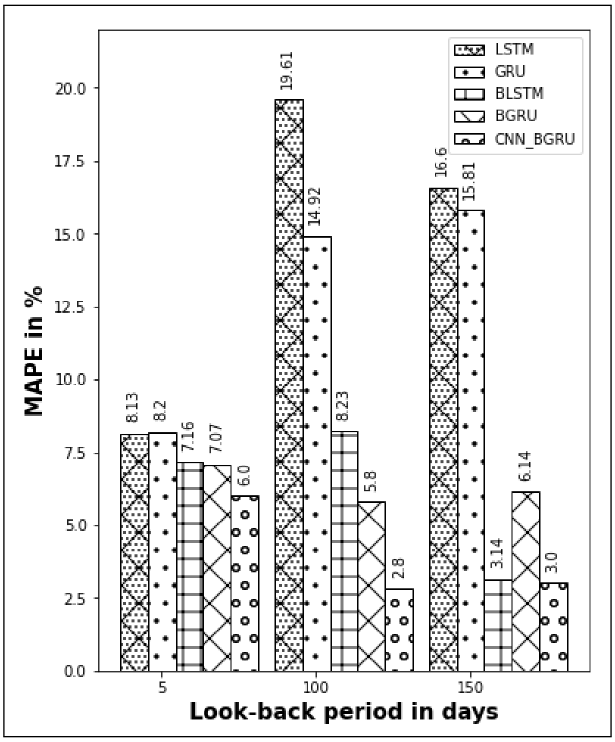

Figure 6: Comparison of various models for long-term prediction for different look-back periods

The increase in look-back period has different effect on different models. The best result is obtained for mid-term forecast when CNN-BGRU model is used with the look-back period of 100 days. There is decrease in average MAPE for CNN-BGRU model with increase look-back period till 100 days after which average MAPE period seems to increase.

Table 3: Average Training time in seconds for long term forecast using different models for different look-back period

	Models				
Look-back	LSTM	GRU	BLSTM	BGRU	CNN-BGRU
5	10.12	10.26	12.16	13.40	19.38
100	69.89	69.09	115.55	117.57	178.76
150	64.193	65.695	105.57	111.1104	243.89

For long-term prediction, LSTM model with look-back period of 5 days has least training time while CNN-BGRU model has longest training time for all three different look-back periods used in the experiment.

4.4 Effect of look-back period on MAPE and train time for proposed model

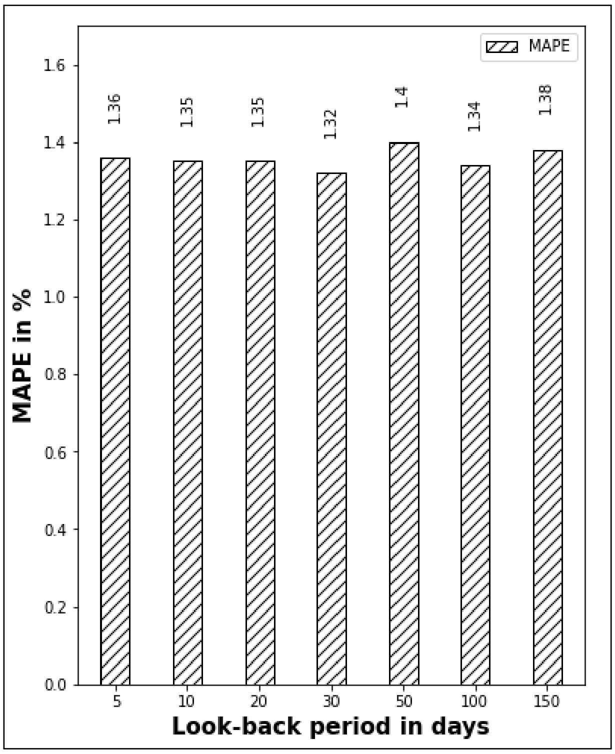

Figure 7: Effect of look-back in MAPE on short term forecast for CNN-BGRU model

From Figure 7 it can be observed that lowest MAPE of 1.32% was achieved when look back period was 30 but it is not significantly different than MAPE achieved with look-back period of 5, 10, 20 and 100 which are 1.36%, 1.35%, 1.35% and 1.34% respectively.

Form Figure 8 it can be observed that the training time increases significantly with increase in look-back period.

So, from Figure 7 and Figure 8 it can be concluded that using look-back period of 5 is best when one aims to predict the closing price of stock of following day.

From Figure 9 the effect of look-back period can be clearly observed. The MAPE decreases with the increase in look-back period till look-back period of 75 days is reached. After the the progress seems to saturate. The training time increases significantly with increase in look-back period. The training time of look-back period of 30 days is greater than that of 50 days period. This might be due to problem with

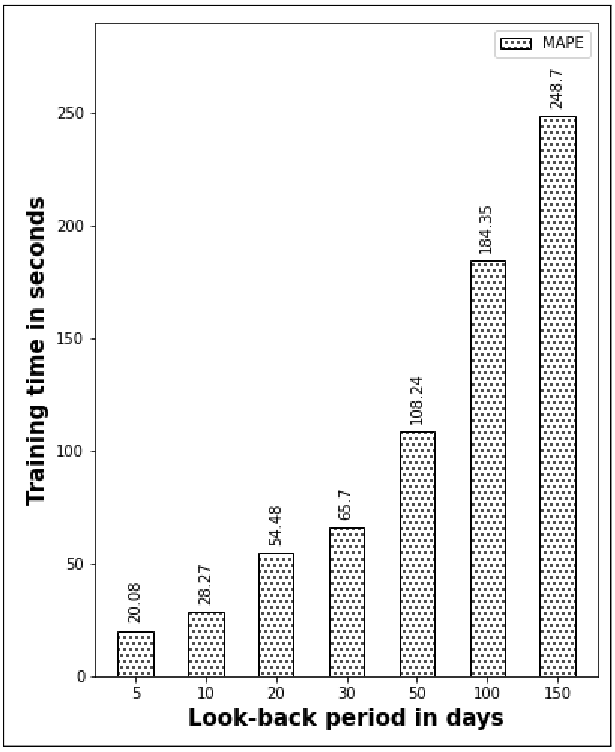

Figure 8: Effect of look-back period on training time for short term prediction for CNN-BGRU model.

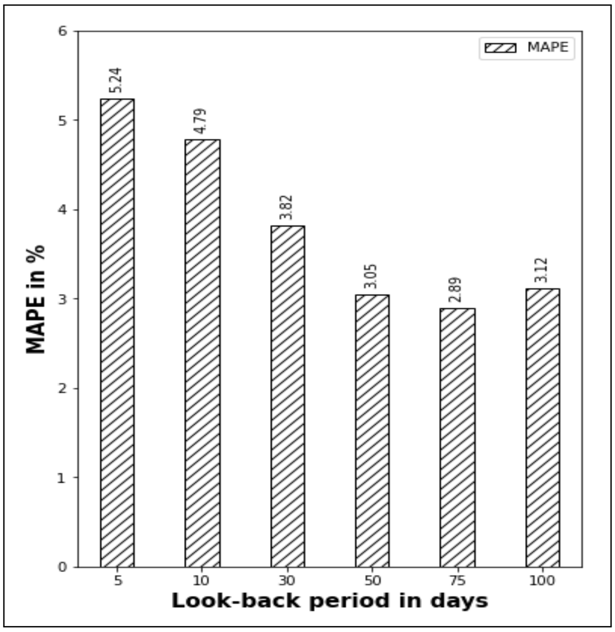

Figure 9: Effect of look-back on MAPE on mid term forecast for CNN-BGRU model

network during that time period as whole experiment process was carried out using Google Colab. From Figure 9 and Figure 10 it can be concluded that look-back period of 75 days is best when one want to predict closing price of stock after 15 days.

From Figure 11 the effect of look-back period can be

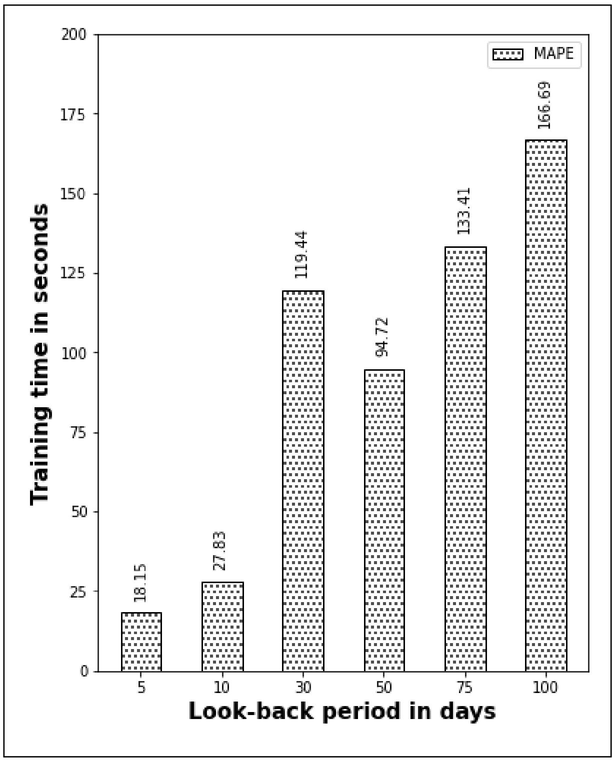

Figure 10: Effect of look-back in training time for mid term prediction of CNN-BGRU

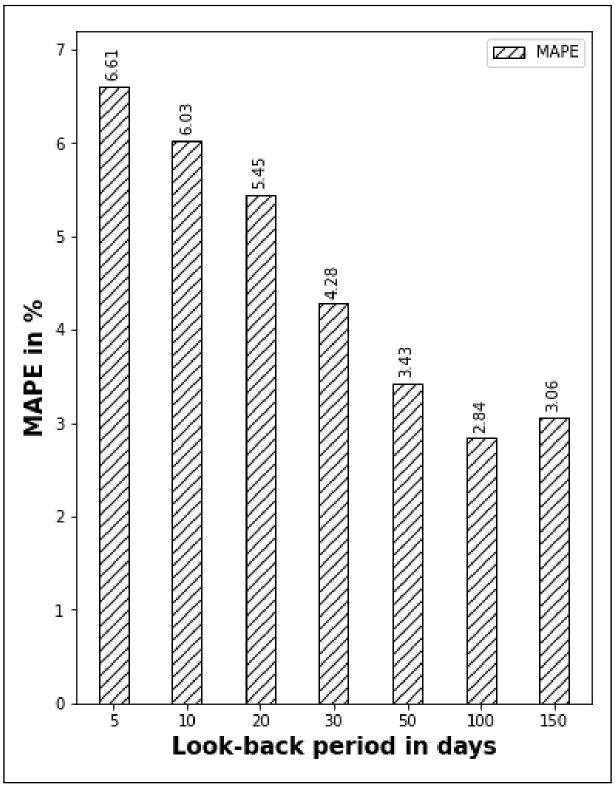

Figure 11: Effect of look-back on MAPE on long term forecast for CNN-BGRU model.

clearly observed. When the proposed model is trained with the look-back period of 100 days we get average

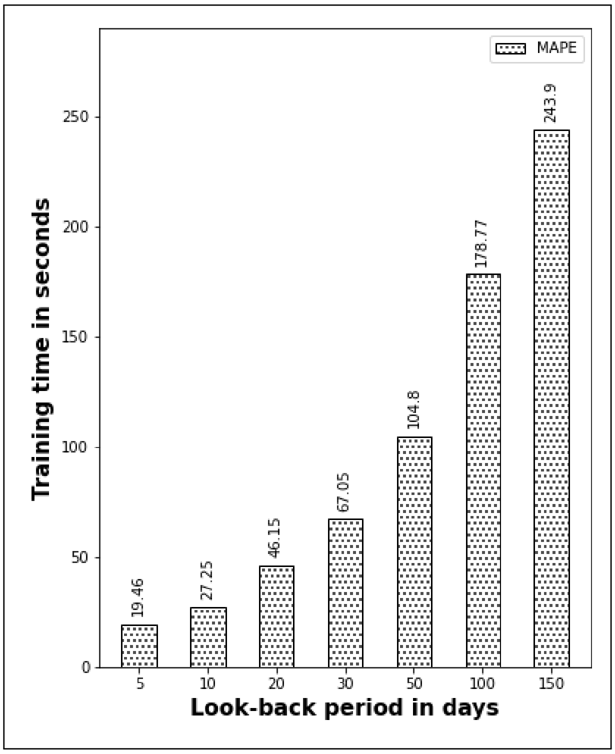

Figure 12: Effect of look-back on training time for long term prediction of CNN-BGRU.

error of nearly three percentage and it takes around 3 minutes to train the model. The MAPE decreases with the increase in look-back period till look-back period of 100 days is reached. After the the progress seems to saturate. The training time increases significantly with increase in look-back period.

From Figure 11 and Figure 12 it can be concluded that look-back period of 100 days is best when one want to predict closing price of stock after 30 days using the proposed model.

5. Conclusion

This paper proposes a CNN-BGRU method to predict the stock closing price of the next day, after 15 days and after 30 days. This method uses the close price of the stock data as input. CNN is used to extract the features of the input data. BGRU is used to learn and predict the extracted feature data. The experimental results show that the CNN-BGRU has the lowest MAPE when compared to LSTM, GRU, BLSTM and BGRU. From the experimental results it can also be concluded that increase in size of look-back period does not improve the prediction accuracy for short term prediction. This might be because stock price of following day does not fluctuate too much from present day unless company decides to split the stocks. This is also because rules of NEPSE does not allow price of stock to fluctuate more than 10% in a trading day.

Although the size of the look-back period has a significant impact for mid term prediction and long term prediction. There is trade-off between accuracy of forecast and amount of time taken by model to train. For mid-term forecast i.e forecast of stock prices after 15 trading days the look-back period of 75 days is recommended and for forecast of stock prices after 30 trading days the look-back period of 100 days is recommended.

References

[1] Bang Xiang Yong, Mohd Rozaini Abdul Rahim, and Ahmad Shahidan Abdullah. A stock market trading system using deep neural network. In *Asian Simulation Conference*, pages 356–364. Springer, 2017.

[2] Arjun Singh Saud and Subarna Shakya. Analysis of look back period for stock price prediction with rnn variants: A case study on banking sector of nepse. *Procedia Computer Science*, 167:788–798, 2020.

[3] Khaled A Althelaya, El-Sayed M El-Alfy, and Salahadin Mohammed. Evaluation of bidirectional lstm for short-and long-term stock market prediction. In *2018 9th international conference on information and communication systems (ICICS)*, pages 151–156. IEEE, 2018.

[4] Dennis Murekachiro, Thabang Mokoteli, and Hima Vadapalli. Predicting emerging and frontier stock markets using deep neural networks. In *Proceedings of SAI Intelligent Systems Conference*, pages 899–918. Springer, 2019.

[5] Huy D Huynh, L Minh Dang, and Duc Duong. A new model for stock price movements prediction using deep neural network. In *Proceedings of the Eighth International Symposium on Information and Communication Technology*, pages 57–62, 2017.

[6] Wenjie Lu, Jiazheng Li, Jingyang Wang, and Lele Qin. A cnn-bilstm-am method for stock price prediction. *Neural Computing and Applications*, pages 1–13, 2020.

[7] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.