Journal of IoT in Social, Mobile, Analytics, and Cloud (ISSN: 2582-1369)
www.irojournals.com/iroismac/

Stock Index Prediction with Financial News Sentiments and Technical Indicators

Subin Maharjan[1], Subarna Shakya[2]

[1,2]Department of Electronics and Computer Engineering, Institute of Engineering, Tribhuvan University, Kathmandu, Nepal

E-mail: [1]076mscsk016.subin@pcampus.edu.np, [2]drss@ioe.edu.np

Abstract

The price of a stock in the market can be influenced by many factors, out of which the sentiment of the investors plays a vital role. Most often, the sentiment of investors depends on the sentiment of the news headlines. Therefore, news headlines also play an important role in the fluctuation of the stock index. This paper uses the combination of Bidirectional Encoder Representations from Transformers (BERT) and Bidirectional Gated Recurrent Unit (BiGRU) algorithms for the prediction of news sentiment scores based on national news headlines and financial news data. Technical indicators like Relative Strength Index (RSI), Exponential Moving Average (EMA), Moving Average Convergence Divergence (MACD), Stochastic Oscillator along with normal stock indicators like 'Date', 'Open', 'Close', 'High', 'Low' and 'Volume' data can be used to predict the short-term momentum of the stock value. This paper uses the BiGRU algorithm to predict the stock index value (a) with technical indicators only and (b) with technical indicators and news sentiment scores. Keeping all the hyperparameters constant, the BiGRU algorithm provided better prediction results when news sentiment scores were added to the dataset along with technical parameters as an input.

Keywords: BERT, BiGRU, Stock, News Sentiments, Technical Indicators, Prediction

1. Introduction

Making money is the common goal of every trader on the stock market but because of the stock market's volatility, profit is not always guaranteed; there could also be significant losses. Investing without sufficient expertise could lead to unanticipated loss. But there is also a strong probability of making money from stocks if one can analyze the market trend and make investments utilizing a sound strategy. Technical analysis techniques are used by traders or investors who make short-term investments. Technical analysis is unconcerned

DOI: https://doi.org/10.36548/jismac.2022.3.006
Received: 03.08.2022, received in revised form: 22.08.2022, accepted: 09.09.2022, published: 19.09.2022

with a company's basic metrics like position in the market, financial reports, profit or loss, assets, liabilities, earning per share (EPS), price to earnings (P/E) ratio, dividend history etc.

To forecast future prices, technical analysis looks at historical volume and price patterns. To analyze the market's structure and forecast whether the stock price will rise or decline in the near future, technical analysts utilize a variety of charts, trend lines, moving averages etc. News sentiment, in addition to technical indications, is also a key factor in how much the market price changes. This paper also focuses on short term price prediction of the stock based on news sentiment scores and technical indicating parameters.

This paper uses the combination of BERT-BiGRU algorithm for predicting news sentiment scores from the news title. Predicted sentiment scores are then added to the dataset containing general parameters of stock like date, open, close, high, low and volume and technical indicators like Relative Strength Index (RSI), Exponential Moving Average (EMA), Moving Average Convergence Divergence (MACD), Stochastic Oscillator. The BiGRU algorithm is then used for stock index value prediction with and without financial news sentiment score as an input.

1.1 News Sentiment Analysis

The stock market patterns are undoubtedly influenced by news headlines [1]. The majority of the time when there is good news, like the announcement of a bonus or financial metrics increasing in the quarterly report of the company, the increment in the price of the stock is seen. Additionally, it has been seen that unfavorable news, such as lower corporate growth or the inability to pay bonuses for the fiscal year, is a significant factor in the company's stock price decline.

Sentiment analysis is a component of Natural Language Processing (NLP). The main function of sentiment analysis is to identify the emotion (for instance, positive, negative, or neutral emotions) present in a particular textual sequence. Sentiment analysis algorithms can classify the given text into the required class based on its sentence structure and has several applications. For instance, e-commerce websites can examine the sentiment of customer feedback to determine whether they are satisfied or not. Businesses can improve their product by analysing the sentiment of the comments provided by their customers. Similarly, we can use news titles or headlines to find out if it gives positive, negative or neutral sentiment.

1.1.1 Bidirectional Encoder Representations from Transformers (BERT).

Bidirectional Encoder Representations from Transformers or BERT is a transformer based bidirectional deep learning model pretrained on 2500 M words form Wikipedia and 800 M words from Books Corpus with 2 variants: Base with 12 encoders and Large with 24 encoders [2]. As BERT can acquire the contextual representation from both ends of phrases, it differs slightly from other language models [3].

We can fine-tune pretrained BERT to perform better on NLP tasks like sentiment analysis, named entity recognition etc. by using fewer resources on smaller datasets with an additional output layer [2]. Compared to conventional models like word2vec, BERT has more advantages because word2vec has a fixed representation for each word regardless of context but BERT gives different representation tothe same word based on the context of the sentence [4].

1.1.2 Bidirectional Gated Recurrent Unit (BiGRU)

Bidirectional Gated Recurrent Unit (BiGRU) is the combination of 2 Gated Recurrent Unit (GRU) layers. GRU is a variant of Recurrent Neural Network (RNN). RNN has short-term memory problems due to the vanishing gradient problem. If a sequence fed to RNN is very long, it will be hard for RNN to carry the information from previous steps and might forget important information in the network. GRU solves the short-term memory problem of RNN. Out of 2 layers, 1^{st} GRU layer processes the data in the front direction while the other in backward direction.

1^{st} GRU will process the information from first to last and vice-versa and at any given time, it will have information of both front and back [5]. The quantity of trainable parameters in BiGRU leads to an increase in the amount of information available to the network, giving it better comprehension of context and the ability to more accurately predict the long sequence of data [5].

1.2 Technical Indicators

Technical indicators are the instruments that forecast future stock trend based on historical data. Different kinds of technical indicators can be used to forecast whether the market will experience an uptrend or a downtrend in the near future for the stock price. This paper uses following technical indicators:

1.2.1 Relative Strength Index (RSI)

Relative Strength Index forecasts future price movement using historical price data. The RSI indicator shows an overbought or oversold condition of stock over a period of time. The RSI value ranges from 0 to 100. The stock is said to be oversold when the RSI is lower than 30 and overbought when the RSI greater than 70. RSI is given by following equation:

$$RSI = 100 - \frac{100}{1 + \frac{Avg.\ Gain}{Avg.\ Loss}}$$

where,

$$Avg.Gain/Loss = \frac{Sum\ of\ all\ the\ gain/loss\ over\ N\ days}{N}$$

1.2.2 Exponential Moving Average (EMA)

Exponential Moving Average predicts upward or downward trend of the market. EMA is given by following formula:

$$EMA = ((Today's\ price - EMA\ of\ last\ day) * Multiplier) + EMA\ of\ last\ day$$

where,

$$Multiplier = \frac{2}{1 + No.of\ days}$$

1.2.3 Moving Average Convergence Divergence (MACD)

Moving Average Convergence Divergence illustrates the direction of the market trend and the likelihood of the stock price to move in either an upward or downward manner over time. MACD has 3 parts and are given by following equations:

$$Macd\_Line = 12day\ EMA - 26days\ EMA$$

$$Signal\_Line = 9day\ EMA\ of\ MACD - Line$$

$$Macd\_Histogram = Macd\_Line - Signal\_Line$$

1.2.4 Stochastic Oscillator

Stochastic Oscillator is also called as a momentum oscillator whose value ranges from 0 to 100 and is used to indicate the overbought and oversold condition of the stock.

$$Stochastic\ Oscillator\ (\%K) = \left(\frac{LTP - LOW_X}{HIGH_X - LOW_X}\right) * 100$$

where,

LTP = Last transaction price

LOW_X = The minimum LTP over X-day period $HIGH_X$ = The maximum LTP over X-day period

2. Related Work

G. Jariwala et al. in their paper [1] write that the sentiments of the participants in the market determines the direction of the stock price and their sentiment can be by affected the news headlines and have used K-Mean clustering, Naïve Bayes, and Support Vector Machine (SVM) for sentiment analysis and found K- Means clustering has less accuracy than SVM and Naïve Bayes.

K. L. Tan et. al. [3] used robustly optimized BERT along with Long Short Term Memory (LSTM) for the purpose of sentiment analysis and achieved the f1- scores of 93% on IMDB, 91% on Twitter US Airline Sentiment and 90% on Sentiment140 dataset.

R. Cai et. al [4] used the combination of BERT and Bidirectional Long Short Term Memory (BERT-BiLSTM) for analysis of the sentiment got 86.2% accuracy with recall of 70.78%.

T. B. Shahi et. al. in their paper [6] used LSTM and GRU algorithms to forecast stock price and concluded that both LSTM and GRU shows better prediction after addition of financial news sentiment as an input along with other stock features.

A. S. Saud et. al. [7] used Vanilla RNN, LSTM and GRU for price prediction of the stock and found that GRU had better performance over others. M. E. Karim et. al. [5] used BiLSTM and BiGRU algorithms to predict the stock price and concluded that BiGRU model outperforms BiLSTM model on same dataset and with same hyperparameters.

3. Proposed Work

This paper is divided into 2 parts: Part-I deals with news sentiment analysis and Part-II deals with stock index value prediction.

3.1 Part-I Financial News Sentiment Analysis

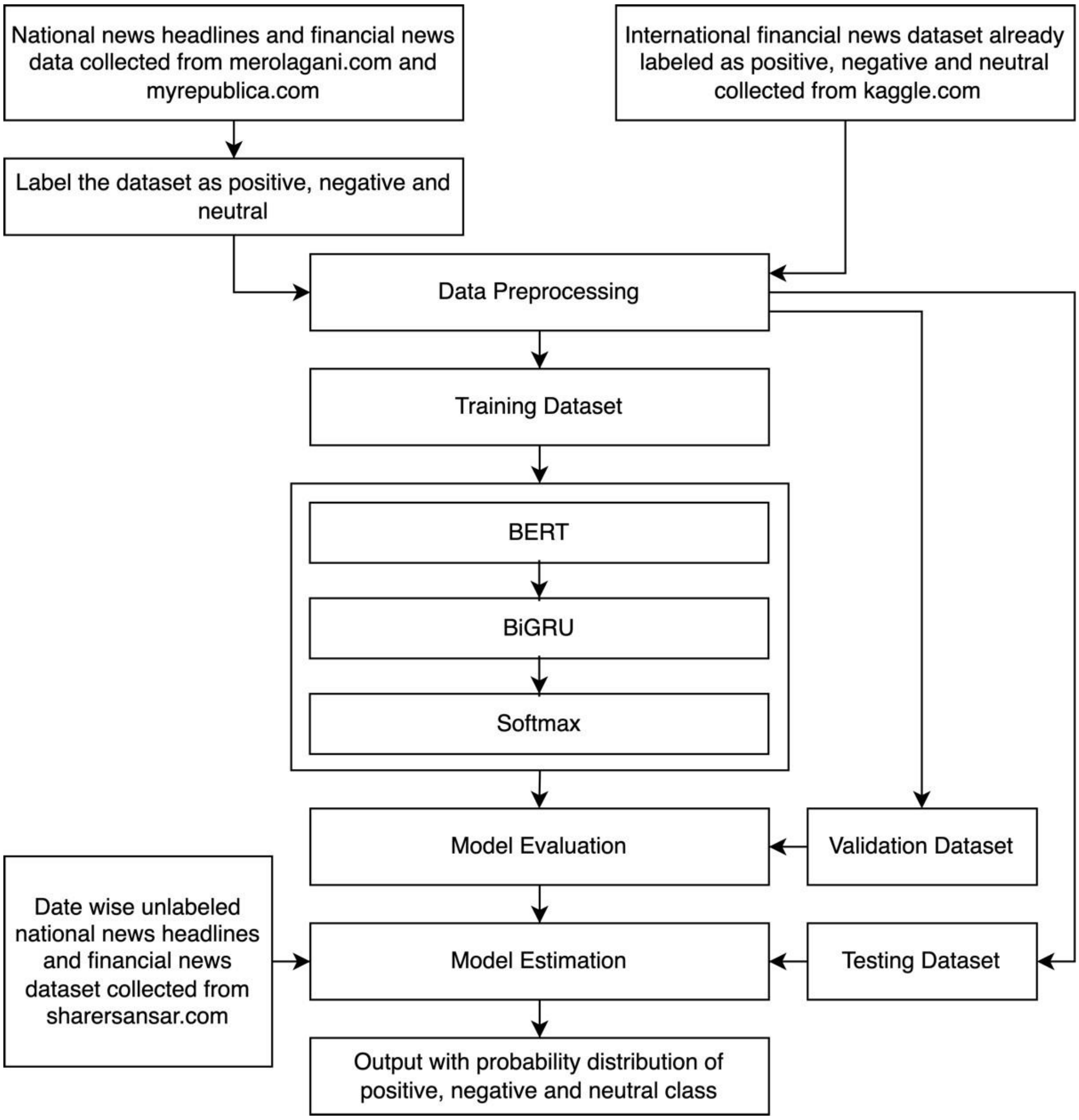

Figure 1. Flow diagram of news sentiment analysis process

3.1.1 Data Collection

National news headlines and financial news data were collected from merolagani.com and myrepublica.com. International news data was collected from kaggle.com.

3.1.2 Data Pre-processing

Unwanted columns like ’date’, ’Image URL’, ’News URL’ etc. and null valued data were removed from the dataset. Html tags like <p>, </p>,
, , etc. were also removed. International dataset was pre-labelled into 3 sentiment classes: Positive, Negative and Neutral. National news dataset wasn’t labelled so it was manually labelled into 3 sentiment classes.

3.1.3 Tokenizing and Word Embedding with BERT

Tokenizing breaks the sentence into a sequence of words. When news headlines are tokenized with BERT, a special token [CLS] is added at the beginning to indicate the sentence has started and [SEP] is added after full stop to notify that sentence has ended. For example, if our news title was ’NEPSE decreased by 10 points today’.

Then after tokenizing our sentence would be [’[CLS]’, ’NEPSE’, ’decreased’, ’by’, ’10’, ’points’, ’today’, ’.’, ’[SEP]’]. Since the length of the news title differs from sentence to sentence, padding is used to make all sentences of equal length. BERT embedding was then used to convert our tokens to their respective integer value.

3.1.4 Data Splitting

70% of the total dataset was used for training, while 15% was used for validation and 15% was used for testing of the model. The dataset was in the ratio of 2:2:1 for positive, neutral and negative sentiment class respectively.

Random oversampling has been used in this paper to balance the imbalance dataset. Random oversampling was done only after splitting the dataset into testing and validation set to avoid the repetition of oversampled data in testing and validation set.

3.1.5 Training and Testing

The BERT-BiGRU model was then trained and tested with the splitted dataset. The model uses SoftMax activation function to interpret the results of BERT-BiGRU model as probability distribution of 3 classes: positive, neutral and negative.

After completion of model testing, date-wise news headlines to be used as input parameters for stock price prediction were collected from sharesansar.com. Such dataset was then passed through our BERT-BiGRU model to obtain the probability distribution of positive, neutral and negative for each news title. This result was then used as input parameter to the part-II of this paper.

3.2 Part-II Stock Index Value Prediction

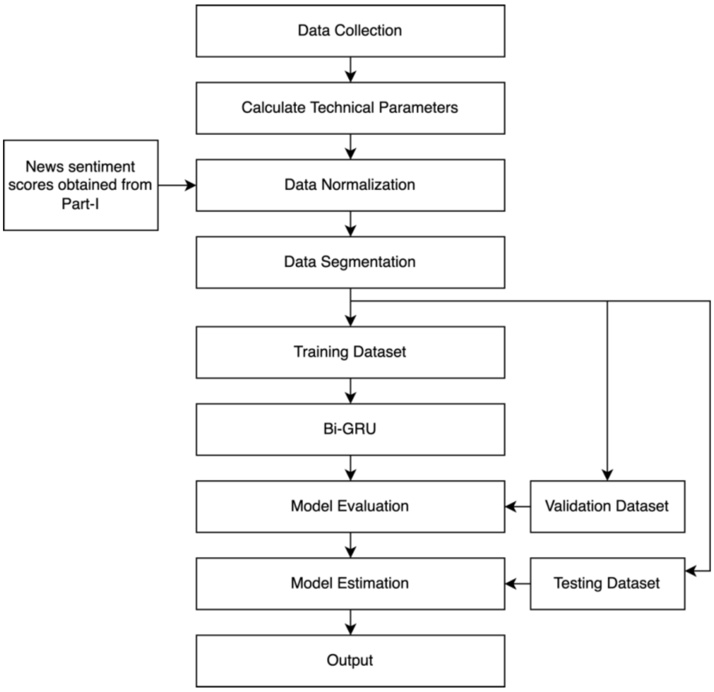

Figure 2. Flow diagram of stock index value prediction

3.2.1 Data Collection

Data for NEPSE index was collected from nepsealpha.com with columns ’Date’, ’Open’, ’Close’, ’High’, ’Low’ and ’Volume’. News sentiment scores for each date was obtained from 1st part of the paper.

3.2.2 Calculate Technical Parameters

Technical parameters: RSI, EMA, MACD and Stochastic Oscillator were calculated and added to the existing dataset.

3.2.3 Data Normalization

Min-Max normalization was used for normalizing data from the scale of 0 to 1.

3.2.4 Data Segmentation

Window sliding technique was used for data segmentation. In this method, if the window size is 7, the window uses the data from day 1 to day 7 to forecast the price on the 8th day. For calculating the price of the 9th day, the window is then moved by 1 day to get an additional 7 days of data and the process is repeated [8]. The window size is also called look back period. The look back period of 5, 7, 10, 12 and 15 days were used in this paper.

3.2.5 Data Splitting

70% of the total dataset was used for training, and 15% was used for validation. 15% of the dataset was used for testing of the model.

3.2.6 Training and Testing

The BiGRU model was then trained and tested with the splitted dataset for all look back periods (5, 7, 10, 12 and 15 days) twice:

- Once, with technical indicating parameters only
- Next, with news sentiment scores along with technical indicating parameters

4. Results and Discussion

The combination of BERT-BiGRU model was able to predict the news sentiments with 82.67% accuracy. The precision, recall and F1-socres are represented in Table 1. Confusion matrix and corresponding ROC curves for positive, neutral and negative classes are shown in Figures 3 and 4 respectively.

Table 1. Result of BERT-BiGRU news sentiment analysis model

Accuracy	**Precision**	**Recall**	**F1-score**
0.8267	0.8353	0.8267	0.8274

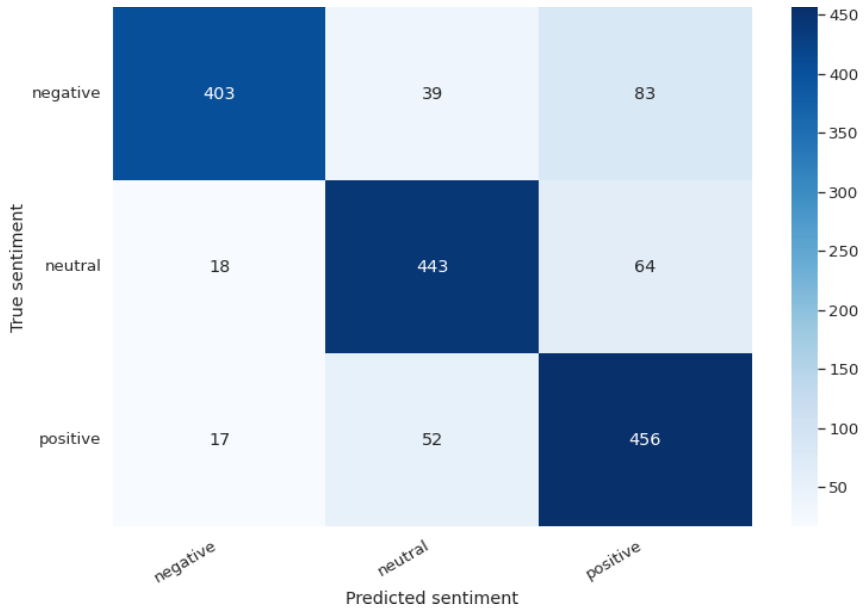

Figure 3. Confusion matrix

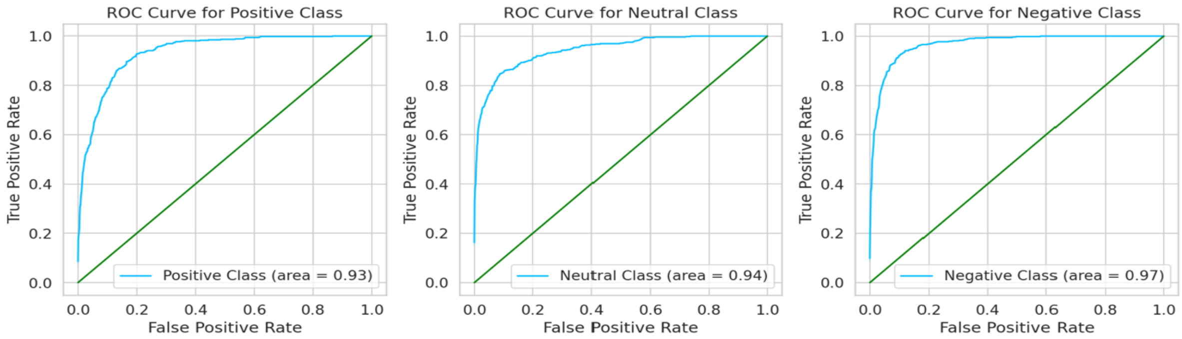

Figure 4. ROC curve for positive, neutral and negative classes

Actual vs predicted closing value of NEPSE index with only technical indicators and with both technical indicators and news sentiment scores are shown in Figures 5 and 6 respectively. It is seen that the graph has improved with the addition of news sentiment score as an input parameter along with other technical parameters.

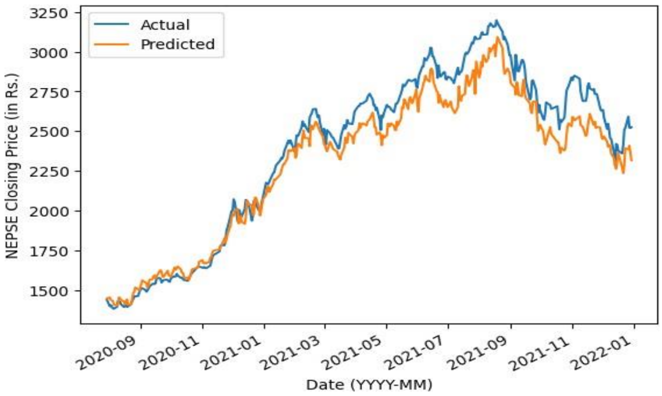

Figure 5. Actual vs Predicted close price for NEPSE (without news)

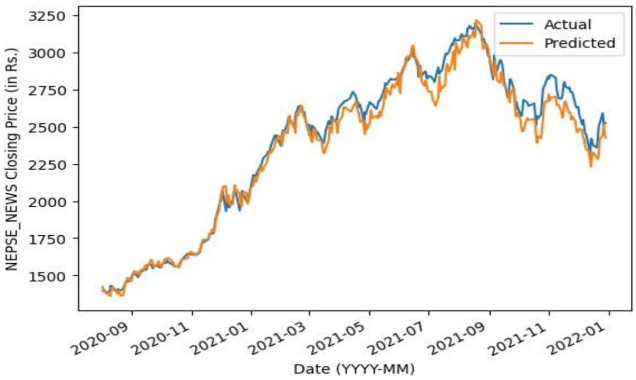

Figure 6. Actual vs Predicted close price for NEPSE (with news)

For all look back periods of 5, 7, 10, 12 and 15 days, MAE value has decreased with the addition of news sentiment scores to the dataset. Figure 7 illustrates the contrast between the MAE value with and without news sentiment scores as a source parameter.

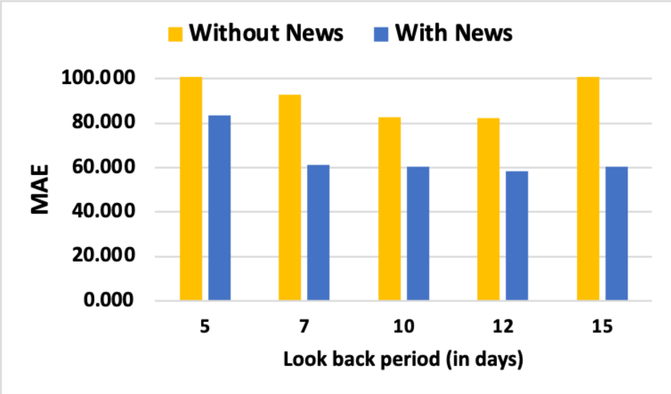

Figure 7. MAE values with and without news sentiment scores for NEPSE

The comparison of RMSE value with and without news sentiment scores as a source parameter is shown in Figure 8. RMSE value also has decreased, for all look back periods, when news sentiment scores are added to the dataset.

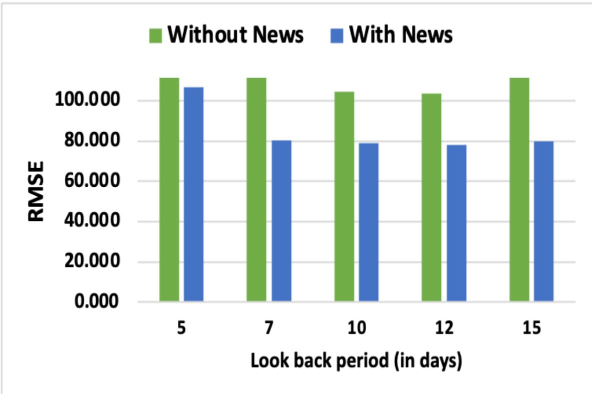

Figure 8. RMSE values with and without news sentiment scores for NEPSE

Figure 9 shows that R^2 value has increased for all look back periods with the addition of news sentiment score as an input parameter.

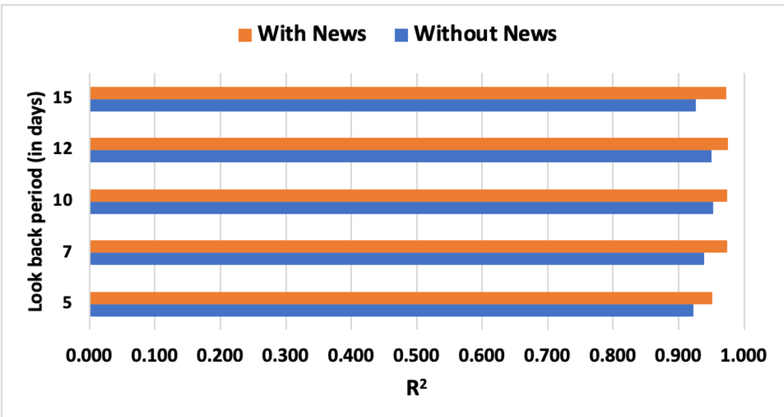

Figure 9. R^2 values with and without news sentiment scores for NEPSE

The decreasing values of MAE and RMSE and increasing value of R^2 signifies that the BiGRU model has performed better when news sentiment score along with technical indicating parameters were added as an input.

5. Conclusion

Addition of news sentiment score (obtained from BERT-BiGRU model) along with technical indicating parameters to our Bi-GRU model have increased the performance of the model for all look back periods of 5, 7, 10, 12 and 15 days. One model is said to perform better thanother if there is decrement in MAE and RMSE values and increment in R^2 value. It is seen from the Figure 7 and Figure 8 that MAE and RMSE values respectively have decreased and from Figure 9, R^2 value have increased after the addition of news sentiment scores for all look back periods.

Therefore, it is concluded that addition of news sentiment score along with technical parameters have better results than just using technical parameters for stock price prediction using Bi-GRU algorithm.

Acknowledgement

The authors of this paper would like to acknowledge Department of Electronics and Computer Engineering, Pulchowk Campus, Institute of Engineering.

References

[1] Gaurav Jariwala, Harshit Agarwal, and Vrai Jadhav. "Sentimental Analysis of News Headlines for Stock Market". In 2020 IEEE International Conference for Innovation in Technology (INOCON), pp. 1–5, 2020.

[2] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. "Bert: Pre-Training of Deep Bidirectional Transformers for Language Understanding". 2018.

[3] Kian Long Tan, Chin Poo Lee, Kalaiarasi Sonai Muthu Anbananthen, and Kian Ming Lim. "Roberta-LSTM: A Hybrid Model for Sentiment Analysis with Transformer and Recurrent Neural Network". IEEE Access, vol. 10, pp. 21517– 21525, 2022.

[4] Ren Cai, Bin Qin, Yangken Chen, Liang Zhang, Ruijiang Yang, Shiwei Chen, and Wei Wang. "Sentiment Analysis About Investors and Consumers in Energy Market Based on BERT-BiLSTM". IEEE Access, vol. 8, pp. 171408– 171415, 2020.

[5] Md. Ebtidaul Karim and Sabrina Ahmed. "A Deep Learning-Based Approach for Stock Price Prediction Using Bidirectional Gated Recurrent Unit and Bidirectional Long Short Term Memory Model". In 2021 2nd Global Conference for Advancement in Technology (GCAT), pp. 1–8, 2021.

[6] Tej Bahadur Shahi, Ashish Shrestha, Arjun Neupane, and William Guo. "Stock Price Forecasting with Deep Learning: A Comparative Study". Mathematics, 2020.

[7] 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, vol. 167, pp. 788–798, 2020.

[8] Dilip Singh and Bhupendra Kumar Gupta. "Closing Price Prediction of Nifty Stock Using LSTM With Dense Network". In Advances in Distributed Computing and Machine Learning, pp. 382–392, 2022.

Author's biography

Subin Maharjan is currently pursuing his MSc. Degree in Computer Systems and Knowledge Engineering from Pulchowk Campus, Institute of Engineering, Tribhuvan University.

Subarna Shakya received MSc and PhD degrees in Computer Engineering from the Lviv Polytechnic National University, Ukraine, in 1996 and 2000, respectively. Currently, he is a professor at the Department of Electronics and Computer Engineering, Pulchowk Campus, Institute of Engineering, Tribhuvan University. His research interest includes egovernment system, distributed and cloud computing, and software engineering and information system, Deep Learning, and Data Science