Stock Market Prediction using Long Short Term Memory Algorithm Combined with Genetic Algorithm

Durgesh Kumar Maurya[*1], Deepesh Dewangan[2], Komal Yadav[3]

[1]Shri Rawatpura Sarkar University Raipur, India

[2]Assistant Professor, Shri Rawatpura Sarkar University Raipur, India

[3]Assistant Professor, Shri Rawatpura Sarkar University Raipur, India

ABSTRACT

Article Info

In this article, Long Short Term Memory (LSTM) algorithm with optimized value of windows size and number of units was used to predict the Nifty – IT daily stock values. Genetic algorithm technique was utilized for optimization. Window Size=36, Number of Units=2 was found to be the best suitable values for prediction using LSTM. Various parameters such as cosine similarity (CS), Root mean square Error (RMSE), R2 etc. were used to determine the quality of the prediction done. The value of CS, RMSE and R2 was calculated as 0.99992, 0.00374 and 0.97713 respectively. The CS and R2 values are closed to 1.0, whereas the value of RMSE is very small, which is an indicative that the actual and predicted data are much closed to each other.

Keywords :- Genetic Algorithm, LSTM, Prediction, Stock, Nifty.

I. INTRODUCTION

The Nifty meaning is a derivation from the mix of two words, i.e. "National Stock Exchange" and "fifty". It is an abbreviation of the National Stock Exchange Fifty. It is a collection of top performing 50 equity stocks that are actively trading in the index. However, 51 stocks are currently trading on Nifty. Hence, Nifty is also known as Nifty50 or CNX Nifty [1]. Information Technology (IT) industry has played a major role in the Indian economy. In order to have a good benchmark of the Indian IT sector, NIFTY IT sector index has been developed. NIFTY IT provides investors and market intermediaries with an appropriate benchmark that captures the performance of the IT segment of the market in India [2]. Now a

Day finance market has a great significant impact on many areas like Technology and business. Shareholders used preferred two main strategies to make decision on stock market, invest money and earn more returns with low risk. To make the right decision investors used stock market prediction tools to market forecasting. The stock market indices are calculated based on their market capitalization. Accurate forecasting of the stock market is therefore a very difficult task by changing the market world. The researchers and market analysts have been keen on developing and testing of stock market behavior. Therefore, different types of statistical techniques were used for stock market prediction [3].

Still, with LSTM, inadequately selected initial weights could sometimes lead to inaccurate performance. The

motivation behind this study was to consider a strategy to such an extent that LSTM could begin with a set of appropriately chosen initial weightings to reduce forecasting error. Specifically, the taken in to consideration hybrid approach (GA-OLSTM) is liable for both the learning and memorization intelligence using the LSTM Network along with the searching ability of the genetic algorithm [4-8]. Factors to judge the quality of prediction with their significance were MSE, RMSE, CS and r.

The results show that the proposed GA-LSTM technique could cope with high variations of stock market values with great accuracy as compared to other approaches

II. DATASET COLLECTION

The dataset was collected from the website (https://www.kaggle.com/datasets/sudalairajkumar/nifty-indices-dataset). The access date was 05 August 2022 [9].

III. RESULTS AND DISCUSSION

In the experimental results, we started experimenting with the time series Uni-Variant data by applying the GA LSTM model. With 100 epochs various factors such as R2, MSE, RMSE, MAE, NORM, RSR and CS, were calculated to investigate the accuracy of forecasting data was found as 0.97713, 0.00001, 0.00374, .0020205, .092220, 0.151196 and 0.999920

respectively which is illustrated in Table 01. Various parameters such as cosine similarity (CS), Root mean square Error (RMSE), R2 etc. were used to determine the quality of the prediction done. The value of CS, RMSE and R2 was calculated as 0.99992, 0.00374 and 0.97713 respectively. The CS and R2 values are closed to 1.0, whereas the value of RMSE is very small, which is an indicative that the actual and predicted data are much closed to each other [8]. The graph obtained for window size and number of units (36, 2) which gave the best possible result is shown in Fig. 01. In this figure, a great resemblance between the actual and forecast data can be observed.
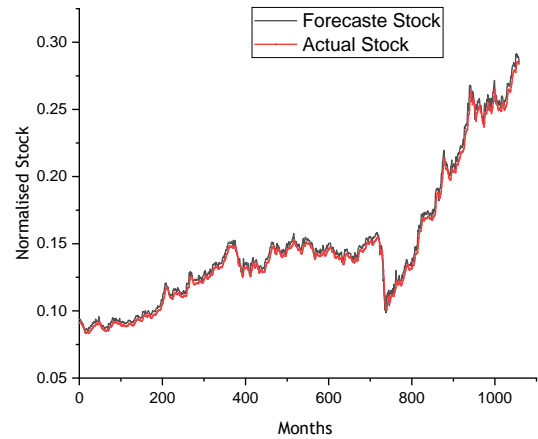

Figure 01 Comparison between actual and forecast Stock (100epoch (s), optimal window size = 36, number of LSTM units = 2) using GA-LSTM algorithm.

Table 1 GA-LSTM Model-based analysis of uni-variant time-series data

Model	Optimal Parameter	Epoch	R2	MSE	RMSE	MAE	NORM	RSR	CS
GA-LSTM	Window Size=36 , Number of Units=2	100	0.97713	0.00001	0.00374	0.00202	0.0922	0.1512	0.99992

IV.CONCLUSION

In present research paper, Deep learning approach for stock forecasting based on the time series analysis is incorporated using PYTHON V 3.9. An attempt has been made to forecast the monthly stock data applying GA-LSTM algorithm for 100 epochs. The factors R2, MSE, RMSE, MAE, NORM, RSR and CS were calculated to judge the quality of forecasting . These parameters were found 0.97713, 0.00001, 0.00374, .0020205, .092220, 0.151196 and 0.999920 respectively. For investigating the best possible combination window Size and Number of units chosen was (36,2) respectively. These experimental results showed that the GA-LSTM approach gave the best result in predicting stock. To improve the result obtained, a hybrid model may be studied in future.

V. REFERENCES

[1]. https://scripbox.com/pf/what-is-nifty/ Accessed on (01.08.2022)

[2]. https://www.niftyindices.com/indices/equity/sectoral-indices/nifty-it (Accessed on 01.08.2022)

[3]. D Kumar, P K Sarangi, and R Verma, Materials Today: Proceedings, https://doi.org/10.1016/j.matpr.2020.11.399 (2020)

[4]. Kim Y, Roh JH, Kim H, Sustainability 10:34–53 (2017)

[5]. Armano G, Marchesi M, Murru A, Inf Sci 170:3–33 (2005)

[6]. Pal SK, Wang PP, Genetic algorithms for pattern recognition. CRC Press, Boca Raton, p 336 (1996)

[7]. Santra SA, Lin LJ Integrating long short-term memory and genetic algorithm for short-term load forecasting. Energies 12:2040 (2019). https://doi.org/10.3390/en12112040

[8]. Thakur N, Karmakar S, and Soni S, Int. j. inf. tecnol.,14(4):1961–1966 (2022)

[9]. https://www.kaggle.com/datasets/sudalairajkumar/nifty-indices-dataset (Accessed on 01.08.2022)