Prediction of Stock Market Using Recurrent Neural Network

Sadman Bin Islam , Mohammad Mahabubul Hasan and Mohammad Monirujjaman Khan
Department of Electronics and Computer Engineering
North South University
Bashundhara, Dhaka, Bangladesh
monirujjaman.khan@northsouth.edu

2021 IEEE 12th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON) | 978-1-6654-0066-4/21/$31.00 ©2021 IEEE | DOI: 10.1109/IEMCON53756.2021.9623206

***Abstract*—In the financial realm, stock price forecasting is becoming increasingly popular. Shares price prediction is important for increasing the interest of speculators in putting money in a company's stock in order to grow the number of shareholders in the stock. Successfully predicting the price of a stock in the future could yield significant profit. When it involves forecasting, various methodologies are used. This paper uses a recently introduced model for predicting stock price. This proposed model is a well-liked model named is the Recurrent Neural Network (RNN) model. One of the variant of RNN is Long Short Term Memory (LSTM) model. It are often shown from the simulation results that utilizing these RNN models such as LSTM, and constructing with proper hyper-parameter tuning, these expected models can estimate the future stock market with the maximum percentage of accuracy. The RMSE for a LSTM model was calculated by changing the amount of epochs, the variation between predicted stock price and actual stock price. The model is trained and classified for accuracy with different sizes of knowledge. The computations are conducted by exploiting a widely accessible datasets for stock markets containing date, volume, opening price, highest price, lowest price, and closing prices. The major goal of this article is to determine to what degree a Machine Learning algorithm can anticipate the stock market price with greater accuracy.**

Keywords— Deep learning; LSTM; RNN; Stock market; Bangladesh; Prediction.

I. INTRODUCTION

Stock market prediction is an object of studies for several years, but given its high complexity, no of variables and sources included it's been proven to be a really difficult task. With the rise in use of machine learning algorithms and techniques during a big variety of disciplines, stock exchange prediction has also been improved using different machine learning techniques. This paper explains how of finding trends available market prices using historical data while (RNNs) are considered to be one among the foremost powerful models for processing sequential data like statistic data [1], they are not able to learn long-term dependencies because interactions between time steps that are several steps apart are difficult to stay track of.The Long short term memory (LSTM) may be a kind of recurrent neural network that's going to solve this problem.

Data comprising the closing stock price of multiple companies from 1999 to 2021 are going to be used as sources of data for the network. Prices were taken from the Stock Bangladesh website. Utilizing that dataset, the model is going to be trained and assessed to plot a graph which will be used to analyze the prediction within the closing stock price of those companies. Section III outlines how the task provided in this paper was done, section IV shows the outcomes, and section V draws research to a close and briefly covers future prospects.

II. RELATED WORKS

There are numerous papers, articles and blogs posted about the prediction of stock prices, from different disciplines like economics, physics, computing and statistics. In 2012, quite 80% of trades within the US stock exchange were done by different algorithms [3].The Efficient Random Walk Theory states that this price of an asset usually reflects all of the previous data related to it [4]. On the contrary, the stochastic process hypothesis [5] states that stock price is independent of its past, to put it another way, the current stock price is solely determined by these parameters and has no bearing on previous data..Concerning Different machine learning models, Support Vector Machines (SVM) are used for statistical learning [6].

III. THEORY

Schmidhuber et al. proposed the LSTM network model in 1997 [8]. The LSTM network model was created to deal with the long-standing problems with gradient expansion and disappearance in RNN [9, 10]. Because it's memory and may perform pretty accurate forecasts, it's been frequently utilized in voicedetection, sentimental analysis, and text analysis [11], [12] it's also been utilized in the realm of stock exchange forecasting in recent years [13]. A standard RNN has simply one repeating module and an easy internal structure. It's nothing but a tanh layer. Four of the LSTM layers, on the opposite hand, are just like ordinary RNN modules and add a singular interactive mode [14]. In [15-22] different deep learning and machine learning approaches have been proposed for various applications. The LSTM memory unit is formed of three components: a forget layer, a memory unit, and therefore an output layer as shown in Figure 1. Figure 2 shows the block diagram of LSTM.

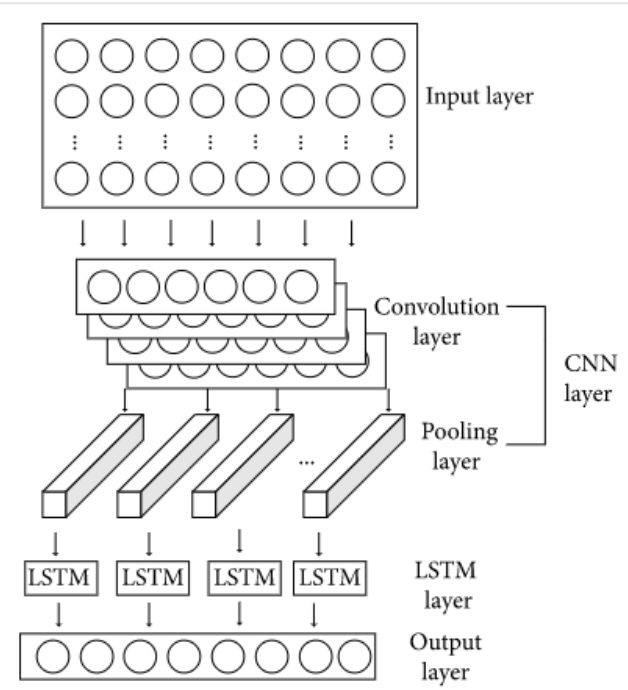


Figure 1: Construction of LSTM Unit

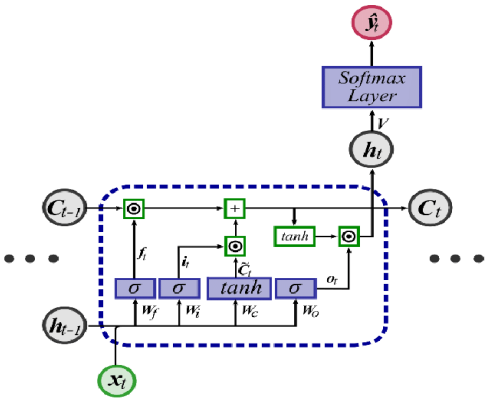
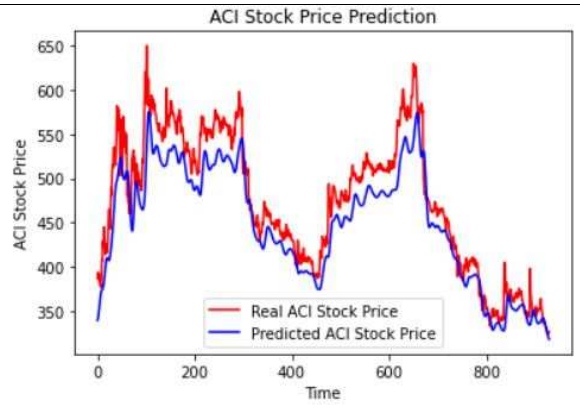
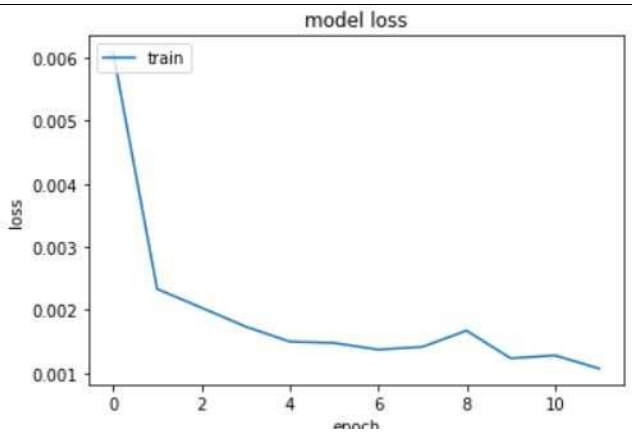
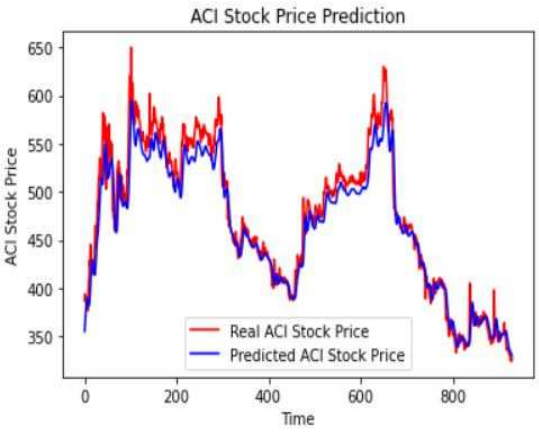
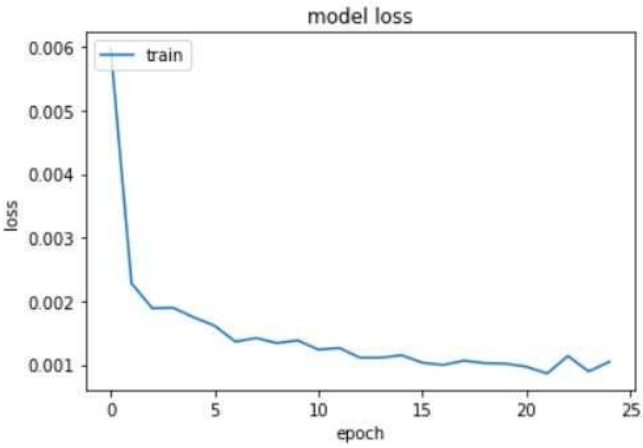


Figure 2: Block Diagram of LSTM Neural Neural Network

IV. METHODOLOGY

Our target was to train our model with the dataset of reputed companies which have stock market share. We used the dataset of ACI, BEXIMCO, ISLAMI BANK companies. We collected the data from the Stock Bangladesh website and sorted the dataset according to our model. Our initial difficulty was to find a better-suited model for the prediction. After trying three models we used one of these pre-trained models to train with our own dataset as it was better fitted for the prediction [15]. We had to modify the model code and imported different libraries from KERAS to train the model properly. Our initial target was to find out the perfect epoch for the model which will work on every company's dataset prediction. We trained our model with a training dataset and used the testing dataset to predict the price. After increasing the epoch up to 100 we can get a better prediction of all the stock market prices of companies we worked on.

Epochs	Prediction (Test Data)	Loss (Train Data)
25		
50		

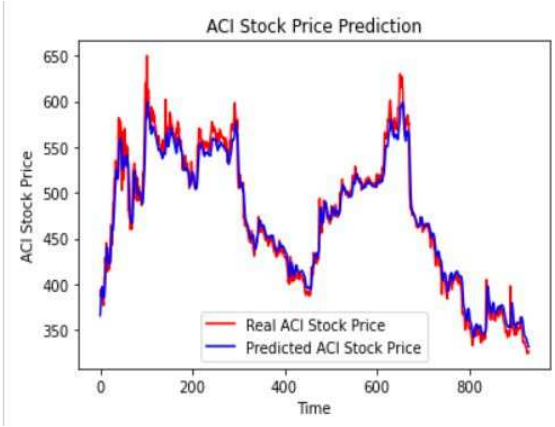
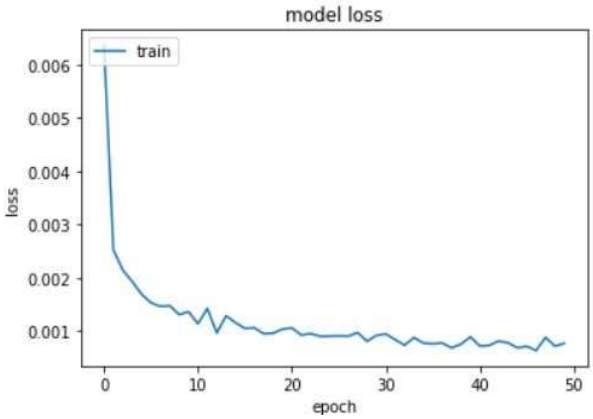
100		

Figure 4: The reduction of loss function after increasing the epoch

V. DATASET AND FEATURES

A. Raw Data:The dataset provides daily quotes for the companies going back to 1999 and is downloaded using the official website of Stock Bangladesh. It comes with the "Date", "Open", "Highest Price", "Lowest_Price","Close_Price" and "Volume" for each day. Overall, the raw dataset is formed of more or less 4000 rows and 6columns.

B. Scaling:Before training, all features are rescaled to (i) make the cost function's curve more proportional and making easier to optimize, (ii) make training less sensitive to feature scale, and (iii) ensuring regularization doesn't respond distinctly for various scaling. Two well-known approaches for rescaling data are normalization and standardization. We tried Normalization which works best for our application.

C. Normalization:Normalization scales all numeric variables within the range [0,1]:

D. Separating Test Data:We employed the walk-forward-cross-validation with a dev size of 10% for our model development as K-Fold Cross Validation fails in finance [16].Then we restrained the final 20% of our data as a separate test set in order to provide aneutralassessment of the final model. This set will be used after training the model perfectly.

VI. RESULT AND DISCUSSION

After training the NN with the sorted dataset of all companies we were able to predict the accuracy of 88% prediction. The model has been trained with datasets from 1999-01-02 to 2021-04-31. And 20% of the dataset is used for testing the dataset.Assembling the data and sweeping was most likely one of the toughest part because the info wasn't served in accurate order.

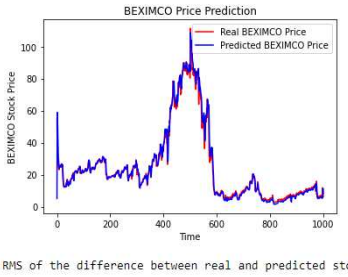


Figure 6.1-Accuracy percentage of BEXIMCO stock price

The graph which is showed above depicts BEXIMCO, with the timeline on the parallel axis and the share price on the perpendicular. From 1999 until 2021 is the timeframe.Here the LSTM was successfully implemented because the accuracy rate is more than 90% which is exceptional for the prediction of stock price.

Similar to Beximco, we also trained our model with ACI, ISLAMI BANK, NAVANACNG, and also got an average accuracy of 85%.

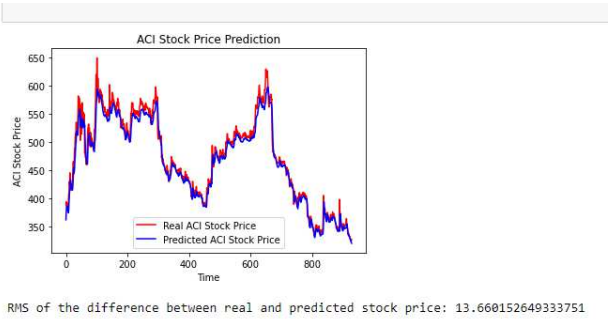


Figure 6.2-Accuracy percentage of ACI stock price

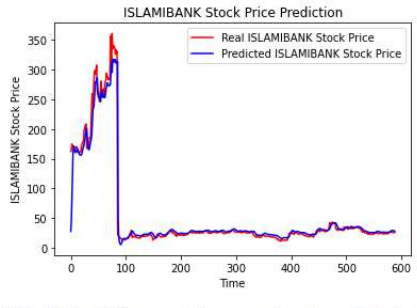


Figure 6.3-Accuracy percentage of ISLAMI BANK stock price

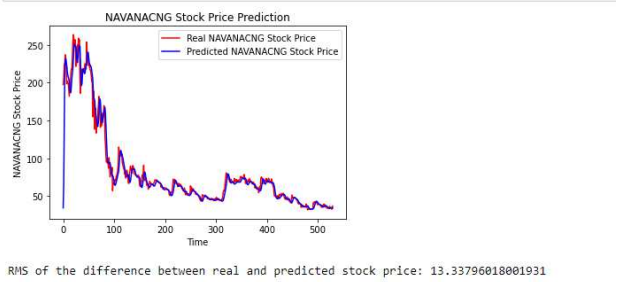


Figure 6.4-Accuracy percentage of NAVANACNG stock price

CONCLUSIONS AND FUTURE WORK

The outcome of this study ensures that machine learning approaches are proficient for predicting the stock exchange performance. Even after simplifying the problem to a binary price trend and using a sophisticated deep learning model like LSTM, we can conclude that stock price forecasting remains a difficult endeavor. Nonetheless, this knowledge will aid future research, particularly in selecting appropriate feature sets and model designs. We have shown that it might need more than Price/Volume data to predict future returns. The potential future direction for this paper may include**:**

- Enriching our dataset. We should include other companies' datasets to train our model more accurately.
- Exploring new models. We should experiment with different models such as attention. resnet-101.
- Experimenting with various time frames. We can explore a variety of different window sizes.

ACKNOWLEDGEMENTS

The author of this paper would like to show courtesy to the Stock Bangladesh for providing the specified dataset which is crucial part of this project. The author would also wish to be grateful to the supervisor for his uninterrupted support and direction.

References

[1] Saad, Emad W., Danil V. Prokhorov, and Donald C. Wunsch. "Comparative study of stock trend prediction using time delay, recurrent and probabilistic neural networks." Neural Networks, IEEE Transactions on 9.6 (1998): 1456-1470

[2] Hochreiter, Sepp, and Jürgen Schmidhuber. "Long short-term memory." Neural computation 9.8 (1997): 1735-1780

[3] M. Glantz and R. Kissell, Multi-asset Risk Modeling, 1st ed. Academic Press, 2013.

[4] E. F. Fama and B. G. Malkiel, Efficient capital markets: A review of theory and empirical work,in The Journal of Finance, vol. 25, no. 2, pp. 383–417, 1970.

[5] B. G. Malkiel, A Random Walk Down Wall Street. Norton, New York, 1973

[6] K. Kim, "Financial time series forecasting using support vector machines," Neurocomputing, vol. 55, no. 1–2, pp. 307 –319, 2003, support Vector Machines.

[7] B. Batres-Estrada, "Deep learning for multivariate financial time series," 2015

[8] https://github.com/AMoazeni/Machine-Learning-Stock-Market-Prediction?fbclid=IwAR2VsCkoSJqoJimXJv2QF2aLoQsLR1eRGuNxvyR7m78Zjrfp8SuXUFKsDgk

[9] Media Filter, [Online]. Available: http://fourier.eng.hmc.edu/e161/lectures/smooth_sharpen/node2.html

[10] G. Hinton," Root Mean Square Propagation " [Online].

[11] Stock Market Prediction Using LSTM Recurrent Neural Network by Adil Moghar and Mhamed Hamiche. [Online]

[12] S. Hochreiter and J. Schmidhuber, Long short-term memory , in *MIT Press*, vol. 9, no. 8, pp. 1735–1780, 1997.

[13] V.-D. Ta, C.-M. Liu, and D. A. Tadesse"Portfolio optimization-based stock prediction using long-short term memory network in quantitative trading," *Applied Sciences*, vol. 10, no. 2, pp. 437–457, 2020.

[14] O. Zarrad, M. A. Hajjaji, and M. N. Mansouri, Hardware application of hybrid wind-solar energy system for pumping water based on artificial neural network controller , in *Studies in Informatics and Control*, vol. 28, no. 1, pp. 35–44, 20

[15] M. D. Kamrul Hasan, Sakil Ahmed, Z. M. Ekram Abdullah, Mohammad Monirujjaman Khan, Mehedi Masud et al., **,** "Deep Learning Approaches for Detecting Pneumonia in COVID-19 Patients by Analyzing Chest X-Ray Images," *Mathematical Problems in Engineering, Hindawi, Volume 2021, Article ID 9929274, PP. 1-8*

[16] Morshedul Bari Antor, A. H. M. Shafayet Jamil, Maliha Mamtaz, Mohammad Monirujjaman Khan et al., "A Comparative Analysis of Machine Learning Algorithms to Predict Alzheimer's Disease," *Journal of Healthcare Engineering,Hindawi, Impact Factor 2.682, Scopus Indexed. vo. 2021, no. 9917919, pp. 1-12. https://doi.org/10.1155/2021/9917919*, 2021

[17] Sumaiya Tarannum Noor, SyedaTasmiah Asad, Mohammad Monirujjaman Khan et al., "Predicting the Risk of Depression Based on ECG Using RNN," *Computational Intelligence and Neuroscience, (Impact Factor 3.633, Scopus Indexed), Article ID 1299870. vol. 2021, pp. 1-12. https://doi.org/10.1155/2021/1299870*, 2021

[18] Azher Uddin, Bayazid Talukder, Mohammad Monirujjaman Khan et al., "Study on Convolutional Neural Network to Detect Covid-19 from Chest X-Rays," *Mathematical Problems in Engineering, vol. 2021, Article ID 3366057, 11 pages, 2021. https://doi.org/10.1155/2021/3366057*

[19] Shalini STALIN, Vandana Roy SA , Prashant Kumar Shukla, Piyush Kumar Shukla, ATEF ZAGUIA and Mohammad Monirujjaman Khan, "A Machine Learning based Big EEG Data Artifact Detection and Wavelet Based Removal: An Empirical Approach," *Mathematical Problems in Engineering, vol. 2021, Article ID 2942808, 11 pages, 2021. https://doi.org/10.1155/2021/2942808*

[20] Md. Sajjad Mahmud Khan, Mahiuddin Ahmed, Raseduz Zaman Rasel and Mohammad Monirujjaman Khan, "Cataract Detection Using Convolutional Neural Network With VGG-19 Model," *2021 IEEE World AI IoT Congress (AIIoT), 2021, pp. 0209-0212, doi: 10.1109/AIIoT52608.2021.9454244.*, 2021

[21] Mohammad Monirujjaman Khan et al., "Stock Market Prediction Using Deep Learning by LSTM," *IEEE 11th Annual Computing and Communication Workshop and Conference (IEEE CCWC), 27th-30th January 2021, USA.pp. 0180-0183, doi: 10.1109/CCWC51732.2021.9375835.*, 2021

[22] Mehran Chowdhury, Md. Hasibur Rahman Khan, Md. Nuruzzaman Pranto, Ratil H. Ashique, Mohammad Monirujjaman Khan, "High Speed Tracking with Machine Learning," *IEEE 17th International Conference on Smart Communities: Improving Quality of Life Using ICT, IoT and AI (HONET-2020), December 14-16, 2020, The University of North Carolina at Charlotte, North Carolina, USA pp. 44-48, doi: 10.1109/HONET50430.2020.9322830.*, 2020