Proceedings of the International Conference on Electronics and Renewable Systems (ICEARS 2022)
IEEE Xplore Part Number: CFP22AV8-ART; ISBN: 978-1-6654-8425-1

Stock Price Prediction using Machine Learning

B N Varaprasad
Information Technology
Velagapudi Ramakrishna Siddhartha Engineering College
Vijayawada, India
varaprasad3996@gmail.com

Ch.Kundan Kanth
Information Technology
Velagapudi Ramakrishna Siddhartha Engineering College
Vijayawada, India
kundankanthch99@gmail.com

G.Jeevan
Information Technology
Velagapudi Ramakrishna Siddhartha Engineering College
Vijayawada, India
jeevang85@gmail.com

Y.Kalyan Chakravarti
Department of Information Technology
Velagapudi Ramakrishna Siddhartha Engineering College
Vijayawada, India
klyn.518@gmail.com

2022 International Conference on Electronics and Renewable Systems (ICEARS) | 978-1-6654-8425-1/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICEARS53579.2022.9752248

Abstract- In the current era stock price prediction plays a key role for prediction of future data with respect to training the past data by using machine learning or deep learning technologies. Building a model and then passing the past data as input that is as training data to the model based on the results acquired need to consider an algorithm which gives better accuracy and response time and segmentation. In this paper for estimating the stock values we are considering LSTM and Regression models of Machine Learning. Factors considered are opening values of stock; closing values of stock, lower and higher values of stock and volume.

Keywords:-Stock Price Prediction, Regression, Long and Short-Term Memory Network, Segmentation, Close, High, Low Open, Accuracy.

I. INTRODUCTION

During the recent times we are able to see the growth of the stock market. From the last ten years the growth is tremendous. Estimation of the stock accurately can generate huge revenues for the customers. Considering the historical data insights need to be drawn to predict the future value. It is very difficult to predict the price of a stock as stock market is more volatile. There are many machine learning/ deep learning models which solves these requirements but need to choose a model from them based on the requirements or observations from the historical data [1]. Collection of data plays a vital role in choosing a model. Supervised Machine Learning is applied on the dataset which was collected from Google stock. The attributes or variables or parameters in the dataset are as follows: opening value of stock, closing value of stock, lower and higher values of stock on that day and the amount of stocks sold or purchased on that day. Bid Prices in stock vary with respect to time. Volume is the number of transfers that takes with different customers and also the data being generated.

Considering Simple Linear Regression and Long Short-Term Memory while building the model, where regression is used to minimize the error rate and LSTM [3]is used to remember the data and their dependencies for long span of time. In the II Part literature survey is mentioned. III part is about the data set that is used in the paper to run our proposed model.Part-4 describes Regression and LSTM layers and their work on the test datasets, Similarly Part-5 describes the statistical representations of data for easier understanding of the scenarios for taking better decisions and the final part contains conclusion and future work.

II. LITERATURE SURVEY

This section briefly discuss about various machine learning algorithms used for the prediction of stock market. In [1], identifies a relationship between performance of the stock market and growth of the economy. In [2], ANN and Random forest methods are used to predict the prices of the selected stocks from different sectors. Like the same way of predicting Stocks they tried to predict Bitcoin using LSTM.Based on the results accuracy and the other required parameters are calculated in [3]. In the work of [4], tried to compare various linear and nonlinear models to predict the stock price. They have not only limited to single market but also applied to other markets using the trained neural networks. In [6], developed a system for the prediction of the stock in the future neural network. They collected the old data of some of the companies as well as other parameters which are necessary for calculating the values have taken. In the work of [7], forecast the value of a stock, instead of statistical techniques we may replace them with neural networks as they gives good accuracy. In [8], LSTM network is used to predict future of stock prices based on the history, besides technical analysis indicators. Using this type of algorithm improvements are able to seen compared with machine learning algorithms.

From the literature survey, many challenges are identified. Some these include: the algorithms are mostly based on the linear models, there are many method specific parameters, the method have performed on the normal machines, some of the methods mostly used the standard feature extraction techniques followed by a simple linear regression or neural networks. This paper addresses the some of these challenges and implemented the deep learning architecture for efficient price prediction of the stock market.

III. DESCRIPTION OF THE DATASET

We have taken the data set from www.nseindia.com. The data we took are the price of the stock, low, high, open, close and

volume. The data took from the data set is from January 1st 2000 to July 31st 2020 for 50 stocks. In our paper we run the model on Adaniport and Bpcl Stocks. We are very much thankful to the NSE for providing the dataset.

TABLE I. SAMPLE DATA FROM THE DATA SET FOR ADANI PORT STOCK

Date	*Open*	*High*	*Low*	*Close*	*Volume*
11/27/2007	770	1050	770	962.9	27294366
11/28/2007	984	990	874	893.9	4581338
11/29/2007	909	914.75	841	884.2	5124121
11/30/2007	890	958	890	921.55	4609762
12/3/2007	939.75	995	922	969.3	2977470
12/4/2007	985	1056	976	1041.45	4849250
12/5/2007	1061	1099.5	1050	1082.45	2848209
12/6/2007	1089	1109.7	1051	1081.3	1749516
12/7/2007	1100	1134	1078	1102.4	2247904
12/10/2007	1110	1110	1061.1	1075.4	1012350
12/11/2007	1081	1089	1041	1047.65	810464
12/12/2007	1032	1065	1016	1036.8	744799
12/13/2007	1040	1150	1030.25	1129.95	3067687
12/14/2007	1139.9	1140	1101.1	1110.5	1070737
12/17/2007	1140	1168	1021.5	1044.25	1404955
12/18/2007	1045	1109.9	1031.55	1074.95	1226984
12/19/2007	1091	1116	1046.3	1066.9	845666

IV. PROPOSED METHOD

The problem with stock market prediction is to handle streaming data. The solution to the problem can be addressed by machine learning methodologies, by relating previous data to the streaming data and then training the model which helps in achieving better results in terms of accuracy and time. There are a number of approaches in machine learning for solving the requirement for this there are two models which are better when compared with the other models. They are regression based model and LSTM based network model.

2.1 Regression- Based Model

Regression Analysis contains two different types of variables they are dependent and independent variables. Let us consider the equation

$$D=gB+k \quad (1)$$

Where D is dependent variable, B is independent variable. Regression is mainly intended to predict continuous values through values which are not dependent [5]. For minimizing the error rate algorithm called gradient descent is used to minimize error rate.

$$V = c + d\,L + \text{error} \quad (2)$$

where, *V* is a continuous variable; L – Not dependent value; and, c, *d* are constants.

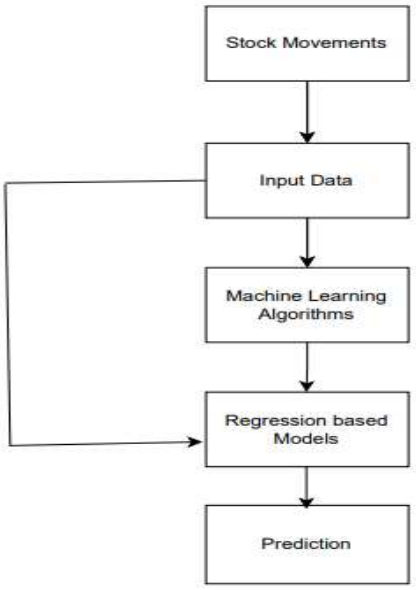


Fig. 1. Flow Chart for Regression Model

$$V = c + d\,L + \text{error} \quad (2)$$

where, *V* is a continuous variable; L – Not dependent value; and, c, *d* are constants. Using the library of panda we carried out our work with CSV format of the data and the parameter that is to be forecasted is calculated, the price of the stocks with respect to the time. The data is to be cut up into dissimilar train sets for cross-corroborate to avoid over fitting. The validation set is kept generally at 20% of the whole dataset. Simple Linear regression as shown in given above equation is implemented on data followed by forecasted were made, which were graphically represented on market prices and time [6].

2.2 Long Short Term Memory (LSTM) Network - Based Model

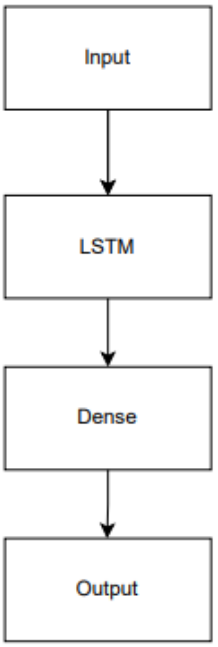


Fig. 2. Layers in LSTM

TABLE II. SUMMARY OF THE PROPOSED LSTM

Layer (type)	Output Shape	Param #
lstm_132 (LSTM)	(None, 60, 50)	10400
dropout_132 (Dropout)	(None, 60, 50)	0
lstm_133 (LSTM)	(None, 60, 50)	20200
dropout_133 (Dropout)	(None, 60, 50)	0
lstm_134 (LSTM)	(None, 60, 50)	20200
dropout_134 (Dropout)	(None, 60, 50)	0
lstm_135 (LSTM)	(None, 60, 50)	20200
dropout_135 (Dropout)	(None, 60, 50)	0
lstm_136 (LSTM)	(None, 60, 50)	20200
dropout_136 (Dropout)	(None, 60, 50)	0
lstm_137 (LSTM)	(None, 50)	20200
dropout_137 (Dropout)	(None, 50)	0
dense_24 (Dense)	(None, 1)	51

Total params: 111,451
Trainable params: 111,451
Non-trainable params: 0

LSTM has more advanced features when compared with that of RNN. In LSTM the information related to the previous state will be available. RNN finds relationships between recent and current information and also involves long term dependencies. In comparison among RNN and LSTM, interval of information is smaller in RNN. Considering this approach for building model is that predictions depend on larger datasets and are dependent on historical data[6]. To attain higher values of accuracy LSTM retains information for older stages [7].

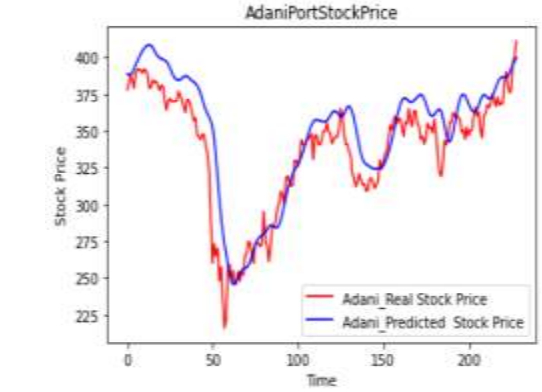
As the data in stock market is streaming, Learning Rate of the system need to monitored with respect to the gradients concerning the weight matrix which may lead to Vanishing Gradient problem. By using LSTM this problem will get solved. As represented diagrammatically the stages involved in LSTM are for cell remembrance and three gates which are i/p, o/p and forget gate. For propagating long term the value has to be remembered and the above mentioned gates will regulate them [8].

In our work, Two LSTM layers stacking have been done to create a sequential model with an o/p value of 256. [0] and [1] are the two different types of input layers. In order to increase the speed of training and to avoid over fitting dropout value is fixed to 0.3. Every neuron in the dense layer of the core is fitted and carried to the side layers. Accuracy is the parameter which we are observing for a change that is increase in this process model is executed with mean square cost function [9].

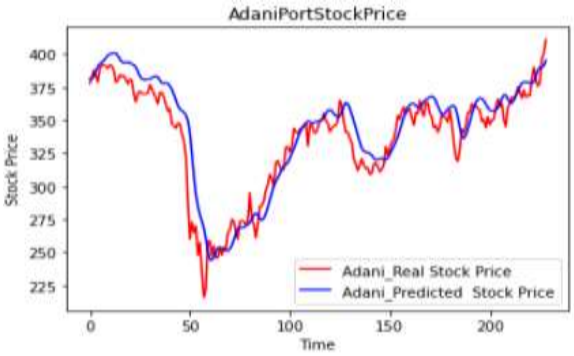
V. EXPERIMENTAL RESULTS

NSE Stock data is used as the dataset for training and testing purposes. The data is split into training and testing as 70 and 30 certain percentages. We took the AdaniPort Stock Price and BPCL to run our model. We run the model for epoch sizes 10,20 and 30 and compared the three results.

(a)

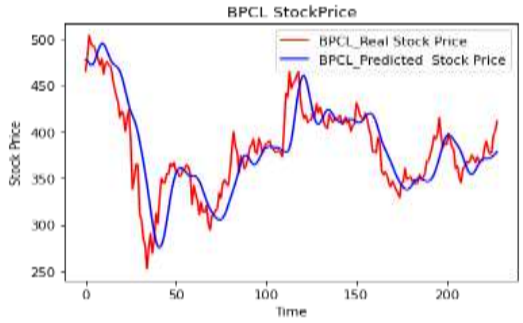


(b)

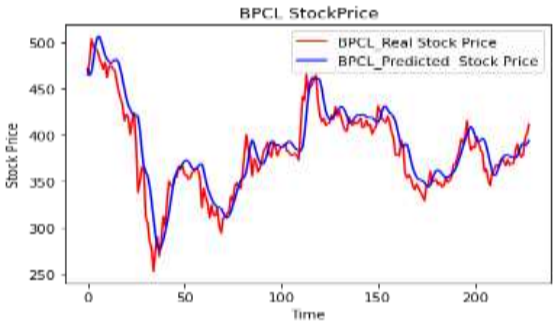


(c)

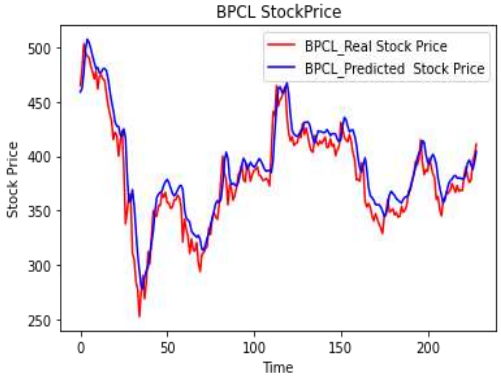
Fig. 3. *AdaniPort stock prediction for (a) 10 Epochs (b) 20 Epochs (c) 30 Epochs*



(a)



(b)



(c)

Fig. 4. *BPCL stock prediction for (a) 10 Epochs (b) 20 Epochs (c) 30 Epochs*

Data having size of the batch as 60 and 10, 20 and 30 epochs, we have plotted. Red color line indicated the prediction and blue color line indicates the actual trend. The two lines indicate the efficiency of LSTM model. When considerable amount of time is passed the trends can be observed. If we want the maximum precision we have to train the system more. If the number of epochs is more, then the actual and predicted stock price is overlapping. Where as in the case less epochs the model is little bit deviating which is not giving actual result.

VI. CONCLUSION

This paper proposed a deep leaning based stock market price prediction. Here, Long Short Term Memory (LSTM) network is proposed and compared the performance of the algorithm with machine learning models. These models are tested on the bench mark dataset taken from NSE Website. We tested the proposed method on two stocks AdaniPortStock and BPCL.The proposed method has shown good performance and which help to obtain the optimistic results. LSTM and regression analysis both show an increase in accuracy and yield optimistic results. Exploited newly introduced machine learning techniques for the prognosticate of stocks have yielded promising results and thereby marked the use them in successful giving and taking schemes. By using Machine Learning models it is observed that the accuracy is attained to its maximum. In future need to implement the models on higher dimensionality datasets and check for accuracy and time.

REFERENCES

[1]. Masoud, Najeb MH. (2017) "The impact of stock market performance upon economic growth." International Journal of Economics and Financial Issues 3 (4) : 788–798.

[2]. Selvin, Sreelekshmy, R. Vinayakumar, E. A. Gopalakrishnan, Vijay Krishna Menon, and K. P. Soman. (2017) "Stock price prediction using LSTM, RNN and CNN-sliding window mode." International Conference on Advances in Computing, Communications and Informatics (ICACCI): 1643-1647.

[3]. Andi, Hari Krishnan. "An Accurate Bitcoin Price Prediction using logistic regression with LSTM Machine Learning model." Journal of Soft Computing Paradigm 3, no. 3 (2021): 205-217.

[4]. Anand, C. "Comparison of Stock Price Prediction Models using Pre-trained Neural Networks." Journal of Ubiquitous Computing and Communication Technologies (UCCT) 3, no. 02 (2021): 122-134

[5]. Mizuno, Hirotaka, Michitaka Kosaka, Hiroshi Yajima, and Norihisa Komoda. (1998) "Application of neural network to

technical analysis of stock market prediction." Studies in Informatic and control 7 (3) : 111-120

[6]. Murkute, Amod, and Tanuja Sarode. (2015) "Forecasting market price of stock using artificial neural network." International Journal of Computer Applications 124 (12) : 11-15.

[7]. Refenes, A.N., Zapranis, A., Francis, G., "Stock Performance Modeling Using Neural Networks: A Comparative Study with Regression Models", Neural Networks, vol. 7, No. 2, pages 375-388, 1994.

[8]. David MQ Nelson, Adriano CM Pereira, and Renato A de Oliveira. Stock market's price movement prediction with lstm neural networks. In 2017 International Joint Conference on Neural Networks (IJCNN), pages 1419–1426. IEEE, 2017

[9]. Yetis, Yunus, Halid Kaplan, and Mo Jamshidi. (2014) "Stock market prediction by using artificial neural network." in2014 World Automation Congress (WAC): 718-722.

[10]. Andrew Pulver, Siwei Lyu, "LSTM with Working Memory" published in 2017 International Joint Conference on Neural Networks (IJCNN)

[11]. Y. Yeti, H. Kaplan, M. Jamshidi, "Stock Market Prediction by Using Artificial Neural Network," IEEE World Automation Congress (WAC) Conference, Waikoloa, 2014, pp.718–722.

[12]. Tom M. Mitchell, "Machine Learning", McGraw-Hill Science/Engineering/Math,1997.

[13]. Zabir Haider Khan, Tasnim Sharmin Alin, Md. Akter Hussain "Price Prediction of Share Market using Artificial Neural Network (ANN)", In the Journal of International Journal of Computer Applications , Volume 22 - no 2, 2011.

[14]. Chang Sim Vui, Gan Kim Soon, Chin Kim On, Rayner Alfred and Patricia Anthony, "A Review of Stock Market Prediction with Artificial Neural Network (ANN)", In the International Conference on Control System, Computing and Engineering , IEEE, 2013.