IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

AI Based Stock Market Predictive Analysis

Author 1
Gargeya Kharat
Student , Sandip University
Nashik
gargeya.kharat15@gmail.com

Author 2
Samruddhi Savale
Student, Sandip university
Nashik
samruddhisavale@gmail.com

Author 3
Siddhika Hase
Student, Sandip University
Nashik
siddhihase66@gmail.com

Author 4
Apurva Bhurhade
Student,Sandip University
Nashik
Bhurhadeapurva12@gmail.com

Author 5
Avinash Taskar
Professor,Sandip University
Nashik
avinash.taskar@sandipuniversity.edu.in

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.

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.

I. 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 , 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 . In the work of , 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 , 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 , forecast the value of a stock, instead of statistical techniques we may replace them with neural networks as they gives good accuracy. In , 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.

II. 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

III. PROPOSED METHOD

The problem with **job** prediction is **related** to **data flow . This** solution can be **solved through** machine learning by **comb ining** previous data **with** streaming data and then training the **mo del, this will help to get a better result** in accuracy and time. Th ere are **many ways to solve this problem** in machine learning **an d** there are two models **that** are better **than others. These** are reg ression based **models** and LSTM based network **models.**

2.1 Regression- Based Model

There are two different types of variables **in regression analysis:** dependent **variables** and independent variables. **Let's** c onsider **this** equation

D=gB+k

where D is **the** dependent **variable and** B is **the** independent varia ble. Regression **basically aims** to predict continuous values **from uncorrelated values[5]. To reduce** the **error, an** algorithm c alled gradient descent is used to **reduce the error.**

V = c + d L + error

Among them, V is a **constant** variable; L **- independent** value ;

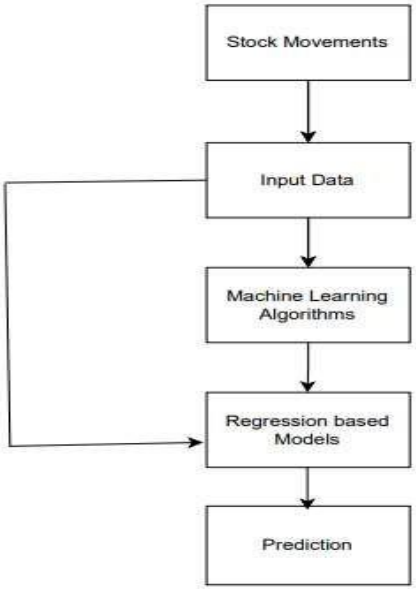


Fig. 1. Flow Chart for Regression Model

Among them, V is a fixed variable; L - independent value; We wor k with data in CSV format using the Pa nda library and calculate the parameter to be estimated, that is, the pri ce of the stock over time. To avoid problems, the information will be divided into different training sessions. A simple linear regression is performed on the data as shown in the equation given above, and the n predictions are made and the market value and time are graphically represented

V = c + d L + error

Long Short Term Memory (LSTM) Network- Based Model

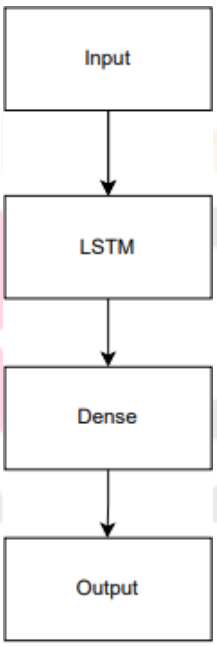


Fig. 2. Layers in LSTM

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 **higher** features compared **to** RNN. In **LSTM,** information **about** previous **states** will be available. RNN **discovers long-term dependencies as well as** relationships between **current and** recent **data. Compared to** RNN and LSTM, **RNN has less data latency.** Considering this **modeling style, the prediction** is **based** on larger **data** and **is based** on **past data [6].** To **obtain more accurate results,** LSTM **collects data from** the **original stage [7]. gradient problem. This problem can be solved using LSTM. As seen** in the **figure,** the stages in LSTM **consist of memory unit** and three **gates, namely** i/p, o/p and forget gate. For **long-term propagation, this** value **needs** to be remembered and the above **gate** will **check it [8]. Model** 256. **The output** value is fixed to **0.3 to increase the training speed and prevent overfitting. Each** neuron in the **thick** layer 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

EXPERIMENTAL RESULTS

NSE Stock data is used as dataset for training and testing purposes. **Data** is split into training and testing **at ratios of** 70 and **30.** We run our **model using AdaniPort stock price and BPCL.** We run the model for sizes **10, 20** and 30 and **compare our** results.

(a)

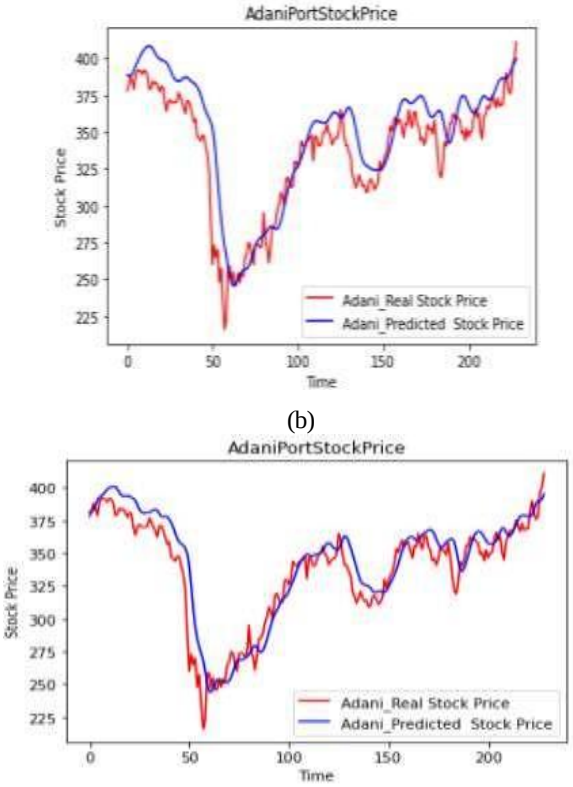
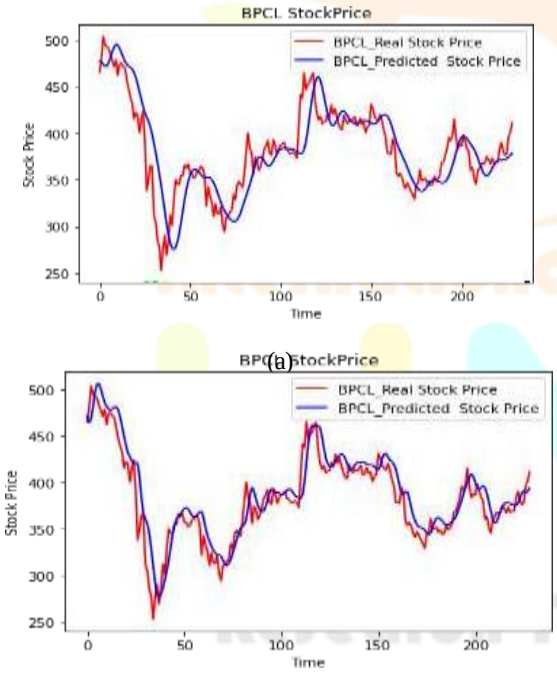


Fig. 3. *AdaniPort stock prediction for (a) 10 Epochs (b) 20 Epochs (c) 30 Epochs*



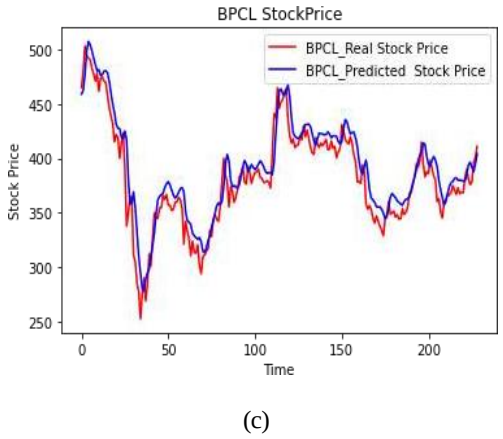


Fig. 4. *BPCL stock prediction for (a) 10 Epochs (b) 20 Epochs (c) 3 Epochs*

data having length of the batch as 60 and 10, 20 and 30 epochs, we've got plotted. crimson color line indicated the prediction and blue color line suggests the real trend. the two lines indicate the performance of LSTM version. when sizeable quantity of time is surpassed the tendencies may be discovered. If we need the maximum precision we must teach the device extra. If the quantity of epochs is greater, then the real and predicted inventory price is overlapping. where as within the case much less epochs the model is little bit deviating which isn't always giving real end result.

IV. CONCLUSION

The stock market price **prediction was proposed in the paper. The proposed** Long Short Term Memory network is compared with machine learning models. **The** models are tested on the bench **mark. The** proposed method **showed** good performance and **helped** to **get** the optimistic results. **The** regression analysis **shows** an increase in **accuracy. The use of** machine learning techniques in giving and taking **schemes has been marked by promising results. The accuracy is achieved by** using Machine Learning **models. In** the future need to check for accuracy and **time with the models.**

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.