iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 **Issue:** VIII **Month of publication:** August 2021

DOI: https://doi.org/10.22214/ijraset.2021.37791

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VIII Aug 2021- Available at www.ijraset.com

Prediction of Stock Price using RNN's LSTM-Based Deep Learning Model

Asmita Pandey[1], Roshni Salian[2]
[1, 2]*Student, Department of Information Technology, B. K. Birla College of Arts, Science and Commerce (Autonomous), Kalyan 421 304, Maharashtra, India*

Abstract: ***Stock Market is referred to as a trading platform where trading of listed companies share price is exchanged. It is a place where individuals can buy or sell shares of the publicly listed companies. The prediction of stock market that how it will perform, its movement is one of the challenging tasks to do. Stock market prediction involves determining the future movement of the stock value of a financial exchange. In this paper the prediction of the stock prices using deep learning's LSTM (Long Short-Term Memory) which is the extension of Recurrent Neural Network is done. The previous two years historical dataset from 31/7/2019 to 13/8/2021 is taken for the prediction purpose. The prediction is based on the time series analysis of data, since it can help us to get an idea of the stock price pattern and also it is considered to be the best tool for understanding the pattern of the previously observed values and make the predictions based on it. For a greater accuracy of the predictions, we should consider past happenings or events as the past affects the future. Since for stock market prediction the data will be in time series and LSTM performs well when the information or the data is of the past and the prediction is to be made for the future then we can say that LSTMs are quite capable of doing the prediction for the stock market values.***
Keywords: Stock Market, prediction, LSTM, Recurrent Neural Network, time series analysis

I. INTRODUCTION

A stock market, equity market or share market is basically a collection of buyers and sellers of stocks where investors connect to buy and sell investments (most commonly stocks). Stocks, also known as equities, symbolizes a part of ownership of a company by an individual or a group of people [6], and the stock market is a place where ownership of investible assets can be bought and sold by the investors. Stock market is a great way for individuals to put their idle investments to work by investing it in quality companies. It defines as the place where shares of public listed companies are traded. The rise and fall in the stock prices will reflect with the relation to a number of different factors which includes economy change altogether, occurrence of political events, environmental changes, changes in the industries and also the changes within the company can cause stock prices to move in either directions. A stock exchange opens the door for stock brokers to trade company stocks and other securities. Buying and selling of stocks might be done only if those stocks are listed on an exchange. Hence, stock exchange is referred to as the meeting place of stock buyers and sellers. India's most leading stock exchanges are the Bombay Stock Exchange (BSE) and the National Stock Exchange (NSE).

The attempts for forecasting the future value of stock exchange of an individual stock or a particular sector or the market, or the market as a whole is known as a stock market prediction. A prediction accounts as an essential opinion for investors to consider for having successful trades leading in confidence of knowledge regarding when to buy and sell the stocks in appropriate time frames. The un-predictable nature of the stock market results in hard decision making on predicting the stock prices [17]. According to the efficient market hypothesis (EMH) it is impossible to predict stock prices accurately and there are many advocates who support this hypothesis as well, but then there are many theories and experiments done by researchers and many experts which demonstrates that precise modeling and designing of relevant variables may produce models to predict or forecast stock prices approximately or maybe even accurately [21]. Stock market attracts investors leading into economic growth. Conventionally there are three techniques for predicting stock prices that is fundamental analysis, technical analysis and machine learning but now there are additions to these techniques such as market mimicry, time series aspect, deep learning, sentiment analysis etc. For stock market analysis, data sizes are usually huge and also non-linear in nature. So, to deal with it we need a variety of data efficient models which can identify hidden patterns and complex relations in all those large data sets taken for prediction purpose [14].

Stock prices are not casually determined values rather it can be handled as a separate time series model. This paper focuses on building an efficient LSTM based deep learning model to analyze time series data and predict the future stock price movement. Time series analysis is helpful for getting an idea about stock price pattern movements. It is used to predict future values based on the data values which are previously observed. It is believed to be one of best tools for future prediction and trend analysis.

©IJRASET: All Rights are Reserved

Prediction of closing price of the Tesla stock given the highest, lowest, and closing price by using Deep Learning, RNN-LSTM is done. Long Short-Term Memory neural networks usually (LSTM) are a special kind of Recurrent neural network (RNN) because they are capable of learning long-term dependencies which the traditional RNN couldn't. LSTM cell (known as the memory part of LSTM unit) has 5 essential components which allows it to model both long-term and short-term data. Components include cell state, Hidden state, Input gate, Forget gate, and Output gate. Gates of the LSTM are the regulators which manages the flow of information inside the LSTM unit.

II. RELATED WORK

Every single day 5000+ trade companies that are listed in Bombay stock Exchange (BSE) offers an average of 24,00,00,000+ stocks thus making an approximate of 2000Cr+ Indian rupees in investments. Hence analysis of such a huge market would be beneficial to all the stakeholders of the system [11]. The researchers Reddy et al. [1], proposed a Support Vector Machine model based on Machine Learning which worked on the large dataset value which was collected from different global financial markets. Compared to the selected benchmarks the model generated higher profit. Usmani et al. [2] did a Prediction of Karachi Stock Exchange on a day closing price using Machine learning algorithms where the attributes used as an input where the oil rates which was found to be relevant. They tried finding whether the combination of different techniques that includes statistical, analytical and data mining techniques can predict stock market or not and up to what accuracy level. The methods or models that had used were Single Layer Perceptron (SLP), Multi-Layer Perceptron (MLP), Radial Basis Function (RBF), Support Vector Machine (SVM) etc. The best technique out of all mentioned was found to be MLP algorithm.

Researchers Sriram et al. [8] performed an assessment of methods for predicting the accuracy in stock price variations. They focused on predicting stock values using logistic regression-based machine learning algorithm which was found to be ideal among the different algorithms for predicting the market price of a stock based historical data. Researchers Mankar et al. [5] tried building a system that forecasts the stock price movements of various companies using sentiment analysis method. Sentiment analysis here was based on tweets which were collected using twitter API and also the closing values of various stocks were taken into consideration. The most feasible model in predicting stock price movement in favor of sentiments of the tweets was found to be the SVM model. Researchers Vijayarani et al. [7] proposed a framework for the stock market price prediction utilizing the random forest algorithm and predicted the object variable, which is the price for a given day. They measured accuracy of various algorithms and the most reasonable one which is dependent on different data focused from the historical data was the random forest algorithm. Hernández-Álvarez et al. [9] conducted an accuracy analysis to determine how useful supervised machine learning algorithms could be in the financial field and found out that the forecast using the Random Tree Algorithm gave better results for the close attribute for a test set unlike MLP.

Parmar et al. [3] used Regression and LSTM models to predict stock values, where Regression involved minimizing the error and LSTM contributed for remembering the data and results for the long run. Improvement in the accuracy of predictions were seen for both the techniques, where the results were positive. Researcher Dr. Karunakar Pothuganti [10] presented a Recurrent Neural Network and Long Short-Term Memory based models to deal with the anticipated stock market files in order to predict the daily returns of each company. LSTM was considered the best foreseeing algorithm for given datasets. Authors Mandi & Skanda [19] considered a benchmark model of linear regression and compared it with basic as well as improved LSTM and found that the improved version of LSTM worked the best. A Comparison study was made between ANN and RF models by researchers Vijh et al. [14], minimizing forecasting error by treating the stock market analysis was focused. Mokalled et al. [13] studied about applying AI techniques using market and news data and proposed an automated trading system which integrates mathematical functions, ML and other external factors for better stock prediction accuracy. A Model independent approach was proposed by Researchers Selvin et al. [16] by not fitting data to a specific model but rather identifying the latent dynamics existing in the data. RNN, LSTM and CNN were used. Reason behind using three models were to identify whether there is any long-term dependency existing in the given data which can eventually be identified from the performance of these models. Sliding window approach with data overlap was used for prediction of future values on short term basis. CNN here proves to give more accurate results than other two models since it only uses the current window for prediction. During the covid-19 pandemic, the stock price volatility got increased majorly leading to the need of critical price forecasting. Researchers Mottaghi, & Farhangdoost [17] evaluated how well different forecasting models i.e. the Autoregression model, Long Short-Term Memory (LSTM), XGBoost, and Last-Value performs against the data during the Covid-19 pandemic and found Autoregression and the Last Value models to be the best.

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VIII Aug 2021- Available at www.ijraset.com

III. METHODOLOGY

This paper aims to propose a model which gives higher accuracy for predicting stock prices. Proposed model is based on a special kind of Recurrent Neural Network that is Long Short-term Memory. We tried making the model predict the stock prices using historical datasets with its most possible accuracy.

A. Recurrent Neural Networks

Recurrent networks differ from the traditional feed-forward networks in the sense that they don't only have neural connections on a single direction that means neurons can pass the data to a previous or the same layer. Thus, data doesn't flow in a single way in this type of networks and hence Recurrent neural networks are known as short term memory neural networks [22]. RNN can use their internal memory to process arbitrary sequence of inputs, which the feed-forward networks couldn't. Each of the RNN's computing unit consist of varying real valued activation and modifiable weights. LSTM's are networks having loops in them which allows information to pass on. The creation of RNNs is done by applying the same weights over a graph-like structure in a loop [16]. RNNs are networks having loops in them which allows information to pass on. In RNN, series of repeated operations are done over and over on a set of data. Patterns and sequences are identified and unfold and are then interpreted to predict upcoming likely outcomes. The output of RNN unit depends on the current input as well as the previous hidden state which carries the past information. Vanishing gradient problem occurs in RNN because of the repeated updation of the hidden states in the feedback loop. To overcome this limitation of vanishing gradient, variant of RNN LSTM (Long Short-Term Memory) was introduced.

B. Long Short – Term Memory

Long Short-Term memory usually called LSTMs which is used in this study is a type of recurrent neural network (RNN) [20]. LSTMs basically are the gated recurrent neural networks which is designed to overcome the vanishing gradient problem by storing the information from the past layers in the cell state. The cell state is an essential part of LSTM's any information flows very easily along it and remains unchanged. It is also known as the memory cell. LSTM's have the ability to add or remove information to the cell state by carefully monitoring it by structures called gates. A typical design of LSTM contains an input gate, an output gate, a forget gate and a memory cell. we can take the input design from the input gate, forget Gate decides which part of the information is to be taken to proceed and an Output Gate contains generated output by the LSTM.[16] LSTM's can be used in time series analysis. The past time variable of the LSTM are used as the observations. It consists of multiple components some of the main are memory cell, forget gate, and input gate [17]. As stock market prediction involves processing of huge amount of data, gradients with respect to the weight matrix have chances to get degraded and become very small and thus leading to the problem of vanishing gradient, But LSTM makes sure to prevent this problem [3]. LSTM is a kind of sequential model.

C. LSTM-Based Prediction Model

The procedure for building the model to predict stock prices involves various steps. This section will cover the description of the proposed model. The first step involves the data collection and its analysis. The historical stock data of TSLA is considered for prediction purpose. For getting this data from the web, a python library called pandas data_reader is used which reads the data from Yahoo! Finance. Yahoo! Finance is useful site for obtaining variety of fundamental data of any company we are looking for. It is a great website to see the current state of the market or for looking up for certain stock information. The data contains 515 records of daily stock prices of the tesla from the dates 31/07/2019 to 13/08/2021. Every record consists the information of high, low, open and close and adj close values of stocks and also the volume of the stocks sold on the respective days. Below Figure 1a and Figure 1b shows the first 5 and last 5 records of TSLA stock data.

Date	High	Low	Open	Close	Volume	Adj Close
2019-07-31	49.335999	47.330002	48.599998	48.321999	45891000.0	48.321999
2019-08-01	48.902000	46.354000	48.529999	46.770000	41297500.0	46.770000
2019-08-02	47.254002	45.846001	46.270000	46.868000	30682500.0	46.868000
2019-08-05	46.273998	45.155998	45.919998	45.664001	35141500.0	45.664001
2019-08-06	46.500000	45.150002	46.375999	46.150002	27821000.0	46.150002

Fig. 1a

©IJRASET: All Rights are Reserved

Date	High	Low	Open	Close	Volume	Adj Close
2021-08-09	719.030029	705.130005	710.169983	713.760010	14715300.0	713.760010
2021-08-10	716.590027	701.880005	713.989990	709.989990	13432300.0	709.989990
2021-08-11	715.179993	704.210022	712.710022	707.820007	9800600.0	707.820007
2021-08-12	722.799988	699.400024	706.340027	722.250000	17459100.0	722.250000
2021-08-13	729.900024	714.340027	723.710022	717.169983	16698900.0	717.169983

Fig. 1b

The complete information of the dataset is available in Figure 1c.

```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
DatetimeIndex: 515 entries, 2019-07-31 to 2021-08-13
Data columns (total 6 columns):
 #   Column     Non-Null Count  Dtype
---  ------     --------------  -----
 0   High       515 non-null    float64
 1   Low        515 non-null    float64
 2   Open       515 non-null    float64
 3   Close      515 non-null    float64
 4   Volume     515 non-null    float64
 5   Adj Close  515 non-null    float64
dtypes: float64(6)
memory usage: 28.2 KB
```

Fig. 1c

The chosen column is the close price which will act as the target for our predictive model. The columns high, low and open can be selected as the target for the prediction as well. Figure 2 shows the movements of the close price column of the Tesla stock data.

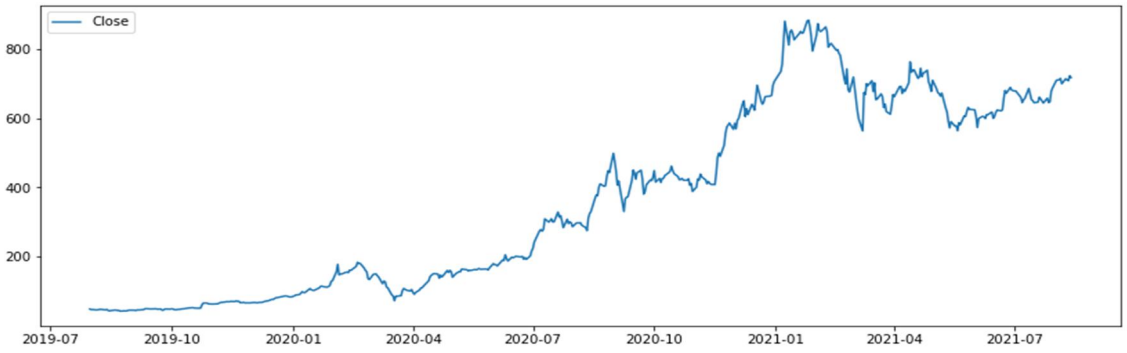


Fig. 2

LSTM's are really sensitive to the scale of the data. So, it's a good practice to always scale or normalize the input data before providing it to the neural network. MinMaxScaler was used to transform the data values between 0 to 1. Suppose if you want to scale down the data between [-1,1] you can give that value as well. For this, sklearn has a great preprocessing library which is capable of doing this i.e. MinMaxScaler.

As it is a time series problem we have taken 60 timesteps for the model to learn. The first 60 closing values goes as the input data and the 61st value is the output data. The next step is the splitting of data where we have considered 80% as training dataset and 20% as testing dataset. The range of the training data is from July 2019 to approx. April 2021 and the test data ends at August 2021. Figure 3 depicts the summary about the built LSTM model.

©IJRASET: All Rights are Reserved

```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
lstm (LSTM)                  (None, 60, 128)           66560
_________________________________________________________________
lstm_1 (LSTM)                (None, 64)                49408
_________________________________________________________________
dense (Dense)                (None, 25)                1625
_________________________________________________________________
dense_1 (Dense)              (None, 1)                 26
=================================================================
Total params: 117,619
Trainable params: 117,619
Non-trainable params: 0
_________________________________________________________________
```

Fig. 3

While compiling the LSTM model, MSE (Mean Squared Error) was used for the loss function and the optimizer used was Adam which is a well performing optimizer. The model was fitted to run on 10 epochs.

IV. RESULTS

Once the model training is completed, the loss per epoch which the model stores after every epoch was plotted. The graph of Figure 4 depicts the same.

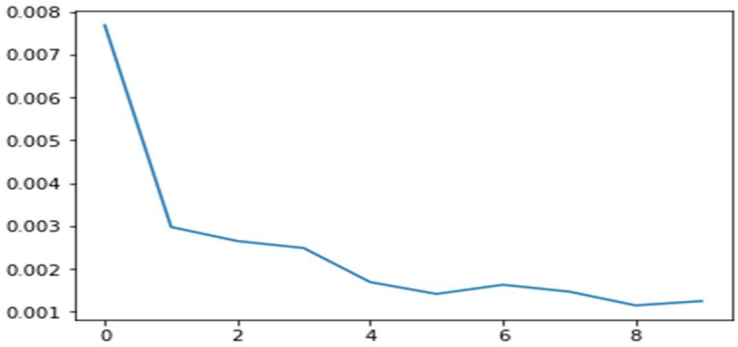


Fig. 4-Loss per Epoch

It can be seen that the loss has decreased with every epoch. It started with a loss of 0.0077 and after running all the epochs the loss was found to be 0.0012 which is a great thing.
The predictions were done on the scaled values, so the prediction was then reversed to obtain the below graph.

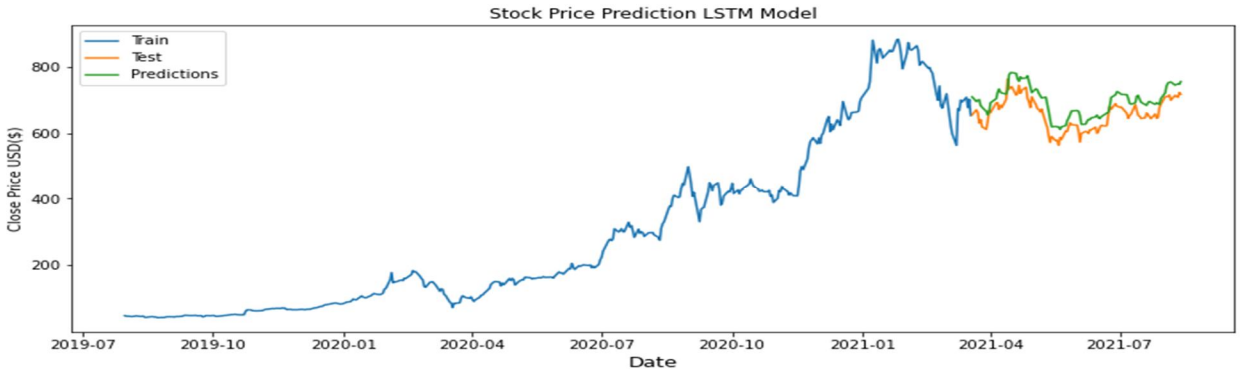


Fig. 5-Actual vs Predicted

In Figure 5, the blue color indicates the close price of the training data, while the orange color shows the close price of the testing data. And at last, the green color indicates the predictions done on the testing data which appears to be quite accurate. The metrics used to determine the accuracy of the model are RMSE (Root Mean Squared Error) and the MAPE (Mean Absolute Percentage Error). The RMSE of the predicted value is 23.1701 and the MAPE value is 2.7654 providing us a great accuracy.

©IJRASET: All Rights are Reserved

V. CONCLUSION

In this paper, we predicted the historical stock prices of Tesla using a deep learning LSTM model which is a type RNN. The strategy to divide the stock data into 80% training and 20% testing improves the accuracy results of the proposed model. In our proposed model, RMSE and MAPE acts as our model evaluation metrics since both are useful for measuring prediction accuracy and we got the values as 23.1701 and 2.7654 respectively giving us a major level of accuracy. And as we can see from Figure 5 the graph of the predicted and actual price is pretty close which overall brings us to the conclusion that LSTM's are a great choice for stock price predictions. In the future, the stock price prediction can be done much accurately with the usage of a much bigger dataset than that of what we have used in this paper and also by adding more parameters and elements to our dataset will also help improving the accuracy rate of prediction [7].

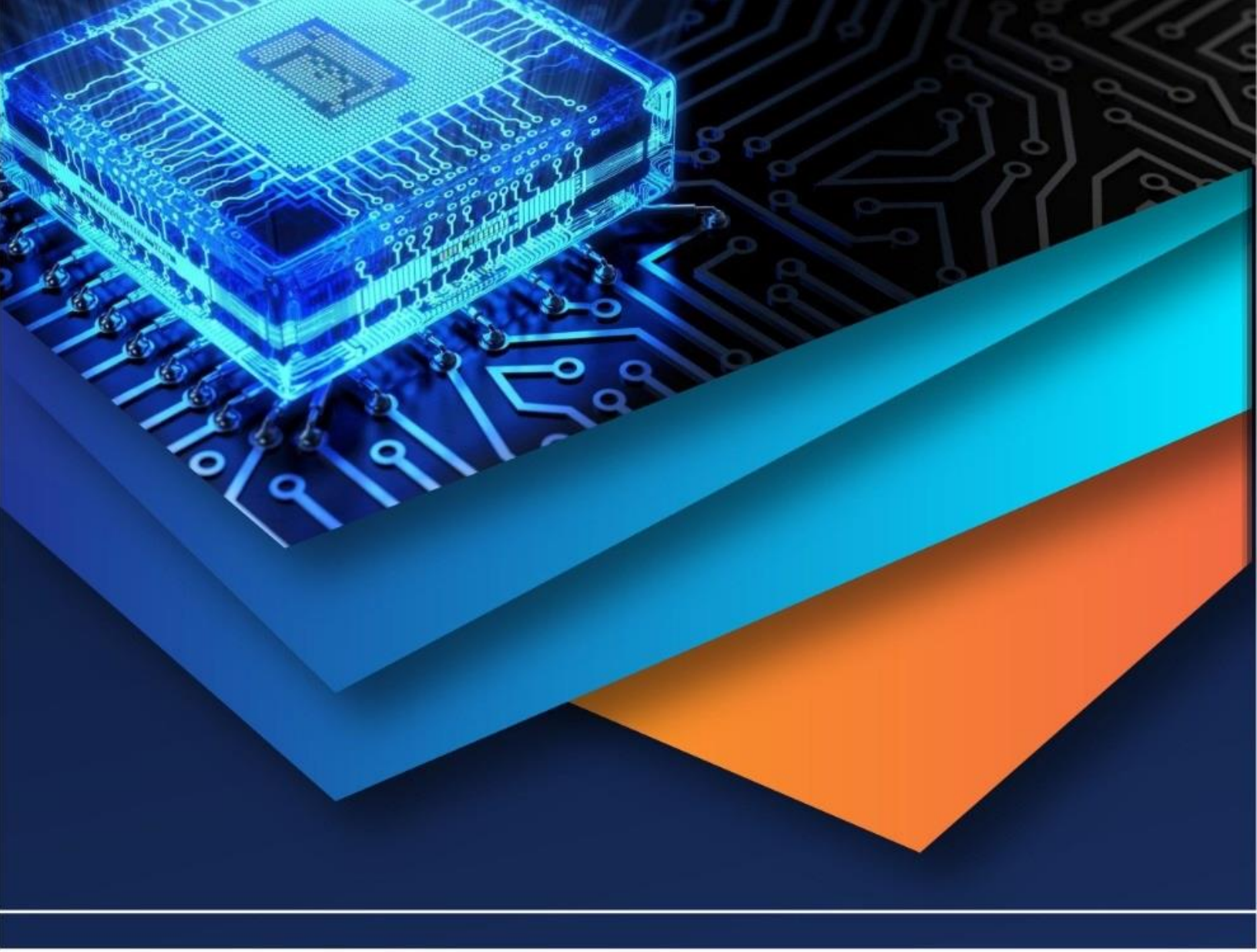
VI. ACKNOWLEDGMENT

We are truly thankful to B.K. Birla College Arts, Science and Commerce (Autonomous), Kalyan for providing us an opportunity to write a research paper as our academic performance. We also would like to thank Professor Swapna Augustine Nikale for her constant assistance during the course of this research.

REFERENCES

[1] Reddy, V. K. S. (2018). Stock market prediction using machine learning. International Research Journal of Engineering and Technology (IRJET), 5(10), 1033-1035.
[2] Usmani, M., Adil, S. H., Raza, K., & Ali, S. S. A. (2016, August). Stock market prediction using machine learning techniques. In 2016 3rd international conference on computer and information sciences (ICCOINS) (pp. 322-327). IEEE.
[3] Parmar, I., Agarwal, N., Saxena, S., Arora, R., Gupta, S., Dhiman, H., & Chouhan, L. (2018, December). Stock market prediction using machine learning. In 2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC) (pp. 574-576). IEEE.
[4] Agrawal, S., Thakkar, D., Soni, D., Bhimani, K., & Patel, C. (2019). Stock market prediction using machine learning techniques. International Journal of Scientific Research in Computer Science, Engineering and Information Technology, 5(2), 1099-1103.
[5] Mankar, T., Hotchandani, T., Madhwani, M., Chidrawar, A., & Lifna, C. S. (2018, January). Stock market prediction based on social sentiments using machine learning. In 2018 International Conference on Smart City and Emerging Technology (ICSCET) (pp. 1-3). IEEE.
[6] Emioma, C. C., & Edeki, S. O. (2021). Stock price prediction using machine learning on least-squares linear regression basis. Journal of Physics: Conference Series, 1734, 012058. https://doi.org/10.1088/1742-6596/1734/1/012058
[7] Vijayarani, S., Suganya, E., & Jeevitha, T. PREDICTING STOCK MARKET USING MACHINE LEARNING ALGORITHMS. IRJMETS, December-2020.
[8] Sriram, S., & Rangarajan, S. Stock Market Prediction using Logistic Regression Analysis-A Pilot Study.
[9] Hernández-Álvarez, M., Hernández, E. A. T., & Yoo, S. G. (2019, February). Stock Market Data Prediction Using Machine Learning Techniques. In International Conference on Information Technology & Systems (pp. 539-547). Springer, Cham.
[10] Pothuganti, K. (2021). Long Short-Term Memory (LSTM) Algorithm Based Prediction of Stock Market Exchange. International Journal of Research Publication and Reviews, 2(1), 90-93.
[11] Pahwa, N., Khalfay, N., Soni, V., & Vora, D. (2017). Stock prediction using machine learning a review paper. International Journal of Computer Applications, 163(5), 36-43.
[12] Kompella, S., & Chakravarthy Chilukuri, K. C. C. (2020). Stock Market Prediction Using Machine Learning Methods. International Journal of Computer Engineering and Technology, 10(3), 2019.
[13] Mokalled, Wassim El-Hajj Mariam, and Mohamad Jaber. "Automated stock price prediction using machine learning." In Proceedings of the Second Financial Narrative Processing Workshop (FNP 2019), pp. 16-24. 2019.
[14] Vijh, M., Chandola, D., Tikkiwal, V. A., & Kumar, A. (2020). Stock closing price prediction using machine learning techniques. Procedia Computer Science, 167, 599-606.
[15] Parray, I. R., Khurana, S. S., Kumar, M., & Altalbe, A. A. (2020). Time series data analysis of stock price movement using machine learning techniques. Soft Computing, 24(21), 16509-16517.
[16] Selvin, S., Vinayakumar, R., Gopalakrishnan, E. A., Menon, V. K., & Soman, K. P. (2017, September). Stock price prediction using LSTM, RNN and CNN-sliding window model. In 2017 international conference on advances in computing, communications and informatics (icacci) (pp. 1643-1647). IEEE.
[17] Mottaghi, N., & Farhangdoost, S. (2021). Stock Price Forecasting in Presence of Covid-19 Pandemic and Evaluating Performances of Machine Learning Models for Time-Series Forecasting. arXiv preprint arXiv:2105.02785.
[18] Mehtab, S., & Sen, J. (2020). Stock price prediction using convolutional neural networks on a multivariate timeseries. arXiv preprint arXiv:2001.09769.
[19] Mandi, R., & Skanda, G. (2021). Analysis of Stock Market using Machine Learning. International Journal of Engineering and Applied Physics, 1(2), 162-167.
[20] Karmiani, D., Kazi, R., Nambisan, A., Shah, A., & Kamble, V. (2019, February). Comparison of predictive algorithms: backpropagation, SVM, LSTM and Kalman Filter for stock market. In 2019 Amity International Conference on Artificial Intelligence (AICAI) (pp. 228-234). IEEE.
[21] Sen, J., Mehtab, S., & Dutta, A. (2021). Stock Price Prediction Using Machine Learning and LSTM-Based Deep Learning Models.
[22] Nelson, D. M., Pereira, A. C., & de Oliveira, R. A. (2017, May). Stock market's price movement prediction with LSTM neural networks. In 2017 International joint conference on neural networks (IJCNN) (pp. 1419-1426). IEEE.

©IJRASET: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)