# Stock Price Prediction: LSTM Based Model

Ranjan Kumar Roy[1*], Koyel Ghosh[2], Dr. Apurbalal Senapati[2]

[1] Ezyequity, Kolkata.
[2] Department of Computer Science and Engineering, Central Institute of Technology, Assam.

*Corresponding author
doi: https://doi.org/10.21467/proceedings.115.19

**ABSTRACT**

Stock price prediction is a critical field used by most business people and common or retail people who tried to increase their money by value with respect to time. People will either gain money or loss their entire life savings in stock market activity. It is a chaos system. Building an accurate model is complex as variation in price depends on multiple factors such as news, social media data, and fundamentals, production of the company, government bonds, historical price and country's economics factor. Prediction model which considers only one factor might not be accurate. Hence incorporating multiple factors news, social media data and historical price might increase the model's accuracy. This paper tried to incorporate the issue when someone implements it as per the model outcome. It cannot give the proper result when someone implements it in real life since capital market data is very sensitive and news-driven. To avoid such a situation, we use the hedging concept when implemented.

**Keywords:** Stock price prediction, Yahoo Finance, LSTM, Machine learning

## 1 Introduction

In the financial domain, stock prediction is a very crucial task. Based on this prediction, future transactions influence a lot. There are others several external factors also involved here. Predicting the market's up and down is very difficult and challenging because other numerous factors may influence the stock price. Since long back, researchers from various disciplines have tried to address this issue. It involves disciplines including domain knowledge of the market, statistics, computer science, economics, operations research, etc. Generally, the domain expert analyzes the market behaviour and tries their business strategies accordingly. But for them, it is difficult to handle such market factors from various sources. To overcome such difficulties needs an automated process to incorporate all such factors, and lots of research is in progress in this domain.

## 2 Related Work

Stock market analysis prediction is an ongoing research area. Their results and predictions are directly related to the investment in the market. Several approaches have been used to predict stock movement in the past. Artificial Neural Network (ANN) is used in the paper [1] along with windowing operators; which is extremely beneficial for working with time series data for predicting stock market price and trend and it is done on Wal-Mart Stores Inc. (WMT) a listed company of the New York Stock Exchange. Five years historical dataset (2010-2015) is used to try the experiments of this study. Artificial Neural Network (ANN) can produce a rational result with a small error according to the result of this study.

To improve stock market prediction, a new automated stock trading method based on deep learning approaches is offered here [2] . The practice of machine learning methods applies Logistic Regression, Support Vector Regression, Random Forest for more accurate detection and efficiency. Historical data will be fed and a



Proceedings DOI: 10.21467/proceedings.115; Series: AIJR Proceedings; ISSN: 2582-3922; ISBN: 978-81-947843-5-7

prediction of the stock price and its movement will be outputted. In [3], they use a stock prediction analysis system using machine learning based on support vector machines (SVM), linear regression and reinforcement learning. Here, they apply SVMs for both linearly and nonlinearly separable data sets. Correlation is applied between different companies' stock prices to predict the price of a stock by using technical indicators of highly correlated stocks, not only stock to be predicted.

The final aim of paper [4] is to predict the behavior of Bombay Stock Exchange (BSE). They have taken Commodity Prices (crude oil, gold, silver), Market History, and Foreign exchange rate (FEX) that control the stock trend, as input attributes for various machine learning models to predict the behavior of Bombay Stock Exchange (BSE). Hybrid Deep Neural Network (HDNN) has been developed in [5] to predict stock market movement. They have integrated the Sugeno fuzzy inference system with the approached HDNN. The day's closing price of a stock has been predicted using HDNN based on certain factors as parameters that affect a stock's price. They have tested that model in the prediction of seven stocks and compared the result of prediction with popular similar models like Multilayer Perceptron (MLP), Generalized Linear Model (GLM), Random Forest Model (RF), Gradient Boost Model (GBM), and the Deep Neural Network (DNN). Surprisingly, the HDNN model has the best performance among all in the same task.

This paper [6] explores a popular stock message board and gets slight daily predictability using supervised learning algorithms when combining daily sentiment with historical price information. This article [7] describes, briefly and simply, the theory of random walks and some of the important issues it raises concerning market analysts' work.

## 3 Dataset

We have collected data from Yahoo finance[13] for the last ten years of data from 2010-01-01 to 2021-02-25. Yahoo! Finance data included the date open, high, low, close, volume, and adjusted close for a given day. We divided our dataset into training dataset 80% and test dataset 20%. Figure 1 shows some sampled rows of the datasets.

| | Date | Open | High | Low | Close | Adj Close | Volume |
|---|---|---|---|---|---|---|---|
| 0 | 2019-12-09 | 31356.800781 | 32105.050781 | 31352.099609 | 32014.250000 | 32014.250000 | 920200 |
| 1 | 2019-12-16 | 32159.000000 | 32443.349609 | 31897.750000 | 32384.949219 | 32384.949219 | 1796800 |
| 2 | 2019-12-23 | 32381.000000 | 32502.800781 | 31963.250000 | 32412.349609 | 32412.349609 | 1394500 |
| 3 | 2019-12-30 | 32486.599609 | 32613.099609 | 31960.400391 | 32069.250000 | 32069.250000 | 722600 |
| 4 | 2020-01-06 | 31910.449219 | 32347.199219 | 30899.550781 | 32097.400391 | 32097.400391 | 1607800 |
| ... | ... | ... | ... | ... | ... | ... | ... |
| 60 | 2021-02-01 | 30976.349609 | 36615.199219 | 30906.449219 | 35654.500000 | 35654.500000 | 1277400 |
| 61 | 2021-02-08 | 36073.851563 | 36477.148438 | 35428.148438 | 36108.898438 | 36108.898438 | 0 |
| 62 | 2021-02-15 | 36501.398438 | 37708.750000 | 35584.601563 | 35841.601563 | 35841.601563 | 0 |
| 63 | 2021-02-22 | 35874.300781 | 37232.199219 | 34658.699219 | 34803.601563 | 34803.601563 | 728800 |
| 64 | 2021-03-01 | 35374.250000 | 36455.148438 | 34983.750000 | 35802.500000 | 35802.500000 | 0 |

**Figure 1:** *Banknifty dataset downloaded from Yahoo! Finance*

[13] www.yahoofinance.com

## 4 Methodology

In this section, the LSTM model will be explained shortly. The total procedure is shown in Figure 2. We have divided our task into three parts.

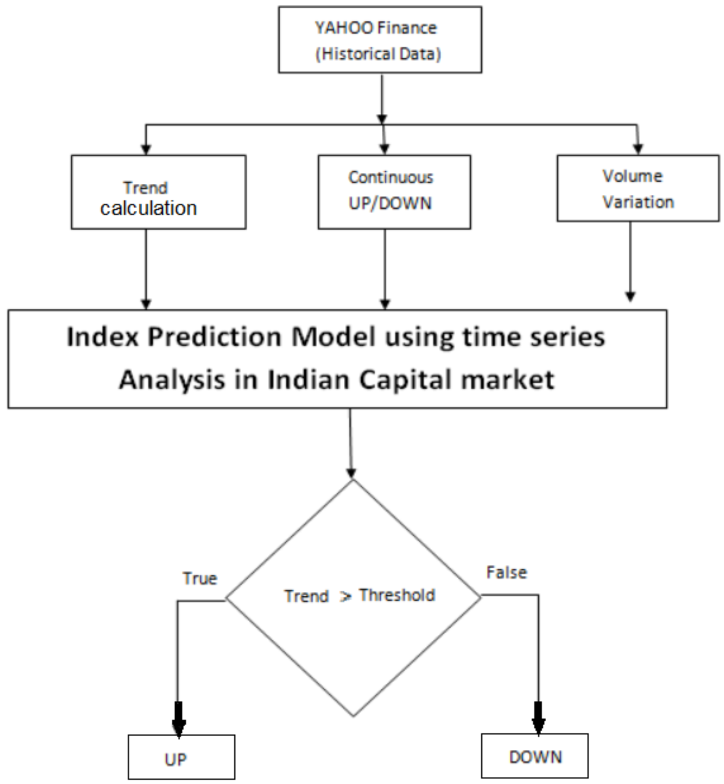


**Figure 2:** *The methodology followed for Stock price prediction*

### 4.1 Trend calculation

Closed price variations with continuous up/down for three days are considered a pattern. **Algorithm 1**, trend on each day is calculated. It is calculated by subtracting today's close price from yesterday's close price. If the obtained value is a non-negative number, then the trend is positive, otherwise negative.

**Algorithm 1:** Algorithm to calculate Trend on Each Day

```
Data:  Closing Price Vector (cPV)
Result: Trend on Each Day (T)
cV=Difference of (cPV);
append 0 to cV;
j = 0;
for i ← 0 to cV do
    if i>0 then
        t[j]=1; ;
    else
        t[j]=0 ;
    end
    increment j;
end
```

### 4.2 Identification of Continuous three days up/down

**Algorithm 2**, continuous up/down has been discussed. Continuous up/down is calculated by considering the trend of the last three days.

**Algorithm 2:** Algorithm to Check Continuous Days Up/Down

```
Data: Date vector (d)
Result: Vector Containing Relationship between Trend and Volume Traded (t)
cVt : Continuous Volume traded
i = 0;
j = 0;
for i,i+1, i+2 ← 0 to d do
    for j,j+1, j = 2 ← 0 to t do
        if j == j+1&&j == j+2&&j == j+3 then
            cVt=1; ;
        else
            cVt=0 ;
        end
    end
end
```

### 4.3 Check Market Movement Using LSTM

In the data provided by YAHOO, Index/stock volume traded on a daily basis is available. Volume traded on each day is compared with trend on the same day to get volume variation patterns. Downloaded a Historical index quote from yahoo finance then Visualises the index quote, and creates the new DataFrame with the closing price. Later scale the data to create the scaled training data set. Lastly, split the data into training and test sets.

To predict the Index price (especially Bank Nifty Index), we use the Long Short Term Memory( LSTM ) [8] model. LSTM is a special type of network adopted from the Recurrent Neural Network (RNN) [9]. LSTM is evolved to overcome the limitations of the RNN model, i.e. it is capable of learning the long-term dependencies. These networks work remarkably well in a wide variety of domains [10]. LSTMs are explicitly designed to avoid the problem of long-term dependency. The accomplishment of LSTMs remembers information for long periods being this in their basic behaviour. The LSTM neural networks contain three gates and a cell memory state. Figure 3 shows the basic structure of a standard LSTM,

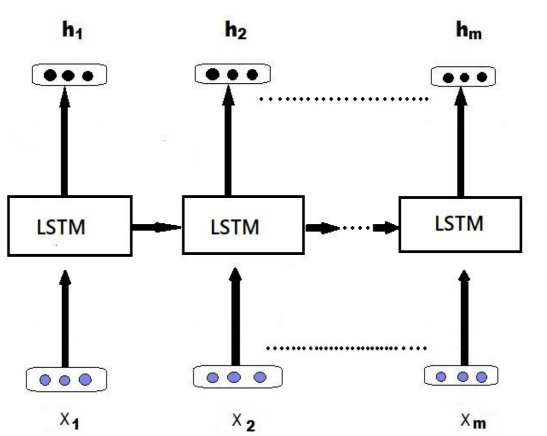


**Figure 3:** *The architecture of basic LSTM*

where {x1,x2,x3,....,xk,,....,,xm} denotes the input vector in a text, {h1,h2,h3,....,ht,...,hm} represents the hidden vector.where For a single LSTM cell, it can be computed as

$$X = \frac{h_t - 1}{w^v_{i,k}} \quad (1)$$

$$f_t = \sigma(W_f.X + b_f) \quad (2)$$

$$i_t = \sigma(W_i.X + b_i) \quad (3)$$

$$o_t = \sigma(W_o.X + b_o) \quad (4)$$

$$c_t = f_t * c_{t-1} + i_t * tanh(W_c.X + b_c) \quad (5)$$

$$h_t = o_t * tanh(c_t) \quad (6)$$

Where Wf,Wi,W0 are the weight matrices and bf,bi,b0 are the bias of LSTM cell during training. $\sigma$ Denotes the sigmoid function. $w^v_{i,k}$ is the word embedding vector as input unit to LSTM, ht is the hidden vector. Now, after calculating the market trend with the outcome of the LSTM model along with trend and volume calculation, it has been observed that market movement of 1% or more of the index (Banknifty/Nifty) the return would always be positive.

## 5 Experiments and Result

After Experiment of **Algorithm-1** (Trend calculation), **Algorithm-2** (Continuous Days Up/Down) and **Algorithm-3** (LSTM model), the outcome of the result shown below in Figure 4 using Python language to visualize understandably. Here the blue colour tells how the machine has learned with the historical data. The red colour shows the validation or testing of data on how the algorithm works, and finally, the yellow colour depicts the actual result. So in the actual result and validation result, there are no significant changes. We have taken the last 64 weeks data of Bank nifty, and it clearly shows in Figure 5 that in every week, most of the cases % of index movement is 1% or more, the output is positive. We give the threshold value more than 1% for a week, and the correlation coefficient is more than 70%-80% based on the US market and Indian capital Market of Bank nifty index movement along with the **Algorithm 1** and **Algorithm 2**.

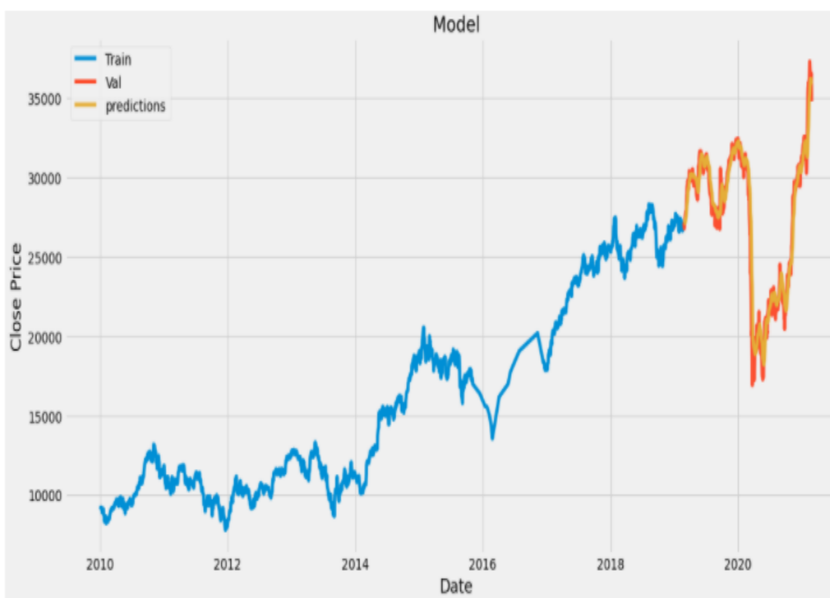


**Figure 4:** *Outcome after three steps such as Trend calculation, Continuous Days Up/Down, LSTM model*

| | Date | Open | High | Low | Close | Adj Close | Volume | Change | Prcnt_of_change |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 2019-12-09 | 31356.800781 | 32105.050781 | 31352.099609 | 32014.250000 | 32014.250000 | 920200 | 753.0 | 2.352096 |
| 1 | 2019-12-16 | 32159.000000 | 32443.349609 | 31897.750000 | 32384.949219 | 32384.949219 | 1796800 | 546.0 | 1.685966 |
| 2 | 2019-12-23 | 32381.000000 | 32502.800781 | 31963.250000 | 32412.349609 | 32412.349609 | 1394500 | 540.0 | 1.666050 |
| 3 | 2019-12-30 | 32486.599609 | 32613.099609 | 31960.400391 | 32069.250000 | 32069.250000 | 722600 | 653.0 | 2.036234 |
| 4 | 2020-01-06 | 31910.449219 | 32347.199219 | 30899.550781 | 32097.400391 | 32097.400391 | 1607800 | 1448.0 | 4.511325 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 60 | 2021-02-01 | 30976.349609 | 36615.199219 | 30906.449219 | 35654.500000 | 35654.500000 | 1277400 | 5709.0 | 16.012229 |
| 61 | 2021-02-08 | 36073.851563 | 36477.148438 | 35428.148438 | 36108.898438 | 36108.898438 | 0 | 1049.0 | 2.905093 |
| 62 | 2021-02-15 | 36501.398438 | 37708.750000 | 35584.601563 | 35841.601563 | 35841.601563 | 0 | 2124.0 | 5.926009 |
| 63 | 2021-02-22 | 35874.300781 | 37232.199219 | 34658.699219 | 34803.601563 | 34803.601563 | 728800 | 2574.0 | 7.395702 |
| 64 | 2021-03-01 | 35374.250000 | 36455.148438 | 34983.750000 | 35802.500000 | 35802.500000 | 0 | 1471.0 | 4.108709 |

**Figure 5:** *Analysis of Bank nifty of 64 Weeks Data*

To view the performance, which was executed on 19.02.2021, we applied the implementation of the strategy on 19.02.2021 when the market turned negative using time series analysis of index movement by the LSTM model. Figure 6 shows the Bank Nifty result.

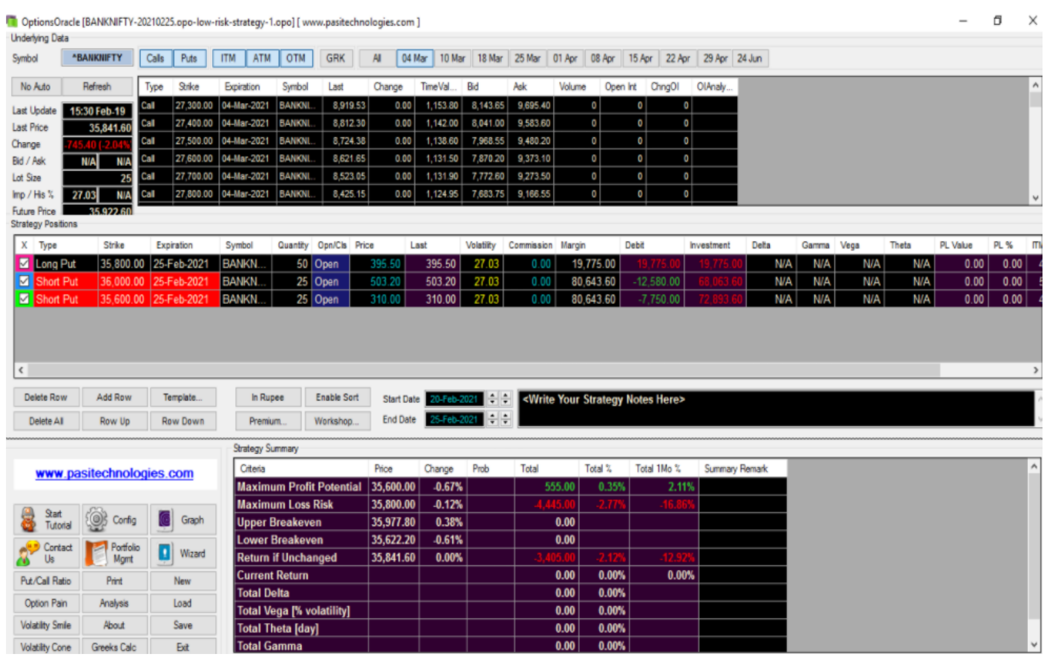


**Figure 6:**

*BUY 2 LOT PUT@395.50 OF STRIKE PRICE 35800*

*SELL 1 LOT PUT@503.20 OF STRIKE PRICE 36000*

*SELL 1 LOT PUT@310.00 OF STRIKE PRICE 35600*

After expire date on 25th Feb 2021-02-25 Figure 7
Profit: Rs 555/-
Investment amount: Rs 35,000/- – 40,000/-
Rate of Investment (ROI): 1-1.5% per week (in 7 days)

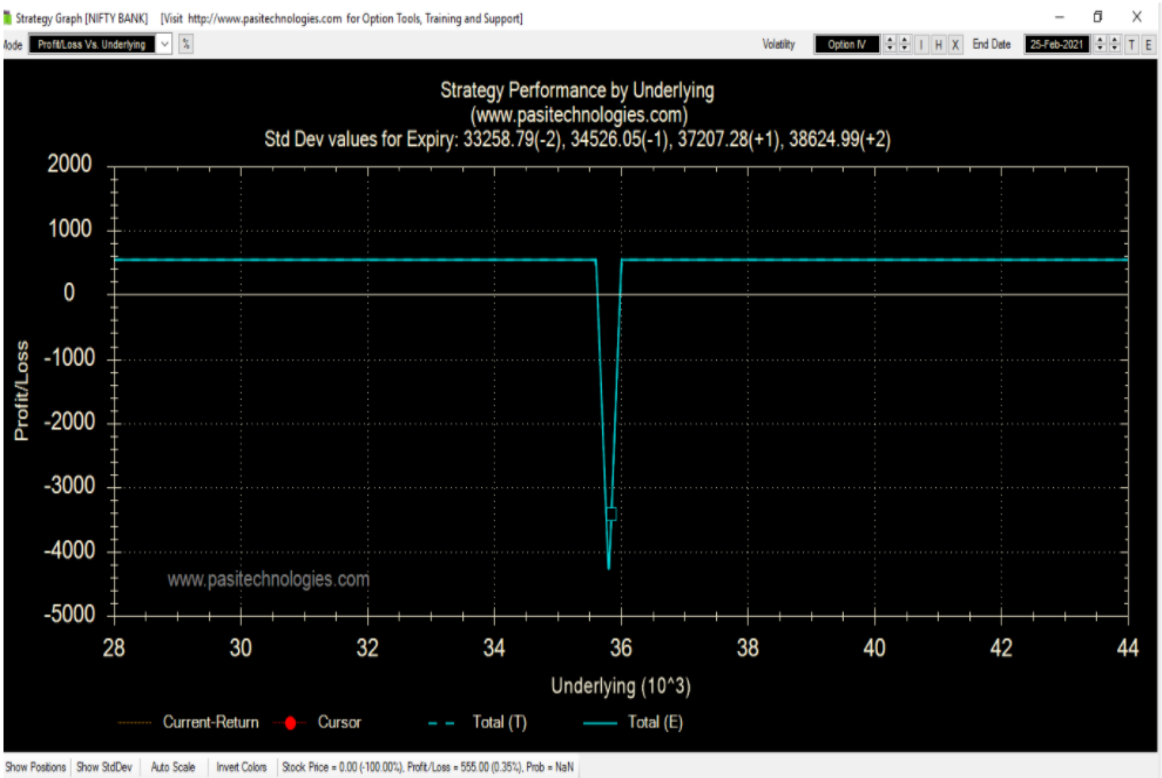


**Figure 7:** *Profit/Loss Graph*

At the time of implementation on 19.02.2021, data shows the negative movement of the Banknifty index. It shows previous days US market down and following days, i.e. on 19.02.2021 Indian market also down since correlation coefficient more than 70% but with negative bias. As we decided on the 19.02.2021 Indian market should be down, we use the PUT option. A simple term PUT option is used when the market is down, and the CALL option is used when the market is up.

**Put - Right to Sell (with example)**

Say for example a builder goes to Mr A on 1st January and pays Rs 1 lakh to builder and took commitment that Mr A will buy the house at Rs 50 lakhs on 30th March by making the payment. Builder has purchased the right to sell. By accepting Rs 1 lakh, A has sold the right. Builder may or may not exercise the right.

**Scenario 1** -Worth of house is Rs 30 lakhs on 30 th March.Builder will Exercise his right A has to pay Rs 50 lakhs and buy the house.Loss to A Rs 19 lakhs. Builder saves 19 lakhs.

**Scenario 2** – The price of the house is 70 lakhs Builder will not exercise his right. A will keep 1 lakh with him. Builder makes 19 lakhs profit. No obligation on the part of Builder to sell the house to A.

Hence, A put option is a contract granting the right to the buyer of the option to sell the underlying asset on or before a specific day at an agreed upon price, but not the obligation to do so. It is the seller who grants this right to the buyer of the option.

**Call - Right to Buy (with example)**

Say for example Mr A went to the builder on 1st January and paid Rs 1 lakh and booked a house worth of Rs 50 lakhs to be purchased on 30th March by making a balance payment.

**Scenario 1 on 30th March:**

On 30th March, if the house is worth Rs 30 lakhs only, what will Mr A do?

OPTIONS with Mr A

1. Pay the balance amount and buy the house.

- This means he is Exercising his right

2. Let his 1 lakh go and buy a similar house at Rs 30 lakhs.

- This means he is not exercising his right.

**Scenario 2 on 30th March**

– The price of the house is 70 lakhs

Mr A will pay the balance amount and buy the house.

- Means he is Exercising his right
- Made a profit of Rs 20 lakhs.
- Builder made a loss of Rs 20 lakhs. He cannot refuse to sell as A has purchased the Right to Buy.

As we implemented the strategy on Banknifty option strategy, It is always needed to do in lot size. 1 Lot size of Banknifty is 25 quantities. There are three terms used in the option.

ATM = At the money

ITM = In the money

OTM = out of the money.

Always ATM Buy/Sell should be double the quantity of ITM/OTM.

Here we take the ATM strike price to be 35800. We take 200 points ITM and OTM price. If market movement more than 200 points

(positive/negative), It should always be profitable. In Figure 7 of the Profit/loss graph, it depicts that also. Here[14] to see the performance of the result that has been implemented.

## 6 Conclusions

In the past few years, it has been observed that most people are investing in the stock market to make money quickly. At the same time investors have a high chance of losing all money invested. So an efficient predictive model is required for the user to understand the future market trend. Many predictive models tell about the market trend, whether it is up or down, but they fail to give accurate results. An attempt has been made to build an efficient predictive model of the stock market where the next day's trend is predicted. By considering various

Proceedings DOI: 10.21467/proceedings.115
ISBN: 978-81-947843-5-7

Series: AIJR Proceedings
ISSN: 2582-3922

patterns like continuous up/down, the volume traded per day, and including the company's sentiment, a model has been built and tested with various stock market data available open-source. The dataset which has been considered for sentiment analysis may be sparse, which means we may not have news/tweet for a particular company for many days. In such cases, Principal component analysis with multiple factors can be applied. The impact of intraday price movement for the next day stock price can be considered to improve accuracy.

So, in future work would be to tune the model how to predict the market movement in either positive or negative direction accurately with high conviction.

## References

[1] N. Hasan and R. I. Rasel, "Artificial neural network approach for stock price and trend prediction,"International Conference on Advanced Information & Communication Technology, Bangladesh, 05 2016

[2] A. Verma, A. Ruiz, H. Singh, and K. Elbahloul, "Automated stock trading using machine learning approaches," 07 2019.

[3] P. Gaur, R. Bhardwaj, V. Bansal, N. Kumari, and S. Gupta,"Stock market prediction using machine learning,"International Journal of Scientific Research in Computer Science, Engineering and Information Technology, pp. 206–210, 05 2019.

[4] P. Kohli, S. Zargar, S. Arora, and P. Gupta, "Stock Prediction Using Machine Learning Algorithms: SIGMA 2018, Volume 1", Advances in Intelligent Systems and Computing, 01 2019, pp. 405–414

[5] V. Kumar and M. L.,"Comparison of hdnn with other machine learning models in stock market prediction,"International Journal of Computer Applications, vol. 182, pp. 1–9, 07 2018.

[6] M. Rechenthin, N. Street, and P. Srinivasan,"Stock chatter: Using stock sentiment to predict price direction,"Algorithmic Finance, vol. 2, 01 2013.

[7] E. Fama, "Random walks in stock market prices,"Financial Analysts Journal, vol. 51, 01 1995

[8] S. Hochreiter and J. Schmidhuber, "Long short-term memory,"Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997. [Online]. Available:https://doi.org/10.1162/neco.1997.9.8.1735

[9] H. K. Sulehria and Y. Zhang, "Hopfield neural networks: A survey," in Proceedings of the 6th Conference on 6th WSEAS Int. Conf. on Artificial Intelligence, Knowledge Engineering and Data Bases - Volume 6, ser. AIKED'07. Stevens Point, Wisconsin, USA: World Scientific and Engineering Academy and Society (WSEAS), 2007, p. 125–130.

[10] B. Nath Saha and A. Senapati, "Long short term memory (lstm) based deep learning for sentiment analysis of english and spanish data," in 2020 International Conference on Computational Performance Evaluation (ComPE), 2020, pp. 442–446.