# Financial Models for Indian Stock Prices Prediction Using LSTM and Bi-Directional LSTM Model

**Manish Choubisa[1], Manish Dubey[2]**

*[1,2]Assistant Professor*

*[1,2]Department of Computer Engineering*

*[1,2]Poornima College of Engineering, Jaipur*

*Email: manish.choubisa@poornima.org, manish.dubey@poornima.org*

**Abstract**— *The Stock market is a large ocean of investors who sell and buy the stocks on daily basis resulting in drastic changes in the stock prices. The factors that affect the price of stock are the principles of demand and supply. The task of predicting the stock prices has been a tedious art on which the researchers and analysts have been working for years. The investors show a lot of interest in this field so that they can invest their assets in the right place. This can be done by knowing the future circumstances of the stock market. Precisely predicting the stockprice variation in market is a massive economic benefit. This task is mostly accomplished by analyzing by any organization; is called as basic investigation. Another strategy, which is going through a great deal of exploration work as of late, is to make a predictive algorithmic model utilizing machine learning and deep learning.*

*This paper tends to build a model using architecture of Recurrent Neural Network(RNN) named Long Short-Term Memory (LSTM) to make predictions on future values of the stock. Deep Neural Networks, being the most extraordinary advancement in Machine Learning, have been used to foster an expectation model for securities exchange. This paper also discusses about two distinct sorts of Recurrent Neural Network, LSTM and Bi-Directional LSTM model.*

**Keywords**—*Stock Market, Recurrent Neural Network, Long Short-Term Memory, Economics*

## INTRODUCTION

The Stock market is a large ocean of investors who sell and buy the stocks on daily basis resulting in drastic changes in the stock prices. The factors that affect the price of stock are the principles of demand and supply. The main reasons of buying shares of reputed companies are to sell them whenever the prices rise. The stock market has a great influence on the world of economics. The rise and fall of share prices are closely related to some Key Performance Indicator (KPI's). The five conventionally used KPI's are the opening stock price ('Open'), closing price ('Close'), lowest price of the day ('Low'), highest price of the day (`High'), and total amount of stock traded during the day(`Volume').

In the recent years, machine learning has gained a lot of attention in very field which led the investors to apply machine learning in the field of stock market and some of the trials have given very promising result.

The proposed model will act as a model in which there will be a dataset which will be used to train and test the model. The main purpose of the prediction is to reduce uncertainty associated to investment decision making.

## EXISTING SYSTEM AND THEIR DRAWBACKS

Much research work has been done in the field of stock price prediction. Some of them have been highlighted here:

In order to make predictions on trends of stock, Support Vector Machines were used to build a regression model using the historical stock data. To robustly predict the stock values, the parameters of Support Vector machines were optimized using the Particle Swarm Optimization. This study shows that SVM is a good method for prediction but the calculations of particle swarm optimization are time consuming [1].

The market emotions play a vital role in making the predictions. Hence to improve the performance of prediction Long Short Term Memory was combined with Naïve Bayesian method to work on market emotions parameter. Using this method, the trends in the financial market can be figured out which can be of completely different time scales. The opening prices of the stocks are predicted using a time series model which is a combination of emotional analysis model and LSTM time serieslearning model. This study has proved that LSTM is a good method for stock prediction [2].

There were some logic defects in the previous studies which were coped using a combination of Realtime Wavelet Denoising and Long Short Term Memory. This model was tested on Asian Stock index to predict the future index. When the results of this model were compared with the results of original LSTM model, this comparison showed that this combination is better than the previous combination [3].

To predict the stock prices a feed-forward multi-layer neural network was built using a hybrid method. This hybrid method comprises of technical analysis variable and basic analysis variables of stock market indicators and BP algorithm. This method achieves high accuracy than the technical analysis method [4].

Considering the above mentioned research works, we know that some factors and indicators of stock are related to each other; a multi-value associated neural network model will be capable of processing various related prices of the same stock and output these factors and indicators at the same time. This can be achieved by a neural network called Recurrent Neural Network. An architecture of RNN is used which is named as Long Short Term Memory (LSTM) is used for making the prediction on stock prices.

## PROPOSED SYSTEM

This paper proposes an algorithm for predicting the closing price of the day using architecture of Recurrent Neural Network called Long Short Term Memory.

### *LSTM – An Overview*

LSTM is architecture of RNN that retrieves context specific temporal dependencies for a longer time interval. Every LSTM neuron acts as a memory cell that stores information. The difference between RNN and LSTM is that the neurons of RNN accept previous hidden state and current input to process the output but the neurons of LSTM takes the old cells state and output its new cell state. The LSTM block architecture shown in the figure 1.

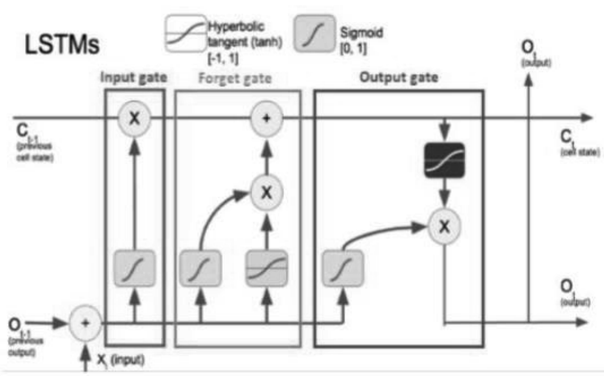


**Figure 1: LSTM Architecture**

An LSTM cell has three components:

- **Input gate:** This gate figures out the criteria for storing information in the cell state based on the input which comprises of previous output current input and previous cell state.
- **Output gate:** This gate decides which information to send forward to the adjacent node of the network based on the input and cell state.
- **Forget gate:** The forget gate decides which information we need to keep and which information should be replaced by recent information. The output value closer to 1 indicates that this piece of information should be kept and the output value closer to 0 indicates that this piece of value should be replaced.

Therefore, LSTM models are best for figuring out that how changes in a particular stock price can affect the trends of several other stocks for a longer time interval. They also decide how long should the past information be kept and when to replace it.

### *Advantages of LSTM*

The key factor about LSTM it's that it can learn context-specific temporal dependencies. Every unit of LSTM stores the information for a longer period or a shorter period without taking into account the activation function under the boundaries of recurrent components.

The most important aspect of LSTM is that any cell state is multiplied only by the output of the forget gate, which ranges in between 0 and 1. Which means the forget gate holds the responsibility of the weights of the input and activation function of the cell state. Hence information from one cell can be passed down to another cell easily without being tampered. No changes such as increasing or decreasing exponentially at each time step or layer happen. The weights get ample amount of time to converge into an optimal value. This solves the vanishing gradient problem of LSTM as the information stored does not get modified on any iteration and the gradient does not vanish when it is trained using backpropagation. LSTM does not get affected by the gaps unlike other RNN's.

### *Terminologies Used*

Below mentioned are the summaries of the f=different terminologies related to our proposed project:

- **Training set:** A part of the original dataset which is used to train the model in order to make the desired predictions.
- **Test set:** A part of the original dataset which is used to make predictions of the output values which are

then compared with the actual values in order to check the performance of the model.

- **Activation Function:** Activation function is the summation of the product of the input and their respective node.

Here the sigmoid and ReLU(Rectified Linear Unit) functions were tested to optimize the model.

a. **Sigmoid:** Represented as-

$$S(x) = \frac{1}{1+e^{-x}} \tag{1}$$

b. **ReLU:** Represented as –

$$Y = \max\ (0, x) \tag{2}$$

**Batch Size:** Number of data samples considered by the model before making further changes in the weights of the parameter.

**Epochs:** A complete iteration through the dataset done by the algorithm.

Root Mean Square Error (RMSE): This measures the difference between the predicted value and the actual value. It is calculated using the summation of the squares of the differences between to values predicted and the actual values and dividing it by the total number of samples.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(\text{Predicted} - \text{Actural})^2}{N}} \tag{3}$$

## METHODOLOGY

There are various machine learning and deep learning algorithms. But for the proposed project Recurrent Neural Network and Long Short-Term Memory is being used. This part of the paper discusses the methodology of our proposed model. The stages of the model are as follows:

### Stage 1: Raw Data

The dataset used is of Apple Inc. from May, 2015 to May 30, 2020 (this is a series of data points indexed in time order or a time series)

### Stage 2: Data Preprocessing

This stage includes-

a. Data Cleaning: The data is searched for any null values and is filled of removed using various methods.

b. Data Transformation: If the scale of the data varies a lot then the data set is normalized. Here the data is transformed using MinMax Scaler.

After cleaning and transforming the data, the dataset is divided into two parts, the training set and the testing set. The training set comprises 70% of the data whereas the testing set comprises the 30% of the dataset.

### Stage 3: Feature Extraction:

In this stage the features which are not relevant to the prediction purpose are removed and only those features are selected which are to be further fed into the model.

### Stage 4: Training Neural network

In this stage the model is trained is using the training dataset for making predictions by assigning random bias and weights. Our LSTM model is a stacked LSTM. Our LSTM model is composed of a sequential input layer followed by 3 LSTM layers and dense layer.

### Stage 5: Output Generation

At this stage, the output value generated by the output layer of the RNN is compared with the desired value. The error or the difference between the target and the obtained output value is minimized by using back propagation algorithm which adjusts the weights and the biases of the network.

## EVALUATION MEASURE

In order to evaluate the efficiency of our model we have used the Root Mean Square Error (RMSE). The difference between the desired value and the received value is decreased using RMSE value. RMSE is the square root of the average of the square of all the error. RMSE is highly recommended evaluation measure and is forms a very good error metric when the desired values are numbers. Compared between RMSE and Mean Absolute Error, RMSE perform better.

The RMSE value for the proposed project:

```
### Calculate RMSE performance metrics
import math
from sklearn.metrics import mean_squared_error
math.sqrt(mean_squared_error(y_train,train_predict))

143.77566088610203
```

```
### Test Data RMSE
math.sqrt(mean_squared_error(ytest,test_predict))

241.76692380820094
```

**Fig 2: RMSE value screen shot**

## VISUALIZING RESULTS

The proposed method is applied on the described data set and after the preprocessing the data set, we obtained the basic graph for better visualization. In the figure 3 the data set plot is shown for used dataset.

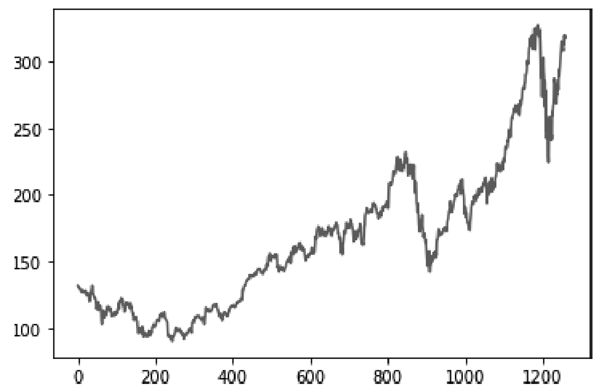

**Fig 3: Dataset on graph**

Figure 4 and 5 represented the graph after splitting the dataset into the training set and the testing set. The orange line shows the training dataset whereas the green line shows the test dataset. The orange line in the below graph shows the predicted data before inverse transformation.

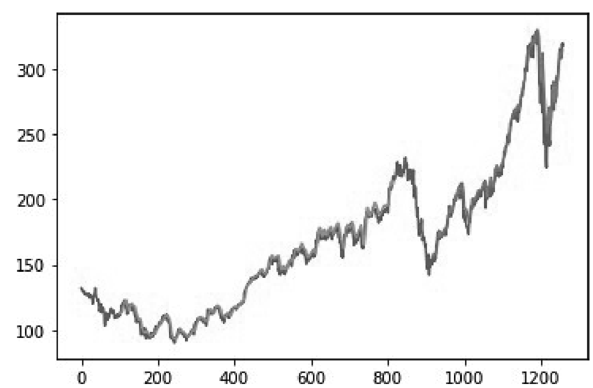

**Fig 4: Splitting dataset**

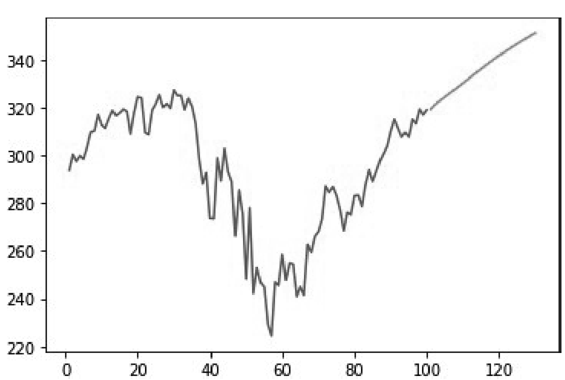

**Fig 5: Predicted values**

This is the final graph with the predictions and after inverse transformation.

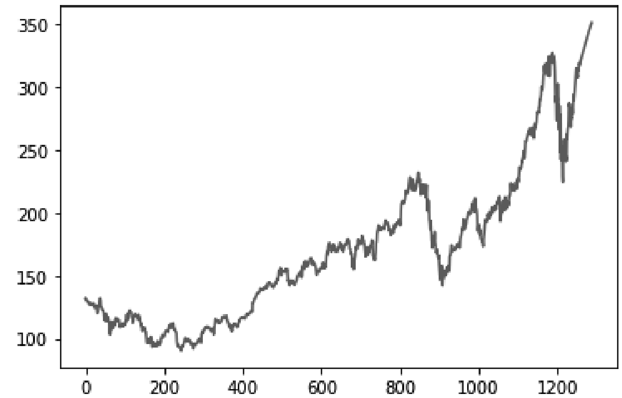

**Fig 6: Final output**

## CONCLUSION AND FUTURE SCOPE

The fame of stock market is growing day by day grabbing the attention of the researchers towards it and building models for prediction using new techniques. These models help the researchers as well as the investors to get an idea of the future scenario of the stocks so that they can use their assets wisely. Hence in order to make correct predictions, a reliable model is required having better accuracy.

The solution proposed in this paper uses architecture of Recurrent Neural Network called Long Short Term Memory which is considered to be the best for time series dataset. This helps investors and the other people who invest in the stock market by giving them a good picture of the future scenarios.

In future more parameters can be added to this project like financial ratios, multiple instances, etc. This can increase the accuracy of the model. The news about the company can also be considered as a factor in order to determine the relationship between the customer and the company.

## REFERENCES


[1] Nandakumar, Raghav, et al. "Stock price prediction using long short term memory." International Research Journal of Engineering and Technology 5.03 (2018).

[2] K. Hiba Sadia, Aditya Sharma, Adarrsh Paul, SarmisthaPadhi, Saurav Sanyal, "Stock Market Prediction Using Machine Learning Algorithms" International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Volume-8 Issue-4, April 2019

[3] Murtaza Roondiwala, Harshal Patel, Shraddha Varma, "Predicting Stock Prices Using LSTM" International Journal of Science and Research (IJSR) ISSN (Online): 2319-7064 Index Copernicus Value (2015): 78.96 | Impact Factor (2015): 6.391

[4] Zhang, Yongjie, Gang Chu, and Dehua Shen. "The role of investor attention in predicting stock prices: The long short-term memory networks perspective." Finance Research Letters 38 (2021): 101484.

[5] Huang, Yusheng, et al. "A new financial data forecasting model using genetic algorithm and long short-term memory network." Neurocomputing 425 (2021): 207-218.

[6] Chen, Qian, Wenyu Zhang, and Yu Lou. "Forecasting stock prices using a hybrid deep learning model integrating attention mechanism, multi-layer perceptron, and bidirectional long-short term memory neural network." IEEE Access 8 (2020): 117365-117376.

[7] Moghar, Adil, and Mhamed Hamiche. "Stock market prediction using LSTM recurrent neural network." Procedia Computer Science 170 (2020): 1168-1173.

[8] Hu, Zexin, Yiqi Zhao, and Matloob Khushi. "A survey of forex and stock price prediction using deep learning." Applied System Innovation 4.1 (2021): 9.

[9] Khalid, Rabiya, et al. "Electricity load and price forecasting using Jaya-Long Short Term Memory (JLSTM) in smart grids." Entropy 22.1 (2020): 10.