International Journal of Advances in
Soft Computing and Intelligent Systems (IJASCIS)
2022, Vol 01, Issue 02, 41-49



Original Paper

# STOCK MARKET PREDICTION USING MACHINE LEARNING

Darsh Maniar[1], Mukesh Rangwani[2] and Dhairya Dutt[3]
[1]Department of Electronics and Communication Engineering, Nirma University, Ahmedabad, India
19bec069@nirmauni.ac.in
[2]Department of Electronics and Communication Engineering, Nirma University, Ahmedabad, India
19bec074@nirmauni.ac.in
[3]Department of Electronics and Communication Engineering, Nirma University, Ahmedabad, India
19bec031@nirmauni.ac.in

## *Abstract*

*Stock trading is one of the most essential activities in the financial sector. The technique to analyse past data and find a matching pattern in that by training a machine is called stock market prediction using machine learning. This paper demonstrates how Machine Learning may be used to forecast the performance of a stock. The majority of stockbrokers utilise technical and fundamental analysis, as well as time series analysis when making stock decisions. To forecast the outcome, the computer language is employed. In general, python is used for the prediction of various stocks. In the stock market, which is a sophisticated and difficult process, prediction plays a very critical role. Old school techniques such as statistical and foundational analysis may not guarantee the prediction's accuracy. It suggests a Machine Learning (ML) technique that will be both efficient and effective. It has trained the machine using publicly available market data and intelligence. Using a trained computer algorithm(Machine Learning) is always preferable since it will give you advice based only on facts, numbers, and data and will not factor in emotions or prejudice. Machine learning for stock market forecasts can assist financial institutions in better managing their clients' portfolios and making educated decisions in order to maximise earnings.*

## *Keywords*

*Machine Learning, Neural Networks, LSTM, Streamlit, RNN model.*

## 1. Introduction

Traders who have a lot of money in the stock market sector buy stock derivatives and equities at a low price and then sell them at a high price. Although the tendency in stock market prediction is not new, numerous groups continue to debate the topic. Before investing in a company, investors do two types of analysis. The first is fundamental analysis, in which investors look at the inherent worth of companies as well as the performance of the industry, economy, and political atmosphere to determine whether to invest or not, and the second is a future analysis, where the role of this prediction model comes. Machine Learning analysis, on the other hand, is the examination of statistics created by market activity, such as historical prices and volumes, to determine the rising demands in stock. Many traders have been inspired to apply machine learning techniques to the industry due to the rising significance of machine learning methods in many industries in the recent past, and few of them have yielded pretty promising results. This study will construct a financial data predictor software in which a dataset including all historical stock prices will be used as the program's training sets[1]. The prediction's major goal is to lessen the amount of ambiguity around investment decisions. By

extrapolating previous data, predictive analytics for various stocks may predict how that particular stock will behave in the near future. After using the preceding analytics methodologies, prescriptive analytics acts as a feedback mechanism. Only predictive factors are considered in this article. This study explains how to use Machine Learning (ML) algorithms to conduct predictive analysis for stock prediction using data from the recent past [2]. However, there are a variety of approaches to do predictive analysis. Various methods such as Artificial Neural Networks, Deep Learning, and Long Short Term Memory (LSTM) are employed in this article. Prediction has gotten more precise and efficient as a result of the use of these algorithms. The prime aim of this project is to forecast the stock price of a Nifty 50 share [3]. The data is predicted using existing stock data, which includes the security's prior opening price (OP), closing price (CP), highest price, lowest price, adjusted closing price, and volume traded. The suggested research project's goal is to use current data to create a model using machine learning methods. It will be beneficial in predicting future results for a certain stock. The development is being done from the perspective of the consumer, so that they may invest in the stock market while minimising risk

# 2. Different Methods For Implementation of the Model

## 2.1. Regular Neural Network (RNN)

A Recurrent Neural Network (RNN) is basically part of Deep Learning involved in the Machine Learning process. RNN is a type of sophisticated neural network with internal memory that allows it to analyse extended sequences. This makes RNN ideal for stock price prediction, which requires a large amount of previous data. The three subsections that follow brim with information.

Here in equation (1) one can see that x1 is the input vector at a given time t. Also, there are two bias namely b and c (in equation (2)). In the equation W denotes input to hidden weight matrices, U denote weight to hidden weight matrices and V denotes hidden-to-output weight matrices.

$$S1 = tanh(Wx_1 + US_{t-1} + b) \quad (1)$$

$$O_t = c + VS_t \quad (2)$$

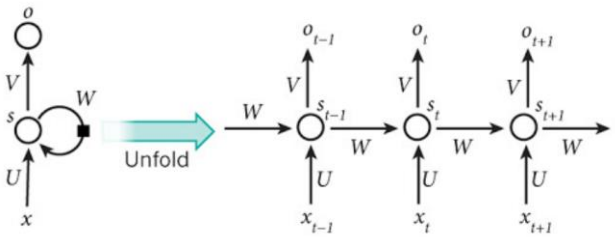


Figure 1. RNN Representation (Source:- www.researchgate.net)

## 2.2. Long Short term memory (LSTM)

The Long Short term memory (LSTM) network is basically a modified form of RNN that makes it simpler to recall facts from the past. In Regular Neural Network there was a problem of vanishing gradient. This problem is resolved by Long Short Term memory (LSTM). The architecture consists of five main parts:-

- Cell State (C1):- The values in the cell state are basically represented in the form of a vector containing arbitrary values in it. The data that was present after the previous state in the memory, the same data is present in the Cell state.

- Forget gate(f1):- It alters the state of the cell with the goal of removing unimportant values from prior time steps. This aids the Long Short term Memory (LSTM) network in forgetting unimportant and unwanted things.

- Input Gate (It):- In order to add a fresh dataset about the ongoing time step, the cell state is changed. It includes new information that might have an impact on the fluctuation of the price of various stocks.Output gate:- It helps to determine the forthcoming hidden state.

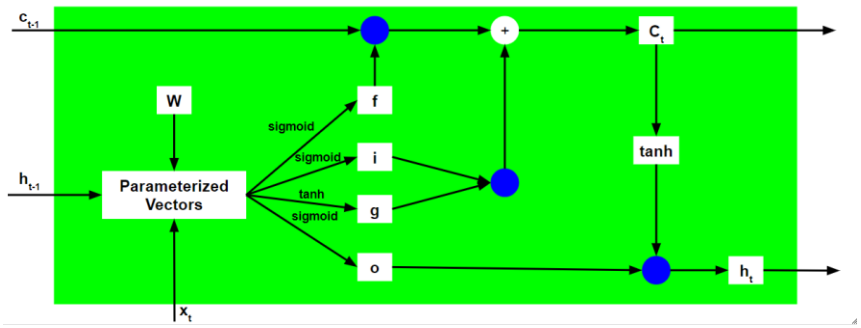


Figure 2. LSTM Execution Flowchart (Geeks For Geeks)

The above image can be understood by the below mathematical equations:- In the above figure, the blue circle represent element-by-element multiplication.

$$i_t = \sigma(W^i x_1 + U^i h_{t-1} + b^i) \quad (3)$$

$$f_t = \sigma(W^f x_1 + U^f h_{t-1} + b^f) \quad (4)$$

$$O_t = \sigma(W^o x_1 + U^o h_{t-1} + b^o) \quad (5)$$

$$c_t = (i_t)(u_t) + (f_t)(c_{t-1}) \quad (6)$$

$$h_t = (o_t)\, tanh(c_t) \quad (7)$$

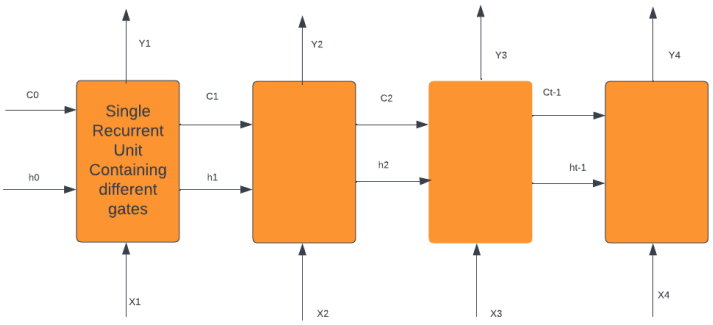


Figure 3.. LSTM Execution Flowchart-2

# 3. Significance of Models or Tools Used in the Work

## 3.1. LSTM RNN Model

The LSTM RNN is used to forecast future stock values for various companies like Facebook and TATA in this Project. Our job, which is a time series issue, is to forecast fluctuations in stock prices for a span of a few days. The LSTM model is quite popular in time-series prediction, which is why it was picked for this purpose. The historical prices of SBIN are automatically gathered using Python's nsepy module. From January 2014 to December 2018, this work includes five years of historical pricing data.

## 3.1. Streamlit

Streamlit is basically a Python tool for building Machine Learning web apps that is open-source. Using Streamlit, it can quickly create web apps and deploy them. Streamlit makes it possible to develop an app in the same manner that you would write python code. Working on the interactive cycle of coding and watching outcomes in the web app is made simple using Streamlit [6].

# 4. Advantages of LSTM Over Other Models

Long Short Term Memory (LSTM) model has the added advantage of storing information (data) for the past many years; This feature was almost missing in the RNN model. RNN could store the past value, but this model almost failed when the gap increased. LSTM has an advantage over alternative RNNs, hidden Markov models, and other sequence learning approaches due to its relative insensitivity to gap length [7]. This is crucial for a Stock Market prediction model since only by analysing a long-term dataset will it be able to generate an accurate and exact conclusion. Furthermore, by maintaining a consistent error flow throughout the backpropagation, LSTM is capable of coping with more complicated situations than RNN. In all of the instances, the LSTM outperforms the SVM [8]. This is because of its superior capacity to recall and forget data compared to SVM. Overall, the LSTM model with moving averages applied to the pooled data set was shown to be the best effective model for forecasting future stock prices. Using the LSTM advanced model with moving averages, it was also confirmed that there is a statistically significant change in predicting stock prices. As a consequence, the LSTM advanced model with moving averages is found to be the most effective in predicting stock values. However, this comes at the cost of increased complexity

and operating costs.

# 5. Algorithm Implemented

LSTMs are a modified form of Recurrent Neural networks.LSTMs are very powerful in sequence prediction problems because they can store past information and forget the information that is not important. This is important because the previous price of a stock is crucial in predicting its future price. So, the booming stock price prediction will be an excellent asset for stock market analysis. Here LSTM Sequential model has been used to predict the stock prices, which is part of the stacked items, which means different layers of the neural network are used to predict the stock price. Different optimisers can be used. Still, the optimiser that is the best in the case of LSTM is adam optimiser, which is a stochastic gradient method used to find local minima [5]. Here, the specified number of epochs, when increased, will lead to a more accurate result, but after some time, it will lead to saturation, so it can find the best optimal epochs by comparing the values of the loss function.

# 6. Results and Simulation

## 6.1. Response prediction


Figure 4. Training dataset (Closing Price)

In the above figure, the closing price data from 2014 to 2018 is provided as training data. Based on this data, an algorithm that can accurately predict the future price of that particular stock has been designed. This data is taken from the NSE(National Stock Exchange) website for Tata Beverages' stock price.

## 6.2. Predicted vs Actual Response 1


Figure 5. Predicted Value - Actual Value result 1 (With less epox value)

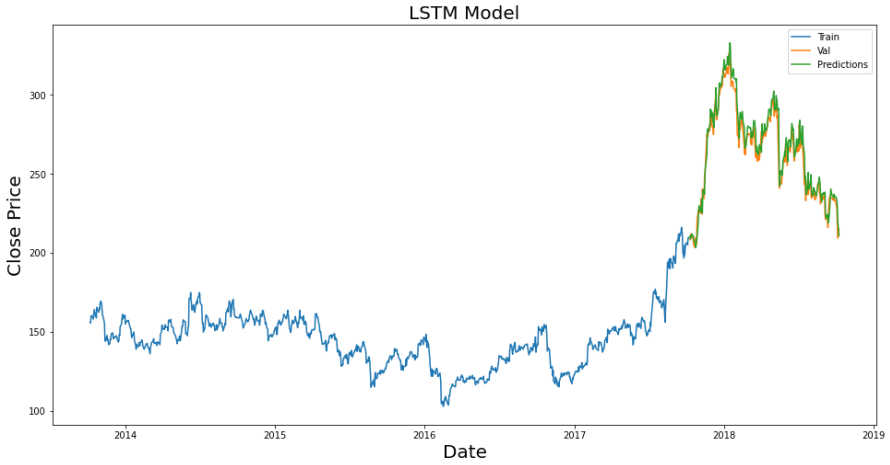


Figure 6. Predicted Value - Actual Value result 1

In the above figure 5 the number of epochs or iterations for the LSTM model is two. In contrast, in the second figure, the number of epochs or iterations are four. As a result the graph clearly shows the difference and the RMSE value for the first is 14.27 while that of the second graph is 7.19 that means the second one is more accurate because the lower the RMSE higher is the accuracy, so the increase in the number of epochs or iterations will help to increase the accuracy of the LSTM model.

## 6.3. Predicted vs Actual Response 2

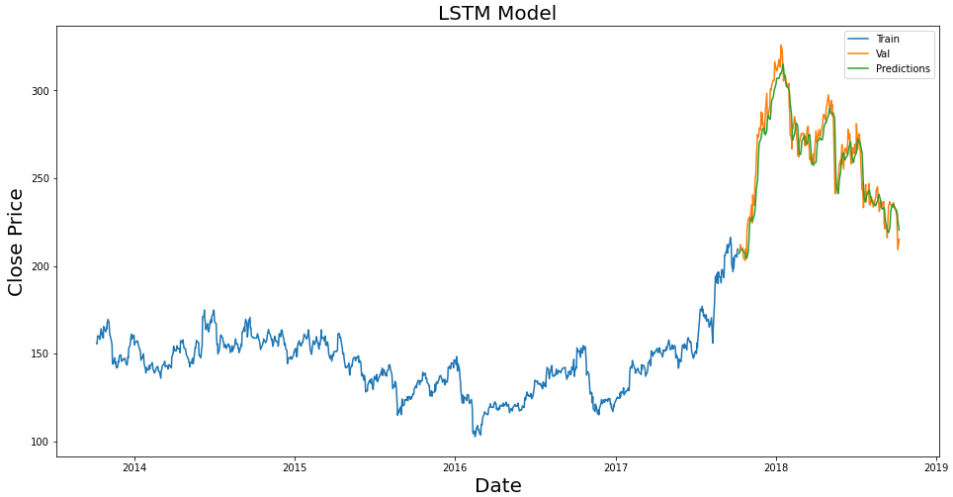


Figure 7. Predicted Value - Actual Value result 2

In figure 5 the number of hidden layers for the LSTM model is two while for the second figure 7 the number of hidden layers is three, so the graph clearly shows the difference, and the RMSE value for the first is 14.27 while that of the second graph is 8.50 that means the second one is more accurate because lower the RMSE higher is the accuracy, so increase in the number of hidden layers will help us to increase the accuracy.

### 6.4. Predicted vs Actual Response 3

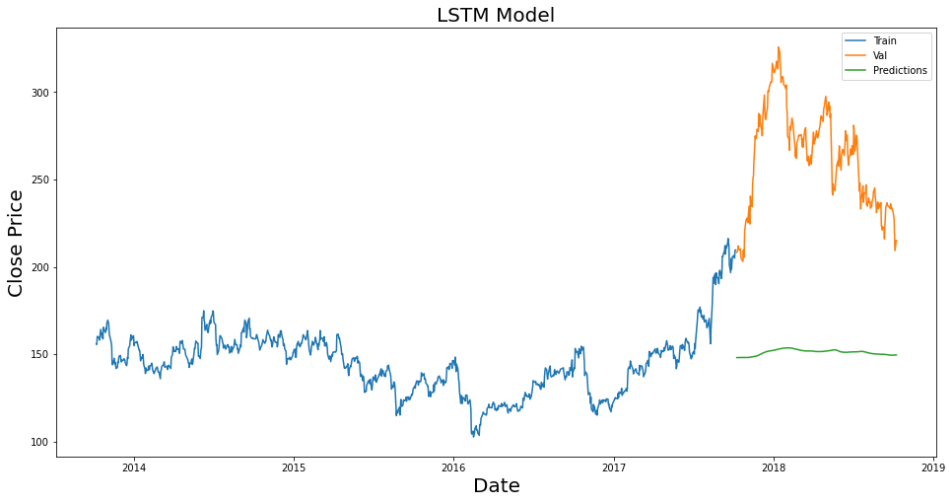


Figure 8. Predicted Value - Actual Value result 3

There are various optimisers such as Stochastic Gradient Descent(SGD),Nesterov accelerated gradient(NAG),Adagrad and adam for the purpose of machine learning especially when LSTM model is implemented. In the figure6 and figure7 result of Adam and Adagrad optimisers is compared. For adam optimiser the obtained result of RMSE is 14.27 while for Adadrad optimiser the obtained result is close to 114. As a result, Adam is a better optimiser for this particular application.

## 7. PROPOSED FUTURE SCOPE

As digital media has evolved dramatically, fluctuations in the prices of stocks is highly affected by this. As a result, Sentimental Analysis has become an important part of the Stock market prediction model [4]. There is also a strong potential for developing a more comprehensive prediction, accurate, and well-to-do machine learning model that is educated by varied sorts of information such as tweets, hashtags, sometimes fake news, news, and other text-based data by merging the newest sentiment analysis techniques with feature various other machine learning techniques [9].

## 8. Conclusion

This paper proposed combining data from several global financial markets with machine learning algorithms to forecast stock index movements. The work study's goal is to assist stock brokers and investors in making stock market investments. Due to the dynamic fluctuation nature of the stock market, an accurate prediction plays a critical part in the stock market company, which is a sophisticated and difficult procedure. Furthermore, the paper demonstrated that, despite the stacked-LSTM model's more intricate model structure compared to the single LSTM model, the stacked LSTM model does not outperform the single LSTM model owing to the risk of overfitting. Rather than providing yet another perfect LSTM model, the paper suggested a fine-tuned and customised deep learning and neural network prediction system approach with thorough feature engineering and merged it with LSTM to conduct prediction, which is unique customisation compared to earlier efforts. It intends to fill the differences between investors and academics by presenting a feature extension approach before recursive feature reduction and gain a considerable boost in model performance by

investigating observations from prior studies. Despite the fact that our proposed approach yielded a satisfactory result, this research offers more possibilities for future research. LSTM is a good technique for Stock Market Prediction. Still, it has several downsides, such as taking longer to train, using a very huge amount of memory to train, and being simply very easy to overfit.

## Acknowledgements

We would like to express our heartfelt gratitude to Mr Manish Patel, Professor at the Institute of Technology, Nirma University for helping us with understanding the topic and guiding us throughout the entire process of this research paper cum article. We could not have imagined having a better advisor and mentor for this research paper. Besides our mentor, We would like to thank Mr Sachin Gajjar, for his valuable insights and guidance in the completion of this research paper. Without his feedback, the paper wouldn't be completed. Our sincere thanks also go to the Department of Electronics and Communication at Nirma University for providing us with the material and constant support for the completion of this research work. Last but not the least, The successful completion of this Internship would not be possible without the dedicated support from all my mentors, family and friends.

## REFERENCES

[1] L. Mathanprasad and M. Gunasekaran, "Analysing the Trend of Stock Market and Evaluate the performance of Market Prediction using Machine Learning Approach," 2022 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI), 2022, pp. 1-9, doi: 10.1109/ACCAI53970.2022.9752616.

[2] Y. Bing, J. K. Hao, and S. C. Zhang, "Stock Market Prediction Using Artificial Neural Networks," Adv. Eng. Forum, vol. 6–7, pp. 1055– 1060, 2012, doi: 10.4028/www.scientific.net/aef.6-7.1055.

[3] Kumar, Gourav Sharma, Vinod. (2016). Stock Market Index Forecasting of Nifty 50 Using Machine Learning Techniques with ANN Approach. International Journal of Modern Computer Science. 4. 22- 27.

[4] T. Mankar, T. Hotchandani, M. Madhwani, A. Chidrawar and C. S. Lifna, "Stock Market Prediction based on Social Sentiments using Machine Learning," 2018 International Conference on Smart City and Emerging Technology (ICSCET), 2018, pp. 1-3, doi: 10.1109/ICSCET.2018.8537242.

[5] Z. Chang, Y. Zhang and W. Chen, "Effective Adam-Optimized LSTM Neural Network for Electricity Price Forecasting," 2018 IEEE 9th International Conference on Software Engineering and Service Science (ICSESS), 2018, pp. 245-248, doi: 10.1109/ICSESS.2018.8663710.

[6] Tawale, Shubham Jawalkar, Gaurav pathare, Sanket. (2022). STOCK PRICE PREDICTION IN PYTHON USING STREAMLIT. International Journal of Engineering Applied Sciences and Technology. 6. 170-174. 10.33564/IJEAST.2022.v06i11.032.

[7] P. Somani, S. Talele and S. Sawant, "Stock market prediction using Hidden Markov Model," 2014 IEEE 7th Joint International Information Technology and Artificial Intelligence Conference, 2014, pp. 89- 92, doi: 10.1109/ITAIC.2014.7065011.

[8] B. Panwar, G. Dhuriya, P. Johri, S. Singh Yadav and N. Gaur, ”Stock Market Prediction Using Linear Regression and SVM,” 2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), 2021, pp. 629-631, doi: 10.1109/ICACITE51222.2021.9404733.

[9] Ranco, Gabriele Aleksovski, Darko Caldarelli, Guido Grcar, Miha ˇ Mozetic, Igor. (2015). The Effects of Twitter Sentiment on Stock Price ˇ Returns. PloS one. 10. e0138441. 10.1371/journal. pone.0138441.

## Authors

Darsh Maniar

I am Darsh Maniar currently a final year undergraduate student at Nirma University pursuing my majors in Electronics and Communication Engineering. Currently, I am Interning at ISRO, Ahmedabad. At ISRO I have performed various real-time tasks related to Machine Learning. This inspired me to explore the field of Machine Learning even further which eventually lead to this Research Article.

Mukesh Rangwani

Myself Mukesh Rangwani currently I am in my fourth year pursuing my major in electronics and communication from the institute of technology, Nirma University. Recently I completed my summer internship as a software developer intern at a company namely V360 which works in a diamond market and I worked there to design a website for them which was a 360-degree diamond viewer and it will go into production. My personal interests include competitive coding, data structures and algorithms, and exploring new technologies.

Dhairya Dutt,

Myself Dhairya Dutt, am currently in my final year pursuing my BTech degree in electronics and communication. I have had 2 internships so far, one at 'eInfochips' (2 months- Software Role) and the other at 'tech table'(4 months - Associate Developer). My interests include cs fundamentals and solving real-life problems.