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

Stock Price Prediction Using LSTM

[1]Bhargavi Jaiswal, [2]Tanu Priya Kumari, [3]Vibhor Gupta

[1][2] Student, [3]Assistant Professor, Galgotias University, Greater Noida, India

[1]School of Computing Science and Engineering,
[1]Galgotias University, Greater Noida, India

Abstract: A strong algorithmic foundation is needed to determine longer-term stock prices since predicting stock values is challenging. It will be challenging to anticipate the cost because stock prices are associated with one another inside the market's nature. The suggested approach analyses market data to anticipate the stock price using machine learning methods such as a recurrent neural network called Long Short Term Memory, where the process weights for each data point are adjusted using stochastic gradient descent. In contrast to stock price prediction algorithms now in use, this method will deliver accurate findings. With various sizes of input data, the network will be trained and assessed in order to force the graphical outputs.

Keywords: Machine Learning Stock Price Prediction, Long Short-term Memory, Stock Market, Artificial Neural Networks, National Stock Exchange

INTRODUCTION

The stock market is where shares of companies with public listings are exchanged. The stock market volatility necessitates a thorough examination based on past data.

The historical time series of stock data is used by the earlier stock trend prediction algorithms, and the statistical analysis of stock data is the main emphasis of traditional scientific approaches for stock price prediction. In this study, a stock price prediction programme for a specific stock is constructed utilising historical stock prices and data as training sets. The programme then creates a methodology. This model uses as suggested by [3] the RNN (Recurrent) approach known as Long Short Term Memory (LSTM) and takes into consideration the past share price of a firm. The suggested method makes forecasts regarding a certain attribute after taking into account the stock's historical data that is currently available. The starting price, daily high, daily low, prior day price, closing price, day of trade, total trading volume, and turnover are the attributes of the stocks. The suggested approach employs time series analysis to forecast a stock price for a specific time period. The National Stock Exchange of India Limited (NSE) is the suggested design for the Indian stock market.

The National Stock Exchange (NSE) was the first stock exchange in India to deploy a contemporary trading platform.

The NSE was the first stock exchange in India to provide investors with cutting-edge services. It allows investors to trade from wherever in India and is utterly contemporary. Greater openness, convergence, and efficiency in the capital market are all benefits of the NSE's pivotal role in transforming the Indian stock market. Investors in India and throughout the world frequently utilise the CNX NIFTY, the common index of the NSE. In the equities and bond markets, as well as derivatives, it offers space for exchange, settlement, and clearing. For trading in Mazuma, currencies, and index options globally, it is one of the biggest exchanges in India. The exchange is participated in by several domestic and ecumenical businesses. Maruti Suzuki, WIPRO, HDFC, and YES BANK Ltd. are examples of regional businesses. Some critical holdings of City Party, Mauritius Limited, and Tiger Ecumenical Five holdings are included in the Pilgrim Investors group.

This is the order of the paper. Related works are mentioned in Section II. The suggested system is shown in Section III, while the proposed algorithm is represented in Section IV. The results are given in Section V, and conclusions are made in Section VI.

RELATED WORK

The information on stock market forecasting systems currently in use is considered when doing the literature review.

The calculation of stock returns has grown significantly as a field of study during the past 20 years. The majority of the time, researchers tried to demonstrate a direct link between macroeconomic information components and stock returns. However, researchers' attention has switched to the nonlinear expectation of stock returns due to the discovery of nonlinear skews in stock returns in financial markets. Although several letters claim that stock returns have been shown in a nonlinear, quantifiable manner,

most of them demanded that the nonlinear model be shown before the estimation. For the reason that financial stock market returns are impulsive, unpredictable, perplexing, and nonlinear in nature. The parameters are forecasted using a variety of methods.
The most notable of them are brown, hyperbolic sigmoid, binary threshold, and linear threshold.
A machine-learning approach to the study of stock market prediction has been mentioned. An increasingly hot topic of interest is stock market forecasting. One of them is a single evaluation, but it doesn't yield findings that can be trusted, therefore it's critical to establish methods for doing assessments that are steadily increasingly accurate. All of the backslide's methods have unique ideal circumstances and challenges to overcome from various allies. Straight backslide models always function by using the least squares approach to fit them, but they can also function in other ways, such as by reducing the "non-appearance of fit" in another standard or by reducing a disabled variant of least squares setback work. Again, fitting nonlinear models may be done using the least squares method.
An increasing example is the influence of financial ratios and technical analysis on the forecasting of stock prices utilising random forests, AI, and frameworks developed by humans. Every day, an increasing number of professionals devote their time to researching methods that might increase the stock prediction model's precision even further. Because there are so many options, it is possible to determine the best approach to predict the price of a stock in a number of different ways, and not all methods function similarly.
Regardless of whether a comparative education file is employed or not, the yield varies between methods. In the aforementioned study, the stock valuation was completed using the self-aware Timberland numbers to calculate the price of the shares using the fiscal ramifications of the structure of the upcoming quarter. This is simply one method for aesthetically crushing the situation by applying a revealing model that uses erratic soils to forecast the stock's future price based on the provided data. The accuracy of the stock value prediction model can be increased by using the financial size close to a model that can separate assumptions significantly. Nevertheless, there are constantly new factors that affect the price of the stock, such as suspicions of the financial authority, general estimates of the association, news from different bodies, and even events that change the entire trade protection. Additionally, [1] emphasises how laborious it is to anticipate stock prices using multi-source, multi-instance learning techniques. However, the Internet has proven to be a useful tool for making this task less challenging, as it is undoubtedly not difficult to immediately evacuate certain propensities due to the related approach of the data, and it is less difficult to establish associations between different variables, in most cases a case of adventure. By using a variety of options for specific legitimate data and employing various strategies, such as using a feelings analyzer to suggest a notable relationship between people's emotions and how they are influenced by express stock enthusiasm, household trading information can be adequately predicted.
Extrapolating significant news events from the web to see how they affect stock prices has been one of the more prominent applications of the want strategy. Additionally, it is claimed that historical data analysis is used as trade prediction protection.
Chronological data may be used to estimate the cost of stock or supplies, albeit it is really required to use counts to do so. The standard frameworks only care about the type of element selected for forecasting. The latter is typically accomplished with the aid of genetic algorithms (GA) or artificial neural networks (ANNs) [5], but these tools fail to create a connection between their inventory costs as far-off temporary dependencies.
Based on financial index data, RautSushrut et al. [2] proposed using supervised learning classifiers to forecast stock price fluctuations and assess their performance. For portfolio modelling in the financial market, analytical computational methodologies were applied. The usage of the SVM technique was demonstrated in the article and tactical strategies for stock price prediction were also demonstrated. Statistical AI methodology was discussed.
It was discovered by Manoj S. Hegde et al. [3] that Long Short Term Memory (LSTM) networks are a subset of Recurrent Neural Networks (RNNs) that can solve linear problems. Recurrent Neural Networks (RNNs) may also be used to forecast stock values.
According to M. Roondiwala et al. [4], the most common RNN design is long short-term memory. The LSTM inserts a memory cell—a processing unit—in the hidden network layer, taking the role of traditional artificial neurons. Networks are well suited for dynamically collecting data structure over time with a high prediction bound because these memory cells make it possible for networks to link memory and distant inputs effectively in time. The study demonstrates that it is possible to anticipate stock prices for the NIFTY50 stocks.
One of the most crucial tasks is gathering data. After training our model, it is critical to evaluate the algorithm by using different data sets. The following parts explain our process.
According to Kim & H. Y. Kim et al. [5].The marvel of the detonating volatile slope, which occurs when loads of an enormously huge system either become too vast, too inconsequential (or become too little), or both, is a serious issue with basic ANNs for stock forecasting and severely weakens their link to the ideal value. This is frequently caused by two **factors: the loads** self-initiate, and the loads at the end of the system also tend to alter significantly more than the loads at the beginning.
The usage of LSTM networks may be used to forecast stock prices, the research further suggests.
As discussed by S. Selvin et al. [6], the typical approaches to financial market investigation and stock value prediction include a thorough examination of previous stock exhibitions and the organization's overall credibility, as well as a measurable investigation that deals exclusively with the calculation and detection of stock value projections. It is also mentioned in the paper that various types of analysis can be carried out to predict the stock value.
The volatility character of the stock market is an area that requires several analyses based on historical data, according to Loke. K.S et al [7]. Traditional stock price technical prediction methods are based on statistical data analysis, whereas traditional stock trend

prediction algorithms employ historical time series data. The author also discusses developments and improvements in stock price prediction utilising AI and machine learning techniques in the study.

There is no universal answer that can be used even though several studies have been done to discover an appropriate model to anticipate stock prices.

Stock markets are important to how the economy works in contemporary culture, according to Xi Zhang1 et al. in their study [8]. The report also suggests a strategy for using various sources of data to anticipate stock prices and notes that analysis may be done on data that comes from a reliable source.

Long Short Term Memory (LSTM), a recurrent (RNN) approach, is used in Tao Xing and Yuan Sun et al.'s [9] model to take into account the previous stock prices of a firm.

The suggested method applies the forecast to a certain attribute while taking into account the stock's historical data that is currently accessible. The starting price, daily high, daily low, prior day price, closing price, and trading date are the stock's features. The suggested approach employs time series analysis to forecast a stock price for a specific time period.

To apply sentiment analysis to stock prediction, Jordan Prosky et al. [10] suggested CNN techniques and their usage in stock price prediction. It is a more open-ended variation of the gated recurrent system, as stated by X. Shao and D. Ma [11]. LSTMs eliminate the evanescent gradient problem that is inherent in RNNs, making them less dangerous than other deep learning techniques like RNNs or conventional feed-forward neural networks. It also describes how to combine LSTM with K-means algorithms for a short-term stock prediction system.

PROPOSED SYSTEM

As stated in the preceding section, the first step is to get historical market data. The next stage is to extract the feature that is needed for data analysis, then divide the data into training and test sets, train the algorithm to forecast prices, and finally visualise the results. The suggested system's architecture is shown in Fig. 1.

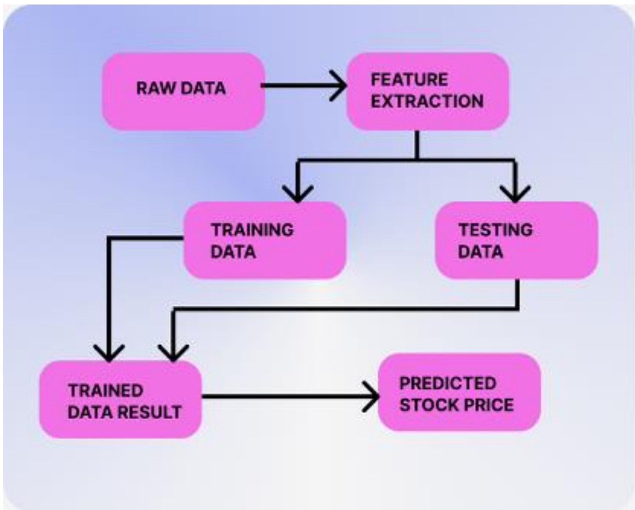

Figure 1: System Architecture

A cell, an info door, an entry door, and a door with a view make up a standard LSTM unit. The three inputs control the flow of data into and out of the cell, which gathers values across arbitrary periods. The LSTM's key benefit is its capacity to learn temporal relationships that are particular to a given situation. Without explicitly applying the activation function inside the recurrent components, each LSTM unit gathers data for either a lengthy or brief duration (thus the name). A crucial fact is that the output of the ignored input door, which fluctuates between 0 and 1, uniquely elevates each cell state. In other words, the LSTM cell's unnoticed door is in charge of both the loads and the capability to start the cell state. As a result, loads can achieve their optimal quality in a reasonable length of time, and data from a previous cell state can flow through a cell unmodified as opposed to growing or shrinking exponentially with each time step or shift. As a result, the evaporating slope problem may be handled by LSTMs since the value stored in the memory cell is not modified repeatedly. When produced using reverse engineering, taking marketplaces like NSE and BSE as Indian trading units for our analyses, the slope does not vanish.

PARAMETERS USED

Table 1 contains a list of the variables and symbols used in this article.

Table 1:Parameters Used

Parameter Used	Meaning
Date	Date of stock price
Open	Open price of a share
Close	Closing price of a share
Volume/ trade quantity	Number of shares traded
High	Highest share value for the day
Low	Lowest share value for the day
Turnover	Total Turnover of the share

A STOCK PRICE PREDICTOR USING LTSM

The suggested framework uses long short-term memory (LSTM) to determine the stock's closing price in order to learn online. In contrast to conventional feed-forward neural networks, long short-term memory (LSTM) is a design for an intermittent neural system (RNN) [1] used in deep learning. The approach concentrates not just on discrete information (such as photos), but also on comprehensive collections of information (such as a speech or a video). For tasks like unpartitioned associated writing recognition, speech recognition, and the identification of irregularities in organised traffic or IDS (Interruption Location Frameworks), for instance, LSTM is helpful.

Algorithm 1: Stock prediction using LSTM

Input: Historic stock data

Output: prediction of the stock price using price variation

Step 1: Start.
Step 2: Data preprocessing after obtaining historical data from the market for a given stock.
Step 3: Import the data set into the data structure and read the open price.
Step 4: Perform feature scaling on the data so that the data values are between 0 and 1.
Step 5: Create a data structure with 60 timestamps and 1 output.
Step 6: Build the RNN (recurrent neural network)
Step 5: data set and initialize the RNN using the sequential repressor.
Step 7: Adding the first LSTM layer and dropout regularization to remove unwanted values.
Step 8: Add the output layer.
Step 9: Compile the RNN by adding the Adam optimization and loss as mean_squared_error.
Step 10: Creating the predictions and visualizing the results using graphing techniques.

A crucial step that must be taken before processing the data is gathering market data.

The most crucial phase in our suggested system is information gathering, which is done by importing data from well-known clearing organisations like BSE (Bombay Stock Exchange) and NSE (National Stock Exchange). The data collection that will be utilised to determine market expectations has to be divided depending on many factors. By including more outside data, the information selection also aids in the dataset's expansion. The majority of our data comes from stock prices from the preceding year. NSEpy is one of the Python packages that may be used to obtain data from the NSE.

RESULT AND DISCUSSIONS

The suggested LSTM model's implementation in Python, uses past data to forecast the price of Maruti Suzuki Motars's stock in the future. The visualisation of the Maruti Suzuki Share prediction is shown in the image below. The creation of an algorithm that forecasts the share price of a company for a specific time period is the subject of our study. The forecasted price of Maruti Suzuki Motors is shown on the graph below, which is the result of our method. Plotting the output of our method using 96 LSTM units for accuracy yields the result displayed in the graph below.

The results of comparing the trained model using the technique described in the preceding section with the original dataset are displayed in Fig. 2, which is also derived from the original dataset. Share price makes up the "x" axis. Days are on the "y" axis. The data is displayed in Fig. 3 over 1500 days.

The next phase is preprocessing data, which entails converting raw data into a fundamental structure. Preprocessing data is a crucial step in information mining. The information gleaned from the source is incoherent, fractured, and tainted with mistakes. The information is cleaned during the preprocessing stage, and highlight scaling is then necessary to focus on a smaller set of parameters. Crossapproval, a highly well-founded, anticipated execution of the model utilising the preparation information, is incorporated into the model's preparation.

The goal of the tuning models is to explicitly modify the calculation training and enhance the calculation itself with additional information. The test sets are flawless since a model shouldn't be selected based on secret knowledge. Increase the information to reflect the expenses of the real offer. The last step is to visualise the data using an approach that helps demonstrate how the data might vary depending on how our algorithm performs.

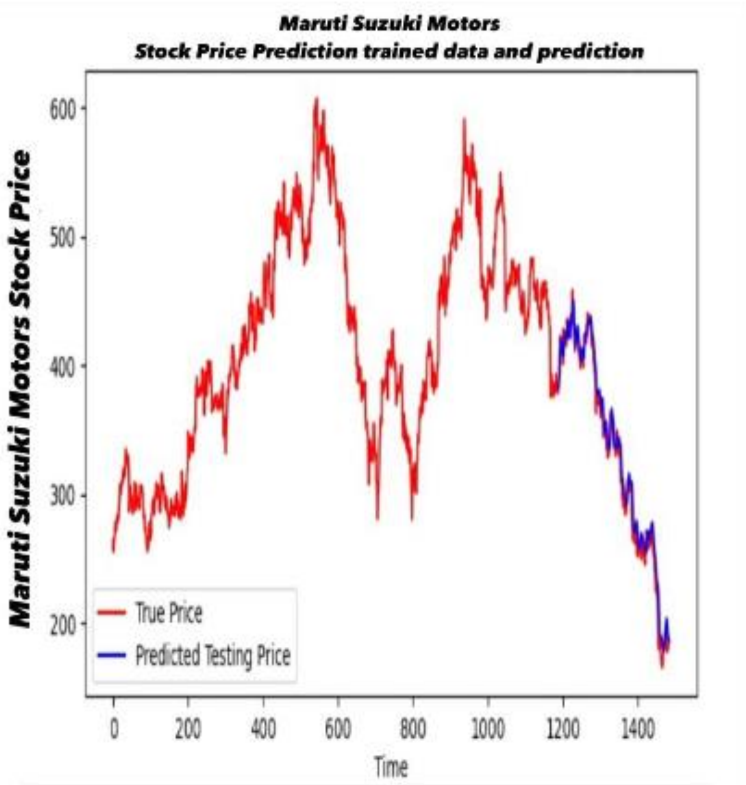

Figure 2: Predicted testing stock price

The data in Fig. 3 is taken from the original dataset, and the outcome is demonstrated by contrasting its accuracy with that of the trained model produced by the method described in the preceding section. The share price is on the "x" axis. Days are shown on the "y" axis. Figure 3 displays the data for 300 days.

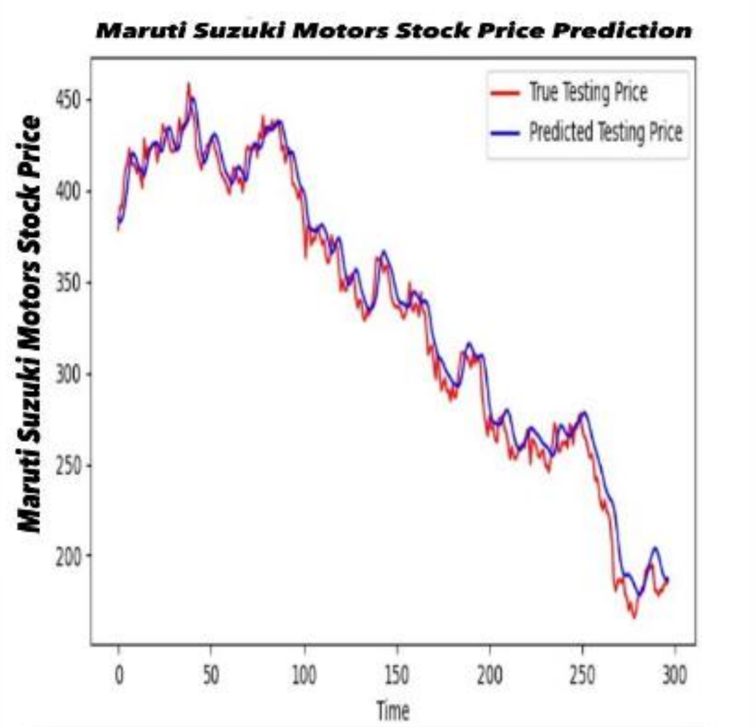

Figure 3: Predicted testing stock price

The whole data set and a portion of the trained data have both been plotted on the graph in Figure 2. The graph displays the starting price of the Maruti Suzuki Motors share for the 1484th day with a very small loss. The algorithm successfully generated a graph with the predicted price testing price (blue) and true price (red). There is a small difference in the predicted price between the predicted price testing price (blue) and true price (red), demonstrating that our algorithm is capable of predicting the lowest loss rate for the provided complete data set of a particular share.

The graph in Fig. 3 displays the starting price of the Maruti Suzuki Motors share for the 300th day with very little loss. The algorithm successfully created a graph with the predicted price for testing (blue) and the actual price for testing (red). There is a small difference in the predicted price between the predicted price for testing (blue) and the actual price for testing (red), demonstrating that our algorithm is capable of predicting with a minimum loss rate of 0.0024.

The suggested approach can predict the share price with a very low loss and error rate; however, the training will be more effective if the epoch batch size is increased. In the section above, we utilised an epoch batch size of 50 to predict the stock prices.

The figures (figs. 2 and 3) from the preceding part of the suggested algorithm can estimate the price with a loss: The open price for the 0.0024th 300th day was 172 Indian rupees, while our forecast price for the share is 166 rupees.

CONCLUSION

In this work, the study on the share is done, and it may be done in the future for a lot more shares.

If the model trains new data sets using more layers, more processing power, and LSTM modules, the prediction could be more accurate.

The market's perception of a specific stock's price change may be ascertained through sentiment research on social media. Facebook is a well-known social network with a multitude of user-posted market trend data, therefore this might be achieved by integrating Twitter and Facebook API into our application.

REFERENCES

[1] Tao Xing, Yuan Sun, Qian Wang, Guo Yu. "The Analysis and Prediction of Stock Price", 2013 IEEE International Conference on Granular Computing (GrC), ISBN: 978-1-4799-1282-7, DOI: 10.1109/GrC.2013.6740438.

[2] Jordan Prosky, Andrew Tan, Xingyou Song, Micael Zhao. "Sentiment Predictability for Stocks" - arXiv:1712.05785v2 [cs.CL], 18 Jan 2018.

[3] . Xi Zhang1, Siyu Qu1, Jieyun Huang1, Binxing Fang1, Philip Yu2, “Stock Market Prediction via Multi-Source Multiple Instance Learning.” IEEE 2018.

[4] X. Shao, D. Ma, Y. Liu, Q. Yin. "Short-term forecast of stock price of multi-branch LSTM based on K-means", 2017 4th International Conference on Systems and Informatics (ICSAI).

[5] Hiba Sadia, Aditya Sharma, Adarrsh Paul, SarmisthaPadhi, Saurav Sanyal- “Stock Market Prediction Using Machine Learning Algorithms”, IJEAT, 2019.

[6] Loke.K.S. “Impact Of Financial Ratios And Technical Analysis On Stock Price Prediction Using Random Forests”, IEEE, 2017.

[7] Raut Sushrut Deepak, Shinde Isha Uday, Dr. D. Malathi, “Machine Learning Approach In Stock Market Prediction”, IJPAM 2017