International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 06 | June - 2022 Impact Factor: 7.185 ISSN: 2582-3930

STOCK PRICE PREDICTION USING LSTM

Nidhi Agnihotri
Department of Computer Science and Engineering,
Maharshi Dayanand University,

Dr. Renuka Arora
Associate Professor,
Department of Computer Science and Engineering,
Maharshi Dayanand University

Abstract

In the financial sector, stock price prediction is the most widely used. Because the stock market is inherently volatile, it is impossible to forecast stock prices. This is a time series problem. Stock price prediction is a challenging task where there are no rules to predict the price of the stock in the stock market. There are numerous strategies for forecasting stock prices. Logistic Regression Model, SVM, ARCH model, RNN, CNN, Backpropagation, Nave Bayes, ARIMA model, and others are some of the prediction approaches. Long Short-Term Memory (LSTM) is the best method for time series issues in these models. The major goal is to forecast current market trends and be able to accurately predict stock values. To accurately anticipate stock prices, we deploy LSTM recurrent neural networks. The accuracy of the predictions is over 93 percent, according to the findings.

Keywords: LSTM, CNN, ML, DL, Trade Open, Trade Close, Trade Low, Trade High.

1. INTRODUCTION

Stock market forecasting entails predicting a company's present developments as well as the value of its stocks, whether they are rising or falling. The stock market is a marketplace where investors may buy and sell firm stock. A stock is a type of investment in a company that reflects ownership. You may purchase and sell equities on the stock market. Buying a firm's stock entitles you to a minor stake in the company.

We are predicting the stock prices using the machine learning algorithm to develop a model which forecasts the stock price effectively based on the current market trends. We have used LSTM recurrent neural networks to successfully forecast stock values. There were two sorts of stocks available, one of which was intraday trading, sometimes known as day trading. Intraday trading implies that all positions are squared off before the market closes for the day, and there is no way to change ownership after the market closes. LSTMs are crucial because they can retain prior or past information, making them extremely strong in sequence prediction issues. This is critical in stock prediction since we need to retain and interpret prior stock data as well as properly estimate future stock values.

2. REVIEW OF LITERATURE

In machine learning areas, stock price prediction may be predicted using AI and machine learning models. Stock price prediction using the SVM model. SVM is a classification method that works with machine learning techniques. It is used to get a new text as an output. Applying Multiple Linear Regression with Interactions to predict the trend in stock prices (Osman Hegazy et al. 2013 [20]; V Kranthi Sai Reddy, 2018 [8]; a Banerjee et al. 2020 [21]; Lufuno Ronald Marwala [13]). Random Walk Hypothesis, which is proposed by Horne, j. C et al 1997 [27] which is used to predict stock prices, Horne j.c [27] said that the stock values are changes random and the past price values are not dependent on current values. The EMH differs from the Random Walk hypothesis in that it predicts stock prices using short-term trends.

Manh Ha Duong Boris's Siliverstovs, 2006 [11] search the abstraction between equity prices and combined finances in Key Eu nations like UK and Germany. Acceleration in European country

investments is likely to result in a stronger link between European nation equity prices. This operation may potentially lead to a merging in financial growth across EU states, if developments in stock markets impact genuine financial products, such as investing and Consuming. Fahad Almudhaf et al., 2012 [22], examine CIVETS' weak-form market efficiency from 2002 to 2012. CIVETS employs the random walk hypothesis process. When it comes to the future price, the equity values in an efficient stock market must follow the random walk hypothesis, which means that the values change randomly and unexpected. Everyday returns for rising and improved markets have been tested for random walks.

To encode data, the LSTM technique uses a Recurrent Neural Network. Economic news headlines from Bloomberg and Reuters are fed into the algorithm. In the stock market, the LSTM with embedded layer and the LSTM with automated encoder are used to forecast stock prices. Xiongwen Pang and colleagues [4]. LSTM layers were used to vectorize the data utilizing an automated encoder and embedded layer. Stock correlation coefficients are chosen at random and forecasted using ARIMA and neural networks. The RNN and LSTM algorithms are used in this application. M. [17] Nabipour et al. Random Forest, decision tree, and neural networks were among the machine learning and deep learning techniques used to forecast stock prices. LSTM gives the best results while predicting stock prices with the least error rate (Hyeong Kya Choi,2018 [16]; Huicheng Liu, 2018 [15]; M. Nabipour et al,2020 [17]; Xiongwen Pang et al, 2020 [4]).

Pranab Bhat, 2020, employed convolution neural networks to forecast market prices. In this model, learning is completed by computing the mean square blunder for each subsequent observation, and the model with the least blunder and the highest predictive power is chosen. Mohammad Mekayel Anik et al, 2020 [23], implemented a linear regression algorithm for future stock price prediction. They fulfilled their objectives in that the model's prediction accuracy is excellent, and it may be utilized to forecast stock prices. Xiao Ding et al. 2020 [14] used an easy and effective interface to add common sense knowledge to the process while learning of events.

The LMS filter is used for solving linear problems. The idea of the filter is to find the filter coefficients and to minimize a system by reducing the least mean square of the error value (Asep Juarna, 2017 [24]; Eleftherios Giovanis, 2018 [25]) They constructed a model utilizing deep regression based on CNN and utilized a hybrid model to forecast stock prices using deep learning and ML techniques. Here they used CNN for parameters, thereby increase the no of loops will stabilize the validation loss. They also tested using DL and a hybrid ML algorithm for stock price prediction. Vivek Rajput and Sarika Bobde

[26] used sentiment analysis from online posts or multimedia and data mining is used. They aim to derive emotion, either good or negative, based on textual information accessible on social networks in sentiment analysis. Sentiment analysis for predicting the stock market to get more accurate and efficient results.

3. DATA COLLECTION

For the experimental study, we downloaded live datasets namely google, nifty, reliance, etc. from the Yahoo Finance website (https://finance.yahoo.com/).

Table 3.1 Google

Attribute Name	Min	Max
Open	87.74	1005.49
Low	86.37	996.62
High	89.29	1008.61
Close	87.58	1004.28

Table 3.2 Nifty50

Attribute Name	Min	Max
Open	7735.15	12932.5
Low	7511.1	12819.35
High	8036.95	12948.85
Close	7610.25	12938.25

Table 3.3 Reliance

Attribute Name	Min	Max
Open	205.5	3298
Low	197.15	3141.3
High	219.5	3298
Close	203.2	3220.85

Sample Input

Table 3.1

Date	Trade Open	TradeLow	Trade High	Trade Close
11-Jun-2021	2,524.92	2,498.29	2,526.99	2,513.93
10-Jun-2021	2,494.01	2,494.00	2,523.26	2,521.60
09-Jun-2021	2,499.50	2,487.33	2,505.00	2,491.40
08-Jun-2021	2,479.90	2,468.24	2,494.50	2,482.85
07-Jun-2021	2,451.32	2,441.07	2,468.00	2,466.09
04-Jun-2021	2,422.52	2,417.77	2,453.86	2,451.76
03-Jun-2021	2,395.02	2,382.83	2,409.75	2,404.61
02-Jun-2021	2,435.31	2,404.20	2,442.00	2,421.28

4. METHODOLOGIES

4.1 LSTM Algorithm

The RNN technique is used by LSTM, which has the capacity to memorize. There are three gates in each LSTM cell: input, forget, and output. When data enters the LSTM network, the useful data is maintained, while the unneeded data is discarded via the forget gate. Weather forecasting, NLP, speech recognition, handwriting recognition, time-series prediction, and other applications can all benefit from LSTM.

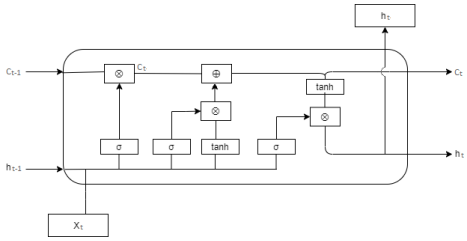

Fig 4.1.1: LSTM Architecture

As shown in Fig. 4.1.1, the inputs to the current cell state (Ct) are the previous hidden state(ht-1), previous cell state (Ct-1) and present input (Xt). The cell consists of three gates i.e., forget gate, input gate and output gate.

Forget Gate:

A forget gate is a device that removes superfluous data from the state of a cell.

- The information that is less important or not required for the LSTM to understand things is removed by performing multiplication of hidden state by a sigmoid function.
- This step is necessary to optimize the performance of the model.
- It has two inputs, h(t-1) and xt, with h(t-1) representing the previous cell hidden state output and xt representing the current cell input.

$$Ft = \sigma (Wfx * Xt + Wfh * ht-1 + bf)$$

Input Gate:

- This cell is responsible for regulating the data that is added to the cell from the input. Forget gate is used to filter some input.
- Using the tanh function, a vector is created by adding all possible values from the previous cell hidden state h(t-1) and the current cell input Xt. In the ranges [-1, 1], the output of the tanh function.
- Finally, the sigmoid and tanh functions' outputs are multiplied, and the result added to the cell state.

$$It = \sigma (Wix * Xt + Whh * ht\text{-}1 + bi) + \tanh (Wcx * Xt + Wch * ht{-}1 + bi)$$

Output Gate:

- Tanh function is applied to the cell state to create a vector with all possible values.
- Sigmoid function is applied to previous cell hidden state h(t-1) and current cell input xt to filter necessary data from the previous cell.
- The outputs of the sigmoid and tanh functions are now multiplied, and the result is passed to the next cell as a hidden state.

$$Ot = \sigma (Wox * Xt + Whh * ht\text{-}1 + Woc * Ct\text{-}1 + bi)$$

Intermediate cell state (Ct) is obtained by the multiplication of Forget gate (Ft) with previous cell state (Ct-1). Then this intermediate state is added to the output of the input gate.

$$Ct = Ft * Ct\text{-}1 + It$$

Current hidden/output state is obtained by multiplying output gate and tanh of cell state.

$$ht = Ot * \tanh (Ct)$$

4.2 SYSTEM ARCHITECTURE

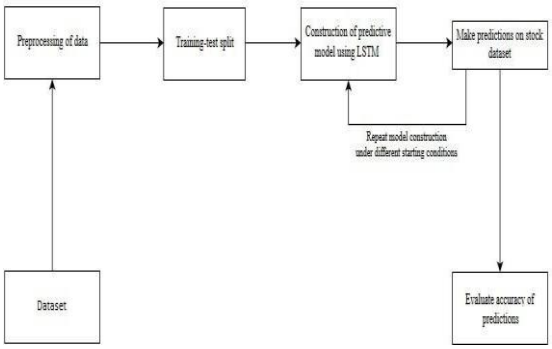

Fig 4.2.1: Overall Architecture

Data Selection: The initial step is to choose data for a company and divide it into training and testing categories. We have used 75% for training and 25% for testing purposes.

Pre-processing of data: In pre-processing, we are selecting attributes required for the algorithm and the remaining attributes are neglected. Trade Open, Trade High, Trade Low, Trade Close, and Trade Volume are the qualities that have been chosen. We use normalisation in pre-processing to acquire numbers in a specific range.

Prediction using LSTM: In this system, The LSTM algorithm is being used to forecast stock prices. Initially, the training data is passed through the system and train the model. Then in the testing phase, the predicted values are compared with the actual values.

Evaluation: In the evaluation phase we are calculating the Accuracy, Mean Square Error (MSE) and Root Mean Square Error (RMSE) values for comparison.

5. EXPERIMENTAL RESULTS

5.1 Google

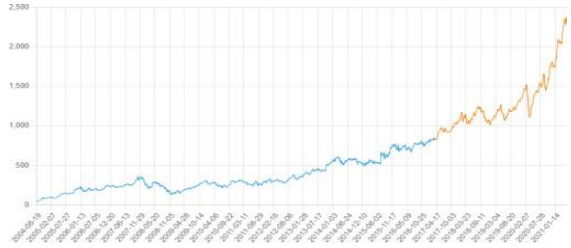

Fig 5.1.1 Google Graph

epochs	Accuracy	MSE	RMSE
10	93.00717	207.6578	14.41034
20	94.01166	156.3873	12.50549
30	95.64188	105.3248	10.26279
40	95.59026	99.17409	9.958619
50	96.99466	62.24641	7.88964

Table 5.1.2 Google Epochs

In the results, as we have shown in Fig 5.1.1, the graph shows Trade Close value for the google dataset. The blue line in this graph represents the training data, while the yellow hue represents the projected values from the test data.

Table 5.1.2 shows the accuracy, MSE and RMSE values for no of iterations (epochs).

5.2 Reliance

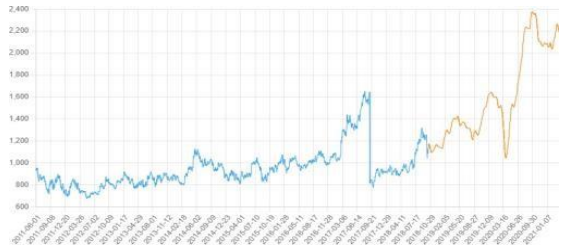

Fig 5.2.1 Reliance Graph

epochs	Accuracy	MSE	RMSE
10	96.25328	4839.5690	69.56701
20	97.63884	2653.1278	51.50852
30	98.19937	1650.3337	40.62430
40	98.13571	1616.9295	40.21106
50	98.37254	1361.8098	36.90270

Table 5.2.2 Reliance Epochs

Above graph 5.2.1 shows Trade Close value for the Reliance dataset and table 5.2.2 shows the MSE, RMSE and accuracy values for the Reliance dataset.

6. CONCLUSION

We use the LSTM algorithm to estimate the closing stock price of any particular company. We used statistics from Google, the Nifty50, TCS, Infosys, and Reliance Stocks, and we were able to attain above 93 % accuracy. In the future, we can extend this application for predicting cryptocurrency trading and, we can add sentiment analysis for better predictions.

7. REFERENCES

[1] Stock Price Prediction Using LSTM on Indian Share Market by Achyut Ghosh, Soumik Bose1, Giridhar Maji, Narayan C. Debnath, Soumya Sen

[2] S. Selvin, R. Vinayakumar, E. A. Gopalkrishnan, V. K. Menon and K. P. Soman - Stock price prediction using LSTM, RNN and CNN-sliding window model - 2017.

[3] Murtaza Roondiwala, Harshal Patel, Shraddha Varma, "Predicting Stock Prices Using LSTM" in Undergraduate Engineering Students, Department of Information Technology, Mumbai University, 2015.

[4] Xiongwen Pang, Yanqiang Zhou, Pan Wang, Weiwei Lin, "An innovative neural network approach for stock market prediction", 2018

[5] Ishita Parmar, Navanshu Agarwal, Sheirsh Saxena, Ridam Arora, Shikhin Gupta, Himanshu Dhiman, Lokesh Chouhan Department of Computer Science and Engineering National Institute of Technology, Hamirpur – 177005, INDIA - Stock Market Prediction Using Machine Learning.

[6] Pranav Bhat Electronics and Telecommunication Department, Maharashtra Institute of Technology, Pune. Savitribai Phule Pune University - A Machine Learning Model for Stock Market Prediction.

[7] Anurag Sinha Department of computer science, Student, Amity University Jharkhand Ranchi, Jharkhand (India), 834001 - Stock Market Prediction Using Machine Learning.

[8] V Kranthi Sai Reddy Student, ECM, Sreenidhi Institute of Science and Technology, Hyderabad, India - Stock Market Prediction Using Machine Learning.

[9] Asset Durmagambetov currently works at the mathematics, CNTFI- 'The Riemann Hypothesis-Millennium Prize Problems' - stock market predictions.

[10] Mariam Moukalled Wassim El-Hajj Mohamad Jaber Computer Science Department American University of Beirut - Automated Stock Price Prediction Using Machine Learning.

[11] Manh Ha Duong Boriss Siliverstovs June 2006 - The Stock Market and Investment.

[12] Dharmaraja Selvamuthu, Vineet Kumar and Abhishek Mishra Department of Mathematics, Indian Institute of Technology Delhi, Hauz Khas, New Delhi 110016, India - Indian stock market prediction using artificial neural networks on tick data.

[13] Lufuno Ronald Marwala A dissertation submitted to the Faculty of Engineering and the Built Environment, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Master of Science in Engineering - Forecasting the Stock Market Index Using Artificial Intelligence Techniques.

[14] Xiao Ding, Kuo Liao, Ting Liu, Zhongyang Li, Junwen Duan Research Centre for Social Computing and Information Retrieval Harbin Institute of Technology, China - Event Representation Learning Enhanced with External Common-sense Knowledge.

[15] Huicheng Liu Department of Electrical and Computer Engineering Queen's University, Canada - Leveraging Financial News for Stock Trend Prediction with Attention-Based Recurrent Neural Network.

[16] Hyeong Kyu Choi, B.A Student Dept. of Business Administration Korea University Seoul, Korea = Stock Price Correlation Coefficient Prediction with ARIMA-LSTM Hybrid Model.

[17] M. Nabipour Faculty of Mechanical Engineering, Tarbiat Modares University, 14115-143 Tehran, Iran; Mojtaba.nabipour@modares.ac.ir - Deep Learning for Stock Market prediction.

[18] Lavanya Ra SRM Institute of Science and Technology | SRM · Department of Computer Science - Stock Market Prediction.

[19] M. Mekayel Anik · M. Shamsul Arefin (B) Department of Computer Science and Engineering, Chittagong University of

Engineering and Technology, Chittagong, Bangladesh - An Intelligent Technique for Stock Market Prediction.

[20] Osman Hegazy, Omar S. Soliman and Mustafa Abdul Salam Faculty of Computers and Informatics, Cairo University, Egypt - A Machine Learning Model for Stock Market Prediction

[21] Sharanya Banerjee, Neha Dabeeru, R. Lavanya SRM Institute of Science and Technology, Chennai in Computer Science and Engineering - Stock Market Prediction

[22] Fahad Almudhaf, Yaser A. Alkulaib from Kuwait University - Are Civets Stock Markets Predictable?

[23] Mohammad Mekayel Anik, Mohammad Shamsul Arefin and M. Ali Akber Dewan, Department of Computer Science and Engineering - An Intelligent Technique for Stock Market Prediction

[24] Asep Juarna, Departemen of Informatics, Gunadarma University, Jakarta-Indonesia - ONE YEAR STOCK PRICE PREDICTION AND ITS VALIDITY USING LEAST SQUARE METHOD IN MATLAB

[25] Giovanis, Eleftherios - Applications of Least Mean Square (LMS) Algorithm Regression in Time-Series Analysis

[26] Vivek Rajput and Sarika Bobde - Stock market prediction using hybrid approach

[27] James C. Van Horne and George G. C. Parker – The Random-walk Theory: An Empirical Test