STOCK PRICE PREDICTION USING LSTM

S. Dinesh[1], A.M.S. Rama Raju[1], S. Rahul[1], O. Naga Sandeep[1],
Mr. N D S S Kiran Relangi[2]

[1]*Final year students of Department of CSE, Anil Neerukonda Institute of Technology and Sciences (A), Visakhapatnam-531162, India*
[2]*Assistant Professor at Department of CSE, Anil Neerukonda Institute of Technology and Sciences (A), Visakhapatnam-531162, India*

Abstract

Stock price prediction is the most significantly used in the financial sector. Stock market is volatile in nature, so it is difficult to predict stock prices. This is a time series problem. Stock price prediction is a difficult task where there are no rules to predict the price of the stock in the stock market. There are so many existing methods for predicting stock prices. The prediction methods are Logistic Regression Model, SVM, ARCH model, RNN, CNN, Backpropagation, Naïve Bayes, ARIMA model, etc. In these models, Long Short-Term Memory (LSTM) is the most suitable algorithm for time series problems. The main objective is to forecast the current market trends and could predict the stock prices accurately. We use LSTM recurrent neural networks to predict the stock prices accurately. The results show that prediction accuracy is over 93%.

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

1. INTRODUCTION

Stock market prediction means forecasting the current trends of a company and predict the value of stocks whether it's going up or down. Stock market is the place where a company's shares are traded. A stock is an investment in an institution where it represents ownership in a company. Stock market is a place where those stocks are purchased. Purchasing a stock of a company is owning a small share of an institution.

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 predict the stock prices accurately. You would find two types of stocks, one of them was Intraday trading, which is known to us by the term day trading. Intraday trading is that which means all positions are squared-off before the market closes then and there and there would be no possibility of changing the ownership after the day end. LSTM's are very important, as they are very powerful in sequence prediction problems because they could store previous or past information. This is very important in stock prediction as we need to store and read the previous stock information as well to forecast the stock prices accurately in the future.

The rest of the paper is organized as follows. Section 2 introduces the research status of stock price prediction. Section 3 introduces the methodologies. Section 4 consists of the experimental results and the analysis of the results. Section 5 concludes the paper.

2. LITERATURE SURVEY

Stock price prediction can be predicted using AI and machine learning models in machine learning fields. Using the SVM model for stock price prediction. SVM is one of the machine learning algorithms which works on classification algorithms. 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. EMH is different from the Random walk hypothesis but the EMH works mainly on Short term patterns for predicting stock prices.

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 Eu nations investments is apt to results successful even Stronger correlation between the different Eu nations and equity prices. This operation may also lead to a merge in financial development between EU nations, if advancements in stock markets affect real financial instruments, such as investing and Consuming. Fahad Almudhaf et al, 2012 [22], tests the weak-form market efficiency of CIVETS over the period 2002–2012. The random walk hypothesis process is used in CIVETS. In an efficient stock market, the equity values must follow a random walk hypothesis, when it comes to the future price, the values are changing randomly and unpredictable. Everyday returns for rising and improved markets have been tested for random walks.

LSTM algorithm consists of a Recurrent Neural network to encode data. The algorithm inputs are economic news headings infusion From Bloomberg and Reuters. Long Short-term Memory with embedded layer and the LSTM with the automatic encoder in the stock market for predicting stock values. The Xiongwen Pang et al [4]. Used an automatic encoder and embedded layer to vectorizing the values by using LSTM layers. Correlation coefficients in stocks are selected randomly and predicted using ARIMA and the neural network approach. In this RNN and LSTM algorithms are implemented. M. Nabipour et al [17]. Used different machine learning and deep learning algorithms for predicting stock values such as random forest, decision tree and neural networks. LSTM gives the most accurate results and it has the best ability to fit. 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]).

Recently, Pranab Bhat, 2020 used convolution neural networks for predicting stock values, in this model learning is finished by computing the mean square blunder for each consequent perception and a model is picked that has the least mistake and high prescient power. In this paper, they are utilizing CNN for anticipating stocks and incentives for the following day. Mohammad Mekayel Anik et al, 2020 [23], implemented a linear regression algorithm for future stock price prediction. In this they achieved their goals in predicting accuracy of the model is very good and it might be used for predicting stock values. 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 a type of adaptive filter which 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 used a hybrid model for predicting the stock values by using deep learning and ML methodologies and they built a model using deep regression based on CNN. 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. In sentiment analysis, they are trying to get emotion either positive or negative based on the textual information available on social networks. 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	87.74	1005.49
Low	86.37	996.62
High	89.29	1008.61
Close	87.58	1004.28

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 Sample Input

Date	Trade Open	Trade Low	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

LSTM uses the RNN approach which has the ability to memorize. Each LSTM cell has three gates i.e. input, forget and output gates. While the data that enters the LSTM's network, the data that is required is kept and the unnecessary data will be forgotten by the forget gate.

LSTM can be used in many applications such as for weather forecasting, NLP, speech recognition, handwriting recognition, time-series prediction, etc.

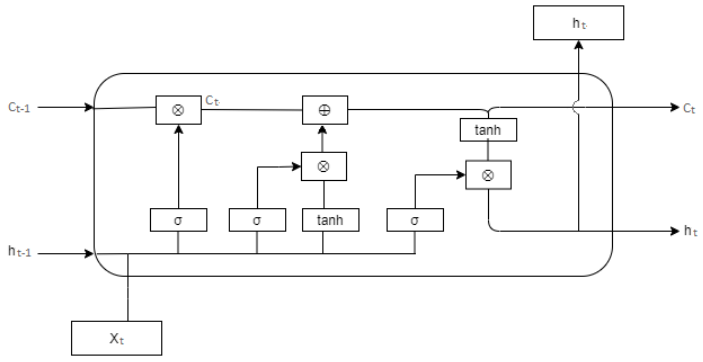

Fig 4.1.1: LSTM Architecture

As shown in Fig. 4.1.1, the inputs to the current cell state (C_t) is the previous hidden state (h_{t-1}), previous cell state (C_{t-1}) and present input (X_t). The cell consists of three gates i.e. forget gate, input gate and output gate.

Forget Gate:

A forget gate will remove unnecessary data from the cell state.

- 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 takes two inputs i.e., $h_{(t-1)}$ and x_t, where $h_{(t-1)}$ is the previous cell hidden state output and x_t is the current cell input.

$$F_t = \sigma\ (W_{fx} * Xt + W_{fh} * h_{t-1} + b_f)$$

Input Gate:

1. 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.
2. A vector is created by adding all the possible values from the previous cell hidden state $h_{(t-1)}$ and current cell input X_t by using the tanh function. The output of the tanh function in the ranges of [-1, 1].
3. Finally, the outputs of sigmoid and tanh functions are multiplied and the output is added to the cell state.

$$I_t = \sigma\ (W_{ix} * X_t + W_{hh} * h_{t-1} + b_i) + \tanh(W_{cx} * X_t + W_{ch} * h_{t-1} + b_i)$$

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 x_t to filter necessary data from the previous cell.
- Now, the outputs of sigmoid and tanh functions are multiplied and this output is sent as a hidden state of the next cell.

$$O_t = \sigma\ (W_{ox} * X_t + W_{hh} * h_{t-1} + W_{oc} * C_{t-1} + b_i)$$

Intermediate cell state (C_t) is obtained by the multiplication of Forget gate (F_t) with previous cell state (C_{t-1}). Then this intermediate state is added to the output of the input gate.

$$C_t = F_t * C_{t-1} + I_t$$

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

$$h_t = O_t * \tanh(C_t)$$

4.2 SYSTEM ARCHITECTURE

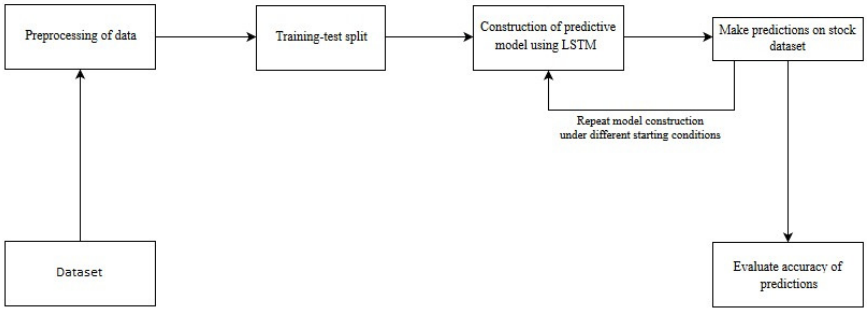

Fig 4.2.1: Overall Architecture

Data Selection: The first step is to select data for an organization and split the data into training and testing. 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. The selected attributes are Trade Open, Trade High, Trade Low, Trade Close, Trade Volume. In pre-processing, we are using normalization to get values in a particular range.

Prediction using LSTM: In this system, we are using the LSTM algorithm for predicting stock values. 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

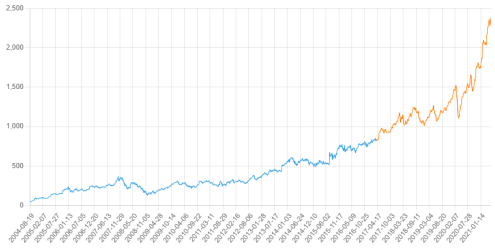

Fig 5.1.1 Google Graph

Table 5.1.2 Google Epochs

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

In the results, as we have shown in Fig 5.1.1, the graph shows Trade Close value for the google dataset. In this graph blue line indicates the training data and the yellow color

shown is the predicted values from the test data. Table 5.1.2 shows the accuracy, MSE and RMSE values for no of iterations (epochs).

5.2 Reliance

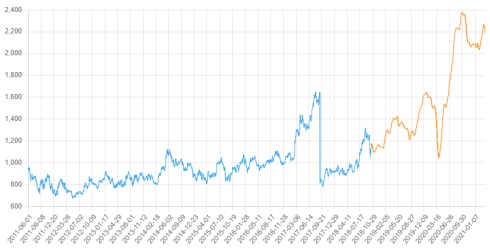

Fig 5.2.1 Reliance Graph

Table 5.2.2 Reliance Epochs

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

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.

CONCLUSION

we are predicting the closing stock price of any given organization, we have developed an application for predicting close stock price using LSTM algorithm. We have used datasets belonging to Google, Nifty50, TCS, Infosys and Reliance Stocks and achieved above 93% accuracy for these datasets. In the future, we can extend this application for predicting cryptocurrency trading and also, we can add sentiment analysis for better predictions.

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