



Prediction of Stock Price Based on LSTM

¹Umer Pinjari, ²Nishant Shinde, ³Prathamesh Jadhav, ⁴Sumer Shinalkar

^{1,2,3,4} Computer Engineering, Dhole Patil College of Engineering, Pune, India

ABSTRACT

Investment in the stock market is one of the most sophisticated and complex forms of commerce. The stock market is inherently unpredictable because stock values are constantly changing due to a variety of reasons, making stock forecasting a challenging and incredibly complex endeavor. Today's investors are intensely engaged in the market and require quick access to precise information. Stock price prediction is a research area with rapidly expanding technology advancements, recognizing the stock's pattern. Predicting a company's future growth and financial development will be very helpful in determining its stock price. This focuses on the use of the Long Short Term Memory (LSTM) recurrent neural network (RNN) based machine learning method. To forecast stock values, use LSTM.

1. INTRODUCTION

Predictions of stock prices are highly valued by the public and many businesspeople. From a work in the stock market, people can make a lot of money or lose their source of income. Future forecasts can be implemented using models and algorithmic predictions to earlier data. Future prediction has proven to be a challenging endeavor that many people have found challenging to comprehend. These kinds of When prediction involves money and dangers, such in stock market speculation, it becomes even more alluring. scientists are doing research on stock market predictions from a range of disciplines, including business and computer science. Researchers have experimented with numerous methodologies, algorithms, and combinations of indications to anticipate the market.

A prediction model's attribute is determined by market performance factors. Short-Term Memory (LSTM) is one of several RNN structures. LSTM transforms traditional artificial neurons in the hidden network layer into the most useful memory cells. With these memory cells, networks can better associate memory with remote input over time, which is why it is important to understand the formation of strong data over time with high predictive power. A great deal of research has been done on daily stock price forecasts, using various data sources with many built-in models such as news articles, Twitter data, Google data, and Wikipedia data.

All of these various factors, when combined with stock prices and stock technology indicators, have influenced stock price movements. In today's society, the question of how to improve the accuracy of stock prices remains unanswered. Time series data is a sequence of data derived from the irregular actions of different fields such as social science, finance, engineering, physics, and economics. Finance, engineering, physics, and economics are all disciplines. Such complexities make forecasting price trends extremely difficult. The main goal of predicting a series of time series is to build future value simulation models based on their previous attributes.

2. LITERATURE SURVEY

Very prime goal of our literature review was to assess various algorithms and models to see if stock price predictions could be made using actual stock prices. However, because we were unable to detect a potential change in this stock price forecast, we decided to review current plans, evaluate major issues, and improve ourselves. A quick search for common solutions to the aforementioned problem led us to LSTM. After deciding to use the LSTM neural network to forecast stock prices, time series data from stock firm prices and related macroeconomic variables are collected over a decade.

3. PROPOSED SYTEM

Figure 1 is depicting the sequence of the process that is followed for each model.



Figure 1 Process

We have proposed that the LSTM (Long Short Term Memory) algorithm be used to provide precise and reliable stock price predictions.

3.1 LSTM - an overview

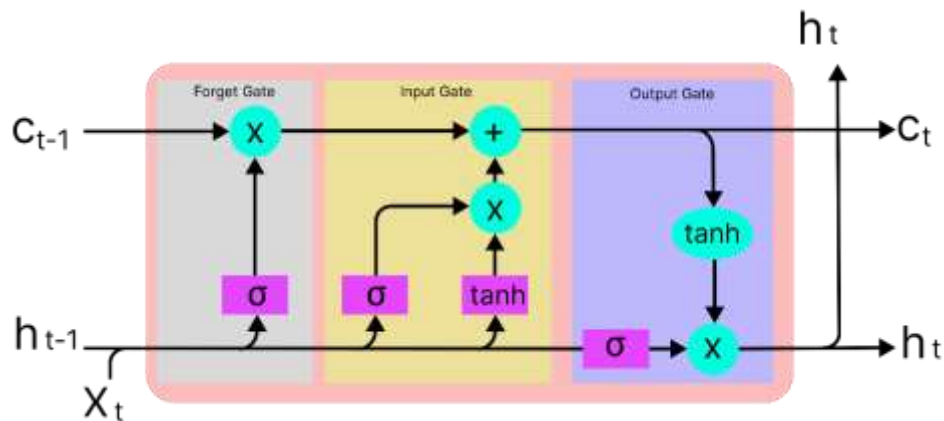


Figure 2 LSTM Memory Cell

A special type of RNN, which can learn long-term dependence, is called Long-Short Term Memory (LSTM). LSTM enables RNN to remember long-term inputs. Contains information in memory, similar to computer memory. It is able to read, write and delete information in its memory. This memory can be seen as a closed cell, with a closed description, the cell decides to store or delete information. In LSTM, there are three gates: input, forget and exit gate. These gates determine whether new input (input gate) should be allowed, data deleted because it is not important (forget gate), or allow it to affect output at current timeline (output gate)

1. **Forget gate:** The forget gateway controls when certain parts of the cell are updated with newer information. It deducts nearly one in parts of the cell state to be retained and zero in values to be neglected.
2. **Input gate :** This network category reads the conditions under which any information should be stored (or updated) in the state cell based on the input (e.g., previous output $o(t-1)$, input $x(t)$, and previous state of cell $c(t-1)$).
3. **Output gate:** This component determines which information is forwarded in the next location in the network based on the input mode and cell.

3.2 ADVANTAGE OF LSTM

The power of LSTM to read intermediate context is its main advantage. Without explicitly leveraging the activation function within the recurring components, each unit remembers details for a long or short period of time. An important fact is that any cell state can be repeated only by releasing the forget gate, which has a value between 0 and 1. That is, the forgetting gateway in the LSTM cell is in charge of both the hardware and the function of cell state activation. Thus, instead of explicitly increasing or decreasing in each step or layer, the data from the previous cell can pass through the unaffected cell, and the components can convert to their suitable value over a specified duration.

Due to the fact that the quantity stored in the memory cell is not transformed in a repeating fashion, the gradient does not disappear when the LSTM is trained to disperse back.

3.3 RECURRENT NEURAL NETWORK

Recurrent Neural Networks (RNN) are a type of neural network that is specifically designed to handle sequential data. There are two kinds of RNNs: discrete-time RNNs and continuous-time RNNs. RNNs that run indefinitely They are built with a cyclic design. Connection architecture, allowing them to update their given the previous states and current input data. RNNs are typically artificial neural networks that consist of standard recurrent cells. These neural structures Networks are well-known for their accuracy in problem solving. It specialises in handling a series of values. $\chi_1 \dots \chi_n$ where n denotes the total number of features, and χ are the features, such as time-series data.

Scaling images with large width and height, as well as processing images of varying sizes, are both possible to a large extent. Furthermore, most RNNs are capable of processing variable-length sequences. However, as demonstrated by a study contacted by Yashoua et al, RNNs lack the ability to learn long-term dependencies. In order to deal with these long-term dependencies, Hoch Reiter and Schmid Huber proposed Long Short-Term Memory in 1997. (LSTM)

4. SYSTEM ARCHITECTURE

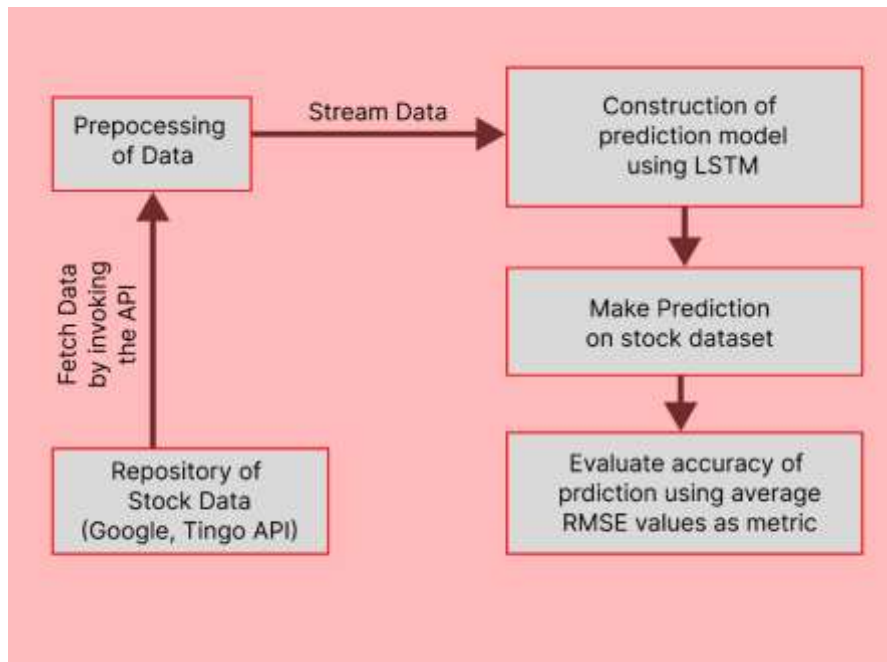


Figure 3 System Model

4.1 Collecting and pre-processing the dataset

The collected data contained five features:

1. Date: Date of stock price.
2. Opening price: When trading begins each day this is opening price of stock.
3. High: The highest price at which the stock was traded during a period(day).
4. Low: The Lowest price at which the stock was traded during a period(day).
5. Volume: How much of a given financial asset has traded in a period of time.
6. Close Interest: The last price at which a particular stock traded for the trading session.

Important sites including Tiingo API, Yahoo, and Google Finance provide stock market information. These websites provide APIs through which stock dataset from multiple firms can be acquired by giving specifications.

The following procedures are used to transform the data into a format that may be used with a prediction model:

1. Transformation of time-series data into input-output components for supervised learning.
2. Scaling the data to the $[-1, +1]$ range.

5. CONCLUSION

Many individuals all over the world are interested in stock investment. Making a selection, however, is a challenging undertaking because there are numerous considerations. Investors who have done well in their investments are keen to forecast the direction of the stock market. The impact of even a small performance gain can be huge. By offering supplementary data like future stock price guidance, a competent forecasting system will assist investors in making investments that are more profitable and accurate. Other relevant information, such as politics, economic development, financial issues, and the climate on social media, may also have an impact on prices in addition to historical data. Emotional analysis has a substantial impact on future prices, according to numerous research. Therefore, combining technical analysis with fundamental analysis might lead to highly accurate predictions.

6. ACKNOWLEDGMENT

We would like to thank all the working staff of Dhole Patil College of Engineering Pune and a great thank to our project guide Prof. Supriya Agre who gave her best to make us do this great project. Also, many thanks to all the writers whose books and papers we consulted for this effort. We never would have finished our job without the assistance of all these resources.

References

- [1]. J. Fernando. Price-to-Earnings Ratio – P/E Ratio <https://www.investopedia.com/terms/p/priceearningsratio.asp>. Nov 13, 2020
- [2]. Roondiwala, Murtaza, Harshal Patel, and Shraddha Varma. "Predicting stock prices using LSTM." *International Journal of Science and Research (IJSR)* 6.4 (2017): 1754-1756.
- [3]. Nandakumar, R., Uttamraj, K. R., Vishal, R., & Lokeswari, Y. V. (2018). Stock price prediction using long short term memory. *International Research Journal of Engineering and Technology*, 5(03).
- [4]. Rajput, Vivek, and Sarika Bobde. "Stock market forecasting techniques: literature survey." *International Journal of Computer Science and Mobile Computing* 5.6 (2016): 500-506.
- [5]. Pahwa, Kunal, and Neha Agarwal. "Stock market analysis using supervised machine learning." *2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)*. IEEE, 2019.
- [6]. Yu, Pengfei, and Xuesong Yan. "Stock price prediction based on deep neural networks." *Neural Computing and Applications* 32.6 (2020): 1609-1628.
- [7]. Singh, S., Madan, T. K., Kumar, J., & Singh, A. K. (2019, July). Stock market forecasting using machine learning: Today and tomorrow. In *2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICT)* (Vol. 1, pp. 738-745). IEEE.
- [8]. [8] Nayak, Aparna, MM Manohara Pai, and Radhika M. Pai. "Prediction models for Indian stock market." *Procedia Computer Science* 89 (2016): 441-449.
- [9]. Ghosh, Achyut, Soumik Bose, Giridhar Maji, Narayan Debnath, and Soumya Sen. "Stock price prediction using LSTM on Indian share market." In *Proceedings of 32nd international conference on*, vol. 63, pp. 101-110. 2019.
- [10]. Wei, Dou. "Prediction of stock price based on LSTM neural network." *2019 International Conference on Artificial Intelligence and Advanced Manufacturing (AIAM)*. IEEE, 2019.
- [11]. Sadia, K. H., Sharma, A., Paul, A., Padhi, S., & Sanyal, S. (2019). Stock market prediction using machine-learning algorithms. *Int. J. Eng. Adv. Technol*, 8(4), 25-31.
- [12]. J. Jagwani, M. Gupta, H. Sachdeva, and A. Singhal. Stock price forecasting using data from yahoo finance and analysing seasonal and nonseasonal trend. pages 462–467, 06 2018.
- [13]. Cheng, J.; Huang, K.; Zheng, Z. towards Better Forecasting by Fusing Near and Distant Future Visions. In *Proceedin*
- [14]. Usmani, Mehak, Syed Hasan Adil, Kamran Raza, and Syed Saad Azhar Ali. "Stock market prediction using machine learning techniques." In *2016 3rd international conference on computer and information sciences (ICCOINS)*, pp. 322-327. IEEE, 2016.