

Stock Price Prediction Based on Deep Learning Using Long Short-Term Memory



R. Shanmuga Priya and C. Sruthi

Abstract Predicting the stock market is either the easiest or the toughest task in the field of computations. There are many factors related to prediction—physical factors versus physiological, rational and irrational behaviour, capitalist sentiment, market rumours, etc. All these aspects combine to make stock costs volatile and are extremely tough to predict with high accuracy. The prices of a stock market depend very much on demand and supply. High-demand stocks will increase in price, while heavy selling stocks will decrease. Fluctuations in stock prices affect investor perception, and thus, there is a need to predict future share price and to predict stock market prices to make more acquaint and precise investment decisions. We examine data analysis in this domain as a game-changer. This paper proposes that historical value bears the impact of all other market events and can be used to predict future movement. Machine learning techniques can detect paradigms and insights that can be used to construct surprisingly correct predictions. We propose the long short-term memory (LSTM) model to examine the future price of a stock. This paper is to predict stock market prices to make more acquaint and precise investment decisions.

Keywords Machine learning · Deep learning · RNN · LSTM

1 Introduction

Financial markets are one of the most significant inventions of our time. They have a notable impact on business, education, employment, technology and thus the economy in many areas. Stock market forecasting has always attracted researchers and analysts. It acts like a voting machine in the short term, but it is actually weighing machine in the long term, and therefore has the scope to predict market movements over a longer period. The stock market fluctuations are incredibly violent which leads to more complex financial indicators. Also, the greatest advantage of price prediction is to maximize the benefits of purchase options simultaneously minimizing its risk. [1, 2] The “Efficient Market Hypothesis” (EMH) suggests it is not possible to

R. S. Priya (✉) · C. Sruthi
Madras Institute of Technology, Anna University, Chennai, Tamil Nadu, India

make the market low with the usage of the same information because the current overall stock price reflects all available information about a firm. Despite this, in the proposed model historical data is used to predict future stock price. In recent years, the increasing prominence of machine learning has produced quite promising results in various fields and industries that have enlightened many analysts, researchers and traders to use machine learning techniques. LSTM is the widely preferred RNNs architecture. LSTM has a memory cell, which acts as a unit for computation. This method is well suited for classifying, processing and predicting statistical data. The paper intends to predict stocks using LSTM. The main objective of forecasting is to reverse the uncertainty associated with investment determination.

2 Existing Approaches

Many existing systems are used to predict stock prices such as linear regression, k-nearest neighbour, Auto Arima and prophet. To estimate the stock market, two-way back-propositional neural networks, auto-regressive moving average models as well as multi-layer perception have been used [3]. In existing systems, the performance of a company, in terms of its stock price movement, is foreseen by internal communication patterns. In these systems, accuracy is unreliable when determining multiple levels of stock movements. For example, using the decision tree as a classifier, the average prediction accuracy was 63.7, 31.92 and 12% for “two levels”, “three levels” and “five levels”, respectively [4]. For each input knowledge, it performs the same function with its output of the current state of input that depends on the previous calculations; from this, it shows recurrent neural networks (RNN) have recurring properties. There exist a vanishing gradient and exploding gradient problem. This result clearly shows that stock prices are not predictable using these methods [5].

3 Proposed Method

Accuracy is vital in market forecasting. Although many algorithms are available for this purpose, making precise selections continues to be the authentic work to achieve the most constructive results. In this work, an effective and reliable stock market forecasting system is hypothesized by combining informative input variables along with RNN. Primarily, the problem of processing the concatenation of data, such as audio, video and text can be solved by RNN. Recurrent neural networks can be discernible of feed-forward neural networks that pose an attached memory. For each input knowledge, it performs the same function with its output of the current state of input that depends on the previous calculations; from this, it shows RNN has recurring properties. Once the output is produced, it is cloned and forwarded to the recurrent networks. From the previous inputs, it can learn contemporaneous input and output which is in turn used to create a call. To process a sequence of inputs, RNNs

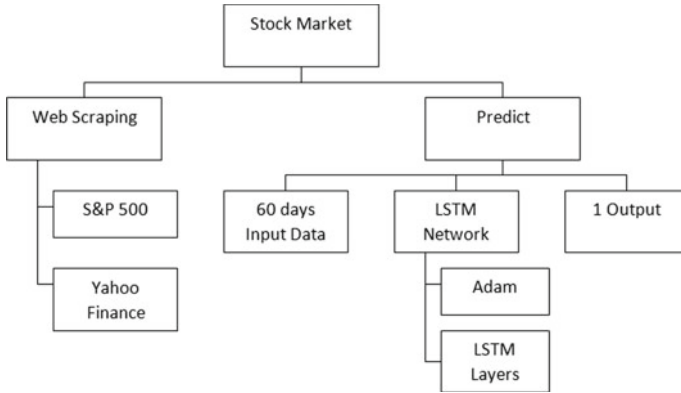


Fig. 1 Architecture diagram

uses its internal state memory which is disparate to feed-forward neural networks. All inputs are independent of every instance input in other neural networks. Whereas, in RNN, whole inputs are interconnected to each other.

In addition, a novel stock market price prediction model based on LSTM using deep learning and stock market basic data is being proposed, and further validation is done. Figure 1. Architecture Diagram. The justification on which they work is that the LSTM is ready to store previous information that is important and neglect information that is not [6, 7].

3.1 Architecture Diagram

See Fig. 1.

3.2 LSTM Approach

Long short-term memory (LSTM) networks are distinctively different from traditional RNNs and in neuron formations. In traditional RNN and normal neural networks, it does not pose any difference between its neurons. But, LSTM connects previous information to the current errand by its apiece neuron which is also a “memory cell”. In LSTM, a unit of computation is being introduced by the memory cell that supplants conventionally in the hidden layer of the network by artificial neurons. Networks can effectively able to connect to its memory and input gradually using these memory cells, therefore with a higher predictive capacity to suit the structure of knowledge over time. RNN poses a vanishing gradient problem that is solved here. Figure 2. LSTM architecture.

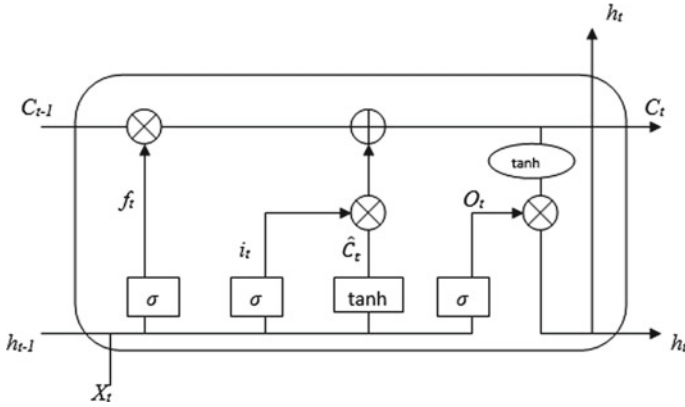


Fig. 2 LSTM architecture

3.3 LSTM Architecture

Functions in LSTM:

tanh: Function’s second derivative can be maintained for a prolonged distance before zero is required to overcome the vanishing gradient problem. The suitable function for the above property can be *tanh*.

Sigmoid: Since the sigmoid output can either be 0 or 1, it is accustomed to obliterate or recall its knowledge. By using such LSTM, unit’s information can be passed.

LSTM unit has three main components:

1. *Forget gate*—Search for details to be dismissed from the block. Here sigmoid layer determines input as $X(t)$ and $h_{(t-1)}$ and elects the parts that can be removed from the previous output (giving an output of 0). This is known as forget gate $f(t)$. The output is $f(t) * c(t-1)$.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{1}$$

2. *Input gate*—To modify its memory which value is to be used from the input can be identified. A value that passes over 0 and 1 is decided by the sigmoid function. The weightage for the values that are passed is given by tanh function, setting its significance levels from -1 to 1. Along with old memory $c(t-1)$, the new memory is computed to offer $c(t)$.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{2}$$

$$\tilde{c}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{3}$$

3. *Output gate*—To determine the output, input and block memory is being used. A value that passes over 0 and 1 is decided by the sigmoid function. The weightage for the values that are passed is given by tanh function, setting its significance levels from -1 to 1, which is multiplied by the sigmoid function's output.

$$O_t = \sigma(W_o[h_{t-1}, x_t]) + b_o \quad (4)$$

$$h_t = O_t * \tanh(C_t) \quad (5)$$

It is observed that RNN and LSTM have a major architectural difference between them. During its long-term memory basis these prototype pedagogies which information is to be stocked and eliminated in the LSTM.

4 Methodologies

The data used here is from Yahoo Finance, S&P 500 component shares. This includes open, high, low, close, volume values of 500 companies.

This paper uses a close price to estimate the future close price. This section discusses the functioning of our system. There are several stages in this system.

They are:

- *Stage I: Raw data:* For forecasting stock prices in future, historical stock data is collected and employed in this phase.
- *Stage II: Data preprocessing:* A portion of data is reduced but with specific criteria, especially for numerical data this involves the preprocessing stages and knowledge file integration. After the transformation of the dataset into a proper data fields, it has been segregated into training data and testing data for evaluation. The test data is retained for about 5 to 10% of the entire datasets.
- *Stage III: Feature extraction:* During this stage, only the selected features are fed into the neural network. Here, the close price is selected.

In the recurrent network, the neural network is being used, where the sequence data is formed by the transformation of input data. By using 60 width sliding windows, segmentation of input data is done. This process is described in Fig. 3. Input Sequence Diagram.

- *Stage IV: Training neural network:* Here, our model assigns random biases and weights by feeding information to the current neural network, and predictions are also done by training the network. In our model, the neural network has four layers including the input layer, two LSTM layers (hidden layer) and the output layer (refer Fig. 4. Prediction Model. All units in adjacent layers are connected by each unit in a layer. The LSTM layer consists of 46 LSTM nerve entities. The fully connected layer is formed by the output layer which has only one unit.

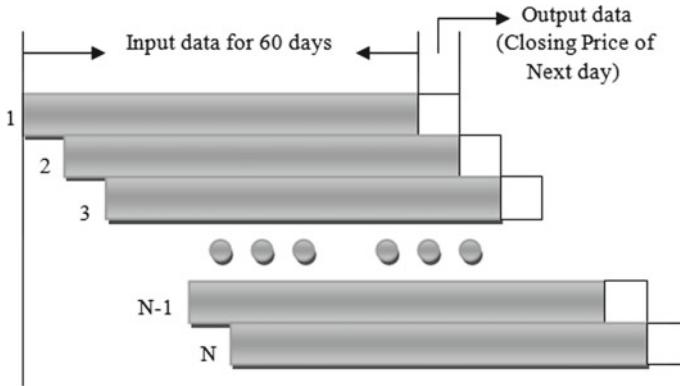
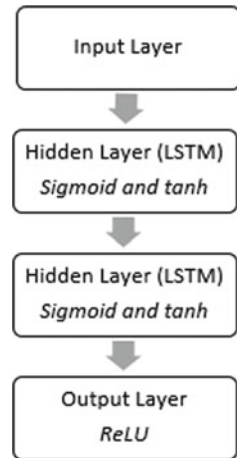


Fig. 3 Input sequence diagram

Fig. 4 Prediction model



In the stock price prediction prototype, it poses three types of activation functions: rectified linear unit (ReLU), hard sigmoid and *tanh*. The ReLU is present in the output layer; σ (hard sigmoid) and *tanh* are shown in Sect. 3.3.

- *Stage V: Output generation:* To the target value, the output value is being compared which was in turn generated by the output layer of our model. A back-propagation algorithm is used to find inaccuracy and dissimilarities that ally the targeted and received output value which are being minimized by adjusting its weights; therefore, it is the bias of a network.

5 Analysis

For predicting stock index, a predictive model requires less error that can be taken into account of processing the input data. In our model, rectified linear unit (ReLU) activation is used. The activation function $f(x) = \max(0, x)$ is applied in ReLU to remove the vanishing gradient and uses element activation when it limits to 0. However, ReLU has linear properties so it will speed up the sigmoid function / tanh being compared to a stochastic gradient. Moreover, it creates an activation function matrix when its threshold bound to 0 and exponential operations like *sigmoid / tanh* are not being used. Neuron becomes re-indolent in data points because the major flow of gradients in ReLU occurs to weight gain where it becomes fragile during its training. The flow of gradient will bound to zero perpetually if this proceeds over that point. From this, ReLU can be terminated irreparably while its training and it can debilitate numerous data units. Using LSTM, unfamiliar values for the next day can be found, and also mediocre values can be estimated on the next day. Investors, analysts or anyone can invest in the stock market in terms of their own interest. This paves the way for analysts to be aware of the stock market in the future.

Also, a series of tests and results are being conducted by measuring a few criteria to evaluate the neural network model’s performance. MSE, MAE, RMSE and R-squared metrics are primarily used to evaluate regression error rates and model performance in regression analysis.

- The AAE (average absolute error): represents to extract the absolute difference of the dataset on an average basis by calculating its difference from the original and predicted values (Eq. 6).
- The MSE (mean squared error): shows the squaring mean difference of the dataset to extract values by finding its difference from the original and predicted values (Eq. 7).
- RMSE (root mean square error): is an error rate obtained by the square root of MSE (Eq. 8).
- The R-squared (coefficient of determination): represents the value that fits perfectly comparing to original values (Eq. 9). Values from 0 to 1 are interpreted as a percentage. The higher the value, the better the model.

$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_i - \hat{Y}| \tag{6}$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y})^2 \tag{7}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y})^2} \tag{8}$$

Table 1 Prediction performance of LSTM for S&P 500

Company	MAE	MSE	RMSE	R ²
AAPL	1.8533	3.8756	1.9686	0.9745
GOOGL	0.5524	0.0679	0.26	0.9927
AMZN	0.8278	0.975	0.987	0.9836

$$R^2 = 1 - \frac{\sum(Y_i - \hat{Y})^2}{\sum(Y_i - \bar{Y})^2} \tag{9}$$

where Y_i —actual value, \hat{Y} —predicted value and \bar{Y} —mean of Y .

Table 1 Prediction performance of LSTM for S&P 500. Figure 5, 6, 7 Close predictions of AAPL (Apple), GOOGL (Google) and AMZN (Amazon), respectively. From the table prediction performance of three companies AAPL, GOOGL and AMZN, it is clear that the accuracy of our model is more than 97%.



Fig. 5 Close prediction of AAPL



Fig. 6 Close prediction of GOOGL



Fig. 7 Close prediction of AMZN

6 Conclusion

The stock market’s trading popularity is increasing swiftly, and it encourages scholars to discover intact methods in forecasting using new terms and techniques. Also, the researchers, investors and individuals are being helped by this forecasting technology in exchange. A forecasting method with ethical accuracy is used to predict stock indices. LSTM outperforms other models because this model learns by using long-term dependency. LSTM is a step forth to RNN; its ability to update, obliterate and recollect information is more effective. In this work, one of the most accurate precision forecasting techniques is used that helps investors, analysts and individuals eagerly to invest by granting good knowledge for the future state of exchange in the stock market exchanges. This can be further improved by an LSTM model with multiple inputs that can extract relevant information by employing additional input gates with low correlated factors to discard their noise controlled by convincing factors called mainstream.

References

1. Silva TR, Li AW, Pamplona EO (19–24 Jul 2020) “Automated trading system for stock index using LSTM neural networks and risk management”. International joint conference on neural networks (IJCNN), United Kingdom
2. Sarkar A, Sahoo AK, Sah S, Pradhan C (13–14 Mar 2020) “LSTMSA: a novel approach for stock market prediction using lstm and sentiment analysis”. International conference on computer science, engineering and applications (ICCSEA), India
3. Pahwa K, Agarwal N (14–16 Feb 2019) “Stock market analysis using supervised machine learning”. International conference on machine learning, big data, cloud and parallel computing (Com-IT-Con), India
4. Wei D (16–18 Oct 2019) “Prediction of stock price based on LSTM neural network. International conference on artificial intelligence and advanced manufacturing (AIAM)
5. Shin HG, Ra I, Choi YH (16–18 Oct 2019) “A deep multimodal reinforcement learning system combined with CNN and LSTM for stock trading”. International conference on information and communication technology convergence (ICTC), Korea

6. Lai CY, Chen RC, Caraka RE (7–10 Jul 2019) “Prediction stock price based on different index factors using LSTM”. International conference on machine learning and cybernetics (ICMLC), Japan
7. Zhang Y, Yang S (18–20 Oct 2019) “Prediction on the highest price of the stock based on PSO-LSTM neural network”. 3rd international conference on electronic information technology and computer engineering (EITCE), China