

SCIENCE | ENGINEERING | TECHNOLOGY

e-ISSN: 2319-8753 | p-ISSN: 2347-6710

EDIA

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 11, Issue 5, May 2022

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.118

9940 572 462

6381 907 438

 \bigcirc

e-ISSN: 2319-8753, p-ISSN: 2320-6710 www.ijirset.com | Impact Factor: 8.118



Volume 11, Issue 5, May 2022

DOI:10.15680/IJIRSET.2022.1105067

Stock Price Prediction Using Deep Learning

Pranshu Bansal¹, Sarthak Bassi², Shoubhikroy Samanta³, Anshu Kumari⁴, Preeti Walia⁵

U.G. Student, Department of Computer Science Engineering, HMR Institute of Technology and Management,

Delhi, India¹

U.G. Student, Department of Computer Science Engineering, HMR Institute of Technology and Management,

Delhi, India²

U.G. Student, Department of Computer Science Engineering, HMR Institute of Technology and Management,

Delhi, India³

U.G. Student, Department of Computer Science Engineering, HMR Institute of Technology and Management,

Delhi, India⁴

Assistant Professor, Department of Computer Science Engineering, HMR Institute of Technology and Management,

Delhi, India⁵

ABSTRACT: The usage of deep learning for predicting stock market prices and trends in today's scenario has been gaining huge popularity day after day. Through our project, we have aimed to implement an approach for forecasting stock prices through the use of deep learning. Our goal is to forecast the stock prices so that we get the ability for making informed and accurate investment decisions. Hence, we present a forecasting system for stock prices by combining mathematical functions, deep learning approaches, and various external factors with the aim of gaining higher accuracy in predicting stock prices, therefore obtaining profitable trades. Two main categories are there that most stocks fall into. Intraday trading is more commonly known as "day trading". Traders involved in these trades hold positions in the securities lying in a time frame ranging from at least one day to the next, which can be for several days, weeks, or months in some cases. LSTMs are well suited to sequence prediction problems as they have the capability to retain information from the past. This plays a significant role in our scenario, as the previous price of a stock is essential for forecasting the movement of the price of the stock. Having obtained a design in detail, and evaluating the prediction term lengths, following the feature engineering processes, and methodologies for pre-processing the data, our project provides a contribution to the research community involved in stock analysis present in technical and financial domains.

KEYWORDS: Deep learning, Stock market, LSTM, Forecasting, Ticker, Profit

I. INTRODUCTION

A financial market is a place that lets people trade a variety of currencies, stocks, derivatives, and equities over a number of virtual platforms that are supported by brokers. The stock market facilitates investors the ability and chance to own shares of publicly traded companies through trading which can be done by the method of exchange or over-the-counter markets. The stock market gives the investors the chance to earn and gain money, to live an affluent lifestyle through the means of investing varying amounts of money while facing the risk that may be lower than what one may face in the risk of opening a new business or requiring a high salaried career. Stock markets are highly sensitive and can be affected by various factors which can cause uncertainty and high volatility in the market. While both humans and automated trading systems (ATS) can take orders and submit them to the market alive but the automated trading systems are operated by the implementations of computer programs which allow them to have better performance than humans in terms of higher momentum in submitting orders.^[5]

However, the performance of ATS's still needs evaluations and control over the performance, and certain risk strategies and safety measures need to be implemented based on human judgment. Certain factors need to be incorporated during the development process of ATS, for instance, various trading strategies have to be adopted, machine learning algorithms

よう 総 IJIRSET | e-ISSN: 2319-8753, p-ISSN: 2320-6710| <u>www.ijirset.com</u> | Impact Factor: 8.118|

Volume 11, Issue 5, May 2022

DOI:10.15680/IJIRSET.2022.1105067

that predict the future stock values, mathematical functions that evaluate the state of specific stocks, and taking in consideration any news related to the stock under analysis.

Time-series prediction is one of the commonly used techniques which is used in real-world applications like weather forecasts and financial market predictions. Through the usage of continuous data during a certain period of time to predict the result in the next time unit. A variety of time series algorithms have shown that they are quite effective in practice. Commonly used algorithms are based on Recurrent Neural Networks (RNN) and their advanced versions-Long-short Term Memory (LSTM) and Gated Recurrent Unit (GRU). The stock market is one of the typical fields that present time-series data and it is studied by many researchers that propose a wide number of models for it.^[3]

II. METHODOLOGY

Long Short-Term Memory:

Long Short-Term Memory is the advanced version of recurrent neural networks (RNN). In the RNN process, the output from the last step is taken as the input in the current step. Hochreiter & Schmidhuber are the designers of LSTM.LSTM is the advanced version of RNN, which tackles the issue of long-term dependencies which stopped RNN from making predictions regarding the words stored in long-term memories but still displays more accurate predictions even from the recent information. The increase in gap length decreases the efficient performance of RNN.LSTM has the ability to retain information for long periods of time by default, which facilitates its use in processing, prediction, and classification on the basis of time-series data.



Structure Of LSTM:

The structure of LSTM is a chain which contains four neural networks and a collection of various blocks of memory, which are known as cells.^[4]

The information is preserved by the cells, and the manipulations to the memory are done by the means of gates. The three types of gates are as follows -

| e-ISSN: 2319-8753, p-ISSN: 2320-6710| <u>www.ijirset.com</u> | Impact Factor: 8.118|



Volume 11, Issue 5, May 2022

DOI:10.15680/IJIRSET.2022.1105067

1. Forget Gate: The main purpose of this gate is to remove the information that is no longer useful in the cell state. Two inputs x_t (input at the particular time) and h_t-1 (previous cell output) are to be provided at the gate and then these need to be multiplied with weight matrices, after which the addition of bias is performed. The result obtained is then passed through an activation function which provides us with a binary output. In case a particular cell state the output is 0, the piece of information is forgotten and for output 1, the information is retained for future use.

2. Input gate: The purpose of this gate is to add the information that is useful to the cell state. In the first step, the information is regulated through the means of the sigmoid function and the values that are to be remembered are filtered, which is similar to the forget gate utilising inputs h_{t-1} and x_t . Further, a vector is formed by using the tanh function which provides an output ranging from -1 to +1, containing all the possible values in the range of h_{t-1} and x_t . At last, the values of the vector and the regulated values are then multiplied for getting the useful information

3. Output gate: The output gate is responsible for executing the process of extracting information that is of use from the present state of the cell to be provided as output. Tanh function is applied on the cell for generating a vector. In the next step, regulation of information takes place through the means of the sigmoid function and then these are filtered according to the values which are to be remembered using inputs h_t-1 and x_t . Finally, the values of the vector and the regulated values are multiplied, which are then sent as an output and input to the next cell.

III. WORKFLOW

Data Description:

First, we require the dataset before making any predictions and training the neural network The category of data that is required, time series: sequence of numbers in a chronological order.

You can either pick daily adjusted or weekly adjusted, open/high/low/close/volume values, adjusted close values, and historical split/dividend events of the global equity specified.

We need the following attributes:

- open price
- the highest price of that day
- the lowest price of that day
- closing price
- adjusted close price (this is used in this project)
- volume

To prepare a training dataset for our neural network, we will be using adjusted close stock prices. It implies that we can aim to forecast the closing price in future. The below graph describes the weekly closing prices of Microsoft Corporation over 20 years:

Validation of Data before training:

Supervised learning is to be used for this project, which involves providing the data into the neural network, so that it can learn through mapping the input data to the output labels. A common method for preparing the training dataset is by extracting the moving average from the given time-series dataset.^[6]

Simple Moving Average, commonly known as SMA, is a way for identifying the direction of trends over certain intervals of time. This can be done through the means of observing the mean of all the values that fall within that time window. The count of prices in a time window is then selected through experiments.

The training data can be made using the data of the weekly stock prices and the SMA which is calculated. Considering the window size to be 50, it implies that we are going to be using the closing price of every 50 consecutive weeks as our training features (X), and the SMA as training label (Y).

Further, the dataset is split into two sets, the training and validation sets. If 70% of the data has been used for training, then 30% is to be used for validation. The API returns us approximately 1000 weeks of data, so that leaves us with 700 for training, and 300 for validation.

| e-ISSN: 2319-8753, p-ISSN: 2320-6710| www.ijirset.com | Impact Factor: 8.118|



Volume 11, Issue 5, May 2022

DOI:10.15680/IJIRSET.2022.1105067

Model Training Phase:

Now that the training data is ready, it is time to create a model for time series prediction, to achieve this we will use the Tensorflow.js framework. For deploying the machine learning capability in a web browser and training machine learning models in JavaScript, we use Tensorflow.js library.^[2]

A sequential model is selected that simply connects each layer and passes the data from the input to the output during the training process. In order for the model to learn time-series data that are sequential, a recurrent neural network (RNN) layer is created and a number of LSTM cells are added to the RNN, which makes it a LSTM model.

The model will be trained using Adam (research paper), a popular optimization algorithm for machine learning. Root mean square error which will determine the difference between predicted values and the actual values, so the model is able to learn by minimizing the error during the training process.^[1]

The hyper-parameters that are being used in the training process and can be tweaked with in the frontend are:

- Training Dataset Size (%): the quantity of data that is used in training process and the remaining part will be used for validation process.
- Epochs: It denotes how many times training dataset has been trained by the LSTM model
- Learning Rate: the pace with which the weights are adjusted during training in each step
- Hidden LSTM Layers: the number of layers of LSTM that are stacked on top of each other to increase the complexity

Validation of Data after training:

After training the model, it is used for predicting the future values which in our case is the moving average. We shall use the machine learning model and the prediction function in Tensorflow.js

The data is divided into 2 parts-the training part and the validation part. The training part is used to train the model and the validation part is used validate whether the predictions of the model map closely to the true values. It will be beneficial for the model to predict values that are closer to true values since this model does not have any validation dataset.^[7]

The remaining data can be used for prediction which will permit us to observe and compare the predicted values with the actual values.

Prediction Description:

The model is validated at last, and the obtained predictions are mapped closely to the true values that they have, and hence, these can be used for forecasting the future values. model. Predict method can be used once again considering the last 50 data points provided as input, as we have kept the window size as 50. As the training data is incremented on a daily basis, we can provide the data of the preceding 50 days as input for forecasting the output for 51st day.^[8]

IV. PREDICTION STEPS

Fetch Data According to Symbol

	1011	
AADI		
AAF L	AL02300AOW243N0N	
Demo API key only allows 'MSFT'		
FETCH DATA		

| e-ISSN: 2319-8753, p-ISSN: 2320-6710| <u>www.ijirset.com</u> | Impact Factor: 8.118|



Volume 11, Issue 5, May 2022

| DOI:10.15680/IJIRSET.2022.1105067 |

Loaded Stocks Data

Symbol: AAPL (last refreshed 2022-05-06)



Set Criteria for Training

Training Dataset Size (%)	Epochs	
75	10	
Learning Rate	Hidden LSTM Layers	
0.01	4	
Typically range between 0.01 and 0.1	Number of LSTM lavers	

Progress and Loss Function



e-ISSN: 2319-8753, p-ISSN: 2320-6710 www.ijirset.com | Impact Factor: 8.118

Volume 11, Issue 5, May 2022

DOI:10.15680/IJIRSET.2022.1105067

Compare True Values to Predicted Value







V. CONCLUSION

With the introduction of Deep Learning and its strong algorithms, the most recent market research and Stock Market Prediction advancements have begun to include such approaches in analysing stock market data. The Opening Value of the stock, the Highest and Lowest values of that stock on the same days, as well as the Closing Value at the end of the day, are all indicated for each date. Furthermore, the total volume of the stocks in the market is provided. With this information, it is up to us to look at the data and develop different algorithms that may help in finding appropriate stocks values.

Predicting the stock market was a time-consuming and laborious procedure a few years or even a decade ago. However, with the application of Deep learning for stock market forecasts, the procedure has become much simpler. It not only saves time and resources but also outperforms people in terms of performance. It will always prefer to use a trained computer algorithm since it will advise you based only on facts, numbers, and data and will not factor in emotions or prejudice.

| e-ISSN: 2319-8753, p-ISSN: 2320-6710| www.ijirset.com | Impact Factor: 8.118|



Volume 11, Issue 5, May 2022

DOI:10.15680/IJIRSET.2022.1105067

In this research, we study the problem of stock market forecasting using Recurrent Neural Network (RNN) with Long Short-Term Memory (LSTM). The purpose of this research is to examine the feasibility and performance of LSTM in stock market forecasting.

REFERENCES

- [1] https://www.sciencedirect.com/science/article/pii/S1877050918307828
- [2] https://www.tensorflow.org/js
- [3] https://www.geeksforgeeks.org/introduction-deeplearning/#:~:text=Deep%20learning%20is%20a%20branch,deep%20learning%20is%20not%20new.
- [4] https://en.wikipedia.org/wiki/Long_short-term_memory
- [5] Stock Price Prediction Using LSTM on Indian Share Market by Achyut Ghosh, Soumik Bose1, Giridhar Maji, Narayan C. Debnath, Soumya Sen
- [6] 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," in International Conference on Advances in Computing, Communications and Informatics, 2017.
- [7] Heaton J. B., Polson N. G., and Witte J. H. (2017). "Deep learning for finance: deep portfolios." Applied Stochastic Models in Business and Industry 33 (1): 3-12.
- [8] Jabin S. (2014). "Stock market prediction using feed-forward artificial neural network". growth 99 (9)





Impact Factor: 8.118





INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY





www.ijirset.com