

STOCK MARKET ANALYSIS USING LONG SHORT-TERM MEMORY

Sushanth M K¹

Computer Science and Engineering Department, Srinivas institute of technology, Valachil

Namratha Pai²

Computer Science and Engineering Department, Srinivas institute of technology, Valachil

Alwyn Edison Mendonca³

Assistant Professor, Computer Science and Engineering Department, Srinivas institute of technology, Valachil

Abstract: *Stock Market refers to collection of markets and exchanges where regular activities of the buying, selling and shares of publicly held companies takes place. The paper focuses on the use of Long Short Term Memory (LSTM) tensorflow model based on machine learning using open source libraries to predict the future value of the financial stocks of a company. In this paper we will see the implementation bringing the acceptable results. The result will be based on the available dataset.*

KEYWORDS: Stock Market, Data Analysis, Implementation, LSTM, Machine Learning

1. INTRODUCTION

A stock market is a public market for the dealing of company stock and derivatives at an acknowledged price. Shares portray an ownership claim on the company, and consequently a claim on future revenue. The stock market is also called the secondary market as it involves trading between two investor. Stock market gets investors together to buy and sell their share. Shareholders in the stock market want to maximize their returns by buying or selling their investments at a suitable time. Once the shareholder has gained exposure to a stock, they can make money in the following ways. Firstly, they can wait for the price of each share to rise and then sell it. The difference between the purchase and sell price, minus any expenses, is the investor's profits. These profits are known as capital gains. The other way they can make money, is through dividends. The cost of a stock is set through demand and supply. Each buy and sell order go in an order book on the exchange deciding the bid and ask. Through the best bid and ask, the trade works out the mid price, which is then quoted as 'the price'. The basic principle is quite simple. Companies will list their shares in the companies as small products called Stocks. They do so in order to raise money for the firm. A company lists its stock at a price called the IPO or initial public offering. This is the offer price at which the company sells the stock and raises money. After which these stock are the property of the owner and he may sell them at any price to a buyer at an Exchange such as BSE (Bombay Stock Exchange) or NSE(National Stock Exchange).[1]

Demand and supply of a stock can be affected by a number of factors. Short term fluctuations can be due to anything from investor sentiment, to rumors and newspaper articles. Over the long term however, it is often said that the share price will gravitate over its long term averages; which is guided by the stock's value. Since stock market data are highly time-variant and are normally in a nonlinear pattern, forecasting the price of a stock is highly demanding task. The stock's value is determined by the company's intrinsic value as well as other external factors (e.g. being the market leader). That is, how much is the company worth today taking into consideration today's and future profits. Prediction provides knowledgeable details regarding the current status of the stock price variations. Thus this can be utilized in decision making for customers in finalizing whether to buy or sell the particular given stock.

2. RELATED WORK

Artificial neural networks are generally considered to have high predictive power. Time series Model [1] using neural network uses the back propagation feed forward network with one input, output and hidden layer each. Predicting the closing price one day in advance is done by this technique. This method uses the Gradient descent algorithm as the learning function and the sigmoid function as the activation function. This model works well when the noise is less and prediction accuracy is reduced considerably when the noise is increased. In this method the parameter used is closing price of stock. Advantages of this technique is lower prediction error and its performance is better than regression. Disadvantages of this technique is Prediction gets worse when the noise variation is increased.

The process is used to predict the market close value and open value of a share. This model advises the user to either withhold buy or sell a share resulting in good decision capability. Data is obtained from Yahoo Finance and it is represented as a time series model. The model uses artificial neural network method [2] which has three layers namely: input layer, hidden layer and output layer and considers maximum epochs to be 5000. This model takes the mean of both the actual data and predicted data from the training set. Based on these values the model suggests the investor to buy a particular stock or else not to buy the stock wait for particular time to buy the same.

The main purpose of this process is to predict the stock prices of companies listed under L1X15 index of National Stock Exchange. This paper uses NNT MATLAB and MLP neural network algorithm [3]. This algorithm is trained for different combinations of data and using the result of this algorithms stock price is predicted. Stocks that are in the demand will increase their price, whereas as stocks that are being heavily sold will decrease their price. In this method, prediction will provide knowledgeable information regarding the current status of the stock price movement. This system can be used to study the current stock market trend and collect trend data. This system provides result and outputs that are taken for all the companies that are listed under L1X15 based on different combination of data and parameters.

The process mainly focuses to improve the effectiveness of national economy. The stock market forecasting is essential to analyse the future stock's returns on investment by using the present stock values. The process applies certain process of flow such as data acquisition process in which acquiring different types of data such as textual data and numerical data from servers, data classification process in which organising data into different categories for its more efficient use, data decomposition process which splits the data into smaller tasks. At the end of the process, data is filtered seasonally by applying saesadj function which means seasonal adjustment. Finally, an automated arima model [4] will be used to predict the future data approximately. This model can be used to predict the stock price on short term basis.

Linear regression in concurrence with the least square method for determining the parameters and moving average method is used to forecast the stock market time series data. The reason for selecting this method is that it tends to reduce the fluctuation in market and obtain the value of trend of time series with high accuracy. The percentage earning which is the ratio of current market price and the earning per share which is used to make the regression equation. This system is able to predict the monthly and weekly dynamic movement of stock prices. The main benefit of this approach is that it yields a numeric forecasting method [5] which has advantage and is easy to understand in comparison to some words such as high, low, medium, flat, which may be ambiguous if their range is not mentioned. The parameter used in this technique is percentage earnings ratio calculated from daily and weekly activities whose data is extracted from stock exchange. The advantage of this method is predict six-month future value and moving average reduces the fluctuations.

The main objective of this process is to predict the value of the stock market. The first thing we have taken into account is the dataset of the stock market prices from previous year and various social media data. The dataset was preprocessed and turned up for real analysis. The technique employed in this

paper is linear regression [6]. This paper predicts the target as a weighted sum of the feature inputs. This model does not perform well when there is a change in the operating environment.

3. MACHINE LEARNING AND TENSORFLOW

Machine learning is an application of artificial intelligence (AI) which enables systems the capacity to automatically learn and improve from experience without being directly programmed. Machine learning focuses on the development of computer programs that can use and retrieve data and use it learn for themselves.

TensorFlow is an end-to-end open source programming platform for machine learning. It has a complete, flexible ecosystem of tools, libraries and resources that helps researchers push the state-of-the-art in ML and developers easily build and deploy ML powered applications. TensorFlow eases the process of acquiring data, training models, serving predictions, and refining future results. TensorFlow packs together a slew of machine learning and deep learning (aka neural networking) models and algorithms and makes them useful by way of an ordinary metaphor. It utilizes python to provide a suitable front-end API for building applications with the framework.

TensorFlow enables developers to create dataflow graphs structures that defines the data movement through a graph or a series of processing nodes. Every node in the graph constitutes a mathematical operation and each connection or edge between nodes is a multidimensional data array or tensor.

4. PREDICTION MODEL

4.1 Data Analysis Stage

In this stage, first we shall look at raw data available to us. The data that we are taking for our model is taken using web scrapping method. This method uses python language for implementation and it will download the data from yahoo finance for the respective stock company name. This method also gives historical data of stock company which are stored in .csv format. This file contains five attributes. They are

- Open: Opening price of the day
- Close: Closing price of the day
- High: Highest price at an instance of time
- Low: Lowest price at an instance of time
- Volume: It is the actual number of shares traded of the particular company that particular day

4.2 Training and Testing Stage

In this stage, the datasets downloaded in the .csv format is used. From the dataset all the NaN data values are removed and then the dataset is scaled. The scaled data is divided into 80% training data and 20% testing data using train_test_split function from sklearn library.

Firstly, the training data is normalized and then it is trained into our LSTM model. Normalization is achieved by training 60days data at a time and considering 61st data as the target.

The normalization is carried till the number of datasets included in the .csv file. The model at a time uses dataset of single attribute for e.g. 'Open' and then it will divide this dataset into train data and test data based on scaling and normalization.

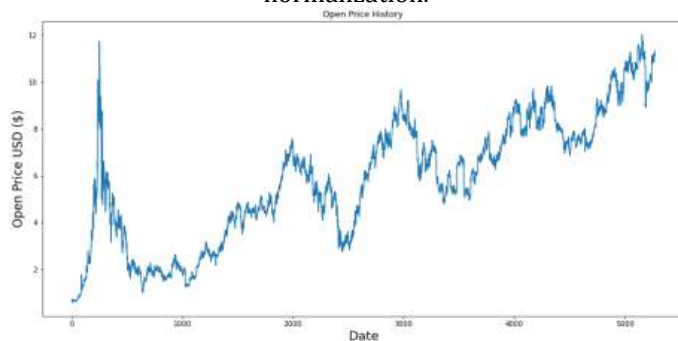


Fig. 3.1: Graph represents Dataset

Fig. 3.1 shows the plotting of training and testing dataset of Open value. This graph is plotted using matplotlib library.

4.3 Implementation and Result

LSTM is an artificial recurrent neural network(RNN) architecture used in the deep learning. LSTM has the feedback connections. It not only cant process single data points, but also entire sequences of data. For example, LSTM is applicable to tasks such as unsegment, handwriting recognition, speech recognition and anomaly detection in network traffics.

A common LSTM unit is composed of a cell, an input gate, an output gate and forgot gate. The cell remembers values over arbitrary time intervals and three gates regulate the flow of information into and out of the cell.

LSTM networks are well matched in identifying, processing and making forecasts based on time series data, since there can be lags of unknown duration between key events in a time series. LSTMs were developed to deal with the vanishing gradient problem that can be encountered when training traditional RNNs. Relative insensitivity to gap length is the precedence of LSTM over RNNs, hidden markov models and other sequence learning methods in various applications.

In this stage, LSTM model is used to predict the future value. The LSTM model can be implemented using tensorflow library. This model contains three layers namely input layer,

hidden layer and output layer (dense layer). In this shape of training data is the input shape of the model. Activation such as sigmoid and relu are used in the implementation of this model. The model has one output layer and predicts the future value based on train data and test data of the model.

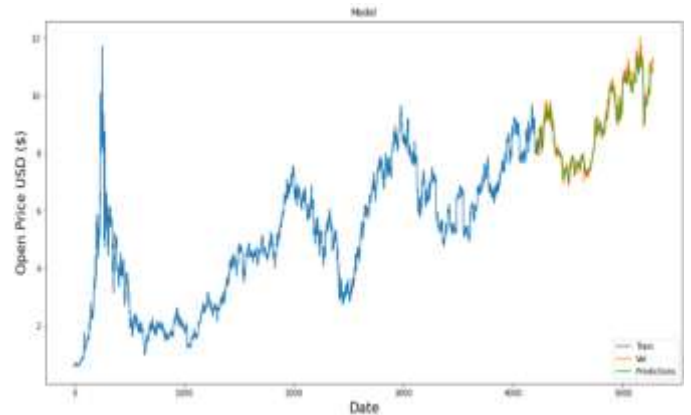


Fig. 3.2: Graph represents model prediction

Fig. 3.2 shows the result which is test data versus prediction data of Open attribute. This same model can be used to predict the other attributes of stock market.

Conclusion

Machine learning and artificial neural network are the tools for solving complex problems and in the creation of some great applications. Machine learning have found tremendous application and has evolved further into deep learning and neural networks, but the core idea is more or less the same for all of them. In the paper we have used LSTM which is part of the artificial neural network. The LSTM is highly dependent on the available dataset. The end result is dependent on the variation in the dataset available. The LSTM algorithm is highly capable of predictions with great accuracy rates. In the paper the dataset is accessed directly from yahoo finance. The normalized data is given as the input to the algorithm for prediction. This paper provides us with smooth understanding of machine learning and tensorflow.

References

- [1] Mrityunjay Sharma, "Survey on Stock Market Prediction and Performance Analysis", *International Journal of Advanced Research in Computer Engineering and Technology* Vol. 3, Issue 1, Jan 2014.
- [2] Mayankumar B Patel, Sunil R Yalamalle, "Stock Price Prediction Using Artificial Neural Network", *International Journal of Innovative Research in Science, Engineering and Technology* Vol. 3, Issue 6, June 2014.
- [3] Nirbhey Singh Pahwa, Vidhi Soni, Neeha Khalfay, Deepali Vora, "Stock Prediction Using Machine Learning a Review Paper", *International Journal of Computer Application* Vol.163, Issue 5, April 2017.

- [4] Dr A.Haritha, Dr PVS Lakshmi, G.Lakshmi, E.Revathi, A.G S Srinivas Deekshith, "Stock Market Prediction Using Arima Model", *International Journal of Computer Engineering and Applications Vol. 12, Issue 4*, April 2018.
- [5] V Kranthi Sai Reddy, "Stock Market Prediction Using Machine Learning", *International Research Journal of Engineering and Technology Vol.5, Issue 10*, Oct 2018.
- [6] K.Hiba Sadia, Aditya Sharma, Adarsh Paul, SarmisthaPadhi, Saurav Sanyal,"Stock Market Prediction Using Machine Learning Algorithms", *International Journal of Engineering and Advanced Technology Vol.8, Issue 4*, April 2019.