International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 04 Issue: 08 | August -2020 ISSN: 2582-3930

STOCK MARKET PREDICTION USING MACHINE LEARNING

[1] [2] [3] [4] [5]
Yashaswini S , Kushal N , Siri R Janige, Veena P , Shaikh Mohammad Nishad
[1] [2] [3] [4] [5]
Assistant professor , UG Student , UG Student , UG Student , UG Student .
[1][2][3][4][5]
Department of CSE, Cambridge institute of technology .

Abstract

Stock price prediction contributes to the development of effective strategies for stock exchange transactions. Predicting stock prices is an important objective in the financial world, since a reasonably accurate prediction has the possibility to yield high financial benefits and edge against market risks. The accessing of stock market data base is relatively easy due to internet and computation power, each operation has increased to perform in fraction of seconds. Stock prices predictor is a system that learns about the performance of a company and predicts future stock prices. The benefit of analyzing the stock is that it has shorter feedback cycles which make it easier to validate predictions. Historical prices is used to predict the future stock price using Linear Regression.

Key words- Stock market, close, LSTM model, linear regression, computation power and prediction.

Introduction

Stock is the part of ownership in a company. It provides a share of the company's in return for the capital invested. Stock market prediction is very important for the investor to invest huge amount of money on the right stock. In order to help these investors to make profit on their investments, stock market prediction was introduced. With the help of correct prediction it can give huge profit for both the investor and the seller. This also helps the further growth of the company. The main aim of this prediction is to predict the correct future value of the financial stock of the company. There is also a stock price prediction, were it learns about the performance of a company and predicts future stock price. Example: If you own a company and have 1000 shares, and I buy 1 share of the company, then I have $1/1000^{th}$ ownership of the company. Stock in simple words is just the product output of the company. Advantages are two techniques have been utilized in this paper: LSTM and Regression, on the Yahoo finance dataset. Both the techniques have shown an improvement in the accuracy of predictions, there by yielding positive results. Use of recently introduced machine learning techniques in the prediction of stocks have yielded promising results and thereby marked the use of them in. Profitable exchange scheme and more efficient.

Stock market is a place were buying and selling of stock takes place. Stock market have 2 types: Primary market and Secondary market. In Primary market, the company is allowed to sell or share the stock among common public people. Example: IPO (Initial Public Offering). In Secondary market, the investor's trade among themselves, (i.e. The Company is not involved). Example: If you buy a company stock from amazon, you are only dealing with the investors who owns the shares in amazon, the company is not directly involved in thetransaction.

The prediction uses the history of respective stock market, carefully analyze it and then predicts the future value. Therefore this increases the accuracy. By introducing Machine Learning into stock market prediction helped us giving more efficient and accurate measurements. Prediction of stock market is used from many years, but by seeing the results Machine Learning approach has given us the best results from all the other methods used sofar.

LITERATURE SURVEY

Many methods have been used in prediction, but Machine Learning technique gave us more accurate results so far. LSTM (Long Short Term Memory) is a type of recurrent neural network capable of learning order dependence in a sequence prediction problems [1]. This paper demonstrates the power of LSTM in stock market prediction in stocks, which is mechanical yet much more unpredictable. Compared with random prediction method, our LSTM model improved the accuracy of stock returns prediction from 14.3% to 27.2%. The historical data of stock market were transformed into 30-days-long sequences with 10 learning features and 3-day earning rate labeling. The model was fitted by training on 900000 sequences and tested using the other 311361 sequences.

Linear regression is another technique used in prediction methods. It helps us to compare or predicts a dependent variable value based on given independent variable. It is a regression technique which find out the linear relation between the dependent and the independent value obtained. The LSTM helps to predict value whereas the Regression technique helps us to give the accuracy [2]. This paper, proposes a novel method to predict the stock closing price based on the deep belief networks (DBNs) with intrinsicplasticity.

In the experiments, the stock in S&P 500 is used to examine the performance. The back propagation algorithm is used for output training to make minor adjustments of structure parameters. The intrinsic plasticity (IP) is also applied into the network to make it have adaptive ability. It is believed that IP learning for adaptive adjustment of neuronal response to external inputs is beneficial for maximizing the input-output mutual information. Our results show that the application of IP learning can remarkably improve the prediction performance. Moreover, the effects of two kinds of IP rules on the performance of prediction are examined. Compared with Trish's IP and without IP, DBN with Li's IP learning has much better prediction performance than the others. These results may have important implications on the modeling of neural network for complex time series prediction[3].

Proposed system

In this proposed system, we focus on predicting the stock values using machine learning algorithms like Linear Regression and LSTM model. We proposed the system "Stock Market Price Prediction" using the Linear Regression Algorithm. In this proposed system, we were able to train the machine from the various data points from the past to make a future prediction. We took data from the previous yearstocks

to train the model. We majorly used two machine-learning libraries to solve the problem. The first one was numpy, which was used to clean and manipulate the data, and getting it into a form ready for analysis. The other was scikit, which was used for real analysis and prediction.

System Architecture

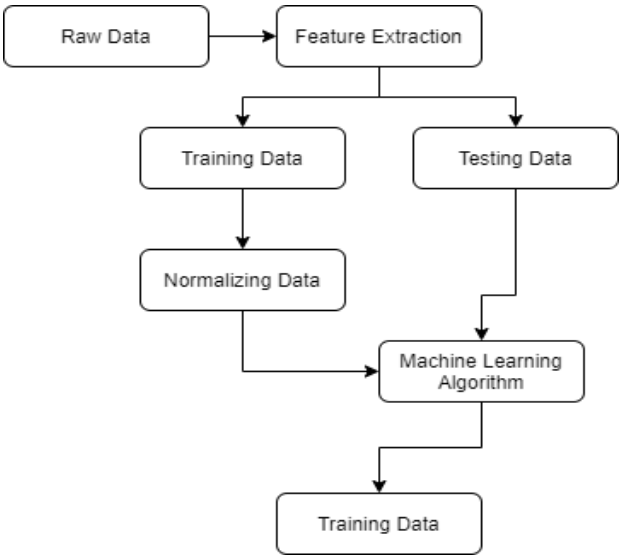

Fig: Flow Chart of System Architecture

The first step is the conversion of this raw data into processed data. This is done using feature extraction, since in the raw data collected there are multiple attributes but only a few of those attributes are useful for the purpose of prediction. So the first step is feature extraction, where the key attributes are extracted from the whole list of attributes available in the raw dataset. Feature extraction starts from an initial state of measured data and builds derived values or features. These features are intended to be informative and non-redundant, facilitating the subsequent learning and generalization steps. Feature extraction is a dimensionality reduction process, where the initial set of raw variables is diminished to progressively reasonable features for ease of management, while still precisely and totally depicting the first informational collection.

Methodology

The dataset we used was from the previous year's stock markets collected from the public database available online, 80 % of data was used to train the machine and the rest 20 % to test the data. The basic approach of the supervised learning model is

to learn the patterns and relationships in the data from the training set and then reproduce them for the test data. We used the python pandas library for data processing which combined different datasets into a data frame. The resulting data frame allowed us to prepare the data for feature extraction. The data frame features were date and the closing price for a particular day. We used all these features to train the machine on LSTM model and predicted the object variable, which is the price for a given day. We also quantified the accuracy by using the predictions for the test set and the actual values. The proposed system touches different areas of research including data pre-processing, Linear Regression, and so on. Regression and LSTM models are engaged for this conjecture separately. Regression involves minimizing error and LSTM contributes to remembering the data and results for the longrun.

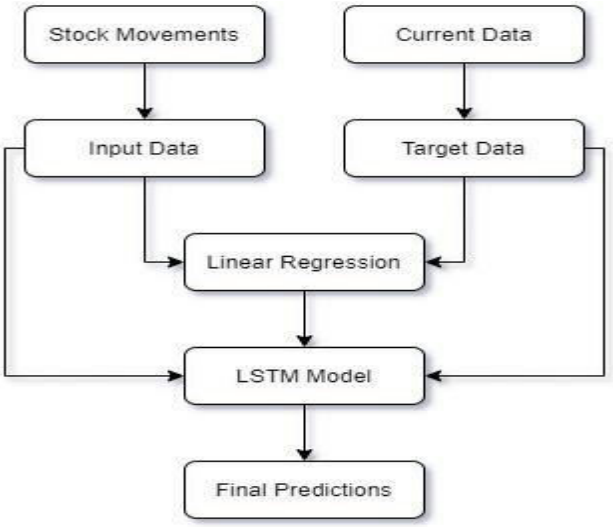

Fig: Flow Chart of the Proposed System

Linear Regression Based Model

Linear regression is a Supervised Learning algorithm which uses ensemble learning method for classification and regression. If the goal is prediction, or forecasting, or error reduction, linear regression can be used to fit a predictive model to an observed data set of y and X values. After developing such a model, if an additional value of X is then given without its accompanying value of y, the fitted model can be used to make a prediction of the value of y. Regression predicts a numerical value. Regression performs operations on a dataset where the target values have been defined already. And the result can be extended by adding new information model can be used to make a prediction of the value of y. Regression predicts a numerical value. Regression performs operations on a dataset where the target values have been defined already. And the result can be extended by adding new information. First, we divide the data into two parts of training and testing. Then we use thetraining section for starting analysis and defining the model. Regression uses a given linear function for predicting continuousvalues:

$$V = a + bK + error$$

Where, V is continuous value; K represents known independent values; and, a, b are coefficients.

LSTM Model

LSTM are very powerful in sequence prediction problems because they are able to store the past information. This is important in our case because the previous price of a stock is crucial in predicting its futureprice.

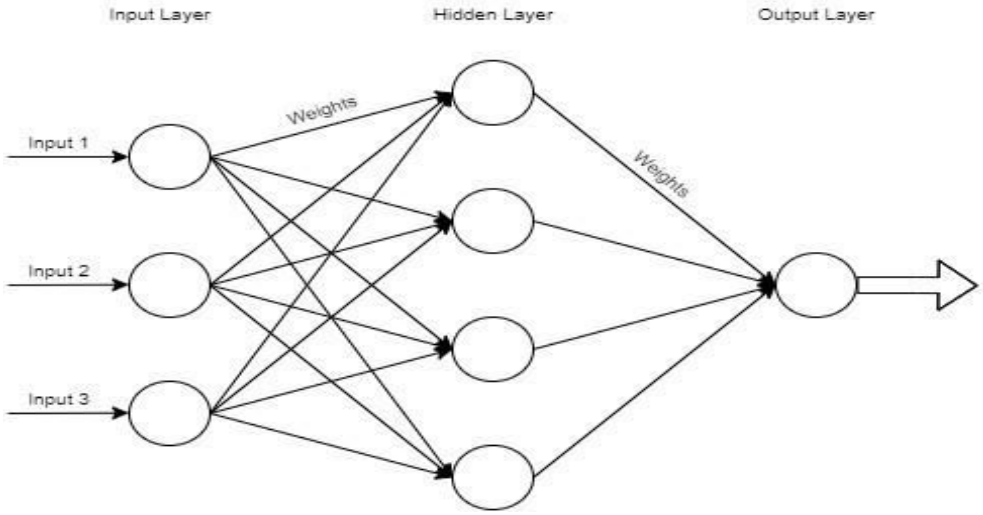

Fig: Schematic of a typical Artificial Neural Network (ANN) architecture

Long-Short-Term Memory Recurrent Neural Network belongs to the family of deep learning algorithms. It is a recurrent network because of the feedback connections in its architecture. It has an advantage over traditional neural networks due to its capability to process the entire sequence of data. Its architecture comprises the cell, input gate, output gate, and forget gate. The cell remembers values over arbitrary time intervals, and the three gates regulate the flow of information into and out of the cell. The cell of the model is responsible for keeping track of the dependencies between the elements in the input sequence. The input gate controls the extent to which a new value flows into the cell, the forget gate controls the extent to which a value remains in the cell, and the output gate controls the extent to which the value in the cell is used to compute the output activation of the LSTMunit.

Results

The proposed system is trained and tested over the dataset taken from Yahoo Finance. It is split into training and testing sets respectively and yields the following results upon passing through the different models:

Data Set (Closed Valve)

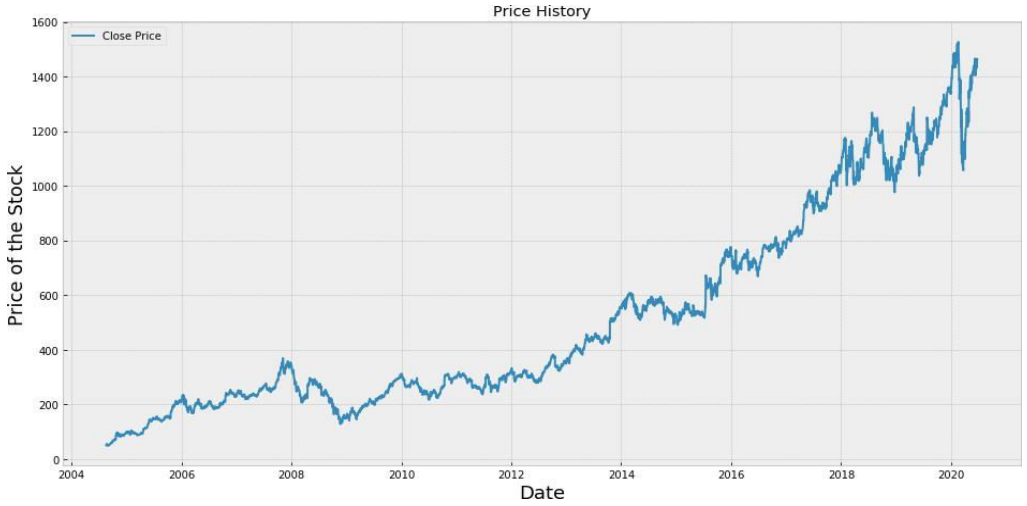

Fig 5.1: Plot between Price and Date

Regression Based Model Results

Fig 5.2: Plot between Price and Date Using Regression

The plot in figure5.1 and 5.2 is the result of application of linear regression algorithm on the dataset to predict varying prices with respect to the time.
The above graph from fig 5.2 is plot over the data having batch size 1 and epochs. The R-square confidence test resulted in a confidence score of0.86625.

LSTM Based Model Results

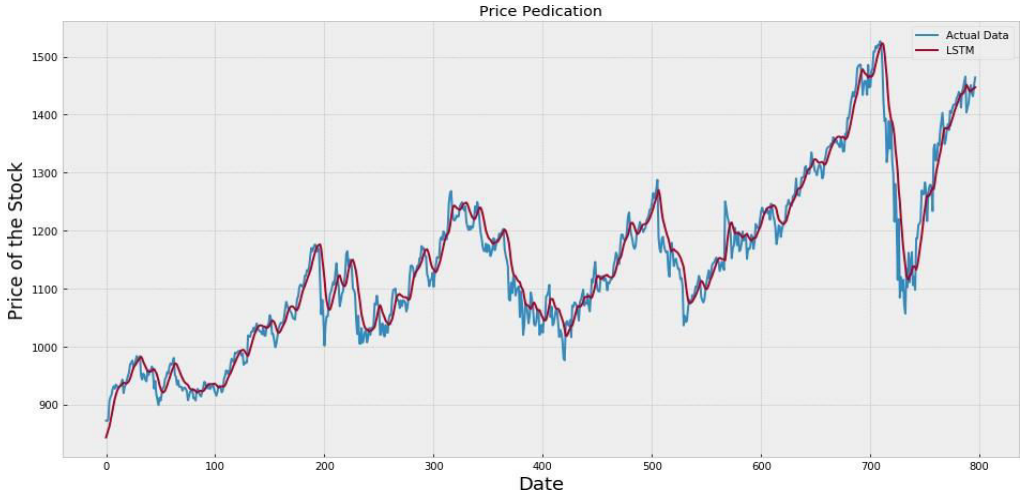

Fig 5.3: Plot between Price and Date Using LSTM

The above fig 5.3 shows Actual & Predicted Trend of LSTM. The prediction is shown by red line and the actual trend is shown by blue. The proximity of these two lines tells, how efficient the LSTM based model is. The prediction approximates real trend when a considerable amount of time has passed. The model resulted in accuracy matrix is 0.93. The more the system is trained and the greater the size of the dataset utilized the greater the accuracy which will be attained. The LSTM Model offered more accuracy than the Regression based Model.

Conclusion

In this paper, we study the use of LSTM and Linear Regression to predict financial movement direction. LSTM is a promising type of tool for financial forecasting. LSTM is superior to the other individual classification methods in forecasting daily movement direction. This is a clear message for financial forecasters and traders, which can lead to a capital gain. However, each method has its own strengths and weaknesses. We also observed that the choice of the indicator function can dramatically improve/reduce the accuracy of the prediction system. Also a particular Machine Learning Algorithm might be better suited to a particular type of stock, say Technology Stocks, whereas the same algorithm might give lower accuracies while predicting some other types of Stocks.

Future Enchantments

Two algorithms, LSTM and Linear Regression were used in this study and only one dataset from Yahoo Finance was applied to train and test the models. The system can only predict the direction (up/down) for the next trading day giving the accuracy of 80%. The model can be further improved using neural network or a hybrid model, consisting of LSTM for classifying the related data points. A model can be developed to train and test stock prices using deep learning technique ANN, the prediction of future stocks can be done by linear regression. However the accuracy may vary depending upon the dataset complexity.

References

[1] A. Sharma, D. Bhuriya and U. Singh, "Survey of stock market prediction using machine learning approach," 2017 International conference of Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, 2017, pp. 506-509, doi:10.1109/ICECA.2017.8212715.
[2] Y. E. Cakra and B. Distiawan Trisedya, "Stock price prediction using linear regression based on sentiment analysis," 2015 International Conference on Advanced Computer Science and Information Systems (ICACSIS), Depok, 2015, pp. 147-154, doi:10.1109/ICACSIS.2015.7415179.
[3] W. Chen, Y. Zhang, C. K. Yeo, C. T. Lau and B. S. Lee, "Stock market prediction using neural network through news on online social networks," 2017 International Smart Cities Conference (ISC2), Wuxi, 2017, pp. 1-6, doi: 10.1109/ISC2.2017.8090834.

- https://ieeexplore.ieee.org/document/8212715
- https://ieeexplore.ieee.org/document/7415179
- https://www.sciencedirect.com/science/article/pii/S1877050920304865
- https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0227222