International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 04 Issue: 05 | May -2020 ISSN: 2582-3930

Stock Price Predictions With ML Using Python

Puneet Singh[1], Prem Kumar Yadav[2], Shashank Chaurasia[3] ,Shubham Bhardwaj[4]

[1234] *Fourth Year, Department of Computer Science and Engineering, RKGIT, Ghaziabad, U.P.*

Abstract -In this paper we analyse existing method of stock market prediction. Different approaches at the problem are the applications of Machine Learning. It is seen that Machine Learning could be used to guide an investor's decisions. The stock market is a transformative, non-straight dynamical and complex system. Long term investment is one of the major investment decisions. This paper is to present comparison of machine learning aided algorithm to evaluate stock prices in the future to analyze market behavior. Stock market price data is generated in huge volume and it changes every second. The main objective is to find the best model to predict the value of the stock market. The project also presents a machine learning model to predict the longevity of stock in a competitive market. The successful prediction of the stock will be a great asset for the stock market institutions and will provide real-life solutions to the problems that stock investors face.

Key Words:linear regression, stock market, machine learning, , LSTM, support vector machine

1. INTRODUCTION

Predicting the Stock Market has been the bane and goal of investors since its existence. Everyday billions of dollars are traded on the exchange, and behind each dollar is an investor hoping to profit in one way or another. Entire companies rise and fall daily based on the behaviour of the market. Should an investor be able to accurately predict market movements, it offers a tantalizing promises of wealth and influence. It is no wonder then that the Stock Market and its associated challenges find their way into the public imagination every time it misbehaves. If there was a common theme among those productions, it was that few people knew how the market worked or reacted. Perhaps a better understanding of stock market prediction might help in the case of similar events in the future.

Exchanging the stocks on money markets is one of the significant speculation exercises. Already, scientists developed different stock examination system that could empower them to envision the bearings of stock esteem development. Predicting and foreseeing of significant worth future cost, in perspective of the present cash related information and news, is of colossal use to the financial pros. Financial masters need to know whether some stock will get higher or lower over particular time-period. To obtain the accurate output, the approach used is to implemented is machine learning along with supervised learning algorithms. Results are tested using different types of supervised learning algorithms with a different set of a features.[8]

2. PROPOSED WORK

In this proposed system, we focus on predicting the stock values using machine learning algorithms like Linear Regression and Moving Averages. We proposed the system "Stock market price prediction" we have predicted the stock market price using the LSTM 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 year stocks to train the model. We majorly used 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 sci-kit, which was used for visualizations, real analysis and prediction. The data set we used was from the previous years 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 tuned up dataframe allowed us to prepare the datafor feature extraction. The dataframe features were date and the closing price for a particular day. We used all these features to train the machine on random forest 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, random forest, and so on.

3. METHODOLOGY

3.1 Linear regression

Before knowing what is linear regression, let us get ourselves accustomed to regression. Regression is a method of modelling a target value based on independent predictors. This method is mostly used for forecasting and finding out cause and effect relationship between variables. Regression techniques mostly differ based on the number of independent variables and the type of relationship between the independentanddependent variables.[5]

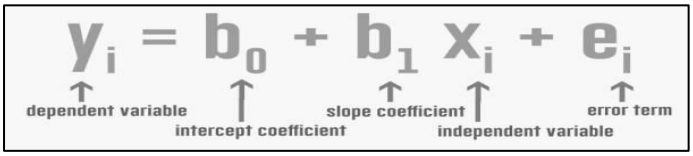

Fig-1: relationship between the independentanddependent variables.

Simple linear regression is a type of regression analysis where the number of independent variables is one and there is a linear relationship between the independent(x) and dependent(y) variable. The redline in the below graph is referred to as the best fit straight line. Based on the given data points, we try to plot a line that models the points the best. The line can be modelled based on the linear equation shownbelow.

The motive of the linear regression algorithm is to find the best values for a_0 and a_1.

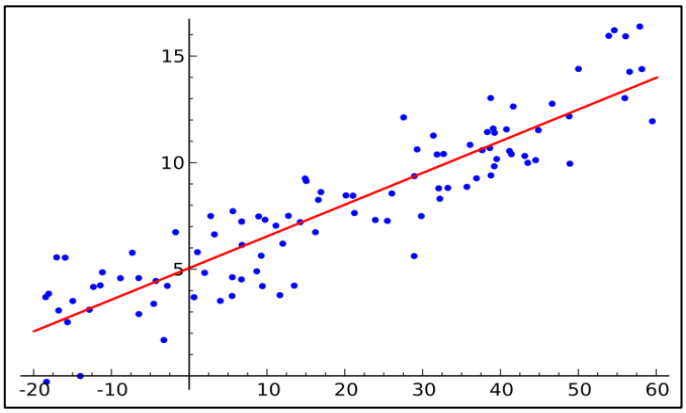

Fig-2: linear regression graph

3.2 Moving Averages

Moving averages are used and discussed quite commonly by technical analysts and traders alike. If you've never heard of a moving average, it is likely you have at least seen one in practice. A moving average can help an analyst filter noise and create a smooth curve from an otherwise noisy curve. It is important to note moving averages lag because they are based on historical data, not current price.The most commonly used Moving Averages (MAs) are the simple and exponential moving average.The Simple Moving Average formula is a very basic arithmetic mean over the number of periods.

$$SMA = \frac{A_1 + A_2 + \ldots + A_n}{n}$$

Fig-3: Simple Moving Average formula

The Exponential Moving Average (EMA) is a wee bit more involved. First, you should find the SMA. Second, calculate the smoothing factor. Then, use your smoothing factor with the previous EMA to find a new value. In this way, the latest

prices are given higher weights, whereas the SMA assignsequal weight to all periods. We'll see this clearer in our graphs below.

$$EMA_{Today} = (Value_{Today} * \left(\frac{Smoothing}{1 + Days}\right)) + EMA_{Yesterday} * (1 - \left(\frac{Smoothing}{1+Days}\right))$$

Fig-4:Exponential Moving Average (EMA)

Certain periods on a moving average are widely used. Many technical traders and market participants will cite the 10, 20, 50, 100, or 200 day moving averages. It all depends on preference or desired granularity. Breaks above and below the moving average are important signals and trigger active traders and algorithms to execute trades depending on if the break is above or below the moving average.[6]

3.3 LSTM

A Long short-term memory (LSTM) is a type of Recurrent Neural Network specially designed to prevent the neural network output for a given input from either decaying or exploding as it cycles through the feedback loops. The feedback loops are what allow recurrent networks to be better at pattern recognition than other neural networks. Memory of past input is critical for solving sequence learning tasks and Long short-term memory networks provide better performance compared to other RNN architectures by alleviating what is called the vanishing gradient problem.

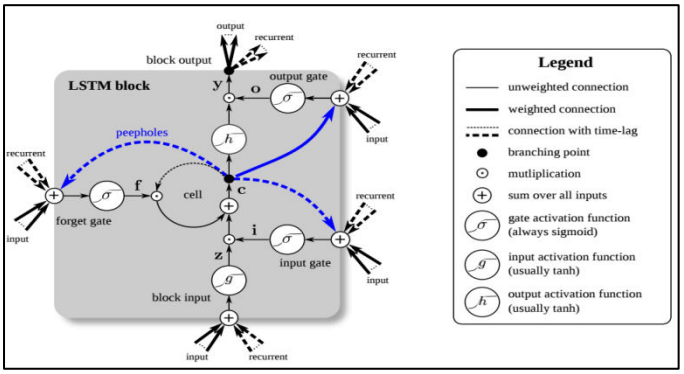

Fig-5: LSTM block

The Long Short-Term Memory Architecture consists of linear units with a self-connection having a constant weight of 1.0. This allows a value (forward pass) or gradient (backward pass) that flows into this self-recurrent unit to be preserved and subsequently retrieved at the required time step. With the unit multiplier, the output or error of the previous time step is the same as the output for the next time step. This self-recurrent unit, the memory cell, is capable of storing information which lies dozen of time-steps in the past. This is very powerful for many tasks.[4]

4. RESULT AND CONCLUSION

The stock market prediction has become an increasingly important issue in the present time. Generally, investments are made using predictions which are obtained from historical stock price after considering all the factors that might affect it. By studying the different algorithms, it is found that different companies stock prices varies randomly. Algorithm used for one company stocks to predict future prices with great accuracy may not be accurate for stock of other company. In other words, the output varies for each technique even if the same data set is being applied. The stock market predictions process is filled with uncertainty and can be influenced by multiple factors. Therefore, the stock market plays an important role in business and finance. Technical analysis is done using by applying machine learning algorithm on historical stock price data. Data like previous year stock prices are also considered and a prediction is made using these points.

On the comparative study of the stock market on different historical stock data. The algorithms and models used to predict the stock values wereanalyzed . A number of algorithms have been used to predict prices which gives promising result too. It is seen that LSTM model is the best among its competitors. Linear regression model is also good for predictions and is faster than the LSTM model but the accuracy is less than the LSTM model. SVM model is seen good for small data-sets but for big data-sets the model is seen to struggle maintaining the accuracy. Moving average model is dependent on the last values of the dataset, based on that the result is generated. So different model are found to be good in

different circumstances and they all have some pros and cons of there own. The use of different algorithms will purely depend upon the situation and person.

REFERENCES

[1] H. Ince, Support Vector Machine for Regression and Applications to Financial Forecasting, *Ieee-Inns-Enns International Joint Conference on Neural Networks IEEE Computer Society, vol6(2000), pp. 348-353.*

[2] Akita, Ryo, et al. "Deep learning for stock prediction using numerical and textual information." *IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS). IEEE*, 2016.

[3] Sak, Ha¸sim, Andrew Senior, and Françoise Beaufays. "Long short-term memory recurrent neural network architectures for large scale acoustic modeling." *Fifteenth annual conference of the international speech communication association*, 2014.

[4] Gers, Felix A., Nicol N. Schraudolph, and Jürgen Schmidhuber. "Learning precise timing with LSTM recurrent networks." *Journal of Machine Learning Research* 3.Aug (2002): 115-143.

[5] M. P. Mali, H. Karchalkar, A. Jain, A. Singh, and V. Kumar, "Open Price Prediction of Stock Market using Regression Analysis," *Ijarcce, vol. 6, no. 5, pp. 418–421, 2017.*

[6] P. Li, C. Jing, T. Liang, M. Liu, Z. Chen, and L. Guo, "Autoregressive moving average modeling in the financial sector," *ICITACEE 2015 - 2nd Int. Conf. Inf. Technol. Comput. Electr. Eng. Green Technol. Strength. Inf. Technol. Electr. Comput. Eng. Implementation, Proc., no. 4, pp. 68–71, 2016.*

[7] H. L. Siew and M. J. Nordin, "Regression techniques for the prediction of stock price trend," *ICSSBE 2012 - Proceedings, 2012 Int. Conf. Stat. Sci. Bus. Eng. "Empowering Decis. Mak. with Stat. Sci., pp. 99–103, 2012.*

[8] A. Sharma, D. Bhuriya, and U. Singh, "Survey of Stock Market Prediction Using Machine Learning Approach," *pp. 506–509, 2017.*

[9] Hakob GRIGORYAN, "A Stock Market Prediction Method Based on Support Vector Machines (SVM) and Independent Component Analysis (ICA)", DSJ 2016.

[10] Farhad Soleimanian Gharehchopogh,Tahmineh Haddadi Bonab and Seyyed Reza Khaze(2013),*A linear regression approach to prediction of stock market trading volume :a case study*,http://airccse.org/journal/mvsc/papers/4313ijmvsc03.pdf.