

Stock Price Prediction Using Machine Learning

Nalawade Sakshi¹, Mule Sarita², Adsare Monika³ Gunjal Shubhangi⁴

¹Nalawade Sakshi Computer Engineering & Jaihind College of Engineering Kuran

²Mule Sarita Computer Engineering & Jaihind College of Engineering Kuran

³Adsare Monika Computer Engineering & Jaihind College of Engineering Kuran

⁴Gunjal Shubhangi Computer Engineering & Jaihind College of Engineering Kuran

Abstract -The stock market or inventory marketplace is one of the most complex and complex ways of doing commercial enterprise. Small belongings, brokerage corporations, the banking region, all rely on this equal body for the distribution of profits and risk; a very complicated version however, this document proposes to use a machine learning set of rules to expect the destiny price of shared assets for change the use of open supply libraries and pre-existing algorithms to assist make this business layout from unpredictable to predictable. We are able to see how this simple software will bring acceptable outcomes. The end result is primarily based completely on numbers and assumes a large variety of axioms which can or might not follow inside the actual world as the time of prediction.

Key Words: basic concepts, data analysis, fundamental, implementation, linear regression, stock market, supervised machine learning

1.INTRODUCTION

The stock market is one of the oldest techniques by which a ordinary man or woman trades stocks, makes investments and earns a few money from corporations that promote part of themselves in this platform. This device seems to be a ability investment scheme if finished accurately. but, the rate and liquidity of this platform are fairly unpredictable and that is wherein we convey the generation to help us. machine mastering is one of those equipment that facilitates us get what we want. The following three paragraphs will briefly provide an explanation for the important thing components of this document: The inventory market, as we recognize, is a very

critical buying and selling platform that affects anybody in my opinion and nationally[2]. The basic precept is pretty simple, corporations will listing their shares in companies as small commodities known as stocks. They do this to elevate cash for the employer. A enterprise lists its stocks at a price referred to as an IPO or preliminary public presenting. that is the offer rate at which the business enterprise sells the shares and raises cash. After that these shares are owned by way of the owner and you may promote them at any fee to a consumer on a stock trade inclusive of the BSE or the Bombay inventory exchange. investors and buyers preserve to sell those stocks at their very own rate, however the company can simplest hold the cash earned throughout the IPO. the ongoing wish of the hare apart other than figuring out more profits, it translates into a selected boom in the share price after every profitable transaction. but, if the business enterprise problems greater stocks at a lower preliminary public providing, the marketplace charge for the exchange falls and investors go through a loss. This precise phenomenon is the purpose for the concern that humans have

while investing in stock markets and the cause for the fall and upward push in stock charges in a nutshell. Now, if we try to plan a chart of the inventory market fee over the term (say 6 months), is it clearly difficult to expect the following final results at the chart?

2.PREDICTION MODEL

A. Data Analysis Stage

www.quandl.com, a international-class dataset shipping platform. The dataset taken is for GOOGL from WIKI and can be extracted from quandl the usage of the "WIKI/GOOGL" token. We extracted and used about 14 years of statistics.in

addition, the extracted features are very particular to the theme used and could actually vary from subject matter to subject. Generalization is possible if, and handiest if, the data of the alternative challenge are collected with the identical consistency as the preceding problem.

B. Training and testing stage

Schooling and trying out phase on this section, we can use what we have extracted from our facts and applied in our machine getting to know model. We are able to implement the SciPy, Scikit-learn and Matplotlib libraries in Python to program our model, train them with the capabilities and labels we have extracted after which check them with the equal facts. First we are able to technique the facts to make the statistics that consists of: Moved values of the label attribute of the share that you want to predict.

right here, in our case, the classifier sees the functions and certainly looks at their label and recalls it. don't forget the mixture of functions and their respective label which in our case is the stock fee a few days later. Then pass ahead and find out which model is accompanied with the aid of the capabilities to produce the respective label. that is how supervised device gaining knowledge of works [10]. For checks in supervised

machine mastering, we positioned a combination of capabilities into the skilled classifier and check the output of the classifier with the real tag. This facilitates us decide the accuracy of our classifier. Which is very important for our version. A classifier with an accuracy of much less than ninety five% is practically useless. Accuracy is a completely essential element in a system mastering model. it's miles essential to understand what accuracy means and the way to increase its accuracy within the subsequent subtopic.

3. Result

Once the model is prepared, we use the template to get the favored effects in any manner we want. In our case, we are able to plot a graph of our consequences (fig. 1) in line with our necessities that we mention eds in advance on this record .It must be according to our wishes, and as stated above, a version with an accuracy of less than ninety

five% is nearly vain. There are some popular strategies for calculating

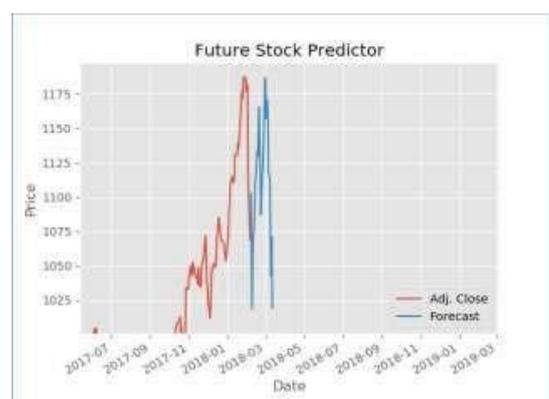


Fig. 1. Graph showing stock price of GOOGL from year 2005 till July 2018. Red is the line representing given data and blue is representing the forecasted or the predicted value of stock .

Accuracy is the issue that every device getting to know developer is usually dedicated to contributing to. After the model has been advanced, there may be an limitless attempt to optimize the version for more and more correct consequences. There are some very commonplace and simple methods to increase the performance of the version and they had been mentioned above. however, let's examine a number of the standard ways to optimize a gadget studying set of rules:

- Unrestricted optimization
 - o first rate gradient
 - o Newton's approach
 - o Batch mastering
 - o respectable stochastic gradient
- restricted optimization
 - o Lagrange duality
 - o SVM in number one and double bureaucracy
 - o confined techniques

maximum gadget gaining knowledge of issues are, in the end,



optimization issues, wherein we decrease a feature issue to some restrictions.

4. SOME COMMON MISTAKES

We mention some of the common mistakes made by professionals in this field, which you should avoid[12]:

- Poor annotation of training and test datasets
- Poor understanding of algorithm assumptions
- Poor understanding of algorithm parameters
- Lack of understanding of the goal
- Do not understand the data
- Prevent leaks (features, information)
- Insufficient data to train the classifier
- Use machine learning where you don't need it

5. CONCLUSIONS

Machine learning as we have seen so far, is a very powerful tool and therefore avoidable, it has a great application. So far we have seen that machine learning relies heavily on data. Therefore, it is important to understand that data is quite valuable, and as simple as it may seem, analyzing the data is not an easy task. Machine learning has found an amazing application and has evolved further into deep learning and neural networks, but the basic idea is pretty much the same for all of them. This document provides a fluid view of how to implement machine learning. There are various ways, methods, and techniques available to manage and solve various problems, in different imaginable situations. This document is limited to supervised machine learning only and seeks to explain only the fundamentals of this complex process.

ACKNOWLEDGEMENT

This article is written on the basis of the author's summer project during his degree course in 2018 after the valuable guidance of Neha Agarwal (Asst. Prof.) AUUP, Noida. The author of this article, does not claim rights to any of the algorithms, codes, data, formulas used, definitions, problem solving approach, as its property.

REFERENCES

- [1] Andrew McCallum, Kamal Nigam, Jason Rennie, Kristie Seymore "A Machine learning approach to Building domain-specific Search engine", IJCAI, 1999 - Citeseer
- [2] Yadav, Sameer. (2017). STOCK MARKET VOLATILITY - A STUDY OF INDIAN STOCK MARKET. Global Journal for Research Analysis. 6. 629-632.
- [3] Montgomery, D.C., Peck, E.A. and Vining, G.G., 2012. Introduction to linear regression analysis (Vol. 821). John Wiley & Sons.
- [4] Draper, N.R.; Smith, H. (1998). Applied Regression Analysis (3rd ed.). John Wiley. ISBN0-471-17082-8.
- [5] Robert S. Pindyck and Daniel L. Rubinfeld (1998, 4th ed.). Econometric Models and Economic Forecasts
- [6] "Linear Regression", 1997-1998, Yale University <http://www.stat.yale.edu/Courses/1997-98/101/linreg.htm>
- [7] Agarwal (July 14, 2017). "Introduction to the Stock Market". Intelligent Economist. Retrieved December 18, 2017.
- [8] Jason Brownlee, March 2016, "Linear Regression for machine learning", Machine learning mastery, viewed on December 2018, <https://machinelearningmastery.com/linearregression-for-machine-learning/>
- [9] Google Developers, Oct 2018, "Descending into ML: Linear Regression", Google LLC, <https://developers.google.com/machinelearning/crash-course/descending-into-ml/linearregression>
- [10] Fiess, N.M. and MacDonald, R., 2002. Towards the fundamentals of technical analysis: analysing the information content of High, Low and Close prices. Economic Modelling, 19(3), pp.353-374.