iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 8 **Issue:** VII **Month of publication:** July 2020

DOI: https://doi.org/10.22214/ijraset.2020.30744

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue VII July 2020- Available at www.ijraset.com

Stock Market Prediction using Logistic Regression Analysis - A Pilot Study

Shreenidhi Sriram[1], Sanjana Rangarajan[2]
[1]*Student, Department of Computer Science & Engineering, College of Engineering, Anna University, Guindy, Chennai, India.*
[2]*Student, Department of Information Technology, College of Engineering, Anna University, Guindy, Chennai, India.*

Abstract: ***Stock market is a major factor that plays a vital role in deciding a Nation's economy. With the increase in use of technology in stock trading, the volatility in stock prices also increases. The aim of Stock Market Prediction is to forecast the future values of the financial stocks of a company. The recent trend in stock market prediction technologies is the use of machine learning which makes predictions based on the current values in stock market indices by training on their previous values. Machine learning itself employs different models to make prediction easier and authentic. This study focuses on use of logistic regression based machine learning to predict stock values. Factors considered are open, close, low, high, last, total trade quantity and turnover. The successful prediction of stock will be the great asset for the stock market institutions and will provide real-life solutions to the problems that stock investors face.***
Keywords: stock market, logistic regression, prediction, machine learning, analysis

I. INTRODUCTION

Of the various factors that decide the economy of a country, stock market plays a pivotal role. It also serves as a great opportunity for the investors and various companies to make an investment and enable them to grow many folds [1]. While the older trend was to stay invested in a company for long years, the recent trend is the intraday market and short term positional holdings, where the traders aim to make money without much investment. Such traders use various tools to predict the stock market behaviour, which facilitates them to make a decision to either hold or sell the stock that they possess. Various factors that influence the stock price, including the social media, financial news and operator gaming might influence the growth or fall of any stock, making the stock prices much more volatile. Such uncertainties in stock prices desist people from making investments in stocks. Thus there is a growing need to accurately predict the stock price movement which can be used in a real-life scenario to plan either further accumulation of a stock or make an early exit and minimize loss.

Machine learning is an ever-expanding technology that focuses on involving artificial intelligence to empower the system and also to learn and make progress from previous trials without the need to be programmed repeatedly [2]. One can see a linear type of curve, especially when one observes the stock price movement over a long period of time. People have an inherent tendency to buy stocks when the stock price is expected to rise in future. There are various methods to accurately predict stock market price movement. This includes a time series forecasting along with technical analysis, modelling, machine learning and prediction of variable stock market. The datasets that are widely used in the prediction model include opening price, closing price, data and other variables that are needed to predict the object variable which is the price on any given day. Stock market forecasting is the act of demanding to conclude the future price of a company stock or other financial instrument traded on an exchange. The successful forecast of a stock's future value might help accrue profit. The efficient-market hypothesis suggests that stock prices reveal all currently existing information and any price changes that are not based on newly exposed information thus are intrinsically unpredictable [3, 4]. Stock price prediction is possible by Machine learning algorithms. Artificial intelligence forms the major component of machine learning that facilitates to empower the system. The future prices of the stock are predicted by Regression model [5].

II. PROBLEM STATEMENT

It is too intricate and complicated to comprehend the problem statement related to stock market. Stock prices are so volatile that one cannot predict the progression in stock prices with that much of accuracy. There are always two groups of people, with one group feeling that stock market price progression can be accurately predicted and the other group that feels that it is so very highly unpredictable and volatile. The quality of the stock alone may not be the sole criterion to decide on buying or selling the shares of that company. It is all the more important to assess the correct levels at which a buying should be made. This pilot study aims at assessment of methods of accurate prediction of stock price variations that would help retailers and institutions in deciding on buying or selling a stock.

©IJRASET: All Rights are Reserved

III. PROCESS INVOLVED

The feature is extracted manually with the raw data from the dataset. After this process, the data obtained is trained and the data is tested simultaneously. Finally, from the trained data result is obtained from the testing phase. The output of the desired model is obtained which is the predicted value of the particular model.

IV. OBJECTIVE

The objectives of our study are to identify the factors that affect the share market, to generate the patterns from large set of data of stock market for prediction, to predict an approximate value of share price and to provide analysis for users through logistic regression models.

V. REVIEW OF LITERATURE

A. *Survey Of Stock Market Prediction*

The stock market prediction has become an increasingly important issue in the present time. One of the methods employed is technical analysis, but such methods do not always yield accurate results. So it is important to develop methods for a more accurate prediction. The technique that was employed in this instance was a regression. Since financial stock marks generate enormous amounts of data at any given time a great volume of data needs to undergo analysis before a prediction can be made. One of the noteworthy techniques that were mentioned was logistic regression. The way logistic regression models work is that they are often fitted using the least squares approach, but they may alternatively be also be fitted in other ways, such as by diminishing the "lack of fit" in some other norm, or by diminishing a handicapped version of the least squares loss function [6].

B. *Review On Machine Learning Techniques For Stock Market Prediction*

Machine learning techniques intend to consequently learn and perceive patterns in huge information. There are large numbers of well-known machine learning algorithms that can be utilized to categorize an issue given a set of peculiarities. In this area some of these algorithms that are especially utilized as a part of classifying stock market information into "up" or "down" periods given a set of inputs originated through macro-economic information and technical analysis has been presented [7].

C. *Approach For Logistic Regression*

Lv et al. [2010] used LR as a comparative method in order to build a better model for predicting stock returns effectively and efficiently [8]. A 30 times holdout method was used in the assessment, along with the two commonly used methods in the top 10 data mining algorithms (the support vector machine and k nearest neighbour) and the two baseline benchmark methods from the statistical area (MDA and LR).

VI. METHODOLOGIES

In our study, we studied the efficacy of Logistic regression model in predicting the stock market price movements. The programming language used is Python to predict the movements in the stock market price. Accuracy plays the major role in stock market prediction. Although many algorithms are available for this purpose, selecting the most accurate one continues to be the fundamental task for getting best results. The algorithm used in our study is Logistic Regression analysis. This involves training and execution of algorithms and getting the desired outputs.

Regression analysis is used to determine the magnitude of relationships between variables as well as to model relationships between variables and for predictions based on the models. Regression includes performing operations on the dataset where target values are pre-defined. The relations which regression establishes between predictor and target values can make a pattern. This pattern can be used on other datasets which their target values are not known. Thus, the data needed for regression are two part, first section for defining model and the other for testing model. First, we divide the data into two parts of training and testing. We use training section for starting the analysis and the testing section for prediction of the desired outputs.

A. *Logistic Regression*

Logistic Regression (LR) is helpful in predicting the presence or absence of a characteristic or outcome, based on values of a set of predictor variables. It is a multivariate analysis model. LR has the advantage of being less affected than discriminant analysis when the normality of the variable cannot be assumed. It has the capacity to analyse a mix of all types of predictors. LR allows prediction of a distinct outcome, such as group membership from a set of variables that may be discrete, continuous or dichotomous. In instances where the independent variables are categorical, or a mix of continuous and categorical, logistic regression is preferred.

©IJRASET: All Rights are Reserved

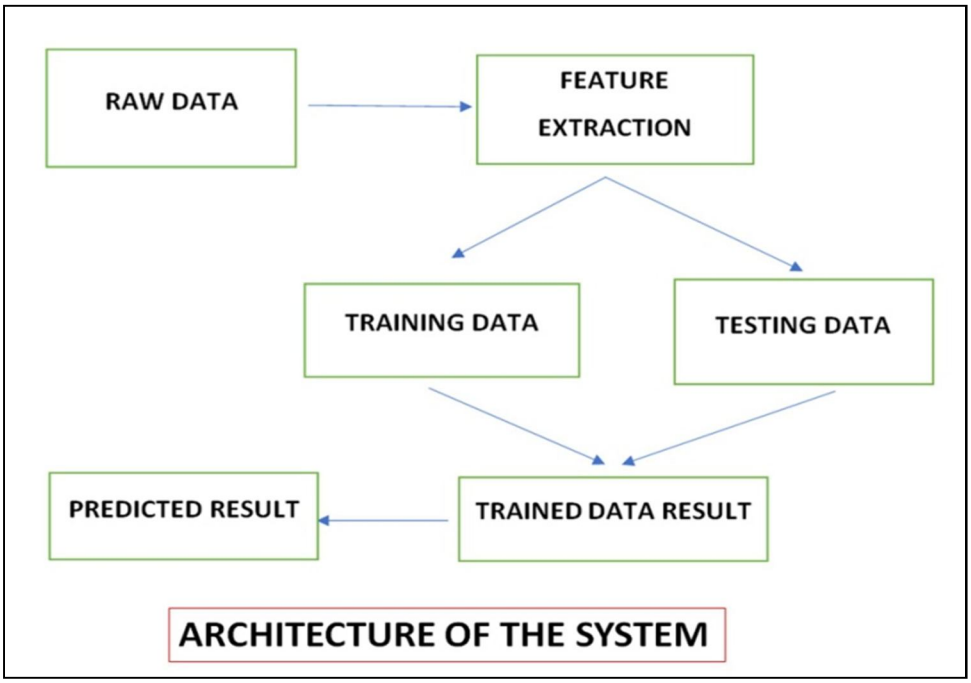
LR uses a more complex cost function and this cost function can be defined as the '**Sigmoid function**' or also known as the 'Logistic function'. The LR model extends the linear regression model by linking the range of real numbers to the 0-1 range. The programming language used was Python and the tools used were Python IDE, Anaconda and Jupiter Notebook.

Dataset is taken from the Kaggle website, Sample data obtained from Tata Company under Tata Global Beverages. Though it includes stocks from different industries has some distinguishing characteristics such as sensitivity, predictability, scalability, to name a few. We have used all the indices for future calculations. A sample of data obtained is shown in the Table 1. Date, Open, High, Low, Last, Close, Total Trade Quantity and Turnover.

Table 1 illustrates the sample data for our study.

DATE	OPEN	HIGH	LOW	LAST	CLOSE	TTQTY	TURNOVER IN LAKHS
08-10-18	208.00	222.25	206.85	216.00	215	4642146	10062.83
05-10-18	217	218.5	205.5	210.25	209.05	3519515	7407.06
04-10-18	223.3	227	216	217	218	1728786	3815.79
01-10-18	234	234.5	221.8	230	230	1534749	3486.50
03-10-18	230	237	225	226	227.6	1708590	3960.27
27-9-18	234	236	231	233	233	5082859	11859.95
26-9-18	240	240	232.5	235	234.5	2240909	5248.6
25-9-18	233	236	232	236	236	2349368	5503.90
24-9-18	233.50	239.20	230	234	233	3423509	7999.55

Fig 1 describes the architecture of the system studies.



©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue VII July 2020- Available at www.ijraset.com

VII. RESULTS

Logistic Regression model Is used to find the probability of how much chance there is for cases such that the event is either success or failure. LR can be implemented when the dependent variable is binary in nature, that is it can have at most two values. Logistic regression is quite similar to Linear regression. The only difference arises here is the sigmoid/logistic function is used.
LOGISTIC FUNCTION: $y=1/1+e^{-x}$ is used in regression.
The Scattered plot graph is obtained by taking Low(x axis) and High(Y axis).

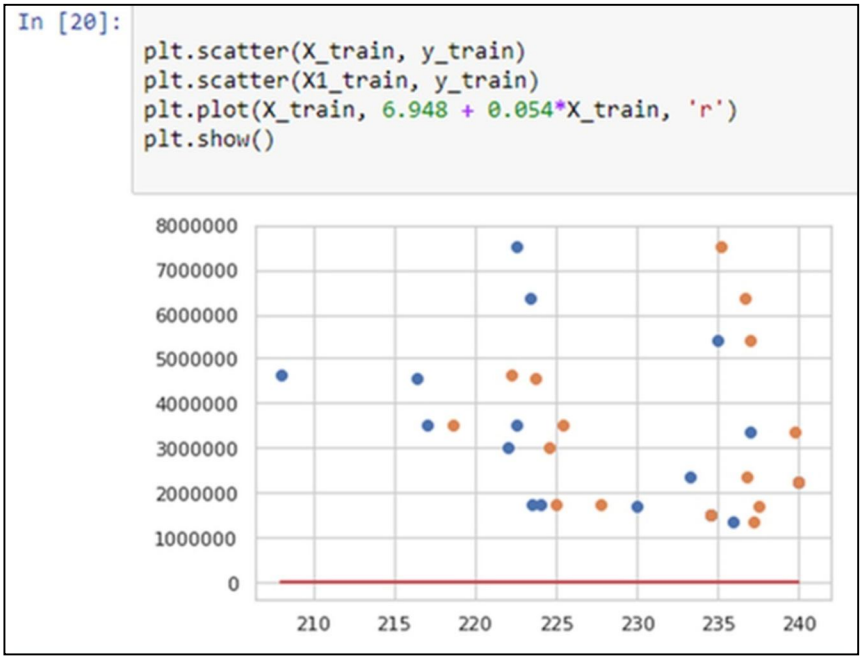
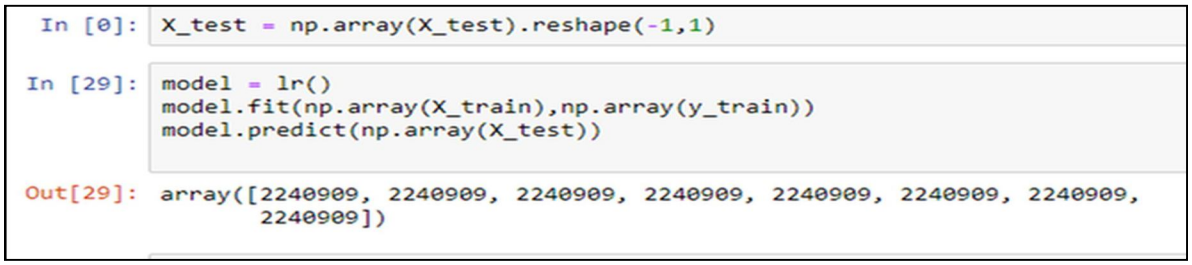


Fig 2 illustrates the scattered plot graph using Regression. After training and testing process, the model returns predicted value between Low and High value which gives turnover (predicted value). Fig 3 illustrates the predicted value of Logistic Regression.



The above image shows the predicted value of stock market using Logistic Regression.

VIII. CONCLUSION AND FUTUREWORK

A. Conclusion

By measuring the accuracy of the different algorithms, we found that the logistic regression algorithm is ideal for predicting the market price of a stock based on various data points from the historical data. The algorithm will be a great asset for brokers and investors for investing money in the stock market since it is trained on a huge collection of historical data and has been chosen after being tested on a sample data.

B. Future Work

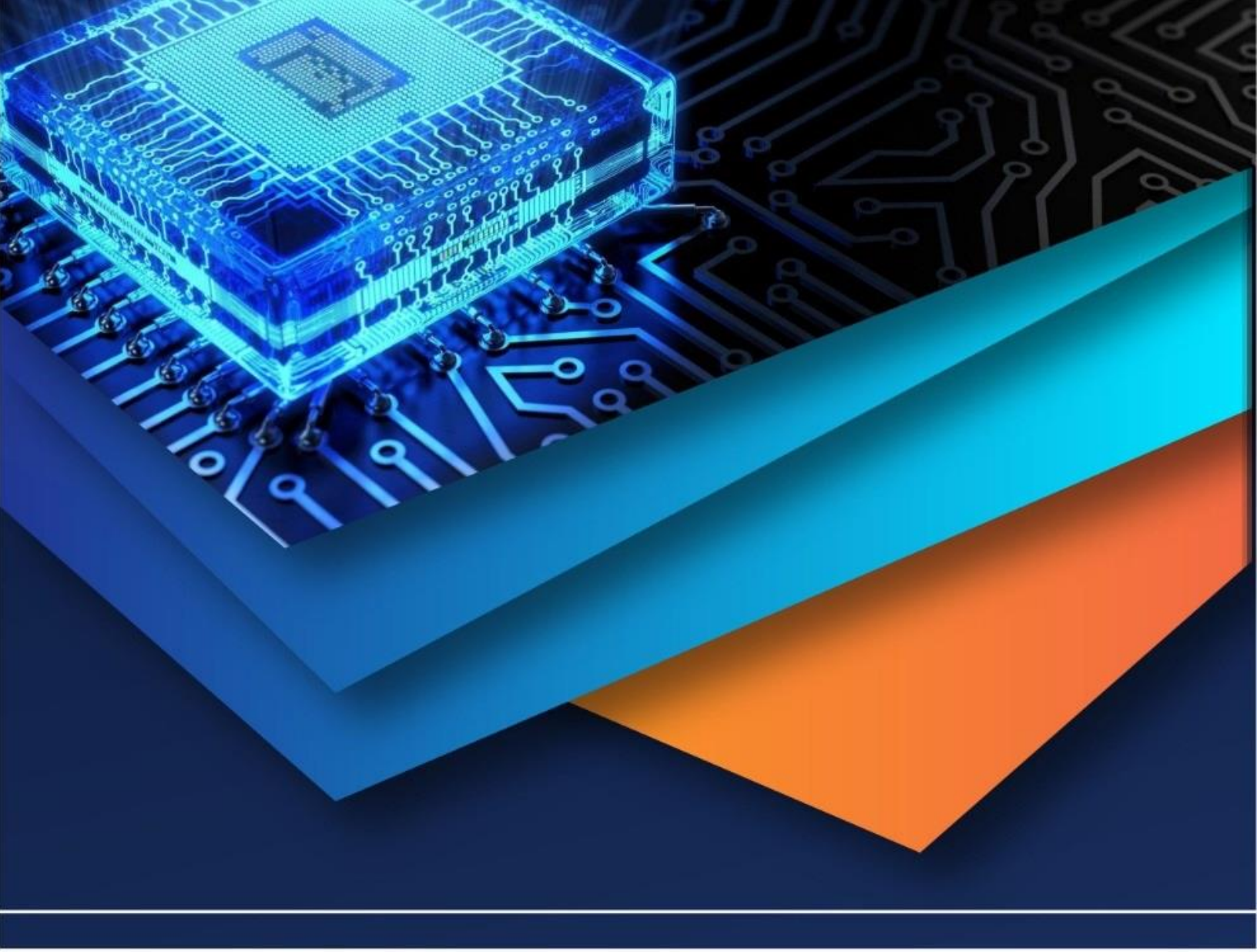
Our future work aims at improvements in prediction of stock market price movements by utilizing a much bigger dataset than the one being utilized currently. This would immensely help to increase the accuracy of our prediction models. Furthermore, further studies are underway in various other models like Linear regression models and Support vector machines (SVM) to check for the accuracy rate and identify the best model for an accurate stock market prediction.

©IJRASET: All Rights are Reserved

REFERENCES

[1] Ondabu, Ibrahim. (2015). Stock Market Investment Incentives: A Gift or a Motivator? Evidence from Literature. Business, Management and Economics Research. Vol. 1, No. 5, p: 54-62.
[2] Velmurugan. S, Ramesh. G P. "Application of Information Science and Technology in Academic Libraries: An Overview", Recent Trends and Advances in Artificial Intelligence and Internet of Things, Springer, Cham, Intelligent Systems Reference Library, Vol. 172, p. 75-83, 2020.
[3] Ashish Sharma, Dinesh Bhuriya, Upendra Singh. "Survey of Stock Market Prediction Using Machine Learning Approach", ICECA 2017.
[4] Lee, Chien-Chiang & Lee, Jun-De & Lee, Chi-Chuan. (2010). Stock prices and the efficient market hypothesis: Evidence from a panel stationary test with structural breaks. Japan and the World Economy. 22. 49-58.
[5] Vanukuru, Kranthi. (2018). Stock Market Prediction Using Machine Learning. International Research Journal of Engineering and Technology 5(10). 1032-1035.
[6] M. Usmani, S. H. Adil, K. Raza and S. S. A. Ali, "Stock market prediction using machine learning techniques," 2016 3rd International Conference on Computer and Information Sciences (ICCOINS), Kuala Lumpur, 2016, pp. 322-327.
[7] K. Hiba Sadia, Aditya Sharma, Adarrsh Paul, SarmisthaPadhi, Saurav Sanyal. (2019). Stock Market Prediction Using Machine Learning Algorithms. International Journal of Engineering and Advanced Technology (IJEAT). 8(4). p: 25-31.
[8] Lv D, Huang Z, Li M, Xiang Y (2019) Selection of the optimal trading model for stock investment in different industries. PLoS ONE 14(2): e0212137.

©IJRASET: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)