Project Report
On

Stock Price Prediction Using Machine Learning

Submitted to: Mr. Kireet Joshi
Submitted by: Krishna Pratap Singh
Section : D

Class roll no : 17
University roll no : 2016821

Declaration

- I, **Krishna Pratap Singh** student of B-tech, Semester 4, Department of Computer Science and Engineering, Graphic Era Deemed University, Dehradun, declare that the technical project work entitled "Stock price prediction using ML" has been carried out by me and submitted in partial fulfillment of the course requirements for the award of degree in B- tech of Graphic Era Deemed University during the academic year 2021-2022.

Date: 16/7/22

<u>Acknowledgement</u>

This project would not have taken shape without the contributions of many who have extended help in one or the way .

I would like to take this opportunity to express my thanks to my father and mother who have provided the entire moral backing and encouragement all through the course duration .

A discussion on the project work, with my seniors in the college was also great source of help and without which it would not have been possible to achieve the results reported in this project .

I would like to express my sincere gratitude to my mentors,

MR. Kireet Joshi (Project supervisor) for his useful comments on the subject matter and for the knowledge I gained by sharing ideas with him.

Finally, it is pleased to acknowledge the indebtedness to all those who devoted themselves directly or indirectly to make this project report success.

Problem Statement

- We have to predict the price of a certain stock at a certain using core machine learning algorithms.

- In this type of model various algorithms can be used to predict the price of a stock, It is Important to understand that apart form doing this we have to make sure that the price of the changes time to time meaning the changes in the graph that occur every time should be seen in the program.

- It is a time series program and good way of getting a deep understanding of various algorithms of machine learning including deep learning.

Abstract

More people invest their money in the stock market. However, this kind of investment possesses a lot of risks. Therefore, many works have been done to build a model

using Machine Learning algorithm to try to predict the stock price values. In this work, Multiple regression and Long-Short Term Memory (LSTM) techniques are used to predict the closing price from five different companies.

The main factor in predicting a stock market is a high level of accuracy and precision. With the introduction of artificial intelligence and high computational capacity, efficiency has increased. In the past few decades, the highly theoretical and speculative nature of the stock market has been examined by capturing and using repetitive patterns. Various machine learning algorithms like Multiple Linear Regression, Polynomial Regression, etc. are used here.

Introduction

- The modern advances in artificial intelligence have led to the creation of new mathematical tools like deep learning and reinforcement learning. Businesses use data science and analytics to obtain solutions for various business issues. Stock trading is one of the most important activities in the finance world. Stocks are an equity investment that denotes a part of ownership in an organization or a company; it entitles you to be a part of that company's earnings and assets. Stock market prediction can be defined as trying to determine the future value of a stock or other financial instrument that is traded on a financial exchange.

- In Indian stock markets, a leading part of trading decisions is made using computer programs. For decades, computer algorithms have been built and tweaked to try to predict the stock market and make the appropriate investment at the right time. Machine Learning can additionally implement algorithms to identify unusual patterns of behaviour based on past behaviours. The stock market has an extremely volatile nature. The major goal is to minimize the uncertainty of the returns by accurately predicting the future stock prices and also identifying their fluctuations in advance to reduce risks.

ALGORITHMS USED

A. **Multiple Linear Regression**

Multiple linear regression is an extended version of the simple linear regression algorithm. The objective is the prediction of the value of a variable based on the value of two or more other variables. The independent variables are needed to predict the value of the dependent variable. The input independent variables could be of continuous or categorical type. The variable for the prediction is called the dependent variable. The regression coefficient means how the dependent variable changes due to a unit change in the independent variable.

B. **Decision Tree Regressor**

A Decision Tree is a very prominent practical approach for supervised learning. It can be employed to solve both Regression and Classification tasks. Both continuous and categorical output variables can be predicted. This algorithm is very useful for resolving decision-related problems. Applications are evaluating growth opportunities for businesses, use of demographic data to find potential clients, etc. The features of an object are analyzed to train a model as a tree and produce meaningful continuous output.

It mainly uses mean squared error to decide whether to split a node into two or more sub-nodes. The accuracy is dependent on the decision of making strategic splits. The top-most decision node in a decision tree is the root node. The parent node that splits into one or more child nodes is called a decision node. Removing the child nodes of a decision node is known as pruning. Bottom nodes that don't split any further are the leaf nodes. The plotting is typically done in an upside-down manner so that the root node is at the top and the leaf nodes are at the bottom. When the dependent variable is continuous, this can be implemented.

Algorithms	Advantages	Disadvantages
Multiple linear regression	1.Find the relative influence of 1 or more predictor variables on the output variable. 2.Ability to find out the anamolies or outliers.	1.It does not work if there is a correlation between error terms or independent variables. 2.Incomplete data can cause many errors.
Decision Tree Regressor	1.It is easy to interpret, understandable for beginners and can visualize faster. 2.It can easily identify the relationship between dependent and independent variables. 3.There is no need to use feature scalability. 4.It can be used for linear as well as nonlinear problems.	1.It does not give suitable results if the dataset is small. 2.Overfitting is possible. 3.The calculations might be hard so the time complexity might increase . 4.A small change can cause a big change in the final answers.

Methodology

The entire code is in the format of the Python Notebook. Python libraries like Pandas, Numpy load the dataset and perform the mathematical calculations on the dataset. Sklearn is used to implement the four different machine learning algorithms. The historical data of the last 5 years was downloaded from the Yahoo finance website.

The dataset available has the following attributes:

1)Date

2)Open

3)High

4)Low

5)Close

6)Volume

7)Adj. Close

a)   **Open Price**: First price of the stock at the beginning of a trading day

b)   **Closing Price**: The price of the stock at the end of a trading day

c)   **Adj Close**: Stock's value after distributing dividends (True value of the stock)

d)   **High Price**: The highest price of the stock in the trading day

e)   **Low Price**: Lowest price of the stock in the trading day

f)   **Volume**: Number of stocks traded for security in all the markets during a given time.

.

Conclusion

- Machine learning has applications related to the recommendation of financial products, customer sentiment analysis, etc. In order to predict the prices of stocks, we need the historical data of the stock. This paper analyses the Data Set with 7 attributes and makes a prediction using different regressors to find the future price. It can be seen that the highest accuracy is obtained using the Decision

- We can summarize by saying that the Decision Tree Regressor gave the best results out of all the models used for the stock price prediction.

References

-Wikipidia

-Geeks For Geeks

-Youtube

-Google