Stock Prices Prediction Using Machine Learning and Deep Learning Techniques

Statement of Work

Course Name	AI Algorithms I
Course Code	AIDI 11688
Course Facilitator	Marcos Bittencourt
Student	Riddhi Thakkar (100799412)
Date	01/11/2020

DURHAM COLLEGE SCHOOL OF BUSINESS, IT AND MANAGEMENT

Contents

Executive Summary

Predicting how the stock market will perform is one of the most difficult things to do. There are so many factors involved in the prediction – physical factors vs. physiological, rational and irrational behavior, etc. All these aspects combine to make share prices volatile and very difficult to predict with a high degree of accuracy

we will work with historical data about the stock prices of a publicly listed company. We will implement a mix of machine learning algorithms to predict the future stock price of this company, starting with simple algorithms like averaging and linear regression, and then move on to advanced techniques like Naïve Bayes, SVM and LSTM etc.

Problem Statement

The stock market appears in the news every day. You hear about it every time it reaches a new high or a new low. The rate of investment and business opportunities in the Stock market can increase if an efficient algorithm could be devised to predict the short-term price of an individual stock.

Broadly, stock market analysis is divided into two parts – Fundamental Analysis and Technical Analysis.

- Fundamental Analysis involves analysing the company's future profitability on the basis of its current business environment and financial performance.
- Technical Analysis, on the other hand, includes reading the charts and using statistical figures to identify the trends in the stock market.

Data

As you might have guessed, our focus will be on the technical analysis part. We'll be using a dataset from Quandl (you can find historical data for various stocks here) and for this particular project, I have used the data for 'Tata Global Beverages'. Time to dive in!

The data can be accessed at the following link: https://cdn.analyticsvidhya.com/wp-content/uploads/2019/03/NSE-TATAGLOBAL11.csv"

Data Requirements

There are multiple variables in the dataset – date, open, high, low, last, close, total_trade_quantity, and turnover.

- The columns *Open* and *Close* represent the starting and final price at which the stock is traded on a particular day.
- *High*, *Low* and *Last* represent the maximum, minimum, and last price of the share for the day.
- *Total Trade Quantity* is the number of shares bought or sold in the day and *Turnover (Lacs)* is the turnover of the particular company on a given date.

Another important thing to note is that the market is closed on weekends and public holidays. Notice the table again, some date values are missing – 2/10/2018, 6/10/2018, 7/10/2018. Of these dates, 2nd is a national holiday while 6th and 7th fall on a weekend.

The profit or loss calculation is usually determined by the closing price of a stock for the day, hence we will consider the closing price as the target variable

Model/Architecture Approach

- Machine Learning Model/Endpoint

For this component, I will use below feature to solve the problem of Stock prediction.

- We would be using algorithms like Naïve Bayes, SVM, Linear Regression and LSTM and compare accuracy to use the best algorithm.
- The logic to preprocess data, machine learning model will be written in Python 3.0 using Pandas

Data Frames. The model steps will be written in Jupiter notebook for easy review and analysis.