Machine Learning Engineer Nanodegree

Capstone Proposal

Mostafa ElGhandour

March, 31st, 2019

Proposal

Domain Background

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, which makes it a very project to implement.

Problem Statement

Stock market analysis is divided into two parts – Fundamental Analysis and Technical Analysis.

Fundamental Analysis: involves analyzing 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.

our focus will be on the technical analysis part.

Datasets and Inputs

I will be using data sets from https://www.quandl.com/ to predict gold price in London.

Solution Statement

I will be trying a few models (most likely linear regression) to predict the stock prices and using the data sets from the following website i will be able to have some test and evaluation data that I can evaluate.

Benchmark Model

There are a lot of benchmark models out there with linear regression algorithms that can be used to validate my results relative to the real stock price predicted data.

Evaluation metric

The evaluation metric would be comparing the predicted number to the real stock price number at a certain time.

Project Design

1-i will be pre-processing the dataset to achieve usable raw data.

2-i will choose a suitable machine learning model (possibly more than one).

3-i will train the data using a part of my dataset.

4-i will test the data using another test data of my dataset.

5-i will calculate the accuracy.

6-i will tune my model parameters and calculate the model's accuracy till i achieve the highest accuracy i can achieve.

7-i will try the model on a real stock data if applicable.