Predicting if the closing price of stock will rise or fall the next trading day

Brittany Miller
Bolu Aiki-Raji
Samman Bikram Thapa

The milestone report should consist of

a. **your project proposal (just include your project proposal):**

This project is an attempt to try to answer the question “Will the stock price of a stock increase or decrease in future?” to some degree of surety by using Machine Learning. We are trying to predict if the close price of stock will rise or fall on the next trading day, so we have approached this as a classification problem.

b. **your progress so far,**

   i. **how much you have done:**

In our project plan, we had set the following milestones:

We are going to:

1) Get the input for the algorithm includes (IV, independent variables):
    a) Moving average of historical close prices
    b) Trading volume, and open, highest, lowest, and close prices of the present trading day
    c) financial indicators, e.g., DecisionPoint Price Momentum Oscillator (PMO), Money Flow Index (MFI), Percentage Price Oscillator (PPO), and etc
    d) A price movement trend based on local Taylor expansion and spline fitting
2) Perform Data cleaning and Normalizing the input
3) Train the KNN classifier (described below)
4) Predict based on the classification of the model:
    a) the stock price will rise (1)
    b) the stock price will fall (-1)

Of the milestones we had set, we have completed getting our Independent Variables partially. Apparently Quandl does have a lot of data because it maps data from various sources, but we need to figure out which data source is helpful to us. So far, we have identified Walt Disney CO (NYSE: DIS), Dr Dl GldMnIdx Bl3x (NYSEARCA: NUGT), Apple Inc. (NASDAQ: AAPL), and Yahoo! Inc. (NASDAQ: YHOO) as our target stocks which we will be learning from. Their data have been found in multiple sources through Quandl and we have been able to get their data to a certain date.

ii. **include the results so far of the things that you have implemented (even it didn't work, and give your reasoning to why it didn't),**

We have implemented the KNN classifier so far in python. We went through the online tutorial at http://machinelearningmastery.com/tutorial-to-implement-k-nearest-neighbors-in-python-from-scratch/ to get the basics of KNN and implement it for the provided data.

iii. **the challenges you are facing,**

The challenges we are facing at this time is that we are not getting the DecisionPoint Price Momentum Oscillator (PMO), Money Flow Index (MFI), Percentage Price Oscillator (PPO) data easily for all the four stocks we chose. And these information are part of our IV, so we are working on finding other sources

iv. **the lessons you learned, and,**

The lessons we have learned so far is that KNN is not model based, so it does not lose any detail and compares every training sample to give the prediction. However, kNN is pretty slow when you have a lot of observations, since it does not generalize over data in advance.

v. **how you are moving forward to finish the project,**

We are working together on the same screen as that seems to be a better. Since we are only using KNN classifier and not another classifier, and since we already have the code implemented, we should be able to finish the learning and testing in a week, at most. We should be able to finish by the time we are expected to submit our deliverable

vi. **each member's contribution.**

All of us shared the same project folder and worked on it together. The coding part was also done on the same screen, so we all have contributed equally.