**Strategy Learner**

CS 4646

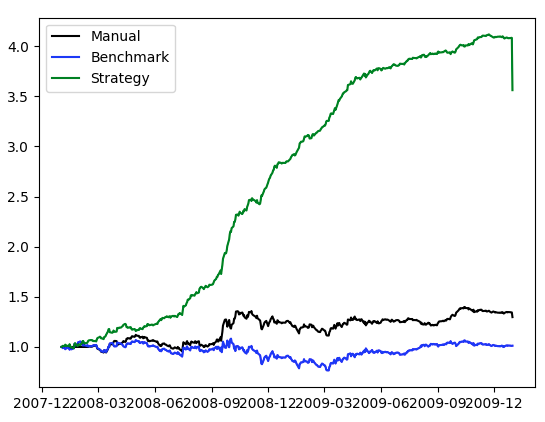
Hardik Sangwan

For the manual strategy for this trading problem, I used Price/SMA, Bollinger Band and Momentum (different from previous assignment).   
SMA is simply the sum of values over some period/divided by the discrete units of that period. In this case it is the sum of prices over last 14 days. Price/SMA ratio gives an indication of whether stock should be bought/sold based on whether the value is less than or greater than 1. A value greater than 1 would indicate that opportunity to buy since price is below SMA and vice versa for less than 1. Bollinger Band described in the earlier assignment. Momentum compares current price data to previous days (14 days in my case) and increases and decreases proportionally to the rate of change of prices. A stable price will mean a momentum closer to 0 while an uptick or downward trend in price will give a bigger positive or negative value. For the strategy, an amount equal to 1000-holdings is bought when current holdings are less than 1000, price/sma < 1.1, Bollinger band < .15 and momentum <0. An amount equal to 1000+hold is sold when momentum is greater than 0, holdings are greater than -1000, price/sma > 0.9 and Bollinger band > 0.95. Nothing is done in any other case.  
  
For the learner’ strategy, a Q Leaner was used. Problem set up as MDP to calculate best action at given state for best long term reward. The reward can be seen as our portfolio gain/return. The data is discretized with values from 0-9 to limit the number of available values for the algorithm and reduce time and space complexity. Current state is calculated based on discretized indicators like holdings and reward, to produce a q table. For testing instead of training, we take the q table we already produced to check for best action given our state and policy.

**Experiment 1: Manual vs Learning Strategy**

|  |  |
| --- | --- |
| Strategy | Cumulative Return |
| Benchmark | 0.01 |
| Manual | 0.31 |
| Strategy Learner/Q Learning | 2.56 |

Q Learner performed significantly better than the benchmark and the manual strategy since it constantly looks for the better possible long term reward compared to the manual strategy which is derived through unchanging parameters based more on theoretical intuition and less on actual data.   
(Fees subtracted at the end of the run for this experiment)



**Experiment 2: Effect of Impact on Strategy**

Table below showing cumulative return for different impacts for the q learning based strategy. It can be seen that a similar return is seen at most values with a local maxima found around 0.05, most likely specific to the symbol being traded and the time period. The outlier near 1 where a net 0 return is seen gives us a view into how impact is affecting our trading behavior. A value of 1 for impact implies that no long or short position will be selected at any moment since the penalty for ‘moving’ is high. On the other hand there is no penalty near an impact of 0 and taking a short/long position is encouraged.

|  |  |
| --- | --- |
| Impact | Cumulative Return |
| 0 | 3.12 |
| 0.005 | 3.26 |
| 0.01 | 3.24 |
| 0.05 | 3.39 |
| 0.1 | 3.37 |
| 0.2 | 3.21 |
| 0.6 | 3.29 |
| 1 | 0.0 |

