Project 8: Strategy Learner

OMSCS CS7646: Machine Learning for Trading

Clement Henghui Li (cli620)

Set up:

I am using the Q-learner algorithm to train and predict the trading day action for a given stock. The possible actions to take are, a short position (0), no action (1) or a long position (2). The possible states are 10 to the power of number of indicators (3); which shows a total number of 1000 of total states. The base is 10 because of discretizing each indicator by separating each indicator into 10 bins. If the indicator is a nan, then we will default the value to a 4 to try and simulate the action of doing nothing. The final parameter becomes momentum*100 + volatility * 10 + Bollinger band. This will give us 1000 possible combinations of all three indicators. The Q learner will update a map of possible states versus possible actions containing the probability of each action given the state.

The alpha and gamma is set to 0.1 and 1.35 respectively. The alpha was set low because 20 window size that was chosen to determine the indicators. This window makes us less confident in the earlier values, so a slower learning curve is required. Gamma is selected as 1.35 because we value the importance of future events like the argument earlier, we do not trust the start values of the indicators. I kept a random action probability to 0.9 for training purposes only. This is intended to prevent us from being too overfitted with our train data.

Experiment 1:

Here we compare the difference of the manual strategy and the q-learner strategy learner.

In Sample:

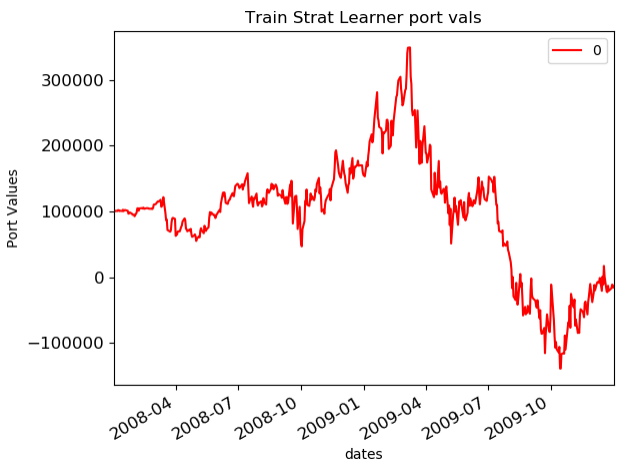
Date Range: 2008-01-01 00:00:00 to 2009-12-31 00:00:00

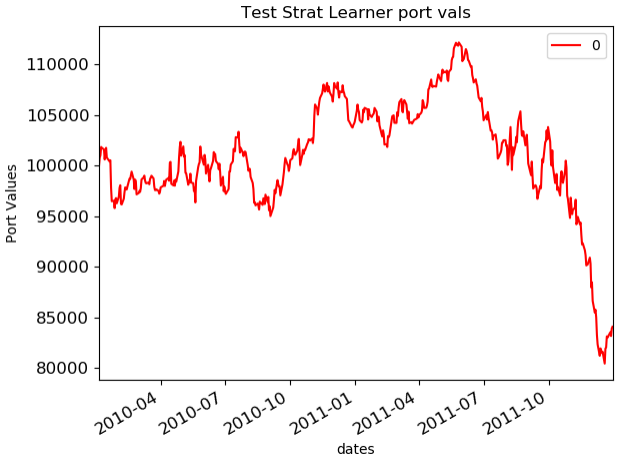
Cumulative Return SL of JPM: 0 -2.3379	Cumulative Return MS of JPM: 0 0.4331
Standard Deviation SL of JPM: 0 3.786526	Standard Deviation MS of JPM: 0 0.01325
Mean SL of JPM: 0 -0.081835	Mean MS of JPM: 0 0.000801
Sharpe Ratio SL of JPM: 0 -0.343081	Sharpe Ratio MS of JPM: 0 0.960271
Final SL Portfolio Value: [-133790.]	Final MS Benchmark Value: [143310.]

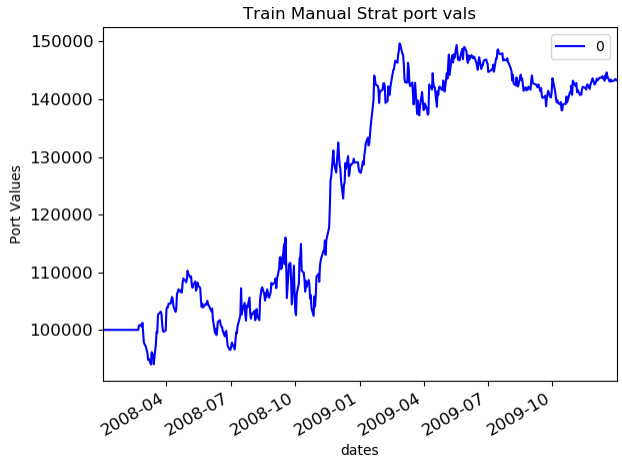
Out Sample:

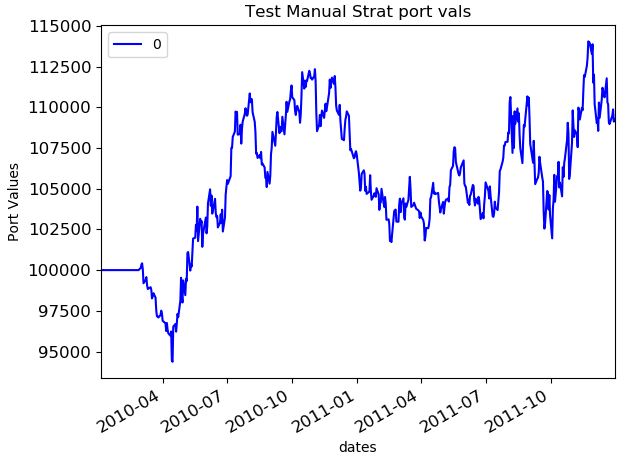
Date Range: 2010-01-01 00:00:00 to 2011-12-31 00:00:00

Cumulative Return SL of JPM: 0 0.6891	Cumulative Return MS of JPM: 0 0.4331
Standard Deviation SL of JPM: 0 0.024888	Standard Deviation MS of JPM: 0 0.01325
Mean SL of JPM: 0 0.001347	Mean MS of JPM: 0 0.000801
Sharpe Ratio SL of JPM: 0 0.85949	Sharpe Ratio MS of JPM: 0 0.960271
Final SL Portfolio Value: [97020.]	Final MS Portfolio Value: [109280.]









This shows that the performance of the strategy learner is worst than the manual method. The Sharpe Ratio is lower for the strategy learner with the volatility being higher and the mean portfolio value being higher. The strategy learner performance is good until the around July 2011.