Project 8: Strategy Learner
Name: Jihoon 'Jay' Song
Course: CS7646

Describe the steps you took to frame the trading problem as a learning problem for your learner. What are your indicators? (They should be the same ones used for Manual Strategy assignment) Describe how you discretized (standardized) or otherwise adjusted your data. If not, tell us why not.

Given the choices, regression/classification, reinforcement and optimization learner, I chose the reinforcement (QLearner) learner as it posed the most interesting problem for me. And just as did in the Manual Strategy assignment, I chose the three indicators, SMA, Bollinger Bands and Stochastic Oscillator to use in my learning problem for my learner. As described in the assignment guideline, I created the StrategyLearner class and divided it into two main functions, addEvidence and testPolicy.

The addEvidence function, as indicated by its name, was used to add the evidence/training data to the learner. In order to normalize and facilitate the usage of the vast indicator datapoints, the indicators were discretized into 10 different bins. In order to make sure it can accommodate any unforeseen datapoint, the bins were written to be dynamic. For instance, for the SMA (rolling mean) indicator, I took the minimum value and the maximum value and took the difference between the two. I then divided it into 10 bins equally. This was done for each indicator. By the end of the discretization, every timeframe had a four-digit string which was the concatenated numbers between 0 and 9, which indicated the bins in which the indicator datapoints fell in the given index.

Using the discretized indicators, we trained the learner using the QLearner which was already complete from our previous assignment, "QLearner". To frame the problem in a way the learner could understand, we used daily returns as the "reward". And in order to make sure we account for the magnitude of the reward, we multiplied it by the net holdings of the shares of the stock. For each index, the action of whether to buy, hold or sell the stock was stored. And depending on the net holdings present (logic taken from manual strategy project), the number of shares purchased or sold was also stored for each index. This was iterated 200 times or as much as it took before the results converged and stayed constant.

The testPolicy function, as indicated by its name was used to test the performance of the learner. Like the addEvidence funcition, the indicators were collected and discretized using the same functions in the given timeframe. In a similar fashion minus the usage of rewards, given this function was used to test the performance of the learner, each action of buy, hold, sell and the associated number of shares were collected and stored.

This dataframe contained all the resultant list of orders for the given timeframe. Using the compute_portvals and get_portfolio_stats function which was previously written in the marketsimcode.py, the total portfolio value for at each index, and the resulting overall portfolio performance statistics were provided.

Experiment 1: Using exactly the same indicators that you used in manual_strategy (trade JPM), compare your manual strategy with your learning strategy in sample. You can use the same impact (.005) as was used for Project 6 or use 0 for both. Be sure to add in an author method.

- **Describe your experiment in detail: Assumptions, parameter values and so on.**

Experiment #1 assumed the impact is 0; the in sample/development period is January 1, 2008 to December 31 2009; the out of sample/testing period is January 1, 2010 to December 31 2011; the starting value was $100,000; the trades have to be in 1000 share increments and could not be over 2000 shares long or below 2000 shares short. The stock used for comparison was 'JPM'. The "windows" for the rolling means were 20 days.

- **Describe the outcome of your experiment.**

As shown in **Figure 1.0***,* for the out-sample test, the Manual Strategy Portfolio as well as the Benchmark performed significantly better than the Strategy Learner. It should be noted that although the Strategy Learner performed significantly worse by the end of the experiment, for a period of January 2011 and June 2011, it was significantly outperforming the other two portfolios. As shown in **Table 1.0,** it looks like the strategy learner is willing to take many more risks than the manual strategy and the benchmark as indicated by the significantly higher standard deviation at 11.9146.

- **Would you expect this relative result every time with in-sample data? Explain why or why not.**

I would not expect this result every time with in-sample data. Given that the appropriate rewards were provided to the learner in the training phase, the learner should perform much better and more consistently in the in-sample data. As shown below in ***Figure 2.0***, the Strategy Learner did indeed perform better over the in-sample data.

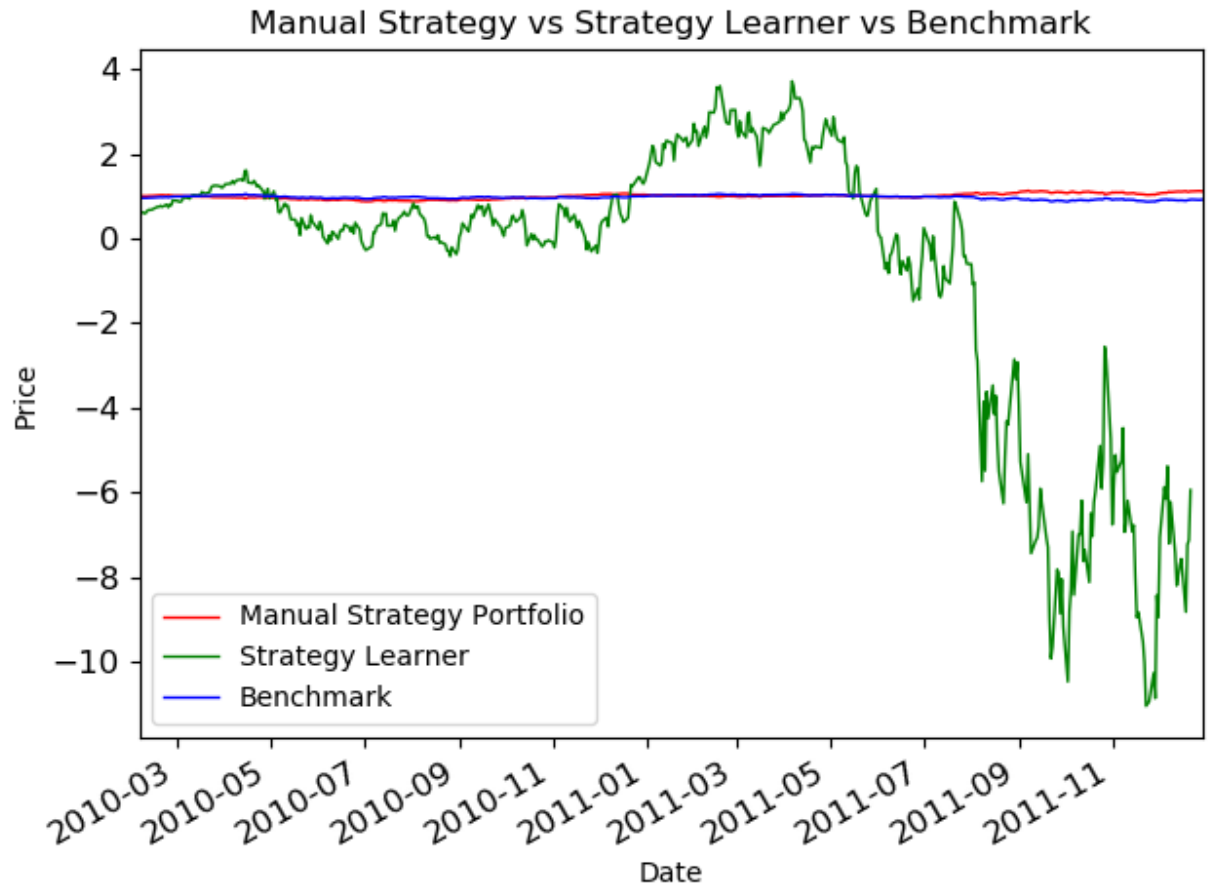

FIGURE 1.0

	Manual Strategy	Benchmark	Strategy Learner
Sharpe Ratio	0.519	-0.257	-0.8495
Cumulative Return	0.116	-0.083	-8.3735
Average Daily Return	0.00026	-0.00014	-0.6376
Standard Deviation	0.00796	0.00848	11.9146
Final Portfolio Value	$111,610.0	$91,660.0	-$737,350.0

TABLE 1.0

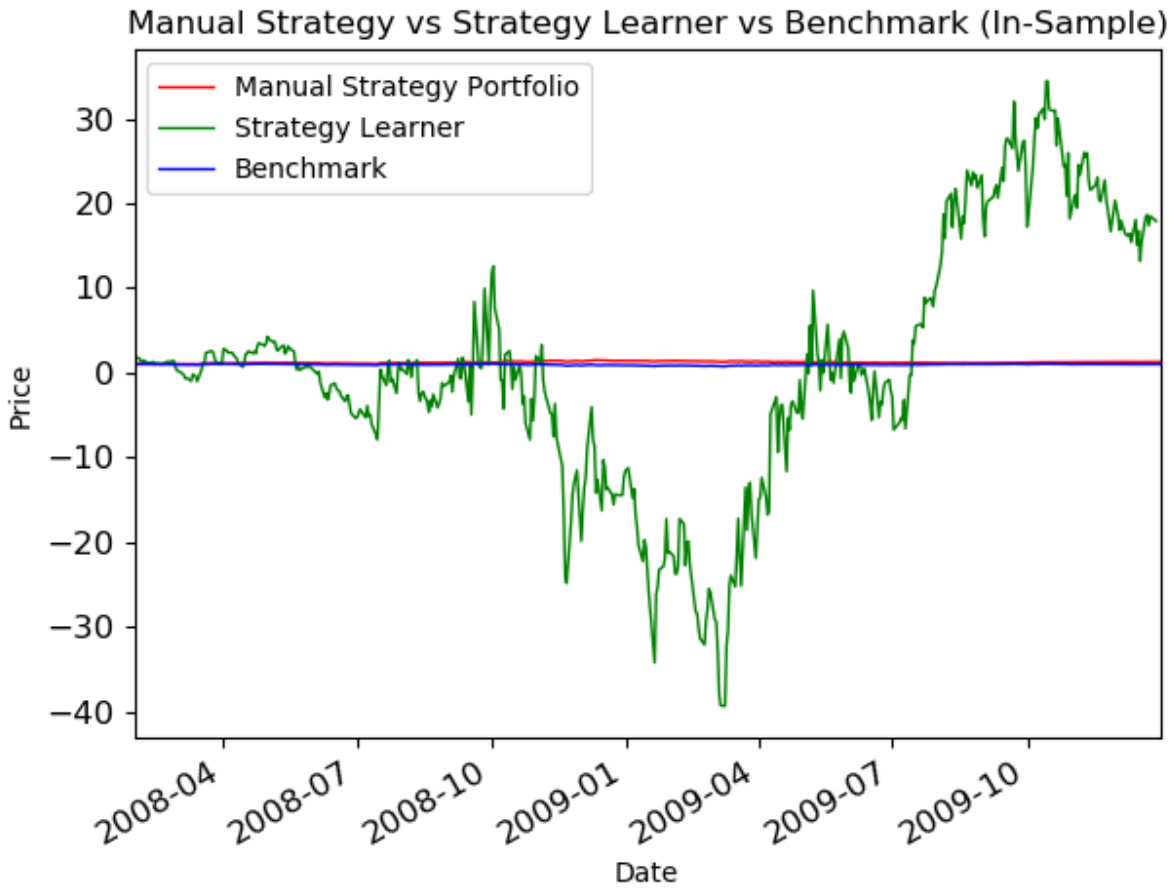

FIGURE 2.0

EXPERIMENT 2: PROVIDE A HYPOTHESIS REGARDING HOW CHANGING THE VALUE OF IMPACT SHOULD AFFECT IN SAMPLE TRADING BEHAVIOR AND RESULTS (PROVIDE AT LEAST TWO METRICS). CONDUCT AN EXPERIMENT WITH JPM ON THE IN SAMPLE PERIOD TO TEST THAT HYPOTHESIS. PROVIDE CHARTS, GRAPHS OR TABLES THAT ILLUSTRATE THE RESULTS OF YOUR EXPERIMENT. THE CODE THAT IMPLEMENTS THIS EXPERIMENT AND GENERATES THE RELEVANT CHARTS AND DATA SHOULD BE SUBMITTED AS EXPERIMENT2.PY. BE SURE TO ADD IN AN AUTHOR METHOD.

My hypothesis is that if the value of impact is increased, the standard deviation of the strategy learner portfolio should decrease, and the average daily return should decrease. I believe this is because increasing the impact would decrease the reward leading to fewer trades occurring and thus decrease in standard deviation, but also a decrease in average daily return.

From looking at **Figure 3.0,** it looks as if there wasn't a significant change when the impact was increased from 0.0 to 0.05. However, looking more closely at **Table 2.0,** we can see that the standard

deviation decreased from 11.91 to 2.36. In addition, the average daily return decreased from -0.64 to -0.044. Therefore, my hypothesis that the standard deviation would decrease looks to be correct. However, my hypothesis that the average daily return would decrease was incorrect. The “increase” in average daily return is a decrease in the magnitude of the loss. This is likely because trades were made less frequently which reduced the risk. This likely led to decrease in the magnitude of the loss.

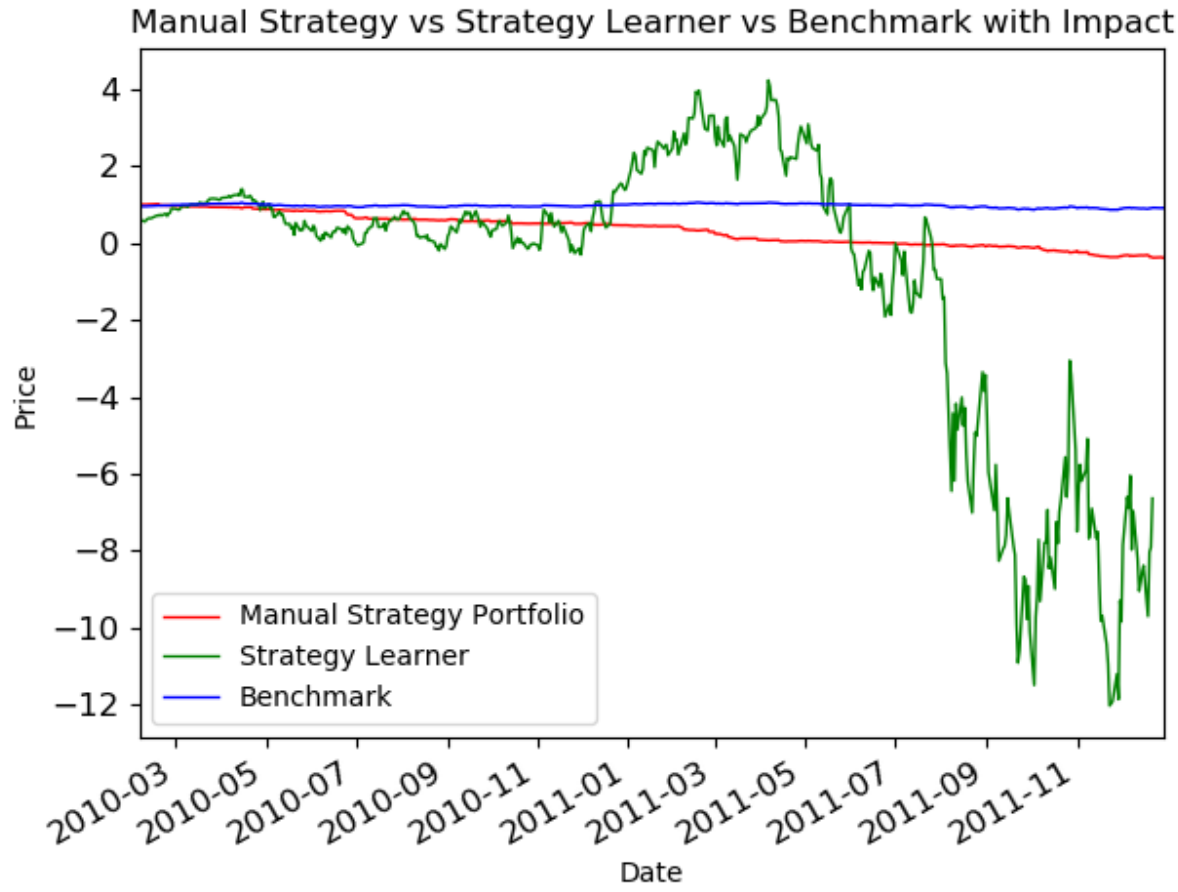

Figure 3.0

	Manual Strategy	**Benchmark**	**Strategy Learner**
Sharpe Ratio	-0.46667	-0.31963	-0.30175
Cumulative Return	-1.392	-0.1017	-7.2341
Average Daily Return	-0.0133	-0.000175	-0.04479
Standard Deviation	0.45266	0.00871	2.35656
Final Portfolio Value	-$38,497.5	$87,990.0	-$636,463.5

Table 2.0