CS 7646

# Project 6

December 1, 2019

## Strategy Learner Description

The Strategy learner generates signals for trades on a particular day. The signals are of the form +1 (BUY), -1 (SELL) and 0 (DO NOTHING). These signals are generated using indicators as training data. For this assignment, I am using Bag Learner as a learner with indicators indicators sma_ratio, bb_ratio and momentum as features. I have also modified the manual strategy to include momentum since only sma_ratio and bb_ratio were used earlier in Project 6.

**Indicators**

- Price/SMA ratio

$$sma[t] = \frac{1}{N} \sum_{i=0}^{i=N-1} price[t-i]$$

$$sma\_ratio[t] = \frac{price[t]}{sma[t]}$$

- BB Ratio

$$std[t] = \sqrt{\frac{1}{N} \sum_{i=0}^{i=N-1} (price[t-i] - sma[t])^2}$$

$$BU[t] = sma[t] + 2 * std[t]$$

$$BL[t] = sma[t] - 2 * std[t]$$

$$bb\_ratio[t] = \frac{price[t] - sma[t]}{2 * std[t]}$$

- Momentum

$$momentum[t] = \frac{price[t] - price[t-N]}{price[t-N]}$$

All the indicators are used with a lookback window of 14 days same as manual strategy.

**Initialization**

An instances of bag learner is initialized with the following parameters

- Leaf size: 5
- Learner: Random Tree Learner
- Bags: 20 (implies 20 instances of Random Tree Learner)

Since this is a classification problem with classes -1, 0 and 1, the learners have been changed to return the mode instead of mean. This means on reaching a leaf node, the class with the maximum frequency is selected as result.

**Training**

- The indicators are used as features with a lookback window of 14 days, same as used in the manual strategy. All the days for which the indicator value is not present are dropped, i.e. the first 13 days.
- To calculate the trade signal that is target variable y, I have used N day returns where N = 14. It calculates the return for today based on the stock price 14 days after. The last few days for which we cannot calulate the return have been dropped.

$$return[t] = \frac{return[t+N]}{return[t]} - 1$$

- The logic for generating trade signal i.e. y is following:
  - If $return[t] > 0.05 + impact$, go long i.e. +1

- If $return[t] < -0.03 + impact$, go short i.e. -1
- In all other cases, do nothing, i.e. 0

- As mentioned before, mode is used for accumulating the results in the leaf of a random tree learner and for all the trees in a bag learner.

**Testing**

- The features are calculated for the test data using the same method as training phase. The days for which indicator values are not present are dropped.
- The trade signals are predicted by the learner by querying the bag learner, which returns -1, 0, 1 for each trading day.
- In the final step, trades dataframe is created using the above signals, with positions ranging from 1000 long to 1000 short. This implies the buy orders can be 2000 and sell orders can be -2000.

**Data Adjustment**

I did not take any steps to adjust the data. This stems from the choice of learner used. Decision trees do not require any scaling or discretization of the data as it splits on the feature value. I am using median to split the node into left and right subtrees.

## Modifications to Manual Strategy

My manual strategy used in Project 6 used only 2 indicators sma_ratio and bb_ratio. I modified that to include sma_ratio, bb_ratio and momentum to include 3 indicators as suggested in the comments. I will be using this modified version for the rest of the analysis. The lookback period is the same for all indicators as used in Project 6 i.e. 14 days.

I used a logical combination of sma_ratio and bb_ratio and momentum to generate trade signals The trading strategy is as follows:

- If ($sma\_ratio > 1$ AND $bb\_ratio > 1$) OR ($momentum > 0.2$), go short
- If ($sma\_ratio < 1$ AND $bb\_ratio < -1$) OR ($momentum < -0.2$), go long
- If none of the above conditions satisfy, do nothing

sma_ratio greater than 1 indicates divergence and the price is going to drop, hence a sell signal. Similary sma_ratio less than 1 indicates divergence in opposite direction, the price is went lower than it should, indicating a buy signal. bb_ratio works similarly to sma_ratio but with thresholds 1 and -1. Momentum's high values are used as a divergence signal. Whenever the momentum is more than 0.2, it indicates an unusual rise in price and a reversal trend, hence a sell signal. Similarly, if the momentum is less than -0.2, it indicates an unusual drop in price, indicating price would rise eventually, hence a buy signal.

Since sma_ratio gave a good notion of trend but was very jittery, I combined it with bb_ratio using an AND operator to confirm that the trend. Momentum is combined with the result of sma_ratio and bb_ratio using an OR indicator to generate additional signals which might have been missed by the other two indicators.

The statistics and plots for the manual strategy are presented alongwith the Strategy Learner in the following section for both in-sample and out-sample periods.

# 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.**

In this experiment, I compare the performance of Strategy Learner with the Manual Strategy.

**Describe your experiment in detail: Assumptions, parameter values and so on. Describe the outcome of your experiment.**

### Description

The Strategy learner is used to predict trade signals using the above setup. For impact value 0.005, strategy learner learns the trading signals (+1, -1 and 0) on in-sample data and then tests the policy learned on in-sample data and out-sample data. The test policy generates a list of trades which are then converted into a dataframe of orders. This order dataframe is used as input for computing the portfolio value with a starting cash of 100,000. The portfolio values are normalized and used for computing statistics and generating a comparison plot.

### Assumption

My assumption is that using the indicators sma_ratio, bb_ratio and momentum, the Strategy learner would be able to learn a strategy and give better results than Manual Strategy atleast for in-sample data.

### Setup

- Impact: 0.005
- Symbol: JPM
- In sample period: Jan 1, 2008 to Dec 31, 2009
- Out sample period: Jan 1, 2010 to Dec 31, 2011
- commission: 0.0
- Impact value for benchmark is set to 0.0
- Leaf size: 5
- Number of Bags: 20
- Starting cash: 100,000
- Position: 1000 shares long, 1000 shares short, 0 shares
- Buy threshold for Strategy learner is 5% gain + impact
- Sell threshold for Strategy learner is 3% loss + impact
- The plots shown and statistics shown in this experiment are all corresponding to a random seed of 42.

## Results

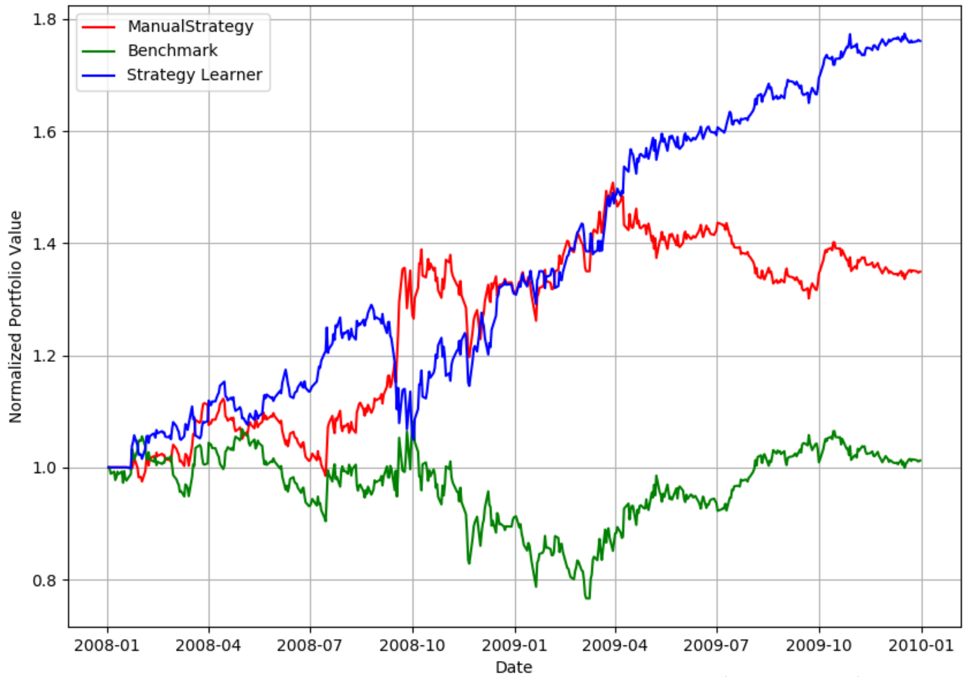

Figure 1: Strategy vs Manual vs Benchmark (In Sample)

Table 1: In-sample statistics

| | Strategy | Manual | Benchmark |
|---|---|---|---|
| Cumulative Return | 0.7605 | 0.3492 | 0.0123 |
| Avg Daily Return | 0.0012 | 0.0007 | 0.0002 |
| Std Daily Return | 0.0126 | 0.0125 | 0.0170 |
| Sharpe Ratio | 1.5119 | 0.8890 | 0.1868 |

It can be observed from the above plot and table that Strategy Learner outperforms both Manual Strategy and Benchmark on in-sample data. It gives both high cumulative returns and Sharpe Ratio. We can also observe that the standard deviation of daily returns is higher for Strategy learner compared to Manual Strategy. This can be attributed to the dip in plot observed of Strategy learner around Oct 2008. The above plot and chart are for seed value 42.

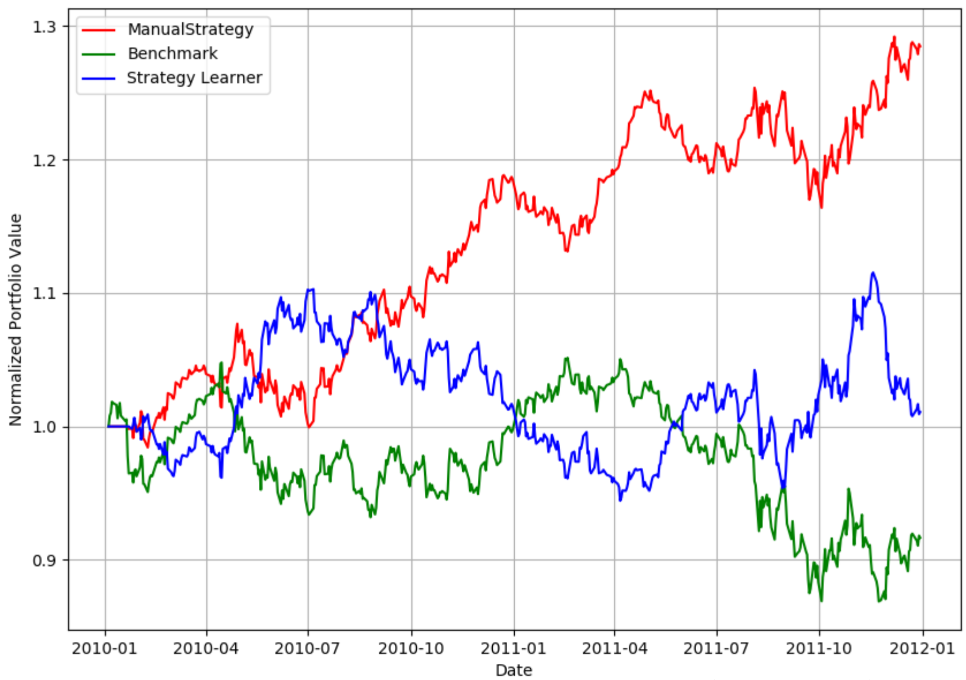

Figure 2: Strategy vs Manual vs Benchmark (Out Sample)

Table 2: Out-sample statistics

| | Strategy | Manual | Benchmark |
|---|---|---|---|
| Cumulative Return | 0.0108 | 0.2847 | -0.0834 |
| Avg Daily Return | 0.0000 | 0.0005 | -0.0001 |
| Std Daily Return | 0.0076 | 0.0068 | 0.0085 |
| Sharpe Ratio | 0.0000 | 1.1672 | -0.1868 |

As we can observe from the plot, that while Strategy learner performs very well on in-sample data it's performance dips on out-sample data. It performs relatively better than benchmark but poorly compared to Manual Strategy. This can be attributed to multiple reasons, that the Strategy learner overfitted on the in-sample data or manual interpretations are more stable compared to learning-based strategy as they involve a trader's expertise. It should be noted though that the strategy learner might perform worse than the benchmark for a different seed value. Above plots are for seed = 42.

**Would you expect this relative result every time with in-sample data? Explain why or why not.** I would expect the Strategy learner to beat the benchmark everytime and also beat the manual strategy for most of the cases on in-sample data. This is because manual strategy is devised using a simple combination of indicators compared to a decision random tree learner which is relatively complex. However, the performance of the Strategy learner might vary for different runs owing to the randomness introduced by Random Tree Learner in the ensemble technique. It might not give the best result in every run.

## 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).**

### Hypothesis

Impact is the cost that the trader has to bear while buying or selling a stock due to the movement of the stock price against the trader. It is determined by the liquidity of a stock. Therefore, a low impact cost would be favourable to the trader. My hypothesis is that as the strategy learner performance increaes as the impact value decreases.

I will be using the metrics cumulative return and sharpe ratio for testing this hypothesis. According to the hypothesis, a decrease in impact value should correspond to increase in cumulative return and sharpe ratio. Also, it would impact the number of trades made in a period.

**Conduct an experiment with JPM on the in sample period to test that hypothesis.**

### Setup

- Impact values used: 0.0005, 0.005, 0.05
- Symbol: JPM
- In sample period: Jan 1, 2008 to Dec 31, 2009
- commission: 0.0
- Impact value for benchmark is set to 0.0
- Leaf size: 5
- Number of Bags: 20
- Starting cash: 100,000
- Position: 1000 shares long, 1000 shares short, 0 shares
- Buy threshold for Strategy learner is 5% gain + impact
- Sell threshold for Strategy learner is 3% loss + impact
- The plots shown and statistics shown in this experiment are all corresponding to a random seed of 42.

### Description

The Strategy learner is used to predict trade signals using the above setup. For each impact value, strategy learner learns the trading signals (+1, -1 and 0) on in-sample data and then tests the policy learned on in-sample data only. The test policy generates a list of trades which are then converted into a dataframe of orders. This order dataframe is used as input for computing the portfolio value with a starting cash of 100000. The portfolio values are normalized and used for computing statistics and generating a comparison plot.

### Results

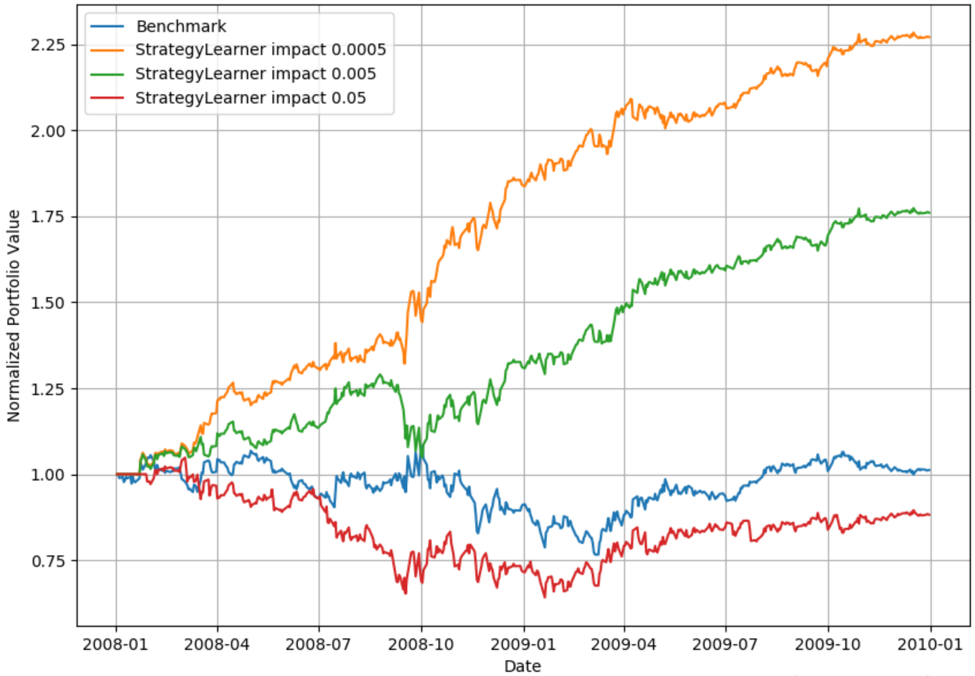


Figure 3: Strategy Learner Performance with Impacts (In Sample)

As we can observe from the plot, the portfolio value increases as the impact decreases. This confirms the hypothesis. We can also see that the learner with impact value 0.05 performs even worse than the benchmark. In general,

$$p\_value(0.0005) > p\_value(0.005) > p\_value(0.05)$$

Table 3: In-sample statistics

| | impact=0.0005 | impact=0.005 | impact=0.05 | Trend as impact decreases |
|---|---|---|---|---|
| Cumulative Return | 1.2711 | 0.7605 | -0.1178 | Increases |
| Avg Daily Return | 0.0017 | 0.0012 | 0.0000 | Increases |
| Std Daily Return | 0.0101 | 0.0126 | 0.0214 | Decreases |
| Sharpe Ratio | 2.6719 | 1.5119 | 0.0000 | Increases |

The impact value also impacts the number of trades the learner predicts. Below are the charts that represent the trades with vertical lines for impact values 0.0005 and 0.05. Vertical blue lines indicating LONG entry points. Vertical black lines indicating SHORT entry points. It is evident from the plots that the number of trades decreases with increase in impact value.

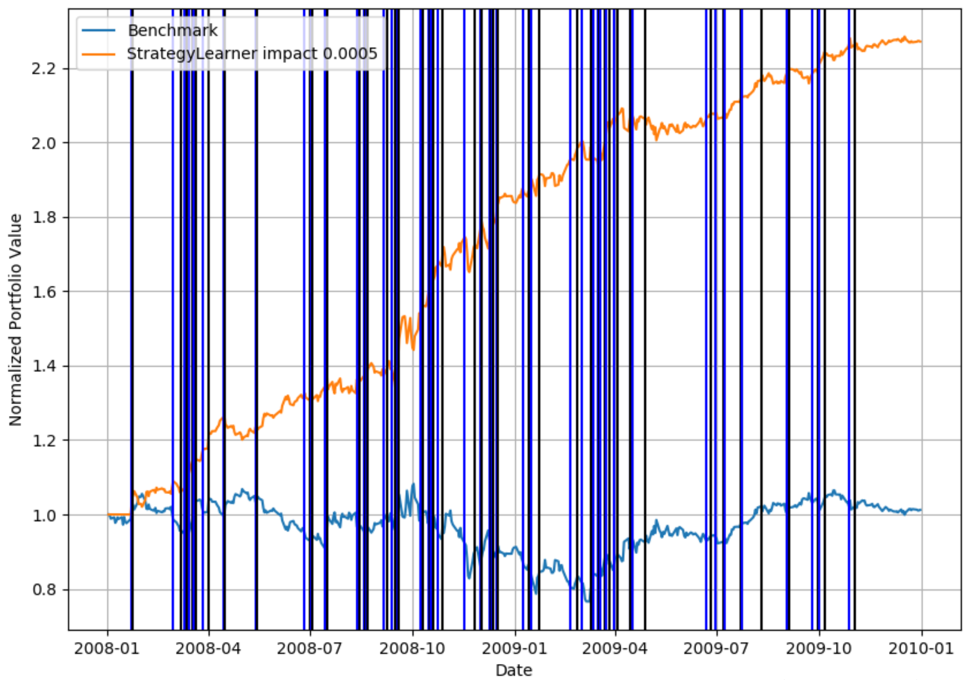


Figure 4: Strategy Learner trades with impact 0.0005 (In Sample)

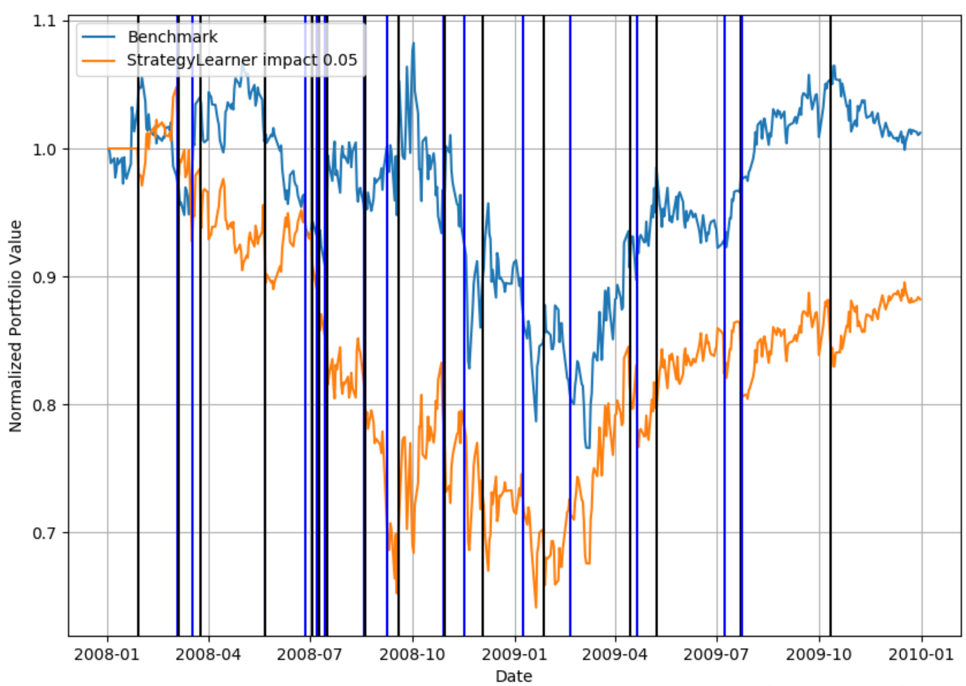


Figure 5: Strategy Learner trades with impact 0.05 (In Sample)