Machine Learning for Trading
Project 8 Strategy Learner

Author Name
lzhang699@gatech.edu

***Abstract*—** Comparison on the performance of 3 trading strategies including Benchmark, Manual Strategy, and Strategy Learner.

1 INTRODUCTION

In this project, 3 different trading strategies are compared and evaluated over two time periods using the ticker JPM, 2008/01/01 - 2009/12/31 (in-sample), 2010/01/01 - 2011/12/31 (out-of-sample). The benchmark strategy assumes holding 1,000 long positions throughout the trading time. The Manual Strategy implements a hard-coded algorithm with 4 indicators which generate buy/sell signals. The Strategy Learner is implemented with Q Learner that is trained with in-sample data, then evaluated with both in-sample and out-of-sample data. Performance is mainly measured by the cumulative return. The hypothesis is that the performance follows the order from the best to the worst, Strategy Learner, Manual Strategy, and benchmark strategy.

2 INDICATOR OVERVIEW

All 5 indicators in the project 6 report are used to build Manual Strategy and Strategy Learner: Simple Moving Average (SMA), Percent B, Momentum, Golden/Death Cross, Moving Average Convergence Divergence (MACD). SMA, Percent B and Momentum are devised slightly to generate more reliable trading signals.

SMA is divided from daily stock prices. If the result is greater than 1.05, then it is a sell signal. If the result is smaller than 0.95, then it is a buy signal. If the result falls between 0.95 and 1.05, then no trade should be executed.

Percent B is combined with Momentum. If the Momentum is negative and Percent B is positive, then it is a sell signal. If the Momentum is positive and Percent B is negative, then it is a buy signal.

Golden/Death Cross and MACD each generate 2 metrics with different window sizes, values of which are compared by each 2 consecutive days/ rows. For

Golden/Death Cross, if 50 SMA crosses above 200 SMA, then it is a buy signal. If 50 SMA crosses under 200 SMA, then it is a sell signal. For MACD, if MACD Raw crosses above MACD Signal, it is a buy signal. If MACD Raw crosses under MACD Signal, it is a sell signal.

Therefore 4 indicators are used to suggest trading signals: price / SMA, Percent B & Momentum, Golden/Death Cross, MACD

3 MANUAL STRATEGY

In order to produce an optimized performance, a weight is assigned to each indicator.

$$0.2 * price\ /\ SMA\ +\ 0.2 * Percent\ B\ \&\ Momentum\ +\ 0.3 * Golden/Death\ Cross\ +\ 0.3 * MACD$$

The trading signals - buy, hold, sell are represented by 1, 0, -1 from each indicator thus the result above also returns a positive, zero or negative number. A positive result means a buy signal, a zero result means hold, and a negative result means sell. Through trials with different weighting on each indicator, this combination returned the best performance as shown in Figure 1. This is an effective strategy because it looks for potential overbought / oversold price points and price points where an upward / downward trending is formulating.

The performance of the Manual Strategy is significantly better than that of the Benchmark. Benchmark's performance is solely based on the price of the given stock JPM due to its initial long and hold position. On the other hand, the Manual Strategy enters long, short, hold positions multiple times based on indicator sums calculated as the formula shown above. Therefore it is able to capture price uptrends / oversold positions to long and price downtreads / overbought positions to short.

By comparing the performance of the Manual Strategy on in-sample against out-of-sample data, it is clear that the in-sample run shows a better performance since the weights on each individual indicator are curated to the in-sample dataset. When used to trade with a different dataset, this weight-based indicator sum does not drive the best performance.

Table 1—Manual Strategy vs Benchmark Metrics In Sample

Name	Manual Strategy	Benchmark
Cumulative Returns	1.0649	0.0123
STD of Daily Returns	0.0108	0.0170
Avg of Daily Returns	0.0015	0.0002

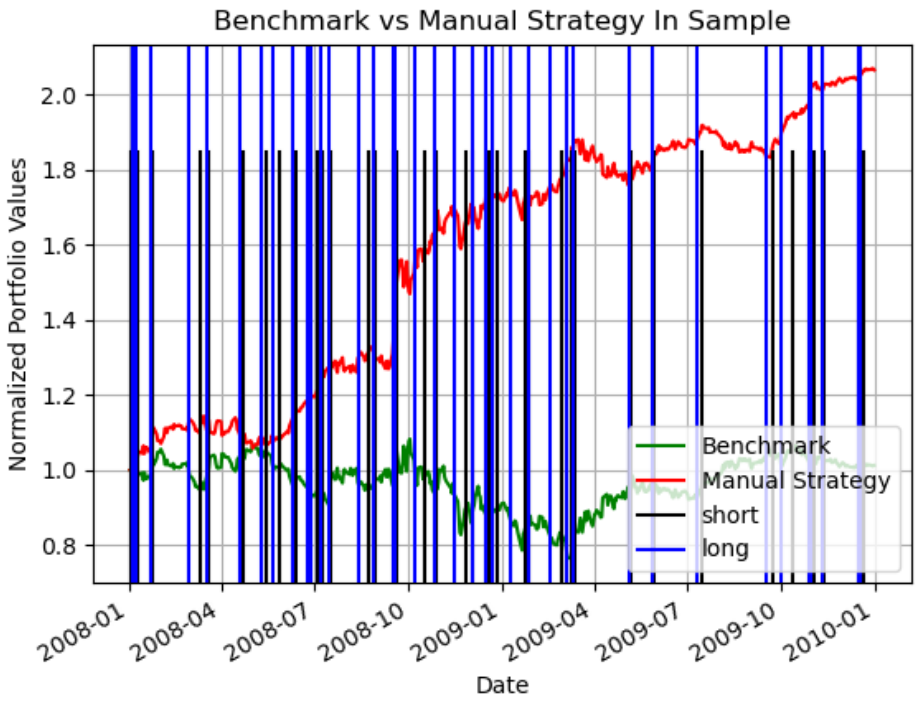

Figure 1—Benchmark vs Manual Strategy In Sample

Table 2—Manual Strategy vs Benchmark Metrics Out of Sample

Name	Manual Strategy	Benchmark
Cumulative Returns	0.4211	-0.0836
STD of Daily Returns	0.0069	0.0085
Avg of Daily Returns	0.0007	-0.0001

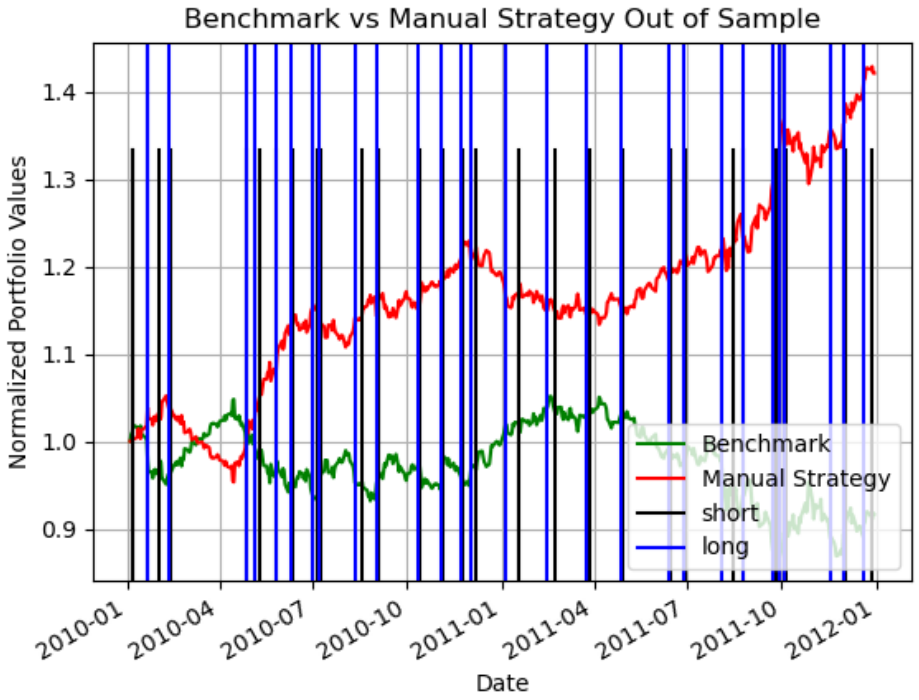

Figure 2—Benchmark vs Manual Strategy Out of Sample

4 Strategy Learner

QLearner is used to implement the Strategy Learner in this project. In order to abstract a problem that a QLearner can solve from a trading problem, state, action, reward are the necessary variables. Training of the learner is done in a loop fashion. Each epoch is a complete iteration through the in-sample data. The portfolio value is captured at the end of each epoch and compared to that from

the previous. If the portfolio values are equal then a convergence occurs and the loop ends. A timeout of maximum 20 epoch is added to prevent an endless loop.

The QLearner parameters are alpha (learning rate) 0.2, gamma (discount rate) 0.9, rar (random action rate) 0.98, radr (random action decay rate) 0.9. radr is tuned to be lower so that random actions can be avoided earlier in the training phase for a shorter convergence time.

4.1 State

The state is determined by the number of indicators and their possible number of permutations. Since each of the indicator can suggest buy / sell / hold actions, total number of possible states is

$$3^4 = 81$$

Buy / sell / hold are represented as 0, 1, 2. Each permutation is discretized as below which results in a number between 0 to 80

$$price/SMA + Percent\ B\ \&\ Momentum * 3 + Golden/Death\ Cross * 9 + MACD * 27$$

4.2 Action

The action is fairly straightforward, sell / hold / buy each is represented as 0 / 1 / 2.

4.3 Reward

Reward is essentially the return of all holdings on a given day. A long position with a price increase means a positive reward and a long position with a price decrease means a negative reward. A short position with a price decrease means a positive reward and a short position with a price increase means a negative reward

5 Experiment 1

In this experiment, the performance of the Benchmark, Manual Strategy and Strategy Learner on in-sample data are compared in Figure 3. The hypothesis is that the Strategy Learner will perform the best, Manual Strategy the second, and the Benchmark the last. The Benchmark and Manual Strategy normalized portfolio is generated with JPM 2008/01/01 - 2009/12/31. The Strategy Learner is trained and tested with the same dataset. From Figure 3, Benchmark shows the

worst performance but Manual Strategy beats Strategy Learner. Strategy Learner is designed as a more general approach to generate orders for an optimal portfolio, whereas the Manual Strategy is designed to specifically optimize a portfolio in a given period of time. Therefore it is reasonable that in some cases the in-sample performance favors the Manual Strategy over the Strategy Learner. The weighting system is a good example of such targeted optimization.

I'd generally expect the same result every time with in-sample data because of the above mentioned reason

Figure 3—Benchmark vs Manual Strategy vs Strategy Learner In Sample

6 EXPERIMENT 2

In this experiment, Strategy Learner is trained with 3 different impact values 0, 0.05, 0.1. Impact is the percent cost at the price of a buy of sell transaction. An impact of 0.05 indicates that at each transaction, there is an extra cost of 0.05 * number of bought shares * 0.05. In the Strategy Learner, the effect of impact is

added by adding the extra cost to (subtracting it from) a reward caused by any buy / sell action. Therefore the hypothesis is that the total number of trades will decrease as impact increases. Besides, the decrease in the number of trades will lower the cumulative return.

From Table 3, we can see that cumulative return indeed decreases as the value of impact increases. From Figure 4, 5, 6, the black vertical lines represent each trade. It is clear that there are fewer trades as the value of impact increases.

Table 3—Manual Strategy vs Benchmark Metrics Out of Sample

Name	**Impact = 0**	**Impact = 0.05**	**Impact = 0.1**
Cumulative Returns	0.6385	0.1708	0.022

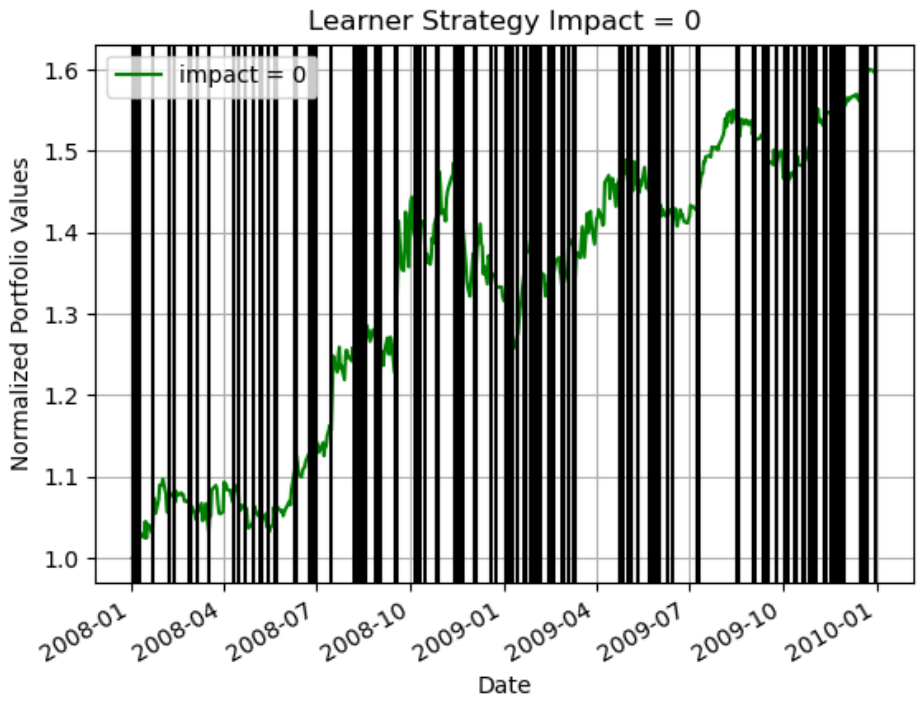

Figure 4—Learner Strategy Impact = 0

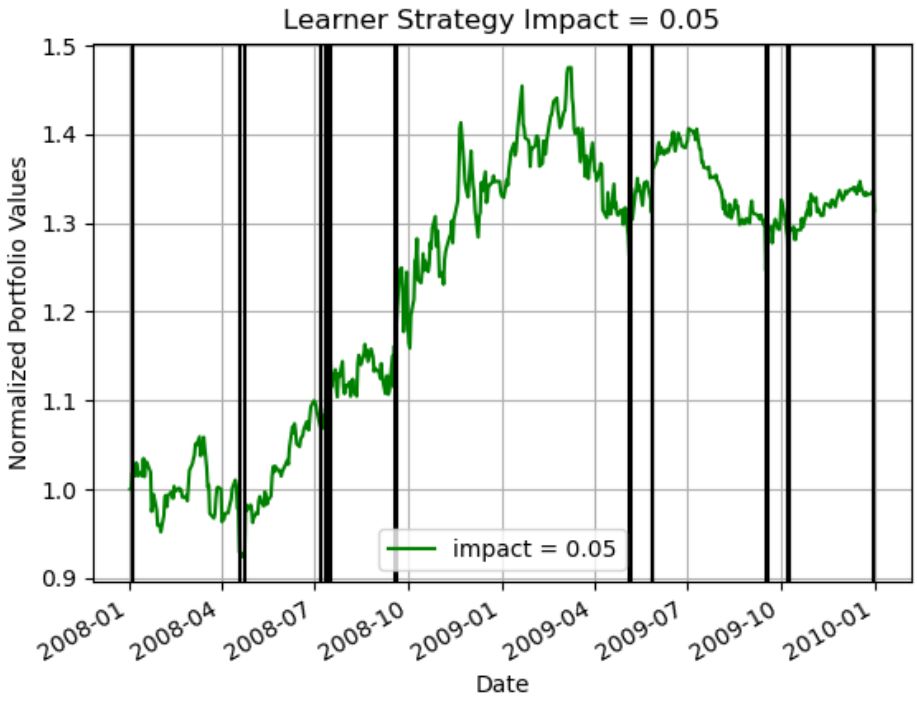

Figure 5—Learner Strategy Impact = 0.05

Figure 6—Learner Strategy Impact = 0.1