CS7646 Project8: strategy evaluation

Anlu Zhou
azhou90@gatech.edu

1 INTRODUCTION

This This project analyzed the trading strategies according to selected indicators' signals. Holding constrains are long 1000 shares, short 1000 shares, or 0 share.

For manual strategy, it decided to buy or sell when some intuition-based conditions matched by giving thresholds to indicators. Benchmark is buying 1000 shares at the start, hold them for the entire trading period, and sell these shares at the end of the period. Manual strategy would be compared with benchmark for both in-sample period and out-sample period. My hypothesis is that intuition-based manual strategy would beat the benchmark for in-sample period. But probability has similar performance with benchmark for out-sample period.

Strategy learner use a random forest learner to determine buy or sell according to the value of selected indicators. It's compared with manual strategy and benchmark. My hypothesis is that strategy learner would beat both intuition-based manual strategy and the benchmark for in-sample period. Additionally, by setting the transaction cost (commission) to 0, the influence of impact to the trading strategies was also analyzed in this project. My hypothesis is that the increasing of impact would hamper the strategy learner's performance.

2 INDICATOR OVERVIEW

Following indicators were used for both manual strategy and strategy learner. The rolling window for momentum, simple moving average (SMA) and Bollinger Bands Percentage are both set to 20 days (discussed in manual strategy part).

2.1 Momentum

Momentum looks at the price changes within selected lookback period. Momentum is calculated as following

$$Momentum = \frac{P_t - P_{t-n}}{P_{t-n}}$$

P_{t-n}: the price at time t-n

P_t: the price at time t

In both Manual strategy and strategy learner, momentum value calculated as above would be used as an indicator.

2.2 Price/Simple moving average (Price/SMA)

By selecting n-days window, n-days SMA will calculate the average of stock price for these n consecutive days, and these stock prices are equally weighted when averaging. The SMA is calculated as following

$$SMA = \frac{P_1+P_2+\cdots+P_n}{n}$$

P_n : the price for period n
n : the number of periods (selected)

In both Manual strategy and strategy learner, Price/SMA calculated as above is used as an indicator.

2.3 Bollinger Bands Percent (%B)

We use a standard deviation of stock prices stand for prices' volatility when we are looking at price excursion. And use Bollinger Brands to see how far the price goes outside the bands.

Bollinger Bands %B at a particular day t is calculated as following

$$BBP_t = \frac{P_t - SMA_t}{2 * stdev_t}$$

P_t : price at t
SMA_t: SMA at t
$stdev_t$: rolling standard deviation at t

In both Manual strategy and strategy learner, Bollinger Bands Percent (%B) calculated as above would be used as an indicator.

2.4 Moving average convergence divergence (MACD)[1]

Exponential moving average (EMA) weight more on prices of the most recent dates than EMA, And MACD shows the difference between two different period

EMA, MACD is positive when 12-period EMA is above 26-period EMA. and a 9-period EMA is the signal line of MACD. MACD is calculated as following

$$MACD = 12 - period\ EMA - 26 - period\ EMA$$

In both Manual strategy and strategy learner, MACD -signal is used as an indicator.

3 MANUAL STRATEGY

For manual strategy, I used 4 indicators described in above part: 1. momentum; 2. Price/SMA; 3. Bollinger band percent(%B) 4. MACD – signal. And I use "OR" condition for every indicator: this means if at least one indicator meet conditions under thresholds constrains, the strategy will suggest a buy or sell signal.
Usually[2], typical rolling window values are 10-day as short term; 20-day as medium term; 50-day as long term. Therefore, I chose a number 20-day window to perform a medium-term indicator analysis, to avoid signal changing too frequently if use short term, or smooth out signals if use long term

3.1 Momentum

Usually, when there is a momentum with a negative value, the trend means the price is going down, and indicate a selling signal. And when there is a momentum with a positive value, the trend means the price is going up, and indicate a buying signal. But this method usually combined with other indicators, such as looking at price excursion.

My intuition-based momentum thresholds considered momentum alone, so it treats value between -0.1 and 0.1 as a randomly fluctuation and not regard it as a trading signal.

Since price fluctuated frequently, when momentum is larger than 0.1, this could be trading at a level above its intrinsic or fair value. Therefore, the price would drop and it's a signal for a short position

when momentum is less than -0.1, this could be caused by this stock has traded lower in price and has the potential for a price bounce. Therefore, it's a signal for a long position

3.2 Price/SMA

This is an indicator that consider the comparison between asset price and SMA.

When price/SMA is less than 0.4, this means that the stock price is less that 0.4*SMA, and the price would increase to meet the SMA line, it's a signal for a long position.

When price/SMA is larger than 1.2, this means that the stock price is larger than 1.2*SMA, and the price would drop to meet the SMA line, it's a signal for a short position.

3.3 Bollinger Bands Percent (%B)

For Bollinger Brands, we first calculate the rolling standard deviation, then add 2 standard deviations both above and below the SMA. Therefore, we can use these two bands to see when the price goes outside and retrieve back the bands before we make a trading decision.

When B% is less than -0.6, it's might indicate the SMA just the price goes outside the lower band and retrieve back the bands, the price is going to increase, it's a signal for a long position.

When B% is larger 0.6, it's might indicate the SMA just the price goes outside the upper band and retrieve back the bands, the price is going to drop, it's a signal for a short position.

3.4 MACD – signal

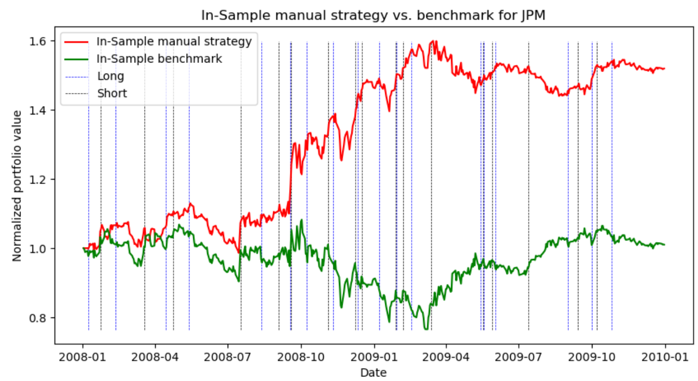
For MACD, When MACD cross above its signal line, means the price is going to increase, it's a buying signal. When MACD cross below its signal line, means the price is going to decrease, it's a selling signal.

Therefore, when MACD – signal is larger than 0.005, this indicate MACD cross above its signal line, the price is going to increase, it's a signal for a long position.

when MACD – signal is less than -0.005, MACD cross below its signal line, the price is going to drop, it's a signal for a short position.

3.5 In-sample manual strategy vs. benchmark for JPM:

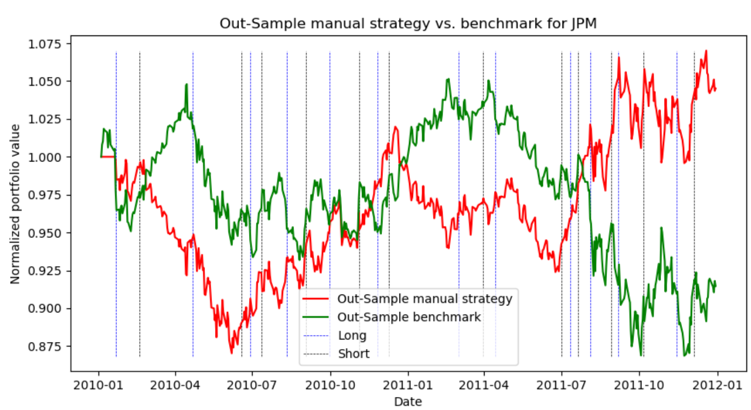
start value is $100,000 cash, investment period is from 2008/1/1 to 2009/12/31, commission =9.95, impact = 0.005



From the figure above, the manual strategy performs much better than benchmark for JPM in In-sample period. This means the conditions set with the indicators to produce trading signals are work well in In-sample period: The long positions of manual strategy are almost at the price increasing points, such as the points at around October 2008. And short positions of manual strategy are almost at the price dropping points, such as the points at around March 2009.

3.6 Out-sample manual strategy vs. benchmark for JPM:

start value is $100,000 cash, investment period is from 2010/1/1 to 2011/12/31, commission =9.95, impact = 0.005



From the figure above, the manual strategy's performance in Out-sample period is not as good as the performance for JPM in In-sample period. This means the conditions set with the indicators to produce trading signals only work in In-sample period, because I tweak the indicator threshold value only according to

In-sample performance: The long positions of manual strategy are not at the price increasing points, such as the points at around May 2010. And short positions of manual strategy are not at the price dropping points, such as the points at around August 2010.

3.7 Compare In-sample and Out-sample performance

Investment Period	Strategy	Cumulated return	Standard deviation	mean	Sharpe ratio
In-sample	manual strategy	0.51942	0.01246	0.00091	1.15648
In-sample	benchmark	0.01014	0.01704	0.00016	0.15339
Out-sample	manual strategy	0.04514	0.00798	0.00012	0.23783
Out-sample	benchmark	-0.08531	0.00850	-0.0014	-0.26362

From the table, In In-sample period, the manual strategy outperformance with 0.51942 cumulated return and 0.00091 mean daily return than benchmark with 0.01014 cumulated return and 0.00016 mean daily return, because the indicators help the strategy buy when price increasing and sell when price decreasing in In-sample. manual strategy's standard deviation is close to benchmark because these standard deviations mostly come from the price randomly fluctuation, but manual strategy's standard deviation is a little lower due to strategies. And manual strategy's Sharpe ratio is larger than benchmark due to higher return. However, the manual strategy's cumulated return Out-sample is 0.04514 less than 0.1*In-sample cumulated return, the manual strategy's mean daily return Out-sample is 0.00012 about one-tens In-sample cumulated return. Indicate that the indicators tweaked In-sample are not suitable for Out-sample price. Indicators are not fixed for every condition.

4 STRATEGY LEARNER

4.1 Steps to frame the trading problem as learning problem

4.1.1 *add_evidence ():*

- Create the training data frame
    - X_train: calculated the same 4 indicators' value that used in manual strategy for In-sample 1. momentum; 2. Price/SMA; 3. Bollinger band percent(%B) 4. MACD – signal. Combine them as a X_train data frame
    - n-days return: calculated the In-sample n-days return

  - Y_train: based on the In-sample n-days return, set upper and lower thresholds together with impact value (default 0.005) to determine whether to buy or sell
- Train the learner
  - When call add_evidence(), the random forest learner will be trained

4.1.2 *test_policy()*

- Create the X_test: calculated the same 4 indicators' value that used in manual strategy for Out-sample 1. momentum; 2. Price/SMA; 3. Bollinger band percent(%B) 4. MACD – signal. Combine them as a X_test data frame
- Predicting
  - Given X_test, the learner will predict the best position Y_pred: long=1; short=-1; out=0
- Output data frame
  - According to Y_pred, create a data frame with the trading straegy: buy with positive value, sell with negative value, the absolute value is the shares for trading

4.2 Hyperparameters

For the Strategy Learner's hyperparameter, I used a random forest with random tree classification and set the bag size to 25 and leaf size to 5. Since tree-based classifications usually need deal with overfitting problems, thus a larger bag size and not too small leaf size could help to get rid of overfitting.

For Indicators' rolling window, I used 20 days to keep consistency with it in the manual strategy. The reason is stated in manual strategy part

For n-days return, I choose 7 days, because smaller days would capture too many random fluctuations and give false positive. And too larger days would not capture the price change tendency.

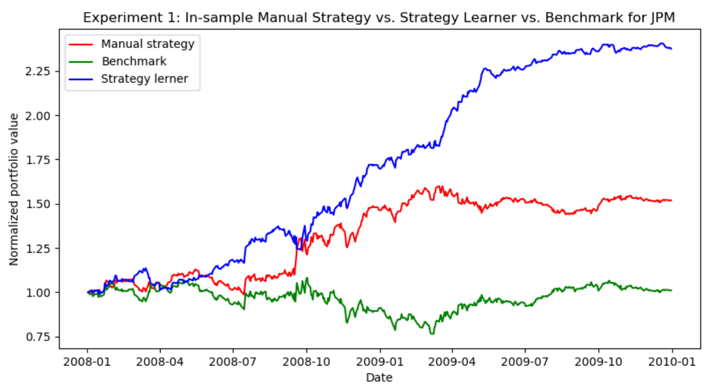
4.3 Adjust data

I didn't adjust the indicators data or price return data because tree-based models can be immune to redundant variables or variables with high correlation which may lead to overfitting in other learning algorithms. Trees also have very few parameters to tune for when training the model and performs relatively well with outliers or missing values in a dataset.

5 EXPERIMENT 1

In this experiment, I compared the In-sample performance for Strategy learner, manual strategy, and benchmark for JPM.

My hypothesis: the strategy learner will have better performance than manual strategy



start value is \$100,000 cash, investment period is from 2008/1/1 to 2009/12/31, commission =9.95, impact = 0.005

Investment Period	Strategy	Cumulated return	Standard deviation	mean	Sharpe ratio
In-sample	manual strategy	0.51942	0.01246	0.00091	1.15648
In-sample	benchmark	0.01014	0.01704	0.00016	0.15339
In-sample	Strategy learner	1.37459	0.01037	0.00177	2.71151

From the table, In In-sample period, the strategy learner outperformance with 1.37459cumulated return and 0.00177mean daily return than manual strategy with 0.51942cumulated return and 0.00091mean daily return, because the learner helps the strategy buy when price increasing and sell according to n-days return. standard deviations are close because these standard deviations mostly come from the price randomly fluctuation, but strategy learner standard deviation is a little lower due to strategies. And strategy learner Sharpe ratio is larger than manual strategy due to higher return.

6 EXPERIMENT 2

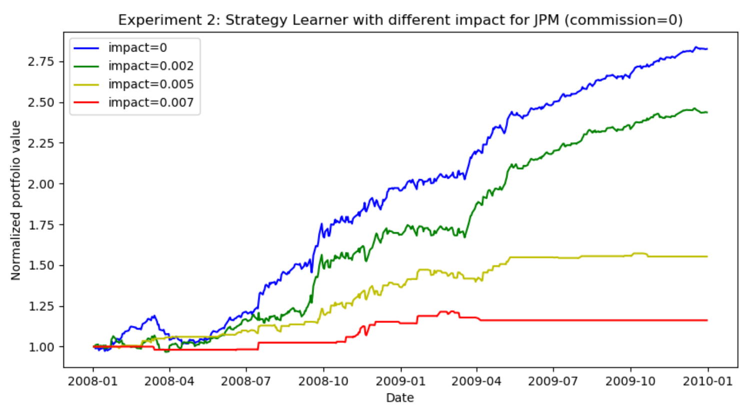
In this experiment, I compared the learner's In-sample performance under different impact with 0 commission.

Impact for 4 different strategy learners: 0, 0.002, 0.005, 0.007

Inside strategy learner, when it decided to buy or sell, it considered YBUY + impact and YSELL - impact as thresholds.

My hypothesis: larger impact will hamper the strategy learner to provide trading signals, the learner tends to trade less, and the performance of the learner will be worse as the impact getting larger. The portfolio value will also be influence by impact because the investment would earn less with larger impact

start value is $100,000 cash, investment period is from 2008/1/1 to 2009/12/31, commission =0.



From the figure above, strategy with smaller impact tend to have better performance, and when impact is 0.007 the strategy learner tends to leave the market. The portfolio line become horizontal because it might be a out position without any long or short, just cash.

Investment Period	Impact	Cumulated return	Standard deviation	mean	Sharpe ratio
In-sample	0	1.8253	0.00975	0.00211	3.43712
In-sample	0.002	1.43611	0.0107	0.00183	2.70765
In-sample	0.005	0.55231	0.0074	0.0009	1.93182
In-sample	0.007	0.16132	0.00479	0.0003	1.02181

From the table above, the larger impact, the less cumulated return, mean daily return and Sharpe ratio. As hypothesis state larger impact will hamper the strategy learner to provide trading signals, the learner tends to trade less, and the performance of the learner will be worse as the impact getting larger.

Investment Period	Impact	Cumulated transaction numbers
In-sample	0	168
In-sample	0.002	179
In-sample	0.005	122
In-sample	0.007	36

From the table above, compare with impact 0, impact of 0.02 didn't increase, this might be caused by strategy due to the stock returns. But when the impact continuously increases, the strategy learner's transaction number decrease from 179 to 36, mean the learner tend to avoid giving trading signals when impact getting larger.

7 REFERENCES

1. Appel, G. (2003). Become your own technical analyst: How to identify significant market turning points using the moving average convergence-divergence indicator or macd. *The Journal of Wealth Management*, 6(1), 27-36.

2. Williams, O. (2006). Empirical optimization of Bollinger Bands for profitability. Available at *SSRN* 2321140.