PROJECT 8: STRATEGY EVALUATION

Malcolm Nathaniel Ng Bao Kun
mkun6@gatech.edu

1 Introduction

The report looks into comparing the performance of a Manual Strategy (MS) vs Strategy Learner (SL) , which determines trading behavior based on an AI model. The initial hypothesis on the paper is that the AI model is capable of performing better trades than the MS.

2 Indicators Overview

The 4 indicators used to devise the strategies are Momentum, Bollinger ™ Bands% (BB%), Moving Average Convergence Divergence (MACD) and Relative Strength Index (RSI).

Momentum indicators are used to determine the strength or weakness of a stock's price changes. Bollinger™ Bands (BB) introduce +/- 2 stdev away from a SMA(20) of a stock's price that comprises the upper band and lower band, we use the BB% variation to fit our use case. The RSI is a momentum oscillator that measures the speed and change of price movements between 0 and 100. The MACD oscillator is a simple and effective momentum indicator that also can be used for trend following.

For Manual Strategy, all the indicators will return -1, 0, or 1, corresponding to a "short," "out" or "long" position. For Strategy Learner, the indicators will return the raw values dataframe and use it to train the model. For both strategies the Momentum and RSI, a 14 days lookback period is used as default optimization. For BB%, a 20 days lookback period is used as default as well. Lastly, the MACD uses EMA(12,26) to generate the MACD line while EMA(9) for the signal line.

Table 1—Indicator rules overview

	Short (Sell)	**Out (Hold)**	**Long (Buy)**
Momentum	> 0.1	0	< 0.1
BB%	< 0	0	> 1
MACD	Signal line > MACD line	0	MACD line > Signal line
RSI	> 70	> 30 & < 70	< 30

3 Manual Strategy

For the MS, generating the overall signal was a trial and error exercise. My initial trading strategy was to combine MACD, RSI and BB% to generate a buy/sell signal with all 1 indicating a enter position and -1 to exit. The cumulative returns of the portfolio was not ideal, due to the hard rule applied there was less than 3 trades during the trading period. In order to introduce more trades, a more flexible rule based approach is applied, and by adding another indicator it gives the strategy more leeway to confirm a long or short position.

The final combined indicator consists of 3 momentum indicators namely RSI, Momentum and MACD which combines trend and momentum. And BB%, a security's volatility indicator. I believe this is an effective strategy as the model gives flexibility to use either the price momentum or the recent change of price like RSI to determine if it is a potential long/short opening. And with a buy/sell signal from the MACD and BB% volatility giving a more confirmed signal.

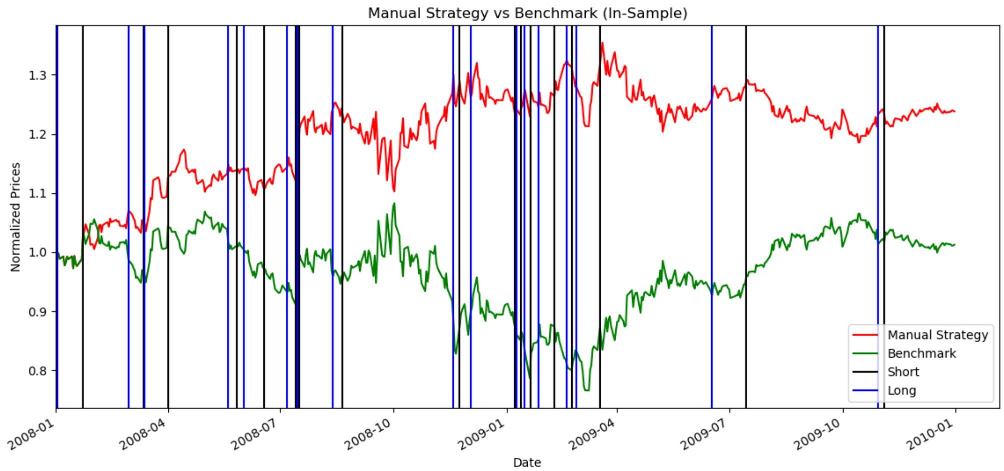

Figure 1a — JPM Normalized Price Manual Strategy vs Benchmark In sample, Time period between Jan 2008 to Dec 2009

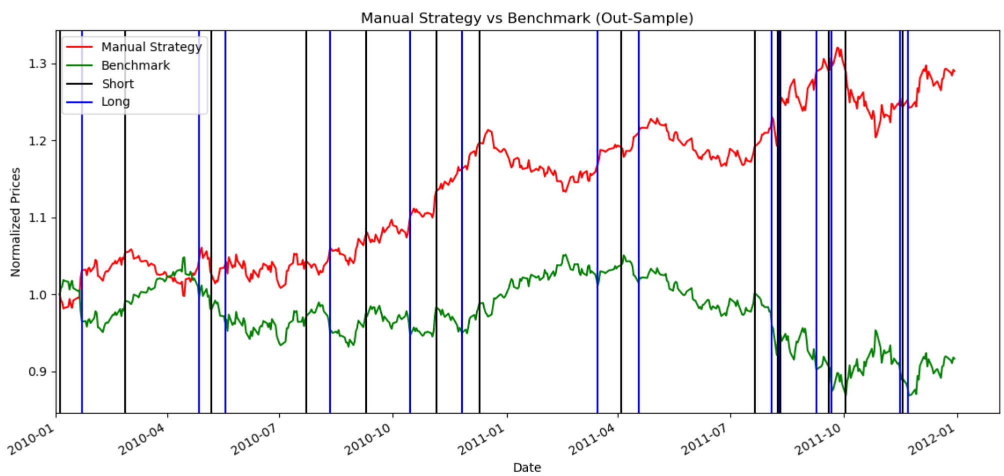

Figure 1b — JPM Normalized Price Manual Strategy vs Benchmark Out sample, Time period between Jan 2010 to Dec 2012

For the performance of the MS in both the in-sample and out-of-sample time periods performed significantly better than the benchmark. Also, we noted that the long/short positions were rather accurate to some extent with clear uptrend/downtrend in the trading period for both samples. An example of a good and bad trade, if we zoom in on the in-sample period 2008-09 a short position was taken and price did follow a downtrend but the MS did not put a long position on 2008-10 during the recovery. It is interesting to note that 2008 to

2009 was during the aftermath of the financial crisis, but we see that technical analysis can be an excellent tool to navigate us during a turbulent climate.

Table 2—Manual Strategy and benchmark performance metrics

	Manual Strategy	**Benchmark**
Cumulative returns (In-sample, Out-of-sample)	0.2379, 0.2898	0.0123, -0.0836
Mean of daily returns (In-sample, Out-of-sample)	0.0005, 0.0005	0.0002, -0.0001
Stdev of daily returns (In-sample, Out-of-sample)	0.0133, 0.0069	0.01704, 0.0085

The cumulative returns (cr) of the MS in-sample and out-of-sample far outperform the benchmark by 1834% and 247% respectively. Also to note that the out-of-sample's cr is negative which tells us the overall portfolio is losing money.

The mean of the daily returns of the MS is positive indicating healthy daily returns while the benchmark for in-sample is lower and out-of-sample is negative indicating negative returns.

The stdev of daily returns of the MS and benchmark is < 1 which can be considered low variance where most of the data points are clustered around the mean.

4 Strategy Learner

As aforementioned at the indicators overview, in order to train a classification learner the indicators return values must not be -1,0,1. We have employed the use of BagLearner with RTLearner which is essentially a Random Forest (RF) algorithm. In order to frame the trading problem, it is necessary to convert the existing decision tree regression learner to a classification learner.

4.1 Training & Testing

First, we tune the RF algorithm to use mode instead of mean to train and test the model. Secondly, we introduce a parameter known as strategy to the indicators to return a dataframe of X features in their raw values. Thirdly, we build the Y (predicted value) of the features based on N day returns of the stock price. We note that the X features are based on current day value while the Y is derived from the future value of the data. We tuned the N day return to be 14 days because it is a decent trade range that is not too far into the future and not too short term. Fourthly, in order to build Y we need to determine the YBUY and YSELL. We use the Numpy percentile function to compute the q-th percentile of the N day return values. We tuned the YBUY to 75th percentile while YSELL to the 25th percentile to be optimal for the learner. The use of the percentile function is perfect for interpolation of the return values as it is statistically unknown to us what is a performant YBUY and YSELL value. So to generate a buy signal the current N day return value has to > 1 while a sell < -1. Lastly, we will pass the training datasets X & Y to the BagLearner.

Once the SL is trained, we will use the learner to predict the buy/sell signals by introducing the testing dataset X features and finally build the trade order dataframe.

4.2 Hyperparameter of Strategy Learner

For the RF algorithm, the RTLearner uses a leaf size of 5 to avoid overfitting the in-sample. For the BagLearner we used bagging of 5. The decision for these parameters were based on the performance of the SL when tested on the out-of-sample. We found that when bagging >5 performance starts to deteriorate. Also not tuning the leaf size to higher as we are bagging with our RT so overfitting of individual trees is less concerning.

5 Experiment 1 (Manual Strategy / Strategy Learner)

Experiment 1 attempts to evaluate the performance between a MS vs SL and the benchmark. The initial experimental hypothesis is that the SL portfolio will outperform both the benchmark and MS portfolios. In order to ensure fairness and repeatability, there is no tuning of the indicator parameters such as lookback period or using a set of different indicators. The indicator overview above will give a good overview of the trading strategy for both strategies. The in-sample period is January 1, 2008 to December 31, 2009. We will be using JPM in-sample start and end datetime for both strategies. Also, the transaction costs for all strategies are 0.005 impact and 9.95 commission per trade. Finally we will obtain the normalized portfolio value, cumulative returns, average daily returns and stdev of daily returns metrics for evaluation.

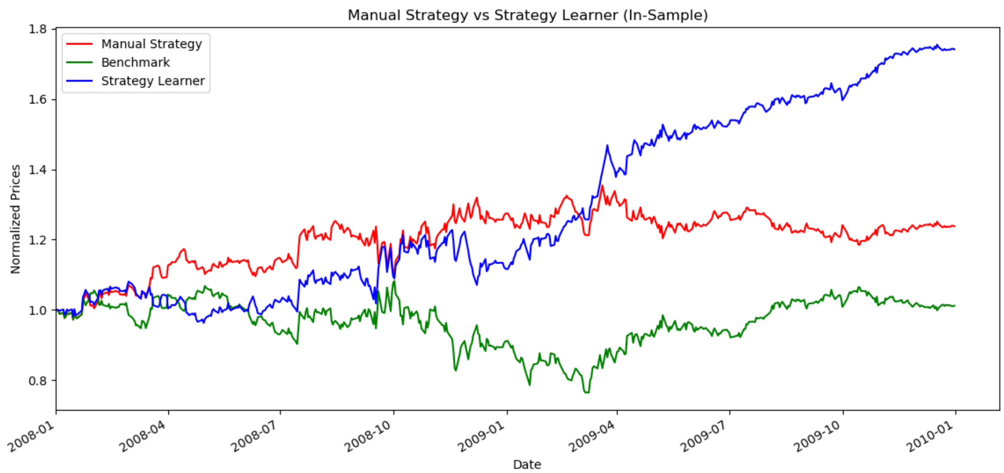

Figure 2— JPM Normalized Price Manual Strategy vs Strategy Learner in-sample, Time period between Jan 2008 to Dec 2010

From figure 2, The SL have clearly outperformed both the MS and benchmark significantly. SL's cumulative returns outshine MS by ~211% and the mean of the daily returns doubles of MS. When we put SL against the benchmark the cumulative returns are astonishing with over ~5922% of the benchmark (Table 2).

Table 3—Manual Strategy and Strategy Learner performance metrics

	Manual Strategy	**Strategy Learner**
Cumulative returns	0.2379	0.7407
Mean of daily returns	0.0005	0.0012
Stdev of daily returns	0.0133	0.01352

We can expect this relative result every time with in-sample data as both the training and testing set is using the same dataset so it is more likely for the trading problem to be generalized and result in overfitting and will perform well during testing. However, given a set of well optimized parameters for the indicators we can expect a decent performance when tested over a different testing set.

6 Experiment 2 (Strategy Learner)

Experiment 2 attempts to study the effects of impact value changes with in-sample trading behavior. Our initial hypothesis for this experiment is that when impact is lower the portfolio performance is better. In order to reduce the randomness effect, we have conducted 20 trials with varying impact values [0.0, 0.005, 0.01]. In order to produce the result, each learner will be re-trained and tested with different impact values while the commission stays constant at 0.0. We will collect the relevant portfolio statistics to discuss our findings.

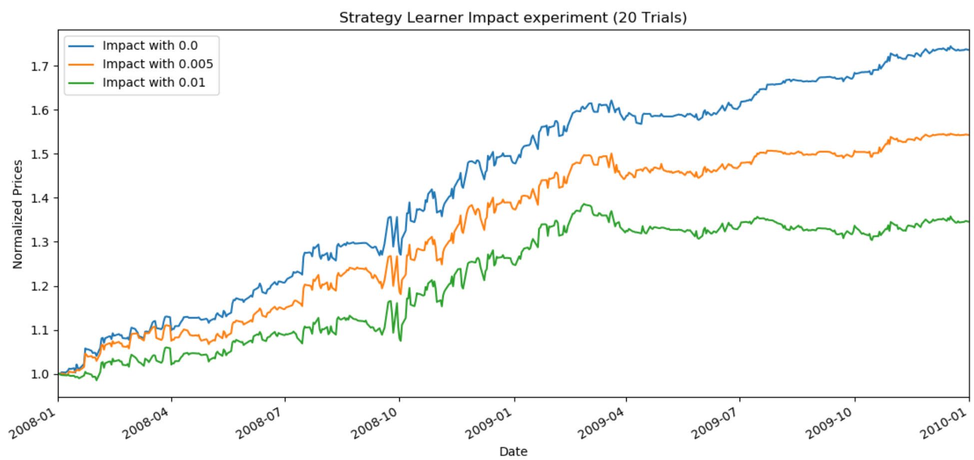

Figure 3— JPM Normalized Price in-sample impact analysis with 20 trials, Time period between Jan 2008 to Dec 2010

Table 4—Strategy Learner Impact performance metrics

	Impact 0.0	**Impact 0.005**	**Impact 0.01**
Cumulative returns	0.7357	0.5426	0.3457
Mean of daily returns	0.0011	0.0009	0.0006
Stdev of daily returns	0.0119	0.0126	0.0138

The pattern is pretty obvious from the findings. In figure 3, we notice that the highest portfolio value belongs to the one with the least impact 0.0 while the lowest portfolio value belongs to the one with the highest impact of 0.01. In table 4, the portfolio metrics also give us a more conclusive pattern that the portfolio with the least impact of 0.0 results in the best cumulative returns and mean of daily returns. While, similarly the portfolio with highest impact of 0.01 resulted in the lowest cumulative returns and mean of daily returns. We also notice that the stdev of daily returns for the least impact 0.0 has the least volatility and variance and an increasing volatility and variance when the impact value increases. This also means that the daily returns are further away from the mean when stdev is higher. However, stdev averages for all the impact values remain insignificant.

In conclusion, we can say that the impact value during trades can have a significant impact on the portfolio performance. And a lower impact value is usually preferred. In reality, depending on trade sizes the a trade can encounter high price impact when the pool is illiquid and it depends on the available liquidity in the market.