CS 7646 Machine Learning for Trading
Project 8: Strategy Evaluation Report

Renyu Zhang
rzhang605@gatech.edu

1 INTRODUCTION

In this project, a manual rule-based strategy and a strategy learner have been implemented to generate trades data frame and evaluate their performances. The Manual Strategy combines indicators with a set of rules to create an overall signal for trading. The Strategy Learner is a classification-based learner using Random Forest. Both strategies use three (3) technical indicators that have been developed in the project 6. In addition, two experiments have been conducted to compare the performance of strategies. The first hypothesis is that the Strategy Learner outperforms the Manual Strategy during the in-sample period. Another hypothesis is that a higher impact could negatively affect the performance of the Strategy Learner during the in-sample period.

2 INDICATOR OVERVIEW

Three technical indicators have been selected to build the manual rule-based strategy and the strategy learner. They are price to simple moving average ratio (price/SMA), Bollinger Bands percent (BB%), and price to exponential moving average (price/EMA).

2.1 Price to Simple Moving Average Ratio

Price to SMA ratio represents the prices of a given stock relative to its simple moving average (SMA). SMA has a lookback period of n days (i.e., n-day window) to calculate the average of stock price values. The mathematical equation of SMA is shown below:

$$\text{SMA}[t] = \frac{\text{Price}[t]}{\text{Price}[t-n:t].\text{mean}()} - 1$$

After tests and trials, I found that the days parameter should be set to 15 for the price/SMA ratio indicator leading to the best performance of Manual Strategy, and the days parameter should be 5 for the price/SMA ratio indicator using the Strategy Learner.

2.2 Bollinger Bands Percent

The Bollinger Bands percent (BB%) is a technical indicator derived from Bollinger Bands. It is calculated by the following formula:

BB% = (price – bottom band) / (top band – bottom band)

The top and bottom bands are calculated by adding a band above and below two standard deviations from the SMA as discussed above.

top band = SMA + 2 * standard deviation

bottom band = SMA - 2 * standard deviation

After the optimization, I found that the days parameter should be set to 15 for the BB% indicator resulting in the best performance of Manual Strategy, and the days parameter should be 20 for the BB% indicator in the Strategy Learner.

2.3 Price to Exponential Moving Average

The price to exponential moving average (price/EMA) ratio is chosen as a technical indicator in this project. It measures the prices of a given stock relative to its moving average. EMA is similar to SMA regarding the average stock price over a lookback period of n days (i.e., n-day window). However, EMA focuses more on recent price data and reacts quicker to price changes than SMA. The implementation of EMA using the Pandas library is shown below:

```
ema = price_df.ewm(span = n).mean()
```

After the optimization, I found that the days parameter should be set to 15 for the price/EMA ratio indicator using the Manual Strategy, and the days parameter should be 5 for the price/EMA ratio indicator in the Strategy Learner.

3 MANUAL STRATEGY

The Manual Strategy follows a set of rules and using the three (3) technical indicators discussed above to generate a data frame of trades. The indicators determine the long or short opportunities based on past stock prices. It would be an effective strategy because it identifies the direction of a current price trend, confirms the significant market moves, reflects the stock volatility, and predicts long-term price movements.

3.1 Methodology

To implement the Manual Strategy, we trade the stock JPM during the in-sample period from January 1, 2008 to December 31, 2009 and the out-of-sample period from January 1, 2010 to December 31, 2011. The starting cash value is $100,000.00, and each trade has a commission for $9.95 and an impact of 0.005. The allowable positions are long 1000 shares, short 1000 share, or do nothing (holding cash).

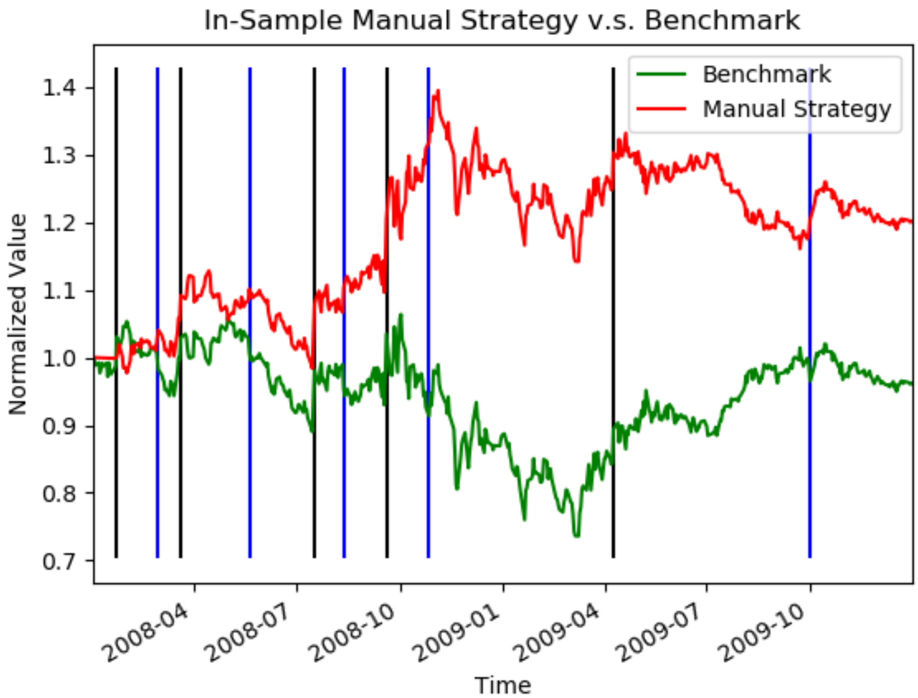
The Manual Strategy utilizes the following rules to long or short stocks:

- Long signal:
  When the ratio is less than 0.95, it means that the stock price is lower than the average price, showing a signal to long the stock.
  When BB% is less than zero (0), it indicates that the stock price is lower than the average price. The stock is oversold, showing a signal to long the stock.
  When the EMA ratio is less than 0.95, it means that the stock price is lower than the average. The stock is oversold, showing a signal to long the stock.
- Short signal:
  When the price/SMA ratio is greater than 1.05, the stock is overbought. The current stock price is too high; thus, it is a good time to short the stock.
  When BB% is greater than 1, the stock is overbought. The current stock price is too high; thus, it is a good time to short the stock.
  When the EMA ratio is greater than 1.1, the stock is overbought. The current stock price is too high; thus, it is a good time to short the stock.

3.2 Comparative Analysis

To evaluate the performance of the Manual Strategy, a benchmark portfolio has been created. It starts with $100,000 cash and invests in 1000 shares of JPM on the first day of trading, and then holds that position until the end of the trading period. The following metrics table and chart show the comparison between the Manual Strategy and the benchmark performance.

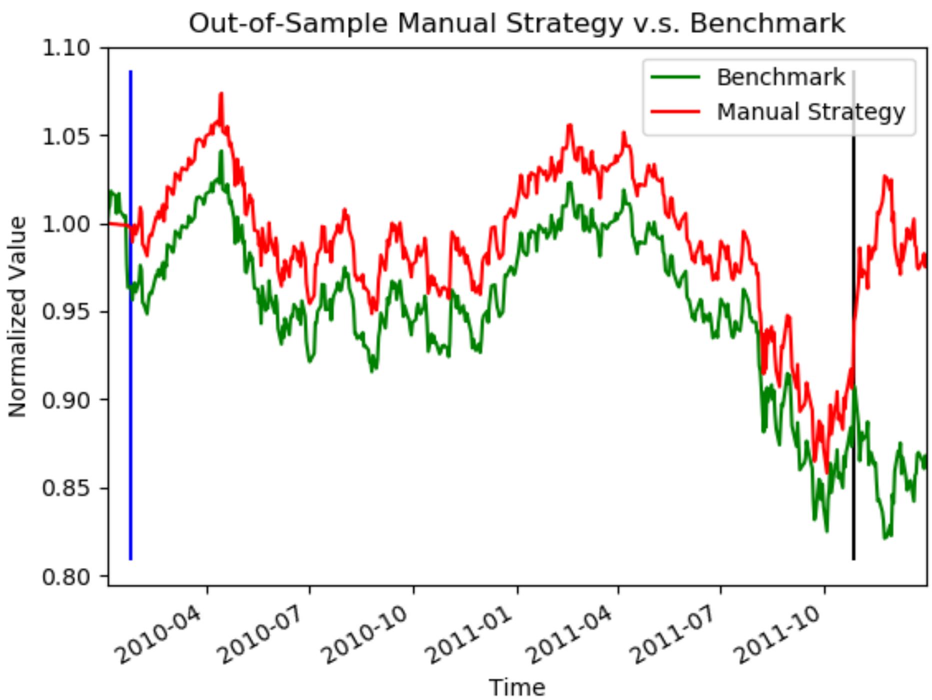
Performance metrics	Benchmark	Manual Strategy (in-sample)
Sharpe ratio	0.0683	0.5382
Cumulative returns	-0.0379	0.2015
Standard deviation of daily returns	0.0175	0.0134
Mean of daily returns	0.0001	0.0005
Final portfolio value	$ 96,012.90	$ 120,142.95



From the above table and chart, we can see that the manual rule-based strategy has outperformed the benchmark over the in-sample time period. The Manual Strategy provides a positive cumulative return, whereas the benchmark has a negative cumulative return during the in-sample time period.

Next, the same trading rules have been applied to the same symbol during the out-of-sample period. Could the Manual Strategy achieve the similar performance over the out-of-sample period as the in-sample period? The following metrics table and chart shows the comparison between the Manual Strategy and the benchmark performance during the out-of-sample period.

Performance metrics	Benchmark	Manual Strategy (out-of-sample)
Sharpe ratio	-0.4468	-0.0257
Cumulative returns	-0.1337	-0.0234
Standard deviation of daily returns	0.0088	0.0082
Mean of daily returns	-0.0002	0.0000
Final portfolio value	$86,440.85	$97,650.30



The above table and chart show that the manual rule-based strategy lost less money than the benchmark during the out-of-sample period. The Manual Strategy still outperformed the benchmark; however, the out-of-sample performance of the strategy has failed to beat its in-sample performance. The main reason is that the Manual Strategy is designed based on the in-sample data. The parameters and thresholds for each indicator are optimized during the in-sample period, but not the out-of-sample period. This leads to somewhat overfitting. Therefore, JPM performed less well over the out-of-sample time period than the in-sample time period.

4 STRATEGY LEARNER

The Strategy Learner chooses the classification-based learner to generate trading data frames. It creates Random Forest with the RTLearner and the BagLearner developed in project three (3). The BagLearner has 10 bags with a leaf size of 5. The values of hyperparameters are determined based on experiments in order to avoid overfitting of the random forest. The learner creates different random decision trees based on various indicator values.

The same indicators discussed above are used to train the Random Forest learner. To optimize results, parameters are set differently for each technical indicator. After tests and trials, I found the following optimal values for the parameter:

- The optimal lookback window for the price/SMA ratio should be 5 days.
- The optimal lookback window for the Bollinger Bands percent is 20 days.
- The optimal lookback window for the price/EMA ratio should be 5 days.

To implement the Strategy Learner, we train the Random Forest learner by trading the stock JPM during the in-sample period from January 1, 2008 to December 31, 2009. The indicators are now features (i.e. xtrain), and the labels (i.e. ytrain) are +1, -1, or 0 representing the trading signals for buying, selling, or no trading respectively. After the learning process, the model should be tested during the out-of-sample period from January 1, 2010 to December 31, 2011. The xtest values are passed to the learner to predict y values. Then, the trades data frame could be determined using the signals of predicted y values.

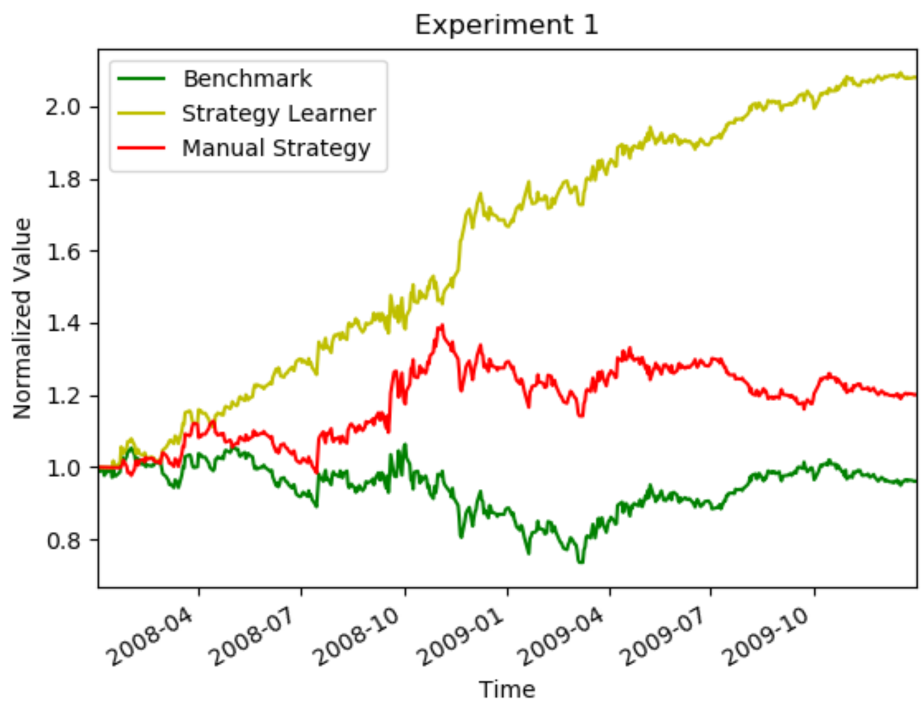
The starting cash value for the Strategy Learner is also $100,000.00, and each trade has a commission for $9.95 and an impact of 0.005. The allowable positions are long 1000 shares, short 1000 share, or do nothing (holding cash). Since the random forest learner has been used to develop the Strategy Learner, it is unnecessary to discretize the data.

5 EXPERIMENT 1

In the first experiment, we compare the Manual Strategy with the Strategy Learner during the in-sample period trading JPM. The experimental hypothesis is that the Strategy Learner outperforms the Manual Strategy during the in-sample period.

In this experiment, the starting cash value is $100,000.00, and each trade has a commission for $9.95 and an impact of 0.005 for both trading strategies. The allowable positions are long 1000 shares, short 1000 share, or do nothing (holding cash). In addition, the benchmark portfolio starts with $100,000 cash and invests in 1000 shares of JPM on the first day of trading, and then holds that position until the end of the in-sample period. When calculating the portfolio values for the Manual Strategy, Strategy Learner, and benchmark, the value should be normalized to 1.0 at the start.

The following chart and table show the outcome of the first experiment:



Performance metrics	Benchmark	Manual Strategy (in-sample)	Strategy Learner (in-sample)
Sharpe ratio	0.0683	0.5382	2.2423
Cumulative returns	-0.0379	0.2015	1.0789
Standard deviation of daily returns	0.0175	0.0134	0.0107
Mean of daily returns	0.0001	0.0005	0.0015
Final portfolio value	$ 96,012.90	$ 120,142.95	$207,864.40

The above chart shows that the Strategy Learner outperforms the Manual Strategy, and the Manual Strategy beats the performance of the benchmark during the in-sample period. The table shows the same result that the benchmark portfolio has a negative cumulative return; however, the other two strategies have the positive cumulative returns. The Strategy Learner did so well because there might be somewhat overfitting during the in-sample period.

6 EXPERIMENT 2

The second experiment is conducted with the Strategy Learner to examine how changing the value of impact should affect in-sample trading behavior using the symbol JPM. The hypothesis is that a higher impact could negatively affect the performance of the Strategy Learner during the in-sample period.

In this experiment, the starting cash value is $100,000.00 for the Strategy Learner, and each trade has a commission of $0.00. The allowable positions are long 1000 shares, short 1000 share, or do nothing (holding cash). The Strategy Learner is tested with the impact values of 0.0, 0.005, and 0.05. The performance metrics are shown in the table below.

Strategy Learner	impact = 0.0	impact = 0.005	impact = 0.05
Sharpe ratio	2.1586	2.0174	1.0254
Cumulative returns	1.0445	0.9484	0.4809
Standard deviation of daily returns	0.0109	0.0109	0.0135
Mean of daily returns	0.0015	0.0014	0.0009
Final portfolio value	$204,427.30	$194,819.50	$148,071.90

As we can see from the table, when the impact value increases, the cumulative returns and the average daily returns decrease. Also, the returns become more volatile with an increasing standard deviation as the impact value increases.

The following chart shows the outcome of the experiment 2. It corroborates the experimental hypothesis. The higher the impact value, the lower the performance.