Georgia Institute of Technology
CS7646 Machine Learning for Trading
Project 8 Strategy Evaluation

Xiang Zhang
xzhang934@gatech.edu

Abstract—In the final project of course CS7646 machine learning for trading, both of manual strategy and strategy learner methods were successfully established to forecast the trading operations. Two experiments were conducted with the help of manual strategy and strategy learner python files to study the performance of classification-based learner, manual strategy with technical indicators, and the influence of the impact factors.

1 INTRODUCTION

In this strategy evaluation project, the main focus was to study the performance of classification-based learner and manual strategy with technical indicators. In experiment 1, decision tree method was selected to build up the learner model and generate predictions on trading operations. The initial assumption of experiment 1 was that the classification-based learner could generate better performance comparing with manual strategy and benchmark.

In the experiment 2, the influence of different impact factors was studied. As impact factor was part of the transaction cost of trading operations, the assumption of experiment 2 was that the number of trades decreased with the increment of impact factor, and the cumulative return of portfolio increased with the increment of impact factor, as the increment of the impact factor increased the transaction cost of each trading operation in the stock market.

2 INDICATORS

In project 6, five indicators, namely simple moving average (SMA), exponential moving average (EMA), Bollinger band (Bollinger band percent), momentum, and volatility were selected to investigate the potential trading opportunities in the stock market. In project 8, three indicators, EMA, Bollinger band percent, and momentum were selected from these five indicators to generate the manual trading strategy.

2.1 Exponential moving average (EMA)

As discussed in project 6, the EMA indicator was estimated with higher weightage on the recent data points. Hence, EMA gave more timely information on the market or stock price changes (EMA Investopedia, n.d).

In the manual strategy in project 8, the threshold value of EMA was set to be 5%. It means that when the current price is 5% lower than the EMA, it is the buying opportunity, and the trader should buy in stocks, as the stock price may increase to reach the EMA value. On the other hand, when the current price is 5% higher than the EMA, it is the selling opportunity as the stock price may decrease to reach the EMA value.

2.2 Bollinger bands percent (BBP)

Bollinger band percentage (BBP) is an effective technical indicator to evaluate the current stock price comparing with the upper or lower Bollinger band (Barchart, n.d).

In this project, the manual strategy was set that when the BBP value is above 0.6, it is the signage of overbought, and it is the right time to sell current holding stocks. On the other hand, when the BBP value is too low, it is the signage of oversold, and the trader should buy in stocks.

2.3 Momentum

Momentum is a powerful indicator to measure the acceleration rate of the current stock price trend. For example, if the momentum is high, it is highly possible that the stock price maintains current trend.

In this project, the threshold values of momentum were set to be 0.1 and -0.1. When the value is greater than 0.1, it can be considered that the current stock price trend has relatively strong momentum to maintain its current trend. When the value is less than -0.1, it's possible that the current stock price will be reversed (Momentum Investopedia, n.d).

3 MANUAL STRATEGY

In the manual strategy, three indicators were combined together to develop a trading strategy. After large amount of trials, the EMA, BBP, and momentum were combined as below:

If (BBP > 0.6) and (momentum < -0.1): Short

If (price > 1.05 * EMA) and (momentum < -0.1): Short

Elif (price < 0.95 * EMA) and (momentum > 0.1): Long

The logic of above manual trading strategy was that when the price was relatively high comparing with its average prices, such as BBP or EMA, and the momentum was low, it was the right time to short the stock, as the price was highly possible to drop. On the other hand, the long trading operation was conducted when the stock price was relatively low, and momentum was high. The stock price was highly possible to increase.

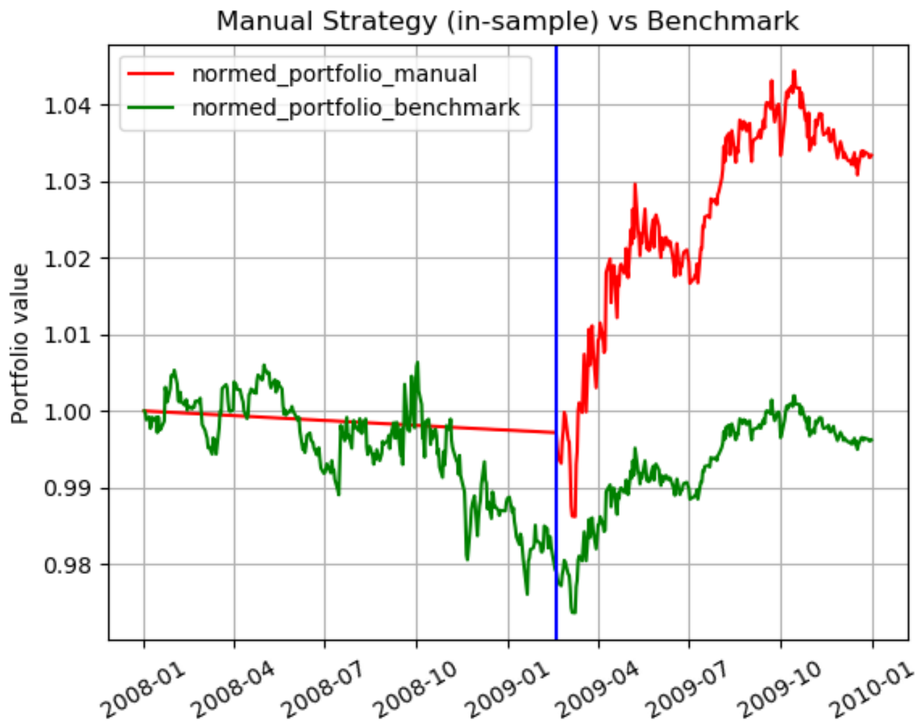

Figure 1 — Manual Strategy VS Strategy Learner (in sample).

As shown in figure 1 above, the manual strategy effectively increased the portfolio return comparing with the benchmark for the in-sample period.
In figure 1, the blue vertical line indicated the long entry point. As the threshold values of technical indicators in the manual trading strategy were set conservatively, there was only on long entry point identified for the in-sample data. Although there were not many trading entries, the portfolio return was improved by 3.3% comparing with the benchmark.

In order to further study the performance of manual strategy, both of manual strategy and benchmark were compared based on the out-of-sample data as shown in figure 2 below.

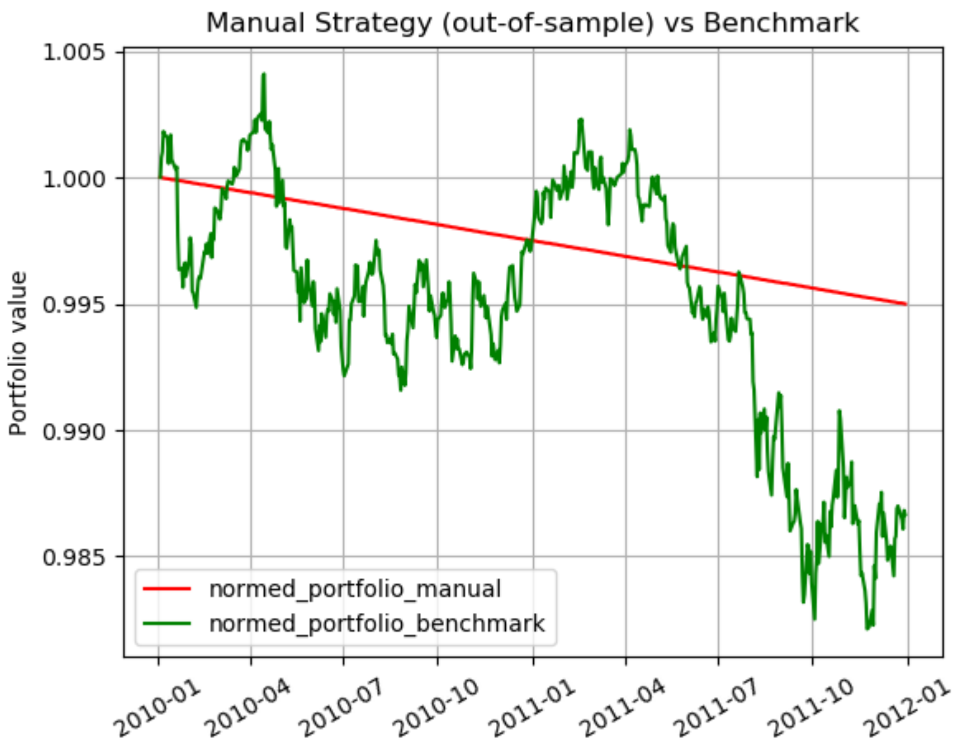

Figure 2 — Manual Strategy VS Strategy Learner (out of sample).

As shown in figure 2, the performance of manual strategy on out-of-sample was not as good as its performance on in-sample data. The manual strategy was not able to identify any entry points during the out-sample data period.

Table 1 — Manual strategy and benchmark in-sample and out-of-sample performance.

Performance	**In-sample-manual**	**In-sample-bench**	**Out-sample-manual**	**Out-sample-bench**
cumulative return	0.03338	-0.00379	-0.005	-0.01335
Std dev of daily return	0.00163	0.00162	0.000	0.000815
Mean of daily return	0.00006	-0.000006	0.000	-0.000028

Table 1 displayed the statistics of the manual strategy performance comparing with the benchmark. Obviously, the manual strategy had better performance on the in-sample period regarding to the cumulative return. The reason was that the technical indicators were set and adjusted based on the in-sample data. The manual strategy was not able to adjust itself to perform well on the unseen data, i.e. out-of-sample data. In order to generate a better performance on the out-of-sample data, the machine learning models were necessary to be developed.

4 STRATEGY LEARNER

Strategy learner was established with machine learning techniques to enhance the performance of predications on the stock trading. In the strategy learner, three indicators, EMA, BBP, momentum were calculated based on the in-sample data. These three indicators were combined together as x_data_train to train the learner. At the same time, based on the historical stock price data, the y_data_train was established as the actual training strategy.

Both of x_data_train and y_data_train were used as inputs for the classification-based learner, RTLearner in this project, to establish the machine learning model, a random tree. The second part of the strategy learner was to test the out-of-sample data from 1st of January 2010 to 31st of December 2011. The testPolicy function within the strategy learner generated the trades data frame for the out-of-sample period.

As for the hyperparameters, in the process of developing y_data_train data frame for the learner to establish the model, 5 day return was set to calculate the return ratio. The window_size was set to be 20 days, since 20 days exponential moving average(EMA) was commonly used as a technical indicator. 20 day window_size was also good for Bollinger band percent and momentum.

In this project, the classification-based learner was employed to develop the forecasting model instead of the Q-learner. In order to discretize the data to generate y_data_train to develop the learner, the threshold was set to be 0.2% of profit rate except for the impact cost. As the impact factor was part of the transaction cost of each trading operation, long operation was only implemented when the 5 day stock return is more than 0.2% of current stock price and also the impact cost of each transaction. Similarly, the short operation was only implemented when the 5 day stock return is less than 0.2% of current stock price and including the impact cost incurred. Therefore, the y_data_train only had 1, -1, 0 three integers. Number 1 stands for long operation. During long operation, the trader would buy necessary amount of stocks to make sure the current stock holding position is +1000 stocks, the maximum holding position. Number 0 stands for doing nothing. There was no trading operation when the value was 0. Similarly, number -1 stands for short operation. During short operation, the trader would buy necessary amount of stocks to make sure the current stock holding position is -1000 stocks, the minimum holding position.

5 EXPERIMENT 1

In the experiment 1, in-sample data was run on both of the manual strategy and strategy learner established with classification-based method (RTLeaner). As required by the project 8 description, the in-sample data period is from 1st of January 2010 to 31st of December 2011. These two dates were used as inputs for the experiment1 python file to conduct necessary experiment. The initial hypotheses and assumption of experiment 1 was that both of the manual strategy and strategy learner would perform better than the benchmark, since both of manual strategy and strategy learner were trained or adjusted based on the in-sample data.

In figure 3 below, the graph was successfully generated for experiment 1 to demonstrate the portfolio value of manual strategy, strategy learner, and benchmark. Same as the hypothesis mentioned above, strategy learner generated the highest portfolio return, and manual strategy also generated better portfolio return comparing with the benchmark.

As the random tree learner was used to establish the strategy learner, the variations on the portfolio return of strategy learner were expected. With the help of bagging technique, the variations could be largely reduced and the strategy learner would be able to generate relatively stable outputs. As both of the strategy learner and manual strategy were trained or adjusted on in-sample data, and manual strategy technical indicators threshold values were set conservatively, it can be expected that the strategy learner can always generate better portfolio return than manual strategy as strategy learner could have more frequent trading operations to retain the profit.

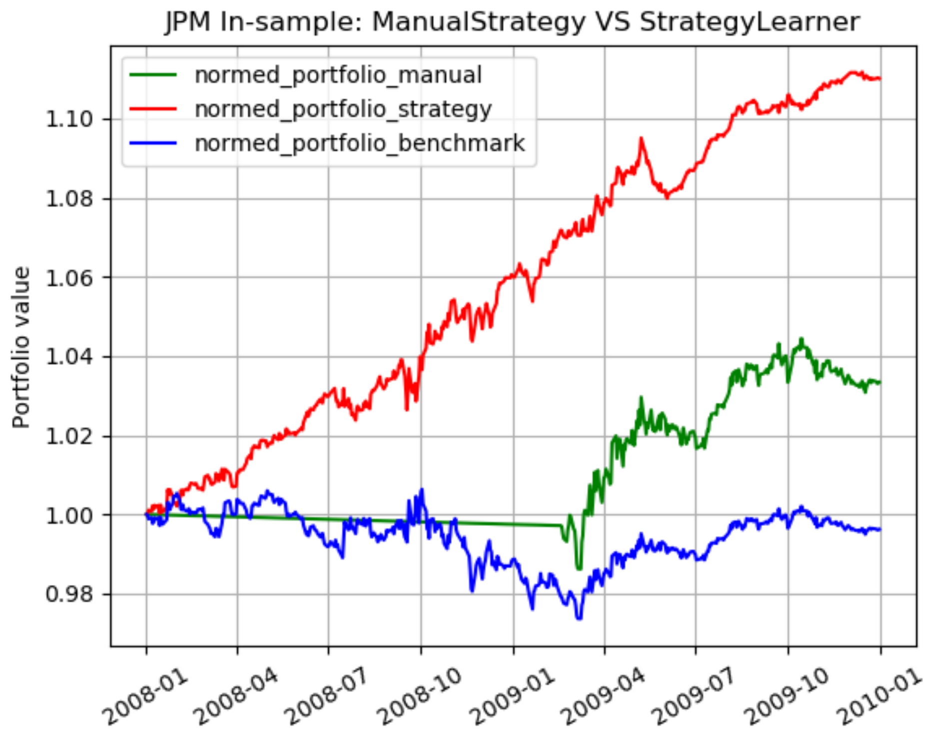

Figure 3 — JPM in-sample: Manual Strategy VS Strategy Learner .

6 EXPERIMENT 2

In the experiment 2, the influence of impact factor was studied based on the in-sample period. Three different impact factors, 0.00, 0.20, 0.30 were selected for the experiment. Two metrics, normalized portfolio value and number of trades were selected to quantitatively measure the influence of impact factor variations.

As clearly demonstrated in figure 4 below, with the increment of impact factor, it could be expected that the portfolio value was decreasing. The reason of this decrement could be interpreted that the increment of the impact factor increased the transaction cost of each trading operation. Therefore, with higher transaction cost, the portfolio value was expected to be lower.

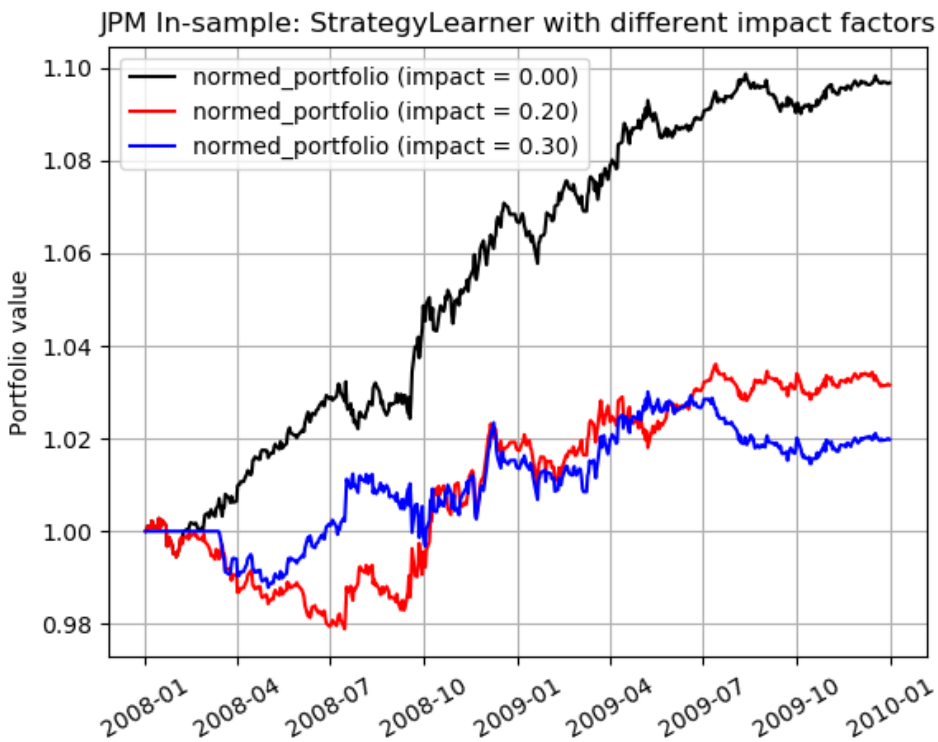

Figure 4— JPM in-sample: influence of different impact factors with Strategy Learner

The second metric was the number of trades. As the increment of impact factor can result in the increment of transaction of each trading operation. Therefore, in order to retain larger profit, the number of trades were expected to be reduced with the higher of the impact factor. As demonstrated in the figure 5 below, generated by testProject python file, it clearly described that when the impact factor increased from 0 to 0.3, the number of trades decreased from 144 to 33.

```
impact factor = 0.0, the number of trades is  144
impact factor = 0.20, the number of trades is  81
impact factor = 0.30, the number of trades is  33
```

Figure 5— Number of trades with different impact factors

7 REFERENCES

1. *Exponential Moving Average (EMA)*. (n.d.). EMA Investopedia. Retrieved July 10, 2021, from https://www.investopedia.com/terms/e/ema.asp
2. *Momentum*. (n.d.). Investopedia. Retrieved July 9, 2021, from https://www.investopedia.com/terms/m/momentum.asp
3. *Technical Indicators: Barchart.com Education*. (n.d.). Barchart.Com. Retrieved July 9, 2021, from https://www.barchart.com/education/technical-indicators/bollinger_bands_percent#:%7E:text=Bollinger%20Band%20Percent%20(BB%20%25B,is%20at%20the%20lower%20band
4. *Percent B (%B) - Fidelity*. (n.d.). Fidelity. Retrieved July 9, 2021, from https://www.fidelity.com/learning-center/trading-investing/technical-analysis/technical-indicator-guide/percent-b