Machine Learning for Trading
Report- Strategy learner

Name: Xi Han
Gt ID: xhan306

Implementation

The strategy learner is mainly a random tree learner from the second project. There is the process I implemented this learner:

1. I Utilized bagging function and choose a suitable number of bags 20 according to the previous results in second project. Comparison of performance between different bag numbers can be found in assess_learners project. In addition, according to requirements of this assignment, the leaf_size of the learner was set to 5.

2. I converted my random tree regression learner into a classification learner followed the requirements mentioned in the classification trader hints.

3. I selected three indicators SMA, Bollinger bands and Momentum to build the manual_strategy and strategy learner. By comparing the performance of three indicators in figures, I selected simply moving average (SMA) for my learner, and the threshold for SMA indicator is 0.03. The figures are shown in experiment1.

4. After finishing all the code, we use autograder to test our strategy learner. It passed all the cases.

Experiment 1

In this experiment, I used three indicators and window_size 20 for manual_strategy and trategy learner. SMA indicator has the best performance according to the figures and was selected as our indicator.

SMA indicator:

The Simple Moving Average is calculated by adding the price of a stock over a number of time periods and then dividing the sum by the number of time periods. The SMA is basically the average price of the given time period.

SMA = (sum (price, n))/ n   (n: time period)

SMA_indicator = price / SMA -1

If SMA_indicator > threshold, it's a sell signal

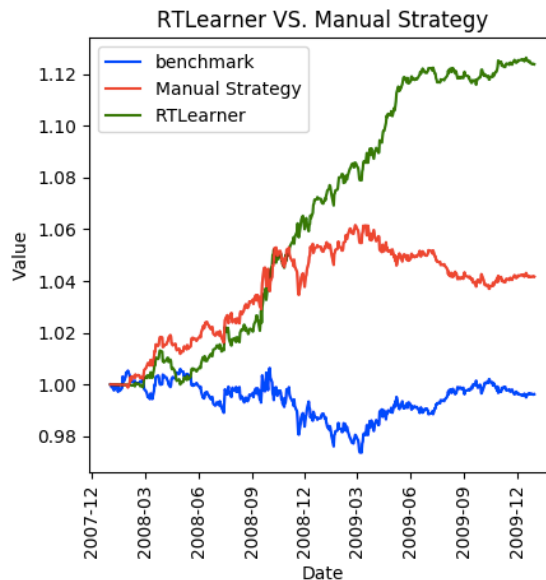If SMA_indicator < threshold, it's a buy signal

Figure 1. Performance of RTLearner with SMA

Bollinger bands:

The Bollinger Band study created by John Bollinger plots upper and lower envelope bands around the price of the stock. The width of the bands is based on the standard deviation of the closing prices from a moving average of price.

Middle Band = n-period moving average (sma is used here)

Upper Band = Middle Band + (2 * n-period standard deviation)

Lower Band = Middle Band - (2 * n-period standard deviation)

When using Bollinger_Bands_indicator(BB_indicator),

BB _indicator = (price- sma)/(2 * moving_std)

If BB_indicator > threshold, it's a sell signal
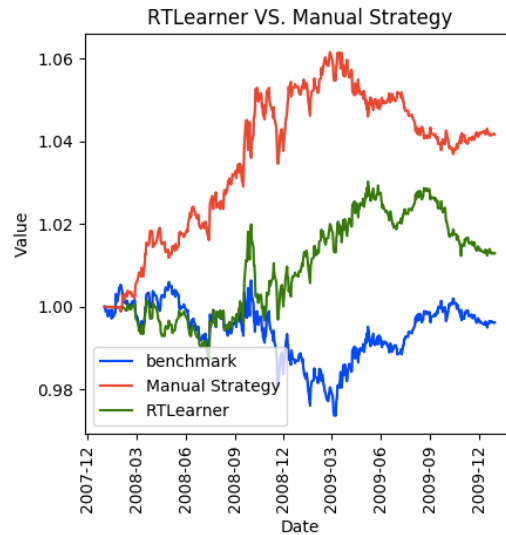
If BB_indicator < threshold, it's a buy signal

Figure 2. Performance of RTLearner with Bollinger Band

Momentum

The Momentum indicator compares the current price with the previous price from a selected number of periods ago.

Momentum = Price - Price of n periods ago

When using Momentum_indicator(MM_indicator),

MM_indicator = (price/price(-n))-1   n is the timeperiod

MM_indicator is calculated by price.diff(window_size)/price.shift(window_size)

If MM_indicator > threshold, it's a buy signal
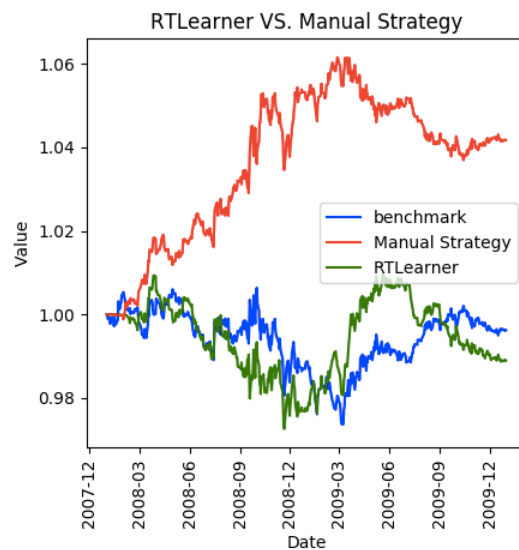
If MM_indicator < threshold, it's a sell signal

Figure 3. Performance of RTLearner with Momentum

Result:

We ingored market impact in expriment1. Different indicators were explored here. It is obvious that SMA has the best performance compared with another two indicators. SMA was selected for following exploration. In addition, it can be observed that the performance of RTLearner_Strategy is a lot better than Manual_Strategy, and the performances of both strategies are better than benchmark.

Experiment 2

In this experiment, SMA indicator was used as indicator and market impact was considered. I investigated how the value of impact affect in sample trading results. The number of trades and the performance of learners with different impacts. Seven values of impact were explored 0, 0.05, 0.10, 0.15, 0.20, 0.25, 0.30.

Result:

The figure below shows that when impact is increasing, the number of trades is decreasing. However, the number of trades increases when impact increases from 0 to 0.05, after that the trend is always downward until impact reaches 0.3.
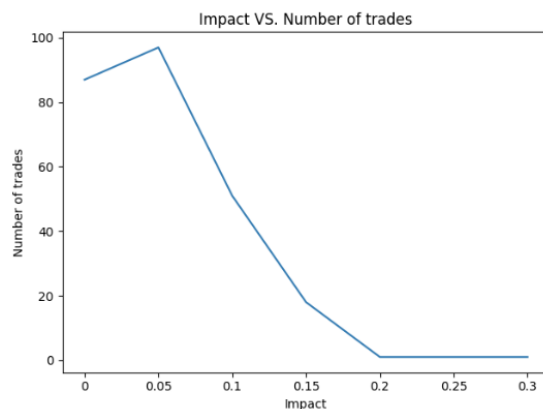

Figure 4. Change of number of trades with different impact

When market impact is small, the trading number is not influenced too much by impact. However, when market impact is bigger, implementing a trade is becoming expensive. That's the reason of decrease of number of trades. When impact increases to 0.25, the cumulative return is very close to baseline according to the figures below. Because the number of trades is close to 0 when impact is 0.25. The cumulative return is not very high when impact can be

ignored. When impact exists, but is not very high, the trade will happen but it results in negative return. In addition, it is obvious that when the impact is 0, the cumulative return is highest.
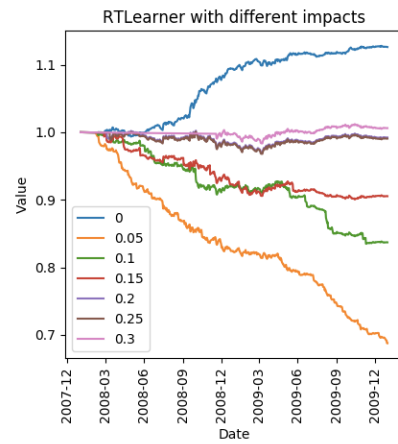

Figure 5. Performance of RTLearner with different impacts