

## **1. Q-Learners Trading Problem**

- I choose Q-Learner to implement my learning trading agent. Q-Learns needs to have actions, states, and rewards clearly defined. We decided to not use dyna in this project as we care more about world time than computational steps.
- Actions are defined as: “Long, Short, Hold”.
- The states are the combination of the different indicators that I use. To determine the exact features to use, I build suits of many features. Then I pick the best 3 based on their correlation with the daily returns.
- Nonetheless, I object similar features keep appearing in multiple examples. They are “momentum2, sma5, Bollinger band”. The features are then discretized into 10 bins to represent the different states.
- Once I have the states and actions, I need to figure out the rewards associated with each states and actions. The idea is to have positive reward if the actions complements the states (BUY action when the states are at HOLD or SHORT, etc.) and 0 or negative rewards when the actions and states oversteps the limit (BUY action when the states are at BUY, and etc.)

## 2. Experiment 1:

- For experiment 1, we trade JPM between 1/1/2008 and 12/31/2009.
  - The four indicators that we used are smap, bbp, stdev, and momentum.
  - To build Manual Strategy we do the following steps:
    - Normalize prices and generate the above technical indicators for 14-day window and trade according the rule below.

<ul style="list-style-type: none"><li>• Oversold signal constituent:<ul style="list-style-type: none"><li>○ Price/SMA (SMA) &lt; 0.95</li><li>○ Bollinger Band % (BBP) &lt; 0</li><li>○ Momentum &lt; 0</li><li>○ Volatility &lt; 0.1</li></ul></li></ul>	<ul style="list-style-type: none"><li>• Overbought signal constituent<ul style="list-style-type: none"><li>○ Price/SMA (SMA) &gt; 1.05</li><li>○ Bollinger Band % (BBP) &gt; 1</li><li>○ Momentum &gt; 0</li><li>○ Volatility &gt; 0.3</li></ul></li></ul>
---	--

3. The idea is that we buy when we see an oversold signal and sell when we see an overbought signal. The idea is we buy when the price has been going down, is traded below the average price and below the bottom band at low volatility. This suggests an imminent turning point which is an upward momentum. This means we should go long. Reversely, we want to sell out position when the price has been going up, is traded above the average and above the top band at high volatility. This signal suggests that the price will drop by reverting towards the mean, indicating we should go short. And we close positions when the symbol crosses through its SMA.

The benchmark is built by buying 1000 shares of JPM on 1/1/2008 and sell 1000 shares of JPM on 12/31/2009

- Then, to build the Strategy Learner, we plug in the 4 indicators above.
- The result is presented in the chart and table below.
- We see that up until 04-2009, Strategy Learner strongly outperform the Manual Strategy and benchmark. However, from 04-2009 to 12-2009, the strategy learner vastly underperforms the other two strategies leading it to eventually have a lower performance.
- This could be attributed to the fact that for a large period, the JPM is exhibiting downward movement behavior. Thus, the Strategy Learner have learned to make those trades. However, afterwards, the stock has an opposite movement in the direction and the learners did not adapt fast enough for the new environment.
- We expect this happen often with the in-sample data as we are practically training and testing over the same data. Thus it leads the data to have similar results often.



Statistics\Portfolio	Manual	Strategy	Benchmark
Cumulative Return	0.283083	-0.335852	0.0123

### **3. Experiment 2:**

In the real world, we would think that impacts will cause the profit of the trades to decrease, therefore, it would de-centivize traders to make many trades as they become costlier. It has been shown in previous homework that increase in impact will decrease in the return. In this experiment I am trying to determine how impacts change the number of trades.

To set up, I run multiple impacts values and observe the number of trades. The result is given below.

<b>Impact</b>	<b>No. of Trade</b>
0	295
0.005	295
0.01	295
0.02	295
0.05	295
0.1	295
0.5	269
1.0	267
1.5	257
2	255
5	251

As we can see, the number of trades decreases with the increase in the impacts. We can see that, impacts does not change the number of trades until it is become very large (0.5% and above).