STRATEGY LEARNER

1. Q-LEARNING TO SOLVE TRADING PROBLEM

I used Q-Learning to solve the trading problem. As its state space, I have used the indicators that are used in the Manual Strategy project. I will explain them shortly since they are stated in previous project.

a. Bollinger Bands : Bollinger Bands are two standard deviations (upper band and lower band) away from a simple moving average.

b. Slow Stochastic: The slow stochastic indicator compares a stock's closing price over its closing prices on a period. In my implementation, I have used 14 days as a period.

c. Moving Average Convergence Divergence (MACD) Histogram: The MACD indicator is a refinement of two Exponential Moving Averages (EMA) systems. I have used 12 and 26 days EMA in this implementation. So MACD is simply 12-day EMA – 26-day EMA And, the signal line of the MACD indicator is the 9-day EMA of the MACD value.

For Q-Learning, the state space needs to be discretized. So, I have discretize each indicator into 10 bins and this resulted in a 1000-state space for the problem. I had already normalized them for Manual Strategy Project and they are standing between -1 and 1.

As action space, I have used 'get long (or stay long)', 'get cash', and 'get short' actions. This made many things easier while coding except when trading impact is involved. I will explain this part at the end.

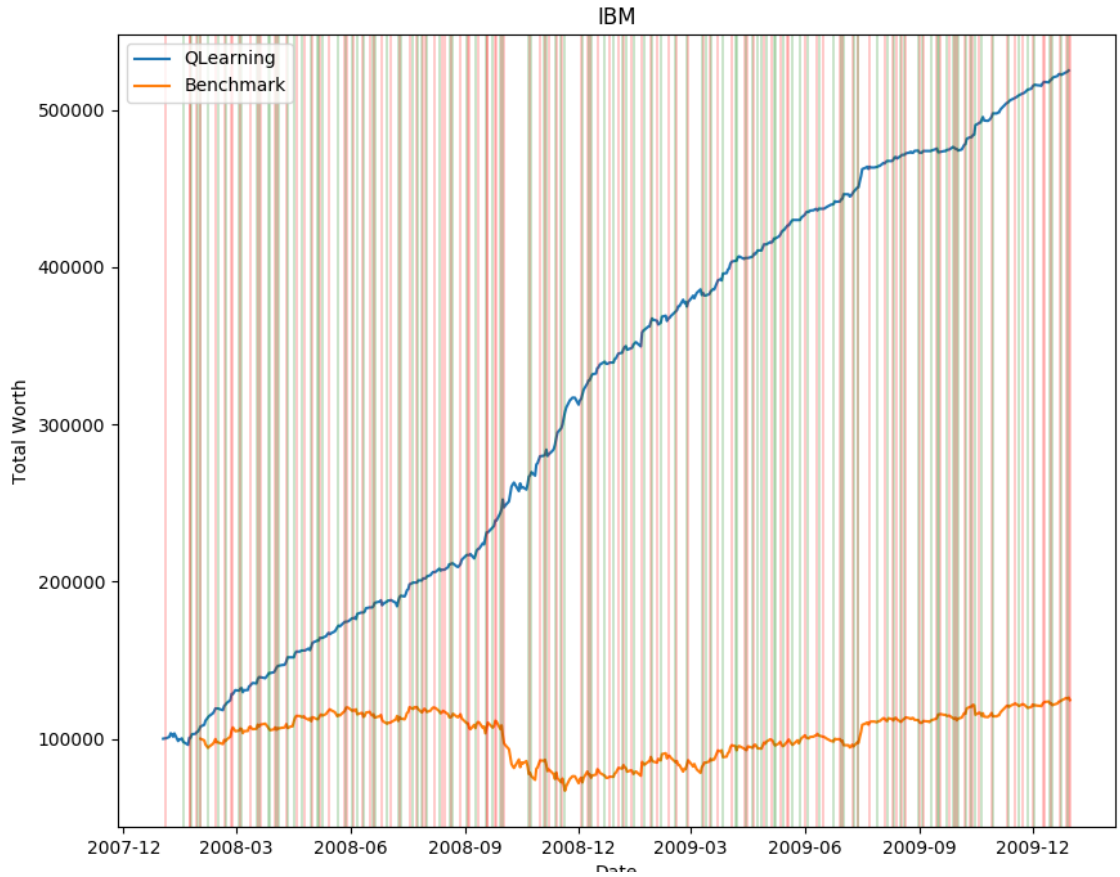
For rewards, I have used classification method. If the stock goes up, it is +1 for longing and if it loses value, it is -1. The opposite values are for shorting for both. The idea behind this implementation is that, if I believe the stock is going up, then I will make money considering I made the correct estimation and it doesn't matter how much I will be making since it is better than not making any money or losing.

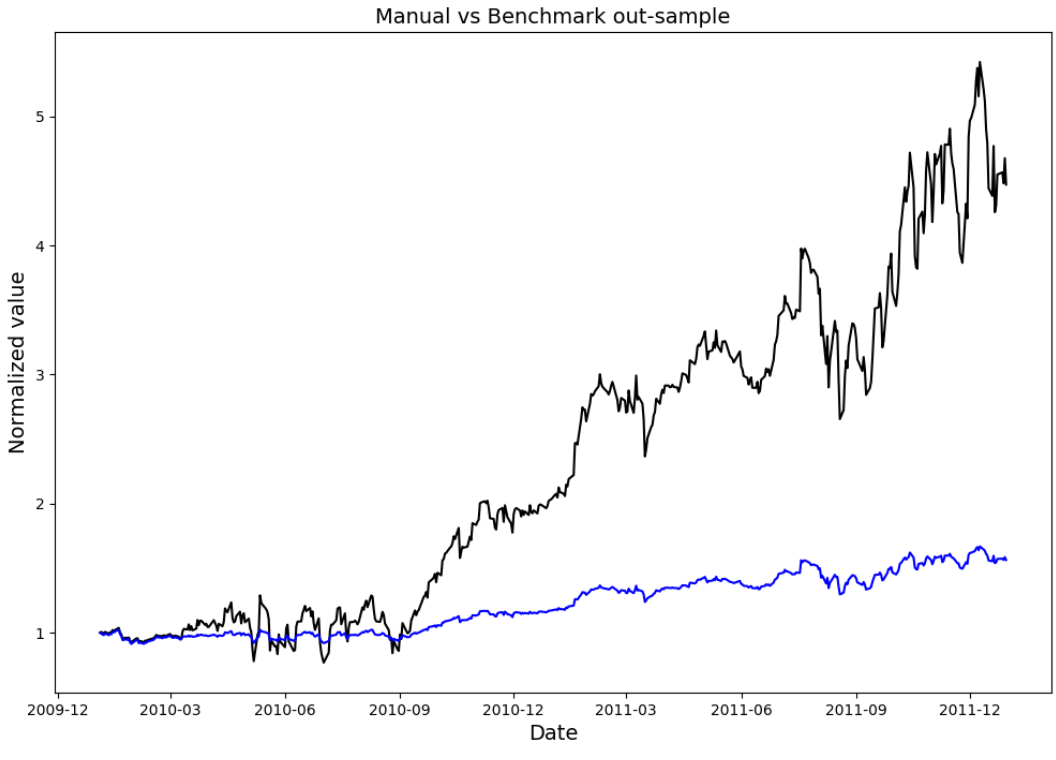
2. EXPERIMENT – 1

I have used the Manual Strategy and Q-Learning Strategy on different stocks and I have observed that at in-sample results, Q-Learning Strategy has a huge advantage over Manual Strategy. This might be caused by many reasons and I think the most important one is overfitting. Q- Learning goes over and over the same dataset and thus it is able to predict the transitions between states and how it will reward it.

In the beginning, I thought this will result in not any better results for out-sample datasets but, after some trials I have seen that Q-Learning is doing better than Manual Strategy at out-sample datasets, too. As well as Q-Learning overfits and can't give similar results at out-sample datasets, Manual Strategy also overfits to what dataset I have constructed it at the previous project.

I have compared each at IBM dataset and Q-Learning gave a return over 5 times as it can be seen below, whereas Manual Strategy can achieve around 4 times. Moreover, it can be understood visually that Q-Learner has a much higher Sharp ratio since it has a steady increase in value and this shows it has much less risk.



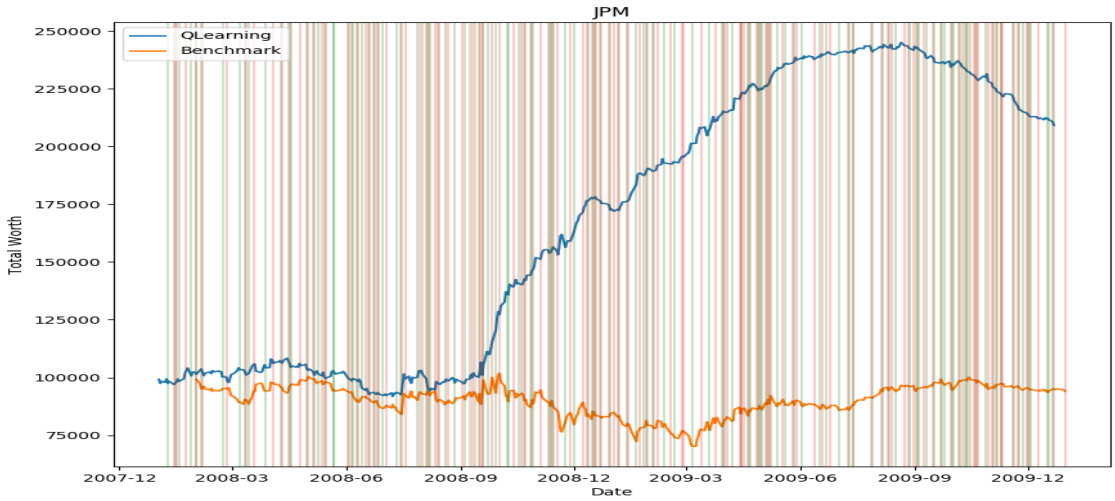


3. EXPERIMENT-2

Market impact creates an inevitable disadvantage for traders and in this project, I have tried to see its effects on my Q-Learner. First of all, the bigger market impact, the less transactions should be done. However, the way I have implemented the project inhibited that. Under high market impact, my Q-Learner tended to stay in cash instead of not moving.

Even though, I have expected market impact to reduce the amount of transaction Q-Learner is making, it increased because my implementation required the Q-Learner go to cash if it thinks it won't make enough profit to cover the market impact. The optimal scenario is not selling what I have in this case. I could change the way I have implemented to solve this problem given I have more time. My Q-Learner did more frequent transactions under market impact.

Second important metric is the cumulative return. Market impact has a decreasing impact on cumulative return. As I have increased the market impact, the cumulative return decreased. Below, you can see a market impact of 0.02. And below that, you can see the same stock and same Q-Learner under 0.05 market impact. Actually, if my Q-Learner behaved appropriately and didn't make more frequent transactions, this effect would be diminished.



4. AAPL and UHN Stock Figures

In this section, you can see the comparative results of Q-Learning vs given benchmark for this two stocks. Green lines show BUY and red ones show SELL commands.