Name: Kok Jian Yu

GTID: 903550380

- Describe the steps you took to frame the trading problem as a learning problem for your learner. What are your indicators? (They should be the same ones used for Manual Strategy assignment) Describe how you discretized (standardized) or otherwise adjusted your data. If not, tell us why not.

The learner that I have used is the Q-Learner. Q-Learn requires that the state and the action be an integer. Therefore, the first thing I have to do to frame the trading problem to a learning problem is to map every state and action to an integer. After doing that, all I have to do to trade my learner will be to pass it actual data (State and Reward) to train with.

The indicators I used are Bollinger Band, Momentum and Stochastic.

Using the values from the training data,
I discretized each of the 3 indicators into 4 equal sized buckets. For Momentum and Stochastic, I simply discretized the raw value returned by the indicators. For Bollinger Band, I discretized subtraction of the middle Bollinger line with the bottom Bollinger line, which essentially returns half of the space of the Bollinger band.

Some additional values I have used are stock_price[i] – stock_price[i-1](Stock Price Diff), and current holdings. (Both of which are data that are accessible in Manual Strategy Assignment)

Holdings is an integer value between -1 to 1, with -1 meaning that it is holding a short position, 0 holding no position and 1 holding a long position.

Stock Price Diff is discretized into 3 equal sized buckets.

Using the buckets obtained from discretizing the training data, I can now discretize any real numbers I get into integers, including values from the test dataset.

One thing to note is that values that are outside the range of the buckets will have the following index:

$$index = numBuckets + 1$$

Therefore, when we split the value in 4 buckets, we can actually get back 5 different values when trying to get which bucket the value lies in.

Meaning that now I have the following values

- Momentum (1 to 5)
- Bollinger Band (1 to 5)
- Stochastic (1 to 5)
- Stock Price Diff (1 to 4)
- Holdings (-1 to 1)

Using the 5 values above, I can then convert a specific state to an integer. I did this by converting all of the 5 values above to start from index 0 and using the following equation.

$$state = momentum + stochastic * 5 + bollinger * 25 + holdings * 125 + price * 375$$

The above equation makes it such that each values have the following specific values

- Momentum (0 to 4)
- Stochastic (0,5,10,15,20)
- Bollinger Band (0,25,50,75,100)
- Holdings (0, 125, 250)
- Stock Price Diff (0, 375, 750, 1125)

This makes it such that every value of the state will only correspond to a single unique combination of features.

Reward is the change in portfolio value for the day.

I then passed this state number and the reward to the q learner to train it.

- Experiment 1: Using exactly the same indicators that you used in manual_strategy (trade JPM), compare your manual strategy with your learning strategy in sample. You can use the same impact (.005) as was used for Project 6 or use 0 for both. Be sure to add in an author method.
  - Describe your experiment in detail: Assumptions, parameter values and so on.
  - Describe the outcome of your experiment.
  - Would you expect this relative result every time with in-sample data? Explain why or why not.

Parameter Value:

- Alpha/Learning Rate = 0.05
- Gamma/Discount Rate = 0.9
- rar/Epsilon = 0.99
- radr/Epsilon Decay = 0.999
- Dyna = 0
- Impact = 0
- Transaction cost = 0

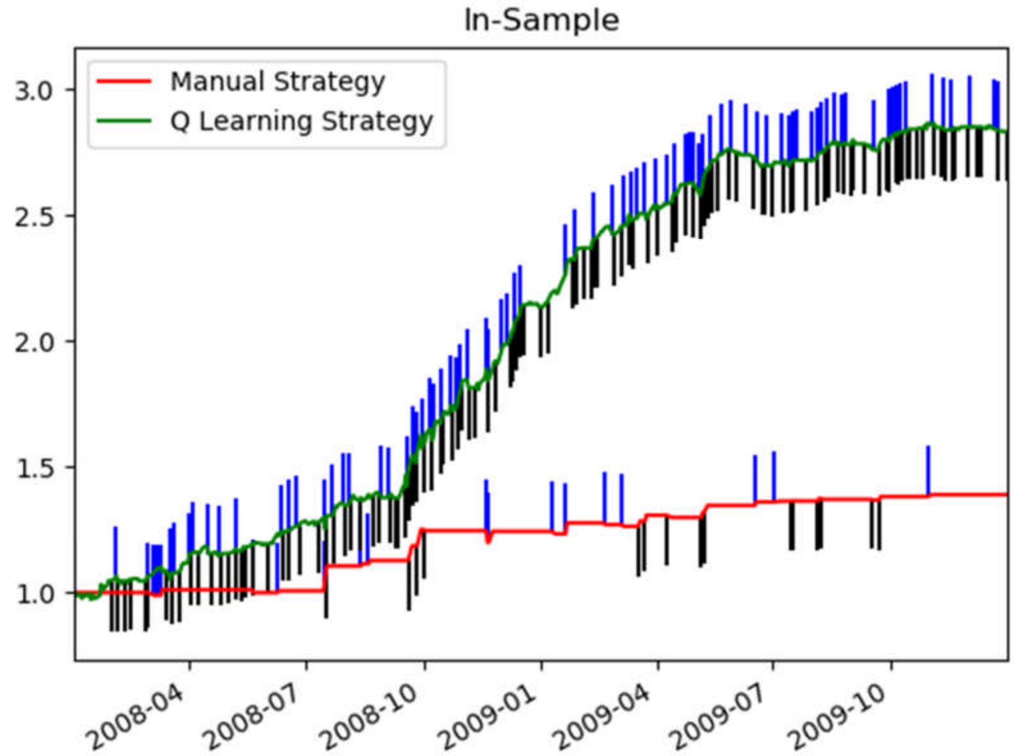
To train the q learner, I ran it through the training data for 30 iterations.

To test the performance of the Q learning strategy, I run the manual strategy and Q learning strategy using their test_policy function, using the in_sample data with an impact of 0.

The outcome of the experiment is that the Q learning algorithm greatly outperforms the Manual Strategy.

- Manual Strategy Final Portfolio = 138940.0,
- Q Learning Strategy Final Portfolio = 283250.0

I will expect similar result every time with in-sample data as the Q learning strategy trained its strategy to maximize the amount earned for the in-sample data. As such, it will perform well for the in-sample data. Since the learner can optimize the values better than I can when I did the manual strategy, I will expect It to perform better than my manual strategy.


In-Sample
Manual Strategy
Q Learning Strategy
3.0
2.5
2.0
1.5
1.0
2008-04
2008-07
2008-10
2009-01
2009-04
2009-07
2009-10

- Experiment 2: Provide a hypothesis regarding how changing the value of `impact` should affect in sample trading behavior and results (provide at least two metrics). Conduct an experiment with JPM on the in sample period to test that hypothesis. Provide charts, graphs or tables that illustrate the results of your experiment. The code that implements this experiment and generates the relevant charts and data should be submitted as `experiment2.py`. Be sure to add in an author method.

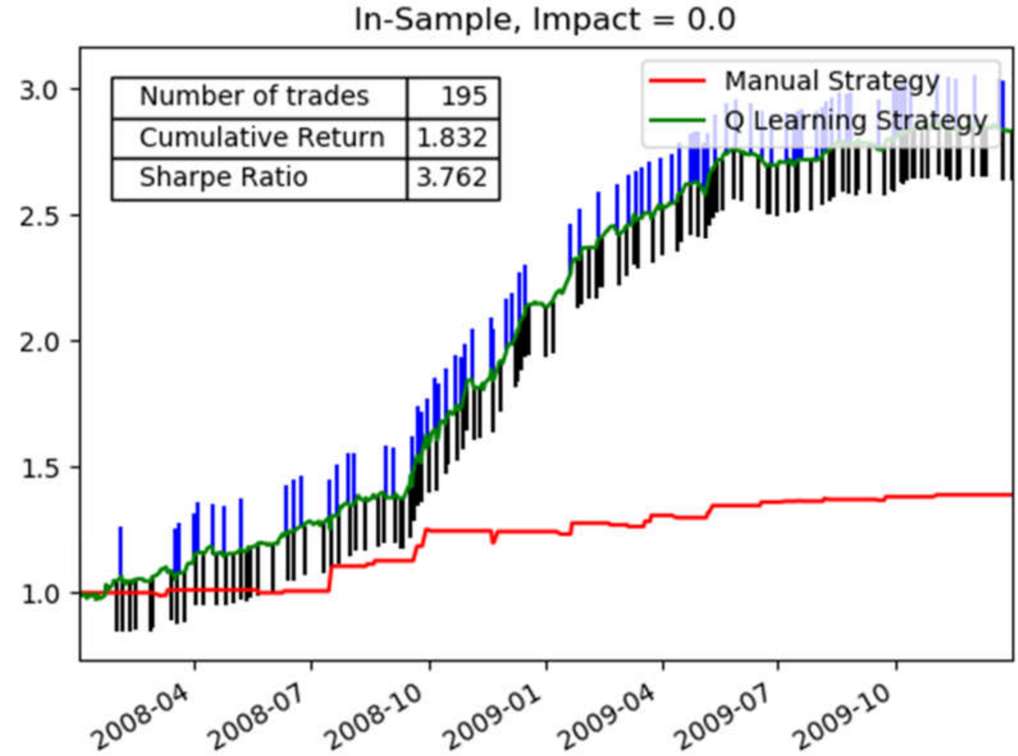
Increasing the impact should change the trading behaviour by making the learner making lesser trade compared to when impact is lower. This should result in a lower cumulative return. In addition to this, since the learner is now making lesser trades to increase the portfolio, this will probably also mean that the portfolio will no longer be constantly increasing, ass there will be several areas where trades don't occur. This means that the slope of the portfolio will be more inconsistent, and this should result in a lower sharpe ratio.

I conducted an experiment to observe this by training the learner on different impact value and seeing how it performs compared to my manual trader given the same impact.

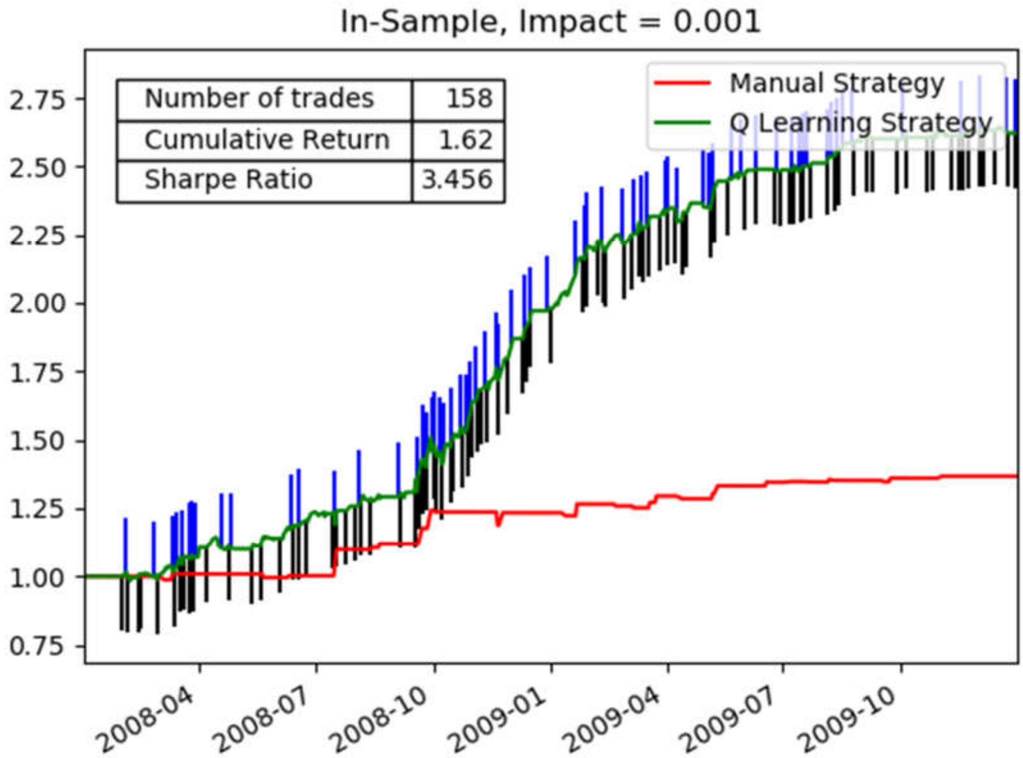
From the experiment, we can indeed see that this is the case. As impact increases, both cumulative return and sharpe ratio decreases. It can also be observed that the number of trades the learner makes also decreases.

Graphs generated for this experiment is on the next page.

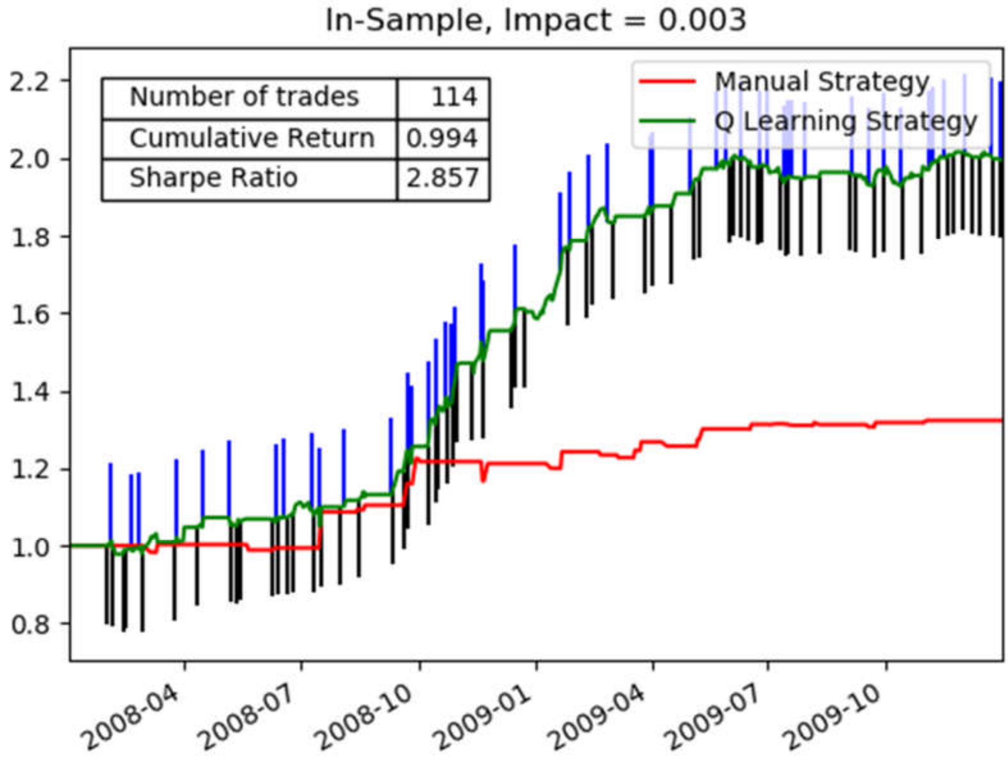
Table in the graphs is showing the statistics for the Q Learning Strategy

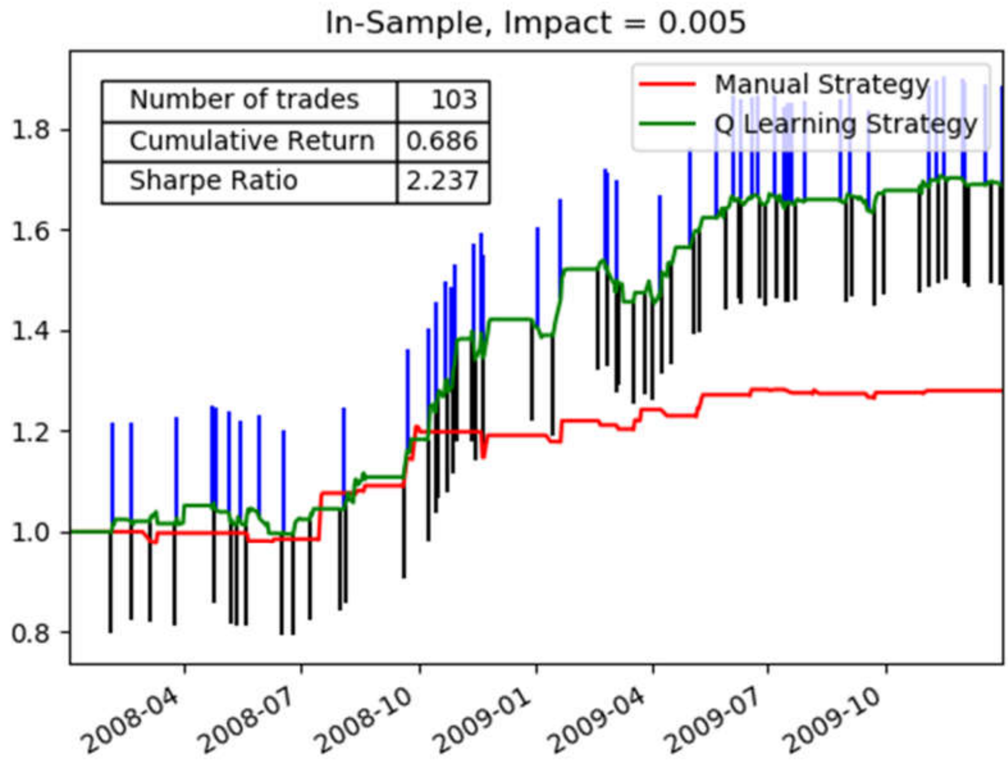


In-Sample, Impact = 0.0
Number of trades | 195
Cumulative Return | 1.832
Sharpe Ratio | 3.762
Manual Strategy
Q Learning Strategy
3.0
2.5
2.0
1.5
1.0
2008-04
2008-07
2008-10
2009-01
2009-04
2009-07
2009-10



In-Sample, Impact = 0.001
Number of trades | 158
Cumulative Return | 1.62
Sharpe Ratio | 3.456
Manual Strategy
Q Learning Strategy
2.75
2.50
2.25
2.00
1.75
1.50
1.25
1.00
0.75
2008-04
2008-07
2008-10
2009-01
2009-04
2009-07
2009-10


In-Sample, Impact = 0.003
Number of trades 114
Cumulative Return 0.994
Sharpe Ratio 2.857
Manual Strategy
Q Learning Strategy
2.2
2.0
1.8
1.6
1.4
1.2
1.0
0.8
2008-04
2008-07
2008-10
2009-01
2009-04
2009-07
2009-10


In-Sample, Impact = 0.005
Number of trades 103
Cumulative Return 0.686
Sharpe Ratio 2.237
Manual Strategy
Q Learning Strategy
1.8
1.6
1.4
1.2
1.0
0.8
2008-04
2008-07
2008-10
2009-01
2009-04
2009-07
2009-10