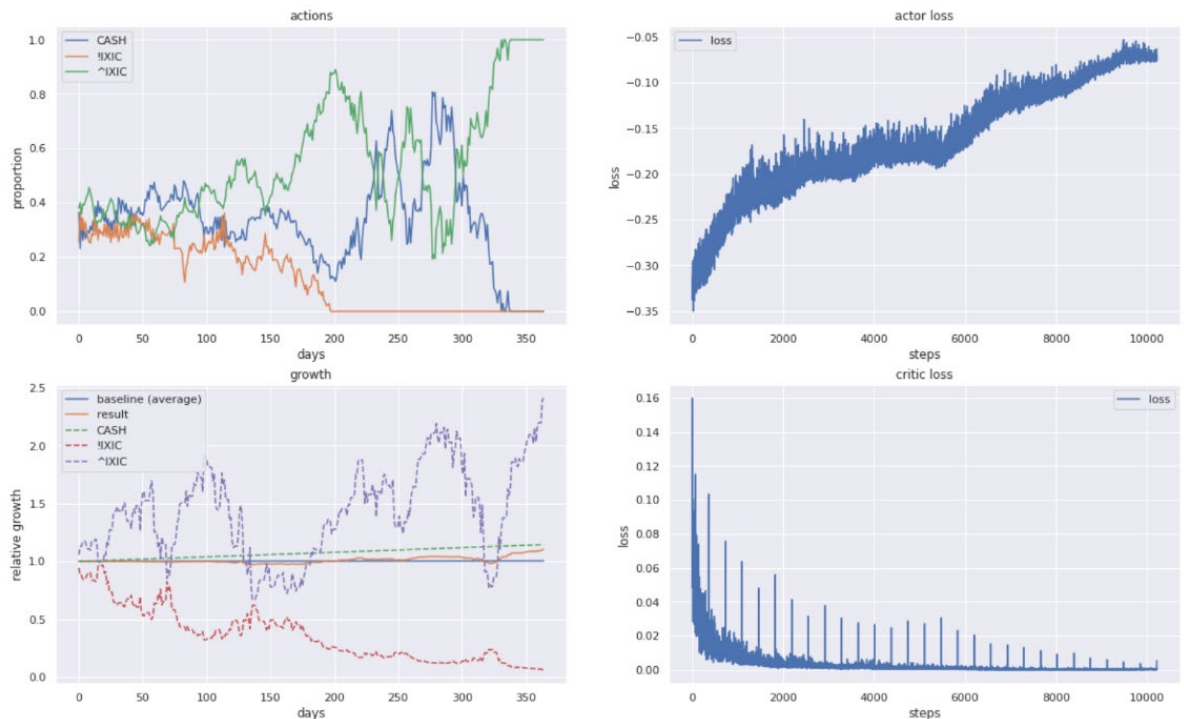


Final project report group3 B08901185 曾啟睿

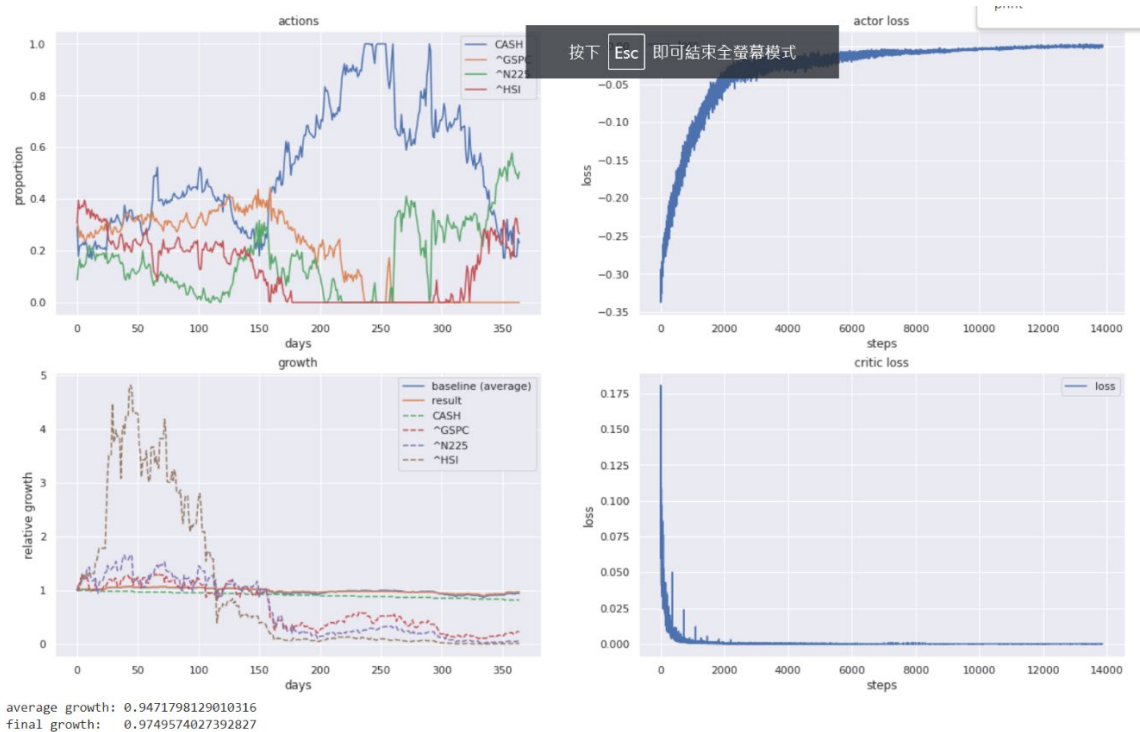
1. Additional simulation result by myself:

The graph below is we try to test our model with fake data at first(so only have two different stock which can be easily found the best strategy by human). For the right part is the loss graph we set the actor loss to backward by the minus of the critic loss. As for the left part, the upper window is the portion of money we allocate(the sum is equal to 1),the blue line represent cash. And the lower window is the stock growth(purple line and red line). We can see from the graph that when the red one started to get down, our model quickly sold it out and buy the purple one as its value is getting larger, this is the graph after two hundred epoch.



average growth: 1.0045306006585295
final growth: 1.102778322225176

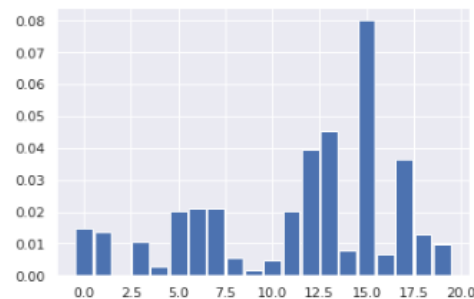
However, when we try to implement for more than four different asset, the result is tend to be hard to train, it seems not taking progress toward our analytic goal, it only learn that while other asset keep losing their value, the model would put much more weight on cash. As you can see for the graph below.



2. Result:

Training

After training for 320 episodes, the performance of last 20 episodes compared with baseline:



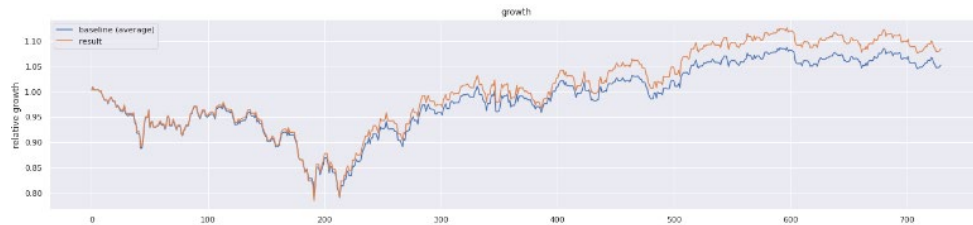
Our model outperforms the baseline by 1.9% per year

Testing on test set

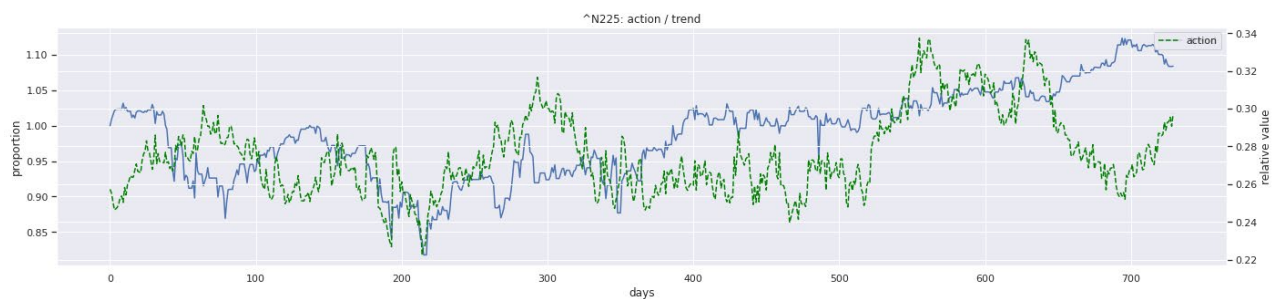
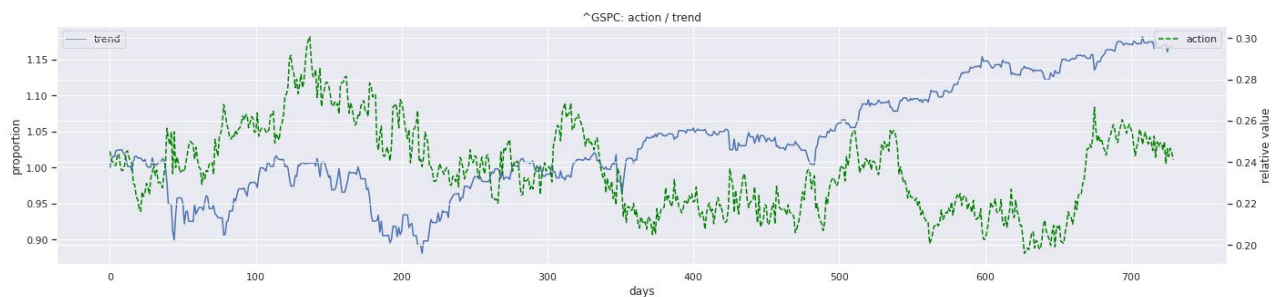
Baseline: 5.2%

Our model: 8.4%

Outperform 3.2% (in 2 years)



The green line is the portion we possess for each stock, the blue line is each stocks' trend. It seem to have an approach that when the price go low -> buy it; price go high -> sold it, (very conservative investor but a little bit short-sighted)



3. Work distribution:

We have five group members: 陳韋旭、莊加旭、顏柏聖、李明翰、曾啟睿

陳韋旭	莊家旭	顏柏聖	李明翰	曾啟睿
30%	30%	20%	10%	10%

Since the code is mainly implement by 陳韋旭 and 莊加旭. Whenever encountered a bug all of us try to fixed it so we try to search for different paper but after our discussion, it will be tide up by 陳韋旭 and 莊加旭 would add into the code, 顏柏聖 and 陳韋旭 fixed a huge portion of bugs. As for me and 李明翰, we only fixed a little, so we work mainly on parameters adjusting. And every presentations are done by all of us in equal portion.(20%,20%,20%,20%,20%)

4. Reflection:

Reinforcement learning is really hard, completely opposite of my thought at first. We face lots of problems, such as exploration and why the model like to keep staying at the same portion of some stocks, and why we implement trading cost but the model still oscillate dramatically... and so on.