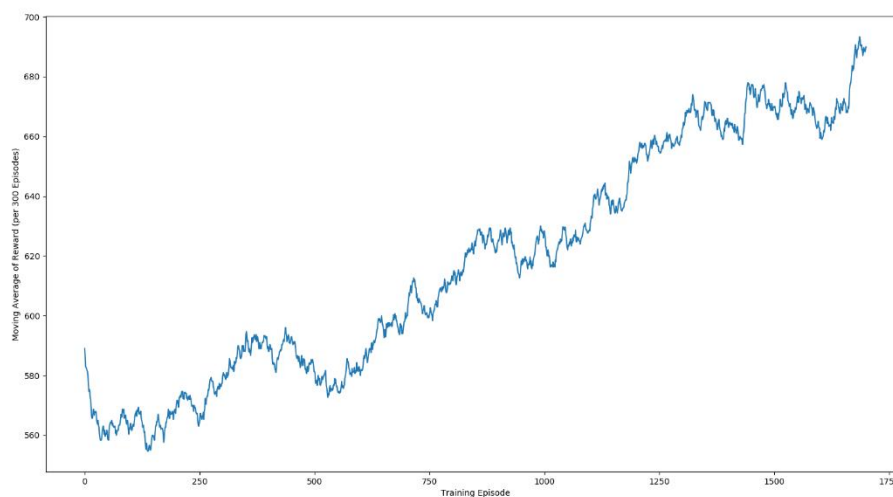


The graph above is my result about StarGunner. The x-axis is training episode and the y-axis is episodic reward. I ran a total of 2000 episodes. As can be seen from the graph, rewards generally show a slight upward trend. If we look at the initial 500 episodes, we will find that there are many rewards below 100 points, and few rewards above 1200 points. In contrast, among the last 500 episodes of the 2000 episodes, few rewards are below 100 points, and many rewards are above 1200 points. Therefore, I think the agent is learning, although the learning speed may be a bit slow.



The graph above is more intuitive than the first graph. The x-axis is training episode and the y-axis is the moving average of rewards. From this graph, we can see that the reward has risen from about 560 points at the beginning to about 690 points at the 2000th episode. Although reward did not exceed 700 points, according to the trend of this line, I suppose that it is only a matter of time to exceed 700 points.