Homework 5

1. Consider the problem of Homework 4 where a company is introducing a new product into the market. If the sales are high, there is a 0.5 probability that they will remain so next month. If they are not, the probability that they will become high next month is only 0.2. The company has the option of launching an advertisement campaign. If it does and the sales are high, the probability that they will remain high next month will increase to 0.8. On the other hand, an advertising campaign while the sales are low will raise the probability to only 0.4.

   If no advertisement is used and the sales are high, the returns are expected to be 10 if the sales remain high next month and 4 if they do not. The corresponding returns if the product starts with low sales are 7 and -2. Using advertisement will result in returns of 7 if the product starts with high sales and continues to be so and 6 if it does not. If the sales start low, the returns are 3 and -5, depending on whether or not they become high. Determine the company's optimal policy over an infinite horizon with discount factor $\alpha = 0.6$ first using policy iteration and then using value iteration. You can write a code for both algorithms. In that case please provide your codes, the outputs of your codes and the number of iterations required for each algorithm.