ENPM808F
HOMEWORK 4

Nithish Kumar S (116316958)

Project Overview:

This report contains details of the Q-learning implementation for the game of dots and boxes.The goal is train the game using self-play and testing its performance with a random agent for a 2x2 grid and 3x3 grid.

Inferences:

The 2x2 grid took lesser number of iterations in the training to give a good winning ratio than the 3x3 grid.

2x2 grid - 2^12 states
3x3 grid - 2^24 states

This is because the 2x2 grid has lesser number of states, so smaller the Q-table size. So it was able to fill the table faster than the other.

Some results:

Greedy policy
Epsilon - 1.0
Decay_rate - 0.01
Discount factor - 0.9
Learning rate - .5

2x2 grid:

Training 100 iterations:
Wins : 52 / 100
Training 1000 iterations:
Wins : 68 / 100
Training 10000 iterations:
Wins : 85 / 100

3x3 grid:

Training 100 iterations:
Wins : 36 / 100
Training 1000 iterations:
Wins : 52 / 100
Training 10000 iterations:
Wins : 79