

Question 4 (1 point)

Policy Iteration Agent against each of the provided agents 50 times

Against Defensive Agent:

Wins: 46 Losses: 0 Draws: 4

Against Aggressive Agent:

Wins: 50 Losses: 0 Draws: 0

Against Random Agent:

Wins: 50 Losses: 0 Draws: 0

This serves as the starting point for policy iteration. Subsequently, the `evaluatePolicy` method calculates the values of each state under the current policy using iterative updates until the values converge within a specified threshold (δ). Finally, the `improvePolicy` method adjusts the policy by performing a single-step expectimax to ensure actions lead to higher state values. The `train` method combines these steps in a loop until the policy stabilizes, yielding an optimal policy for the agent.

Screenshots and live testing you can see in video.