

CS 747 Assignment 2

Manoj Bhadu(170010036)

T1:

1. Value Iteration:

- a. Used numpy allclose to compare value function at t and $t-1$ and set $rtol=1e-13$, $atol=1e-15$ or number of iteration 20,000. Which of the condition reaches first I terminate the iteration.

2. Howard Policy iteration:

- a. For given policy calculated the value function by using iteration and used numpy linlag norm for calculating the difference the previous and current iteration setting the epsilon to $1e-10$.
- b. Initialized the policy with all zero actions.
- c. Calculate Q value for each state and action and then which action has the most difference with the value function. take that action as optimal policy and reiterate loop till action with most difference is same as optimal policy(previous iteration).

3. Linear Programming Formulation:

- a. Formulate the problem as discussed in class as maximize the negative of value function given linear constraints.
- b. Calculated the optimal value function using pulp solver and then calculated optimal policy using the relation between them as mentioned in slides.