Assignment 0

1. You should implement a 4 x 4 (or N x N) Gridworld (as given in the Sutton and Barto [book]) environment. You can also use [Gym] environments, but it is suggested to do so only after you have implemented an environment yourselves.

2. Implement **Policy evaluation, Policy iteration, Value iteration** algorithms. You can use the same environment for all three. For Policy Evaluation, you can use a random policy (which takes each action with the equal probability) to evaluate and compare your output with the one in the book. For the other two algorithms, you can start with a random policy and continue updates until your algorithm converges to optimal behavior.

3. Submit the implementations in .ipynb file format with the output for each cell and also the final output (value function for your policy in Policy Evaluation and Optimal value function and Optimal policy for the other two). Document your code with comments wherever required.

4. Make a folder of your name with all your .ipynb notebooks in it and place it inside the Assignment0 folder. Then create a Pull request in this repository under the [Reinforcement_Learning_SMP branch].

Deadline - 17/05/2020