**HW 11: ECE 601 Machine Learning for Engineers**

**Important Notes:**

(a) When a HW question asks for writing a code, you would need to include the entire code as well as the output of the program as well as any other analysis requested in the question.

(b) Don't panic about the length of the HW assignment. HW assignments are treated as opportunities for improving learning and understanding, so I might include some extra text to help you better understand the concepts or learn about a point that was not covered during the class. The actual work needed from you is indeed manageable.

(c) Note that there are no programming assignments this week, so you can submit your work as a single PDF.

# Question 1: Expected Rewards in an MDP

**Problem:** Let $(\mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \gamma)$ be a finite Markov Decision Process. Suppose from state $s$, the agent can take action $a_1$ or $a_2$, with transition probabilities and rewards:

- $\mathcal{P}(s_1 \mid s, a_1) = 0.6$, $\mathcal{P}(s_2 \mid s, a_1) = 0.4$, $R(s, a_1, s_1) = 1$, $R(s, a_1, s_2) = 0$
- $\mathcal{P}(s_1 \mid s, a_2) = 0.5$, $\mathcal{P}(s_2 \mid s, a_2) = 0.5$, $R(s, a_2, s_1) = 2$, $R(s, a_2, s_2) = 2$

1. Compute the expected immediate reward for each action:

$$\mathbb{E}[R_{t+1} \mid S_t = s, A_t = a_1], \quad \mathbb{E}[R_{t+1} \mid S_t = s, A_t = a_2]$$

2. Assume a discount factor $\gamma = 0.9$ and that the process continues indefinitely. Suppose the agent always chooses the same action from state $s$ and is returned to $s$ after each transition (i.e., $s$ is recurrent with probability 1).

$$R(s, a_1, s) = 0.6, R(s, a_2, s) = 2.0$$

Let $v(s; a_i)$ denote the expected return starting from $s$ and always taking action $a_i$. Compute:

$$v(s; a_1), \quad v(s; a_2)$$

# Question 2: Solving Bellman Equations for State Values

**Problem:** Consider an MDP with $\mathcal{S} = \{s_1, s_2\}$, $\mathcal{A} = \{a\}$, and deterministic transitions:

- $\mathcal{P}(s_2 \mid s_1, a) = 1, R(s_1, a, s_2) = 3$
- $\mathcal{P}(s_1 \mid s_2, a) = 1, R(s_2, a, s_1) = 2$
- $\pi(a \mid s_1) = \pi(a \mid s_2) = 1$
- $\gamma = 0.5$

1. Write the Bellman equations for $v_\pi(s_1)$ and $v_\pi(s_2)$.
2. Solve them to compute the exact values of $v_\pi(s_1)$ and $v_\pi(s_2)$.