AI4ICPS IIT Kharagpur	**IIT KHARAGPUR AI4ICPS I HUB FOUNDATION** **H**ands-on **A**pproach to **AI, Cohort-2, July – October 2024** **Assignment 7: Reinforcement Learning**

Due date: Friday 30th August 2024, EOD – IST.

Important Instructions for submitting solutions

1. Submit the solution to all questions in the assignment should be submitted in a **single PDF file with not more than 500 words**.
2. Any plagiarism if detected will automatically attract **zero marks** for that assignment.
3. It is preferable if the **text of PDF file can be extracted** through a PDF extractor e.g. PyPDF. For example, pictures of handwritten text are not extractable, whereas PDF generated by MS Word, Latex, etc., are.
4. Exceptionally good solutions with extractable text may receive **special appreciation** from the teachers.

1. You are developing a reinforcement learning-based system for an autonomous drone tasked with delivering packages in an urban environment. The drone must navigate through a city, avoiding obstacles such as buildings, trees, and other flying objects. It must also optimize its flight path to minimize delivery time while ensuring safety and battery efficiency. The drone is equipped with sensors for detecting obstacles and GPS for navigation.

   Formulate this problem as an RL task. Specify the following components of the associated MDP:
   - State Space (S): Describe how you would define the state space for this problem.
   - Action Space (*A):* Define the action space available for the drone
   - Reward Function (*R*): Propose a reward function that encourages efficient and safe deliveries

2. Discuss the trade-off between exploration and exploitation in Reinforcement Learning. Explain the ϵ-greedy strategy and how it accomplishes balancing this trade-off.

3. In a simple MDP, an agent is in a state s, and the actions it can take can lead to the following outcomes:
   - With probability 0.4, the agent transitions to state s', with reward $R = 10$, and $v(s') = 5$

- With probability 0.6, the agent transitions to state s'', with reward $R = 2$, and $v(s') = 3$.

The discount factor γ is 0.5. Using Bellman equation, find the expected value of state s.