Centre for infrastructure, Sustainable Transportation and Urban Planning

Indian Institute of Science (IISc), Bengaluru
Summer Internship: Round 1

The aim of this exam is to test your knowledge in the area of reinforcement learning for the purpose of shortlisting for the project:Inverse Reinforcement learning for tolling and parking (supervised by Dr. Tarun Rambha (link)): Follow the instructions below.

- Plagiarism will result in instant disqualification. You must write your own code.
- Create copy of ridehailing package. The problem statement is given in next page. You need to add codes for solving the problem statement without making any changes to the given environment.
- All python codes has to be in '.py' format only. Additionally, create a pdf file to document the type of policy, type of RL algorithm (along with hyperparameters used), python libraries used and any assumptions. The primary intention of this pdf file is to understand your methodology in solving this problem.
- Follow PEP 8 – style guide (reference link) for Python code.
- Make your submission using following google form: submission link. You need to upload a zip file containing all codes and the pdf file. You are allowed to make only one submission for this test.
- Timeline: Test commences on 8^{th} April 2023 (10:00 AM). The last date for submission is 10^{th} April 2023 (10:00 AM). Late submissions will not be accepted.
- It is recommended that you send your work, even if it is not a fully optimized solution.

Problem Statement

Consider a ride-hailing system like that of Uber and Ola. The operator owns multiple cars. In this system, customer request for a ride comes from the customer location to the operator. The operator needs to take action by assigning a car or rejecting the request. The operator wants to maximize profit received by serving the requests minus travel costs over a 24-hour period. Use the SUI(Small Untimed Instances) configuration of the gym environment provided by https://github.com/nkullman/ridehailing_package. Your problem statement is to set up a policy for the operator in order to maximize the rewards. The state space, actions and reward structure which is described in the Github repository should be used for solving this problem. Use any Reinforcement learning algorithm to maximize reward in this environment. To this end, you could use any available packages like Baselines, Tensorflow. Share your solution with all codes along with a pdf file explaining the solution approach and details.
Hint: Using gym v0.17.3 works best on this repository.