THE HONG KONG UNIVERSITY OF SCIENCE & TECHNOLOGY
Department of Computer Science and Engineering
COMP4211: Machine Learning
Fall 2020 Assignment 3
Due time and date: 11:59 pm, Dec 7 (Mon), 2020.

IMPORTANT NOTES

- **Your grade will be based on the correctness and clarity.**

- **Late submission: 25 marks will be deducted for every 24 hours after the deadline.**

- **ZERO-Tolerance on Plagiarism: All involved parties will get zero mark.**

The following figure shows a simple problem (actions are in red, R is the immediate reward, SLEEP is the goal state, and the numbers (0.2,0.4,0.4) are probabilities for the three possible transitions). Let the discount factor γ be 0.9.
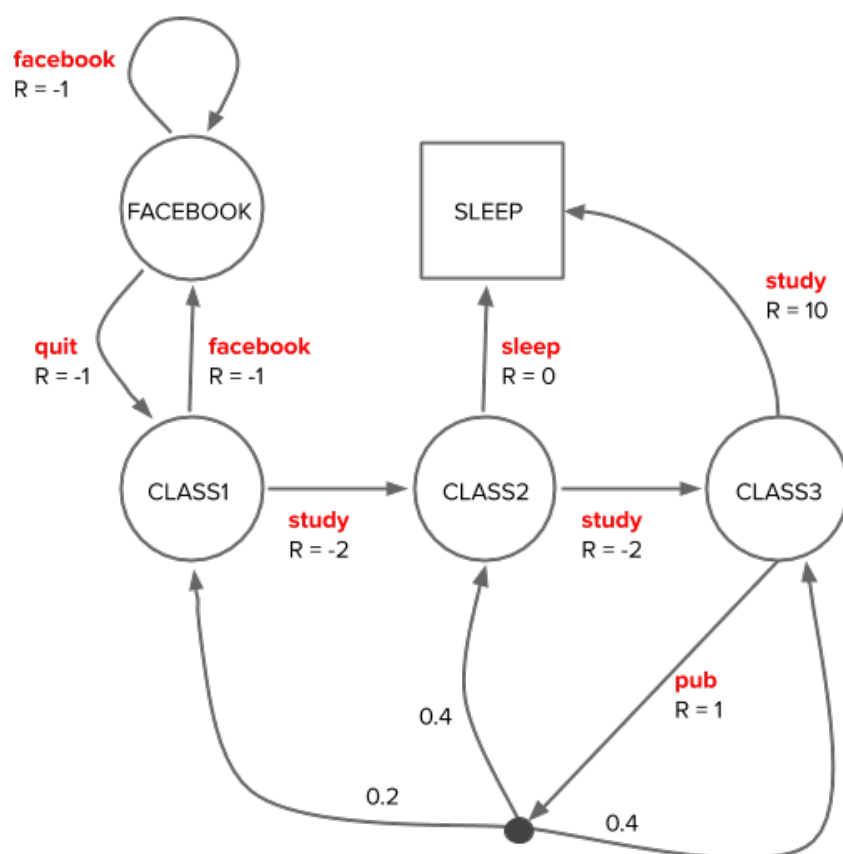
Q1. Consider the following policy π:

- At state FACEBOOK: take action "quit";

- At state CLASS1: take action "study";

- At state CLASS2: take action "study";

- At state CLASS3: take action "pub"

Write down the Bellman equations for the states FACEBOOK, CLASS1, CLASS2 and CLASS3, and then solve for the corresponding V^π values (by hand or by any linear system solver).

Q2. We now try to find the optimal policy by using value iteration.

1. Write down the optimal Bellman condition for the states FACEBOOK, CLASS1, CLASS2 and CLASS3.

2. Use `numpy` to find the optimal policy using value iteration. Show all the state values at each iteration.

Q3. In this question, you try to learn the optimal policy by using Q-learning. The following is a partially learned Q table.

	facebook	quit	study	pub	sleep
FACEBOOK	-1	-1	-	-	-
CLASS1	-1.9	-	-2	-	-
CLASS2	-	-	-1.1	-	0
CLASS3	-	-	10	1	-

You took the following sequence of actions:

$$\text{CLASS3} \xrightarrow{\text{pub}} \text{CLASS3} \xrightarrow{\text{pub}} \text{CLASS2} \xrightarrow{\text{study}} \text{CLASS3} \xrightarrow{\text{pub}} \text{CLASS1} \xrightarrow{\text{study}} \text{CLASS2} \xrightarrow{\text{sleep}} \text{SLEEP}$$

Update the Q table below after each action. Show your steps on how to obtain each table entry clearly.

	facebook	quit	study	pub	sleep
FACEBOOK	?	?	-	-	-
CLASS1	?	-	?	-	-
CLASS2	-	-	?	-	?
CLASS3	-	-	?	?	-

Submission Guidelines

- Note: You should use a `python` notebook for this assignment. Notice that you CANNOT import packages other than `numpy`, `sklearn`, `pytorch`, and `matplotlib`. If you have difficulties with using `markdown` to finish the write-up part, you can also attach an extra pdf file.

- Please submit the Python notebook (ipynb file) with *markdown* summary preserving all outputs. Submit all files to Canvas at once. You do not have to zip them. Note that the submission should be legible, otherwise you may lose some points if the assignment is difficult to read. Plagiarism will lead to zero points.