Exam for course 5

Surname :
Firstname :
Please report your answers on this page only. Questions are on the following pages.

Question 1 (G-5.1) : Exploration / Exploitation

Question 2 (G-5.2) : Supervised versus Reinforcement

1.
2.
3.

Question 3 (G-5.3) : The reward hypothesis

Question 4 (B-5.1) : Gamma

Question 5 (B-5.2) : Policy and Value function

Question 6 (B-5.3) : Q-learning

[Green] Question 1 : Exploration / Exploitation

Indicate which of the following scenari correspond to Exploration :

— Learning the strategy of a particular opponent and try to win by anticipating his moves
— Listen to a playlist of favorite artists from people with different preferences than oneself
— Having a drink once a week in one of your favorite bars
— Buy the same brand of coffee every week

[Green] Question 2 : Supervised versus Reinforcement

For each of the following propositions, report if it corresponds to a case of *supervised* or *reinforcement* learning.

1. The outcome of decisions taken by the algorithm can come much later
2. Learn how to classify medical images of patients versus healthy people by using a feedback provided by experts after each decision.
3. It is essential to consider the temporal dimension when making predictions about new data

[Green] Question 3 : The reward hypothesis

Report the number corresponding to the definition of the reward hypothesis :

1. All goals can be achieved by the maximization of long term rewards
2. All goals can be described by balancing long term and short term rewards
3. All goals can be described by the maximization of expected cumulated reward over time
4. All goals can be described by the minimization of expected cumulated reward over time

[Blue] Question 4 : Gamma

Why choosing a small γ in Q-Learning ? Report the corresponding numbers :

1. To account for stochastic environments
2. To favor short term rewards
3. To account for long term rewards
4. To account for future rewards

[Blue] Question 5 : Policy and Value function

Explain the difference between the policy function and the value function.

[Blue] Question 6 : Q-learning

Define Q-learning.