Introduction to Machine Learning

Week 12

Prof. B. Ravindran, IIT Madras

---

1. (1 Mark) What is the VC dimension of the class of linear classifiers in 2D space?

   (a) 2
   (b) 3
   (c) 4
   (d) None of the above

   Soln. B - Any 3 points can be classified using a linear decision boundary

2. (1 Mark) Which of the following learning algorithms does NOT typically perform empirical risk minimization?

   (a) Linear regression
   (b) Logistic regression
   (c) Decision trees
   (d) Support Vector Machines

   Soln. D - Refer to the lectures

3. (2 Marks) Statement 1: As the size of the hypothesis class increases, the sample complexity for PAC learning always increases.
   Statement 2: A larger hypothesis class has a higher VC dimension.

   Choose the correct option:

   (a) Statement 1 is true. Statement 2 is true. Statement 2 is the correct reason for statement 1
   (b) Statement 1 is true. Statement 2 is true. Statement 2 is not the correct reason for statement 1
   (c) Statement 1 is true. Statement 2 is false
   (d) Both statements are false

   Soln. B - Refer to the lectures

4. (1 Mark) When a model's hypothesis class is too small, how does this affect the model's performance in terms of bias and variance?

   (a) High bias, low variance
   (b) Low bias, high variance
   (c) High bias, high variance
   (d) Low bias, low variance

   Soln. A - Refer to the lectures

5. (1 Mark) Imagine you're designing a robot that needs to navigate through a maze to reach a target. Which reward scheme would be most effective in teaching the robot to find the shortest path?

   (a) +5 for reaching the target, -1 for hitting a wall

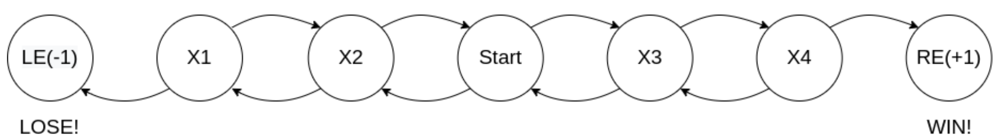
   (b) +5 for reaching the target, -0.1 for every second that passes before the robot reaches the target.

   (c) +5 for reaching the target, -0.1 for every second that passes before the robot reaches the target, +1 for hitting a wall.

   (d) -5 for reaching the target, +0.1 for every second that passes before the robot reaches the target.

   Soln. B - The +5 reward for reaching the target encourages goal achievement, while the -0.1 penalty for each second promotes finding the shortest path. Omitting rewards for hitting walls as question has nothing in this regard.

For the rest of the questions, we will follow a simplistic game and see how a Reinforcement Learning agent can learn to behave optimally in it.
This is our game:



*At the start of the game, the agent is on the **Start** state and can choose to move left or right at each turn. If it reaches the right end(**RE**), it wins and if it reaches the left end(**LE**), it loses.*

*Because we love maths so much, instead of saying the agent wins or loses, we will say that the agent gets a reward of +1 at **RE** and a reward of -1 at **LE**. Then the objective of the agent is simply to maximum the reward it obtains!*

6. (1 Mark) For each state, we define a variable that will store its value. The value of the state will help the agent determine how to behave later. First we will learn this value.

   Let V be the mapping from state to its value.
   Initially,
   V(LE) = -1
   V(X1) = V(X2) = V(X3) = V(X4) = V(Start) = 0
   V(RE) = +1

   For each state $S \in \{X1, X2, X3, X4, Start\}$, with S_L being the state to its immediate left and S_R being the state to its immediate right, repeat:

   $$V(S) = 0.9 \times max(V(S_L), V(S_R))$$

   Till V converges (does not change for any state).

   What is V(X4) after one application of the given formula?

(a) 1
(b) 0.9
(c) 0.81
(d) 0

Soln. B -

$$V(X4) = 0.9 \times max(V(X3), V(RE))$$
$$V(S) = 0.9 \times max(0, +1) = 0.9$$

7. (1 Mark) What is V(X1) after one application of given formula?

(a) -1
(b) -0.9
(c) -0.81
(d) 0

Soln. D -

$$V(X1) = 0.9 \times max(V(LE), V(X2))$$
$$V(S) = 0.9 \times max(-1, 0) = 0$$

8. (2 Marks) What is V(X1) after V converges?

(a) 0.59
(b) -0.9
(c) 0.63
(d) 0

Sol. A - This is the sequence of changes in V:
$V(X4) = 0.9 \rightarrow V(X3) = 0.81 \rightarrow V(Start) = 0.729 \rightarrow V(X2) = 0.656 \rightarrow V(X1) = 0.59$
Final value for X1 is 0.59.