Assignment 1

Introduction to Machine Learning

Prof. B. Ravindran

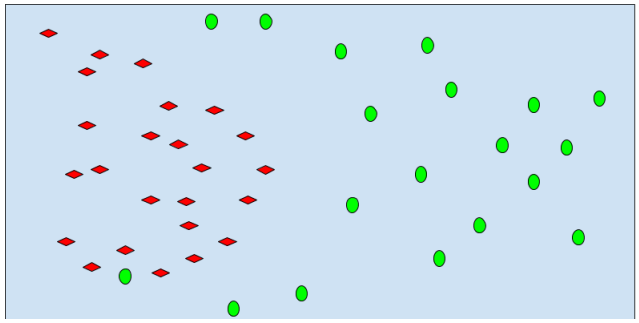
1. Which of the following tasks can be best solved using clustering.

   (a) Predicting the amount of rainfall based on various cues

   (b) Detecting fraudulent credit card transactions

   (c) Training a robot to solve a maze

2. What would be the ideal complexity of the curve which can be used for separating the two classes shown in the image below.



   (a) Linear

   (b) Quadratic

   (c) Cubic

3. **Pavlov's Experiment**

   Pavlov's experiment is a classic experiment conducted by the Russian physiologist Ivan Pavlov. He experiments on dogs shown in the image below.

   Before conditioning dog responds with saliva only in presence of the food but after conditioning it starts salivating just with the bell. Select the correct option(s) about the experiment.

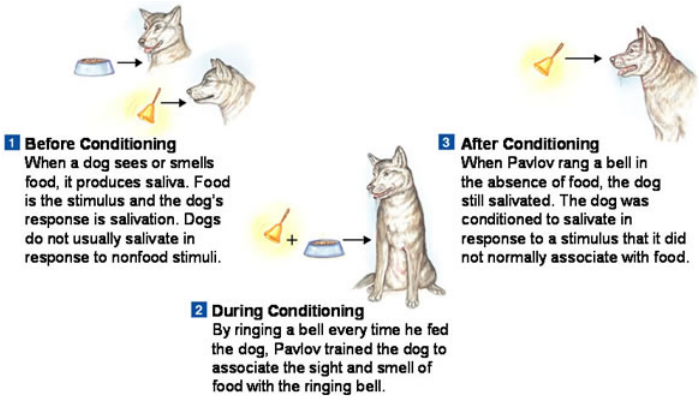
   (a) In this experiment, the dog acts as a Reinforcement learning agent

   (b) In this experiment, the dog learns in a supervised setting

   (c) Comparing this experiment to Reinforcement learning theory, the various states are

   - Presence of just the bell
   - Presence of just food
   - Presence of both food and bell

4. There are n bins of which the kth contains $k-1$ blue balls and $n-k$ red balls. You pick a bin at random and remove two balls at random without replacement. Find the probability that:



(Image source - http://www.goldiesroom.org/)

- the second ball is red;
- the second ball is red, given that the first is red.

(a) 1/3, 2/3
(b) 1/2, 1/3
(c) 1/2, 2/3
(d) 1/3, 1/3

5. A medical company touts its new test for a certain genetic disorder. The false negative rate is small: if you have the disorder, the probability that the test returns a positive result is 0.999. The false positive rate is also small: if you do not have the disorder, the probability that the test returns a positive result is only 0.005. Assume that 2% of the population has the disorder. If a person chosen uniformly from the population is tested and the result comes back positive, what is the probability that the person has the disorder?

   (a) 0.803
   (b) 0.976
   (c) 0.02
   (d) 0.204

6. In an experiment, n coins are tossed, with each one showing up heads with probability p independently of the others. Each of the coins which shows up heads is then tossed again. What is the probability of observing 5 heads in the second round of tosses, if we toss 15 coins in the first round and $p = 0.4$?

   (Hint: First find the mass function of the number of heads observed in the second round.)

   (a) 0.372
   (b) 0.055
   (c) 0.0345
   (d) 0.0488

7. Consider two random variables X and Y having joint density function $f(x, y) = 2e^{-x-y}, 0 < x < y < \infty$. Are X and Y independent? Find the covariance of X and Y.

   (a) Yes, 1/4
   (b) Yes, 1/2
   (c) No, 1/4
   (d) No, 1/2

8. An airline knows that 5 percent of the people making reservations on a certain flight will not show up. Consequently, their policy is to sell 52 tickets for a flight that can hold only 50 passengers. What is the probability that there will be a seat available for every passenger who shows up?

   (a) 0.5101
   (b) 0.81
   (c) 0.6308
   (d) 0.7405

9. Let X have mass function

$$f(x) = \begin{cases} \{x(x+1)\}^{-1} & \text{if x = 1, 2, ...,} \\ 0 & \text{otherwise,} \end{cases}$$

   and let $\alpha \in \mathbb{R}$. For what values of α is it the case that $\mathbb{E}(X^\alpha) < \infty$?

   (a) $\alpha < \frac{1}{2}$
   (b) $\alpha < 1$
   (c) $\alpha > 1$
   (d) $\alpha > \frac{3}{4}$

10. Is the following a distribution function?

$$F(x) = \begin{cases} e^{-1/x} & x > 0 \\ 0 & \text{otherwise} \end{cases}$$

    If so, give the corresponding density function. If not, mention why it is not a distribution function.

    (a) No, not a monotonic function
    (b) Yes, $x^{-2}e^{-1/x}, x > 0$
    (c) No, not right continuous
    (d) Yes, $x^{-1}e^{-1/x}, x > 0$