1. What types of Machine Learning, if any, best describe the following three scenarios:

1. A coin classification system is created for a vending machine. The developers obtain exact coin specifications from the U.S. Mint and derive a statistical model of the size, weight, and denomination, which the vending machine then uses to classify coins.
2. Instead of calling the U.S. Mint to obtain coin information, an algorithm is presented with a large set of labeled coins. The algorithm uses this data to infer decision boundaries which the vending machine then uses to classify its coins.
3. A computer develops a strategy for playing Tic-Tac-Toe by playing repeatedly and adjusting its strategy by penalizing moves that eventually lead to losing.
4. (i) Supervised Learning, (ii) Unsupervised Learning, (iii) Reinforcement Learning
5. (i) Supervised Learning, (ii) Not learning, (iii) Unsupervised Learning
6. (i) Not learning, (ii) Reinforcement Learning, (iii) Supervised Learning
7. (i) Not learning, (ii) Supervised Learning, (iii) Reinforcement Learning
8. (i) Supervised Learning, (ii) Reinforcement Learning, (iii) Unsupervised Learning

* The answer is d.

2. Which of the following problems are best suited for Machine Learning?

(i) Classifying numbers into primes and non-primes.

(ii) Detecting potential fraud in credit card charges.

(iii) Determining the time it would take a falling object to hit the ground.

(iv) Determining the optimal cycle for traffic lights in a busy intersection.

[a] (ii) and (iv)

[b] (i) and (ii)

[c] (i), (ii), and (iii)

[d] (iii)

[e] (i) and (iii)

* The answer is a

3. We have 2 opaque bags, each containing 2 balls. One bag has 2 black balls and the other has a black ball and a white ball. You pick a bag at random and then pick one of the balls in that bag at random. When you look at the ball, it is black. You now pick the second ball from that same bag. What is the probability that this ball is also black?

The answer is d.

Explaination:

Let F = first ball is black

S = second ball is black

Now, we know,

(S|F) = P(S ∩ F) / P(F) = 0.5/(0.5 \* 1 + 0.5\*0.5) = 2/3

Consider a sample of 10 marbles drawn from a bin containing red and green marbles. The probability that any marble we draw is red is µ = 0.55 (independently, with replacement). We address the probability of getting no red marbles (ν = 0) in the following cases:

4. We draw only one such sample. Compute the probability that ν = 0. The closest answer is (‘closest answer’ means: |your answer−given option| is closest to 0):

[a] 7.331 × 10−6

[b] 3.405 × 10−4

[c] 0.289

[d] 0.450

[e] 0.550

We have,

Let P be the probability of getting no red marbles

P = 1- µ

So, P = 1-0.55 = 0.45

So, Probability of getting no red marbles, P = 0.4510 = 3.405 \* 10-4

5. We draw 1,000 independent samples. Compute the probability that (at least) one of the samples has ν = 0. The closest answer is:

[a] 7.331 × 10−6

[b] 3.405 × 10−4

[c] 0.289

[d] 0.450

[e] 0.550

P( at least one v = 0) = 1 – (P(none has v = 0))

= 1 – ( 1 - P that at least one has v = 0)

= 1 – ( 1 - 3.405 \* 10-4 )1000

= 1 – 0.711

= 0.289

6.

Which hypothesis g agrees the most with the possible target functions in terms of the above score?

The remaining points are [101 110 111].  
The output of the target functions be x,y,z where 101 -> x, 110 -> y, 111 -> z.  
  
For hypothesis a.  
-> g returns 1 for all three points,i.e x = 1, y = 1, z = 1.  
  
Target Functions agreeing with the hypothesis on all 3 points = {[ 1 1 1 ]}

Target Functions agreeing with the hypothesis on all 2 points = {[ 1 1 0 ],[ 1 0 1 ],[ 0 1 1 ]}

Target Functions agreeing with the hypothesis on 1 point = {[ 1 0 0 ],[ 0 0 1 ],[ 0 1 0 ]}

Target Functions that do not agree not any points = {[ 0 0 0 ]}

Score for this hypothesis = 3 \* 1 + 2 \* 3 + 1 \* 3 + 0 \* 1 = 12

For hypothesis b.  
-> g returns 0 for all three points.

Hypothesis being opposite of hypothesis a, the score does remain same.  
Score = 3 \* 1 + 2 \* 3 + 1 \* 3 + 0 \* 1 = 12

For hypothesis c.  
-> g is the XOR function applied to x, i.e., if the number of 1s in x is odd, g returns 1; if it is even, g returns 0.

Target Functions agreeing with the hypothesis on all 3 points = {[ 0 0 1 ]}

Target Functions agreeing with the hypothesis on 2 points = {[ 1 0 1 ],[ 0 1 1 ],[ 0 0 0 ]}

Target Functions agreeing with the hypothesis on 1 point = {[ 0 1 0 ],[ 1 0 0 ],[ 1 1 1 ]}

Target Functions agreeing with the hypothesis on 0 points = {[ 1 1 0 ]}

Score for this hypothesis = 3 \* 1 + 2 \* 3 + 1 \* 3 + 0 \* 1 = 12

For hypothesis d.

-> g returns the opposite of the XOR function: if the number of 1s is odd, it returns 0, otherwise returns 1.

This hypothesis being opposite of hypothesis d, the score does remain same.   
Score = 0 \* 1 + 1 \* 3 + 2 \* 3 + 3 \* 1 = 12

For hypothesis e.  
They are all equivalent (equal scores for g in [a] through [d]).

-> Since we got same score for all hypothesis. Hypothesis e is the right answer  
  
-> The answer is e.

7.

<https://github.com/PrekshaK/ML>

Take N = 10. How many iterations does it take on average for the PLA to converge for N = 10 training points? Pick the value closest to your results (again, ‘closest’ means: |your answer − given option| is closest to 0).

* 15

8. Which of the following is closest to P[f(x) 6= g(x)] for N = 10?

* 0.1

9. Now, try N = 100. How many iterations does it take on average for the PLA to converge for N = 100 training points? Pick the value closest to your results.

-> 100

10. Which of the following is closest to P[f(x) 6= g(x)] for N = 100?

-> 0.01