CMPE 257 Machine Learning Spring 2019 HW#1

Submitted By:- Akshata Deo
SJSU ID:- 012565761

Question 2: - What types of Machine Learning, if any, best describe the following scenarios:

a.

Answer:- None. Machine is not learning from the data.

b.

Answer:- Supervised Learning. With the help of given data, algorithm will try to model relationship dependencies between input and output. To be specific it is a binary classification problem.

c.

Answer: - Supervised Learning. With the help of given data, algorithm will try to model relationship dependencies between input and output. To be specific it is a multiclass classification problem.

d.

Answer: - Unsupervised Learning. We do not have output data to model relationship.

e.

Answer: - Reinforcement learning. Agent is learning its behaviour from the reward feedback by performing actions on the environment.

Question 3:

Answer: We know that, $P(A|B) = P(A \cap B)/P(B)$
where, $P(A \cap B)$ is the probability of 2 kittens in a box = ½
and, $P(B)$ is the probability of the white kitten in a box = ¾
So, $P(A|B) = (½ / ¾) = 2/3$

Question 4:

a.

Answer: Given, $\mu = 0.65$.
Probability that any marble we draw is not red = 1-0.65 = 0.35
As each draw is independent (with replacement), the probability to draw one such sample is-

$P = (0.35)^{10} = 0.00002759$

b.

Answer: Probability that at least one v=0 exists among 1000 independent samples will be equal to (1-P[no sample has v=0]).
Probability that no marble drawn is red = 0.35

Probability that no marble drawn is red in a sample = $(0.35)^{10}$

Probability that at least one marble is red = $1-(0.35)^{10}$

probability that (at least) one of the samples has v = 0 will be $(1-(1-(0.35)^{10})^{1000})$ which is equal to 0.0272

Question 5.

Answer: -

Considering the remaining combinations {101, 110, 111} we will get 8 possible target functions as described below: -

X_n	Target Functions f(xn)							
	f(x1)	f(x2)	f(x3)	f(x4)	f(x5)	f(x6)	f(x7)	f(x8)
1 0 1	0	0	0	0	1	1	1	1
1 1 0	0	0	1	1	0	0	1	1
1 1 1	0	1	0	1	0	1	0	1

a. g returns 1 for all three points.

Answer: - As per table above and "score" formula in question, f(x8) is the only target function agreeing with hypothesis on all 3 points. f(x4), f(x6), f(x7) are the 3 functions agreeing with hypothesis on exactly 2 points. f(x2), f(x3), f(x5) are the3 target functions agreeing with hypothesis on exactly 1 point. f(x1) is the only target function agreeing with hypothesis on 0 points.

So,

Score = (1*3) + (3*2) + (3*1) + (1*0) = 12

b. g returns 0 for all three points.

Answer: - f(x1) is the only target function agreeing with hypothesis on all 3 points. f(x2), f(x3), f(x5) are the 3 functions agreeing with hypothesis on exactly 2 points. f(x4), f(x6), f(x7) are the3 target functions agreeing with hypothesis on exactly 1 point. f(x8) is the only target function agreeing with hypothesis on 0 points.

So,

Score = (1*3) + (3*2) + (3*1) + (1*0) = 12

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.

Answer: - f(x2) is the only target function agreeing with hypothesis on all 3 points. f(x1), f(x4), f(x6) are the 3 functions agreeing with hypothesis on exactly 2 points. f(x3), f(x5), f(x8) are the3 target functions agreeing with hypothesis on exactly 1 point. f(x7) is the only target function agreeing with hypothesis on 0 points.

So,

Score = (1*3) + (3*2) + (3*1) + (1*0) = 12

d. g returns the opposite of the XOR function.

Answer: - f(x7) is the only target function agreeing with hypothesis on all 3 points. f(x3), f(x5), f(x8) are the 3 functions agreeing with hypothesis on exactly 2 points. f(x1), f(x4), f(x6) are the3 target

functions agreeing with hypothesis on exactly 1 point. f(x2) is the only target function agreeing with hypothesis on 0 points.

So,

Score = (1*3) + (3*2) + (3*1) + (1*0) = 12

Question 6. The weight update rule is $w(t+1) = w(t) + x(t)y(t)$

(a) Show that $y(t)w^T(t)x(t) < 0$.

Answer. Given, $x(t)$ is misclassified by $w(t)$,

means $w^T(t)x(t)$ & $y(t)$ has the opposite signs. Therefore, product of these two will always be negative.

Hence, $y(t)w^T(t)x(t) < 0$.

(b) Show that $y(t)w^T(t+1)x(t) > y(t)w^T(t)x(t)$.

Answer. According to the update rule:

$w(t+1) = w(t)+y(t)x(t)$

$w^T(t+1) = w^T(t) + x^T(t)y^T(t)$ //take transpose of the whole equation.

$y(t)\, w^T(t+1)\, x(t) = y(t)\, w^T(t)\, x(t) + y(t)\, x^T(t)y^T(t)\, x(t)$ //multiply $y(t)x(t)$ on both sides.

$y(t)w^T(t+1)x(t) = y(t)w^T(t)x(t) + y^2(t)|x(t)|^2$

Since $y^2(t)|x(t)|^2 > 0$,

Therefore, $y(t)w^T(t+1)x(t) > y(t)w^T(t)x(t)$

(c) As far as classifying $x(t)$ is concerned, argue that the move from $w(t)$ to $w(t+1)$ is a move "in the right direction."

Answer. We have already proved-

$y(t)w^T(t)x(t) < 0$,

And $y(t)w^T(t+1)x(t) > y(t)w^T(t)x(t)$

the update rule keep on increases $y(t)w^T(t)x(t)$ until it reaches 0, where $\{x(t),y(t)\}$ is perfectly classified by $w(t)$.

Question 7. Given, the perceptron in 2D: $h(x) = \text{sign}(w^Tx)$

Where $w = [w0, w1, w2]^T$ & $x = [1, x1, x2]^T$.

(a). We express the line by the equation: $x_2 = ax_1 + b$. //equation 1

According to the given data:

$w_0 + w_1x_1 + w_2x_2 = 0$

$x_2 = (-w_1/w_2)x_1 + (-w_0/w_2)$ // equation 2

On comparing equation 1 & 2.

Slope a = $-w_1/w_2$ & Intercept b = $-w_0/w_2$

(b). First, w = $[1,2,3]^T$

So the equation will be:

$1 + 2x_1 + 3x_2 = 0$

$2x_1 + 3x_2 = -1$ // equation 1

If $x_1=0$, then $x_2=-1/3$

& if $x_2=0$, then $x_1=-1/2$

Second, w = $-[1,2,3]^T$

So the equation will be:

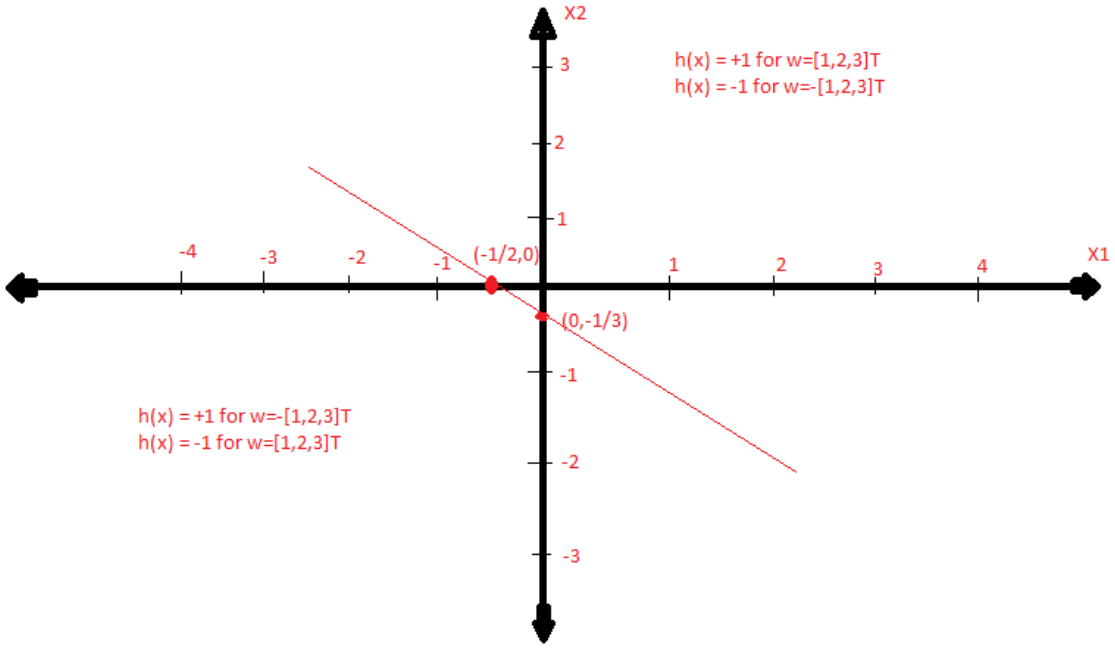
$- 1 – 2x_1 – 3x_2 = 0$

Or $1 + 2x_1 + 3x_2 = 0$

$2x_1 + 3x_2 = -1$ // equation 1

If $x_1=0$, then $x_2=-1/3$

& if $x_2=0$, then $x_1=-1/2$



Question 8. ipynb file attached.

Question 9. ipynb file attached.

Question 10. Choice of Error Measure: Fingerprint Verification: f is +1, if it is authentic user, and f is −1 if it is unauthentic user. Now, error could be of two types: false accept and false reject.

		g	
		+1	-1
f	+1	Correct	false Reject
	-1	false Accept	Correct

Fingerprint Verification for Supermarket:

		g	
		+1	-1
f	+1	0	10
	-1	1	0

$E_{in} = 1/N \sum_{n=1toN}$ [10*false reject + 1*false accept]

Fingerprint Verification for CIA Fingerprint:

		g	
		+1	-1
f	+1	0	1
	-1	1000	0

$E_{in} = 1/N \sum_{n=1toN}$ [1*false reject + 1000*false accept]
