CSE 575: Homework #2

Amey Bhilegaonkar

24/10/2022

Problem 1

a)

Solution: Given $x_1 = (-2, 0)$, $x_2 = (0, 2)$, $x_3 = (2, 2)$, $x_4 = (2, 0)$, $x_5 = (3, -1)$
where x_1, x_2, x_3 are labeled as 1 and x_4, x_5 are labeled as -1.
Primal optimization will be:

$min_{w1,w2,b} \frac{1}{2}||w||^2$; such that;
$t_n(w^T x_n + b) \geq 1$, n = 1, 2..., N and $w = [w1, w2]^T$

Substituting given data points in the above equation we get,

$\min_{w1,w2,b} => -2w + b = \frac{1}{2}w_1^2 + w_2^2; =>$
1. $-2w_1 + b \geq 1,$
2. $2w_2 + b \geq 1,$
3. $-2w_1 - 2w_2 - b \geq 1,$
4. $-2w_1 - b \geq 1,$
5. $-3w_1 + w_2 - b \geq 1,$

b)

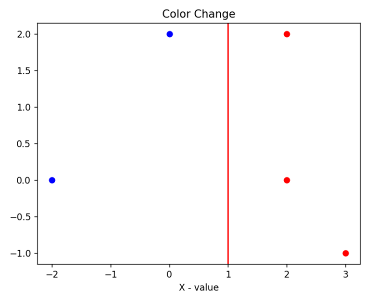
Solution: General dual optimization formulation is as follows:
$argmax_a L(a) = \sum_{n=1}^{N} a_n - \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} a_i a_j t_i t_j x_i^T x_j$ s.t. $a_n \geq 0$ For All n, $\sum_{n=1}^{N} a_n t_n$
Substituting given data points in the above equation we get,

$argmax_{a1a2a3a4a5} a_1 + a_2 + a_3 + a_4 + a_5 - \frac{1}{2}(4a_1^2 + 4a_2^2 + 8a_3^2 + 4a_4^2 + 10a_5^2 + 8a_1a_3 + 8a_1a_4 + 12a_1a_5 - 8a_2a_3 + 4a_2a_5 + 8a_3a_4 + 8a_3a_5 + 12a_4a_5)$

such that, $a1, a2, a3, a4, a5 \geq 0$ and $a_1 + a_2 - a_3 - a_4 - a_5 = 0$

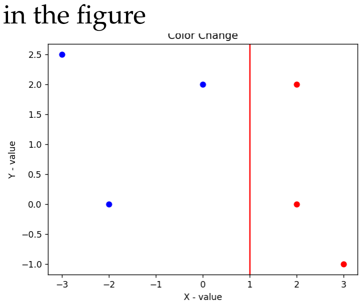
c)

Solution: Decision boundary with graphical representation will be x = 1



d)

Solution: Decision boundary will **NOT** be impacted after the addition of new data point as shown in the figure



Problem 2

a)

Solution: A decision boundary can be drawn with no mis-classifications. Therefore figure (a) satisfies the conditions for a hard-margin linear SVM.

b)

Solution: A soft-margin linear SVM with $C = 0.1$ has more number of mis-classifications compared to higher value of C=10. Therefore figure (d) satisfies the conditions for A soft-margin linear SVM with $C = 0.1$

c)

Solution: A soft-margin linear SVM with C = 10 has a fewer number of misclassifications compared to C = 0.1. Therefore figure (b) satisfies the conditions for a soft-margin linear SVM with C = 10

d)

Solution: A hard-margin kernel SVM with $k(x, z) = (x^T z)^2$ has a decision boundary for a polynomial degree 2 similar to figure (e).

e)

Solution: A hard-margin kernel SVM with $k(x, z) = e^{(-||x-z||^2)}$ has a Gaussian decision boundary with $\gamma = 1$ similar to figure (c).

Problem 3

a)

Solution: The AND logical function can be given as:

$x_1 = 1$ and $x_2 = 1$ then $y = 1$ and for all other combination $y = 0$ Given perception formulation is:
$y(x_1, x_2) = f(w_1x_1 + w_2x_2 + b)$ such that $f(a) = 1$ if $a \geq 0$, $f(a) = 0$, otherwise
We can initialize w_1, w_2, as 1 and b as -1.5, we get from the above equation, the values of y:

x_1, x_2	a	comment
(1, 1)	0.5	$a > 0$
(1, 0)	-0.5	$a < 0$
(0, 1)	-0.5	$a < 0$
(0, 0)	-1.5	$a < 0$

Hence, we can say AND gate is successfully represented with the given weights.

b)

Solution: The NOR logical function can be given as:

$x_1 = 0$ and $x_2 = 0$ then $y = 1$ for other inputs $y = 0$ Given perception formulation is:
$y(x_1, x_2) = f(w_1x_1 + w_2x_2 + b)$ such that $f(a) = 1$ if $a \geq 0$, $f(a) = 0$ otherwise
We can initialize w_1, w_2, as -1 and b as 0.5, we get from the above equation, the values of y:

x_1, x_2	a	comment
(1, 1)	-1.5	$a < 0$
(1, 0)	-0.5	$a < 0$
(0, 1)	-0.5	$a < 0$
(0, 0)	0.5	$a > 0$

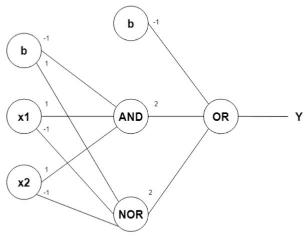
Hence, we can say NOR gate is successfully represented with the given weights.

c)

Solution: The XNOR logical function can be given as:

$y =$ NOT $(x_1 \oplus x_2)$
which simplifies to -
$y =$ OR (AND$(x_1, x_2), NOR(x_1, x_2))$

(i) Since NOR and AND gates both require minimum of one perceptron each, we need atleast 3 perceptrons (one for each - NOR, AND and OR). It would require combination of 3 perceptrons and it can be imagined graphically as:



The truth table for XNOR perceptron:

x_1, x_2	a
(1, 1)	1
(1, 0)	0
(0, 1)	0
(0, 0)	1

(ii).
Given perception formulation is:
$y(x_1, x_2) = f(w_1x_1 + w_2x_2 + b)$ such that $f(a) = 1$ if $a \geq 0$, $f(a) = 0$ otherwise
We can initialize w_1, w_2, as 1 and b as -1 for AND,
We can initialize w_1, w_2, as -1 and b as 1 for NOR,
We can initialize w_1, w_2, as 2 and b as -1 for OR,
we get from the above equation, the values of y:

Result of AND perceptron

x_1, x_2	a
(1, 1)	1
(1, 0)	0
(0, 1)	0
(0, 0)	0

Result of NOR perceptron

x_1, x_2	a
(1, 1)	0
(1, 0)	0
(0, 1)	0
(0, 0)	1

Result of OR perceptron which is the final result of XNOR, here input will be results of AND and NOR

x_1, x_2	a
(1, 0)	1
(0, 0)	0
(0, 0)	0
(0, 1)	1

Hence, we can say XNOR gate is successfully represented with the given weights.

Problem 4

a)

Solution: $z`_1^{[1]} = x_1 w_{1_1}^{[1]} + x_2 w_{1_2}^{[1]} + b_1^{[1]}$

$z`_1^{[1]} = 2.1$
$z_1^{[1]} = \sigma(z`_1^{[1]})$

$z_1^{[1]} = 0.890903$
Similarly,

$z_2^{[1]} = 0.960834$

$z_1^{[2]} = 0.600197$

b)

Solution:
for label (1), the output generated is = 0.600197
$Error(E) = \frac{1}{2}(y(x) - t)^2$
$E = 0.07992133$

(i) $\frac{\partial E}{\partial W_{1_1}^{[2]}} = \frac{\partial E}{\partial z_1^{[2]}} * \frac{\partial Z_1^{[2]}}{\partial z`_1^{[2]}} * \frac{`_1^{[2]}}{\partial W_{1_1}^{[2]}}$
$\frac{\partial E}{\partial W_{1_1}^{[2]}} = -0.08547$

Similarly,

(ii) $\frac{\partial E}{\partial W_{1_2}^{[2]}} = -0.09218$
(iii) $\frac{\partial E}{\partial W_{1_1}^{[1]}} = -0.00233$
(iv) $\frac{\partial E}{\partial W_{1_2}^{[1]}} = -0.00047$
(v) $\frac{\partial E}{\partial W_{2_1}^{[1]}} = -0.00181$
(vi) $\frac{\partial E}{\partial W_{2_2}^{[1]}} = -0.000361$

c)

Solution: $\eta = 0.1$

$w_{1_1}^{[1]} = w_{1_1}^{[1]} - \eta \frac{\partial E}{\partial W_{1_1}^{[1]}}$

$w_{1_1}^{[1]} = 2.500233$
Similarly,

$w_{1_2}^{[1]} = -1.499953$

$w_{2_1}^{[1]} = 3.0000451$

$w_{2_2}^{[1]} = 2.000009026$

$w_{1_1}^{[2]} = 0.5085$

$w_{1_2}^{[2]} = 1.00921$