Home work -3

Part-A

Q1. Multi-Input Feed Forward

Inputs $x_1 = 2,\ x_2 = 1,\ x_3 = 3$

Hidden layer: 2 sigmoid units, 1 sigmoid unit: output

Weights:

$$A = \begin{bmatrix} 0.2 & -0.5 \\ 0.1 & 0.3 \\ -0.2 & 0.8 \\ 0.4 & -0.6 \end{bmatrix} \qquad B = \begin{bmatrix} 0.7 \\ -1.2 \\ 0.5 \end{bmatrix}$$

(a) Hidden Pre-activations and Activations

$$x = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} \qquad z_j = x_1 A_{1j} + x_2 A_{2j} + x_3 A_{3j} + A_{4j}$$

Hidden unit 1: $z_1 = 2(0.2) + (1)(0.1) + 3(-0.2) + 0.4$

$= 0.3$

Hidden unit 2: $z_2 = 2(-0.5) + 1(0.3) + 3(0.8) - 0.6$

$= 1.1$

Activations:

$$a_1 = \sigma(z_1) = \frac{1}{1+e^{-0.3}} = 0.5744$$

$$a_2 = \sigma(z_2) = \frac{1}{1+e^{-1.1}} = 0.7503$$

(b) output Activation y

output pre-activation.

$$z_y = a_1 B_1 + a_2 B_2 + B_3 = 0.5744(0.7) + 0.7503(-1.2) + 0.5$$

$$= 0.0017$$

Activation

$$y = \sigma(z_y) = \frac{1}{1+e^{-0.0017}} = 0.5004$$

Hidden 1 : Pre-activation (0.3), Activation = 0.5744

Hidden 2 Pre-activation (1.1), Activation = 0.7503

output Pre-activation (0.0017) Activation = 0.5004

$z_1 = 0.3$, $z_2 = 1.1$, $a_1 = 0.5744$, $a_2 = 0.7503$

$y = 0.5004$

Q2: XOR with ReLU Network

$h_1 = \text{ReLU}(x_1 + x_2)$, $h_2 = \text{ReLU}(x_1 + x_2 - 1)$ $h_3 = \text{ReLU}(2x_1 - x_2)$

$y = \text{ReLU}(h_1 - 2h_2 + h_3)$, Inputs: (0,0), (0,1), (1,0), (1,1)

(a) outputs for all four XOR inputs, h_1, h_2, h_3, y

(1) (0,0): $h_1 = \text{ReLU}(0+0) = 0$,

$h_2 = \text{ReLU}(0+0-1) = \text{ReLU}(-1) = 0$

$h_3 = \text{ReLU}(2\times 0 - 0) = \text{ReLU}(0) = 0$

$y = \text{ReLU}(0 - 2\times 0 + 0) = \text{ReLU}(0) = 0$

(2) Input (0,1):

$h_1 = \text{ReLU}(0+1) = 1$

$h_2 = \text{ReLU}(0+1-1) = \text{ReLU}(0) = 0$

$h_3 = \text{ReLU}(2*0-1) = \text{ReLU}(-1) = 0$

$y = \text{ReLU}(1-2*0+0) = \text{ReLU}(1) = 1$

$(0,1)$, $h_1 = \text{ReLU}(0+1) = 1$
$h_2 = \text{ReLU}(0+1-1) = 0$, $h_3 = \text{ReLU}(2*0-1)$
$= \text{ReLU}(-1) = 0$

$y = \text{ReLU}(1-2\times0+0) = \text{ReLU}(1) = 1$.

$(1,0)$ $h_1 = \text{ReLU}(1+0) = 1$, $h_2 = \text{ReLU}(1+0-1) = 0$
$h_3 = \text{ReLU}(2*1-0) = \text{ReLU}(2) = 2$.
$y = \text{ReLU}(1-2*0+2) = \text{ReLU}(3) = 3$

$(1,1)$: $h_1 = \text{ReLU}(1+1) = 2$
$h_2 = \text{ReLU}(1+1-1) = \text{ReLU}(1) = 1$
$h_3 = \text{ReLU}(2*1-1) = \text{ReLU}(1) = 1$
$y = \text{ReLU}(2-2*1+1) = \text{ReLU}(1) = 1$

(b) decision boundary with the original 2-hidden unit XOR network

$y_{orig} = \text{ReLU}(h_1 - 2h_2)$

<u>original</u>: output is 1 for $(0,1)$ and $(1,0)$;
0 otherwise (matches XOR)

<u>Extended</u>: Output is 1 for $(0,1)$ and $(1,1)$
3 for $(1,0)$; 0 for $(0,0)$.

So, extended network does not match the original XOR decision boundary.

(c) Does this extension still compute XOR exactly? if not, which inputs differs

A) No it doesnot compute XOR exactly

for input (1,0), output is 3 (should be 1)

For input (1,1) output is 1 (should be 0)

Final answers

(a). outputs for all inputs : 0, 1, 3, 1

(b) The extended network does not match the original XOR decision boundary : (1,0) (1,1) differ.

(c) NO, inputs (1,0), (1,1) differ from correct XOR outputs.

Q3. Decision boundary and Misclassification.

$$y = \begin{cases} 1 & \text{if } w_1 x_1 + w_2 x_2 + b > 0 \\ 0 & \text{otherwise} \end{cases}$$

$w_1 = -1$, $w_2 = 2$, $b = 1$,

Dataset: $(2,1) \rightarrow 1$, $(1,3) \rightarrow 0$, $(3,2) \rightarrow 1$

$(0,1) \rightarrow 0$

(a) Decision boundary: set argument to zero

$w_1x_1 + w_2x_2 + b = 0$

$1 \cdot x_1 + (-2)x_2 + 1 = 0$

$x_1 - 2x_2 + 1 = 0$

$x_1 = 2x_2 - 1$

It is a straight line with slope 2 and intercept -1 in the (x_1, x_2) plane

(b) Classify each point

$x_1 - 2x_2 + 1 > 0$, output is 1, else 0

	output	Label	Correct?
$(2,1) \rightarrow 2 - 2\times1 + 1 = 1$	1	1	yes
$(1,3) \rightarrow 1 - 2\times3 + 1 = -4$	0	0	yes
$(3,2) \rightarrow 3 - 2\times2 + 1 = 0$	0	1	$\boxed{NO}$
$(0,1) \rightarrow 0 - 2\times1 + 1 = -1$	0	0	yes

Misclassified point: $(3, 2)$, → should be 1,

Predicted → 0

(c) Perceptron loss → only $(3,2)$ is misclassified

④ Multi-Layer Forward Pass (Matrix style)

$(x_1, x_2) = (2, 3)$ 2 hidden layers, 2 sigmoid units

(a) Hidden activations in layer 1

$$A^{(1)} = \begin{bmatrix} 0.2 & -0.3 \\ 0.4 & 0.1 \\ 0.5 & -0.6 \end{bmatrix}$$

$$z_1^{(1)} = 0.2 + (-0.3) \cdot 2 + 0.4(3)$$

$$= 0.2 - 0.6 + 1.2$$

$$a_1^{(1)} = \sigma(0.8) = \frac{1}{1+e^{-0.8}} = 0.68997$$

Unit . 2

$$z_2^{(1)} = 0.5 + (-0.6) \cdot 2 + 0.1 \times 3$$

$$= 0.5 - 1.2 + 0.3$$

$$= -0.4$$

$$a_2^{(1)} = \sigma(-0.4) = \frac{1}{1+e^{0.4}} = 0.40131$$

Layer -1: activations, $(0.6897, 0.40131)$

(b) Layer: 2 $A^{(2)} = \begin{bmatrix} 0.7 & -0.5 \\ -0.2 & 0.3 \\ 0.1 & 0.4 \end{bmatrix}$

$$z_1^{(2)} = 0.7 + (-0.5)(0.68997) + (-0.2) \cdot (0.40131)$$

$$= 0.2747$$

$$a_1^{(2)} = \sigma(0.274753) = 0.568$$

Unit-2

$$z_2(2) = 0.1 + 0.4 \cdot 0.68997 + 0.3 \cdot 0.40131$$

$$= 0.496381$$

$$a_2(2) = \sigma(0.496381) = 0.62113$$

Layer 2 activations: $(0.56824, 0.62113)$

(c) Final output

output weights $(1.0, -1.2, 0.3)$

$$y = 1.0 + (-1.2)(0.56824) + 0.3(0.62113)$$

$$\boxed{y = 0.504}$$

Qs. Linear SVM:

Positive points $(y = +1)$: $P_1 = (1,3)$, $P_2 = (2,2)$

Negative points $(y = -1)$: $n_1 = (0,0)$

(1) Augment each point with a bias term

$$P_1 = (1,3) = (1,3,1)$$

$$P_2 = (2,2) = (2,2,1)$$

$$n_1 = (0,0) = (0,0,1)$$

(2) Dual constraint equations for $\alpha_1, \alpha_2, \alpha_3$

α_1, α_2 are positive, α_3 for negative

$$y_i(\bar{w}^T \cdot x_i) = 1$$

$$y_1 = +1, \ y_2 = +1, \ y_3 = -1$$

$$\alpha_1 + \alpha_2 - \alpha_3 = 0$$

(3) solve for dual variables

$$\bar{w} = \sum_{i=1}^{3} \alpha_i y_i \bar{x}_i$$

$$\bar{w} = \alpha_1 [+1] \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix} + \alpha_2 [+1] \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix} + \alpha_3 [-1] \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$= \alpha_1 \cdot 1 + \alpha_2 \cdot 2, \ \alpha_1 \cdot 3 + \alpha_2 \cdot 2, \ \alpha_1 \cdot 1 + \alpha_2 \cdot 1 - \alpha_3$$

Lets denote $w_1 = \alpha_1 + 2\alpha_2$

$$w_2 = 3\alpha_1 + 2\alpha_2$$

$$b = \alpha_1 + \alpha_2 - \alpha_3$$

$$\alpha_1 + \alpha_2 - \alpha_3 = 0 \Rightarrow b = 0$$

$w_1 + 3w$ Now apply margin conditions

$$P_1 = (1, 3, 1), \ y = +1;$$

$$(1, 3, 1) \cdot [w_1, w_2, b] = w_1 + 3w_2 + b = 1$$

$w_1 + 3w_2 = 1$

For $p_2 = (2,2), y = +1$;

$2w_1 + 2w_2 + b = 1$ $\Rightarrow 2w_1 + 2w_2 = 1$

For $n = (0,0), y = -1$

$-(0w_1 + 0w_2 + b) = 1$

$\Rightarrow -b = 1$ $\Rightarrow \boxed{b = -1}$

For possitive points: $w_1 x_1 + w_2 x_2 + b = 1$

For negative points: $w_1 x_1 + w_2 x_2 + b = -1$

For $p_1 = (1,3)$: $w_1 \cdot 1 + w_2 \cdot 3 + b = 1$

$p_2 = (2,2)$: $w_1 \cdot 2 + w_2 \cdot 2 + b = 1$

$n_1 = (0,0) = w_1 \cdot 0 + w_2 \cdot 0 + b = -1$

$$\boxed{b = -1}$$

$w_1 + 3w_2 - 1 = 1 \Rightarrow w_1 + 3w_2 = 2$

$2w_1 + 2w_2 - 1 = 1 \Rightarrow 2w_1 + 2w_2 = 2$

Solving these 2 questions

$2w_1 + 2w_2 = 2 \Rightarrow w_1 + w_2 = 1$

$(w_1 + 3w_2) - (w_1 + w_2) = 2 - 1$

$\Rightarrow 2w_2 = 1$

$w_1 = 1 - w_2 = 1 - 0.5 = 0.5$

$w_1 = 0.5,\ w_2 = 0.5,\ b = -1$

(4) weight vector $\bar{w} = (0.5, 0.5)$

(5) $0.5x_1 + 0.5x_2 - 1 = 0$

or $x_1 + x_2 = 2$

(6) For $p_1 = (1, 3)$

$0.5 \times 1 + 0.5 \times 3 - 1 = 1$

For $p_2\ (2, 2)$.

$0.5 \times 2 + 0.5 \times 2 - 1 = 1 + 1 - 1 = 1$

For $n_1 = (0, 0)$: $0.5 \times 0 + 0.5 \times 0 - 1 = -1$

Answers:

(1) Augmented points: $(1, 3, 1)$, $(2, 2, 1)$, $(0, 0, 1)$

(2) Constraint: $\alpha_1 + \alpha_2 - \alpha_3 = 0$

(3) Solution: $w_1 = 0.5$, $w_2 = 0.5$, $b = -1$

(4) weight vector: $(0.5, 0.5)$

(5) Hyperplane: $x_1 + x_2 = 2$

(6) Margin conditions verified.

⑥ Nonlinear SVM:

$S_1 = (1,0)$, $y_1 = -1$

$S_2 = (2,1)$, $y_2 = +1$

$$\phi_1\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{cases} \begin{pmatrix} 4-x_2+|x_1-x_2| \\ 4-x_1+|x_1-x_2| \end{pmatrix} & \text{if } \sqrt{x_1^2+x_2^2} > 2 \\ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} & \text{otherwise.} \end{cases}$$

$x_1 = 1$, $x_2 = 0$,

$|x_1 - x_2| = |1-0| = 1$

$$\phi(S_1) = \begin{pmatrix} 4-0+1 \\ 4-1+1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 5 \\ 4 \\ 1 \\ 0 \end{pmatrix}$$

For, $x_1 = 2$, $x_2 = 1$

$|x_1 - x_2| = |2-1| = 1$

$$\phi(S_2) = \begin{pmatrix} 4-1+1 \\ 4-2+1 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \\ 2 \\ 1 \end{pmatrix}$$

② Margin equations in the transformed space

Using α_1, α_2

$w = \alpha_1 y_1 \phi_1 + \alpha_2 y_2 \phi_2$

$y_i(w^T\phi_i + b) = 1$ for support vectors

For support vectors

For S_1,

$-1(w^T \phi(S_1) + b) = 1 \Rightarrow w^T \phi(S_1) + b = -1$

For S2

$+1(w^T \phi(S_2) + b) = 1 \Rightarrow w^T \phi(S_2) + b = 1$

③ Solve for α_1, α_2

$$\phi_1 = \phi(S_1) = \begin{pmatrix} 5 \\ 4 \\ 1 \\ 0 \end{pmatrix}$$

$$\phi_2 = \phi(S_2) = \begin{pmatrix} 4 \\ 3 \\ 2 \\ 1 \end{pmatrix}$$

$w = -\alpha_1 \phi_1 + \alpha_2 \phi_2$

$w^T \phi_1 + b = -1 \quad w^T \phi_2 + b = 1$

$-\alpha_1(42) + \alpha_2(34) + b = -1$

$-\alpha_1(34) + \alpha_2(30) + b = 1$

substract 2 from 1

$[-42\alpha_1 + 34\alpha_2 + b] - [-34\alpha_1 + 30\alpha_2 + b] =$

also for SVM, $\sum \alpha_i y_i = 0$

$-\alpha_1 + \alpha_2 = 0,$

$\Rightarrow \alpha_1 = \alpha_2$

$4\alpha_1 - 2\alpha_1 = 1$

$2\alpha_1 = 1 \quad \boxed{\alpha_1 = \alpha_2 = \frac{1}{2}}$

(4)

$w = -\alpha_1 \phi_1 + \alpha_2 \phi_2$

$= \frac{1}{2}(-\phi_1 + \phi_2)$

$w = \frac{1}{2}\left(\begin{pmatrix} 4 \\ 3 \\ 2 \\ 1 \end{pmatrix} - \begin{pmatrix} 5 \\ 4 \\ 1 \\ 0 \end{pmatrix}\right)$

$= \frac{1}{2}\begin{pmatrix} -1 \\ -1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -0.5 \\ -0.5 \\ 0.5 \\ 0.5 \end{pmatrix}$

⑤ $\cdot f(x) = w^T \phi(x) + b$

$w^T \phi_1 + b = -1$

$b = -1 - w^T \phi_1$

$w^T \phi_1 = (-0.5) \times 5 + (-0.5) \times 4 + 0.5 \times 1 + 0.5 \times 0$

$= -4$

$b = -1 - (-4) = 3$

$f(x) = w^T \phi(x) + b = \begin{pmatrix} -0.5 \\ -0.5 \\ 0.5 \\ 0.5 \end{pmatrix} \phi(x) + 3$

$f(x) = 0.5\phi_1(x) - 0.5\phi_2(x) + 0.5\phi_3(x) + 0.5$

⑥ $q = (1, 1)$

$\sqrt{1^2+1^2} = \sqrt{2} = 1.41 < 2$ below threshold

so, $\phi(q)$ is undefined or ignored.

It is only valid for $\|x\| \geq 2$