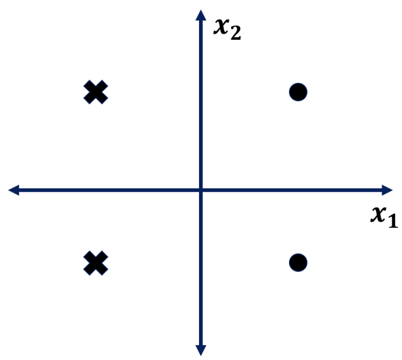
Exam Prep 12 Solutions

Q1. Linear Separability

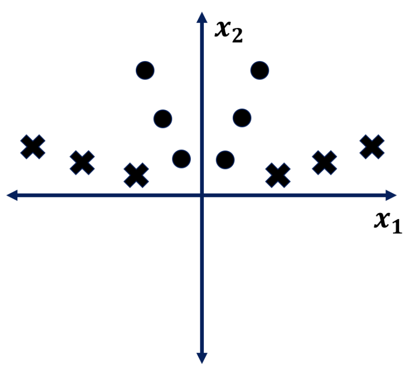
(a) For each of the datasets represented by the graphs below, please select the feature maps for which the perceptron algorithm can perfectly classify the data.

Each data point is in the form (x_1, x_2), and has some label Y, which is either a 1 (dot) or -1 (cross).



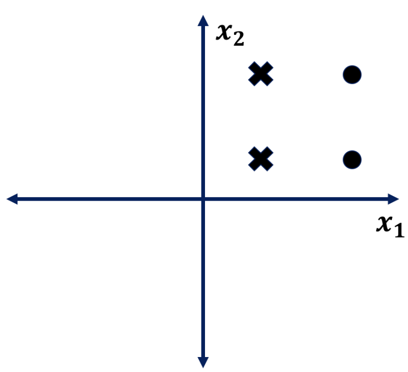
(i)

- ■ $[x_1 \quad x_2 \quad 1]$
- ■ $[x_1 \quad x_2 \quad x_1^2]$
- ■ $[x_1 \quad x_2 \quad |x_1|]$
- ■ $[x_1 \quad x_2 \quad Y]$
- ■ $[x_1 \quad x_2]$



(ii)

- □ $[x_1 \quad x_2 \quad 1]$
- ■ $[x_1 \quad x_2 \quad x_1^2]$
- ■ $[x_1 \quad x_2 \quad |x_1|]$
- ■ $[x_1 \quad x_2 \quad Y]$
- □ $[x_1 \quad x_2]$



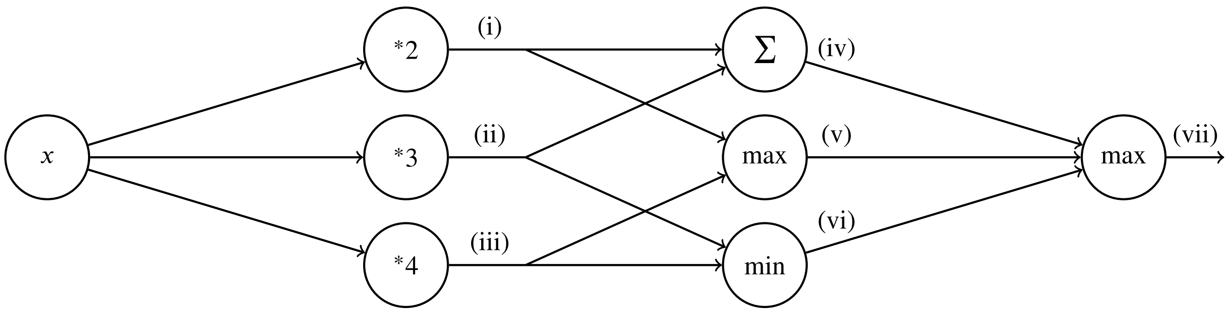
(iii)

- ■ $[x_1 \quad x_2 \quad 1]$
- ■ $[x_1 \quad x_2 \quad x_1^2]$
- □ $[x_1 \quad x_2 \quad |x_1|]$
- ■ $[x_1 \quad x_2 \quad Y]$
- □ $[x_1 \quad x_2]$

Q2. Deep Learning

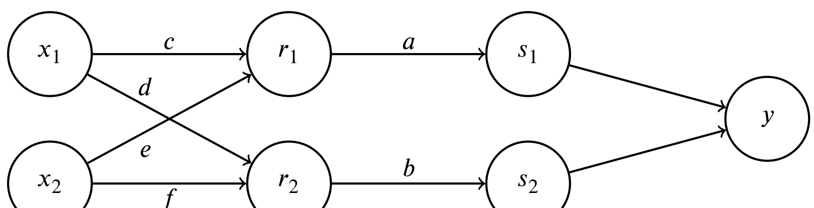
(a) Perform forward propagation on the neural network below for $x = 1$ by filling in the values in the table. Note that (i), ..., (vii) are outputs after performing the appropriate operation as indicated in the node.

(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)
2	3	4	5	4	3	5



(b) **[Optional]** Below is a neural network with weights a, b, c, d, e, f. The inputs are x_1 and x_2.
The first hidden layer computes $r_1 = \max(c \cdot x_1 + e \cdot x_2, 0)$ and $r_2 = \max(d \cdot x_1 + f \cdot x_2, 0)$.
The second hidden layer computes $s_1 = \frac{1}{1+\exp(-a \cdot r_1)}$ and $s_2 = \frac{1}{1+\exp(-b \cdot r_2)}$.
The output layer computes $y = s_1 + s_2$. Note that the weights a, b, c, d, e, f are indicated along the edges of the neural network here.

Suppose the network has inputs $x_1 = 1, x_2 = -1$.
The weight values are $a = 1, b = 1, c = 4, d = 1, e = 2, f = 2$.
Forward propagation then computes $r_1 = 2, r_2 = 0, s_1 = 0.9, s_2 = 0.5, y = 1.4$. Note: some values are rounded.



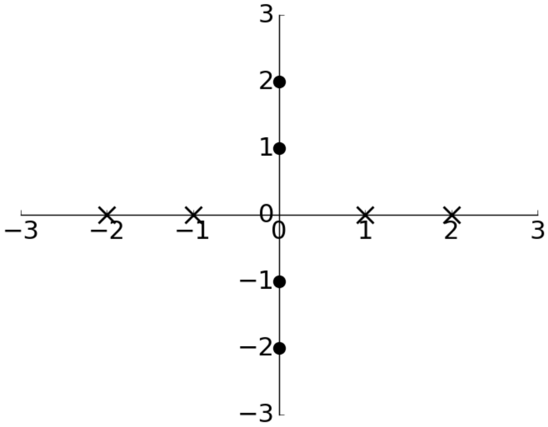
Using the values computed from forward propagation, use backpropagation to numerically calculate the following partial derivatives. Write your answers as a single number (not an expression). You do not need a calculator. Use scratch paper if needed.
Hint: For $g(z) = \frac{1}{1+\exp(-z)}$, the derivative is $\frac{\partial g}{\partial z} = g(z)(1 - g(z))$.

$\frac{\partial y}{\partial a}$	$\frac{\partial y}{\partial b}$	$\frac{\partial y}{\partial c}$	$\frac{\partial y}{\partial d}$	$\frac{\partial y}{\partial e}$	$\frac{\partial y}{\partial f}$
0.18	0	0.09	0	−0.09	0

$$
\begin{aligned}
\frac{\partial y}{\partial a} &= \frac{\partial y}{\partial s_1}\frac{\partial s_1}{\partial a} \\
&= 1 \cdot \frac{\partial g(a \cdot r_1)}{\partial a} \\
&= r_1 \cdot g(a \cdot r_1)(1 - g(a \cdot r_1)) \\
&= r_1 \cdot s_1(1 - s_1) \\
&= 2 \cdot 0.9 \cdot (1 - 0.9) \\
&= 0.18 \\
\frac{\partial y}{\partial b} &= \frac{\partial y}{\partial s_2}\frac{\partial s_2}{\partial b} \\
&= 1 \cdot \frac{\partial g(b \cdot r_2)}{\partial b} \\
&= r_2 \cdot g(b \cdot r_2)(1 - g(b \cdot r_2)) \\
&= r_2 \cdot s_2(1 - s_2) \\
&= 0 \cdot 0.5(1 - 0.5) \\
&= 0 \\
\frac{\partial y}{\partial c} &= \frac{\partial y}{\partial s_1}\frac{\partial s_1}{\partial r_1}\frac{\partial r_1}{\partial c} \\
&= 1 \cdot [a \cdot g(a \cdot r_1)(1 - g(a \cdot r_1))] \cdot x_1 \\
&= [a \cdot s_1(1 - s_1)] \cdot x_1 \\
&= [1 \cdot 0.9(1 - 0.9)] \cdot 1 \\
&= 0.09 \\
\frac{\partial y}{\partial d} &= \frac{\partial y}{\partial s_2}\frac{\partial s_2}{\partial r_2}\frac{\partial r_2}{\partial d} \\
&= \frac{\partial y}{\partial s_2}\frac{\partial s_2}{\partial r_2} \cdot 0 \\
&= 0 \\
\frac{\partial y}{\partial e} &= \frac{\partial y}{\partial s_1}\frac{\partial s_1}{\partial r_1}\frac{\partial r_1}{\partial e} \\
&= 1 \cdot [a \cdot g(a \cdot r_1)(1 - g(a \cdot r_1))] \cdot x_2 \\
&= [a \cdot s_1(1 - s_1)] \cdot x_2 \\
&= [1 \cdot 0.9(1 - 0.9)] \cdot -1 \\
&= -0.09 \\
\frac{\partial y}{\partial f} &= \frac{\partial y}{\partial s_2}\frac{\partial s_2}{\partial r_2}\frac{\partial r_2}{\partial f} \\
&= \frac{\partial y}{\partial s_2}\frac{\partial s_2}{\partial r_2} \cdot 0 \\
&= 0
\end{aligned}
$$

(c) Below are two plots with horizontal axis x_1 and vertical axis x_2 containing data labelled $\times$ and $\bullet$. For each plot, we wish to find a function $f(x_1, x_2)$ such that $f(x_1, x_2) \geq 0$ for all data labelled $\times$ and $f(x_1, x_2) < 0$ for all data labelled $\bullet$.

Below each plot is the function $f(x_1, x_2)$ for that specific plot. Complete the expressions such that all the data is labelled correctly. If not possible, mark "No valid combination".



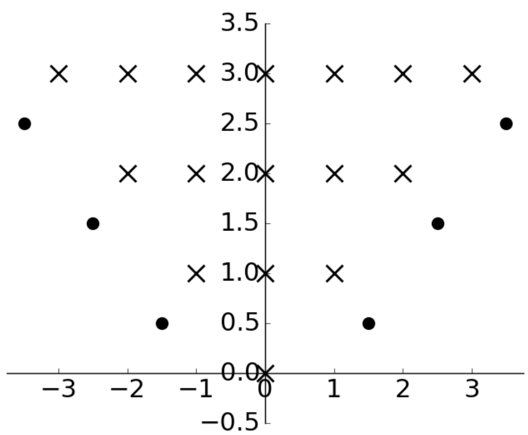
$f(x_1, x_2) = \max(\underline{\textbf{(i)}} + \underline{\textbf{(ii)}}, \underline{\textbf{(iii)}} + \underline{\textbf{(iv)}}) + \underline{\textbf{(v)}}$

(i)	● x_1	○ $-x_1$	○ 0
(ii)	○ x_2	○ $-x_2$	● 0
(iii)	○ x_1	● $-x_1$	○ 0
(iv)	○ x_2	○ $-x_2$	● 0
(v)	○ 1	● -1	○ 0

○ No valid combination

There are two possible solutions:

$$f(x_1, x_2) = \max(x_1, -x_1) - 1$$
$$f(x_1, x_2) = \max(-x_1, x_1) - 1$$



$f(x_1, x_2) = \underline{\textbf{(vi)}} - \max(\underline{\textbf{(vii)}} + \underline{\textbf{(viii)}}, \underline{\textbf{(ix)}} + \underline{\textbf{(x)}})$

(vi)	● x_2	○ $-x_2$	○ 0
(vii)	● x_1	○ $-x_1$	○ 0
(viii)	○ x_2	○ $-x_2$	● 0
(ix)	○ x_1	● $-x_1$	○ 0
(x)	○ x_2	○ $-x_2$	● 0

○ No valid combination

There are four possible solutions:

$$f(x_1, x_2) = x_2 - \max(x_1, -x_1)$$
$$f(x_1, x_2) = x_2 - \max(-x_1, x_1)$$
$$f(x_1, x_2) = -\max(x_1 - x_2, -x_1 - x_2)$$
$$f(x_2, x_2) = -\max(-x_1 - x_2, x_1 - x_2)$$