Birla Institute of Technology and Science, Pilani

Work Integrated Learning Programmes Division

M. Tech. in AI & ML

I Semester 2023-2024

Mid-Semester Test
(EC2 - Regular)

Course Number	AIMLCZG51
Course Name	DEEP NEURAL NETWORK
Nature of Exam	Closed Book
Weight-age for grading	30
Duration	2 hrs

# Pages	2
# Questions	6

Instructions

1. All questions are compulsory.
2. All answers must be directed to the question in short and simple paragraphs or bullet points; use visuals/diagrams wherever necessary.
3. Assumptions made if any, should be stated clearly at the beginning of your answer.

---

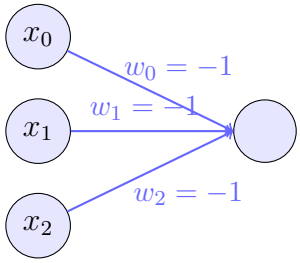
1. Given a truth table, design a perceptron and find the weights and threshold of the perceptron. [5]

x_1	x_2	y
0	0	1
1	0	0
0	1	0
1	1	0

Rubrics and one solution

x_1	x_2	y	h
-1	-1	1	$w_0 + w_1(-1) + w_2(-1) > 0$
1	-1	-1	$w_0 + w_1(1) + w_2(-1) < 0$
-1	1	-1	$w_0 + w_1(-1) + w_2(1) < 0$
1	1	-1	$w_0 + w_1(1) + w_2(1) < 0$

$$w_0 - w_1 - w_2 > 0 \quad (1)$$
$$w_0 + w_1 - w_2 < 0 \quad (2)$$
$$w_0 - w_1 + w_2 < 0 \quad (3)$$
$$w_0 + w_1 + w_2 < 0 \quad (4)$$
$$(2) + (3) \rightarrow 2w_0 < 0 \rightarrow w_0 < 0 \quad (5)$$
$$(2) + (4) \rightarrow 2w_0 + 2w_1 < 0 \rightarrow w_1 < 0 \quad (6)$$
$$(3) + (4) \rightarrow 2w_0 + 2w_2 < 0 \rightarrow w_2 < 0 \quad (7)$$



- Table [1]
- Equations 1 to 4 [1]
- Equations 5 to 7 [1]
- Any value of weight and threshold satisfying equations 5 to 7 [1]
- Perceptron diagram [1]

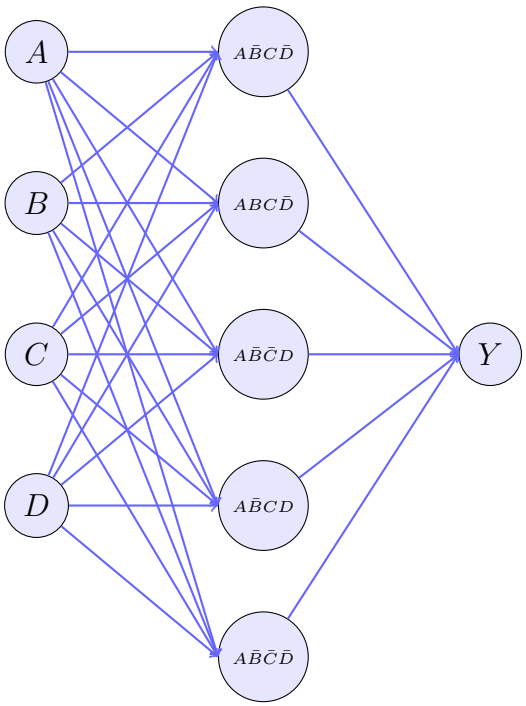
2. Construct an MLP for the given boolean expression. What is the depth and width of the MLP? [5]

A	B	C	D	Y
1	0	1	0	1
1	1	1	0	1
1	0	0	1	1
1	0	1	1	1
1	0	0	0	1

- Write the boolean expression [1]

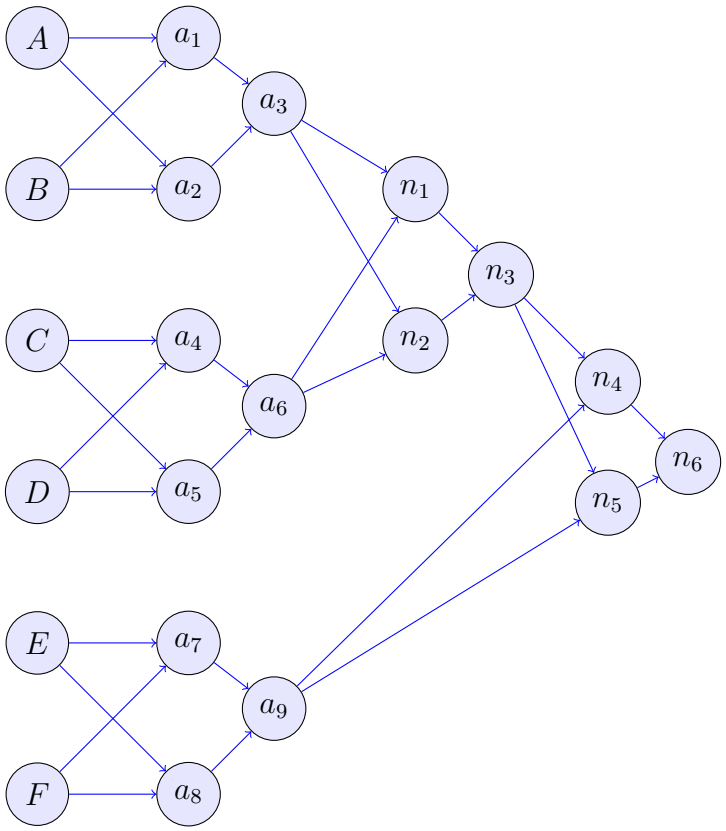
$$Y = A\bar{B}C\bar{D} + ABC\bar{D} + A\bar{B}\bar{C}D + A\bar{B}CD + A\bar{B}\bar{C}\bar{D}$$

- Draw one node for each expression in layer 1 [1]
- Connect all nodes to a single node in layer 2 [1]
- Depth = 2 [1]
- Width = 5 [1]



3. Find the number of layers and number of perceptrons required for an MLP that can be used for computing the XOR of 6 parameters. Draw the MLP network. If a single hidden layer is used, how many perceptrons will be required? [5]

- n = 6
- Number of perceptrons $= 3(6-1) = 3(6-1) = 15$ [1]
- Number of layers $= 2\log_2 6 = 2\log_2 6 = 2 * \log 6/\log 2 = 2 * ceil(2.5) = 2 * 3 = 6$ [1]
- Network - multiple answers expected. [2]
- Single layer; number of perceptrons $= 2^{n-1} = 2^5 = 32$ [1]



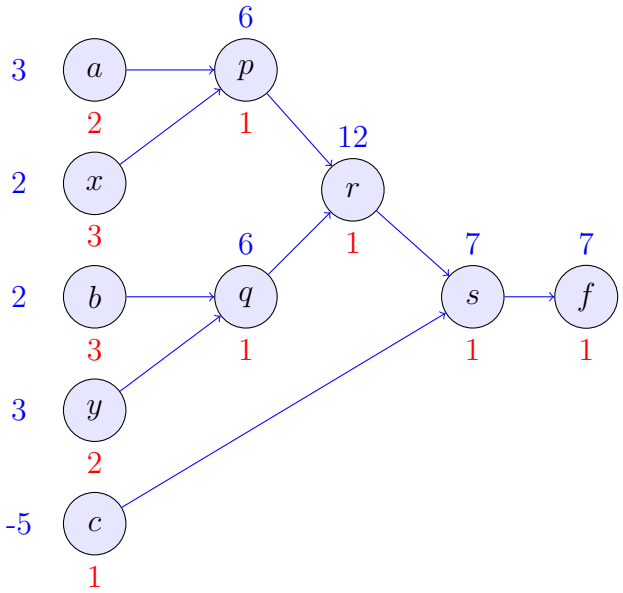
4. Derive the equation for the derivative of binary cross entropy loss L with respect to the weighted sum Z, assuming activation $a = \sigma(Z)$ and the loss L is computed from this activation A. [5]
Each equation, 1 mark.

$$
\begin{aligned}
L &= -[Y \cdot \log(A) + (1-Y) \cdot \log(1-A)] \\
\frac{\partial L}{\partial Z} &= \frac{\partial L}{\partial A} \cdot \frac{\partial A}{\partial Z} \\
\frac{\partial L}{\partial A} &= \frac{-y}{A} + \frac{1-Y}{1-A} \\
\frac{\partial A}{\partial Z} &= A \cdot (1-A) \\
\frac{\partial L}{\partial Z} &= \left(\frac{-Y}{A} + \frac{1-Y}{1-A}\right) A \cdot (1-A) = (A-Y)
\end{aligned}
$$

5. Draw the computational graph for the equation $f = relu(ax+by+c)$. Show the computations of derivatives of f wrt a, b, c in the graph. Using the graph, compute the value of f and the derivatives if $a = 3, b = 2, c = (-5), x = 2$ and $y = 3$. [5]

- FP graph (blue digits) 1 mark
- BP graph (red digits) 1 mark
- Output computation 1 mark
- Gradient computation 1 mark

$$
\begin{array}{lll}
p = ax & \dfrac{\partial p}{\partial a} & = x \\
q = by & \dfrac{\partial q}{\partial b} & = y \\
r = p + q & \dfrac{\partial r}{\partial p} = \dfrac{\partial r}{\partial q} & = 1 \\
s = r + c & \dfrac{\partial s}{\partial r} = \dfrac{\partial s}{\partial c} & = 1 \\
f = relu(s) & \dfrac{\partial f}{\partial s} & = 1
\end{array}
$$



6. Given an error surface,

$$E(p, q, r) = 3p^3 + 3q^2 + 4r + 5$$

compute the optimal value of learning rate that will minimize the error surface, the largest learning rate for convergence and smallest learning rate for divergence. [5]

$$\frac{\partial E}{\partial p} = 9p \quad \frac{\partial E}{\partial q} = 6q \quad \frac{\partial E}{\partial r} = 4 \qquad 1 \text{ mark}$$

$$\eta_p = 1/9 \quad \eta_q = 1/6 \quad \eta_r = 1/4 \qquad 1 \text{ mark}$$

Optimal learning rate for convergence $\eta_{opt} = \min[\eta_p, \eta_q, \eta_r] = 1/9$ 1 mark

Largest learning rate for convergence $= \min[2\eta_p, 2\eta_q, 2\eta_r] = 2/9$ 1 mark

Smallest Learning rate for divergence $> 2\eta_{opt} = 2 * 1/9 = 0.22$ 1 mark