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 - Makeup)

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. Consider a single layer perceptron having 2 inputs and 1 output. Let threshold be 0.5, learning rate be 0.6, bias be -2 and weight values $w_1 = 0.3$ and $w_2 = 0.7$. Given the input patterns in the table, compute the value of the output and train using perceptron learning rule for one epoch. [5]

E.g. #	x_1	x_2	y
1	1	1	+1
2	1	0	+1
3	0	1	-1
4	0	0	+1

Rubrics and answer

- z - 0.25, h - 0.25, Δ - 0.25, new - 0.25 marks
- Each row in table carries 1 mark each.
- Any wrong equation or wrong computation, reduce appropriately, but maintain a minimum of 0.25 for each row.

$$z = w_1x_1 + w_2x_2 + b$$
$$h = sign(z)$$
$$\Delta w_1 = \eta(t-h)x_1$$
$$\Delta w_2 = \eta(t-h)x_2$$
$$\Delta b = \eta(t-h)$$
$$w_{new} \leftarrow w_{old} + \Delta w$$

Epoch 1

x_1	x_2	y	b	w_1	w_2	z	h	y, h	Δ	new b,w_1,w_2
1	1	+1	-2	0.3	0.7	0.3 + 0.7 - 2 =-1	-1	N	$\delta b = 0.6(1-(-1)) = 1.2$ $\delta w_1 = 0.6(1+1)1 = 1.2$ $\delta w_2 = 0.6(1+1)1 = 1.2$	$b = -2 + 1.2 = 0.8$ $w_1 = 0.3 + 1.2 = 1.5$ $w_2 = 0.7 + 1.2 = 1.9$
1	0	+1	0.8	1.5	1.9	1.5 + 0 + 0.8 =2.3	+1	Y		
0	1	-1	0.8	1.5	1.9	0 + 1.9 + 0.8 = 2.7	+1	N	$\delta b = 0.6(-1-1) = -1.2$ $\delta w_1 = 0.6(-1-1)0 = 0$ $\delta w_2 = 0.6(-1-1)1 = -1.2$	$b = 0.8 - 1.2 = -0.4$ $w_1 = 1.5$ $w_2 = 1.9 - 1.2 = 0.7$
0	0	+1	-0.4	1.5	0.7	0+0-0.4 = -0.4	-1	N	$\delta b = 0.6(1+1) = 1.2$ $\delta w_1 = 0.6(1+1)0 = 0$ $\delta w_2 = 0.6(1+1)0 = 0$	$b = -0.4 + 1.2 = 0.8$ $w_1 = 1.5$ $w_2 = 0.7$

2. Derive the equation for the derivative of categorical cross-entropy loss L with respect to the weighted sum Z, for a three-class classification problem. Assume single hidden layer and d input neurons. [5]

Rubrics and answer

$$\text{for class k, output } Z_k = \sum_{i=1}^{3} w_{jk} \cdot a_i$$

$$\text{softmax function } \hat{y}_k = \frac{e^{Z_k}}{\sum_{j=1}^{3} e^{Z_j}} \quad \text{0.5 mark}$$

$$\text{categorical cross-entropy loss } L = -\sum_{k=1}^{3} y_k \cdot \log(\hat{y}_k) \quad \text{1 mark}$$

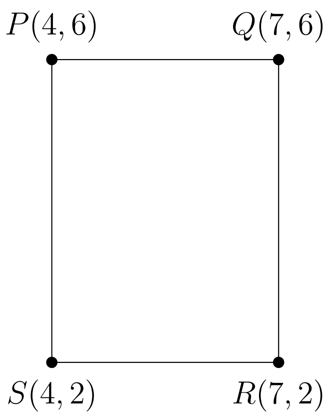
$$\frac{\partial L}{\partial Z_k} = \frac{\partial L}{\partial \hat{y}_k} \cdot \frac{\partial \hat{y}_k}{\partial Z_k} \quad \text{1 mark}$$

$$\frac{\partial L}{\partial \hat{y}_k} = -\frac{y_k}{\hat{y}_k} \quad \text{1 mark}$$

$$\frac{\partial \hat{y}_k}{\partial Z_k} = \hat{y}_k \cdot (1 - \hat{y}_k) \quad \text{1 mark}$$

$$\frac{\partial L}{\partial Z_k} = -y_k \cdot (1 - \hat{y}_k) \quad \text{0.5 mark}$$

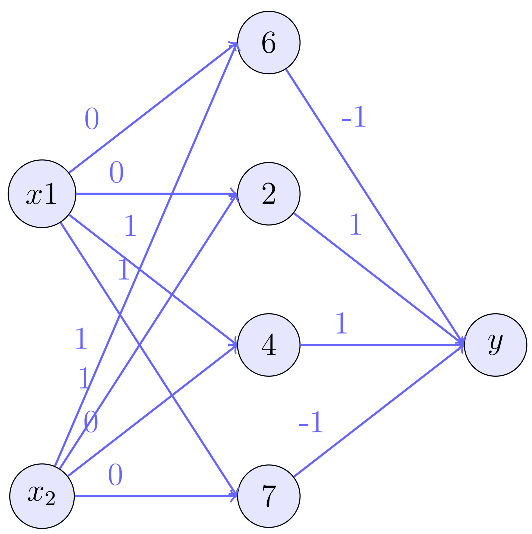
3. Construct an MLP for the given complex decision boundary. [6]



Rubrics and answer

- Each equation, 1 mark each.
- MLP, hidden layer with correct parameters - 1 mark
- MLP, last layer with correct parameters - 1 mark

$$PQ \rightarrow 0x_1 + 1x_2 = 6$$
$$SR \rightarrow 0x_1 + 1x_2 = 2$$
$$PS \rightarrow 1x_1 + 0x_2 = 4$$
$$QR \rightarrow 1x_1 + 0x_2 = 7$$



4. Find the minimum value of p for the equation $t = (2p+3)^2$ using SGD. Assume the initial value of p as 6 and learning rate as 0.1. Do 3 iterations. [4]

Rubrics and answer

- Each equation, 1 mark each.

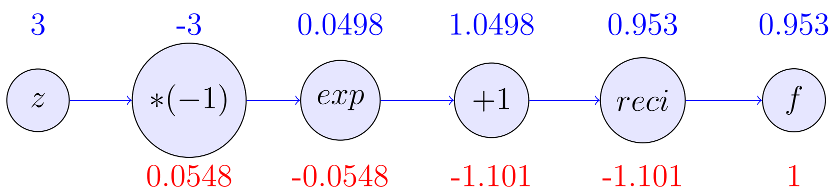
$$dt/dp = 2(2p+3)2 = 8p+12$$
$$p_1 \leftarrow 6 - 0.1(8*6+12) = 0$$
$$p_2 \leftarrow 0 - 0.1(8*0+12) = -1.2$$
$$p_3 \leftarrow (-1.2) - 0.1(8*(-1.2)+12) = -1.44$$

5. Draw the computational graph for the equation $f = 1/(1+e^{(-z)})$. Show the computations of derivatives of f wrt z in the graph. Using the graph, compute the value of f and the derivatives if $z = 3$. [5]

Rubrics and answer

- FP graph 1 mark
- BP graph 1 mark
- Output computation 1 mark
- Gradient computation 1 mark

$$p = -z = -3 \qquad \frac{\partial p}{\partial z} = -1 \qquad \frac{\partial f}{\partial p} = (-1)(-0.0548) = 0.0548$$
$$q = e^p = e^{-3} = 0.0498 \qquad \frac{\partial q}{\partial p} = e^p \qquad \frac{\partial f}{\partial q} = e^{-3}(-1.101) = -0.0548$$
$$r = p + 1 = 0.0498 + 1 = 1.0498 \qquad \frac{\partial r}{\partial p} = 1 \qquad \frac{\partial f}{\partial p} = 1(-1.101) = -1.101$$
$$f = \frac{1}{r} = \frac{1}{1.0498} = 0.953 \qquad \frac{\partial f}{\partial r} = \frac{-1}{r^2} \qquad \frac{\partial f}{\partial r} = \frac{-1}{0.953^2} = -1.101$$



6. Given an error surface, compute the value that minimizes the error with respect to (w_1, w_2, w_3). Compute the minimum possible value of error. [5]

$$E(w_1, w_2, w_3) = (w_1 - w_2)^3 - 2(w_1^2 - w_2) + w_1^2 + w_2^2$$

Rubrics and answer

$$\frac{\partial E}{\partial w_1} = 3(w_1 - w_2)^2 - 2w_1$$
$$\frac{\partial E}{\partial w_2} = -3(w_1 - w_2)^2 + 2w_2 + 2$$
$$\frac{\partial E}{\partial w_3} = 0$$

- Each equation 1 mark
- if the student writes or attempts that the equations have to be equated to zero and compute the value of parameters - 2 marks