Manarat International University (MIU)

Department of Computer Science and Engineering (Fall 2019)

Neural Network and Fuzzy Systems (CSE-433)

Solution of the problem Set of Home Work-1

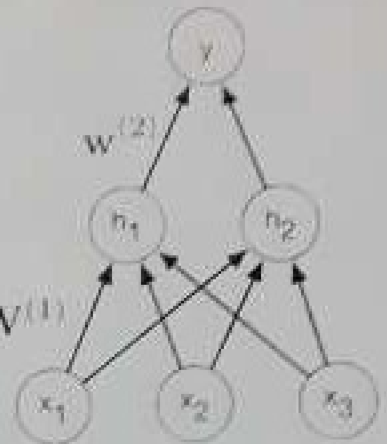
2. Riadul Islam Designs a multi-layer perceptron which receives three binary-valued (i.e. 0 or 1) inputs x_1, x_2, x_3, and outputs 1 if exactly two of the inputs are 1, and outputs 0 otherwise. He uses following activation function for all of the units.

$$W^{(1)} = \underbrace{\begin{pmatrix} \quad & \quad & \quad \end{pmatrix}}_{2\times3 \text{ matrix}}$$

$$b^{(1)} = (\qquad)$$

$$w^{(2)} = (\qquad)$$

$$b^{(2)} =$$



$$z = \begin{cases} 1 & \text{if } z \geq 0 \\ 0 & \text{if } z < 0 \end{cases}$$

Now, specify weights and biases which correctly implement his network.

Note: You do not need to explain your solution.

Hint: One of the hidden units should activate if 2 or more inputs are on, and the other should activate if all of the inputs are on.

<u>Sol^n</u>

$$W^{(1)} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \qquad \vec{b}^{(1)} = \begin{bmatrix} -1.5 \\ -2.5 \end{bmatrix}$$

$$\vec{w}^{(2)} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \qquad b^{(2)} = 0.5$$

<u>Explanation</u> (This part is not required to answer)

<u>Case-1</u> when all inputs are zero

$$h_1 = Z(x_1 \cdot w_{11}^{(1)} + x_2 \cdot w_{12}^{(1)} + x_3 \cdot w_{13}^{(1)} + b_1^{(1)}) = Z(0\times1 + 0\times1 + 0\times1 - 1.5) = Z(-1.5) = 0$$

$$h_2 = Z(x_1 \cdot w_{21}^{(1)} + x_2 \cdot w_{22}^{(1)} + x_3 \cdot w_{23}^{(1)} + b_2^{(1)}) = Z(0\times1 + 0\times1 + 0\times1 - 2.5) = Z(-2.5) = 0$$

$$y = Z(w_1^{(2)} \cdot h_1 + w_2^{(2)} h_2 + b^{(2)}) = Z(0+0-0.5) = Z(-0.5) = 0$$

Similarly,

<u>Case-2</u> when one input is one others zero

$$h_1 = Z(1+0+0-1.5) = Z(-0.5) = 0$$

$$h_2 = Z(1+0+0-2.5) = Z(-1.5) = 0$$

$$y = Z(0+0-0.5) = 0$$

Case-3 when two input is one

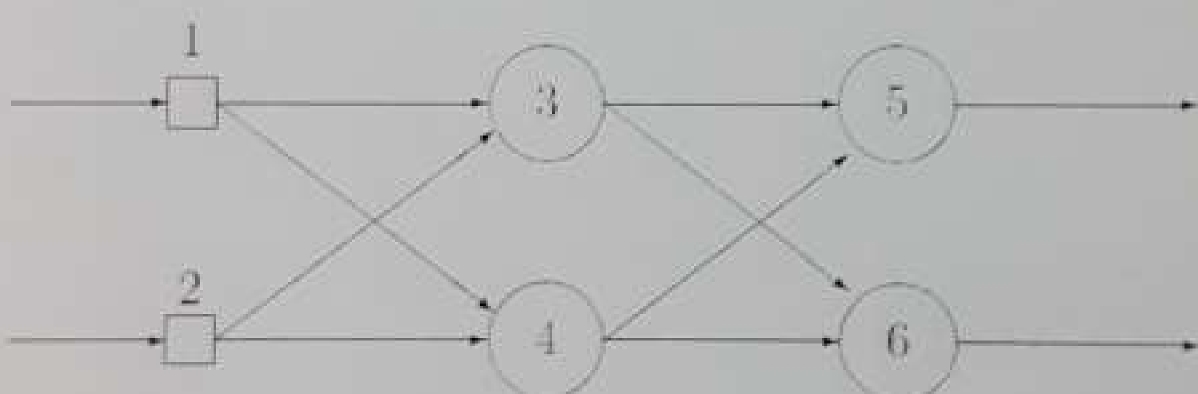
$$h_1 = z(1+1+0-1.5) = z(0.5) = 1 \quad \Big| \quad y = (1+0-0.5) = z(0.5)$$
$$h_2 = z(1+1+0-2.5) = z(-0.5) = 0 \quad \Big| \quad = \boxed{1}$$

Case-4 when all input is one

$$h_1 = z(1+1+1-1.5) = z(1.5) = 1 \quad \Big| \quad y = (1-1-0.5) = z(-0.5)$$
$$h_2 = z(1+1+1-2.5) = z(0.5) = 1 \quad \Big| \quad = 0$$

So, outputs are 1 when exactly two input is 1 (Case-3)
otherwise output is 0 (Case 1,2,4)

3. The following diagram represents a feed-forward neural network with one hidden layer



A weight on connection between nodes i and j is denoted by w_{ij}, such as w_{13} is the weight on the connection between nodes 1 and 3. The following table lists all the weights in the network:

$w_{13} = -2$	$w_{35} = 1$
$w_{23} = 3$	$w_{45} = -1$
$w_{14} = 4$	$w_{36} = -1$
$w_{24} = -1$	$w_{46} = 1$

$$\varphi(v) = \begin{cases} 1 & \text{if } v \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

$\varphi(v)$ is the activation function of all the nodes. Where v denotes the weighted sum of a node. Each of the input nodes (1 and 2) can only receive binary values (either 0 or 1). Calculate the output of the network (y_5 and y_6) for each of the input patterns:

Pattern:	P_1	P_2	P_3	P_4
Node 1:	0	1	0	1
Node 2:	0	0	1	1

$\underline{Sol^n}$ we know,

$$\begin{array}{l|l} v_3 = \omega_{13}x_1 + \omega_{23}x_2 & y_3 = \phi(v_3) \\ v_4 = \omega_{14}x_1 + \omega_{24}x_2 & y_4 = \phi(v_4) \\ v_5 = \omega_{35}y_3 + \omega_{45}y_4 & y_5 = \phi(v_5) \\ v_6 = \omega_{36}y_3 + \omega_{46}y_4 & y_6 = \phi(v_6) \end{array}$$

<u>For P_1</u>: Input pattern $(0,0)$

$$y_3 = \phi(v_3) = \phi(-2\times0 + 3\times0) = \phi(0) = 1$$

$$y_4 = \phi(v_4) = \phi(4\times0 - 1\times0) = \phi(0) = 1$$

$$y_5 = \phi(v_5) = \phi(1\times1 - 1\times1) = \phi(0) = 1$$

$$y_6 = \phi(v_6) = \phi(-1\times1 + 1\times1) = \phi(0) = 1$$

The output of the network is $(1,1)$

<u>For, P_2</u>: Input pattern$(1,0)$

$$y_3 = \phi(v_3) = \phi(-2\times1 + 3\times0) = \phi(-2) = 0$$

$$y_4 = \phi(v_4) = \phi(4\times1 - 1\times0) = \phi(4) = 1$$

$$y_5 = \phi(v_5) = \phi(1\times0 - 1\times1) = \phi(-1) = 0$$

$$y_6 = \phi(v_6) = \phi(-1\times0 + 1\times1) = \phi(1) = 1$$

The output of the network is $(0,1)$

<u>For P_3</u>: Input Pattern $(0,1)$

$$y_3 = \phi(v_3) = \phi(-2\times0 + 3\times1) = \phi(3) = 1$$

$$y_4 = \phi(v_4) = \phi(4\times0 - 1\times1) = \phi(-1) = 0$$

$$y_5 = \phi(v_5) = \phi(1\times1 - 1\times0) = \phi(1) = 1$$

$$y_6 = \phi(v_6) = \phi(-1\times1 + 0\times1) = \phi(-1) = 0$$

The output of the network is $(1,0)$

For P4: Input Pattern (1,1)

$$y_3 = \phi(v_3) = \phi(-2\times1+3\times1) = \phi(1) = 1$$

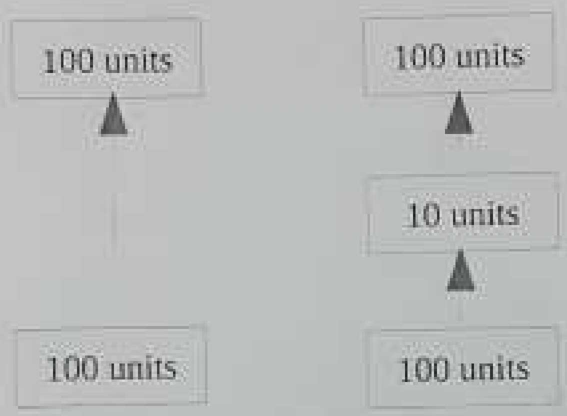
$$y_4 = \phi(v_4) = \phi(4\times1-1\times1) = \phi(3) = 1$$

$$y_5 = \phi(v_5) = \phi(1\times1-1\times1) = \phi(0) = 1$$

$$y_6 = \phi(v_6) = \phi(-1\times1+1\times1) = \phi(0) = 1$$

The output of the network is (1,1)

4. Consider the following two multi-layer perceptrons, where all of the layers use linear activation functions



Network A   Network B

→ bottleneck layer

Connections $(100\times100) = 10{,}000$

→ Connections $(100\times10)+(10\times100) = 2{,}000$

a. Mention the advantages of implementing Network A over Network B.

b. Mention the advantages of implementing Network B over Network A.

a) <u>Advantages</u> of Network-A over Network-B
- → A is more expressive
- → A has higher network capacity
- → A can model complex function

b) <u>Advantages</u> of <u>Implementing</u> Network-B over Network-A
- → B has fewer connections, so less prone to <u>overfitting</u>
- → For B, due to fewer connections backprop requires fewer operations.
- → B has bottleneck layer
- → Learn compact representation like autoencoder