Homework-2

Question 1

Consider a perceptron that computes its output O according to:

$$\begin{aligned} O &= g(h) \\ h &= \sum_{i=1}^{n} w_i x_i \end{aligned}$$

where

$$g(h) = \frac{1}{1+|h|}.$$

Show that the DELTA rule for updating the weights is given by:

$$\text{for} \quad i = 1, \ldots, n, \quad w_i \leftarrow w_i + \epsilon \delta x_i$$

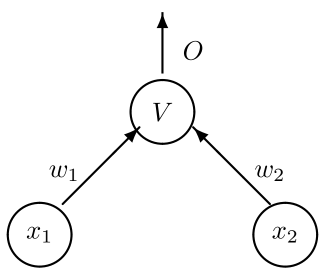
where δ is given by:

$$\delta = \begin{cases} -O^2(y-O) & \text{if } h \geq 0 \\ O^2(y-O) & \text{if } h < 0 \end{cases}$$

Hint: prove

$$g'(h) = \begin{cases} -O^2 & \text{if } h \geq 0 \\ O^2 & \text{if } h < 0 \end{cases}$$

Question 2



	x_1	x_2	y
e_1	1	0	1
e_2	1	1	1
e_3	2	0	0
e_4	2	1	0

You are given a perceptron implemented with a sigmoid, with $\beta = 1$. There are NO bias connections. The initial values of the weights are $w_1 = 0$, $w_2 = 1$.

Part 1

Give explicit expressions to the way the weights change if the network is given the example e_3. Use $\epsilon = 0.1$. You may use temporary variables in your answer, but make sure that they are all specified in terms of the given values. You may use the notation $S(.)$ instead of explicitly computing sigmoid values.

Part 2

a. If you could choose ϵ to be as small as you like, and run back propagation as many epochs as you like with the four examples e_1, e_2, e_3, e_4, do you expect the computed output of the perceptron to be within 0.001 of the desired output (the value of y) for all four examples?

Answer: YES / NO / IMPOSSIBLE-TO-TELL

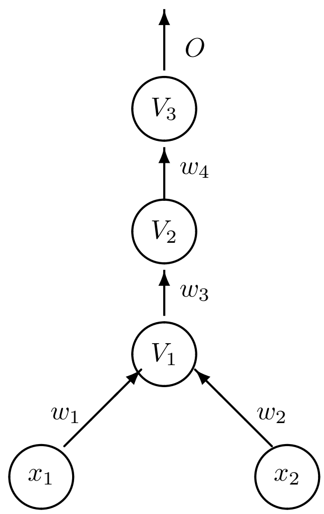
b. Will your answer to a. remain the same if bias connection is added (everything else stays as in a.)?

Answer:

- YES. My answer is exactly the same as in Part a.
- NO. My answer changes. With this new condition
  - my new answer to a. is YES.
  - my new answer to a. is NO.
  - my new answer to a. is IMPOSSIBLE-TO-TELL

Question 3

The following neural network has 3 layers with nodes that compute the sigmoid function. There are no bias connections.



A.1 Give explicit expressions to the values of all nodes in forward propagation when the network is given the input $x_1 = 1$, $x_2 = -1$, with the desired output $y = 0$. Your answer should be in terms of the old weights w_1, w_2, w_3, w_4. You may use the notation $S(.)$ instead of explicitly computing sigmoid values.

Answer

$V_1 =$

$V_2 =$

$V_3 =$

A.2 Give explicit expressions to the output produces by the network when given this example. Your answer should be in terms of the old weights w_1, w_2, w_3, w_4, and the values V_1, V_2, V_3 computed in A.1. If needed, you may use the notation $S(.)$ instead of explicitly computing sigmoid values.

Answer

$O =$

A.3 Give explicit expressions to how the weights change by back propagation when the network is given the same example as above. Use $\epsilon = 0.1$.

Your answer should be in terms of the old weights w_1, w_2, w_3, w_4, the values V_1, V_2, V_3 computed in A.1., and the value O, computed in A.2. If needed, you may use the notation $S(.)$ instead of explicitly computing sigmoid values. You may use temporary variables in your answer, but make sure that they are defined in terms of the above variables.

Answer I am using the following temporary variables in my answer:

The expressions for the new weights are:

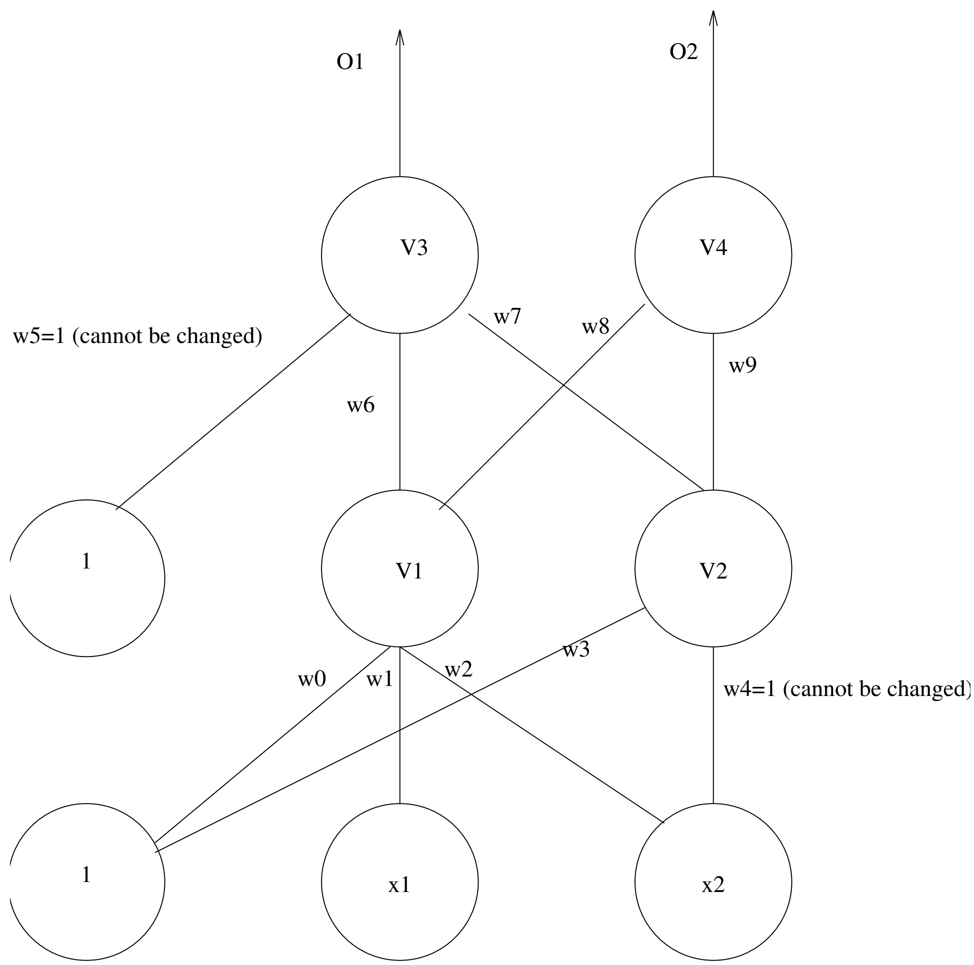
new w_1 =

new w_2 =

new w_3 =

new w_4 =

Question 4



The above neural network has two layers (one hidden layer), two inputs, and two outputs. There is NO bias connection for V_4, and there is NO connection between V_2 and x_1. The weights w_4 and w_5 have the value 1 and cannot be changed. All nodes compute the sigmoid function with $\beta = 1$.

A.1 Give explicit expressions to the values of all nodes in forward propagation when the network is given the input $x_1 = 2$, $x_2 = 3$, with the desired output $y_1 = 0$, $y_2 = 2$ Your answer should be in terms of the old weights w_0, w_1, w_2, w_3, w_6, w_7, w_8, w_9. Do not use any other temporary variables (such as h_i). If needed, you may use the notation $S(.)$ instead of explicitly computing sigmoid values.

Answer

$V_1 =$

$V_2 =$

$V_3 =$

$V_4 =$

A.2 Give explicit expressions to the output produces by the network when given this example. Your answer should be in terms of the old weights w_0, w_1, w_2, w_3, w_6, w_7, w_8, w_9, and the values V_1, V_2, V_3, V_4 computed in A.1. Do not use any other temporary variables (such as h_i). If needed, you may use the notation $S(.)$ instead of explicitly computing sigmoid values.

Answer

$O_1 =$

$O_2 =$

A.3 Give explicit expressions to how the weights change by back propagation when the network is given the same example as above. Use $\epsilon = 0.1$.

Your answer should be in terms of the old weights w_0, w_1, w_2, w_3, w_6, w_7, w_8, w_9. and the values V_1, V_2, V_3, V_4 computed in A.1. and the values O_1, O_2 computed in A.2. If needed, you may use the notation $S(.)$ instead of explicitly computing sigmoid values. You may use temporary variables in your answer, but make sure that they are defined in terms of the above variables.

Answer I am using the following temporary variables in my answer:

The expressions for the new weights are:

new w_0 =

new w_1 =

new w_2 =

new w_3 =

new w_6 =

new w_7 =

new w_8 =

new w_9 =