Tania Soutonglang

CS 484-01

November 19, 2022

Homework 7 Part 1

1.1)

a.

$$xw_i$$
$$z_1 = xw_1 = 4(1) = 4 \rightarrow 4$$
$$z_2 = xw_2 = 4(1) = 4 \rightarrow 4$$
$$z_3 = xw_3 = 4(-1) = -4 \rightarrow 0$$

$$z = (z_1w_4) + (z_2w_5) + (z_3w_6)$$
$$\big(4(0.5)\big) + \big(4(1)\big) + \big(0(2)\big)$$
$$2 + 4 + 0$$
$$6$$

$$\boldsymbol{a} = \sigma(z)$$
$$\frac{1}{1+e^{-x}}$$
$$\frac{1}{1+e^{-6}}$$
$$0.99753 \rightarrow \mathbf{0.998}$$

b.

$$\boldsymbol{loss} = (y - \hat{y})^2$$
$$(y-a)^2$$
$$(0-0.99753)^2$$
$$0.99507 \rightarrow \mathbf{0.996}$$

c.

$$\frac{dL}{da} = -2(y-a)$$
$$-2(0-0.99753)$$
$$-2(-0.99753)$$
$$1.99506$$

$$\frac{da}{dz} = \frac{d}{dz}\big(e(2)\big)$$
$$a(1-a)$$
$$(0.99753)(1-0.99753)$$

$$(0.99753)(0.00247)$$
$$0.0024639$$

$$\frac{dL}{dz} = \frac{dL}{da} * \frac{da}{dz}$$
$$1.99506 * 0.0024639$$
$$0.0049156$$

$$\frac{dC}{dw_4} = \frac{dL}{dz} * \frac{dz}{dw_4}$$
$$(0.0049156)(\mathrm{z}_{\mathrm{i}}')$$
$$(0.0049156)(4)$$
$$0.019662$$

$$\frac{dC}{dw_5} = \frac{dL}{dz} * \frac{dz}{dw_5}$$
$$(0.0049156)(\mathrm{z}_2')$$
$$(0.0049156)(4)$$
$$0.019662$$

$$\frac{dC}{dw_6} = \frac{dL}{dz} * \frac{dz}{dw_6}$$
$$(0.0049156)(\mathrm{z}_3')$$
$$(0.0049156)(0)$$
$$0$$

$$z = z_1' w_4 + z_2' w_5 + z_1' w_6$$
$$\frac{dL}{dz_1} = w_4 \frac{dz}{dz_2'} = w_5 \frac{dz}{dz_3'} = w_6$$
$$\frac{dL}{dz_1} = \frac{dL}{dz} * \frac{dz}{dz_1} = 0.0049156 * 0.5 = 0.0024578$$
$$\frac{dL}{dz_2} = \frac{dL}{dz} * \frac{dz}{dz_2} = 0.0049156 * 1 = 0.0049156$$
$$\frac{dL}{dz_3} = \frac{dL}{dz} * \frac{dz}{dz_3} = 0.0049156 * 2 = 0.0098312$$
$$\frac{dz_1'}{dz_1} = 1$$
$$\frac{dz_2'}{dz_2} = 1$$
$$\frac{dz_3'}{dz_3} = 1$$

$$\frac{dl}{dz_1} = \frac{dL}{dx_1'} * \frac{2z_1'}{2z_1} = 0.0024578 * 1 = 0.0024578$$

$$same\ with \frac{dl}{dz_2} = 0.0049156\ and\ \frac{dl}{dz_3} = 0.0098312$$

$$\frac{dL}{dw_1} * \frac{dL}{dz_1} * \frac{dz_1}{dw_1} = 0.0024578 * 4 = 0.0098312$$

$$\frac{dL}{dw_2} * \frac{dL}{dz_2} * \frac{dz_2}{dw_2} = 0.0049156 * 4 = 0.019662$$

$$\frac{dL}{dw_3} * \frac{dL}{dz_3} * \frac{dz_3}{dw_3} = 0.0098312 * 4 = 0.039325$$

$$\alpha = 0.1$$

$$w_i = w_i - \alpha \frac{dL}{dw_i}$$

$$w_1 = 1 - 0.1(0.0098312) = 0.99902$$

$$w_2 = 1 - 0.1(0.019662) = 0.99803$$

$$w_3 = -1 - 0.1(0.039325) = -1.0039325$$

$$w_4 = 0.5 - 0.1(0.0098312) = 0.49902$$

$$w_5 = 1 - 0.1(0.019662) = 0.99803$$

$$w_6 = 2 - 0.1(0.039325) = 1.99607$$

$$z_1 = xw_1 = 4(0.99902) = 3.99608$$

$$z_2 = xw_2 = 4(0.99803) = 3.99212$$

$$z_3 = xw_3 = 4(-1.0039325) = -4.01573$$

$$z_1' = 3.99608$$

$$z_2' = 3.99212$$

$$z_3' = 0$$

$$3.99608(0.5) + 3.99212(1) + 0(2)$$

$$1.99804 + 3.99212 + 0$$

$$5.99016 \rightarrow \mathbf{5.990}$$

d.

$$a = -(z) = \frac{1}{1+e^{-z}} = 0.997503$$

$$L(y, a) = (y - a)^2$$

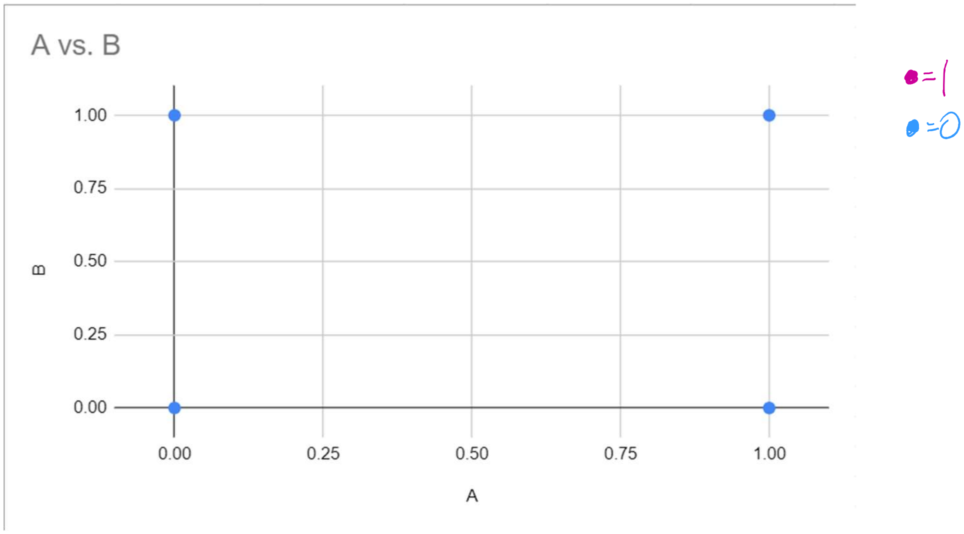
$$= (0 - 0.997503)^2$$

$$0.99501 \rightarrow \mathbf{0.995}$$

e. The accuracy of the first output (b) was 0.996 and the accuracy of the second output (d) was 0.995. From this, the first output is closer to the target.

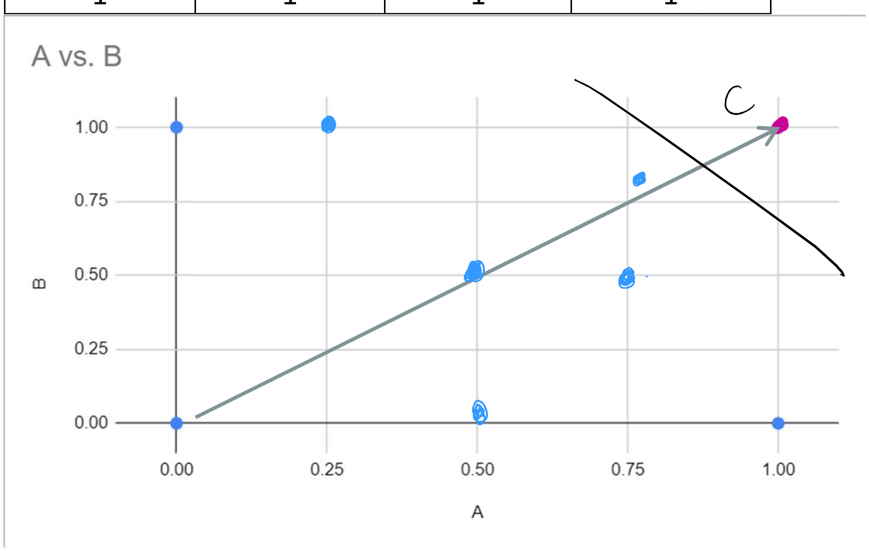
1.2) Tan Chapter 4

4.14)



a) *A* and *B* and *C*

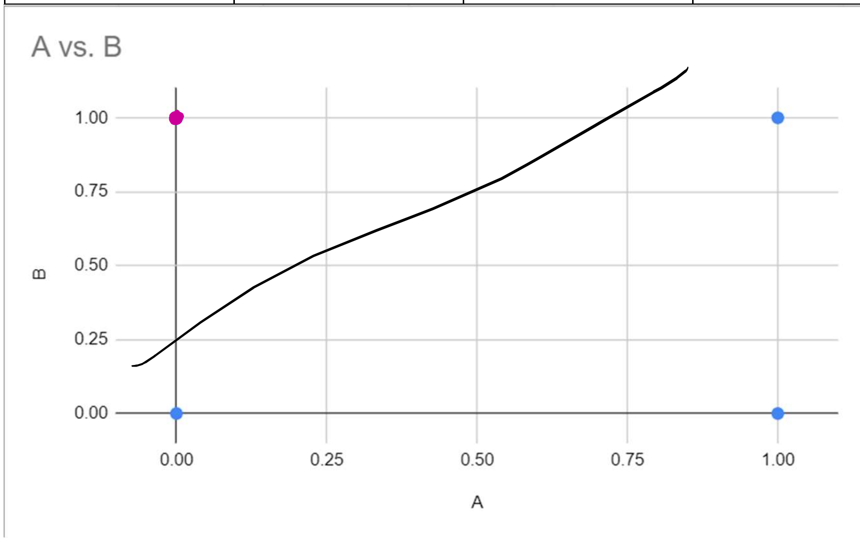
A	B	C	A and B and C
0	0	0	0
1	0	0	0
0	1	0	0
1	1	0	0
0	0	1	0
1	0	1	0
0	1	1	0
1	1	1	1



Linearly separable

b) not *A* and *B*

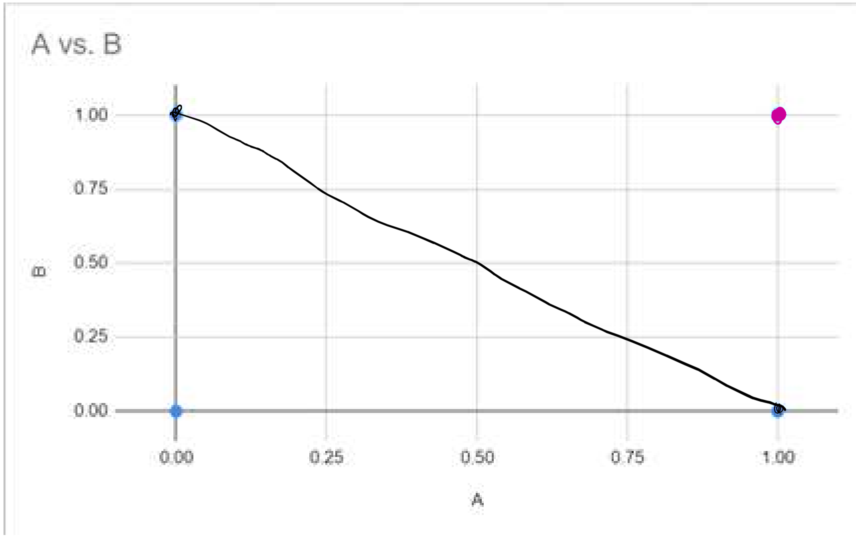
A	B	Ā	B	Ā and B
0	0	1	0	0
1	0	0	0	0
0	1	1	1	1
1	1	0	1	0



Linearly separable

c) (*A* or *B*) and (*A* or *C*)

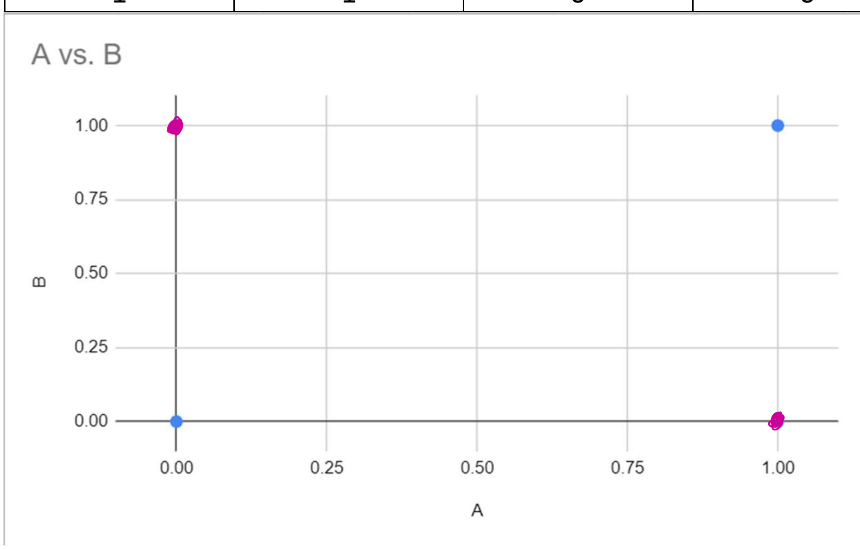
A	B	C	A or B	A or C	(A or B) and (A or C)
0	0	0	0	0	0
1	0	0	1	1	1
0	1	0	1		1
1	1	0	1		1
0	0	1		1	
1	0	1		1	
0	1	1			
1	1	1			



Linearly separable

d) (*A* xor *B*) and (*A* or *B*)

A	B	A xor B	A or B	(A xor B) and (A or B)
0	0	0	0	0
1	0	1	1	1
0	1	1	1	1
1	1	0	0	0



Linearly not separable

4.15)

a) AND uses 2 inputs, *x1* and *x2*, and gives 1 output, *y*. This makes the perceptron function to become

$$y = \Theta(\mathrm{x}_1 w_1 + x_2 w_2 + b)$$

The value of *w1* will be 1, *w2* will be 1, and *b* will be -1.5, causing the function to turn into

$$y = \Theta[x_1(1) + x_2(1) - 1.5] = \boldsymbol{\Theta(x_1 + x_2 - 1.5)}$$

The OR function is similar to the AND function where it uses the same number of inputs, *x1* and *x2*, to get 1 output, *y*, and therefore would also use the function $y = \Theta(\mathrm{x}_1 w_1 + x_2 w_2 + b)$. *x1* will be 1, *x2* will be 1, and *b* will be 0.5, causing the function to become

$$y = \Theta[\mathrm{x}_1(1) + x_2(1) + 0.5] = \boldsymbol{\Theta(x_1 + x_2 + 0.5)}$$

b) The resulting network of an activation function represented linearly is a linear combination of the input elements, making the network as expressive as a perceptron. Also, when the activation function is linear, nesting *n* number of hidden layers in the function wouldn't have an effect on the results.

1.3) n = 8

hidden layers = 3

h1 = 16 neurons

h2 = 8 neurons

h3 = 4 neurons

$$i * h1 + \sum_{k=1}^{n-1} (h_k * h_{k+1}) + h_n * o + \sum_{k=1}^{n} h_k + o$$

$$16 * 8 = 128$$
$$8 * 4 = 32$$
$$4 * 4 = 16$$
$$128 + 32 + 16 = \boldsymbol{176\ parameters}$$