

# Assignment 2 Machine Learning COS4852

Adriaan Louw (53031377)

July 9, 2018

## Contents

<b>1</b>	<b>Question 1</b>	<b>1</b>
1.1	Question 1(a)	1
1.2	Question 1(b)	3
1.3	Question 1(c)	5
1.4	Question 1(d)	5
<b>2</b>	<b>Question 2</b>	<b>5</b>
2.1	Question 2(a)	5
2.2	Question 2(b)	7
2.3	Question 2(c)	10
<b>3</b>	<b>Question 3</b>	<b>12</b>
3.1	Question 3(a)	12
3.2	Question 3(b)	12
3.3	Question 3(c)	12
3.4	Question 3(d)	12
3.5	Question 3(e)	12

## 1 Question 1

### 1.1 Question 1(a)

Firstly we calculate the line

$$x_2 = mx_1 + c \quad (1)$$

for the intersect points (2,0) and (0,6).

Calculating slope m,

$$\begin{aligned} m &= \frac{6 - 0}{0 - 2} \\ &= -3 \end{aligned} \quad (2)$$

$x_2$  intercept  $c$  is 6.

This makes equation 1

$$x_2 = -3x_1 + 6 \quad (3)$$

Nils J Nilsson (1998) gives the equation for the hyperplane as

$$\sum_{i=1}^n x_i \omega_i \geq \theta \quad (4)$$

which in this case gives the equation for the hyperplane to be

$$\omega_1 x_1 + \omega_2 x_2 + \omega_3 = 0 \quad (5)$$

We need to get equation 5 in the form of equation 1

$$\begin{aligned} \omega_1 x_1 + \omega_2 x_2 + \omega_3 &= 0 \\ \omega_2 x_2 &= -\omega_1 x_1 - \omega_3 \\ x_2 &= -\frac{\omega_1 x_1}{\omega_2} - \frac{\omega_3}{\omega_2} \end{aligned} \quad (6)$$

Comparing coefficients m and c from equation 3 to 6 we get

$$\begin{aligned} -\frac{\omega_1}{\omega_2} &= -3 \\ \omega_1 &= 3\omega_2 \end{aligned} \quad (7)$$

and

$$\begin{aligned} -\frac{\omega_3}{\omega_2} &= 6 \\ \omega_3 &= -6\omega_2 \end{aligned} \quad (8)$$

If we choose  $\omega_3 = -2$  then  $\omega_1 = 1$  and  $\omega_2 = \frac{1}{3}$ . This makes the hyperplane equation from equation 5

$$x_1 + \frac{x_2}{3} - 2 = 0 \quad (9)$$

Now we need to test this hyperplane. For positive instance (2,6)

$$\begin{aligned} x_1 + \frac{x_2}{3} - 2 &= \\ 2 + \frac{6}{3} - 2 &= \\ 2 \end{aligned} \quad (10)$$

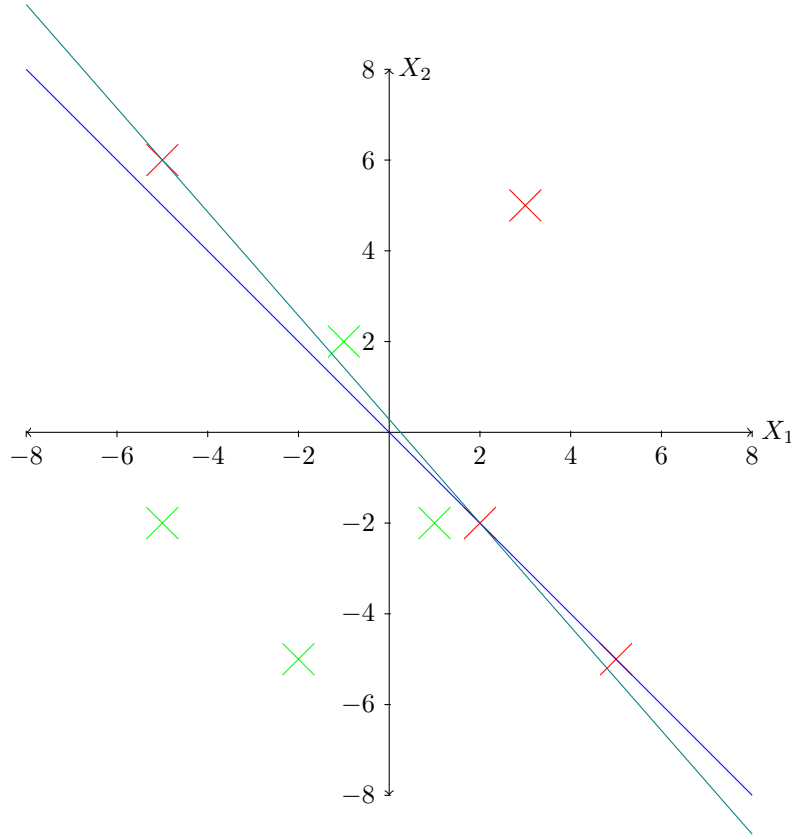
Which is as expected.

And the negative instance (-1,2)

$$\begin{aligned} x_1 + \frac{x_2}{3} - 2 &= \\ -1 + \frac{2}{3} - 2 &= \\ -\frac{7}{3} \end{aligned} \quad (11)$$

This is also as expected. The perceptron now classifies the the data correctly

## 1.2 Question 1(b)



From the above image we can see that any that it is not possible to create a hyperplane that correctly classifies all negative instances and positive instances. The blue line is the line  $x_2 = -x_1$  and the teal line is the line  $x_2 = -\frac{7}{8}x_1 + \frac{2}{7}$ . The any minimum plane that correctly classifies all the negative instances will classify the positive instance  $(-1,2)$  incorrectly as negative.

We can create a hyperplane from regression from all the points close to where the hyperplane should be. Using negative points  $(-5,6), (2,-2), (5,-5)$  and positive points  $(-1,2), (1,-2)$

For the equation of the regressed line  $x_2 = mx_1 + c$

$$m = r \frac{S(x_2)}{S(x_1)}$$

$$m = \frac{\sum((x_1 - \bar{x}_1)(x_2 - \bar{x}_2))}{\sqrt{\sum(x_1 - \bar{x}_1)^2 \sum(x_2 - \bar{x}_2)^2}} \frac{\sqrt{\frac{\sum(x_2 - \bar{x}_2)^2}{n-1}}}{\sqrt{\frac{\sum(x_1 - \bar{x}_1)^2}{n-1}}} \quad (12)$$

where  $r$  is Pearsons Correlation Coefficient and  $S$  is standard deviation of axis  $x_2$  or  $x_1$ .

Here follows the calculation

$x_1$	$x_2$	$x_1 - \bar{x}_1$	$x_2 - \bar{x}_2$	$(x_1 - \bar{x}_1)(x_2 - \bar{x}_2)$	$(x_1 - \bar{x}_1)^2$	$(x_2 - \bar{x}_2)^2$
-5	6	-5.4	6.2	-33.48	29.16	38.44
2	-2	1.6	-1.8	-2.88	2.56	3.24
5	-5	4.6	-4.8	-22.08	21.16	23.04
-1	2	-1.4	2.2	-3.08	1.96	4.84
1	-2	0.6	-1.8	-1.08	0.36	3.24

From the above table we have  $\bar{x}_1 = 0.4, \bar{x}_2 = -0.2, \sum((x_1 - \bar{x}_1)(x_2 - \bar{x}_2)) = -62.6, (x_1 - \bar{x}_1)^2 = 55.2, \text{ and } (x_2 - \bar{x}_2)^2 = 72.8$

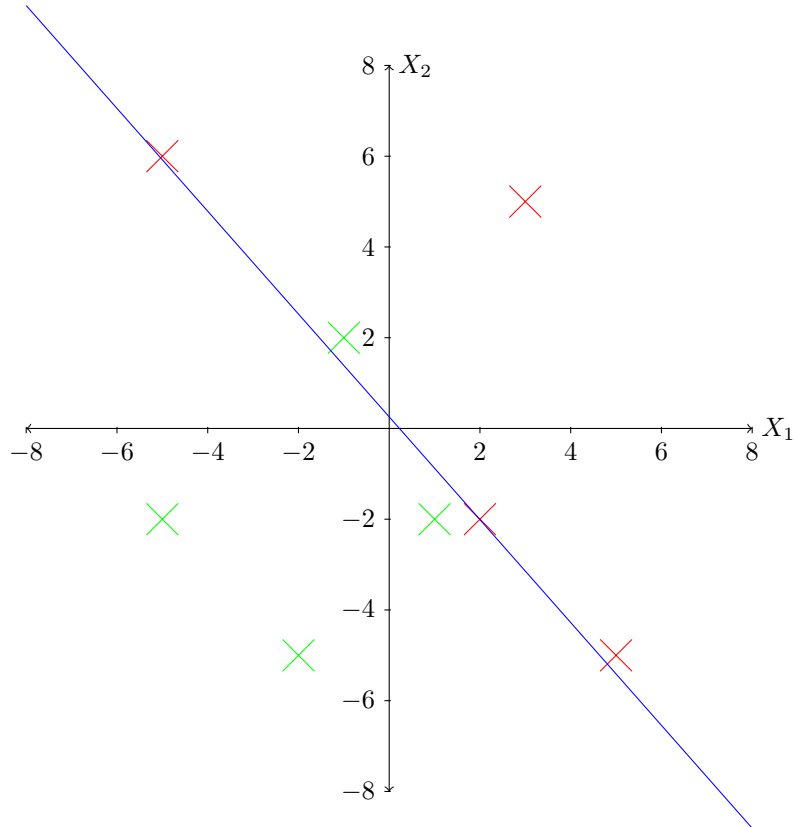
Passing these into equation 12 we get  $m = -1.13$

$$\begin{aligned}
 c &= \bar{x}_2 - m\bar{x}_1 \\
 &= -0.2 - (-1.13)(0.4) \\
 &= 0.25
 \end{aligned} \tag{13}$$

Which gives us equation

$$x_2 = -1.13x_1 + 0.25 \tag{14}$$

Visually it would be



This will minimise the error even though it would incorrectly classify (-1,2)

### 1.3 Question 1(c)

### 1.4 Question 1(d)

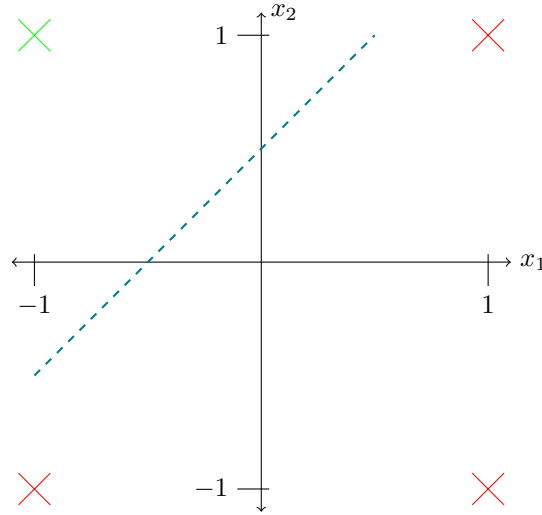
## 2 Question 2

### 2.1 Question 2(a)

The truth table for this function is

$x_1$	$x_2$	$\neg x_1$	$\neg x_1 \vee x_2$
-1	-1	1	-1
-1	1	1	1
1	-1	-1	-1
1	1	-1	-1

Visually this is:



The dashed line represents the function

$$x_2 = x_1 + 0.5 \quad (15)$$

This hyperplane will classify the boolean function correctly because it linearly separates all the positive instances from the negative ones.

The equation for the hyperplane with weights are given by equation 5 and we know from equation 6 what the relationship is from the weight equation to the line equation. Thus for case

$$\begin{aligned} m &= -\frac{\omega_1}{\omega_2} \\ 1 &= -\frac{\omega_1}{\omega_2} \\ \omega_2 &= -\omega_1 \end{aligned} \quad (16)$$

and

$$\begin{aligned}
c &= -\frac{\omega_0}{\omega_2} \\
0.5 &= -\frac{\omega_0}{\omega_2} \\
\omega_0 &= -\frac{\omega_2}{2}
\end{aligned} \tag{17}$$

Now we have the relationship between the weights. If we try  $\omega_0 = 1$  we get  $\omega_1 = 2$  and  $\omega_2 = -2$

Which makes our weight equation

$$2x_1 - 2x_2 + 1 \tag{18}$$

Testing for polarity with point (-1,-1) we get

$$2(-1) - 2(-1) + 1 = 1 \tag{19}$$

which indicates a positive result but we expect a negative result. Thus our polarity is wrong

We try  $\omega_0 = -1$  which gives  $\omega_1 = -2$  and  $\omega_2 = 2$

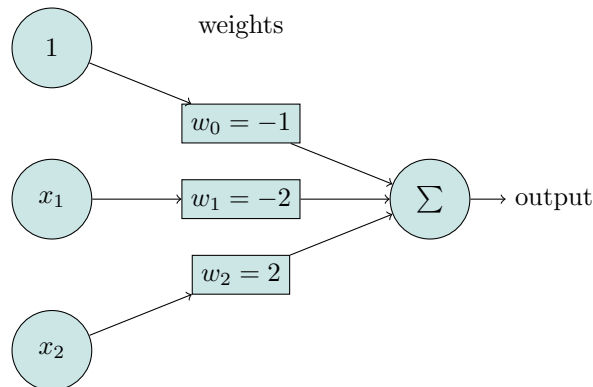
Which makes the weight equation

$$-2x_1 + 2x_2 - 1 \tag{20}$$

Trying all our data we can sum it up in the following table

$x_1$	$x_2$	Result
-1	-1	-1
-1	1	4
1	-1	-5
1	1	-1

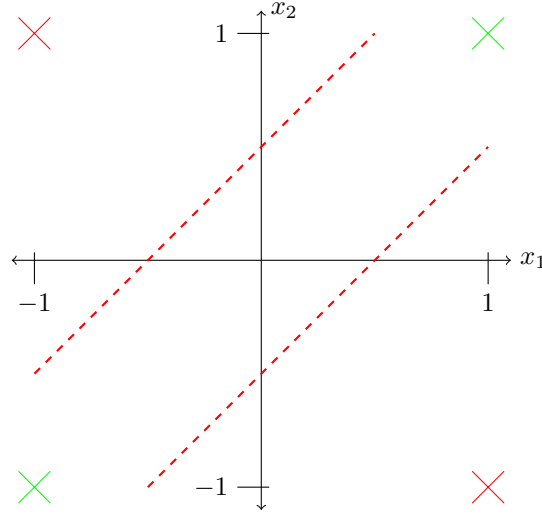
This shows that weights  $\omega_0 = -1$ ;  $\omega_1 = -2$  and  $\omega_2 = 2$  correctly classify the data inputs



## 2.2 Question 2(b)

$x_1$	$x_2$	$x_1 \oplus x_2$	$f_2 = \neg(x_1 \oplus x_2)$
-1	-1	-1	1
-1	1	1	-1
1	-1	1	-1
1	1	-1	1

Visually this is:



We can see from the previous diagram that  $f_2$  cannot be linearly separated by 1 hyperplane. In other words a single perceptron cannot classify this function. We need to try to decompose  $f_2$  into multiple linearly separable functions that can be modelled by multiple perceptrons.

$$\begin{aligned}
 f_2 &= \neg(x_1 \oplus x_2) \\
 &= \neg((x_1 \wedge \neg x_2) \vee (\neg x_1 \wedge x_2)) \\
 &= \neg(x_1 \wedge \neg x_2) \wedge \neg(\neg x_1 \wedge x_2) \\
 &= (\neg x_1 \vee x_2) \wedge (x_1 \vee \neg x_2)
 \end{aligned} \tag{21}$$

We can thus create new function which is equivalent to  $f_2$

$$g(h_1, h_2) = h_1 \wedge h_2 \tag{22}$$

where

$$h_1 = \neg x_1 \vee x_2 \tag{23}$$

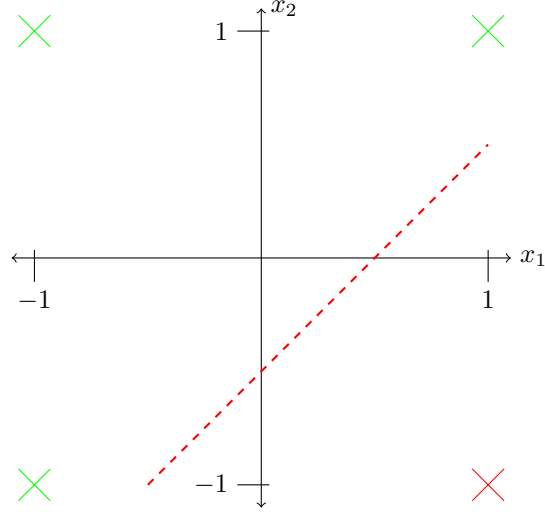
and

$$h_2 = x_1 \vee \neg x_2 \tag{24}$$

The truth table will thus become

$x_1$	$x_2$	$x_1 \oplus x_2$	$f_2 = \neg(x_1 \oplus x_2)$	$h_1$	$h_2$	$g(h_1, h_2)$
-1	-1	-1	1	1	1	1
-1	1	1	-1	1	-1	-1
1	-1	1	-1	-1	1	-1
1	1	-1	1	1	1	1

Drawing  $h_1$  gives



Using equation 5 and knowing this hyperplane goes through points  $(-0.5, 1)$  and  $(1, 0.5)$ . We have

$$\begin{aligned}\omega_1(-0.5) + \omega_2(-1) + \omega_0 &= 0 \\ \frac{-\omega_1}{2} - \omega_2 + \omega_0 &= 0\end{aligned}\tag{25}$$

for the point  $(-0.5, -1)$  and

$$\begin{aligned}\omega_1(1) + \omega_2(0.5) + \omega_0 &= 0 \\ \omega_1 + \frac{\omega_2}{2} + \omega_0 &= 0\end{aligned}\tag{26}$$

for point  $(1, 0.5)$

Equating equations 25 and 26 we get

$$\begin{aligned}\frac{-\omega_1}{2} - \omega_2 + \omega_0 &= \omega_1 + \frac{\omega_2}{2} + \omega_0 \\ \frac{3\omega_1}{2} &= \frac{3\omega_2}{2} \\ \omega_2 &= -\omega_1\end{aligned}\tag{27}$$

Passing equation 27 into 26 we get

$$\begin{aligned}\omega_1 + \frac{-\omega_1}{2} + \omega_0 &= 0 \\ \omega_0 &= -\frac{\omega_1}{2}\end{aligned}\tag{28}$$



If we take  $\omega_0 = 1$  then  $\omega_1 = -2$  and  $\omega_2 = 2$ .

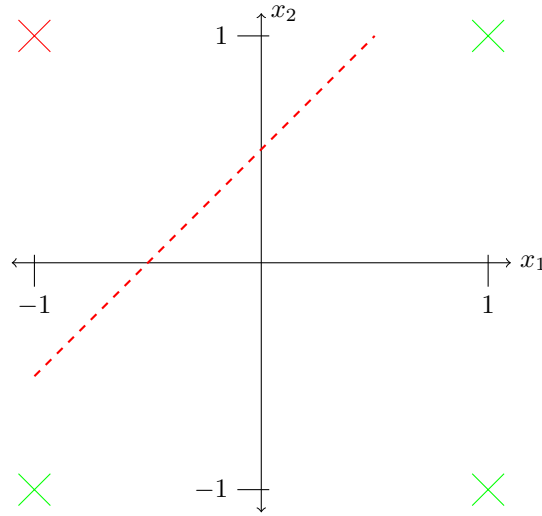
Testing for polarity we use the point  $(-1,-1)$  where we expect a positive answer

For point  $(-1,-1)$

$$\begin{aligned} -2x_1 + 2x_2 + 1 &= \\ -1(-1) + 2(-1) + 1 &= 1 \end{aligned} \quad (29)$$

Which is greater than 0 as expected. Thus the weights for  $h_1$  is  $\omega_0 = 1; \omega_1 = -2; \omega_2 = 2$

Now drawing  $h_2$



This is a similar perceptron from as in Question 2a. It uses the same hyper-plane but the polarity is reversed. Thus from question 2a  $\omega_2 = -\omega_1; \omega_0 = -\frac{\omega_2}{2}$ .

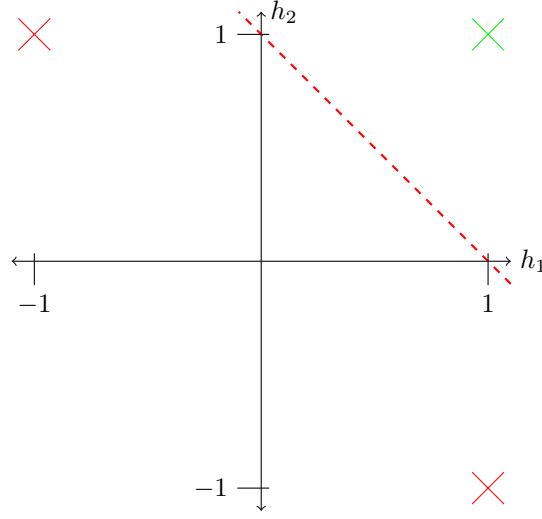
Testing polarity with  $\omega_0 = 1$  then  $\omega_1 = 2; \omega_2 = -2$ . This gives with positive instance  $(-1,-1)$

$$\begin{aligned} 2x_1 - 2x_2 + 1 &= \\ 2(-1) - 2(-1) + 1 &= 1 \end{aligned} \quad (30)$$

Which indicates the preceptron calssifies the point  $(-1,-1)$  correctly as a positive instance

This means  $\omega_0 = 1; \omega_1 = 2; \omega_2 = -2$  for  $h_2$

Now we need to create a perceptron for function g. The domain of g as a function of  $h_1$  and  $h_2$  can be shown as follows



From the above diagram we can see that a hyperplane through the points (0,1) and (1,0) will classify this perceptron

From applying equation 5 to point  $(h_1, h_2) = (1, 0)$

$$\begin{aligned}\omega_1 x_1 + \omega_2 x_2 + \omega_0 &= 0 \\ \omega_1(1) + \omega_2(0) + \omega_0 &= 0 \\ \omega_1 &= -\omega_0\end{aligned}\tag{31}$$

and for point  $(h_1, h_2) = (0, 1)$

$$\begin{aligned}\omega_1 x_1 + \omega_2 x_2 + \omega_0 &= 0 \\ \omega_1(0) + \omega_2(1) + \omega_0 &= 0 \\ \omega_2 &= -\omega_0\end{aligned}\tag{32}$$

Now equation equations 31 and 32 we get  $\omega_1 = \omega_2$

If we let  $\omega_0 = -1$  then from substitution into equations 31 and 32 we get  $\omega_1 = \omega_2 = 1$ .

Testing for polarity we have positive instance (1,1)

$$\begin{aligned}\omega_1 x_1 + \omega_2 x_2 + \omega_0 &= \\ x_1 + x_2 - 1 &= \\ (1) + (1) - 1 &= 1\end{aligned}\tag{33}$$

Which is positive as expected

Figure 1 shows the whole network.

### 2.3 Question 2(c)

$x_1$	$x_2$	$\neg x_1$	$\neg x_2$	$x_1 \vee x_2$	$\neg x_1 \vee \neg x_2$	$(x_1 \vee x_2) \wedge (\neg x_1 \vee \neg x_2)$
-1	-1	1	1	-1	1	-1
-1	1	1	-1	1	1	1
1	-1	-1	1	1	1	1
1	1	-1	-1	1	-1	-1

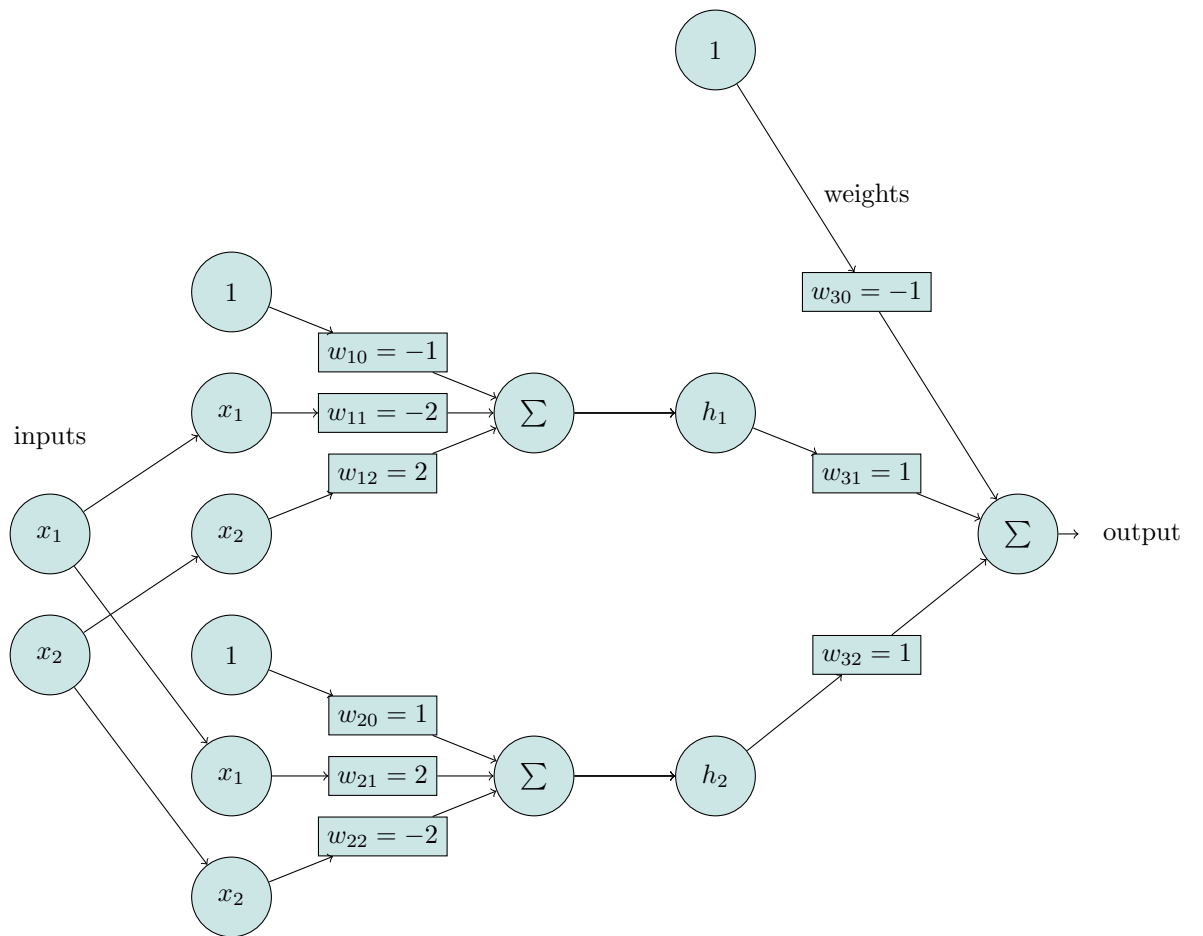
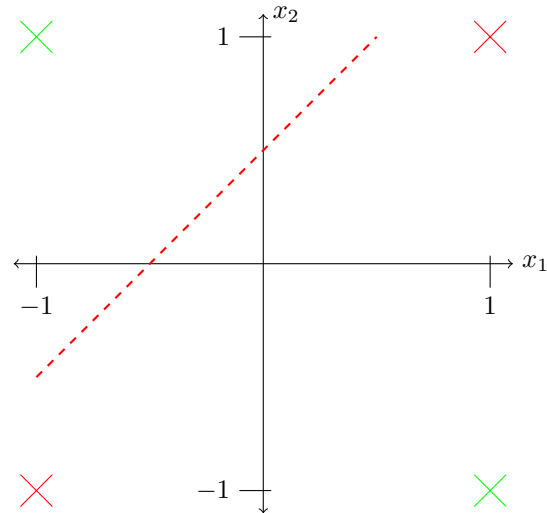


Figure 1: Network for question 2b

Visually this is:



### 3 Question 3

3.1 Question 3(a)

3.2 Question 3(b)

3.3 Question 3(c)

3.4 Question 3(d)

3.5 Question 3(e)

### References

Nils J Nilsson. (1998). *Introduction to Machine Learning*. Retrieved from <http://robotics.stanford.edu/people/nilsson/MLB00K.pdf>