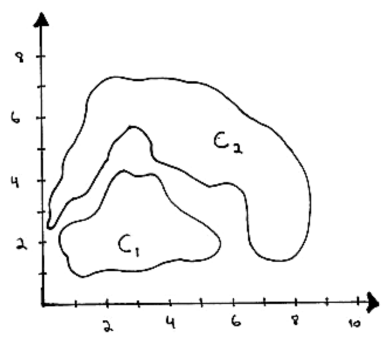
Kieran Cosgrove Student #: 20226841

CISC/CMPE 452/COGS 400
Assignment 1
Theoretical Part

1. **Percptron:** - Consider the classification problem defined below.
$\mathbf{p}_1 = ([-1, 1]^t, 1)$, $\mathbf{p}_2 = ([-1, -1]^t, 1)$, $\mathbf{p}_3 = ([0, 0]^t, 0)$, and $\mathbf{p}_4 = ([1, 0]^t, 0)$.
   (a) Design a single-neuron perceptron to solve this problem. Design the network graphically, by choosing weight vectors that are orthogonal to the decision boundaries.
   (b) Test your solution with all four input vectors.
   (c) Classify the following input vectors with your solution $\mathbf{p}_5 = ([-2, 0]^t)$, $\mathbf{p}_6 = ([1, 1]^t)$, $\mathbf{p}_7 = ([0, 1]^t)$, and $\mathbf{p}_8 = ([-1, -2]^t)$.
   (d) Which of the vectors in part (3) will always be classified the same way, regardless of the solution? Which may vary depending on the solution? Why?

2. **Multilayer Percptron** - You are presented with the following input set:
$\mathbf{x}_1 = [3, 1, 1]^t$, $\mathbf{x}_2 = [4, 0, 1]^t$, $\mathbf{x}_3 = [4, -1, 1]^t$, $\mathbf{x}_4 = [5, 2, 1]^t$,$\mathbf{x}_5 = [5, 3, 1]^t$, $\mathbf{x}_6 = [3, 3, 1]^t$, $\mathbf{x}_7 = [2, 0, 1]^t$, and $\mathbf{x}_8 = [1, 1, 1]^t$. A neural network with two discrete bipolar perceptrons in the hidden layer and a single discrete bipolar output perceptron needs to classify the presented inputs in either of two classes, C_1 and C_2 such that $x_1, x_2, x_3 \in C_1$, with the remaining inputs belonging to C_2.
i– Check whether the weights $\mathbf{w}_1 = [2, 1]^T, \theta_1 = 5, \quad \text{and} \quad \mathbf{w}_2 = [0, 1]^T, \theta_2 = -2$ would provide the linear separation of patterns as required.
ii– Repeat part **i** for the weights $\mathbf{w}_1 = [0, -1]^T, \theta_1 = 1.5, \quad \text{and} \quad \mathbf{w}_2 = [1, 0]^T, \theta_2 = -2.5$
iii– Complete the design of the classifier by using the results from either part **i** or **ii** and compute the weights of the single perceptron at the output.

3. **Constructing a Network** - Construct a multilayer perceptron which will be able to separate the two classes shown in the figure. Use two neurons in the output layer and find a suitable number of hidden layer neurons. Two output of the network should be $[1, 0]^t$ if the input belongs to class C_1 and $[0, 1]^t$ if the input belongs to class C_2. Use the activation function
$$\sigma(net) = \begin{cases} 1 & \text{if } net > 0; \\ 0 & \text{if } net \leq 0. \end{cases}$$
and determine the weight by hand. What is the minimum amount of neurons in the hidden layer required for a perfect separation of the classes?



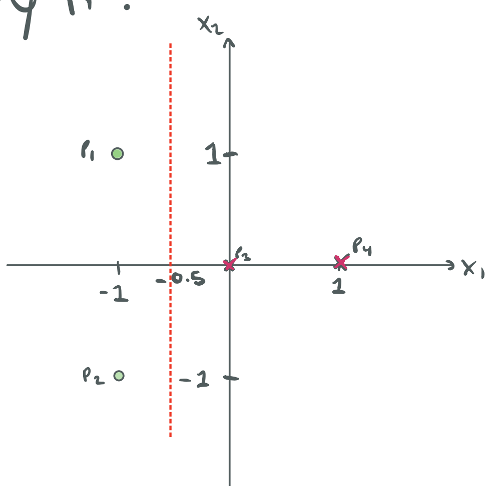
Kieran Cosgrove Student #: 20226841

Assignment 1 - Theoretical Part

Question 1 - Perceptron

(a) $P_1 = ([-1, 1]^t, 1)$, $P_2 = ([-1, -1], 1)$, $P_3 = ([0, 0]^t, 0)$, $P_4 = ([1, 0]^t, 0)$.

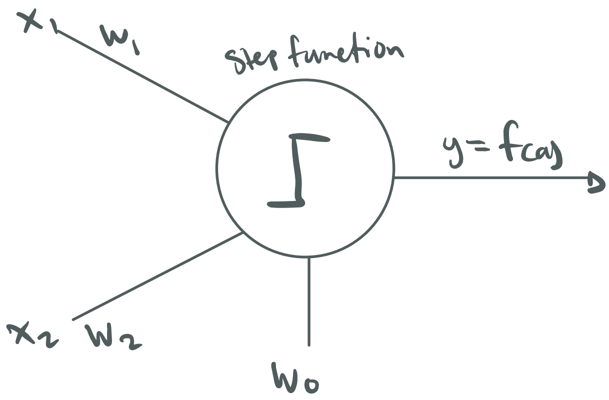
After attempting an iterative approach and graphing, it would take close to 85 iterations to find a solution. Therefore I am under the assumption that I will pick a line and justify it:



Derived from this line I can deduce that the intercepts are $(-\frac{1}{2}, 0)$ and $(0, \infty)$.

Thus appropriate weights must follow the format

$$[\ ((-w_0/w_1), 0) \text{ and } (0, (-w_0/w_2))\ ]$$



$x = (-2, -4, 0)$ where the step function

$\uparrow$ w_0, $\uparrow$ w_1, $\uparrow$ w_2

is 1 if function is greater than 0 and 0 if less than or equal to 0.

$$\sigma(net) \begin{cases} 1 & \text{if } net > 0 \\ 0 & \text{if } net \leq 0 \end{cases}$$

b) i) $P_1 = ([-1, 1]^t, 1)$

$= w_1 x_1 + w_2 x_2$

$= -1(-4) + 1(0)$

$= 4 > 0 \therefore > 1$ ✓

ii) $P_2 = ([-1, -1]^t, 1)$

$= -1(-4) + (-1)(0)$

$= 4 > 0$ ✓

iii) $P_3 = ([0, 0]^t, 0]$

$= 0(-4) + 0(0)$

$= 0 \geq 0$ ✓

iv) $P_4 = [[1, 0]^t, 0]$

$= 1(-4) + 0(0)$

$= -4 \leq 0$ ✓

c) $P_5 = ([-2,0]^T)$ $P_6 = ([1,1]^T)$ $P_7 = ([0,1]^T)$ $P_8 = ([-1,-2]^T)$

$= -2(-4) + 0(0)$ $= (1)(-4) + (1)(0)$ $= (0)(-4) + (1)(0)$ $= (-1)(-4) + (-2)(0)$

$= 8 > 0 \therefore 1$ $= -4 \leq 0 \therefore 0$ $= 0 \leq 0 \therefore 0$ $= 4 > 0 \therefore 1$

Therefore, $P_5 = ([-2,0]^T, 1)$, $P_6 = ([1,1]^T, 0]$, and $P_7 = ([-1,-2]^T, 1)$

d) The vector p5 is always going to be classified the same way (1) as it is to the left of my decision boundary and thus will always result in a unique classification. Similarly, P_6 will always be 1. Alternatively P_7 and P_8 will vary depending on weights or bias of the solution selected. The classification of vectors whose absolute value of x_1 is less than a bias can be particularly susceptible.

Question 2: MLP

- $x_1 x_2 x_3 \in C_1$ else $\in C_2$

i) $W_1 = [2,1]^T, \theta_1 = 5$, $W_2 = [0,1]^T$, $\theta = -2$

$$P_1 = \begin{cases} 1 & \text{if } 2x_1 + 0x_2 > 5 \\ 0 & \text{otherwise} \end{cases} \quad P_2 = \begin{cases} 1 & \text{if } x_1 + x_2 > -2 \\ 0 & \text{otherwise} \end{cases}$$

$X_1 = (2(3) + 0(1) > 5, (1)3 + (1)(1) > -2)$
$= (6 > 5, 4 > -2)$
$\Rightarrow [1, 1]$

$x_2 = (2(4) + 0(0) > 5, (1)(4) + (0)(0) > -2)$
$= (8 > 5, 4 > -2)$
$\Rightarrow [1, 1]$

$X_3 = (2(4) + (0)(-1) > 5, (1)(4) + (1)(0) > -2)$
$= (8 > 5, 3 > -2)$
$\Rightarrow [1, 1]$

$X_4 = (2(5) + (0)(2) > 5, (1)(5) + (1)(2) > -2)$
$= (10 > 5, 7 > -2)$
$\Rightarrow [1, 1]$

$X_5 = (2(5) + (0)(3) > 5, (1)(5) + (1)(3) > -2)$
$= (10 > 5, 8 > -2)$
$\Rightarrow [1, 1]$

$X_6 = (2(3) + (0)(3) > 5, (1)(3) + (1)(3) > -2)$
$= (6 > 5, 6 > -2)$
$\Rightarrow [1, 1]$

$X_7 = (2(2) + (0)(0) > 5, (1)2 + (1)(0) > -2)$
$= (4 < 5, 2 > -2)$
$\Rightarrow [0, 1]$

$X_8 = (2(1) + (0)(1) > 5, (1)(1) + (1)(1) > -2)$
$= (2 < 5, 2 > -2)$
$\Rightarrow [0, 1]$

Given the fact that only $x_1, x_2, x_3 \in C_1$ these weights and thresholds for w_1 and w_2 cannot provide linear seperation of the patterns required.

ii) $P_1 = \begin{cases} 1 & \text{if } 0x_1 + (1)x_2 > 1.5 \\ 0 & \text{otherwise} \end{cases}$ $P_2 = \begin{cases} 1 & \text{if } (-1)x_1 + (0)x_2 > -2.5 \\ 0 & \end{cases}$

$x_1 = (0(3) + 1(1) > 1.5, -1(3) + (0)(1) > -2.5)$
$= (1 > 1.5, -3 > -2.5)$
$= [0, 0]$

$x_2 = (0(4) + 1(0) > 1.5, -1(4) + (0)(0) > -2.5)$
$= (0 > 1.5, -4 > -2.5)$
$\Rightarrow [0, 0]$

$x_3 = (0(4) + 1(-1) > 1.5, -1(4) + (0)(-1) > -2.5)$
$= (-1 > 1.5, -4 > -2.5)$
$\Rightarrow [0, 0]$

$x_4 = (0(5) + 1(2) > 1.5, -1(5) + (0)(2) > -2.5)$
$= (2 > 1.5, -5 > -2.5)$
$\Rightarrow [1, 0]$

$x_5 = (0(5) + 1(3) > 1.5, -1(5) + (0)(3) > -2.5)$
$= (3 > 1.5, -5 > -2.5)$
$\Rightarrow [1, 0]$

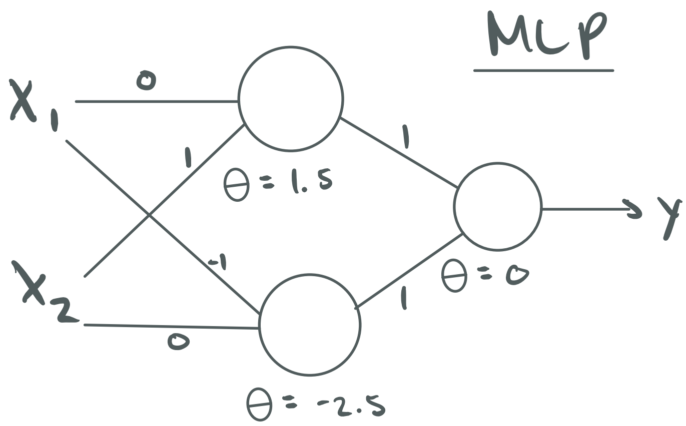
$x_6 = (0(3) + 1(3) > 1.5, -1(3) + (0)(3) > -2.5)$
$= (3 > 1.5, -3 > -2.5)$
$\Rightarrow [1, 0]$

$x_7 = (0(2) + 1(0) > 1.5, -1(2) + (0)(0) > -2.5)$
$= (0 > 1.5, -2 > 2.5)$
$\Rightarrow [0, 1]$

$x_8 = (0(1) + 1(1) > 1.5, -1(1) + (0)(1) > -2.5)$
$= (1 > 1.5, -1 > -2.5)$
$\Rightarrow [0, 1]$

These new weights and thresholds do provide linear seperation between C_1 and C_2.

iii)



Where C_1 is output $y=0$ and C_2 is $y=1$.

Question 3: Constructing a Network

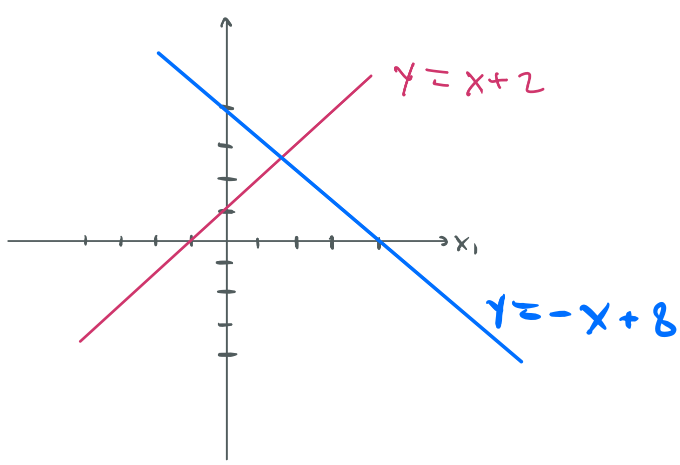
Outputs:

$C_1 : [1,0]^t$, $C_2 : [0,1]^t$

Activation function

$$\sigma(net) = \begin{cases} 1 & \text{if } net > 0 \\ 0 & \text{if } net \le 0 \end{cases}$$

In order to seperate these two classes we will need two linear function:



Recall the formula:

$$x_2 = (-w_1/w_2)x_1 - w_0/w_2$$

For $y = x+2$:

$1 = \frac{-w_1}{w_2}$ and $2 = \frac{-w_0}{w_2}$ $\Rightarrow$ $w_2 = -w_1$ and $2w_2 = -w_0$

Thus $\left.\begin{aligned} w_0 &= -2 \\ w_1 &= -1 \\ w_2 &= 1 \end{aligned}\right\}$ $-\frac{(-1)}{1} = 1$ and $-\frac{(-2)}{1} = 2$ ✓

For $y = -x+8$

$-1 = \frac{-w_1}{w_2}$ and $8 = \frac{-w_0}{w_2}$

Thus $\left.\begin{aligned} w_0 &= -8 \\ w_1 &= 1 \\ w_2 &= 1 \end{aligned}\right\}$ $-1 = -\frac{(1)}{1}$ and $-\frac{(-8)}{1} = 8$ ✓

Given that two lines are necessary to seperate the classes, the minimum number of hidden layers is two.